## Contents – Part II

### Sentiment Analysis

- Dynamic Topic-Based Sentiment Analysis of Large-Scale Online News
  
  *Peng Liu, Jon Atle Gulla, and Lemei Zhang*  
  
- Improving Object and Event Monitoring on Twitter Through Lexical Analysis and User Profiling
  
  *Yihong Zhang, Claudia Szabo, and Quan Z. Sheng*  
  
- Aspect-Based Sentiment Analysis Using Lexico-Semantic Patterns
  
  *Kim Schouten, Frederique Baas, Olivier Bus, Alexander Osinga, Nikki van de Ven, Steffie van Loenhout, Lisanne Vrolijk, and Flavius Frasincar*  
  
- Multilevel Browsing of Folksonomy-Based Digital Collections
  
  *Joaquin Gayoso-Cabada, Daniel Rodriguez-Cerezo, and Jose-Luis Sierra*  

### Ranking in Social Networks

- Faderank: An Incremental Algorithm for Ranking Twitter Users
  
  *Massimo Bartoletti, Stefano Lande, and Alessandro Massa*  
  
- Personalized Re-ranking of Tweets
  
  *Yukun Zhao, Shangsong Liang, and Jun Ma*  
  
- Ranking Microblog Users via URL Biased Posts
  
  *Yongjun Ye, Peng Li, Rui Li, Meilin Zhou, Yifang Wan, and Bin Wang*  
  
- Identifying Implicit Enterprise Users from the Imbalanced Social Data
  
  *Zhenni You, Tieyun Qian, Baochao Zhang, and Shi Ying*  

### Microblog Data Analysis

- Understanding Factors That Affect Web Traffic via Twitter
  
  *Chunjing Xiao, Zhiguang Qin, Xucheng Luo, and Aleksandar Kuzmanovic*  
  
- Analysis of Teens’ Chronic Stress on Micro-blog
  
  *Yuanyuan Xue, Qi Li, Liang Zhao, Jia Jia, Ling Feng, Feng Yu, and David A. Clifton*
Large-Scale Stylistic Analysis of Formality in Academia and Social Media
Thin Nguyen, Svetha Venkatesh, and Dinh Phung

137

Discriminative Cues for Different Stages of Smoking Cessation in Online Community
Thin Nguyen, Ron Borland, John Yearwood, Hua-Hie Yong, Svetha Venkatesh, and Dinh Phung

146

Query Processing

POL: A Pattern Oriented Load-Shedding for Semantic Data Stream Processing
Fethi Belghaouti, Amel Bouzeghoub, Zakia Kazi-Aoul, and Raja Chiky

157

Unsupervised Blocking of Imbalanced Datasets for Record Matching
Chenxiao Dou, Daniel Sun, and Raymond K. Wong

172

Partially Decompressing Binary Interpolative Coding for Fast Query Processing
Xi Fu, Peng Li, Rui Li, and Bin Wang

187

Spatial and Temporal Data

Graph-Based Metric Embedding for Next POI Recommendation
Min Xie, Hongzhi Yin, Fanjiang Xu, Hao Wang, and Xiaofang Zhou

207

Temporal Pattern Based QoS Prediction
Liang Chen, Haochao Ying, Qibo Qiu, Jian Wu, Hai Dong, and Athman Bouguettaya

223

Searching for Data Sources for the Semantic Enrichment of Trajectories
Luiz André P. Paes Leme, Chiara Renso, Bernardo P. Nunes, Giseli Rabello Lopes, Marco A. Casanova, and Vânia P. Vidal

238

On Impact of Weather on Human Mobility in Cities
Jun Pang, Polina Zablotskaia, and Yang Zhang

247

Graph Theory

Minimum Spanning Tree on Uncertain Graphs
Anzhen Zhang, Zhaonian Zou, Jianzhong Li, and Hong Gao

259
A Block-Based Edge Partitioning for Random Walks Algorithms over Large Social Graphs .................................................. 275
   Yifan Li, Camelia Constantin, and Cedric du Mouza

Differentially Private Network Data Release via Stochastic Kronecker Graph .............................................................. 290
   Dai Li, Wei Zhang, and Yunfang Chen

An Executable Specification for SPARQL ........................................ 298
   Mihaela Bornea, Julian Dolby, Achille Fokoue,
   Anastasios Kementsietsidis, Kavitha Srinivas, and Mandana Vaziri

Non-traditional Environments

Towards a Scalable Framework for Artifact-Centric Business Process Management Systems ............................................. 309
   Jiankun Lei, Rufan Bai, Lipeng Guo, and Liang Zhang

Bridging Semantic Gap Between App Names: Collective Matrix Factorization for Similar Mobile App Recommendation .......... 324
   Ning Bu, Shuzi Niu, Lei Yu, Wenjing Ma, and Guoping Long

Summarizing Multimedia Content ............................................. 340
   Natwar Modani, Pranav Maneriker, Gaurush Hiranandani,
   Atanu R. Sinha, Upal, Vaishnavi Subramanian, and Shivani Gupta

Bridging Enterprise and Software Engineering Through an User-Centered Design Perspective ......................................... 349
   Pedro Valente, Thiago Silva, Marco Winckler, and Nuno Nunes

Special Session on Data Quality and Trust in Big Data

Region Profile Based Geo-Spatial Analytic Search .......................... 361
   Xiaofeng Du and Zhan Cui

Segmentation and Enhancement of Low Quality Fingerprint Images ...... 371
   Hasan Fleyeh

Feature Selection and Bleach Time Modelling of Paper Pulp Using Tree Based Learners ..................................................... 385
   Karl Hansson, Hasan Fleyeh, and Siril Yella

Trust Model of Wireless Sensor Networks Based on Shannon Entropy ...... 397
   Jun Hu and Chun Guan
Assessing the Quality and Reliability of Visual Estimates in Determining Plant Cover on Railway Embankments .......................... 404
  Siril Yella and Roger G. Nyberg

Community-Based Message Transmission with Energy Efficient in Opportunistic Networks .................................................. 411
  Sheng Zhang, Xin Wang, Minghui Yao, and William Wei Song

A Multi-Semantic Classification Model of Reviews Based on Directed Weighted Graph ......................................................... 424
  Shaozhong Zhang, William Wei Song, Minjie Ding, and Ping Hu

Data Warehouse Quality Assessment Using Contexts ..................... 436
  Flavia Serra and Adriana Marotta

Author Index ................................................................. 449
Contents – Part I

Social Network Data Analysis

Attribute-Based Influence Maximization in Social Networks
   Jiuxin Cao, Tao Zhou, Dan Dong, Shuai Xu, Ziqing Zhu, Zhuo Ma,
   and Bo Liu
   3

Twitter Normalization via 1-to-N Recovering
   Yafeng Ren, Jiayuan Deng, and Donghong Ji
   19

A Data Cleaning Method for CiteSeer Dataset
   Yan Wang, Hao Zhang, Yaxin Li, Deyun Wang, Yanlin Ma, Tong Zhou,
   and Jianguo Lu
   35

Towards Understanding URL Resources in Recent Sina Weibo
   Yifang Wan, Peng Li, Rui Li, Metlin Zhou, Yongjun Ye, and Bin Wang
   50

Recommender Systems

Nonparametric Bayesian Probabilistic Latent Factor Model for Group
Recommendation Systems
   Nipa Chowdhury and Xiongcai Cai
   61

Joint User Knowledge and Matrix Factorization for Recommender Systems
   Yonghong Yu, Yang Gao, Hao Wang, and Ruili Wang
   77

GEMRec: A Graph-Based Emotion-Aware Music Recommendation Approach
   Dongjing Wang, Shuiguang Deng, and Guandong Xu
   92

Topic Modeling

Domain Dictionary-Based Topic Modeling for Social Text
   Bo Jiang, Jiguang Liang, Ying Sha, Rui Li, and Lihong Wang
   109

Towards an Impact-Driven Quality Control Model for Imbalanced
Crowdsourcing Tasks
   Kinda El Maarry and Wolf-Tilo Balke
   124

Modeling and Analyzing Engagement in Social Network Challenges
   Marco Brambilla, Stefano Ceri, Chiara Leonardi, Andrea Mauri,
   and Riccardo Volonterio
   140
Data Diversity

Select, Link and Rank: Diversified Query Expansion and Entity Ranking Using Wikipedia .................................................. 157
   Adit Krishnan, Deepak Padmanabhan, Sayan Ranu, and Sameep Mehta

Multi-dimension Diversification in Legal Information Retrieval. ............. 174
   Marios Koniaris, Ioannis Anagnostopoulos, and Yannis Vassiliou

Generating Multiple Diverse Summaries ............................................. 190
   Natwar Modani, Balaji Vasan Srinivasan, and Harsh Jhamtani

Diversifying the Results of Keyword Queries on Linked Data ............... 199
   Ananya Dass, Cem Aksoy, Aggeliki Dimitriou, Dimitri Theodoratos, and Xiaoying Wu

Data Similarity

Semantic Similarity of Workflow Traces with Various Granularities .......... 211
   Qing Liu, Quan Bai, and Yi Yang

Intermediate Semantics Based Distance Metric Learning for Video Annotation and Similarity Measurements ................................................. 227
   Wen Qu, Xiangmin Zhou, Daling Wang, Shi Feng, Yifei Zhang, and Ge Yu

A Community Detection Algorithm Considering Edge Betweenness and Vertex Similarity. ............................................................................. 243
   Hongwei Lu, Chang Liu, and Zaobin Gan

Measuring and Ensuring Similarity of User Interfaces: The Impact of Web Layout ................................................................. 252
   Sebastian Heil, Maxim Bakaev, and Martin Gaedke

Context-Aware Recommendation

Semantic Context-Aware Recommendation via Topic Models Leveraging Linked Open Data. ......................................................... 263
   Mehdi Allahyari and Krys Kochut

Optimizing Factorization Machines for Top-N Context-Aware Recommendations ............................................................................. 278
   Fajie Yuan, Guibing Guo, Joemon M. Jose, Long Chen, Haitao Yu, and Weinan Zhang
Taxonomy Tree Based Similarity Measurement of Textual Attributes of Items for Recommender Systems ............................................... 294
*Longquan Tao, Fei Liu, and Jinli Cao*

A Personalized Recommendation Algorithm for User-Preference Similarity Through the Semantic Analysis ................................................... 302
*Haolin Zhang and Feiyue Ye*

**Prediction**

Learning-Based SPARQL Query Performance Prediction .................. 313
*Wei Emma Zhang, Quan Z. Sheng, Kerry Taylor, Yongrui Qin, and Lina Yao*

Can Online Emotions Predict the Stock Market in China? .................. 328
*Zhenkun Zhou, Jichang Zhao, and Ke Xu*

Predicting Replacement of Smartphones with Mobile App Usage ........ 343
*Dun Yang, Zhiang Wu, Xiaopeng Wang, Jie Cao, and Guandong Xu*

Real Time Prediction on Revisitation Behaviors of Short-Term Type Commodities ................................................................. 352
*Xiangzhen Xu, Jinghua Fu, Yuliang Shi, Shijun Liu, and Lizhen Cui*

**Big Data Processing**

Parallel Materialization of Datalog Programs with Spark for Scalable Reasoning ................................................................. 363
*Haijiang Wu, Jie Liu, Tao Wang, Dan Ye, Jun Wei, and Hua Zhong*

A Data Type-Driven Property Alignment Framework for Product Duplicate Detection on the Web ......................................................... 380
*Gijs van Rooij, Ravi Sewnarain, Martin Skogholt, Tim van der Zaan, Flavius Frasincar, and Kim Schouten*

A Semantic Data Parallel Query Method Based on Hadoop ............... 396
*Liu Yang, Liu Yang, Jiangbo Niu, Zhigang Hu, Jun Long, and Meiguang Zheng*

A Strategy to Improve Accuracy of Multi-dimensional Feature Forecasting in Big Data Stream Computing Environments ......................... 405
*Dawei Sun, Hao Tang, Shang Gao, and Fengyun Li*
Cloud Computing

Automatic Creation and Analysis of a Linked Data Cloud Diagram .................. 417
   Alexander Arturo Mera Caraballo, Bernardo Pereira Nunes,
   Giseli Rabello Lopes, Luiz André Portes Paes Leme,
   and Marco Antonio Casanova

Fast Multi-keywords Search over Encrypted Cloud Data ............................ 433
   Cheng Hong, Yifu Li, Min Zhang, and Dengguo Feng

A Scalable Parallel Semantic Reasoning Algorithm-Based on RDFS Rules on Hadoop ................................................................. 447
   Liu Yang, Xiao Wen, Zhigang Hu, Chang Liu, Jun Long,
   and Meiguang Zheng

Cloud Resource Allocation from the User Perspective: A Bare-Bones
   Reinforcement Learning Approach .................................................. 457
   Alexandros Kontarinis, Verena Kantere, and Nectarios Koziris

Event Detection

Event Phase Extraction and Summarization ............................................. 473
   Chengyu Wang, Rong Zhang, Xiaofeng He, Guomin Zhou,
   and Aoying Zhou

ESAP: A Novel Approach for Cross-Platform Event Dissemination
   Trend Analysis Between Social Network and Search Engine .................... 489
   Yan Tang, Pengju Ma, Boyuan Kong, Wenqian Ji, Xiaofeng Gao,
   and Xuezheng Peng

Learning Event Profile for Improving First Story Detection
   in Twitter Stream .............................................................................. 505
   Yongqin Qiu, Rui Li, Lihong Wang, and Bin Wang

A Novel Approach of Discovering Local Community Using Node
   Vector Model .................................................................................. 513
   Jinglian Liu, Daling Wang, Shi Feng, Yifei Zhang, and Weiji Zhao

Data Mining

Labeled Phrase Latent Dirichlet Allocation ............................................... 525
   Yi-Kun Tang, Xian-Ling Mao, and Heyan Huang

WTEN: An Advanced Coupled Tensor Factorization Strategy for Learning
   from Imbalanced Data ...................................................................... 537
   Quan Do, Thanh Pham, Wei Liu, and Kotagiri Ramamohanarao
Mining Actionable Knowledge Using Reordering Based Diversified
Actionable Decision Trees .................................................. 553
Sudha Subramani, Hua Wang, Sathiyaabhamaa Balasubramaniam,
Rui Zhou, Jiangang Ma, Yanchun Zhang, Frank Whittaker, Yueai Zhao,
and Sarathkumar Rangarajan

Improving Distant Supervision of Relation Extraction
with Unsupervised Methods .................................................. 561
Min Peng, Jimin Huang, Zhaoyu Sun, Shizhong Wang, Hua Wang,
Guangping Zhuo, and Gang Tian

Author Index ................................................................. 569
Web Information Systems Engineering - WISE 2016
17th International Conference, Shanghai, China,
November 8-10, 2016, Proceedings, Part II
Zhang, Y. (Eds.)
2016, XXIII, 452 p. 142 illus., Softcover
ISBN: 978-3-319-48742-7