

# **2014 Second International Conference on Advanced Cloud and Big Data**

**(CBD 2014)**

**Huangshan, China  
20-22 November 2014**



**IEEE Catalog Number: CFP1403Y-POD  
ISBN: 978-1-4799-8067-3**

# 2014 Second International Conference on Advanced Cloud and Big Data

## CBD 2014

### Table of Contents

Welcome Message from the General	
Conference Co-Chairs.....	ix
Organizing Committee.....	x
Program Committee.....	xi
Keynote Speeches.....	xiii
Technical Forum Speeches.....	xviii

---

### Session 1: Big Data and Applications I

Data Locality-Aware Query Evaluation for Big Data Analytics in Distributed Clouds .....	1
<i>Qiufen Xia, Weifa Liang, and Zichuan Xu</i>	
SLDP: A Novel Data Placement Strategy for Large-Scale Heterogeneous Hadoop Cluster .....	9
<i>Runqun Xiong, Junzhou Luo, and Fang Dong</i>	
Scaling Information-Theoretic Text Clustering: A Sampling-based Approximate Method .....	18
<i>Zhexi Xu, Zhiang Wu, Jie Cao, and Hengnong Xuan</i>	
An Effective Method for Mining Quantitative Association Rules with Clustering Partition in Satellite Telemetry Data .....	26
<i>Xin Dong and Dechang Pi</i>	
Examination of Ex-Dividend Day Trading Using Big Data of American Depository Receipts .....	34
<i>Bi-Huei Tsai</i>	
A Score Based Approach towards Improving Bayesian Network Structure Learning .....	39
<i>Yan Tang and Zhuoming Xu</i>	

Multi-objective Ant Colony System for Data-Intensive Service Provision .....	45
<i>Lijuan Wang, Jun Shen, and Junzhou Luo</i>	
A Parallel Algorithm to Mine Abnormal Patterns from Satellite Data .....	53
<i>Yuhang Xu and Dechang Pi</i>	

## **Session 2: Cloud Computing and Applications**

Strategy-Proof Auction Mechanism with Group Price for Virtual Machine Allocation in Clouds .....	60
<i>Yonglong Zhang, Bin Li, Zhiqiu Huang, Jin Wang, Junwu Zhu, and Huanfeng Peng</i>	
Deadline Constrained Energy-Efficient Scheduling for Workflows in Clouds .....	69
<i>Wei Zheng and Shouhui Huang</i>	
Data Placement and Task Scheduling Optimization for Data Intensive Scientific Workflow in Multiple Data Centers Environment .....	77
<i>Mingjun Wang, Jinghui Zhang, Fang Dong, and Junzhou Luo</i>	
Efficient Auction Mechanism with Group Price for Resource Allocation in Clouds .....	85
<i>Yiyi Ma, Bin Li, Yonglong Zhang, and Junwu Zhu</i>	
Improved GEP Algorithm for Task Scheduling in Cloud Computing .....	93
<i>LI Kun-lun, Wang Jun, Song Jian, and Dong Qing-yun</i>	
Multi-Q: Multiple Queries Optimization Based on MapReduce in Cloud .....	100
<i>Ding Ding, Fang Dong, and Junzhou Luo</i>	
Cost in the Cloud Rationalisation and Research Trails .....	108
<i>Kahina Hamadache, Paraskevi Zerva, Ariana Polyviou, Viliam Simko, Rustem Dautov, Fotis Gonidis, and Ivan Paez Anaya</i>	
A Monitoring and Prediction Model of Workflow Based Self-Adaptive Software System .....	115
<i>Xiaowei Zhang, Bin Li, and Junwu Zhu</i>	

## **Session 3: Big Data and Applications II**

Near-Optimal Approximate Duplicate-Detection in Data Streams Over Sliding Windows for the Uniform Query Frequency or Membership Likelihood .....	122
<i>Xiujun Wang, Xiao Zheng, Zhe Dang, Xuangou Wu, and Baohua Zhao</i>	
MyBSP: An Iterative Processing Framework Based on the Cloud Platform for Graph Data .....	128
<i>Chao Liu, Hong Yao, Deze Zeng, Qingzhong Liang, Chengyu Hu, and Xuesong Yan</i>	
Energy Consumption Data Based Machine Anomaly Detection .....	136
<i>Hui Chen, Xiang Fei, Sheng Wang, Xin Lu, Guoqin Jin, Weidong Li, and Xuyang Wu</i>	

Protecting and Analysing Health Care Data on Cloud .....	143
<i>Danan Thilakanathan, Yu Zhao, Shiping Chen, Surya Nepal, Rafael A. Calvo, and Abelardo Pardo</i>	

Research of Big Data Space-Time Analytics for Clouding Based Contexts-Aware IOV Applications .....	150
<i>Di Zheng, Kerong Ben, and Hongliang Yuan</i>	

An Image Processing System Based on Heterogeneous Embedded Multicore Processors .....	157
<i>Liang Zhang, Peiyi Shen, Juan Song, Lb Dong, Guoxi Wang, and Yuxin Cai</i>	

## **Session 4: Data Center Networking**

SA-TCP: A Novel Approach to Mitigate TCP Incast in Data Center Networks .....	163
<i>Yongmao Ren, Jun Li, Guodong Wang, Lingling Li, and Shanshan Shi</i>	

Data Center Multicast with High Stability .....	168
<i>Ye Shen, Jing Feng, Zhanfeng Wang, Weijun Ma, Min Yin, and Lei Jiang</i>	

Improved Directed Diffusion Protocol Based on Visible Forwarding Path and Promoted Evaluation Criterion .....	174
<i>Yan Gu, Liang Ma, Jie Guo, and Dahai Jing</i>	

Research on Undirected Double-Loop Data Center Networks .....	180
<i>Yebin Chen, Ying Li, Weipeng Tai, and Xiao Zheng</i>	

SSNF: Shared Datacenter Mechanism for Inter-datacenter Bulk Transfer .....	184
<i>Yang Yu, Wang Rong, and Wang Zhijun</i>	

NVLAN: A Novel VLAN Technology for Scalable Multi-tenant Datacenter Networks .....	190
<i>Yang Hu, Huiba Li, and Yuxing Peng</i>	

A Node-to-Set Disjoint Path Routing Algorithm in DCell Networks .....	196
<i>Xi Wang and Jianxi Fan</i>	

## **Session 5: Cloud Computing and Security**

Entropy-Based Denial of Service Attack Detection in Cloud Data Center .....	201
<i>Jiuxin Cao, Bin Yu, Fang Dong, Xiangying Zhu, and Shuai Xu</i>	

A MapReduce Task Scheduling Algorithm for Deadline-Constraint in Homogeneous Environment .....	208
<i>Yi Yang, Jiao Xu, Fei Wang, Zhaocai Ma, Jingshan Wang, and Lian Li</i>	

Advances and Future Challenges in Electric Power Big Data .....	213
<i>Yang Yindong and Bi Zhongqin</i>	

Inferring Application Type Information from Tor Encrypted Traffic .....	220
<i>Gaofeng He, Ming Yang, Junzhou Luo, and Xiaodan Gu</i>	

Efficient Approximate Search Using VGRAM Over Encrypted Cloud Data .....	228
<i>Lidong Han and Changhui Hu</i>	
Distributed Key Management Scheme Based on ECC for Heterogeneous Sensor Networks .....	235
<i>Jiu-Ru Wang and Hai-Feng Wang</i>	
A Novel Method to Detect Encrypted Data Exfiltration .....	240
<i>Gaofeng He, Tao Zhang, Yuanyuan Ma, and Bingfeng Xu</i>	
An Intrusion Detection Model Based on Deep Belief Networks .....	247
<i>Ni Gao, Ling Gao, Quanli Gao, and Hai Wang</i>	
 <b>Session 6: Social Computing and Service Computing</b>	
Efficient Parallel Community Detection in Large Edge-Intensive Networks .....	253
<i>Guangliang Gao, Zhan Bu, Zhiang Wu, Yuan Li, and Jie Cao</i>	
Social Recommendation Algorithm Dynamically Adaptable to User Profiling for SNS .....	261
<i>Weimin Li, Yikai Ni, Minye Wu, Zhengbo Ye, and Qun Jin</i>	
Geographic Location-Based Service Reliability Prediction .....	267
<i>Haiyan Wang and Jun Qian</i>	
A Personalized Service Recommendation Algorithm for Service Functionality .....	275
<i>Wei Li, Di Hu, and Junzhou Luo</i>	
A Clustering-Based Similarity Measurement for Collaborative Filtering .....	282
<i>Liang Gu, Peng Yang, and Yongqiang Dong</i>	
The SOM Based Improved K-Means Clustering Collaborative Filtering Algorithm in TV Recommendation System .....	288
<i>Zhaocai Ma, Yi Yang, Fei Wang, Caihong Li, and Lian Li</i>	
A Group Member Search Method Based on Dispersion Analysis .....	296
<i>Qiang Yang, Bing Huang, Xiaojun Zeng, and Jianfeng Liu</i>	
Collaborative Filtering for Recommender Systems .....	301
<i>Ruisheng Zhang, Qi-dong Liu, Chun-Gui, Jia-Xuan Wei, and Huiyi-Ma</i>	
 <b>Author Index</b> .....	 309