2017 Fifth International Conference on Advanced Cloud and Big Data (CBD 2017)

Shanghai, China
13-16 August 2017
Welcome Message
Organizing Committee
Program Committee
Reviewers

Session 1: Keynote
Business Intelligence Revisited
  Jianwen Su and Yan Tang

Session 2: Cloud Computing and Applications I
A Bandwidth Allocation Policy for Helpers in Cloud-Assisted P2P
  Guowei Huang, Lingjing Kong, Keke Wu, and Zhi Chen

A Dual-Buffer Based Acceleration Strategy for FW Block Algorithm on GPUs
  Yiwen Li and Jian Yu

A Large Scale Data Transmission Control Mechanism across Data Centers
  Wenjiang Ding, Wei Li, Runhuan Zhang, Qianqian Zhang, and Chenhao Yang

A Novel Solution of Cloud Operating System Based on X11 and Docker
  Zhen Du, Zhuging Xu, Fang Dong, and Dian Shen

Adaptive Resource Prediction in the Cloud Using Linear Stacking Model
  Shasha Liao, Hongjie Zhang, Guansheng Shu, and Jing Li

An Heuristic for Bag-of-Tasks Scheduling Problems with Resource Demands
  and Budget Constraints to Minimize Makespan on Hybrid Clouds
  Yi Zhang, Jin Sun, and Zebin Wu

BigVM: A Multi-Layer-Microservice-Based Platform for Deploying SaaS
  Tianlei Zheng, Yuqun Zhang, Xi Zheng, Min Fu, and Xiao Liu
Cost Optimization for Scheduling Scientific Workflows on Clouds
under Deadline Constraints ........................................................................................................................51
   Wei Zheng, Bugingo Emmanuel, Chen Wang, Yingsheng Qin,
   and Dongzhan Zhang

Session 3: Cloud Computing and Applications II
Dynamic Flow Scheduling for Power Optimization of Data Center Networks ............................................57
   Juan Luo, Song Zhang, Luxiu Yin, and Yaling Guo
Efficient Data and Task Co-Scheduling for Scientific Workflow
in Geo-Distributed Datacenters ..................................................................................................................63
   Jian Chen, Jinghui Zhang, and Aibo Song
EQF: An Explicit Queue-Length Feedback for TCP Congestion Control
in Datacenter Networks ................................................................................................................................69
   Yifei Lu, Xiaoting Fan, and Lei Qian
Estimating SDN Traffic Matrix Based on Online Informative Flow Measurement
Method ..........................................................................................................................................................75
   Dongyang Li, Ningyun Dai, Feng Li, Changyou Xing, and Fei Dai
PCMSim: A Hybrid Memory System Simulator for the Cloud Storage .......................................................81
   Jibin Wang and Bohan Wang
Research and Implementation of a Software Online Testing Platform Model
Based on Cloud Computing ........................................................................................................................87
   Jing Chen, Chunxiao Wang, Fei Liu, and Yinglong Wang
Towards Cost Reduction in Cloud-Based Workflow Management through Data
Replication ..................................................................................................................................................94
   Fei Xie, Jun Yan, and Jun Shen

Session 4: Big Data and Applications I
A Cost-Efficient Data Placement Algorithm with High Reliability in Hadoop ............................................100
   Yao Du, Runqun Xiong, Jiahui Jin, and Junzhou Luo
A Data Replenishment Method for Self Regulating Sail-Assisted Ships
Considering Big Data ....................................................................................................................................106
   Tianyi Zhang, Jianing Zhang, and Chengming Yang
A Decision Support Scheme for Cloud Storage ..........................................................................................111
   Yang Gao, Keqiu Li, and Heng Qi
A Distributed Approach for Constructing Generalized Suffix Tree on Spark
by Using Optimized Elastic Range Algorithm .........................................................................................117
   Yunhao Li, Jiahui Jin, Runqun Xiong, and Junzhou Luo
A Feature-Based Food Recommendation on Apache Spark ...................................................................123
   Wenyuan Xu, Yun Li, Liting Wei, and Zheng Wu
An Online Prediction Algorithm of Traffic in Big Data Based on the Storm ..............................................129
Shangdong Liu, Yimu Ji, Dianchao Zhang, Yongge Yuan, Jian Gong, and Ruchuan Wang

An Ontology-Based Data Organization Method ........................................................................................135
Qian Hao, Yue Li, Li Min Wang, and Mei Wang

Session 5: Big Data and Applications II

Consistent Hash Based Flexible Redundant Data Partition Strategy ..............................................................141
Zhizhu Zhong, Ning Wang, Dun Li, Zhimeng Du, and Zhenfei Wang

Data Deduplication Based on Hadoop ...........................................................................................................147
Dongzhan Zhang, Chengfa Liao, Wenjing Yan, Ran Tao, and Wei Zheng

Detection of Super Nodes for Network Data Streams with Filtering Strategy ............................................153
Aiping Zhou, Huisheng Zhu, Lijun Liu, and Chengang Zhu

Distributed Processing Practice of the 3D City Model Based on HBase .......................................................159
Chaoquan Zhang and Bo Mao

Finding Compositional Skyline Based on Sharing Strategy ........................................................................164
Leigang Dong and Guohua Liu

GeaBase: A High-Performance Distributed Graph Database for Industry-Scale Applications .......................170
Zhishong Fu, Zhengwei Wu, Houyi Li, Yize Li, Min Wu, Xiaojie Chen, Xiaomeng Ye, Benquan Yu, and Xi Hu

Intelligent Anti-Money Laundering Solution Based upon Novel Community Detection in Massive Transaction Networks on Spark .........................................................176
Xurui Li, Xiang Cao, Xuetao Qiu, Jintao Zhao, and Jianbin Zheng

Logic-Partition Based Gaussian Sampling for Online Aggregation ............................................................182
Longbin Zhang, Yuxiang Wang, and Xiaoliang Xu

Session 6: Social Computing and Applications I

A Lexical Updating Algorithm for Sentiment Analysis on Chinese Movie Reviews ...............................................188
Yiwei Song, Kaiwen Gu, Huakang Li, and Guozi Sun

A Multidimensional Comprehensive Recommendation Method Based on Social Network ..............................194
Bo Zhang, Cui Mu, Qin Zhao, Zhangjie Peng, Junhong Ding, and Bo Liu

A Novel Method for Open Relation Extraction from Public Announcements of Chinese Listed Companies ....200
Tianyuan Zhuang, Pengwei Wang, Yaying Zhang, and Zhaohui Zhang

A Real-Time Bus Transfer Scheme Recommendation Systems ........................................................................206
Jun Zhang, Lai Tu, Fan Zhang, and Benxiong Huang
Session 7: Social Computing and Applications II

Cost-Efficient Influence Maximization in Online Social Networks .......................................................... 232
   Jingya Zhou, Jianxi Fan, Jin Wang, Xi Wang, and Baolei Cheng

Efficient Fine-Grained Location Prediction Based on User Mobility Pattern in LBSNs ........................................ 238
   Jixun Cao, Shuai Xu, Xuelin Zhu, Renjun Lv, and Bo Liu

K-Truss Based Top-Communities Search in Large Graphs ........................................................................... 244
   Anzuruni Maulidi Katunka, Cairong Yan, Kossonon Brandou Serge, and Zhaohui Zhang

Location and Motion Prediction of Consumers in a Large Shopping Mall ...................................................... 250
   Zeyu Zhang and Weiping Zhu

Passenger Travel Path Estimation Algorithm Based on High Accuracy Location Data ................................................. 256
   Huali Xiao, Li Sun, Shaohong Kong, Guiwei Gong, and Fan Zhang

VOUt: Mining Sampled GPS Trace for Taximeter Fraud .................................................................................... 261
   Ding Jing and Hongzi Zhu

Session 8: Cloud Security and Applications

Dynamic Cloud Access Control via Accelerometers Using Customized Actions of Occupants ................. 267
   Biaokai Zhu, Jumin Zhao, Deng-Ao Li, Hong Wang, Ruiqin Bai, Yanxia Li, and Bin Yang

Fast Identification of Encrypted Traffic via Large-Scale Sparse Screening .................................................... 273
   Peng Meng, Guopeng Zhou, and Juan Meng

Identifying Mobile Applications for Encrypted Network Traffic ...................................................................... 279
   Gaofeng He, Bingfeng Xu, and Haiting Zhu

Detecting Clocking Attack in Contactless Access Control Systems ............................................................... 285
   Guixiang Tian, Li Lu, and Muhammad Jawad Hussain
Session 9: Machine Learning and Miscellaneous I

A Data Reconstruction Algorithm Based on Neural Network for Compressed Sensing .....................................................................................................................................................291
   Li Tian, Guorui Li, and Cong Wang

A Mining Approach to Obtain the Software Vulnerability Characteristics .................................................................296
   Xiang Li, Jinfu Chen, Zhechaow Lin, Lin Zhang, Zibin Wang, Minmin Zhou, and Wanggen Xie

A New Design of Real-Time Traffic Index Model for Freeway .....................................................................................302
   Fusheng Xu, Zhongxiang Huang, and Xueying Zhu

A Parallel Algorithm for Mining Time Relaxed Gradual Clustering Pattern Based on Spatio-Temporal Trajectories ...................................................................................................308
   Hongyan Sun, Genlin Ji, Bin Zhao, and Xintao Liu

An Improved Algorithm for Frequent Itemsets Mining ...............................................................................................314
   Hao Jiang and Xu He

Deep Neural Network Based on Translation Model for Diabetes Knowledge Graph ........................................................................................................................................................318
   Suna Yin, Dehua Chen, and Jiajin Le

Detecting Spammers in E-Commerce Website via Spectrum Features of User Relation Graph ..........................................................................................................................................324
   Bo Guo, Hao Wang, Zhaojun Yu, and Yu Sun

Improved Candidate Answers Ranking for QA .........................................................................................................331
   Chengchao Yu, Rui Zhao, and Zhihua Wei

Session 10: Machine Learning and Miscellaneous II

PixVid: Capturing Temporal Correlated Changes in Time Series ...............................................................................337
   Yeqiang Lin, Mehrzad Lavassani, Jiyun Li, and Tingting Zhang

RBTA: A Multivariate Time-Series Method for City Incidents Mining and Forecasting ..................................................343
   Jieyi Wang, Yongkun Wang, and Yaohui Jin

Recommendation for Cross-Disciplinary Collaboration Based on Potential Research Field Discovery ................................................................................................................................349
   Wei Liang, Xiaokang Zhou, Suzhen Huang, Chunhua Hu, and Qun Jin

Recommendation Algorithm Based on User Trust and Interest with Probability Matrix Factorization ........................................................................................................................................362
   Chuanmin Mi, Peng Peng, Lin Xiao, and Yue Lu

Search for Top-K Spatial Objects Based on Various-Widths Clustering ........................................................................362
   Shoujian Yu and Guangyi Feng
Semi-SGD: Semi-Supervised Learning Based Spammer Group Detection in Product Reviews
Lu Zhang, Yang Yuan, Zhiang Wu, and Jie Cao ............................................................. 368

Target Distribution Guided Network Sampling
Renjie Fan, Zhiwen Yu, Bin Guo, Liang Wang, and Dingqi Yang ................................... 374

Trajectory Data Driven Transit-Transportation Planning
Yihan Guo, Shaoyong Wang, Lin Zheng, and Mingming Lu ............................................. 380

Session 11: Machine Learning and Miscellaneous III

Two-Phase Neural Network Model for Pollution Concentrations Forecasting
Cairong Yan, Shuhua Xu, Yan Huang, Yongfeng Huang, and Zhaohui Zhang .................. 385

Predicting Object Types in Linked Data by Text Classification
Xiang Zhang, Erjing Lin, and Siyao Pi ................................................................................. 391

Video Object Detection for Tractability with Deep Learning Method
Bing Tian, Liang Li, Yansheng Qu, and Li Yan ................................................................. 397

Your Body Defines Your Fall Detection System: A Somatotype-Based Feature
Selection Method
Xiaoling Fu, Zheng Tan, Chungang Yan, Zhong Li, Cheng Wang,
and Cheng-Xiang Wang ....................................................................................................... 402

An Indoor TV White Space Resource Collecting and APs Locations Selecting Scheme
Chengsi Qi, Guanglin Zhang, Wenqiang Zhang, and Demin Li .......................................... 408

An Outlier Detection Method and Its Application for Earthen Ruins Data
Monitored by WSN
Yun Xiao, Haohan Gao, Jia Ding, Xuanhong Wang, and Xianjia Meng ......................... 414

ORSCA: Optimal Route Selection and Communication Association for Drones
in WSNs
Tao Wu, Panlong Yang, Yubo Yan, Xunpeng Rao, Ping Li, and Wanru Xu ......................... 420

Author Index .......................................................................................................................... 425