

2016 IEEE First International Conference on Data Science in Cyberspace (DSC 2016)

**Changsha, China
13-16 June 2016**



**IEEE Catalog Number: CFP16E91-POD
ISBN: 978-1-5090-1193-3**

**Copyright © 2016 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:	CFP16E91-POD
ISBN (Print-On-Demand):	978-1-5090-1193-3
ISBN (Online):	978-1-5090-1192-6

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
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2016 IEEE First International Conference on Data Science in Cyberspace

DSC 2016

Table of Contents

Message from General Chairs.....	xiv
Conference Organization.....	xv
Program Committee.....	xvi

Main Conference: International Conference on Data Science in Cyberspace

CareDedup: Cache-Aware Deduplication for Reading Performance Optimization in Primary Storage	1
<i>Bin Lin, Shanshan Li, Xiangke Liao, Xiaodong Liu, Jing Zhang, and Zhouyang Jia</i>	
A Measurement Study on Mainline DHT and Magnet Link	11
<i>Zhang Xinxing, Tian Zhihong, and Zhang Luchen</i>	
The Prediction of User Topic Interest Based on Tags and Interaction of Users	20
<i>Lu Deng, Jiuming Huang, Yi Han, Bin Zhou, and Qiang Liu</i>	
Social Recommendation with Tag Side Information	27
<i>Xiang Hu, Wendong Wang, Xiangyang Gong, Bai Wang, Xirong Que, and Hongke Xia</i>	
Chinese Article Classification Oriented to Social Network Based on Convolutional Neural Networks	33
<i>Xiang Zhu, Jiuming Huang, Zhongcheng Zhou, and Yi Han</i>	
Predicate-Oriented Query of RDF Data Based on a Distributed Storage Model	37
<i>Xuling Luo and Bin Wu</i>	
LinkSHRINK: Overlapping Community Detection with Link-Graph	44
<i>Dingyi Yin, Bin-Wu, and Yunlei Zhang</i>	
Finding Experts in Community Question Answering Based on Topic-Sensitive Link Analysis	54
<i>Juan Yang, Shuang Peng, Lin Wang, and Bin Wu</i>	

Link Prediction-Based Multi-label Classification on Networked Data	61
<i>Yinfeng Zhao, Lei Li, and Xindong Wu</i>	
An Experimental Study on Block DCT Coefficient Analysis for Image Splicing Detection	69
<i>Xiang Lin, Shi-Lin Wang, Wei-Jun Huang, and Jun-Yao Lai</i>	
BPPGD: Budgeted Parallel Primal Gradient Descent Kernel SVM on Spark	74
<i>Jinchen Sai, Bai Wang, and Bin Wu</i>	
Parallelization of Latent Group Model for Group Recommendation Algorithm	80
<i>Xuelin Zeng, Bin Wu, Jing Shi, Chang Liu, and Qian Guo</i>	
LSTM-based Deep Learning Models for Answer Ranking	90
<i>Zhenzhen Li, Jiuming Huang, Zhongcheng Zhou, Haoyu Zhang, Shoufeng Chang, and Zhijie Huang</i>	
A Hybrid Document Feature Extraction Method Using Latent Dirichlet Allocation and Word2Vec	98
<i>Zhibo Wang, Long Ma, and Yanqing Zhang</i>	
A Dual Threshold Secret Sharing Scheme among Weighted Participants of Special Right	104
<i>Guozhen Shi, Yunfei Ci, Rongna Xie, Haojie Wang, and Jiqiang Zeng</i>	
Equalized Interval Centroid Based Watermarking Scheme for Stepping Stone Traceback	109
<i>Xiaoqiang Xu, Jing Zhang, and Qianmu Li</i>	
A Topic Model for Hierarchical Documents	118
<i>Yang Yang, Feifei Wang, Fei Jiang, Shuyuan Jin, and Jin Xu</i>	
Integrating Relationships and Attributes: A Model of Multilayer Networks	127
<i>Wen Zhou, Weidong Bao, Xiaomin Zhu, Ji Wang, and Chao Chen</i>	
Cognitive Detection of Multiple Discrete Emotions from Chinese Online Reviews	137
<i>Si Jiang and Jiayin Qi</i>	
Constructions of Uniquely 3-Colorable Graphs	143
<i>Zepeng Li and Jin Xu</i>	
Structural Vulnerability Analysis in Complex Networks Based on Core Theory	149
<i>Kan Zhang, Fei Jiang, Yang Zuo, and Yunyun Niu</i>	
Exploring New Cryptographical Construction of Complex Network Data	155
<i>Hongyu Wang, Jin Xu, and Bing Yao</i>	
The Similarity Analysis of Malicious Software	161
<i>Jing Liu, Yuan Wang, and Yongjun Wang</i>	
Uncovering and Characterizing Internet Water Army in Online Forums	169
<i>Guirong Chen, Wandong Cai, Jiuming Huang, and Xianlong Jiao</i>	

Frequent Pattern Mining Based on Approximate Edit Distance Matrix	179
<i>Dan Guo, Ermao Yuan, and Xuegang Hu</i>	
FTM: Recommending the Right Items for User Temporal Interests with Matrix Factorization through Topic Model	189
<i>Yanmin Shang, Kefu Xu, Yi Han, and Chuang Zhang</i>	
Clustering Product Features of Online Reviews Based on Nonnegative Matrix Tri-factorizations	199
<i>Wang Jiajia, Liu Yezheng, Jiang Yuanchun, Sun Chunhua, Sun Jianshan, and Du Yanan</i>	
A New Scheme Based on HSSL for Solving the Stochastic Point Location Problem	209
<i>Jinchao Huang, Yan Yan, Ying Guo, and Shenghong Li</i>	
A General Strategy for Solving the Stochastic Point Location Problem by Utilizing the Correlation of Three Adjacent Nodes	215
<i>Ying Guo, Hao Ge, Jinchao Huang, and Shenghong Li</i>	
Fingerprinting Web Browser for Tracing Anonymous Web Attackers	222
<i>Xiaofeng Liu, Qixu Liu, Xiaoxi Wang, and Zhaopeng Jia</i>	
On Adjacent Vertex-Distinguishing Total Chromatic Number of Generalized Petersen Graphs	230
<i>Enqiang Zhu, Fei Jiang, Zepeng Li, Zehui Shao, and Jin Xu</i>	
Using Structural Features to Characterize Social Ties	235
<i>Yang Zuo and Kan Zhang</i>	
A Framework of Privacy Decision Recommendation for Image Sharing in Online Social Networks	243
<i>Donghui Hu, Fan Chen, Xintao Wu, and Zhongqiu Zhao</i>	
Mining Individual Mobility Patterns Based on Location History	252
<i>Xiaopeng Chen, Dianxi Shi, Banghui Zhao, and Fan Liu</i>	
Influence Maximization in Social Networks Based on Non-backtracking Random Walk	260
<i>Jingzhi Pan, Fei Jiang, and Jin Xu</i>	
On Improving a Microblog Ranking	268
<i>Jidong Li, Xin Li, Mingming Shi, Meng Zhou, and Linjing Lai</i>	
Max-Flow Rate Priority Algorithm for Evacuation Route Planning	275
<i>Dan Guo, Chen Gao, Wu Ni, and Xuegang Hu</i>	
Role and Time-Based Access Control with Efficient Revocation for Cloud Storage	284
<i>Fenghua Li, Yanchao Wang, Jinbo Xiong, and Rongna Xie</i>	

Session W1: International Workshop on Information Diffusion in Social Networks

Design and Implementation of Scheduling Pool Scheduling Algorithm Based on Reuse of Jobs in Spark	290
<i>Tang Jianchao, Yang Shuqiang, Huang Chaoqiang, and Yan Zhou</i>	
Discovering Latent Influence in Online Social Retweet Behaviors	296
<i>Bo Sun, Chungjin Hu, Wenwen Xu, and Huixing Fan</i>	
OPSDS: A Semantic Data Integration and Service System Based on Domain Ontology	302
<i>Xin Liu, Chungjin Hu, Jianyi Huang, and Feng Liu</i>	
A Bidirectional LSTM Model for Question Title and Body Analysis in Question Answering	307
<i>Yuanping Nie, Chao An, Jiuming Huang, Zhou Yan, and Yi Han</i>	
Information Diffusion Mechanisms in Online Social Networks	312
<i>Shushen Fu, Chungjin Hu, Ying Hu, Bo Sun, Wenrui Ying, and Peng Shi</i>	
Question Similarity Modeling with Bidirectional Long Short-Term Memory Neural Network	318
<i>Chao An, Jiuming Huang, Shoufeng Chang, and Zhijie Huang</i>	
A Police Big Data Analytics Platform: Framework and Implications	323
<i>Hai Yu and Chungjin Hu</i>	

Session W2: International Workshop on Data Analysis and Security in Smart City

Convolutional Neural Network Based on Principal Component Analysis Initialization for Image Classification	329
<i>Xu-Die Ren, Hao-Nan Guo, Guan-Chen He, Xu Xu, Chong Di, and Sheng-Hong Li</i>	
Energy-Efficient Hadoop Green Scheduler	335
<i>Jianhong Zhai, Hongli Zhang, Xiaorou Zhong, Wei Li, Lai Wang, and Zeyu He</i>	
TREST: A Hadoop Based Distributed Mobile Trajectory Retrieval System	341
<i>Jianming Lv, Xingtong Wang, Fengtao Huang, Junjie Yang, Tianfeng Wu, and Qifa Yan</i>	
A Novel Android Malware Detection Method Based on Markov Blanket	347
<i>Xiaotian Zhang, Donghui Hu, Yuqi Fan, and Kui Yu</i>	
Shrinking the Sentiment Analysis for Signed Network Construction	353
<i>Shen Su, Hongli Zhang, Yue Zhang, Dongyang Zhan, and Junxi Guo</i>	

Session W3: International Workshop of Heterogeneous Information Network Analysis and Applications

Attribute Credibility Based Sybil Goup Detection in Online Social Networks	358
<i>Yechao Xia, Li Pan, Liang Shi, and Futai Zou</i>	
A Simple Method for Locating Topic Sources in Uncertainty Diffusion Networks	364
<i>Huang Jianyi, Hu Chungjin, Fang Mingzhe, Wu Tong, and Shi Peng</i>	
RDDShare: Reusing Results of Spark RDD	370
<i>Huang Chao-Qiang, Yang Shu-Qiang, Tang Jian-Chao, and Yan Zhou</i>	
A New Weighted Similarity Method Based on Neighborhood User Contributions for Collaborative Filtering	376
<i>Xuefeng Zang, Tianqi Liu, Shuyu Qiao, Wenzhu Gao, Jiatong Wang, Xiaoxin Sun, and Bangzuo Zhang</i>	
Automatic Threshold Calculation Based Label Propagation Algorithm for Overlapping Community	382
<i>Gongshen Liu, Kui Meng, Hongyi Guo, Li Pan, and Jianhua Li</i>	
Remanufacturing Closed-Loop Supply Chain Model with RFID Technology Saving Recycling Cost	388
<i>Yang Ai-Feng, Wan Si, and Hu Xiao-Jian</i>	
A News Event Detection Algorithm Based on Key Elements Recognition	394
<i>Xiaoting Qu, Juan Yang, Bin Wu, and Haiming Xin</i>	
Online Topic Evolution Modeling Based on Hierarchical Dirichlet Process	400
<i>Tao Ma, Dacheng Qu, Rui Ma, Wei Feng, and Kan Li</i>	

Session W4: International Workshop on Big Search

The Multiple Attribute Association Decision-Making Method to Make Online Advertisements Using Influential Users in Social Network	406
<i>Jianmin He, Long Yu, and Yezheng Liu</i>	
Utilize Item Correlation to Improve Aggregate Diversity for Recommender Systems	412
<i>Liu Yezheng, Wang Jinkun, Jiang Yuanchun, Sun Jianshan, and Sun Chunhua</i>	
Community Detection Based on Variable Vertex Influence	418
<i>Yuntao Yao, Wei Wu, Mingtao Lei, and Xi Zhang</i>	
Efficient Privacy-Preserving Processing Scheme for Location-Based Queries in Mobile Cloud	424
<i>Qingqing Xie and Liangmin Wang</i>	
Distributed Storage Optimization for Small Data with High Density in Internet of Vehicles	430
<i>Hongbo Zhang, Huibing Zhang, and Xiaoli Hu</i>	

Research on Video Anti-hotlinking for OTT	435
<i>Dongyan Zhang, Zhiwen Yang, and Weihua Li</i>	
A Fine-Grained Multiparty Access Control Model for Photo Sharing in OSNs	440
<i>Chao Lee, Wei Wang, and Yunchuan Guo</i>	

Session W5: International Workshop on Data Visualization

Web Service Run-Time Monitoring and Visualization Analysis Based on Probe	446
<i>Liang Chen, Dapeng Xiong, Hua Wang, and Peng Zou</i>	
Research on Visualization Techniques in Large Scale Virtual Battlefield	452
<i>Wu Lingda, Hao Hongxing, Yang Chao, Yu Ronghuan, and Hu Huaquan</i>	
Topology Analysis of Vector Fields and Application Prospect	456
<i>Li Chao, Wu Lingda, Yang Jia, and Zhao Bin</i>	
The Cloud Design of the Cognitive Virus Based on SaaS	462
<i>Sun Yang and Xiong Wei</i>	
Research on Network Simplification by Edge Bundling	466
<i>Yao Zhonghua and Wu Lingda</i>	
Low-Cost INS Velocity Kalman Filtering Based on Rational Fitting	473
<i>Pei Dong and Qin Daguo</i>	
An Expandable Community Division Method for Network Visualization	478
<i>Xiangang Wang and Hanchen Song</i>	
A Study on Route Planning of Helicopter in Low Altitude Area	484
<i>Liyun Hao, Chao Guo, and Lingda Wu</i>	
A Hybrid Modeling Method for Dynamic Liquid Simulation	489
<i>Ling Zou and Guoping Wang</i>	
The Visualization Analysis and Vulnerability Repair Research for the Module Dependency Managerial of VxWorks 5.5 Operating System	493
<i>Peng Wang, Liang Chen, Peng Zou, Li Li, and Junlei Bao</i>	

Session W6: International Workshop on Big Data and Business Analytics

An Empirical Study on the Performance of Enterprises and Investment in the Human Capital	497
<i>Huaming Wu and Jin Hang</i>	
Combination Forecast on Health Status of Residents in China	503
<i>Chunlei Han, Shuangshuang Wang, Kun Han, and Xihou Hu</i>	
A Personalized Microblog Search Model Considering User-Author Relationship	508
<i>Yuanchun Jiang, Yuxiang Xu, and Liang Shao</i>	

An Empirical Business Study on Service Providers' Satisfaction in Sharing Economy	514
<i>Mengyu Zhang, Xusen Cheng, Xuan Luo, and Shixuan Fu</i>	
Complementarity: A Novel Collaborator Recommendation Method for SMEs	520
<i>Wei Xu, Ying Lu, Jing Zhao, and Minghui Qian</i>	
Detection and Defense of SYN Flood Attacks Based on Dual Stack Network Firewall	526
<i>Ding Pengfule, Tian Zhihong, Zhang Hongli, Wang Yong, Zhang Liang, and Guo Sanchuan</i>	
GEV Regression with Convex Loss Applied to Imbalanced Binary Classification	532
<i>Haolin Zhang, Gongshen Liu, Li Pan, Kui Meng, and Jianhua Li</i>	
Behavior Analysis Based SMS Spammer Detection in Mobile Communication Networks	538
<i>Zhang Bin, Zhao Gang, Feng Yunbo, Zhang Xiaolu, Jiang Weiqiang, Dai Jing, and Gao Jiafeng</i>	
Unknown Word Detection in Song Poetry	544
<i>Xia Li, Bin Wu, and Bailing Zhang</i>	
A Real Time EEG Analysis System	550
<i>Jonathan Garza, Yuezhe Li, Yuchou Chang, and Hong Lin</i>	

Session W7: International Workshop on Privacy for Big Data

LRDM: Local Record-Driving Mechanism for Big Data Privacy Preservation in Social Networks	556
<i>Weihao Li and Hui Li</i>	
A Voronoi-Based Dummy Generation Algorithm for Privacy-Aware Location-Based Services	561
<i>Cui Zhang and Fenghua Li</i>	
A Secure Communication Protocol with Privacy-Preserving Monitoring and Controllable Linkability for V2G	567
<i>Rong Jiang, Rongxing Lu, Chengzhe Lai, and Aiping Li</i>	
Real-Time Traffic Status Classification Based on Gaussian Mixture Model	573
<i>Xiong Liu, Li Pan, and Xiaoliang Sun</i>	
New Properties and Bounds of Anti-average Numbers	579
<i>Yangyang Zhou, Jin Xu, and Bing Yao</i>	
A Novel APPs Recommendation Algorithm Based on APPs Popularity and User Behaviors	584
<i>Liu Yezheng, Du Fei, Jiang Yuanchun, Liu Xiao, and Wang Qiudan</i>	

Mobile Authentication System Based on National Regulation and NFC Technology	590
<i>Chengjun Cai, Jian Weng, and Jianan Liu</i>	
Differentially Private Publication Scheme for Trajectory Data	596
<i>Meng Li, Liehuang Zhu, Zijian Zhang, and Rixin Xu</i>	
Survey on Domain Name System Security	602
<i>Futai Zou, Siyu Zhang, Bei Pei, Li Pan, Linsen Li, and Jianhua Li</i>	

Session W8: International Workshop on Social Recommendation Systems

Time-Aware First Story Detection in Twitter Stream	608
<i>Yongqin Qiu, Sixu Li, Wenjing Yang, Rui Li, Lihong Wang, and Bin Wang</i>	
Ranking-Based Music Recommendation in Online Music Radios	614
<i>Yao Lu, Zhi Qiao, Peng Zhang, and Li Guo</i>	
A Fast Algorithm for Competitive Recommendation Marketing Strategy	620
<i>Wenyu Zang, Xiao Wang, and Yue Hu</i>	
Image Super-Resolution with Deep Convolutional Neural Network	626
<i>Xiancai Ji, Yao Lu, and Li Guo</i>	
A Survey of Game Theoretic Methods for Cyber Security	631
<i>Yuan Wang, Yongjun Wang, Jing Liu, Zhijian Huang, and Peidai Xie</i>	
Pursuit Estimator Learning Automata Based Approach for Online Event Pattern Tracking	637
<i>Wen Jiang, Hao Ge, Tianrong Wu, Fanming Wang, and Shenghong Li</i>	

Session W9: International Workshop on Open Source Web Data

Survey on Software Vulnerability Analysis Method Based on Machine Learning	642
<i>Gong Jie, Kuang Xiao-Hui, and Liu Qiang</i>	
Correlation Analysis Using Global Dataset of Events, Location and Tone	648
<i>Kedi Chen, Fengcai Qiao, and Hui Wang</i>	
Sentiment Classification of Chinese Microblogging Texts with Global RNN	653
<i>Jiajun Cheng, Pei Li, Zhaoyun Ding, Sheng Zhang, and Hui Wang</i>	
Finding Influential Papers in Citation Networks	658
<i>Sheng Zhang, Danling Zhao, Ran Cheng, Jiajun Cheng, and Hui Wang</i>	
Probability-Weighted Extreme Learning Machine for Classification with Uncertain Data	663
<i>Hang Gao, Yuxing Peng, and Songlei Jian</i>	

Link Prediction Based on Clustering Information in Scientific Coauthorship Networks	668
<i>Yang Ma, Guangquan Cheng, Zhong Liu, and Xingxing Liang</i>	
Open Relation Extraction from Chinese Microblog Text	673
<i>Jing Xu, Liang Gan, Zhou Yan, Quanyuan Wu, and Yan Jia</i>	
A Topic Label Extraction Method for the University BBS	678
<i>Wenling Tang, Xu Wu, Yuxiao Li, and Jin Xu</i>	
New Cumulative Distributions for Scale-Free Networks	683
<i>Xiaomin Wang, Bing Yao, Hongyu Wang, and Jin Xu</i>	
A Fast and High Quality Approach for Overlapping Community Detection through Minimizing Conductance	688
<i>Yang Gao, Hongli Zhang, and Yue Zhang</i>	
A Fingerprint Identification Algorithm Based on Local Minutiae Topological Property	694
<i>Sun Bei, Luo Wusheng, Du Liebo, and Lu Qin</i>	
Detecting Malicious Server Based on Server-to-Server Realation Graph	698
<i>Zihao Wang, Futai Zou, Bei Pei, Weijia He, Li Pan, Zhaochong Mao, and Linsen Li</i>	
Author Index	703
Combing Semantic and Prior Polarity Features for Boosting Twitter Sentiment Analysis Using Ensemble Learning	709
<i>Zhao Jianqiang</i>	