Contents – Part I

Semantic Web and Knowledge Management

A General Fine-Grained Truth Discovery Approach for Crowdsourced Data Aggregation .................................................. 3
   Yang Du, Hongli Xu, Yu-E Sun, and Liusheng Huang

Learning the Structures of Online Asynchronous Conversations ............ 19
   Jun Chen, Chaokun Wang, Heran Lin, Weiping Wang, Zhipeng Cai, and Jianmin Wang

A Question Routing Technique Using Deep Neural Network for Communities of Question Answering .......................... 35
   Amr Azzam, Neamat Tazi, and Ahmad Hossny

Category-Level Transfer Learning from Knowledge Base to Microblog Stream for Accurate Event Detection ..................... 50
   Weijing Huang, Tengjiao Wang, Wei Chen, and Yazhou Wang

Indexing and Distributed Systems

AngleCut: A Ring-Based Hashing Scheme for Distributed Metadata Management .......................................................... 71
   Jiaxi Liu, Renxuan Wang, Xiaofeng Gao, Xiaochun Yang, and Guihai Chen

An Efficient Bulk Loading Approach of Secondary Index in Distributed Log-Structured Data Stores ......................... 87
   Yanchao Zhu, Zhao Zhang, Peng Cai, Weining Qian, and Aoying Zhou

Performance Comparison of Distributed Processing of Large Volume of Data on Top of Xen and Docker-Based Virtual Clusters .......... 103
   Haejin Chung and Yunmook Nah

An Adaptive Data Partitioning Scheme for Accelerating Exploratory Spark SQL Queries ............................................. 114
   Chenghao Guo, Zhigang Wu, Zhenying He, and X. Sean Wang

Network Embedding

Semi-Supervised Network Embedding ............................................. 131
   Chaozhuo Li, Zhoujun Li, Senzhang Wang, Yang Yang, Xiaoming Zhang, and Jianshe Zhou
CirE: Circular Embeddings of Knowledge Graphs .......................... 148
Zhijuan Du, Zehui Hao, Xiaofeng Meng, and Qiuyue Wang

PPNE: Property Preserving Network Embedding .......................... 163
Chaozhuo Li, Senzhang Wang, Defian Yang, Zhoujun Li, Yang Yang,
Xiaoming Zhang, and Jianshe Zhou

HINE: Heterogeneous Information Network Embedding .................. 180
Yuxin Chen and Chenguang Wang

Trajectory and Time Series Data Processing

DT-KST: Distributed Top-k Similarity Query on Big Trajectory Streams .... 199
Zhigang Zhang, Yilin Wang, Jiali Mao, Shaojie Qiao, Cheqing Jin,
and Aoying Zhou

A Distributed Multi-level Composite Index for KNN Processing on Long
Time Series ................................................................. 215
Xiaqing Wang, Zicheng Fang, Peng Wang, Ruyuan Zhu, and Wei Wang

Outlier Trajectory Detection: A Trajectory Analytics Based Approach ....... 231
Zhongjian Lv, Jiajie Xu, Pengpeng Zhao, Guanfeng Liu, Lei Zhao,
and Xiaofang Zhou

Clustering Time Series Utilizing a Dimension Hierarchical
Decomposition Approach .................................................... 247
Qiuhong Li, Peng Wang, Yang Wang, Wei Wang, Yimin Liu, Jiaye Wu,
and Danyang Dou

Data Mining

Fast Extended One-Versus-Rest Multi-label SVM Classification Algorithm
Based on Approximate Extreme Points ...................................... 265
Zhongwei Sun, Zhongwen Guo, Xupeng Wang, Jing Liu,
and Shiyong Liu

Efficiently Discovering Most-Specific Mixed Patterns from Large
Data Trees ................................................................. 279
Xiaoying Wu and Dimitri Theodoratos

Max-Cosine Matching Based Neural Models for Recognizing
Textual Entailment .......................................................... 295
Zhipeng Xie and Junfeng Hu

An Intelligent Field-Aware Factorization Machine Model .................. 309
Cairong Yan, Qinglong Zhang, Xue Zhao, and Yongfeng Huang
Query Processing and Optimization (I)

Beyond Skylines: Explicit Preferences .............................................. 327
   Markus Endres and Timotheus Preisinger

Optimizing Window Aggregate Functions in Relational Database Systems . . . 343
   Guangxuan Song, Jiansong Ma, Xiaoling Wang, Cheqing Jin, and Yu Cao

Query Optimization on Hybrid Storage ............................................. 361
   Anxuan Yu, Qingzhong Meng, Xuan Zhou, Binyu Shen, and Yansong Zhang

Efficient Batch Grouping in Relational Datasets .................................. 376
   Jizhou Sun, Jianzhong Li, and Hong Gao

Text Mining

Memory-Enhanced Latent Semantic Model: Short Text Understanding for Sentiment Analysis .......................................................... 393
   Fei Hu, Xiaofei Xu, Jingyuan Wang, Zhanbo Yang, and Li Li

Supervised Intensive Topic Models for Emotion Detection over Short Text . . . 408
   Yanghui Rao, Jianhui Pang, Haoran Xie, An Liu, Tak-Lam Wong, Qing Li, and Fu Lee Wang

Leveraging Pattern Associations for Word Embedding Models .................... 423
   Qian Liu, Heyan Huang, Yang Gao, Xiaochi Wei, and Ruiying Geng

Multi-Granularity Neural Sentence Model for Measuring Short Text Similarity ................................................................. 439
   Jiangping Huang, Shuxin Yao, Chen Lyu, and Donghong Ji

Recommendation

Leveraging Kernel Incorporated Matrix Factorization for Smartphone Application Recommendation ......................................................... 459
   Chenyang Liu, Jian Cao, and Jing He

Preference Integration in Context-Aware Recommendation ......................... 475
   Lin Zheng and Fuxi Zhu

Jointly Modeling Heterogeneous Temporal Properties in Location Recommendation ..................................................... 490
   Saeid Hosseini, Hongzhi Yin, Meihui Zhang, Xiaofang Zhou, and Shazia Sadiq
Location-Aware News Recommendation Using Deep Localized Semantic Analysis .......................... 507
  Cheng Chen, Thomas Lukasiewicz, Xiangwu Meng, and Zhenghua Xu

Review-Based Cross-Domain Recommendation Through Joint Tensor Factorization ....................... 525
  Tianhang Song, Zhaohui Peng, Senzhang Wang, Wenjing Fu, Xiaoguang Hong, and Philip S. Yu

Security, Privacy, Sensor and Cloud

A Local-Clustering-Based Personalized Differential Privacy Framework for User-Based Collaborative Filtering ............... 543
  Yongkai Li, Shubo Liu, Jun Wang, and Mengjun Liu

Fast Multi-dimensional Range Queries on Encrypted Cloud Databases .................................. 559
  Jialin Chi, Cheng Hong, Min Zhang, and Zhenfeng Zhang

  Xiao Liu, An Liu, Xiangliang Zhang, Zhixu Li, Guanfeng Liu, Lei Zhao, and Xiaofang Zhou

Supporting Cost-Efficient Multi-tenant Database Services with Service Level Objectives (SLOs) .................. 592
  Yifeng Luo, Junshi Guo, Jiaye Zhu, Jihong Guan, and Shuigeng Zhou

Recovering Missing Values from Corrupted Spatio-Temporal Sensory Data via Robust Low-Rank Tensor Completion ...... 607
  Wenjie Ruan, Peipei Xu, Quan Z. Sheng, Nickolas J.G. Falkner, Xue Li, and Wei Emma Zhang

Social Network Analytics (I)

Group-Level Influence Maximization with Budget Constraint ........................................ 625
  Qian Yan, Hao Huang, Yunjun Gao, Wei Lu, and Qinming He

Correlating Stressor Events for Social Network Based Adolescent Stress Prediction ................................. 642
  Qi Li, Liang Zhao, Yuanyuan Xue, Li Jin, Mostafa Alli, and Ling Feng

Emotion Detection in Online Social Network Based on Multi-label Learning .................................. 659
  Xiao Zhang, Wenzhong Li, and Sanglu Lu
Tutorials

Urban Computing: Enabling Urban Intelligence with Big Data ................. 677
Yu Zheng

Incentive-Based Dynamic Content Management in Mobile Crowdsourcing
for Smart City Applications .................................................. 680
Sanjay K. Madria

Tutorial on Data Analytics in Multi-engine Environments ................. 682
Verena Kantere and Maxim Filatov

Author Index ................................................................. 685
Contents – Part II

Map Matching and Spatial Keywords

HIMM: An HMM-Based Interactive Map-Matching System .......................... 3
  Xibo Zhou, Ye Ding, Haoyu Tan, Qiong Luo, and Lionel M. Ni

HyMU: A Hybrid Map Updating Framework ................................. 19
  Tao Wang, Jiali Mao, and Cheqing Jin

Multi-objective Spatial Keyword Query with Semantics ....................... 34
  Jing Chen, Jiajie Xu, Chengfei Liu, Zhixu Li, An Liu, and Zhiming Ding

Query Processing and Optimization (II)

RSkycube: Efficient Skycube Computation by Reusing Principle ............ 51
  Kaiqi Zhang, Hong Gao, Xixian Han, Donghua Yang, Zhipeng Cai, and Jianzhong Li

Similarity Search Combining Query Relaxation and Diversification .......... 65
  Ruoxi Shi, Hongzhi Wang, Tao Wang, Yutai Hou, Yiwen Tang, Jianzhong Li, and Hong Gao

An Unsupervised Approach for Low-Quality Answer Detection in Community Question-Answering .................................................. 85
  Haocheng Wu, Zuohui Tian, Wei Wu, and Enhong Chen

Approximate OLAP on Sustained Data Streams ................................. 102
  Salman Ahmed Shaikh and Hiroyuki Kitagawa

Search and Information Retrieval

Hierarchical Semantic Representations of Online News Comments for Emotion Tagging Using Multiple Information Sources ..................... 121
  Chao Wang, Ying Zhang, Wei Jie, Christian Sauer, and Xiaojie Yuan

Towards a Query-Less News Search Framework on Twitter .................... 137
  Xiaotian Hao, Ji Cheng, Jan Vosecky, and Wilfred Ng

Semantic Definition Ranking ......................................................... 153
  Zehui Hao, Zhongyuan Wang, Xiaofeng Meng, Jun Yan, and Qiuyue Wang
An Improved Approach for Long Tail Advertising in Sponsored Search  
*Amar Budhiraja and P. Krishna Reddy*

**String and Sequence Processing**

Locating Longest Common Subsequences with Limited Penalty  
*Bin Wang, Xiaochun Yang, and Jinxu Li*

Top-k String Auto-Completion with Synonyms  
*Pengfei Xu and Jiaheng Lu*

Efficient Regular Expression Matching on Compressed Strings  
*Yutong Han, Bin Wang, Xiaochun Yang, and Huaijie Zhu*

Mining Top-k Distinguishing Temporal Sequential Patterns from Event Sequences  
*Lei Duan, Li Yan, Guozhu Dong, Jyrki Nummenmaa, and Hao Yang*

**Stream Data Processing**

Soft Quorums: A High Availability Solution for Service Oriented Stream Systems  
*Chunyao Song, Tingjian Ge, Cindy Chen, and Jie Wang*

StroMAX: Partitioning-Based Scheduler for Real-Time Stream Processing System  
*Jiawei Jiang, Zhipeng Zhang, Bin Cui, Yunhai Tong, and Ning Xu*

Partition-Based Clustering with Sliding Windows for Data Streams  
*Jonghem Youn, Jihun Choi, Junho Shim, and Sang-goo Lee*

CBP: A New Parallelization Paradigm for Massively Distributed Stream Processing  
*Qingsong Guo and Yongluan Zhou*

**Social Network Analytics (II)**

Measuring and Maximizing Influence via Random Walk in Social Activity Networks  
*Pengpeng Zhao, Yongkun Li, Hong Xie, Zhiyong Wu, Yinlong Xu, and John C.S. Lui*

Adaptive Overlapping Community Detection with Bayesian NonNegative Matrix Factorization  
*Xiaohua Shi, Hongtao Lu, and Guanbo Jia*
A Unified Approach for Learning Expertise and Authority in Digital Libraries ........................................... 354
  B. de La Robertie, L. Ermakova, Y. Pitarch, A. Takasu, and O. Teste

Graph and Network Data Processing

Efficient Local Clustering Coefficient Estimation in Massive Graphs ........ 371
  Hao Zhang, Yuanyuan Zhu, Lu Qin, Hong Cheng, and Jeffrey Xu Yu

Efficient Processing of Growing Temporal Graphs .............................. 387
  Huanhuan Wu, Yunjian Zhao, James Cheng, and Da Yan

Effective $k$-Vertex Connected Component Detection in Large-Scale Networks ........................................... 404
  Yuan Li, Yuhai Zhao, Guoren Wang, Feida Zhu, Yubao Wu, and Shengle Shi

Spatial Databases

Efficient Landmark-Based Candidate Generation for $k$NN Queries on Road Networks ........................................... 425
  Tenindra Abeywickrama and Muhammad Aamir Cheema

MinSum Based Optimal Location Query in Road Networks ................. 441
  Lv Xu, Ganglin Mai, Zitong Chen, Yubao Liu, and Genan Dai

Efficiently Mining High Utility Co-location Patterns from Spatial Data Sets with Instance-Specific Utilities ........................................... 458
  Lizhen Wang, Wanguo Jiang, Hongmei Chen, and Yuan Fang

Real Time Data Processing

Supporting Real-Time Analytic Queries in Big and Fast Data Environments ........................................... 477
  Guangjun Wu, Xiaochun Yun, Chao Li, Shupeng Wang, Yipeng Wang, Xiaoyu Zhang, Siyu Jia, and Guangyan Zhang

Boosting Moving Average Reversion Strategy for Online Portfolio Selection: A Meta-learning Approach ........................................... 494
  Xiao Lin, Min Zhang, Yongfeng Zhang, Zhaoquan Gu, Yiqun Liu, and Shaoping Ma

Continuous Summarization over Microblog Threads ....................... 511
  Liangjun Song, Ping Zhang, Zhifeng Bao, and Timos Sellis

Drawing Density Core-Sets from Incomplete Relational Data ............ 527
  Yongnan Liu, Jianzhong Li, and Hong Gao
Big Data (Industrial)

Co-training an Improved Recurrent Neural Network with Probability Statistic Models for Named Entity Recognition ........................................ 545
  Yueqing Sun, Lin Li, Zhongwei Xie, Qing Xie, Xin Li, and Guandong Xu

EtherQL: A Query Layer for Blockchain System ................................. 556
  Yang Li, Kai Zheng, Ying Yan, Qi Liu, and Xiaofang Zhou

Optimizing Scalar User-Defined Functions in In-Memory Column-Store
Database Systems ................................................................. 568
  Cheol Ryu, Sunho Lee, Kihong Kim, Kunsoo Park, Yongsik Kwon,
  Sang Kyun Cha, Changbin Song, Emanuel Ziegler, and Stephan Muench

GPS-Simulated Trajectory Detection ............................................ 581
  Han Su, Wei Chen, Rong Liu, Min Nie, Bolong Zheng, Zehao Huang,
  and Defu Lian

Social Networks and Graphs (Industrial)

Predicting Academic Performance via Semi-supervised Learning with Constructed Campus Social Network ........................................ 597
  Huaxiu Yao, Min Nie, Han Su, Hu Xia, and Defu Lian

Social User Profiling: A Social-Aware Topic Modeling Perspective ......... 610
  Chao Ma, Chen Zhu, Yanjie Fu, Hengshu Zhu, Guiquan Liu,
  and Enhong Chen

Cost-Effective Data Partition for Distributed Stream Processing System .... 623
  Xiaotong Wang, Junhua Fang, Yuming Li, Rong Zhang,
  and Aoying Zhou

A Graph-Based Push Service Platform ......................................... 636
  Huifeng Guo, Ruiming Tang, Yuming Ye, Zhenguo Li, and Xiuqiang He

Edge Influence Computation in Dynamic Graphs ............................. 649
  Yongrui Qin, Quan Z. Sheng, Simon Parkinson,
  and Nickolas J.G. Falkner

Demos

DKGBuilder: An Architecture for Building a Domain Knowledge Graph from Scratch ............................................................... 663
  Yan Fan, Chengyu Wang, Guomin Zhou, and Xiaofeng He
CLTR: Collectively Linking Temporal Records Across
Heterogeneous Sources ................................................. 668
    Yanyan Zou and Kasun S. Perera

PhenomenaAssociater: Linking Multi-domain Spatio-Temporal Datasets..... 672
    Prathamesh Walkikar and Vandana P. Janeja

VisDM–A Data Stream Visualization Platform ................................. 677
    Lars Melander, Kjell Orsborn, Tore Risch, and Daniel Wedlund

Author Index ................................................................. 681