Table of Contents

Session 1: Data Mining (1)

Frequent Subgraph Summarization with Error Control 1
Zheng Liu, Ruoming Jin, Hong Cheng, and Jeffrey Xu Yu

LSA-PTM: A Propagation-Based Topic Model Using Latent Semantic Analysis on Heterogeneous Information Networks 13
Qian Wang, Zhaohui Peng, Fei Jiang, and Qingzhong Li

Improving Semi-supervised Text Classification by Using Wikipedia Knowledge 25
Zhilin Zhang, Huaizhong Lin, Pengfei Li, Huazhong Wang, and Dongming Lu

Time Series Representation: A Random Shifting Perspective 37
Xue Bai, Yun Xiong, Yangyong Zhu, and Hengshu Zhu

Mining Frequent Itemsets from Sparse Data Streams in Limited Memory Environments 51
Juan J. Cameron, Alfredo Cuzzocrea, Fan Jiang, and Carson K. Leung

Human Dynamics Revealed through Log Analytics in a Cloud Computing Environment 58
Sixi Chen, Ning An, Lian Li, Yongwei Wu, Weimin Zheng, and Lin Sun

Session 2: Information Integration and Heterogeneous Systems

Data Fusion: Resolving Conflicts from Multiple Sources 64
Xin Luna Dong, Laure Berti-Equille, and Divesh Srivastava

Entity Resolution on Uncertain Relations 77
Huabin Feng, Hongzhi Wang, Jianzhong Li, and Hong Gao

Imputation for Categorical Attributes with Probabilistic Reasoning 87
Lian Jin, Hongzhi Wang, and Hong Gao

A New Approach to Identify Influential Spreaders in Complex Networks 99
Qingcheng Hu, Yang Gao, Pengfei Ma, Yanshen Yin, Yong Zhang, and Chunxiao Xing

Bibliografische Informationen
http://d-nb.info/1033976806

digitalisiert durch DEUTSCHE NATIONAL BIBLIOTHEK

Presenting XML Schema Mapping with Conjunctive-Disjunctive Views 105
Xuhui Li, Shanfeng Zhu, Mengchi Liu, and Ming Zhong

An ETL Framework for Online Analytical Processing of Linked Open Data 111
Hiroyuki Inoue, Toshiyuki Amagasa, and Hiroyuki Kitagawa

Session 3: Big Data

A Framework for Analyzing Monetary Cost of Database Systems in the Cloud 118
Changbing Chen and Bingsheng He

Efficient Distributed Multi-dimensional Index for Big Data Management 130
Xin Zhou, Xiao Zhang, Yanhao Wang, Rui Li, and Shan Wang

Fast Multi-fields Query Processing in Bigtable Based Cloud Systems ... 142
Haiping Wang, Xiang Ci, and Xiaofeng Meng

ST-HBase: A Scalable Data Management System for Massive Geo-tagged Objects 155
Youzhong Ma, Yu Zhang, and Xiaofeng Meng

EasyControl: Improve Database Throughput under Overloading 167
Chengkai Tao, Qizhi Liu, and Yan Zhang

Combination of In-Memory Graph Computation with MapReduce: A Subgraph-Centric Method of PageRank 173
Qiuhong Li, Wei Wang, Peng Wang, Ke Dai, Zhihui Wang, Yang Wang, and Weiwei Sun

Session 4: Spatial and Temporal Databases

A Human-Machine Method for Web Table Understanding 179
Guoliang Li

Probabilistic k-Skyband Operator over Sliding Windows 190
Xing Feng, Wenjie Zhang, Xiang Zhao, Ying Zhang, and Yunjun Gao

SATURN: A Fast Keyword kNN Search System in Road Networks 203
Nan Zhang, Yang Wang, and Jianhua Feng

Top-K Aggregate Queries on Continuous Probabilistic Datasets 216
Jianwen Chen, Ling Feng, and Jun Zhang

Shortest Path Finder with Light Materialized Path View for Location Based Services 229
Aye Thida Hlaing, Htoo Htoo, Yutaka Ohsawa, Noboru Sonehara, and Masao Sakauchi

RUM+-tree: A New Multidimensional Index Supporting Frequent Updates 235
Yuean Zhu, Shan Wang, Xuan Zhou, and Yansong Zhang

Session 5: Data Mining (2)

Joint Clustering and Feature Selection 241
Liang Du and Yi-Dong Shen

A Self-Supervised Framework for Clustering Ensemble 253
Liang Du, Yi-Dong Shen, Zhiyong Shen, Jianying Wang, and Zhiwu Xu

GCBN: A Hybrid Spatio-temporal Causal Model for Traffic Analysis and Prediction 265
Chaogui Zhang and Jiangtao Ren

An Overlapped Community Partition Algorithm Based on Line Graph 277
Zhenyu Zhang, Zhen Zhang, Wenzhong Yang, and Xiaohong Wu

Distance-Based Feature Selection from Probabilistic Data 282
Tingting Zhao, Bin Pei, Suyun Zhao, Hong Chen, and Cuiping Li

Session 6: Data Mining (3)

Aspect-Specific Polarity-Aware Summarization of Online Reviews 289
Gaoyan Ou, Wei Chen, Peng Liu, Tengjiao Wang, Dongqing Yang, Kai Lei, and Yueqin Liu

Fast Top-k Distance-Based Outlier Detection on Uncertain Data 301
Salman Ahmed Shaikh and Hiroyuki Kitagawa

A Metric Learning Based Approach to Evaluate Task-Specific Time Series Similarity 314
Yang Lu, Wayne Xin Zhao, Hongfei Yan, and Xiaoming Li

Using Coalitional Games to Detect Communities in Social Networks 326
Lihua Zhou, Chao Cheng, Kevin Lü, and Hongmei Chen

A Comparison Study of Clustering Models for Online Review Sentiment Analysis 332
Baojun Ma, Hua Yuan, and Qiang Wei

Session 7: Information Extraction

TopicDSDR: Combining Topic Decomposition and Data Reconstruction for Summarization 338
Zhiming Zhang, Hongjie Li, and Lian'en Huang

CUVIM: Extracting Fresh Information from Social Network 351
Rui Guo, Hongzhi Wang, Kaiyu Li, Jianzhong Li, and Hong Gao

Comments-Oriented Document Summarization Based on Multi-aspect Co-feedback Ranking 363
Lifu Huang, Hongjie Li, and Lian'en Huang

Image Annotation with Weak Labels 375
Feng Tian and Xukun Shen

Session 8: New Hardware and Miscellaneous

Scan and Join Optimization by Exploiting Internal Parallelism of Flash-Based Solid State Drives 381
Wenyu Lai, Yulei Fan, and Xiaofeng Meng

MixSL: An Efficient Transaction Recovery Model in Flash-Based DBMS 393
Yulei Fan and Xiaofeng Meng

Keyword Oriented Bitmap Join Index for In-Memory Analytical Processing 405
Yansong Zhang, Mingchuan Su, Xuan Zhou, Shan Wang, and Xue Wang

Can SSDs Help Reduce Random I/Os in Hash Joins? 417
Liang Huai Yang, Mingchao Liu, Yifan Pan, Weihua Gong, and Simon Stannus

An Index Model for Multitenant Data Storage in SaaS 423
Pang Cheng, Li Qingzhong, and Kong Lanju

Session 9: Query Processing and Optimization

Reverse Top-k Group Nearest Neighbor Search 429
Tao Jiang, Yunjun Gao, Bin Zhang, Qing Liu, and Lu Chen

Efficient Subsequence Search in Databases 440
Rohit Jain, Mukesh K. Mohania, and Sunil Prabhakar

A Novel Data Broadcast Strategy for Traffic Information Query in the VANETs 453
Xinjing Wang, Longjiang Guo, Jinbao Li, and Meirui Ren

A Spatial Proximity Based Compression Method for GML Documents 465
Qingting Wei and Jihong Guan

Efficient MSubtree Results Computation for XML Keyword Queries 472
Junfeng Zhou, Jinjia Zhao, Bo Wang, Xingmin Zhao, and Ziyang Chen

Session 10: Data Mining (4)

Measuring and Visualizing Interest Similarity between Microblog Users 478
Jiayu Tang, Zhiyuan Liu, and Maosong Sun

Correlation Range Query 490
Wenjun Zhou and Hao Zhang

Real Time Event Detection in Twitter 502
Xun Wang, Feida Zhu, Jing Jiang, and Sujian Li

A Data Cleaning Framework Based on User Feedback 514
Hui Xie, Hongzhi Wang, Jianzhong Li, and Hong Gao

Session 11: Social Network and Graphs

Finding Critical Blocks of Information Diffusion in Social Networks 521
Ende Zhang, Guoren Wang, Kening Gao, and Ge Yu

Guide Query in Social Networks 533
Yu-Chieh Lin, Philip S. Yu, and Ming-Syan Chen

Probabilistic Graph Summarization 545
Nasrin Hassanlou, Maryam Shoaran, and Alex Thomo

Blind Chance: On Potential Trust Friends Query in Mobile Social Networks 557
Jinzeng Zhang and Xiaofeng Meng

Sensitive Edges Protection in Social Networks 564
Liangwen Yu, Tao Yang, Zhengang Wu, Jiawei Zhu, Jianbin Hu, and Zhong Chen

Session 12: Information Retrieval

Ontology-Based Semantic Search for Large-Scale RDF Data 570
Xiaolong Tang, Xin Wang, Zhiyong Feng, and Longxiang Jiang

A Unified Generative Model for Characterizing Microblogs' Topics 583
Kun Zhuang, Heyan Huang, Xin Xin, Xiaochi Wei, Xianxiang Yang, Chong Feng, and Ying Fang

A Novel Model for Medical Image Similarity Retrieval 595
Pengyuan Li, Haiwei Pan, Jianzhong Li, Qilong Han, Xiaoqin Xie, and Zhiqiang Zhang

Finding Similar Questions with Categorization Information and Dependency Syntactic Tree 607
Xin Lian, Xiaojie Yuan, Xiangyu Hu, and Haiwei Zhang

Ranking Web Pages by Associating Keywords with Locations 613
Peiquan Jin, Xiaoxiang Zhang, Qingqing Zhang, Sheng Lin, and Lihua Yue

Session 13: Workflow Systems and Service Computing

Behavioral Consistency Measurement and Analysis of WS-BPEL Processes 619
Xuewei Zhang, Wei Song, Jianchun Xing, Qiliang Yang, Hongda Wang, and Wenjia Zhang

Efficient Complex Event Processing under Boolean Model 631
Shanglian Peng, Tianrui Li, Hongjun Wang, Jia He, Tiejun Wang, Fan Li, and An Liu

Bounding Trust under Uncertain Topology Information in Reputation-Based Trust Systems 643
Xi Gong, Ting Yu, and Adam J. Lee

A Novel Service Selection Based on Resource-Directive Decomposition 649
Dongming Jiang, Yuan Jiang, Wuping Xie, and Zhen You

Session 14: Recommender Systems

Collaborative Filtering Based on Rating Psychology 655
Haijun Zhang, Chunyang Liu, Zhoujun Li, and Xiaoming Zhang

Modeling Semantic and Behavioral Relations for Query Suggestion 666
Jimeng Chen, Yuan Wang, Jie Liu, and Yalou Huang

Mining User Interest and Its Evolution for Recommendation on the Micro-blogging System 679
Jianjun Yu, Yi Shen, and Jianjun Xie

Collaborative Filtering Using Multidimensional Psychometrics Model ... 691
Haijun Zhang, Xiaoming Zhang, Zhoujun Li, and Chunyang Liu

Incorporating Social Actions into Recommender Systems 698
Di Ma, Dandan Song, and Lejian Liao

Session 15: Security, Privacy, and Trust

DiffR-Tree: A Differentially Private Spatial Index for OLAP Query 705
Miao Wang, Xiaojian Zhang, and Xiaofeng Meng

K-core-preferred Attack to the Internet: Is it More Malicious Than Degree Attack? 717
Jichang Zhao, Junjie Wu, Mingming Chen, Zhiwen Fang, and Ke Xu

Fuzzy Keyword Search over Encrypted Data in the Public Key Setting 729
Qiuxiang Dong, Zhi Guan, Liang Wu, and Zhong Chen

The Hardness of (ϵ, m)-anonymity 741
Yujia Li, Dong Li, Xianmang He, Wei Wang, and Huahui Chen

Session 16: Semantic Web and Ontology

Annotating Web Images by Combining Label Set Relevance with Correlation 747
Feng Tian and Xukun Shen

Clustering Analysis and Semantics Annotation of 3D Models Based on Users' Implicit Feedbacks 757
Lv Tianyang, Huang Shaobin, Wu Peng, and Lang Dapeng

bCATE: A Balanced Contention-Aware Transaction Execution Model for Highly Concurrent OLTP Systems 769
Xiaogang Shi, Yanfei Lv, Yingxia Shao, and Bin Cui

Personalized News Recommendation Using Ontologies Harvested from the Web 781
Junyang Rao, Aixia Jia, Yansong Feng, and Dongyan Zhao

TinyQP: A Query Processing System in Wireless Sensor Networks 788
Shangfeng Mo, Hong Chen, Xiaoying Zhang, and Cuiping Li

CWePS: Chinese Web People Search 792
Jianhua Yin and Lili Jiang

CONCERT: A Concept-Centric Web News Recommendation System ... 796
Hongda Ren and Wei Feng

DOI Proxy Framework for Automated Entering and Validation of Scientific Papers 799
Kun Ma, Bo Yang, and Guangwei Chen

Effectively Return Query Results for Keyword Search on XML Data 802
Teng He, Guoqing Wu, Qingsong Chen, Junfeng Zhou, and Ziyang Chen

BrackitMR: Flexible XQuery Processing in MapReduce 806
Caetano Sauer, Sebastian Bächle, and Theo Härder

Author Index 809