

# Contents

## Graph and Social Network

An Influence Field Perspective on Predicting User's Retweeting Behavior . . . .	3
<i>Yi Shen, Jianjun Yu, Kejun Dong, Juan Zhao, and Kai Nan</i>	
Realizing Impact Sourcing by Adaptive Gold Questions: A Socially Responsible Measure for Workers' Trustworthiness . . . . .	17
<i>Kinda El Maarry, Ulrich Güntzer, and Wolf-Tilo Balke</i>	
A Coalition Formation Game Theory-Based Approach for Detecting Communities in Multi-relational Networks. . . . .	30
<i>Lihua Zhou, Peizhong Yang, Kevin Lü, Zidong Zhang, and Hongmei Chen</i>	
Individual Influence Maximization via Link Recommendation . . . . .	42
<i>Guowei Ma, Qi Liu, Enhong Chen, and Biao Xiang</i>	
Community Detection Based on Minimum-Cut Graph Partitioning . . . . .	57
<i>Yashen Wang, Heyan Huang, Chong Feng, and Zhirun Liu</i>	
Coherent Topic Hierarchy: A Strategy for Topic Evolutionary Analysis on Microblog Feeds . . . . .	70
<i>Jiahui Zhu, Xuhui Li, Min Peng, Jiajia Huang, Tieyun Qian, Jimin Huang, Jiping Liu, Ri Hong, and Pinglan Liu</i>	
Age Detection for Chinese Users in Weibo . . . . .	83
<i>Li Chen, Tieyun Qian, Fei Wang, Zhenni You, Qingxi Peng, and Ming Zhong</i>	
Batch Mode Active Learning for Networked Data With Optimal Subset Selection . . . . .	96
<i>Haihui Xu, Pengpeng Zhao, Victor S. Sheng, Guanfeng Liu, Lei Zhao, Jian Wu, and Zhiming Cui</i>	

## Information and Knowledge

A Question Answering System Built on Domain Knowledge Base . . . . .	111
<i>Yicheng Liu, Yu Hao, Xiaoyan Zhu, and Jiao Li</i>	
Fast and Accurate Computation of Role Similarity via Vertex Centrality. . . .	123
<i>Longjie Li, Lvjian Qian, Victor E. Lee, Mingwei Leng, Mei Chen, and Xiaoyun Chen</i>	

Multimodal-Based Supervised Learning for Image Search Reranking . . . . .	135
<i>Shengnan Zhao, Jun Ma, and Chaoran Cui</i>	
Novel Word Features for Keyword Extraction . . . . .	148
<i>Yiqun Chen, Jian Yin, Weiheng Zhu, and Shiding Qiu</i>	
A Multi-round Global Performance Evaluation Method for Interactive Image Retrieval . . . . .	161
<i>Jiyi Li</i>	
Resorting Relevance Evidences to Cumulative Citation Recommendation for Knowledge Base Acceleration. . . . .	169
<i>Jingang Wang, Lejian Liao, Dandan Song, Lerong Ma, Chin-Yew Lin, and Yong Rui</i>	
Relevance Search on Signed Heterogeneous Information Network Based on Meta-path Factorization . . . . .	181
<i>Min Zhu, Tianchen Zhu, Zhaohui Peng, Guang Yang, Yang Xu, Senzhang Wang, Xiangwei Wang, and Xiaoguang Hong</i>	
Improving the Effectiveness of Keyword Search in Databases Using Query Logs. . . . .	193
<i>Jing Zhou, Yang Liu, and Ziqiang Yu</i>	
cluTM: Content and Link Integrated Topic Model on Heterogeneous Information Networks . . . . .	207
<i>Qian Wang, Zhaohui Peng, Senzhang Wang, Philip S. Yu, Qingzhong Li, and Xiaoguang Hong</i>	
<b>Recommender Systems</b>	
Learning to Recommend with User Generated Content . . . . .	221
<i>Yueshen Xu, Zhiyuan Chen, Jianwei Yin, Zizheng Wu, and Taojun Yao</i>	
Combining Positive and Negative Feedbacks with Factored Similarity Matrix for Recommender Systems . . . . .	233
<i>Mengshuang Wang, Jun Ma, Shanshan Huang, and Peizhe Cheng</i>	
Review Comment Analysis for Predicting Ratings . . . . .	247
<i>Rong Zhang, Yifan Gao, Wenzhe Yu, Pingfu Chao, Xiaoyan Yang, Ming Gao, and Aoying Zhou</i>	
Adaptive Temporal Model for IPTV Recommendation . . . . .	260
<i>Yan Yang, Qinmin Hu, Liang He, Minjie Ni, and Zhijin Wang</i>	

RPCV: Recommend Potential Customers to Vendors in Location-Based Social Network. . . . .	272
<i>Yuanliu Liu, Pengpeng Zhao, Victor S. Sheng, Zhixu Li, An Liu, Jian Wu, and Zhiming Cui</i>	
Mining Dependencies Considering Time Lag in Spatio-Temporal Traffic Data . . . . .	285
<i>Xiabing Zhou, Haikun Hong, Xingxing Xing, Wenhao Huang, Kaigui Bian, and Kunqing Xie</i>	
Location Semantics Protection Based on Bayesian Inference . . . . .	297
<i>Zhengang Wu, Zhong Chen, Jiawei Zhu, Huiping Sun, and Zhi Guan</i>	
<b>Big Data</b>	
SALA: A Skew-Avoiding and Locality-Aware Algorithm for MapReduce-Based Join . . . . .	311
<i>Ziyu Lin, Minxing Cai, Ziming Huang, and Yongxuan Lai</i>	
Energy-Proportional Query Processing on Database Clusters . . . . .	324
<i>Jia Zhuang Xie, Peiquan Jin, Shouhong Wan, and Lihua Yue</i>	
SparkRDF: In-Memory Distributed RDF Management Framework for Large-Scale Social Data . . . . .	337
<i>Zhichao Xu, Wei Chen, Lei Gai, and Tengjiao Wang</i>	
Distributed Grid-Based K Nearest Neighbour Query Processing Over Moving Objects . . . . .	350
<i>Min Yang, Yang Liu, and Ziqiang Yu</i>	
An Efficient Block Sampling Strategy for Online Aggregation in the Cloud. . . . .	362
<i>Xiang Ci and Xiaofeng Meng</i>	
Computing Probability Threshold Set Similarity on Probabilistic Sets. . . . .	374
<i>Lei Wang, Ming Gao, Rong Zhang, Cheqing Jin, and Aoying Zhou</i>	
Region-aware Top-k Similarity Search . . . . .	387
<i>Sitong Liu, Jianhua Feng, and Yongwei Wu</i>	
A Partition-Based Bi-directional Filtering Method for String Similarity JOINS. . . . .	400
<i>Ying Huang, Baoning Niu, and Chunhua Song</i>	
Fast Multiway Maximum Margin Clustering Based on Genetic Algorithm via the Nyström Method . . . . .	413
<i>Ying Kang, Dong Zhang, Bo Yu, Xiaoyan Gu, Weiping Wang, and Dan Meng</i>	

**Short Papers**

DTMF: A User Adaptive Model for Hybrid Recommendation . . . . .	429
<i>Wenlong Yang, Jun Ma, and Shanshan Huang</i>	
An Adaptive Skew Handling Join Algorithm for Large-scale Data Analysis . . . . .	433
<i>Di Wu, Tengjiao Wang, Yuxin Chen, Shun Li, Hongyan Li, and Kai Lei</i>	
NokeaRM: Employing Non-key Attributes in Record Matching . . . . .	438
<i>Qiang Yang, Zhixu Li, Jun Jiang, Pengpeng Zhao, Guanfeng Liu, An Liu, and Jia Zhu</i>	
Efficient Foreign Key Discovery Based on Nearest Neighbor Search . . . . .	443
<i>Xiaojie Yuan, Xiangrui Cai, Man Yu, Chao Wang, Ying Zhang, and Yanlong Wen</i>	
OntoEvent: An Ontology-Based Event Description Language for Semantic Complex Event Processing. . . . .	448
<i>Meng Ma, Ping Wang, Jun Yang, and Chao Li</i>	
Differential Trust Propagation with Community Discovery for Link-Based Web Spam Demotion . . . . .	452
<i>Xianchao Zhang, Yafei Feng, Hua Shen, and Wenxin Liang</i>	
Inferring User Preference in Good Abandonment from Eye Movements . . . .	457
<i>Wanxuan Lu and Yunde Jia</i>	
Private Range Queries on Outsourced Databases . . . . .	461
<i>Lu Li, Liusheng Huang, An Liu, Yao Shen, Wei Yang, and Shengnan Shao</i>	
Efficient Influence Maximization Based on Three Degrees of Influence Theory. . . . .	465
<i>Yadong Qin, Jun Ma, and Shuai Gao</i>	
Bichromatic Reverse $k$ NN Query Algorithm on Road Network Distance . . . .	469
<i>Tin Nilar Win, Htoo Htoo, and Yutaka Ohsawa</i>	
IIRS: A Novel Framework of Identifying Commodity Entities on E-commerce Big Data. . . . .	473
<i>Qiqing Fang, Yamin Hu, Shujun Lv, Lejiang Guo, Lei Xiao, and Yahui Hu</i>	
Extracting Appraisal Expressions from Short Texts. . . . .	481
<i>Peiquan Jin, Yongbo Yu, Jie Zhao, and Lihua Yue</i>	
Friendship Link Recommendation Based on Content Structure Information . . .	486
<i>Xiaoming Zhang, Qiao Deng, and Zhoujun Li</i>	

Overlapping Community Detection in Directed Heterogeneous Social Network. . . . .	490
<i>Changhe Qiu, Wei Chen, Tengjiao Wang, and Kai Lei</i>	
Efficient MapReduce-Based Method for Massive Entity Matching . . . . .	494
<i>Pingfu Chao, Zhu Gao, Yuming Li, Junhua Fang, Rong Zhang, and Aoying Zhou</i>	
The Role of Physical Location in Our Online Social Networks . . . . .	498
<i>Jia Zhu, Pui Cheong Gabriel Fung, Kam-fai Wong, Binyang Li, Zhixu Li, and Haoye Dong</i>	
Hierarchical Community Evolution Mining from Dynamic Networks . . . . .	502
<i>Yonghui Zhang, Chuan Li, Yanmei Li, Changjie Tang, and Ning Yang</i>	
Exploiting Conceptual Relations of Sentences for Multi-document Summarization . . . . .	506
<i>Hai-Tao Zheng, Shu-Qin Gong, Ji-Min Guo, and Wen-Zhen Wu</i>	
Towards Narrative Information Systems . . . . .	511
<i>Philipp Wille, Christoph Lof, and Wolf-Tilo Balke</i>	
Maximizing the Spread of Competitive Influence in a Social Network Oriented to Viral Marketing. . . . .	516
<i>Hong Wu, Weiyi Liu, Kun Yue, Weipeng Huang, and Ke Yang</i>	
SEMR: Secure and Efficient Multi-dimensional Range Query Processing in Two-tiered Wireless Sensor Networks . . . . .	520
<i>Lei Dong, Jianxiang Zhu, Xiaoying Zhang, Hong Chen, Cuiping Li, and Hui Sun</i>	
A Sampling-Based Framework for Crowdsourced Select Query with Multiple Predicates . . . . .	525
<i>Jianhong Feng, Huiqi Hu, Xueping Weng, Jianhua Feng, and Yongwei Wu</i>	
A Customized Schema Design Framework for Multi-tenant Database . . . . .	530
<i>Jiacai Ni, Jianhua Feng, and Yongwei Wu</i>	
A Multi-attribute Probabilistic Matrix Factorization Model for Personalized Recommendation . . . . .	535
<i>Feng Tan, Li Li, Zeyu Zhang, and Yunlong Guo</i>	
A Novel Knowledge Extraction Framework for Resumes Based on Text Classifier . . . . .	540
<i>Jie Chen, Zhendong Niu, and Hongping Fu</i>	

SimRank Based Top-k Query Aggregation for Multi-Relational Networks. . .	544
<i>Jing Xu, Cuiping Li, Hong Chen, and Hui Sun</i>	
On Dynamic Top-k Influence Maximization . . . . .	549
<i>Hao Wang, Nana Pan, U. Leong Hou, Bohan Zhan, and Zhiguo Gong</i>	
Community Based Spammer Detection in Social Networks . . . . .	554
<i>Dehai Liu, Benjin Mei, Jinchuan Chen, Zhiwu Lu, and Xiaoyong Du</i>	
LDM: A DTD Schema Mapping Language Based on Logic Patterns . . . . .	559
<i>Xuhui Li, Yijun Guan, Mengchi Liu, Rui Cai, Ming Zhong, and Tieyun Qian</i>	
Effective Sampling of Points of Interests on Maps Based on Road Networks . . . . .	563
<i>Ziting Zhou, Pengpeng Zhao, Victor S. Sheng, Jiajie Xu, Zhixu Li, Jian Wu, and Zhiming Cui</i>	
Associated Index for Big Structured and Unstructured Data. . . . .	567
<i>Chunying Zhu, Qingzhong Li, Lanju Kong, Xiangwei Wang, and Xiaoguang Hong</i>	
<b>Demo Papers</b>	
Web Knowledge Base Improved OCR Correction for Chinese Business Cards . . . . .	573
<i>Xiaoping Wang, Yanghua Xiao, and Wei Wang</i>	
Shortest Path and Word Vector Based Relation Representation and Clustering . . . . .	577
<i>Xiaoping Wang, Yanghua Xiao, and Wei Wang</i>	
CrowdSR: A Crowd Enabled System for Semantic Recovering of Web Tables . . . . .	581
<i>Huaxi Liu, Ning Wang, and Xiangran Ren</i>	
A Personalized News Recommendation System Based on Tag Dependency Graph . . . . .	584
<i>Pengqiang Ai, Yingyuan Xiao, Ke Zhu, Hongya Wang, and Ching-Hsien Hsu</i>	
SmartInt: A Demonstration System for the Interaction Between Schema Mapping and Record Matching. . . . .	587
<i>Jun Jiang, Zhixu Li, Qiang Yang, Pengpeng Zhao, Guanfeng Liu, and Lei Zhao</i>	

CDSG: A Community Detection System Based on the Game Theory . . . . . 590  
*Peizhong Yang, Lihua Zhou, Lizhen Wang, Xuguang Bao,  
and Zidong Zhang*

Erratum to: Web-Age Information Management . . . . . E1  
*Xin Luna Dong, Xiaohui Yu, Jian Li, and Yizhou Sun*

**Author Index** . . . . . 593

Web-Age Information Management

16th International Conference, WAIM 2015, Qingdao,

China, June 8-10, 2015. Proceedings

Dong, X.L.; Yu, X.; Li, J.; SUN, Y. (Eds.)

2015, XIX, 596 p. 209 illus., Softcover

ISBN: 978-3-319-21041-4