Contents – Part II

Research Full Paper: Streaming and Real-Time Data Analysis

Flexible and Adaptive Stream Join Algorithm ........................................... 3
   Junhua Fang, Xiaotong Wang, Rong Zhang, and Aoying Zhou

Finding Frequent Items in Time Decayed Data Streams ......................... 17
   Shanshan Wu, Huaizhong Lin, Leong Hou U, Yunjun Gao, and Dongming Lu

A Target-Dependent Sentiment Analysis Method for Micro-blog Streams . . . 30
   Yongheng Wang, Hui Gao, and Shaofeng Geng

Online Streaming Feature Selection Using Sampling Technique and Correlations Between Features ......................................................... 43
   Hai-Tao Zheng and Haiyang Zhang

Real-Time Anomaly Detection over ECG Data Stream Based on Component Spectrum .......................................................... 56
   Meng Wu, Zhen Qiu, Shenda Hong, and Hongyan Li

A Workload-Driven Vertical Partitioning Approach Based on Streaming Framework .......................................................... 68
   Hong Kang, Mengyu Guo, and Xiaojie Yuan

Research Full Paper: Recommendation System

Spica: A Path Bundling Model for Rational Route Recommendation ........... 85
   Lei Lv, Yang Liu, and Xiaohui Yu

FHSM: Factored Hybrid Similarity Methods for Top-N Recommender Systems .......................................................... 98
   Xin Xin, Dong Wang, Yue Ding, and Chen Lini

Personalized Resource Recommendation Based on Regular Tag and User Operation .......................................................... 111
   Sisi Liu, Yongjian Liu, and Qing Xie

Academic Paper Recommendation Based on Community Detection in Citation-Collaboration Networks ......................................................... 124
   Qisen Wang, Wenzhong Li, Xiao Zhang, and Sanglu Lu
Improving Recommendation Accuracy for Travelers by Exploiting POI Correlations
Kai Zhang, Dapeng Zhao, and Xiaoling Wang 137

A Context-Aware Method for Top-k Recommendation in Smart TV
Peng Liu, Jun Ma, Yongjin Wang, Lintao Ma, and Shanshan Huang 150

A Hybrid Method for POI Recommendation: Combining Check-In Count, Geographical Information and Reviews
Xiefeng Xu, Pengpeng Zhao, Guanfeng Liu, Caidong Gu, Jiajie Xu, Jian Wu, and Zhiming Cui 162

Improving Temporal Recommendation Accuracy and Diversity via Long and Short-Term Preference Transfer and Fusion Models
Bei Zhang and Yong Feng 174

An Approach for Cross-Community Content Recommendation: A Case Study on Docker
Yang Yong, Li Ying, Tang Hongyan, Jia Tong, and Shao Wenlong 186

Research Full Paper: Data Quality and Privacy

Scalable Private Blocking Technique for Privacy-Preserving Record Linkage
Shumin Han, Derong Shen, Tiezheng Nie, Yue Kou, and Ge Yu 201

Repair Singleton IDs on the Fly
Xingcan Cui, Xiaohui Yu, and De Guo 214

Modeling for Noisy Labels of Crowd Workers
Qian Yan, Hao Huang, Yunjun Gao, Chen Ying, Qingyang Hu, Tieyun Qian, and Qinming He 227

Incomplete Data Classification Based on Multiple Views
Ming Sun, Hongzhi Wang, Fanshan Meng, Jianzhong Li, and Hong Gao 239

Fuzzy Keywords Query
Shanshan Han, Hongzhi Wang, Hong Gao, Jianzhong Li, and Shenbin Huang 251

EPLA: Efficient Personal Location Anonymity
Dapeng Zhao, Kai Zhang, Yuanyuan Jin, Xiaoling Wang, Patrick C.K. Hung, and Wendi Ji 263

A Secure and Robust Covert Channel Based on Secret Sharing Scheme
Xiaorong Lu, Yang Wang, Liusheng Huang, Wei Yang, and Yao Shen 276
Discovering Approximate Functional Dependencies from Distributed Big Data ................................................................. 289
  Weibang Li, Zhanhuai Li, Qun Chen, Tao Jiang, and Zhilei Yin

Research Full Paper: Query Optimization and Scalable Data Processing

Making Cold Data Identification Efficient in Non-volatile Memory Systems . . . 305
  Binbin Wang and Jiwu Shu

Preference Join on Heterogeneous Data .................................................. 317
  Changping Wang, Chaokun Wang, Hao Wang, Jun Chen, and Xiaojun Ye

Accelerating Time Series Shapelets Discovery with Key Points. ............... 330
  Zhenguo Zhang, Haiwei Zhang, Yanlong Wen, and Xiaojie Yuan

An Adaptive Partition-Based Caching Approach for Efficient Range Queries on Key-Value Data .................................................. 343
  Wei Ge, Min Chen, Chunfeng Yuan, and Yihua Huang

Handling Estimation Inaccuracy in Query Optimization ......................... 355
  Chiraz Moumen, Franck Morvan, and Abdelkader Hameurlain

Practical Study of Subclasses of Regular Expressions in DTD and XML Schema .................................................. 368
  Yeting Li, Xiaolan Zhang, Feifei Peng, and Haiming Chen

Research Short Paper

Chinese Microblog Sentiment Analysis Based on Sentiment Features ......... 385
  Weiwei Li, Yuqiang Li, and Yan Wang

FVBM: A Filter-Verification-Based Method for Finding Top-k Closeness Centrality on Dynamic Social Networks .................................................. 389
  Yiyong Lin, Jinbo Zhang, Yuanxiang Ying, Shenda Hong, and Hongyan Li

Online Hot Topic Detection from Web News Based on Bursty Term Identification .................................................. 393
  Chao Wang, Xue Zhao, Ying Zhang, and Xiaojie Yuan

Grouped Team Formation in Social Networks. ................................. 398
  Ze Lv, Jianbin Huang, Yu Zhou, Heli Sun, and Xiaolin Jia

Profit Maximizing Route Recommendation for Vehicle Sharing Requests. . . 402
  Zhiqiang Zhao, Jianbin Huang, Hua Gao, Heli Sun, and Xiaolin Jia
Ontology-Based Interactive Post-mining of Interesting Co-location Patterns  
Xuguang Bao, Lizhen Wang, and Hongmei Chen

AALRSMF: An Adaptive Learning Rate Schedule for Matrix Factorization  
Feng Wei, Hao Guo, Shaoyin Cheng, and Fan Jiang

A Graph Clustering Algorithm for Citation Networks  
Bo Zhang, Tiezheng Nie, Derong Shen, Yue Kou, Ge Yu, and Ziwei Zhou

A Distributed Frequent Itemsets Mining Algorithm Using Sparse Boolean Matrix on Spark  
Yonghong Luo, Zhifan Yang, Huike Shi, and Ying Zhang

A Simple Stochastic Gradient Variational Bayes for the Correlated Topic Model  
Tomonari Masada and Atsuhiro Takasu

Reasoning with Large Scale OWL 2 EL Ontologies Based on MapReduce  
Zhangquan Zhou, Guilin Qi, Chang Liu, Raghava Mutharaju, and Pascal Hitzler

Purchase and Redemption Prediction Based on Multi-task Gaussian Process and Dimensionality Reduction  
Chao Wang, Xiangrui Cai, Zhenguo Zhang, and Yanlong Wen

Similarity Recoverable, Format-Preserving String Encryption  
Yijin Li and Wendy Hui Wang

RORS: Enhanced Rule-Based OWL Reasoning on Spark  
Zhihui Liu, Zhiyong Feng, Xiaowang Zhang, Xin Wang, and Guozheng Rao

A Hadoop-Based Database Querying Approach for Non-expert Users  
Yale Chai, Chao Wang, Yanlong Wen, and Xiaojie Yuan

A Collaborative Join Scheme on a MIC-Based Heterogeneous Platform  
Kailai Zhou, Hui Sun, Hong Chen, Tianzhen Wu, and Cuiping Li

Pairwise Expansion: A New Topdown Search for mCK Queries Problem over Spatial Web  
Yuan Qiu, Tadashi Ohmori, Takahiko Shintani, and Hideyuki Fujita

Mentioning the Optimal Users in the Appropriate Time on Twitter  
Zhaoyun Ding, Xueqing Zou, Yueyang Li, Su He, Jiajun Cheng, Fengcai Qiao, and Hui Wang

Historical Geo-Social Query Processing  
Xiaoying Chen, Chong Zhang, Yanli Hu, Bin Ge, and Weidong Xiao
WS-Rank: Bringing Sentences into Graph for Keyword Extraction ............ 474
   Fan Yang, Yue-Sheng Zhu, and Yu-Jia Ma

Efficient Community Maintenance for Dynamic Social Networks .......... 478
   Hongchao Qin, Ye Yuan, Feida Zhu, and Guoren Wang

Open Sesame! Web Authentication Cracking via Mobile App Analysis ...... 483
   Hui Liu, Yuanyuan Zhang, Juanru Li, Hui Wang, and Dawu Gu

K-th Order Skyline Queries in Bicriteria Networks ......................... 488
   Shunqing Jiang, Jiping Zheng, Jialiang Chen, and Wei Yu

A K-Motifs Discovery Approach for Large Time-Series Data Analysis ...... 492
   Yupeng Hu, Cun Ji, Ming Jing, and Xueqing Li

User Occupation Prediction on Microblogs .................................. 497
   Xia Lv, Peiquan Jin, and Lihua Yue

Industry Full Paper

Combo-Recommendation Based on Potential Relevance of Items ............. 505
   Yanhong Pan, Yanfei Zhang, and Rong Zhang

Demo Paper

OPGs-Rec: Organized-POI-Groups Based Recommendation ..................... 521
   JiaPeng Li, Yanxia Xu, and Lei Zhao

A Demonstration of Encrypted Logistics Information System .............. 525
   Huakang Li, Xinwen Zhang, Yitao Yang, and Guozi Sun

PCMiner: An Extensible System for Analysing and Detecting Protein
   Complexes ............................................................................. 529
   Danyang Xiao, Jia Zhu, Yong Tang, Lingxiao Chen, and Jingmin Wei

MASM: A Novel Movie Analysis System Based on Microblog ............... 533
   Xingcheng Wu, Jia Zhu, Yong Tang, Rui Ding, Xueqin Lin,
   and Chuanhua Xu

A Text Retrieval System Based on Distributed Representations ............ 537
   Zhe Zhao, Tao Liu, Jun Chen, Bofang Li, and Xiaoyong Du

A System for Searching Renting Houses Based on Relaxed
   Query Answering .................................................................... 542
   Jianfeng Du, Kunxun Qi, and Can Lin

TagTour: A Personalized Tourist Resource Recommendation System ....... 547
   Tian Han, Yongjian Liu, and Qing Xie
Contents – Part I

Research Full Paper: Spatio-temporal, Textual and Multimedia Data Management

Probabilistic Nearest Neighbor Query in Traffic-Aware Spatial Networks

Shuo Shang, Zhewei Wei, Ji-Rong Wen, and Shunzhi Zhu

NERank: Bringing Order to Named Entities from Texts

Chengyu Wang, Rong Zhang, Xiaofeng He, Guomin Zhou, and Aoying Zhou

FTS: A Practical Model for Feature-Based Trajectory Synthesis

Jiapeng Li, Wei Chen, An Liu, Zhixu Li, and Lei Zhao

Distributed Text Representation with Weighting Scheme Guidance for Sentiment Analysis

Zhe Zhao, Tao Liu, Xiaoyun Hou, Bofang Li, and Xiaoyong Du

A Real Time Wireless Interactive Multimedia System

Hong Li, Wei Yang, Yang Xu, Jianxin Wang, and Liusheng Huang

Mining Co-locations from Continuously Distributed Uncertain Spatial Data

Bozhong Liu, Ling Chen, Chunyang Liu, Chengqi Zhang, and Weidong Qiu

An Online Approach for Direction-Based Trajectory Compression with Error Bound Guarantee

Bingqing Ke, Jie Shao, Yi Zhang, Dongxiang Zhang, and Yang Yang

A Data Grouping CNN Algorithm for Short-Term Traffic Flow Forecasting

Donghai Yu, Yang Liu, and Xiaohui Yu

Efficient Evaluation of Shortest Travel-Time Path Queries in Road Networks by Optimizing Waypoints in Route Requests

Detian Zhang, Chi-Yin Chow, Qing Li, and An Liu

Discovering Companion Vehicles from Live Streaming Traffic Data

Chen Liu, Xiongbin Wang, Meiling Zhu, and Yanbo Han

Time-Constrained Sequenced Route Query in Indoor Spaces

Wenyi Luo, Peiquan Jin, and Lihua Yue
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near-Duplicate Web Video Retrieval and Localization Using Improved</td>
<td>141</td>
</tr>
<tr>
<td>Edit Distance</td>
<td></td>
</tr>
<tr>
<td>Hao Liu, Qingjie Zhao, Hao Wang, and Cong Zhang</td>
<td></td>
</tr>
<tr>
<td>Efficient Group Top-k Spatial Keyword Query Processing</td>
<td>153</td>
</tr>
<tr>
<td>Kai Yao, Jianjun Li, Guohui Li, and Changyin Luo</td>
<td></td>
</tr>
<tr>
<td><strong>Research Full Paper: Social Media Data Analysis</strong></td>
<td></td>
</tr>
<tr>
<td>Learn to Recommend Local Event Using Heterogeneous Social Networks .</td>
<td>169</td>
</tr>
<tr>
<td>Shaoqing Wang, Zheng Wang, Cuiping Li, Kankan Zhao, and Hong Chen</td>
<td></td>
</tr>
<tr>
<td>Top-k Temporal Keyword Query over Social Media Data</td>
<td>183</td>
</tr>
<tr>
<td>Fan Xia, Chengcheng Yu, Weining Qian, and Aoying Zhou</td>
<td></td>
</tr>
<tr>
<td>When a Friend Online is More Than a Friend in Life: Intimate</td>
<td>196</td>
</tr>
<tr>
<td>Relationship Prediction in Microblogs</td>
<td></td>
</tr>
<tr>
<td>Yunshi Lan, Mengqi Zhang, Feida Zhu, Jing Jiang, and Ee-Peng Lim</td>
<td></td>
</tr>
<tr>
<td>Community Inference with Bayesian Non-negative Matrix Factorization</td>
<td>208</td>
</tr>
<tr>
<td>Xiaohua Shi and Hongtao Lu</td>
<td></td>
</tr>
<tr>
<td>Maximizing the Influence Ranking Under Limited Cost in Social Network.</td>
<td>220</td>
</tr>
<tr>
<td>Xiaoguang Hong, Ziyan Liu, Zhaohui Peng, Zhiyong Chen, and Hui Li</td>
<td></td>
</tr>
<tr>
<td>Towards Efficient Influence Maximization for Evolving Social Networks.</td>
<td>232</td>
</tr>
<tr>
<td>Xiaodong Liu, Xiangke Liao, Shanshan Li, and Bin Lin</td>
<td></td>
</tr>
<tr>
<td>Detecting Community Pacemakers of Burst Topic in Twitter</td>
<td>245</td>
</tr>
<tr>
<td>Guozhong Dong, Wu Yang, Feida Zhu, and Wei Wang</td>
<td></td>
</tr>
<tr>
<td>Dynamic User Attribute Discovery on Social Media</td>
<td>256</td>
</tr>
<tr>
<td>Xiu Huang, Yang Yang, Yue Hu, Fumin Shen, and Jie Shao</td>
<td></td>
</tr>
<tr>
<td>The Competition of User Attentions Among Social Network Services:</td>
<td>268</td>
</tr>
<tr>
<td>A Social Evolutionary Game Approach</td>
<td></td>
</tr>
<tr>
<td>Jingyuan Li, Yuanzhuo Wang, Yuan Lu, Xueqi Cheng, and Yan Ren</td>
<td></td>
</tr>
<tr>
<td>Mechanism Analysis of Competitive Information Synchronous</td>
<td>280</td>
</tr>
<tr>
<td>Dissemination in Social Networks</td>
<td></td>
</tr>
<tr>
<td>Yuan Lu, Yuanzhuo Wang, Jianye Yu, Jingyuan Li, and Li Liu</td>
<td></td>
</tr>
<tr>
<td>A Topic-Specific Contextual Expert Finding Method in Social Network.</td>
<td>292</td>
</tr>
<tr>
<td>Xiaojin Xie, Yijia Li, Zhiqiang Zhang, Haiwei Pan, and Shuai Han</td>
<td></td>
</tr>
</tbody>
</table>
Budget Minimization with Time and Influence Constraints in Social Network ........................................ 304
  Peng Dou, Sizhen Du, and Guojie Song

B-mine: Frequent Pattern Mining and Its Application to Knowledge Discovery from Social Networks ........................................ 316
  Fan Jiang, Carson K. Leung, and Hao Zhang

An Efficient Online Event Detection Method for Microblogs via User Modeling ........................................ 329
  Weijing Huang, Wei Chen, Lamei Zhang, and Tengjiao Wang

Man-O-Meter: Modeling and Assessing the Evolution of Language Usage of Individuals on Microblogs ........................................ 342
  Kuntal Dey, Saroj Kaushik, Hemank Lamba, and Seema Nagar

Research Full Paper: Modelling and Learning with Big Data

Forecasting Career Choice for College Students Based on Campus Big Data ........................................ 359
  Min Nie, Lei Yang, Bin Ding, Hu Xia, Huachun Xu, and Defu Lian

Online Prediction for Forex with an Optimized Experts Selection Model ........................................ 371
  Jia Zhu, Jing Yang, Jing Xiao, Changqin Huang, Gansen Zhao, and Yong Tang

Fast Rare Category Detection Using Nearest Centroid Neighborhood ........................................ 383
  Song Wang, Hao Huang, Yunjun Gao, Tieyun Qian, Liang Hong, and Zhiyong Peng

Latent Semantic Diagnosis in Traditional Chinese Medicine ........................................ 395
  Wendi Ji, Ying Zhang, Xiaoling Wang, and Yiping Zhou

Correlation-Based Weighted K-Labelsets for Multi-label Classification ........................................ 408
  Jingyang Xu and Jun Ma

Classifying Relation via Bidirectional Recurrent Neural Network Based on Local Information ........................................ 420
  Xiaoyun Hou, Zhe Zhao, Tao Liu, and Xiaoyong Du

Psychological Stress Detection from Online Shopping ........................................ 431
  Liang Zhao, Hao Wang, Yuanyuan Xue, Qi Li, and Ling Feng

Confidence-Learning Based Collaborative Filtering with Heterogeneous Implicit Feedbacks ........................................ 444
  Jing Wang, Lanfen Lin, Heng Zhang, and Jiaqi Tu

Star-Scan: A Stable Clustering by Statistically Finding Centers and Noises ........................................ 456
  Nan Yang, Qing Liu, Yaping Li, Lin Xiao, and Xiaoping Liu
Web Technologies and Applications
18th Asia-Pacific Web Conference, APWeb 2016,
Suzhou, China, September 23-25, 2016. Proceedings,
Part II
Li, F.; Shim, K.; Zheng, K.; Liu, G. (Eds.)
2016, XXII, 601 p. 236 illus., Softcover
ISBN: 978-3-319-45816-8