

# Contents – Part I

## Data Mining

More Efficient Algorithm for Mining Frequent Patterns with Multiple Minimum Supports . . . . .	3
<i>Wensheng Gan, Jerry Chun-Wei Lin, Philippe Fournier-Viger, and Han-Chieh Chao</i>	
Efficient Mining of Uncertain Data for High-Utility Itemsets . . . . .	17
<i>Jerry Chun-Wei Lin, Wensheng Gan, Philippe Fournier-Viger, Tzung-Pei Hong, and Vincent S. Tseng</i>	
An Improved HMM Model for Sensing Data Predicting in WSN . . . . .	31
<i>Zeyu Zhang, Bailong Deng, Siding Chen, and Li Li</i>	
eXtreme Gradient Boosting for Identifying Individual Users Across Different Digital Devices . . . . .	43
<i>Rongwei Song, Siding Chen, Bailong Deng, and Li Li</i>	
Two-Phase Mining for Frequent Closed Episodes . . . . .	55
<i>Guoqiong Liao, Xiaoting Yang, Sihong Xie, Philip S. Yu, and Changxuan Wan</i>	
Effectively Updating High Utility Co-location Patterns in Evolving Spatial Databases. . . . .	67
<i>Xiaoxuan Wang, Lizhen Wang, Junli Lu, and Lihua Zhou</i>	
Mining Top- $k$ Distinguishing Sequential Patterns with Flexible Gap Constraints. . . . .	82
<i>Chao Gao, Lei Duan, Guozhu Dong, Haiqing Zhang, Hao Yang, and Changjie Tang</i>	
A Novel Chinese Text Mining Method for E-Commerce Review Spam Detection . . . . .	95
<i>Xiu Li and Xinwei Yan</i>	

## Spatial and Temporal Databases

Retrieving Routes of Interest Over Road Networks . . . . .	109
<i>Wengen Li, Jiannong Cao, Jihong Guan, Man Lung Yiu, and Shuigeng Zhou</i>	

Semantic-Aware Trajectory Compression with Urban Road Network . . . . .	124
<i>Na Ta, Guoliang Li, Bole Chen, and Jianhua Feng</i>	
Discovering Underground Roads from Trajectories Without Road Network . . .	137
<i>Qiuge Song, Jiali Mao, and Cheqing Jin</i>	
Ridesharing Recommendation: Whether and Where Should I Wait?. . . . .	151
<i>Chengcheng Dai</i>	
Keyword-aware Optimal Location Query in Road Network . . . . .	164
<i>Jinling Bao, Xingshan Liu, Rui Zhou, and Bin Wang</i>	
Point-of-Interest Recommendations by Unifying Multiple Correlations. . . . .	178
<i>Ce Cheng, Jiajin Huang, and Ning Zhong</i>	
Top- <i>k</i> Team Recommendation in Spatial Crowdsourcing . . . . .	191
<i>Dawei Gao, Yongxin Tong, Jieying She, Tianshu Song, Lei Chen, and Ke Xu</i>	
Explicable Location Prediction Based on Preference Tensor Model . . . . .	205
<i>Duoduo Zhang, Ning Yang, and Yuchi Ma</i>	

## Recommender Systems

Random Partition Factorization Machines for Context-Aware Recommendations . . . . .	219
<i>Shaoqing Wang, Cuilan Du, Kankan Zhao, Cuiping Li, Yangxi Li, Yang Zheng, Zheng Wang, and Hong Chen</i>	
A Novel Framework to Process the Quantity and Quality of User Behavior Data in Recommender Systems. . . . .	231
<i>Penghua Yu, Lanfen Lin, and Yuangang Yao</i>	
RankMBPR: Rank-Aware Mutual Bayesian Personalized Ranking for Item Recommendation . . . . .	244
<i>Lu Yu, Ge Zhou, Chuxu Zhang, Junming Huang, Chuang Liu, and Zi-Ke Zhang</i>	
Unsupervised Expert Finding in Social Network for Personalized Recommendation . . . . .	257
<i>Junmei Ding, Yan Chen, Xin Li, Guiquan Liu, Aili Shen, and Xiangfu Meng</i>	
An Approach for Clothing Recommendation Based on Multiple Image Attributes. . . . .	272
<i>Dandan Sha, Daling Wang, Xiangmin Zhou, Shi Feng, Yifei Zhang, and Ge Yu</i>	

SocialFM: A Social Recommender System with Factorization Machines . . . .	286
<i>Juming Zhou, Dong Wang, Yue Ding, and Litian Yin</i>	

Identifying Linked Data Datasets for sameAs Interlinking Using Recommendation Techniques . . . . .	298
<i>Haichi Liu, Ting Wang, Jintao Tang, Hong Ning, Dengping Wei, Songxian Xie, and Peilei Liu</i>	

Query-Biased Multi-document Abstractive Summarization via Submodular Maximization Using Event Guidance. . . . .	310
<i>Rui Sun, Zhenchao Wang, Yafeng Ren, and Donghong Ji</i>	

## Graph Data Management

Inferring Diffusion Network on Incomplete Cascade Data . . . . .	325
<i>Peng Dou, Sizhen Du, and Guojie Song</i>	

Anchor Link Prediction Using Topological Information in Social Networks . . .	338
<i>Shuo Feng, Derong Shen, Tiezheng Nie, Yue Kou, and Ge Yu</i>	

Collaborative Partitioning for Multiple Social Networks with Anchor Nodes . . .	353
<i>Fenglan Li, Anming Ji, Songchang Jin, Shuqiang Yang, and Qiang Liu</i>	

A General Framework for Graph Matching and Its Application in Ontology Matching . . . . .	365
<i>Yuda Zang, Jianyong Wang, and Xuan Zhu</i>	

Internet Traffic Analysis in a Large University Town: A Graphical and Clustering Approach . . . . .	378
<i>Weitao Weng, Kai Lei, Kuai Xu, Xiaoyou Liu, and Tao Sun</i>	

Conceptual Sentence Embeddings . . . . .	390
<i>Yashen Wang, Heyan Huang, Chong Feng, Qiang Zhou, and Jiahui Gu</i>	

Inferring Social Roles of Mobile Users Based on Communication Behaviors. . . . .	402
<i>Yipeng Chen, Hongyan Li, Jinbo Zhang, and Gaoshan Miao</i>	

Sparse Topical Coding with Sparse Groups . . . . .	415
<i>Min Peng, Qianqian Xie, Jiajia Huang, Jiahui Zhu, Shuang Ouyang, Jimin Huang, and Gang Tian</i>	

## Information Retrieval

Differential Evolution-Based Fusion for Results Diversification of Web Search . . . . .	429
<i>Chunlin Xu, Chunlan Huang, and Shengli Wu</i>	

BMF: An Indexing Structure to Support Multi-element Check . . . . .	441
<i>Chenyang Xu, Qin Liu, and Weixiong Rao</i>	
Efficient Unique Column Combinations Discovery Based on Data Distribution . . . . .	454
<i>Chao Wang, Shupeng Han, Xiangrui Cai, Haiwei Zhang, and Yanlong Wen</i>	
Event Related Document Retrieval Based on Bipartite Graph . . . . .	467
<i>Wenjing Yang, Rui Li, Peng Li, Meilin Zhou, and Bin Wang</i>	
SPedia: A Semantics Based Repository of Scientific Publications Data . . . . .	479
<i>Muhammad Ahtisham Aslam and Naif Radi Aljohani</i>	
A Set-Based Training Query Classification Approach for Twitter Search . . . . .	491
<i>Qingli Ma, Ben He, Jungang Xu, and Bin Wang</i>	
NBLucene: Flexible and Efficient Open Source Search Engine . . . . .	504
<i>Zhaohua Zhang, Benjun Ye, Jiayi Huang, Rebecca Stones, Gang Wang, and Xiaoguang Liu</i>	
Context-Aware Entity Summarization . . . . .	517
<i>Jihong Yan, Yanhua Wang, Ming Gao, and Aoying Zhou</i>	
Erratum to: BMF: An Indexing Structure to Support Multi-element Check . . .	E1
<i>Chenyang Xu, Qin Liu, and Weixiong Rao</i>	
<b>Author Index . . . . .</b>	<b>531</b>

## Contents – Part II

### Privacy and Trust

Detecting Anomalous Ratings Using Matrix Factorization for Recommender Systems . . . . .	3
<i>Zhihai Yang, Zhongmin Cai, and Xinyuan Chen</i>	
A Novel Spatial Cloaking Scheme Using Hierarchical Hilbert Curve for Location-Based Services . . . . .	15
<i>Ningning Cui, Xiaochun Yang, and Bin Wang</i>	
Efficient Privacy-Preserving Content-Based Image Retrieval in the Cloud . . .	28
<i>Kai Huang, Ming Xu, Shaojing Fu, and Dongsheng Wang</i>	
Preserving the $d$ -Reachability When Anonymizing Social Networks . . . . .	40
<i>Xiangyu Liu, Jiajia Li, Dahai Zhou, Yunzhe An, and Xiufeng Xia</i>	
Personalized Location Anonymity - A Kernel Density Estimation Approach . . .	52
<i>Dapeng Zhao, Jiansong Ma, Xiaoling Wang, and Xiuxia Tian</i>	
Detecting Data-model-oriented Anomalies in Parallel Business Process . . . . .	65
<i>Ning Yin, Shanshan Wang, Hongyan Li, and Lilue Fan</i>	
Learning User Credibility on Aspects from Review Texts . . . . .	78
<i>Yifan Gao, Yuming Li, Yanhong Pan, Jiali Mao, and Rong Zhang</i>	
Detecting Anomaly in Traffic Flow from Road Similarity Analysis . . . . .	92
<i>Xinran Liu, Xingwu Liu, Yuanhong Wang, Juhua Pu, and Xiangliang Zhang</i>	

### Query Processing and Optimization

PACOKS: Progressive Ant-Colony-Optimization-Based Keyword Search over Relational Databases . . . . .	107
<i>Ziyu Lin, Qian Xue, and Yongxuan Lai</i>	
Enhanced Query Classification with Millions of Fine-Grained Topics . . . . .	120
<i>Qi Ye, Feng Wang, Bo Li, and Zhimin Liu</i>	
A Hybrid Machine-Crowdsourcing Approach for Web Table Matching and Cleaning . . . . .	132
<i>Chunhua Li, Pengpeng Zhao, Victor S. Sheng, Zhixu Li, Guanfeng Liu, Jian Wu, and Zhiming Cui</i>	

An Update Method for Shortest Path Caching with Burst Paths Based on Sliding Windows . . . . .	145
<i>Xiaohua Li, Ning Wang, Kanggui Peng, Xiaochun Yang, and Ge Yu</i>	
Low Overhead Log Replication for Main Memory Database System . . . . .	159
<i>Jinwei Guo, Chendong Zhang, Peng Cai, Minqi Zhou, and Aoying Zhou</i>	
Diversification of Keyword Query Result Patterns. . . . .	171
<i>Cem Aksoy, Ananya Dass, Dimitri Theodoratos, and Xiaoying Wu</i>	
Efficient Approximate Substring Matching in Compressed String . . . . .	184
<i>Yutong Han, Bin Wang, and Xiaochun Yang</i>	
Top-K Similarity Search for Query-By-Humming . . . . .	198
<i>Peipei Wang, Bin Wang, and Shiying Luo</i>	
<b>Social Media</b>	
Restricted Boltzmann Machines for Retweeting Behaviours Prediction. . . . .	213
<i>Xiang Li, Lijuan Xie, Yong Tan, and Qiuli Tong</i>	
Cross-Collection Emotion Tagging for Online News . . . . .	225
<i>Li Yu, Xue Zhao, Chao Wang, Haiwei Zhang, and Ying Zhang</i>	
Online News Emotion Prediction with Bidirectional LSTM . . . . .	238
<i>Xue Zhao, Chao Wang, Zhifan Yang, Ying Zhang, and Xiaojie Yuan</i>	
Learning for Search Results Diversification in Twitter . . . . .	251
<i>Ying Wang, Zhunchen Luo, and Yang Yu</i>	
Adjustable Time-Window-Based Event Detection on Twitter . . . . .	265
<i>Qinyi Wang, Jieying She, Tianshu Song, Yongxin Tong, Lei Chen, and Ke Xu</i>	
User-IBTM: An Online Framework for Hashtag Suggestion in Twitter . . . . .	279
<i>Jia Li and Hua Xu</i>	
Unifying User and Message Clustering Information for Retweeting Behavior Prediction. . . . .	291
<i>Bo Jiang, Jiguang Liang, Ying Sha, Lihong Wang, Zhixin Kuang, Rui Li, and Peng Li</i>	
KPCA-WT: An Efficient Framework for High Quality Microblog Extraction in Time-Frequency Domain. . . . .	304
<i>Min Peng, Xinyuan Dai, Kai Zhang, Guanyin Zeng, Jiahui Zhu, Shuang Ouyang, Qianqian Xie, and Gang Tian</i>	

**Big Data Analytics**

Active Learning Method for Constraint-Based Clustering Algorithms. . . . .	319
<i>Lijun Cai, Tinghao Yu, Tingqin He, Lei Chen, and Meiqi Lin</i>	
An Effective Cluster Assignment Strategy for Large Time Series Data. . . . .	330
<i>Damir Mirzanurov, Waqas Nawaz, JooYoung Lee, and Qiang Qu</i>	
AdaWIRL: A Novel Bayesian Ranking Approach for Personal Big-Hit Paper Prediction . . . . .	342
<i>Chuxu Zhang, Lu Yu, Jie Lu, Tao Zhou, and Zi-Ke Zhang</i>	
Detecting Live Events by Mining Textual and Spatial-Temporal Features from Microblogs . . . . .	356
<i>Zhejun Zheng, Beihong Jin, Yanling Cui, and Qiang Ji</i>	
A Label Correlation Based Weighting Feature Selection Approach for Multi-label Data. . . . .	369
<i>Lu Liu, Jing Zhang, Peipei Li, Yuhong Zhang, and Xuegang Hu</i>	
Valuable Group Trajectory Pattern Mining Directed by Adaptable Value Measuring Model . . . . .	380
<i>Xinyu Huang, Tengjiao Wang, Shun Li, and Wei Chen</i>	
DualPOS: A Semi-supervised Attribute Selection Approach for Symbolic Data Based on Rough Set Theory . . . . .	392
<i>Jianhua Dai, Huifeng Han, Hu Hu, Qinghua Hu, Jinghong Zhang, and Wentao Wang</i>	
Semi-supervised Clustering Based on Artificial Bee Colony Algorithm with Kernel Strategy . . . . .	403
<i>Jianhua Dai, Huifeng Han, Hu Hu, Qinghua Hu, Bingjie Wei, and Yuejun Yan</i>	

**Distributed and Cloud Computing**

HMNRS: A Hierarchical Multi-source Name Resolution Service for the Industrial Internet . . . . .	417
<i>Yang Liu, Guoqiang Fu, and Xinchu Li</i>	
Optimizing Replica Exchange Strategy for Load Balancing in Multitenant Databases. . . . .	430
<i>Teng Liu, Qingzhong Li, Lanju Kong, Lei Liu, and Lizhen Cui</i>	
ERPC: An Edge-Resources Based Framework to Reduce Bandwidth Cost in the Personal Cloud . . . . .	444
<i>Shaoduo Gan, Jie Yu, Xiaoling Li, Jun Ma, Lei Luo, Qingbo Wu, and Shasha Li</i>	

Multidimensional Similarity Join Using MapReduce . . . . .	457
<i>Ye Li, Jian Wang, and Leong Hou U</i>	
Real-Time Logo Recognition from Live Video Streams Using an Elastic Cloud Platform . . . . .	469
<i>Jianbing Ding, Hongyang Chao, and Mansheng Yang</i>	
Profit Based Two-Step Job Scheduling in Clouds . . . . .	481
<i>Shuo Zhang, Li Pan, Shijun Liu, Lei Wu, and Xiangxu Meng</i>	
A Join Optimization Method for CPU/MIC Heterogeneous Systems . . . . .	493
<i>Kailai Zhou, Hong Chen, Hui Sun, Cuiping Li, and Tianzhen Wu</i>	
GFSF: A Novel Similarity Join Method Based on Frequency Vector . . . . .	506
<i>Ziyu Lin, Daowen Luo, and Yongxuan Lai</i>	
Erratum to: Web-Age Information Management (Part I and II) . . . . .	E1
<i>Bin Cui, Nan Zhang, Jianliang Xu, Xiang Lian, and Dexi Liu</i>	

**Demo Papers**

SHMS: A Smart Phone Self-health Management System Using Data Mining. . .	521
<i>Chuanhua Xu, Jia Zhu, Zhixu Li, Jing Xiao, Changqin Huang, and Yong Tang</i>	
MVUC: An Interactive System for Mining and Visualizing Urban Co-locations . . . . .	524
<i>Xiao Wang, Hongmei Chen, and Qing Xiao</i>	
SNExtractor: A Prototype for Extracting Semantic Networks from Web Documents . . . . .	527
<i>Chi Zhang, Yanhua Wang, Chengyu Wang, Wenliang Cheng, and Xiaofeng He</i>	
Crowd-PANDA: Using Crowdsourcing Method for Academic Knowledge Acquisition. . . . .	531
<i>Zhaoan Dong, Jiaheng Lu, and Tok Wang Ling</i>	
LPSMon: A Stream-Based Live Public Sentiment Monitoring System . . . . .	534
<i>Kun Ma, Zijie Tang, Jialin Zhong, and Bo Yang</i>	
DPBT: A System for Detecting Pacemakers in Burst Topics. . . . .	537
<i>Guozhong Dong, Wu Yang, Feida Zhu, and Wei Wang</i>	
CEQA - An Open Source Chinese Question Answer System Based on Linked Knowledge . . . . .	540
<i>Zeyu Du, Yan Yang, Qinming Hu, and Liang He</i>	



OSSRec:An Open Source Software Recommendation System Based on Wisdom of Crowds. . . . .	544
<i>Mengwen Chen, Gang Yin, Chenxi Song, Tao Wang, Cheng Yang, and Huaimin Wang</i>	
<b>Author Index</b> . . . . .	549

Web-Age Information Management

17th International Conference, WAIM 2016, Nanchang,  
China, June 3-5, 2016, Proceedings, Part I

Cui, B.; Zhang, N.; Xu, J.; Lian, X.; Liu, D. (Eds.)

2016, XXI, 534 p. 193 illus., Softcover

ISBN: 978-3-319-39936-2