## Contents – Part I

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

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probabilistic Nearest Neighbor Query in Traffic-Aware Spatial Networks</td>
<td>3</td>
</tr>
<tr>
<td><em>Shuo Shang, Zhewei Wei, Ji-Rong Wen, and Shunzhi Zhu</em></td>
<td></td>
</tr>
<tr>
<td>NERank: Bringing Order to Named Entities from Texts</td>
<td>15</td>
</tr>
<tr>
<td><em>Chengyu Wang, Rong Zhang, Xiaofeng He, Guomin Zhou, and Aoying Zhou</em></td>
<td></td>
</tr>
<tr>
<td>FTS: A Practical Model for Feature-Based Trajectory Synthesis</td>
<td>28</td>
</tr>
<tr>
<td><em>Jiapeng Li, Wei Chen, An Liu, Zhixu Li, and Lei Zhao</em></td>
<td></td>
</tr>
<tr>
<td>Distributed Text Representation with Weighting Scheme Guidance</td>
<td>41</td>
</tr>
<tr>
<td>for Sentiment Analysis</td>
<td></td>
</tr>
<tr>
<td><em>Zhe Zhao, Tao Liu, Xiaoyun Hou, Bofang Li, and Xiaoyong Du</em></td>
<td></td>
</tr>
<tr>
<td>A Real Time Wireless Interactive Multimedia System</td>
<td>53</td>
</tr>
<tr>
<td><em>Hong Li, Wei Yang, Yang Xu, Jianxin Wang, and Liusheng Huang</em></td>
<td></td>
</tr>
<tr>
<td>Mining Co-locations from Continuously Distributed Uncertain Spatial Data</td>
<td>66</td>
</tr>
<tr>
<td><em>Bozhong Liu, Ling Chen, Chunyang Liu, Chengqi Zhang, and Weidong Qiu</em></td>
<td></td>
</tr>
<tr>
<td>An Online Approach for Direction-Based Trajectory Compression</td>
<td>79</td>
</tr>
<tr>
<td>with Error Bound Guarantee</td>
<td></td>
</tr>
<tr>
<td><em>Bingqing Ke, Jie Shao, Yi Zhang, Dongxiang Zhang, and Yang Yang</em></td>
<td></td>
</tr>
<tr>
<td>A Data Grouping CNN Algorithm for Short-Term Traffic Flow Forecasting</td>
<td>92</td>
</tr>
<tr>
<td><em>Donghai Yu, Yang Liu, and Xiaohui Yu</em></td>
<td></td>
</tr>
<tr>
<td>Efficient Evaluation of Shortest Travel-Time Path Queries in Road Networks by Optimizing Waypoints in Route Requests</td>
<td>104</td>
</tr>
<tr>
<td><em>Detian Zhang, Chi-Yin Chow, Qing Li, and An Liu</em></td>
<td></td>
</tr>
<tr>
<td>Discovering Companion Vehicles from Live Streaming Traffic Data</td>
<td>116</td>
</tr>
<tr>
<td><em>Chen Liu, Xiongbin Wang, Meiling Zhu, and Yanbo Han</em></td>
<td></td>
</tr>
<tr>
<td>Time-Constrained Sequenced Route Query in Indoor Spaces</td>
<td>129</td>
</tr>
<tr>
<td><em>Wenyi Luo, Peiquan Jin, and Lihua Yue</em></td>
<td></td>
</tr>
</tbody>
</table>
Near-Duplicate Web Video Retrieval and Localization Using Improved Edit Distance .................................................. 141
   Hao Liu, Qingjie Zhao, Hao Wang, and Cong Zhang

Efficient Group Top-k Spatial Keyword Query Processing ................. 153
   Kai Yao, Jianjun Li, Guohui Li, and Changyin Luo

Research Full Paper: Social Media Data Analysis

Learn to Recommend Local Event Using Heterogeneous Social Networks ... 169
   Shaoqing Wang, Zheng Wang, Cuiping Li, Kankan Zhao, and Hong Chen

Top-k Temporal Keyword Query over Social Media Data ................. 183
   Fan Xia, Chengcheng Yu, Weining Qian, and Aoying Zhou

When a Friend Online is More Than a Friend in Life: Intimate Relationship Prediction in Microblogs ............................... 196
   Yunshi Lan, Mengqi Zhang, Feida Zhu, Jing Jiang, and Ee-Peng Lim

Community Inference with Bayesian Non-negative Matrix Factorization .... 208
   Xiaohua Shi and Hongtao Lu

Maximizing the Influence Ranking Under Limited Cost in Social Network ... 220
   Xiaoguang Hong, Ziyun Liu, Zhaohui Peng, Zhiyong Chen, and Hui Li

Towards Efficient Influence Maximization for Evolving Social Networks ... 232
   Xiaodong Liu, Xiangke Liao, Shanshan Li, and Bin Lin

Detecting Community Pacemakers of Burst Topic in Twitter ................ 245
   Guozhong Dong, Wu Yang, Feida Zhu, and Wei Wang

Dynamic User Attribute Discovery on Social Media .......................... 256
   Xiu Huang, Yang Yang, Yue Hu, Fumin Shen, and Jie Shao

The Competition of User Attentions Among Social Network Services:
A Social Evolutionary Game Approach ...................................... 268
   Jingyuan Li, Yuanzhuo Wang, Yuan Lu, Xueqi Cheng, and Yan Ren

Mechanism Analysis of Competitive Information Synchronous
Dissemination in Social Networks .............................................. 280
   Yuan Lu, Yuanzhuo Wang, Jianye Yu, Jingyuan Li, and Li Liu

A Topic-Specific Contextual Expert Finding Method in Social Network .... 292
   Xiaoqin Xie, Yijia Li, Zhiqiang Zhang, Haiwei Pan, and Shuai Han
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
Aggregating Crowd Wisdom with Instance Grouping Methods
Li’ang Yin, Zhengbo Li, Jianhua Han, and Yong Yu

CoDS: Co-training with Domain Similarity for Cross-Domain Image Sentiment Classification
Linlin Zhang, Meng Chen, Xiaohui Yu, and Yang Liu

Feature Selection via Vectorizing Feature’s Discriminative Information
Jun Wang, Hengpeng Xu, and Jinmao Wei

A Label Inference Method Based on Maximal Entropy Random Walk over Graphs
Jing Pan, Yajun Yang, Qinghua Hu, and Hong Shi

An Adaptive kNN Using Listwise Approach for Implicit Feedback
Bu-Xiao Wu, Jing Xiao, Jia Zhu, and Chen Ding

Quantifying the Effect of Sentiment on Topic Evolution in Chinese Microblog
Peng Fu, Zheng Lin, Hailun Lin, Fengcheng Yuan, Weiping Wang, and Dan Meng

Measuring Directional Semantic Similarity with Multi-features
Bo Liu, Xuanhua Shi, and Hai Jin

Finding Latest Influential Research Papers Through Modeling Two Views of Citation Links
Lu Huang, Hongyan Liu, Jun He, and Xiaoyong Du

Multi-label Chinese Microblog Emotion Classification via Convolutional Neural Network
Yaqi Wang, Shi Feng, Daling Wang, Ge Yu, and Yifei Zhang

Mining Recent High Expected Weighted Itemsets from Uncertain Databases
Wensheng Gan, Jerry Chun-Wei Lin, Philippe Fournier-Viger, and Han-Chieh Chao

Context-Aware Chinese Microblog Sentiment Classification with Bidirectional LSTM
Yang Wang, Shi Feng, Daling Wang, Yifei Zhang, and Ge Yu

Author Index
## Contents – Part II

### Research Full Paper: Streaming and Real-Time Data Analysis

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flexible and Adaptive Stream Join Algorithm.</td>
<td>3</td>
</tr>
<tr>
<td><em>Junhua Fang, Xiaotong Wang, Rong Zhang, and Aoying Zhou</em></td>
<td></td>
</tr>
<tr>
<td>Finding Frequent Items in Time Decayed Data Streams</td>
<td>17</td>
</tr>
<tr>
<td><em>Shanshan Wu, Huaizhong Lin, Leong Hou U, Yunjun Gao, and Dongming Lu</em></td>
<td></td>
</tr>
<tr>
<td>A Target-Dependent Sentiment Analysis Method for Micro-blog Streams.</td>
<td>30</td>
</tr>
<tr>
<td><em>Yongheng Wang, Hui Gao, and Shaofeng Geng</em></td>
<td></td>
</tr>
<tr>
<td>Online Streaming Feature Selection Using Sampling Technique and Correlations Between Features</td>
<td>43</td>
</tr>
<tr>
<td><em>Hai-Tao Zheng and Haiyang Zhang</em></td>
<td></td>
</tr>
<tr>
<td>Real-Time Anomaly Detection over ECG Data Stream Based on Component Spectrum</td>
<td>56</td>
</tr>
<tr>
<td><em>Meng Wu, Zhen Qiu, Shenda Hong, and Hongyan Li</em></td>
<td></td>
</tr>
<tr>
<td>A Workload-Driven Vertical Partitioning Approach Based on Streaming Framework</td>
<td>68</td>
</tr>
<tr>
<td><em>Hong Kang, Mengyu Guo, and Xiaojie Yuan</em></td>
<td></td>
</tr>
</tbody>
</table>

### Research Full Paper: Recommendation System

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spica: A Path Bundling Model for Rational Route Recommendation</td>
<td>85</td>
</tr>
<tr>
<td><em>Lei Lv, Yang Liu, and Xiaohui Yu</em></td>
<td></td>
</tr>
<tr>
<td>FHSM: Factored Hybrid Similarity Methods for Top-N Recommender Systems</td>
<td>98</td>
</tr>
<tr>
<td><em>Xin Xin, Dong Wang, Yue Ding, and Chen Lini</em></td>
<td></td>
</tr>
<tr>
<td>Personalized Resource Recommendation Based on Regular Tag and User Operation</td>
<td>111</td>
</tr>
<tr>
<td><em>Sisi Liu, Yongjian Liu, and Qing Xie</em></td>
<td></td>
</tr>
<tr>
<td>Academic Paper Recommendation Based on Community Detection in Citation-Collaboration Networks</td>
<td>124</td>
</tr>
<tr>
<td><em>Qisen Wang, Wenzhong Li, Xiao Zhang, and Sanglu Lu</em></td>
<td></td>
</tr>
</tbody>
</table>
Improving Recommendation Accuracy for Travelers by Exploiting POI Correlations
Kai Zhang, Dapeng Zhao, and Xiaoling Wang

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

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

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

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

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

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

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

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

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

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

A Secure and Robust Covert Channel Based on Secret Sharing Scheme
Xiaorong Lu, Yang Wang, Liusheng Huang, Wei Yang, and Yao Shen
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. . . 406
  Xuguang Bao, Lizhen Wang, and Hongmei Chen

AALRSMF: An Adaptive Learning Rate Schedule for Matrix Factorization . . 410
  Feng Wei, Hao Guo, Shaoyin Cheng, and Fan Jiang

A Graph Clustering Algorithm for Citation Networks ............................. 414
  Bo Zhang, Tiezhen Nie, Derong Shen, Yue Kou, Ge Yu, and Ziwei Zhou

A Distributed Frequent Itemsets Mining Algorithm Using Sparse Boolean
  Matrix on Spark .................................................................................. 419
  Yonghong Luo, Zhifan Yang, Huike Shi, and Ying Zhang

A Simple Stochastic Gradient Variational Bayes for the Correlated
  Topic Model. ..................................................................................... 424
  Tomonari Masada and Atsuhiro Takasu

Reasoning with Large Scale OWL 2 EL Ontologies Based on MapReduce. . 429
  Zhangquan Zhou, Guilin Qi, Chang Liu, Raghava Mutharaju, and Pascal Hitzler

Purchase and Redemption Prediction Based on Multi-task Gaussian Process
  and Dimensionality Reduction ................................................................. 434
  Chao Wang, Xiangrui Cai, Zhenguo Zhang, and Yanlong Wen

Similarity Recoverable, Format-Preserving String Encryption .................. 439
  Yijin Li and Wendy Hui Wang

RORS: Enhanced Rule-Based OWL Reasoning on Spark .......................... 444
  Zhihui Liu, Zhiyong Feng, Xiaowang Zhang, Xin Wang, and Guozheng Rao

A Hadoop-Based Database Querying Approach for Non-expert Users ......... 449
  Yale Chai, Chao Wang, Yanlong Wen, and Xiaojie Yuan

A Collaborative Join Scheme on a MIC-Based Heterogeneous Platform ....... 454
  Kailai Zhou, Hui Sun, Hong Chen, Tianzhen Wu, and Cuiping Li

Pairwise Expansion: A New Topdown Search for mCK Queries Problem
  over Spatial Web ................................................................................. 459
  Yuan Qiu, Tadashi Ohmori, Takahiko Shintani, and Hideyuki Fujita

 Mentioning the Optimal Users in the Appropriate Time on Twitter ......... 464
  Zhaoyun Ding, Xueqing Zou, Yueyang Li, Su He, Jiajun Cheng, Fengcai Qiao, and Hui Wang

Historical Geo-Social Query Processing .................................................. 469
  Xiaoying Chen, Chong Zhang, Yanli Hu, Bin Ge, and Weidong Xiao
WS-Rank: Bringing Sentences into Graph for Keyword Extraction
Fan Yang, Yue-Sheng Zhu, and Yu-Jia Ma

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

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

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

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

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

Industry Full Paper
Combo-Recommendation Based on Potential Relevance of Items
Yanhong Pan, Yanfei Zhang, and Rong Zhang

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

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

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

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

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

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

TagTour: A Personalized Tourist Resource Recommendation System
Tian Han, Yongjian Liu, and Qing Xie
A Chronic Disease Analysis System Based on Dirty Data Mining

Ming Sun, Hongzhi Wang, Jianzhong Li, Hong Gao, and Shenbin Huang

ADDS: An Automated Disease Diagnosis-Aided System

Zhentuan Xu, Xiaoli Wang, Yating Chen, Yangbin Pan, Mengsang Wu, and Mengyuan Xiong

Indoor Map Service System Based on Wechat Two-Dimensional Code

Chen Guo and Qingwu Hu

Factorization Machine Based Business Credit Scoring by Leveraging Internet Data

Ge Zhu and Lin Li

OICRM: An Ontology-Based Interesting Co-location Rule Miner

Xuguang Bao, Lizhen Wang, and Meijiao Wang

CB-CAS: A CAS-Based Cross-Browser SSO System

Peng Gao, Yongjian Liu, and Qing Xie

A Demonstration of QA System Based on Knowledge Base

Zhenjiang Dong, Hong Chen, Jingqiang Chen, Huakang Li, and Tao Li

An Alarming and Prediction System for Infections Disease Based on Combined Models

Jiahong Li, Hongzhi Wang, Shengqiang Zhang, Xiangyu Gao, Ziqi Qu, and Shenbin Huang

Co-location Detector: A System to Find Interesting Spatial Co-locating Relationships

Xuguang Bao, Lizhen Wang, and Qing Xiao

KEIPD: Knowledge Extraction and Inference System for Personal Documents

Zhaoyang Lv, Yuanyuan Liu, and Xiaohui Yu

Author Index