

Contents – Part I

Semantic Web and Knowledge Management

A General Fine-Grained Truth Discovery Approach for Crowdsourced Data Aggregation	3
<i>Yang Du, Hongli Xu, Yu-E Sun, and Liusheng Huang</i>	
Learning the Structures of Online Asynchronous Conversations	19
<i>Jun Chen, Chaokun Wang, Heran Lin, Weiping Wang, Zhipeng Cai, and Jianmin Wang</i>	
A Question Routing Technique Using Deep Neural Network for Communities of Question Answering	35
<i>Amr Azzam, Neamat Tazi, and Ahmad Hossny</i>	
Category-Level Transfer Learning from Knowledge Base to Microblog Stream for Accurate Event Detection	50
<i>Weijing Huang, Tengjiao Wang, Wei Chen, and Yazhou Wang</i>	

Indexing and Distributed Systems

AngleCut: A Ring-Based Hashing Scheme for Distributed Metadata Management	71
<i>Jiayi Liu, Renxuan Wang, Xiaofeng Gao, Xiaochun Yang, and Guihai Chen</i>	
An Efficient Bulk Loading Approach of Secondary Index in Distributed Log-Structured Data Stores.	87
<i>Yanchao Zhu, Zhao Zhang, Peng Cai, Weining Qian, and Aoying Zhou</i>	
Performance Comparison of Distributed Processing of Large Volume of Data on Top of Xen and Docker-Based Virtual Clusters	103
<i>Haebin Chung and Yunmook Nah</i>	
An Adaptive Data Partitioning Scheme for Accelerating Exploratory Spark SQL Queries	114
<i>Chenghao Guo, Zhigang Wu, Zhenying He, and X. Sean Wang</i>	

Network Embedding

Semi-Supervised Network Embedding	131
<i>Chaozhuo Li, Zhoujun Li, Senzhang Wang, Yang Yang, Xiaoming Zhang, and Jianshe Zhou</i>	

CirE: Circular Embeddings of Knowledge Graphs	148
<i>Zhijuan Du, Zehui Hao, Xiaofeng Meng, and Qiuyue Wang</i>	
PPNE: Property Preserving Network Embedding	163
<i>Chaozhao Li, Senzhang Wang, Dejian Yang, Zhoujun Li, Yang Yang, Xiaoming Zhang, and Jianshe Zhou</i>	
HINE: Heterogeneous Information Network Embedding	180
<i>Yuxin Chen and Chenguang Wang</i>	

Trajectory and Time Series Data Processing

DT-KST: Distributed Top- k Similarity Query on Big Trajectory Streams	199
<i>Zhigang Zhang, Yilin Wang, Jiali Mao, Shaojie Qiao, Cheqing Jin, and Aoying Zhou</i>	
A Distributed Multi-level Composite Index for KNN Processing on Long Time Series	215
<i>Xiaqing Wang, Zicheng Fang, Peng Wang, Ruiyuan Zhu, and Wei Wang</i>	
Outlier Trajectory Detection: A Trajectory Analytics Based Approach	231
<i>Zhongjian Lv, Jiajie Xu, Pengpeng Zhao, Guanfeng Liu, Lei Zhao, and Xiaofang Zhou</i>	
Clustering Time Series Utilizing a Dimension Hierarchical Decomposition Approach	247
<i>QiuHong Li, Peng Wang, Yang Wang, Wei Wang, Yimin Liu, Jiaye Wu, and Danyang Dou</i>	

Data Mining

Fast Extended One-Versus-Rest Multi-label SVM Classification Algorithm Based on Approximate Extreme Points	265
<i>Zhongwei Sun, Zhongwen Guo, Xupeng Wang, Jing Liu, and Shiyong Liu</i>	
Efficiently Discovering Most-Specific Mixed Patterns from Large Data Trees	279
<i>Xiaoying Wu and Dimitri Theodoratos</i>	
Max-Cosine Matching Based Neural Models for Recognizing Textual Entailment	295
<i>Zhipeng Xie and Junfeng Hu</i>	
An Intelligent Field-Aware Factorization Machine Model	309
<i>Cairong Yan, Qinglong Zhang, Xue Zhao, and Yongfeng Huang</i>	

Query Processing and Optimization (I)

Beyond Skylines: Explicit Preferences	327
<i>Markus Endres and Timotheus Preisinger</i>	
Optimizing Window Aggregate Functions in Relational Database Systems . . .	343
<i>Guangxuan Song, Jiansong Ma, Xiaoling Wang, Cheqing Jin, and Yu Cao</i>	
Query Optimization on Hybrid Storage	361
<i>Anxuan Yu, Qingzhong Meng, Xuan Zhou, Binyu Shen, and Yansong Zhang</i>	
Efficient Batch Grouping in Relational Datasets	376
<i>Jizhou Sun, Jianzhong Li, and Hong Gao</i>	

Text Mining

Memory-Enhanced Latent Semantic Model: Short Text Understanding for Sentiment Analysis	393
<i>Fei Hu, Xiaofei Xu, Jingyuan Wang, Zhanbo Yang, and Li Li</i>	
Supervised Intensive Topic Models for Emotion Detection over Short Text . . .	408
<i>Yanghui Rao, Jianhui Pang, Haoran Xie, An Liu, Tak-Lam Wong, Qing Li, and Fu Lee Wang</i>	
Leveraging Pattern Associations for Word Embedding Models	423
<i>Qian Liu, Heyan Huang, Yang Gao, Xiaochi Wei, and Ruiying Geng</i>	
Multi-Granularity Neural Sentence Model for Measuring Short Text Similarity	439
<i>Jiangping Huang, Shuxin Yao, Chen Lyu, and Donghong Ji</i>	

Recommendation

Leveraging Kernel Incorporated Matrix Factorization for Smartphone Application Recommendation	459
<i>Chenyang Liu, Jian Cao, and Jing He</i>	
Preference Integration in Context-Aware Recommendation	475
<i>Lin Zheng and Fuxi Zhu</i>	
Jointly Modeling Heterogeneous Temporal Properties in Location Recommendation	490
<i>Saeid Hosseini, Hongzhi Yin, Meihui Zhang, Xiaofang Zhou, and Shazia Sadiq</i>	

Location-Aware News Recommendation Using Deep Localized Semantic Analysis.	507
<i>Cheng Chen, Thomas Lukasiewicz, Xiangwu Meng, and Zhenghua Xu</i>	

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

Security, Privacy, Sensor and Cloud

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

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

When Differential Privacy Meets Randomized Perturbation: A Hybrid Approach for Privacy-Preserving Recommender System.	576
<i>Xiao Liu, An Liu, Xiangliang Zhang, Zhixu Li, Guanfeng Liu, Lei Zhao, and Xiaofang Zhou</i>	

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

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

Social Network Analytics (I)

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

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

Emotion Detection in Online Social Network Based on Multi-label Learning	659
<i>Xiao Zhang, Wenzhong Li, and Sanglu Lu</i>	

Tutorials

Urban Computing: Enabling Urban Intelligence with Big Data	677
<i>Yu Zheng</i>	
Incentive-Based Dynamic Content Management in Mobile Crowdsourcing for Smart City Applications	680
<i>Sanjay K. Madria</i>	
Tutorial on Data Analytics in Multi-engine Environments	682
<i>Verena Kantere and Maxim Filatov</i>	
Author Index	685

Contents – Part II

Map Matching and Spatial Keywords

HIMM: An HMM-Based Interactive Map-Matching System	3
<i>Xibo Zhou, Ye Ding, Haoyu Tan, Qiong Luo, and Lionel M. Ni</i>	
HyMU: A Hybrid Map Updating Framework	19
<i>Tao Wang, Jiali Mao, and Cheqing Jin</i>	
Multi-objective Spatial Keyword Query with Semantics	34
<i>Jing Chen, Jiajie Xu, Chengfei Liu, Zhixu Li, An Liu, and Zhiming Ding</i>	

Query Processing and Optimization (II)

RSkycube: Efficient Skycube Computation by Reusing Principle	51
<i>Kaiqi Zhang, Hong Gao, Xixian Han, Donghua Yang, Zhipeng Cai, and Jianzhong Li</i>	
Similarity Search Combining Query Relaxation and Diversification	65
<i>Ruoxi Shi, Hongzhi Wang, Tao Wang, Yutai Hou, Yiwen Tang, Jianzhong Li, and Hong Gao</i>	
An Unsupervised Approach for Low-Quality Answer Detection in Community Question-Answering.	85
<i>Haocheng Wu, Zuohui Tian, Wei Wu, and Enhong Chen</i>	
Approximate OLAP on Sustained Data Streams	102
<i>Salman Ahmed Shaikh and Hiroyuki Kitagawa</i>	

Search and Information Retrieval

Hierarchical Semantic Representations of Online News Comments for Emotion Tagging Using Multiple Information Sources	121
<i>Chao Wang, Ying Zhang, Wei Jie, Christian Sauer, and Xiaojie Yuan</i>	
Towards a Query-Less News Search Framework on Twitter	137
<i>Xiaotian Hao, Ji Cheng, Jan Vosecky, and Wilfred Ng</i>	
Semantic Definition Ranking	153
<i>Zehui Hao, Zhongyuan Wang, Xiaofeng Meng, Jun Yan, and Qiuyue Wang</i>	

An Improved Approach for Long Tail Advertising in Sponsored Search	169
<i>Amar Budhiraja and P. Krishna Reddy</i>	

String and Sequence Processing

Locating Longest Common Subsequences with Limited Penalty	187
<i>Bin Wang, Xiaochun Yang, and Jinxu Li</i>	
Top- <i>k</i> String Auto-Completion with Synonyms	202
<i>Pengfei Xu and Jiaheng Lu</i>	
Efficient Regular Expression Matching on Compressed Strings	219
<i>Yutong Han, Bin Wang, Xiaochun Yang, and Huaijie Zhu</i>	
Mining Top- <i>k</i> Distinguishing Temporal Sequential Patterns from Event Sequences	235
<i>Lei Duan, Li Yan, Guozhu Dong, Jyrki Nummenmaa, and Hao Yang</i>	

Stream Data Processing

Soft Quorums: A High Availability Solution for Service Oriented Stream Systems	253
<i>Chunyao Song, Tingjian Ge, Cindy Chen, and Jie Wang</i>	
StroMAX: Partitioning-Based Scheduler for Real-Time Stream Processing System.	269
<i>Jiawei Jiang, Zhipeng Zhang, Bin Cui, Yunhai Tong, and Ning Xu</i>	
Partition-Based Clustering with Sliding Windows for Data Streams	289
<i>Jonghem Youn, Jihun Choi, Junho Shim, and Sang-goo Lee</i>	
CBP: A New Parallelization Paradigm for Massively Distributed Stream Processing	304
<i>Qingsong Guo and Yongluan Zhou</i>	

Social Network Analytics (II)

Measuring and Maximizing Influence via Random Walk in Social Activity Networks	323
<i>Pengpeng Zhao, Yongkun Li, Hong Xie, Zhiyong Wu, Yinlong Xu, and John C.S. Lui</i>	
Adaptive Overlapping Community Detection with Bayesian NonNegative Matrix Factorization	339
<i>Xiaohua Shi, Hongtao Lu, and Guanbo Jia</i>	

A Unified Approach for Learning Expertise and Authority in Digital Libraries	354
<i>B. de La Robertie, L. Ermakova, Y. Pitarch, A. Takasu, and O. Teste</i>	

Graph and Network Data Processing

Efficient Local Clustering Coefficient Estimation in Massive Graphs	371
<i>Hao Zhang, Yuanyuan Zhu, Lu Qin, Hong Cheng, and Jeffrey Xu Yu</i>	
Efficient Processing of Growing Temporal Graphs	387
<i>Huanhuan Wu, Yunjian Zhao, James Cheng, and Da Yan</i>	
Effective k -Vertex Connected Component Detection in Large-Scale Networks	404
<i>Yuan Li, Yuhai Zhao, Guoren Wang, Feida Zhu, Yubao Wu, and Shengle Shi</i>	

Spatial Databases

Efficient Landmark-Based Candidate Generation for k NN Queries on Road Networks	425
<i>Tenindra Abeywickrama and Muhammad Aamir Cheema</i>	
MinSum Based Optimal Location Query in Road Networks	441
<i>Lv Xu, Ganglin Mai, Zitong Chen, Yubao Liu, and Genan Dai</i>	
Efficiently Mining High Utility Co-location Patterns from Spatial Data Sets with Instance-Specific Utilities	458
<i>Lizhen Wang, Wanguo Jiang, Hongmei Chen, and Yuan Fang</i>	

Real Time Data Processing

Supporting Real-Time Analytic Queries in Big and Fast Data Environments	477
<i>Guangjun Wu, Xiaochun Yun, Chao Li, Shupeng Wang, Yipeng Wang, Xiaoyu Zhang, Siyu Jia, and Guangyan Zhang</i>	
Boosting Moving Average Reversion Strategy for Online Portfolio Selection: A Meta-learning Approach	494
<i>Xiao Lin, Min Zhang, Yongfeng Zhang, Zhaoquan Gu, Yiqun Liu, and Shaoping Ma</i>	
Continuous Summarization over Microblog Threads	511
<i>Liangjun Song, Ping Zhang, Zhifeng Bao, and Timos Sellis</i>	
Drawing Density Core-Sets from Incomplete Relational Data	527
<i>Yongnan Liu, Jianzhong Li, and Hong Gao</i>	

Big Data (Industrial)

Co-training an Improved Recurrent Neural Network with Probability Statistic Models for Named Entity Recognition	545
<i>Yueqing Sun, Lin Li, Zhongwei Xie, Qing Xie, Xin Li, and Guandong Xu</i>	
EtherQL: A Query Layer for Blockchain System.	556
<i>Yang Li, Kai Zheng, Ying Yan, Qi Liu, and Xiaofang Zhou</i>	
Optimizing Scalar User-Defined Functions in In-Memory Column-Store Database Systems	568
<i>Cheol Ryu, Sunho Lee, Kihong Kim, Kunsoo Park, Yongsik Kwon, Sang Kyun Cha, Changbin Song, Emanuel Ziegler, and Stephan Muench</i>	
GPS-Simulated Trajectory Detection	581
<i>Han Su, Wei Chen, Rong Liu, Min Nie, Bolong Zheng, Zehao Huang, and Defu Lian</i>	

Social Networks and Graphs (Industrial)

Predicting Academic Performance via Semi-supervised Learning with Constructed Campus Social Network	597
<i>Huaxiu Yao, Min Nie, Han Su, Hu Xia, and Defu Lian</i>	
Social User Profiling: A Social-Aware Topic Modeling Perspective.	610
<i>Chao Ma, Chen Zhu, Yanjie Fu, Hengshu Zhu, Guiquan Liu, and Enhong Chen</i>	
Cost-Effective Data Partition for Distributed Stream Processing System	623
<i>Xiaotong Wang, Junhua Fang, Yuming Li, Rong Zhang, and Aoying Zhou</i>	
A Graph-Based Push Service Platform.	636
<i>Huifeng Guo, Ruiming Tang, Yunming Ye, Zhenguo Li, and Xiuqiang He</i>	
Edge Influence Computation in Dynamic Graphs	649
<i>Yongrui Qin, Quan Z. Sheng, Simon Parkinson, and Nickolas J.G. Falkner</i>	

Demos

DKGBuilder: An Architecture for Building a Domain Knowledge Graph from Scratch.	663
<i>Yan Fan, Chengyu Wang, Guomin Zhou, and Xiaofeng He</i>	

CLTR: Collectively Linking Temporal Records Across Heterogeneous Sources	668
<i>Yanyan Zou and Kasun S. Perera</i>	
PhenomenaAssociater: Linking Multi-domain Spatio-Temporal Datasets.	672
<i>Prathamesh Walkikar and Vandana P. Janeja</i>	
VisDM—A Data Stream Visualization Platform	677
<i>Lars Melander, Kjell Orsborn, Tore Risch, and Daniel Wedlund</i>	
Author Index	681

Database Systems for Advanced Applications

22nd International Conference, DASFAA 2017, Suzhou,
China, March 27-30, 2017, Proceedings, Part I

Candan, S.; Chen, L.; Pedersen, T.B.; Chang, L.; Hua, W.
(Eds.)

2017, XXIII, 688 p. 228 illus., Softcover

ISBN: 978-3-319-55752-6