

# **2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA 2018)**

**Shanghai, China  
9-12 March 2018**



**IEEE Catalog Number: CFP18TPE-POD  
ISBN: 978-1-5386-4795-0**

**Copyright © 2018 by the Institute of Electrical and Electronics Engineers, Inc.  
All Rights Reserved**

*Copyright and Reprint Permissions:* Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

***\*\*\* This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP18TPE-POD
ISBN (Print-On-Demand):	978-1-5386-4795-0
ISBN (Online):	978-1-5386-4794-3

**Additional Copies of This Publication Are Available From:**

Curran Associates, Inc  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: (845) 758-0400  
Fax: (845) 758-2633  
E-mail: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

# 2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA 2018)

## Table of Contents

<i>Preface</i> .....	viii
<i>Committees</i> .....	ix
<b>Big Data Analysis and Methods</b>	
The Effect Analysis of Critical Subgroup Based on K-core on Public Outbreak as Big Data.....	1
<i>Yu Gong, Mingchun Zheng</i>	
College Data Analysis Based on Multi-learning Method.....	5
<i>Jia Shuai, Shaocong Mo, Yunyao Yang, Xu Han, Zhongyuan Chen, Zefeng Xie, Chao Dai</i>	
Leveraging Big Data Analytics to Improve Decision Making in South African public Universities .....	9
<i>Mmatshuene Anna Segooa, Billy Mathias Kalema</i>	
Big Data Analytics Using Association Rules in eLearning.....	14
<i>Korn Poonsirivong, Chanintorn Jittawiriaynukoon</i>	
The Research on Industrial Big Data Information Security Risks.....	19
<i>Feng Xiaorong, Jia Shizhun, Mai Songtao</i>	
Study on the Application of Big Data in Accurate Marketing of Cross-border Ecommerce in China.....	24
<i>Huiqun Liu, Xiao Wang</i>	
An Effective Selecting Approach for Social Media Big Data Analysis –Taking Commercial Hotspot Exploration with Weibo Check-in Data as An Example.....	28
<i>Qingwu Hu, Yuan Zhang</i>	
The Logistics Performance Evaluation Index System in the Transportation Industry Based on Big Data.....	33
<i>Ran Ding, Peng Ren</i>	
Centrality Analysis of Online Social Network Big Data.....	38
<i>Hui Li</i>	
Real-Time Big Data Warehousing and Analysis Framework .....	43
<i>Abbas Raza Ali</i>	
<b>Data Mining</b>	
Mining Spatio-temporal Patterns in Multivariate Spatial Time Series.....	50
<i>Gene P. K. Wu, Keith C. C. Chan</i>	
Association Rules Mining Analysis of App Usage Based on Mobile Traffic Flow Data .....	55
<i>Yang Liu, Ke Yu, Xiaofei Wu, Yazhou Shi, Yaowen Tan</i>	
Analysis and Prediction of Influenza in the UAE based on Arabic Tweets .....	61
<i>Balsam Alkouz, Zaher Al Aghbari</i>	
An Algorithmic Approach to Handle Circular Trading in Commercial Taxation System.....	67
<i>Jithin Mathews, Priya Mehta, Ch. Sobhan Babu, S.V. Kasi Visweswara Rao</i>	

Research on Kano Model Based on Online Comment Data Mining.....	76
<i>Huaming Song, Chao Chen, Qin Yu</i>	
A Brief Analysis of the Key Technologies and Applications of Educational Data Mining on Online Learning Platform.....	83
<i>Wei Zhang, Shiming Qin</i>	
<b>Data Theory and Engineering</b>	
Research and Implementation of Data Fusion Method Based on RDF.....	87
<i>Bin Yu, Chen Zhang, Yitong Li, Jiangyan Sun</i>	
Research hotspots of "Course of Action" Based on Bibliometrics.....	92
<i>Chao Chen, Changjun Fan, Zili Li, Li Zeng</i>	
The Simulation of Open One-side Uncertain Probability for Fusion Model of Data Uncertainty and Data Relation Uncertainty.....	97
<i>Ruiqi Ye, Mingxue Liao, Tianyu Cui, Pin Lv</i>	
Generalizing Application Agnostic Remaining Useful Life Estimation Using Data-Driven Open Source Algorithms.....	102
<i>Bernhard Schlegel, Artur Mrowca, Peter Wolf, Bernhard Sick, Sebastian Steinhorst</i>	
Spatio-temporal Prediction of Shopping Behaviours Using Taxi Trajectory Data.....	112
<i>John Cartlidge, Shuhui Gong, Ruibin Bai, Yang Yue, Qingquan Li, and Guoping Qiu</i>	
Automated Prediction of Shopping Behaviours Using Taxi Trajectory Data and Social Media Reviews.....	117
<i>Shuhui Gong, John Cartlidge, Ruibin Bai, Yang Yue, Qingquan Li and Guoping Qiu</i>	
Reliability Enhanced Data Processing System Capable of Runtime Fault Recovery.....	122
<i>Beichen Shen, Libiao Peng, Yongle Xie</i>	
Evaluate Academic Influence of Research Teams in Scholarly Data.....	128
<i>Kaiyu Liu, Hong Qiao, Mingchun Zheng</i>	
Learning Subconcepts in User-Generated Text from Social Media Sites: A Case of Exploring Food Concept.....	133
<i>Jie Wang</i>	
Effective NBA Player Signing Strategies Based on Salary Cap and Statistics Analysis.....	138
<i>Ramya Nagarajan, Yihang Zhao, Lin Li</i>	
Dimensional Modeling of Medical Data Warehouse Based on Ontology.....	144
<i>Shuxia Ren, Teng Wang, Xu Lu</i>	
Application of Ontology in Medical Heterogeneous Data Integration.....	150
<i>Shuxia Ren, Xu Lu, Teng Wang</i>	
MES: A Memory-Efficient Key-Value Storage With User-Level Network Stack.....	156
<i>Ninggou Yang, Junpeng Liu, Peng Zhang, Chao Zheng, Qingyun Liu</i>	
<b>Big Data Processing and Application</b>	
A Parallel Algorithm for Matrix Fast Exponentiation Based on MPI.....	162
<i>Ji Zhang, Xu-chuan Zhou</i>	
Building the Federation of Cloud Service for Big Data.....	166
<i>Jian Shu, Changyong Liang, Binyou Wang, Jian Xu</i>	
Big Data and Foreign Language E-learning on Production-oriented Approach.....	170
<i>Wu Yanhui</i>	

The Analysis of Dynamic Relationship between the Environmental Regulation and Technical Innovation— Based on the Big Data of Chinese Paper-making Industry .....	175
<i>Yanjia Li, Yuandi Wang, Xiang Wang</i>	
Research on SVM Improved Algorithm for Large Data Classification .....	181
<i>Hong Dai</i>	
SMEs Transformation through Usage and Understanding of Big Data-Case Study: Spanish Restaurant IndustryMarina Mattera.....	186
<i>Marina Mattera</i>	
Research and Application on Occupational Hazards Regulation Big Data.....	190
<i>Yan Chen, Xin Li</i>	
A New Identification Method of the Market Downturns in Chinese A-shares Using Liquidity Network.....	195
<i>Qunce Li, Qinke Peng, Ling Chai, Muhammad Owais</i>	
Questioning the Efficient Markets Hypothesis: Big Data Evidence of Non-Random Stock Prices.....	201
<i>Paul J. Darwen</i>	
Design and Implementation of Monitoring Data Storage and Processing Scheme Based on Distributed Computing.....	206
<i>Yaping Chi, Yintan Yang, Ping Xu, Gefei Li, Shuhao Li</i>	
Load Balancing Greedy Algorithm for Reduce on Hadoop Platform .....	212
<i>Hui Xia</i>	
Hadoop-based College Student Behavior Warning Decision System .....	217
<i>Ge Su-Hui, Bai Cheng-Jie, Wan Quan</i>	
An Extensible Hadoop Framework for Monitoring Performance Metrics and Events of OpenStack Cloud.....	222
<i>Jianxi Yang, Chaoxiao Shen, Yaping Chi, Ping Xu, Wei Sun</i>	
Anti-Crawler Strategy and Distributed Crawler Based on Hadoop.....	227
<i>HongRu Wang, ChunFang Li, LingFei Zhang, MinYong Shi</i>	
<b>Digital Network and Data Analysis</b>	
Domain2Vec: Vector Representation of Mobile Server's Domain Based on Mobile User Visiting Sequences .....	232
<i>Wan Jiang, Ke Yu, Xiaofei Wu, Fang Wei, Binbin Wang</i>	
Population Modeling Approach U sing Complex Network Theory and Queueing Systems .....	237
<i>Minyu Feng, Hongqu</i>	
Noise Threshold Estimation in Spectrum Monitoring Data Analysis Application .....	242
<i>Min Chen, Rong Shi, Binbin He</i>	
Data Filtering Algorithm Based onAttribute Reduction and Gene Expression Programming.....	248
<i>Li Shiyue, Yue Dong, Deng Song , Zhang Liping</i>	
Study of Performance Comparison of Satellite Error Correction Codes for Correcting Big Burst Data Errors .....	254
<i>Bingrui Wang, Qingli Zhang</i>	
A Study on Tactics for College Website at Search Engine Optimization .....	259
<i>Wang Bin, Gao Bingyun, Liu Peishun, Li Xiaoqing</i>	
Detection of Super Nodes Based on Connection Metrics for Network Data Streams .....	264
<i>Aiping Zhou, Huisheng Zhu, Lijun Liu, Chengang Zhu</i>	
Social Network Information Visualization Based on Fruchterman Reingold Layout Algorithm .....	270
<i>Lin Bi, Yue Wang, Jian-ping Zhao, Hui Qi, Ying Zhang</i>	

A Noise Classification Algorithm Based on SAMME and BP Neural Network.....	274
<i>Chen Guoqiang, Zhang Yuying, Tang Deyou</i>	
A Message Authentication Scheme for VANETs Based on Trapdoor Hash Function .....	279
<i>Yang Xiaodong, An Faying, Yang Ping, Xiao Likun, Li Yutong, Ma Tingchun, Wang Caifen</i>	
Research and Simulation Implementation of Mobile Phone Communication Based on SL4A .....	283
<i>Huajun Lei</i>	
<b>Cloud Computing and Recommendation System</b>	
Research and Implementation of POI Recommendation System Integrating Temporal Feature.....	287
<i>Jinxin Liu, Xu Jiao, Youzhi Jin, Xinyu Liu, Li Liu</i>	
Research on Artist Intelligent Recommendation Method Based on Mapping Knowledge Domains-Data Mining Based on Artist Association Analysis .....	293
<i>Ge Su , Ke Ma, Fulian Yin</i>	
Design and Implementation of Scientific Research Information Recommendation System.....	297
<i>Zeja Pan, Ying Jiang, Tong Wang, Yudan Lin, Hongjun Li, Jiting Xu, Xin Zhao</i>	
Access Control of Cloud Users Credible Behavior Based on IPv6 .....	303
<i>Sisi Cao, Wenbo Zheng, Shaocong Mo, Xin Jin, Chao Daik, Zeyu Fan, Yili Qu, Jiangwei Zhou, Fei Deng</i>	
Performance Evaluation of Memory Usage Costs for Commercial Cloud Services .....	307
<i>Takahiro Koita, Yu Manabe</i>	
Attribute-based Proxy Re-encryption Method for Revocation in Cloud Storage: Reduction of Communication Cost at Re-encryption.....	312
<i>Yoshiko Yasumura, Hiroki Imabayashi, Hayato Yamana</i>	
Verification Method of Data Quality in Science and Technology Cloud in Shaanxi Province .....	319
<i>Bin Yu, Chen Zhang, Zhou Hua Tang, JiangYan Sun</i>	
<b>Computer and Information Engineering</b>	
App Usage Prediction Based on an Extended Feature Selection Method $RSPEC_{MIC}$ .....	324
<i>Yazhou Shi, Fang Wei, Ke Yu, Xiaofei Wu</i>	
The Evolution of Network Topology Structure of Chinese Stock Market .....	329
<i>Xuwei Cao, Yongbin Shi, Penghao Wang, LiuJun Chen, Yougui Wang</i>	
Improved Approaches for Financial Market Forecasting Based on Stationary Time Series Analysis.....	334
<i>Hongbo Sun, Jing Xu</i>	
A Topic-based Cross-language Retrieval Model with PLSA and TF-IDF.....	340
<i>Zhuo-Liang Huo, Jun-Feng Wu, Yao Lu, Chang-Zhou Li</i>	
Mixed Feature Expansion-based Short Text Classification.....	345
<i>Zhenrong Deng, Baojun Zhang, Jingjing Zhang</i>	
A Fast Registration from 3D CT Images to 2D X-ray Images .....	351
<i>Hanqian Pan, Chunlin Zhou, Qi Zhu, and Delei Zheng</i>	
Improved Deep Belief Network Model and Its Application in Named Entity Recognition of Chinese Electronic Medical Records.....	356
<i>Wusuo Li, Shenghui Shi, Ziqiao Gao, Wei Wei, Qunxiong Zhu, Xiaoyong Lin, Daguang Jiang, Shang Gao</i>	
The Effect of the Family Planning Policy on Population Loss in China’ s Provinces.....	361
<i>Wang Xiang, Hu Die, Wang Yuandi, Hu Ruifeng</i>	
Enhancing Privacy Using Location Semantics in Location Based Services .....	368
<i>Mingjie Ma</i>	

A Multi-tenant Play Script Management System for SAAS .....	374
<i>XueKe Liu, MinYong Shi, ChunFang Li</i>	
Detection of Dependence Between Two Groups of Variables in the Presence of Another Group .....	379
<i>Ah Hin Pooi</i>	
Analyzing Collusive and Opportunistic Voting Strategies in an Academic Department P & T Decision Using Low-Dimensional Spectral Embedding.....	383
<i>Sanjoy Das</i>	
Image Classification of Fine-grained Fashion Image Based on Style Using Pre-Trained Convolutional Neural Network .....	387
<i>Yian Seo, Kyung-shik Shin</i>	
<b>Power Electronics Technology and Reliability Prediction</b>	
PEMFC Power Prediction Based on Deep Auto-encoder and LS-SVMR .....	391
<i>Sheng Hong, Longlong Sun, JiaWei Yin, Ting Yu, Yifan Wang, Wei Zhu</i>	
An Improved Parallel Maximum Flow Approach for Vulnerability Analysis of Power System.....	397
<i>Yanlin Shi, Dafang Zhang, Zulong Diao</i>	
Early Anomaly Detection in Wind Turbine Bolts Breaking Problem – Methodology and Application.....	402
<i>Chung-Wei Wu, Mei Chen</i>	
Short Term Load Forecasting Using Wavelet Augmented Non-linear Autoregressive Neural Networks-A Single Customer Level Perspective .....	407
<i>Solomon Netsanet, Jianhua Zhang, Dehua Zheng</i>	
Remaining Useful Life Prediction of Electric Vehicle Lithium-Ion Battery Based on Particle Filter Method....	412
<i>Zachary Bosire Omariba, Lijun Zhang, Dongbai Sun</i>	
<b>Data Clustering and Algorithm</b>	
Study on Initial Sensitivity of Fuzzy Clustering Algorithm for Big Data Processing.....	417
<i>Enwen Li, Linong Wang, Bin Song, Ruyi Cai</i>	
K-Means Clustering Algorithm Based on Improved Cuckoo Search Algorithm and Its Application .....	422
<i>Shuce Ye, Xiaoli Huang, Yinyin Teng, Yuxia Li</i>	
Clustering of Public Transport Operation using K-Means .....	427
<i>Rabiah Abdul Kadir, Yasuki Shima, Riza Sulaiman, Fathelalem Ali</i>	
SparkSNN: A Density-based Clustering Algorithm on Spark .....	433
<i>Amar Mani Aryal, Sujing Wang</i>	
State Grid Office System User Clustering Analysis Based on K-Means Algorithm.....	438
<i>Ting Li, Peng Ye, Shikai Zheng</i>	
Protocol Reverse Based on Hierarchical Clustering and Probability Alignment from Network Traces.....	443
<i>Fanzhi Meng, Chunrui Zhang, Guo Wu</i>	
PCA-Based Hierarchical Clustering of AIS Trajectories with Automatic Extraction of Clusters .....	448
<i>Jing Cao, Maohan Liang, Yan Li, Jinwei Chen, Huanhuan Li, Ryan Wen Liu, and Jingxian Liu</i>	
Clustering Driving Trip Trajectory Data Based on Pattern Discovery Techniques .....	453
<i>Gene P. K. Wu, Keith C. C. Chan</i>	
Overlapping community detection algorithm based on Markov chain clustering.....	458
<i>GuoLiang Deng, ZhiXin Ma, XianBo Li</i>	
<b>Author Index</b> .....	463