

Contents

Data Analytics for Targeted Healthcare

Conversation Intention Perception Based on Knowledge Base.	3
<i>Yi-Zheng Chen, Hua-Kang Li, and Yi Liu</i>	
An Empirical Algorithm for Bias Correction Based on GC Estimation for Single Cell Sequencing.	15
<i>Bo Xu, Tengpeng Li, Yi Luo, Ruotao Xu, and Hongmin Cai</i>	
Joint Tree of Combinatorial Maps.	22
<i>Tao Wang, Congyan Lang, and Songhe Feng</i>	
An Optimization Method of Fusing Multiple Decisions in Object Detection . . .	29
<i>Zhu Teng and Baopeng Zhang</i>	

Data Mining and Decision Analytics for Public Health and Wellness

Mining Typical Order Sequences from EHR for Building Clinical Pathways . . .	39
<i>Shoji Hirano and Shusaku Tsumoto</i>	
Aspect-Based Sentiment Analysis of Amazon Reviews for Fitness Tracking Devices	50
<i>Alaa Shafae, Hassan Issa, Stefan Agne, Stephan Baumann, and Andreas Dengel</i>	

Biologically Inspired Techniques for Data Mining

An Efficient Drug-Target Interaction Mining Algorithm in Heterogeneous Biological Networks	65
<i>Congcong Li, Jing Sun, Yun Xiong, and Guangyong Zheng</i>	
Comparison of Hemagglutinin and Neurominidase of Influenza A Virus Subtype H1N1, H5N1, H5N2, and H7N9 Using Apriori Algorithm. . .	77
<i>Dae Young Kim, Hye-Jun Kim, Junhyeok Bae, and Taeseon Yoon</i>	
Privacy Preserving Association Rule Mining Using Binary Encoded NSGA-II. . .	87
<i>Peng Cheng, Jeng-Shyang Pan, and Chun-Wei Lin</i>	
Detecting the Data Group Most Prone to a Specific Disguise Value	100
<i>Wen-Yang Lin and Wen-Yu Feng</i>	

rRNA of Alphaproteobacteria Rickettsiales and mtDNA Pattern Analyzing
with Apriori & SVM. 112
Seung Jae Lim, Shin Hyo Bang, Dae Seop Kim, and Taeseon Yoon

Mobile Data Management, Mining, and Computing on Social Networks

Team Formation with the Communication Load Constraint in Social Networks . . 125
Yui-Chieh Teng, Jun-Zhe Wang, and Jiun-Long Huang

Information Diffusion Pattern Mining over Online Social Media 137
Eric Hsueh-Chan Lu and Hui-Ju Hung

Mining Frequent Progressive Usage Patterns Across Multiple Mobile
Broadcasting Channels. 149
Bijay Prasad Jaysawal and Jen-Wei Huang

Efficient Gossiping Dissemination in Mobile Peer-to-Peer Environment
Based on Neighborhood Detection 156
Yu-Jen Lin, Bo-Heng Chen, Kun-Ta Chuang, and Chao-Chun Chen

Predicting Locations of Mobile Users Based on Behavior Semantic Mining . . . 168
Huei-Yu Lung, Chih-Heng Chung, and Bi-Ru Dai

Query Prediction by Currently-Browsed Web Pages and Its Applications 181
Hui-Ju Hung and Pu-Jen Cheng

An Emergency System in Mobile Devices 194
Chung-Hua Chu and Wen-Hsiang Chang

Study of Mobile Value-Added Services with Augmented Reality 199
Chung-Hua Chu, Shu-Lin Wang, and Bi-Chi Tseng

Mining Uncertain Sequence Data on Hadoop Platform. 204
Zi-Yun Sun, Ming-Che Tsai, and Hsiao-Ping Tsai

Big Data Science and Engineering on E-Commerce

Two Algorithms Under Stochastic Gradient Descent Framework
for Recommender Systems. 219
Tian-Hsiang Huang and Vladimir Nikulin

Scalable Textual Similarity Search on Large Document Collections Through
Random Indexing and K-means Clustering. 231
Ali Cevahir

Complementary Product Selection in E-Commerce 239
Yian Chen, Yuh-Ming Chiu, and Song Han

A Short-Term Bookmarking System for Collecting User-Interest Data	247
<i>Chu-Cheng Hsieh, Yoni Medoff, and Naren Chittar</i>	

Cloud Service Discovery

Automatic Self-Suspended Task for a MapReduce System on Cloud Computing	257
<i>Tzu-Chi Huang, Ce-Kuen Shieh, Sheng-Wei Huang, Chui-Ming Chiu, and Tyng-Yeu Liang</i>	
Tag Completion and Refinement for Web Service via Low-Rank Matrix Completion	269
<i>Lei Chen, Geng Yang, Zhengyu Chen, Fu Xiao, and Xuxia Li</i>	

Mobile Sensing, Mining and Visualization for Human Behavior Inference

A Differential Approach for Identifying Important Student Learning Behavior Patterns with Evolving Usage over Time	281
<i>John S. Kinnebrew, Daniel L.C. Mack, Gautam Biswas, and Chih-Kai Chang</i>	
Mining Top-K Relevant Stay Regions from Historical Trajectories	293
<i>Yung-Hsiang Lin, Chien-Hsiang Lai, and Po-Ruey Lei</i>	
Associative Classification for Human Activity Inference on Smart Phones . . .	305
<i>Yu-Hsiang Peng, Gunarto Sindoro Njoo, Shou-Chun Li, and Wen-Chih Peng</i>	
Personalized Smartphone Wearing Behavior Analysis	318
<i>Yen-Hsuan Lin and Yi-Ta Chuang</i>	
Activity Recognition Using Discriminant Sequence Patterns in Smart Environments	329
<i>Been-Chian Chien and Rong-Sing Huang</i>	

Scalable Data Analytics: Theory and Applications

Using Knowledge Graph to Handle Label Imperfection	345
<i>Yi Liu, Huakang Li, and Yizheng Chen</i>	
Learning to Display in Sponsored Search	357
<i>Xin Xin and Heyan Huang</i>	
Models for Distributed, Large Scale Data Cleaning	369
<i>Vincent J. Maccio, Fei Chiang, and Douglas G. Down</i>	

A Study of Rumor Spreading with Epidemic Model Based on Network Topology	381
<i>Dawei Meng, Lizhi Wan, and Lei Zhang</i>	
GS4: Generating Synthetic Samples for Semi-Supervised Nearest Neighbor Classification	393
<i>Panagiotis Moutafis and Ioannis A. Kakadiaris</i>	
A Scalable Data Analytics Algorithm for Mining Frequent Patterns from Uncertain Data	404
<i>Richard Kyle MacKinnon, Carson Kai-Sang Leung, and Syed K. Tanbeer</i>	
Parallel Time Series Modeling - A Case Study of In-Database Big Data Analytics	417
<i>Hai Qian, Shengwen Yang, Rahul Iyer, Xixuan Feng, Mark Wellons, and Caleb Welton</i>	
A Biclustering-Based Classification Framework for Microarray Analysis	429
<i>Baljeet Malhotra, Daniel Dahlmeier, and Naveen Nandan</i>	
Parallel Implementation of a Density-Based Stream Clustering Algorithm Over a GPU Scheduling System	441
<i>Marwan Hassani, Ayman Tarakji, Lyubomir Georgiev, and Thomas Seidl</i>	
Algorithms for Large-Scale Information Processing in Knowledge Discovery	
Three-way Indexing ZDDs for Large-Scale Sparse Datasets	457
<i>Hiroshi Aoki, Takahisa Toda, and Shin-ichi Minato</i>	
Local Feature Selection by Formal Concept Analysis for Multi-class Classification	470
<i>Madori Ikeda and Akihiro Yamamoto</i>	
Ensemble Clustering of High Dimensional Data with FastMap Projection . . .	483
<i>Imran Khan, Joshua Zhexue Huang, Nguyen Thanh Tung, and Graham Williams</i>	
A General Framework for Parallel Unary Operations on ZDDs	494
<i>Shogo Takeuchi, Takahisa Toda, and Shin-ichi Minato</i>	
On the Size of the Zero-Suppressed Binary Decision Diagram that Represents All the Subtrees in a Tree	504
<i>Norihito Yasuda, Masaaki Nishino, and Shin-ichi Minato</i>	

Data Mining in Social Networks

A Unified Fusion Framework for Time-Related Rank in Threaded Discussion Communities	513
<i>Qiang You, Weiming Hu, Ou Wu, and Haiqiang Zuo</i>	
Multi-group-based User Perceptions for Friend Recommendation in Social Networks	525
<i>Trung L.T. Nguyen and Tru H. Cao</i>	
STC: A Joint Sentiment-Topic Model for Community Identification	535
<i>Baoguo Yang and Suresh Manandhar</i>	
Considerations About Multistep Community Detection	549
<i>Antonio A. Gentile, Angelo Corallo, Cristian Bisconti, and Laura Fortunato</i>	
An Edge-Centric Approach for Change Point Detection in Dynamic Networks . . .	562
<i>Yongsheng Cheng and Xiaokang Lin</i>	
Toward Automatic Censorship Detection in Microblogs	572
<i>Donn Morrison</i>	
Competition Component Identification on Twitter	584
<i>Cheng-Huang Yang, Ji-De Chen, and Hung-Yu Kao</i>	
Social Network Community Detection Using Strongly Connected Components	596
<i>Wookey Lee, James J. Lee, and Jinho Kim</i>	

Data Mining in Biomedical informatics and Healthcare

Automatic Segmentation and Quantitative Analysis of Gray Matter on MR Images of Patients with Epilepsy Based on Unsupervised Learning Methods	607
<i>Rui Wang, Jie Wang, Jing Ding, and Su Zhang</i>	
An Author Topic Analysis of Tobacco Regulation Investigators	616
<i>Ding Cheng Li, Janet Okamoto, Scott Leischow, and Hongfang Liu</i>	
Mining Severe Drug-Drug Interaction Adverse Events Using Semantic Web Technologies: A Case Study	628
<i>Guoqian Jiang, Hongfang Liu, Harold R. Solbrig, and Christopher G. Chute</i>	
Cancer Based Pharmacogenomics Network for Drug Repurposing	639
<i>Liwei Wang, Hongfang Liu, Christopher G. Chute, and Qian Zhu</i>	

Linked Vaccine Adverse Event Data Representation from VAERS for Biomedical Informatics Research	652
<i>Cui Tao, Puqiang Wu, and Yuji Zhang</i>	

Pattern Mining and Application of Big Data

QRM: A Probabilistic Model for Search Engine Query Recommendation. . . .	665
<i>JianGuo Wang and Joshua Zhexue Huang</i>	
Toward Mining User Traversal Patterns in the Indoor Environment.	677
<i>Shan-Yun Teng, Tzu-Yuan Chung, Kun-Ta Chuang, and Wei-Shinn Ku</i>	
Myocardial Infarction Classification by Morphological Feature Extraction from Big 12-Lead ECG Data	689
<i>Julia Tzu-Ya Weng, Jyun-Jie Lin, Yi-Cheng Chen, and Pei-Chann Chang</i>	
Automatic Restaurant Information and Keyword Extraction by Mining Blog Data for Chinese Restaurant Search.	700
<i>Chien-Li Chou, Min-Ho Tsai, Chien-Ho Chao, Hsiao-Jung Lin, Hua-Tsung Chen, Suh-Yin Lee, and Chien-Peng Ho</i>	
The Application of Association Rules in Clinical Disease: The Relationship Between Recovery After Operation of Endovascular Aneurysm Repairing and Chronic	712
<i>Lin Hui, Chun-Che Shih, Huan-Chao Keh, Po-Yuan Yu, Yuan-Cheng Cheng, and Nan-Ching Huang</i>	
Human Action Recognition in Video Under Clutter and Moving Background . . .	722
<i>Der-Jyh Duh, Cheng-Chung Kan, Shu-Yuan Chen, and Chia-Ming Lu</i>	
Big Data Generation: Application of Mobile Healthcare.	735
<i>Huan-Chao Keh, Lin Hui, Kuang-Yi Chou, Yuan-Cheng Cheng, Po-Yuan Yu, and Nan-Ching Huang</i>	
A Comparison of Feature-Combination for Example-Based Super Resolution . . .	744
<i>Shwu-Huey Yen, Jen-Hui Tsao, and Wan-Ting Liao</i>	
Comparative Analysis of 3D-Culture System for Murine Neonatal Heart Regeneration: A Systematic Approach for Big Gene Expression Data	754
<i>Julia Tzu-Ya Weng, Yi-Cheng Chen, Pei-Chann Chang, Shin-Ping Huang, and Yu-Wei Chiu</i>	
Data Collection and Analysis from Social Network – Profile Analyzer System (PAS).	765
<i>Wei-Hao Chang, Bing Li, and Xin Fang</i>	
Analyzing Personality Traits Based on User Behavours of Using Microblog . . .	773
<i>Xin Fang, Bing Li, and Wei-Hao Chang</i>	

BigSAM: Mining Interesting Patterns from Probabilistic Databases of Uncertain Big Data	780
<i>Fan Jiang, Carson Kai-Sang Leung, and Richard Kyle MacKinnon</i>	
MBPD: Motif-Based Period Detection	793
<i>Rasaq Otunba, Jessica Lin, and Pavel Senin</i>	
A Model-Based Multivariate Time Series Clustering Algorithm	805
<i>Pei-Yuan Zhou and Keith C.C. Chan</i>	
Mining Time-Aware Transit Patterns for Route Recommendation in Big Check-in Data.	818
<i>Hsun-Ping Hsieh and Cheng-Te Li</i>	
Author Index	831

Trends and Applications in Knowledge Discovery and
Data Mining

PAKDD 2014 International Workshops: DANTH, BDM,
MobiSocial, BigEC, CloudSD, MSMV-MBI, SDA,
DMDA-Health, ALSIP, SocNet, DMBIH, BigPMA, Tainan,
Taiwan, May 13-16, 2014. Revised Selected Papers
Peng, W.-C.; Wang, H.; Bailey, J.; Tseng, V.S.; Ho, T.-B.;
Zhou, Z.-H.; Chen, A.L.P. (Eds.)

2014, XXI, 833 p. 348 illus., Softcover

ISBN: 978-3-319-13185-6