# Table of Contents – Part II

## Data Mining

Ensemble Pruning: A Submodular Function Maximization Perspective ............................................. 1  
*Chaofeng Sha, Keqiang Wang, Xiaoling Wang, and Aoying Zhou*

Identify and Trace Criminal Suspects in the Crowd Aided by Fast Trajectories Retrieval .................. 16  
*Jianming Lv, Haibiao Lin, Can Yang, Zhiwen Yu, Yinghong Chen, and Miaoyi Deng*

Multi-Output Regression with Tag Correlation Analysis for Effective Image Tagging ........................ 31  
*Hongyun Cai, Zi Huang, Xiaofeng Zhu, Qing Zhang, and Xuefei Li*

The Ranking Based Constrained Document Clustering Method and Its Application to Social Event Detection ........ 47  
*Taufik Sutanto and Richi Nayak*

## Spatio-temporal Data Management

A Skylining Approach to Optimize Influence and Cost in Location Selection .................................... 61  
*Juwei Shi, Hua Lu, Jiaheng Lu, and Chengxuan Liao*

Geo-Social Skyline Queries ................................. 77  
*Tobias Emrich, Maximilian Franzke, Nikos Mamoulis, Matthias Renz, and Andreas Züfle*

Reverse-Nearest Neighbor Queries on Uncertain Moving Object Trajectories .............................. 92  
*Tobias Emrich, Hans-Peter Kriegel, Nikos Mamoulis, Johannes Niedermayer, Matthias Renz, and Andreas Züfle*

Selectivity Estimation of Reverse k-Nearest Neighbor Queries ............................................ 108  
*Michael Steinke, Johannes Niedermayer, and Peer Kröger*

## Graph Data Management

Efficient Sampling Methods for Shortest Path Query over Uncertain Graphs ............................... 124  
*Yurong Cheng, Ye Yuan, Guoren Wang, Baiyou Qiao, and Zhiqiong Wang*
Exploiting Transitive Similarity and Temporal Dynamics for Similarity Search in Heterogeneous Information Networks ................................. 141  
Jiazhen He, James Bailey, and Rui Zhang

Top-k Similarity Matching in Large Graphs with Attributes .................. 156  
Xiaofeng Ding, Jianhong Jia, Jiuyong Li, Jixue Liu, and Hai Jin

On Perspective-Aware Top-k Similarity Search in Multi-relational Networks ................................................................. 171  
Yinglong Zhang, Cuiping Li, Hong Chen, and Likun Sheng

Security, Privacy and Trust

$\rho$-uncertainty Anonymization by Partial Suppression ....................... 188  
Xiao Jia, Chao Pan, Xinhui Xu, Kenny Q. Zhu, and Eric Lo

Access Control for Data Integration in Presence of Data Dependencies ................................................................. 203  
Mehdi Haddad, Jovan Stevovic, Annamaria Chiasera,  
Yannis Velegrakis, and Mohand-Saïd Hacid

Thwarting Passive Privacy Attacks in Collaborative Filtering ............... 218  
Rui Chen, Min Xie, and Laks V.S. Lakshmanan

Privacy-Preserving Schema Reuse ................................................ 234  
Nguyen Quoc Viet Hung, Do Son Thanh, Nguyen Thanh Tam, and  
Karl Aberer

Web and Social Data Management

Any Suggestions? Active Schema Support for Structuring Web Information ................................................................. 251  
Silviu Homoceanu, Felix Getlert, Christian Pek, and Wolf-Tilo Balke

ADI: Towards a Framework of App Developer Inspection .................. 266  
Kai Xing, Di Jiang, Wilfred Ng, and Xiaotian Hao

Novel Community Recommendation Based on a User-Community Total Relation ................................................................. 281  
Qian Yu, Zhiyong Peng, Liang Hong, Bin Liu, and Haiping Peng

User Interaction Based Community Detection in Online Social Networks ................................................................. 296  
Himel Dev, Mohammed Eunus Ali, and Tanzima Hashem
# Keyword Search

Object Semantics for XML Keyword Search ........................................... 311  
*Thuy Ngoc Le, Tok Wang Ling, H.V. Jagadish, and Jiaheng Lu*

Large-Scale Similarity Join with Edit-Distance Constraints .................. 328  
*Chen Lin, Haiyang Yu, Wei Weng, and Xianmang He*

Topical Presentation of Search Results on Database .............................. 343  
*Hao Hu, Mingxi Zhang, Zhenying He, Peng Wang, Wei Wang, and Chengfei Liu*

An Efficient Approach for Mining Top-k High Utility Specialized Query Expansions on Social Tagging Systems ........................................... 361  
*Jia-Ling Koh and I-Chih Chiu*

# Data Stream Management

Novel Techniques to Reduce Search Space in Periodic-Frequent Pattern Mining ................................................................. 377  
*R. Uday Kiran and Masaru Kitsuregawa*

Inferring Road Type in Crowdsourced Map Services ............................. 392  
*Ye Ding, Jiangchuan Zheng, Haoyu Tan, Wuman Luo, and Lionel M. Ni*

Rights Protection for Trajectory Streams ........................................ 407  
*Mingliang Yue, Zhiyong Peng, Kai Zheng, and Yuwei Peng*

Efficient Detection of Emergency Event from Moving Object Data Streams ................................................................. 422  
*Limin Guo, Guangyan Huang, and Zhiming Ding*

# Data Quality

Cost Reduction for Web-Based Data Imputation .................................. 438  
*Zhizu Li, Shuo Shang, Qing Xie, and Xiangliang Zhang*

Incremental Quality Inference in Crowdsourcing ................................ 453  
*Jianhong Feng, Guoliang Li, Henan Wang, and Jianhua Feng*

Repair Diversification for Functional Dependency Violations ............... 468  
*Chu He, Zijing Tan, Qing Chen, Chaofeng Sha, Zhihui Wang, and Wei Wang*
XX Table of Contents – Part II

**Industrial Papers**

BigOP: Generating Comprehensive Big Data Workloads as a Benchmarking Framework ........................................... 483  
Yuqing Zhu, Jianfeng Zhan, Chuliang Weng, Raghunath Nambiar, Jinchao Zhang, Xingzhen Chen, and Lei Wang

A*DAX: A Platform for Cross-Domain Data Linking, Sharing and Analytics .......................................................... 493  
Narayanan Amudha, Gim Guan Chua, Eric Siew Khuan Foo, Shen Tut Goh, Shuqiao Guo, Paul Min Chin Lim, Mun-Thye Mak, Muhammad Cassim Mahmud Munshi, See-Kiong Ng, Wee Siong Ng, and Huayu Wu

Optimizing Database Load and Extract for Big Data Era .................................................. 503  
K.T. Sridhar and M.A. Sakkeer

Web Page Centered Communication System Based on a Physical Property .................................................. 513  
Yuhki Shiraishi, Yukiko Kawai, Jianwei Zhang, and Toyokazu Akiyama

**Demo Papers**

TaxiHailer: A Situation-Specific Taxi Pick-Up Points Recommendation System ........................................... 523  
Leyi Song, Chengyu Wang, Xiaoyi Duan, Bing Xiao, Xiao Liu, Rong Zhang, Xiaofeng He, and Xueqing Gong

Camel: A Journey Group T-Pattern Mining System Based on Instagram Trajectory Data .................................. 527  
Yaxin Yu, Xudong Huang, Xinhua Zhu, and Guoren Wang

Harbinger: An Analyzing and Predicting System for Online Social Network Users’ Behavior .......................... 531  
Rui Guo, Hongzhi Wang, Lucheng Zhong, Jianzhong Li, and Hong Gao

Cloud-Scale Transaction Processing with ParaDB System: A Demonstration .................................................. 535  
Xiaoyan Guo, Yu Cao, Baoyao Zhou, Dong Xiang, and Luyuan Zhao

BSMA-GEN: A Parallel Synthetic Data Generator for Social Media Timeline Structures .......................... 539  
Chengcheng Yu, Fan Xia, Qunyan Zhang, Haixin Ma, Weineng Qian, Mingi Zhou, Cheqing Jin, and Aoying Zhou
A Mobile Log Data Analysis System Based on Multidimensional Data Visualization .......................................................... 543
  Ting Liang, Yu Cao, Min Zhu, Baoyao Zhou, Mingzhao Li, and Qihong Gan

Tutorials

Similarity-Based Analytics for Trajectory Data: Theory, Algorithms and Applications .................................................... 549
  Kai Zheng

Graph Mining Approaches: From Main Memory to Map/Reduce .......... 551
  Sharma Chakravarthy

Crowdsourced Algorithms in Data Management ............................ 553
  Dongwon Lee

Author Index ........................................................................ 555