2014 IEEE Fourth International Conference on Big Data and Cloud Computing

(BdCloud 2014)

Sydney, Australia
3-5 December 2014
2014 IEEE Fourth International Conference on Big Data and Cloud Computing

BDCloud 2014

Table of Contents

Message from BDCloud Chairs.................................................................................................................. xv
BDCloud 2014 Conference Organization.................................................................................................. xvii
BDCloud 2014 Program Committee........................................................................................................ xix
Message from SocialCom Chairs ........................................................................................................... xxii
SocialCom 2014 Conference Organization ............................................................................................... xxiv
SocialCom 2014 Program Committee....................................................................................................... xxv
Message from SustainCom Chairs ........................................................................................................... xxvi
SustainCom 2014 Conference Organization ............................................................................................. xxvii
SustainCom 2014 Program Committee .................................................................................................... xxviii
Message from PriSec Chairs.................................................................................................................... xxix
PriSec 2014 Symposium Organization .................................................................................................. xxxi
PriSec 2014 Program Committee ........................................................................................................... xxxii
Message from GSN Chairs ....................................................................................................................... xxxiii
GSN 2014 Workshop Organization .......................................................................................................... xxxiv
GSN 2014 Program Committee ............................................................................................................ xxv

The Fourth IEEE International Conference on Big Data and Cloud Computing (BDCloud 2014)

Efficient Storage of Big-Data for Real-Time GPS Applications ................................................................. 1
Pavan Kumar Akulakrishna, Lakshmi J, and Nandy S K

Real Time Routing in Road Networks ..................................................................................................... 9
Aakriti Gupta, Lakshmi J., and S. K. Nandy

Tahoe-LAFS Distributed Storage Service in Community Network Clouds ............................................. 17
Mennan Selimi and Felix Freitag
Event Pattern Discovery on IDS Traces of Cloud Services .................................................................25  
Shin-Ying Huang, Yennun Huang, and Neeraj Suri

A New Approach Based on Intelligent Water Drops Algorithm for Node 
Selection in Service-Oriented Wireless Sensor Networks ..............................................................33  
Ahmadreza Vajdi, Gongxuan Zhang, Yongli Wang, Yang Zhao, Dongmei Liu, 
and Tianshu Wang

Automating Deployment of Customized Scientific Data Analytic Environments 
on Clouds ..............................................................................................................................................41  
Chao Jin, Wenjun Wu, and Hui Zhang

A Domain-Driven, Generative Data Model for Big Pet Store ..................................................................49  
Ronald J. Nowling and Jay Vyas

Dynamic Workload Balancing for Hadoop MapReduce .....................................................................56  
Xiaofei Hou, Ashwin Kumar T K, Johnson P Thomas, and Vijay Varadharajan

A Stop Planning Method over Big Traffic Data for Airport Shuttle Bus .................................................63  
Yan Liu, Guochao Jia, Xu Tao, Xiaolong Xu, and Wanchun Dou

A Neural Network Based Pre-Selection of Big Data in Photon Science ......................................................71  
Daniel Becker and Achim Streit

A Cloud Model for Distributed Transport System Integration ..................................................................77  
Regin Paul, Margaret Hamilton, and Daryl D’Souza

Efficient Pre-copy Live Migration with Memory Compaction and Adaptive VM 
Downtime Control ...................................................................................................................................85  
Guangyong Piao, Youngsup Oh, Baegjae Sung, and Chanik Park

Practical Analysis of Big Acoustic Sensor Data for Environmental Monitoring ............................................91  
Anthony Truskinger, Mark Cottman-Fields, Philip Eichinski, Michael Towsey, 
and Paul Roe

RAID-Aware SSD: Improving the Write Performance and Lifespan of SSD 
in SSD-Based RAID-5 System ..................................................................................................................99  
Xiaoquan Wu, Nong Xiao, Fang Liu, Zhiguang Chen, Yimo Du, and Yuxuan Xing

Fault Tolerant Erasure Coded Replication for HDFS Based Cloud Storage ..............................................104  
Aye Chan Ko and Wint Thida Zaw

Optimal Distributed Data Warehouse System Architecture .......................................................................110  
Mehdi Kashfi and Abdolreza Hajmoosaei

Secure Index Construction for Privacy-Preserving Large-Scale Image 
Retrieval ....................................................................................................................................................116  
Bo Cheng, Li Zhuo, Yu Bai, Yuanfan Peng, and Jing Zhang

A Paralleled Big Data Algorithm with MapReduce Framework for Mining 
Twitter Data ..............................................................................................................................................121  
LI Bing and Keith C.C. Chan
### Data-Intensive Workflow Optimization Based on Application Task Graph
Partitioning in Heterogeneous Computing Systems .................................................................129

* Saima Gulzar Ahmad, Chee Sun Liew, M. Mustafa Rafique, Ehsan Ullah Munir, and Samee U. Khan

### Remote Monitoring System Enabling Cloud Technology upon Smart Phones and Inertial Sensors for Human Kinematics ..........................................................137

* M. Sajeewani Karunarathne, Samuel A. Jones, Samitha W. Ekanayake, and Pubudu N. Pathirana

### High-Performance Processing of Large-Scale Parallel Applications in Heterogeneous Cloud Computing Data Centers ..........................................................143

* Awada Uchechukwu, Keqiu Li, and Keqin Li

### Cloud-Based Educational Big Data Application of Apriori Algorithm and K-Means Clustering Algorithm Based on Students’ Information ........................................151

* Jiaqu Yi, Sizhe Li, Maomao Wu, H.H. Au Yeung, Wilton W.T Fok, Ying Wang, and Fang Liu

### Shared I/O Scheduling in Cloud for Structured Data Processing ........................................159

* Baoquan Zhang, Jingmei Li, Tao Xu, Dongsheng Wang, and Nan Ding

### Protecting Access Confidentiality with Data Distribution and Swapping .............................167

* Sabrina de Capitani di Vimercati, Sara Foresti, Stefano Paraboschi, Gerardo Pelosi, and Pierangela Samarati

### Quantifying Failure Risk of Version Switch for Rolling Upgrade on Clouds ..........................175

* Daniel Sun, Len Bass, Alan Fekete, Vincent Gramoli, An Binh Tran, Sherry Xu, and Liming Zhu

### Multilingual Sentiment Classification on Large Textual Data .............................................183

* Jantima Polpinij

### Congestion Score Computation of Big Traffic Data ............................................................189

* Jiwan Lee

### A Hadoop-Based Output Analyzer for Large-Scale Simulation Data ....................................197

* Kangsun Lee and Joonho Park

### Variation-Aware Resource Allocation Evaluation for Cloud Workflows Using Statistical Model Checking ..........................................................201

* Saijie Huang, Mingsong Chen, Xiao Liu, Dehui Du, and Xiaohong Chen

### Extracting Trends of Traffic Congestion Using a NoSQL Database ......................................209

* Titus Irma Damaiyanti, Ardi Imawan, and Joonho Kwon

### Context-Aware Data Prefetching in Mobile Service Environments .......................................214

* Waldemar Hummer, Stefan Schulte, Philipp Hoenisch, and Schahram Dustdar

### A Resource Allocation Strategy for Multimedia Cloud Using Game Theory .......................222

* Yirui Li and Li Zhuo

---

vii
Evaluation of Linux I/O Schedulers for Big Data Workloads .................................................................227
    Abdelmounaam Rezgui, Matthew White, Sami Rezgui, and Zaki Malik

A Data Science Solution for Mining Interesting Patterns from Uncertain Big Data .................................................................235
    Carson Kai-Sang Leung and Fan Jiang

Improve User Experience on Web for Machine Translation System Using Storm .................................................................243
    Rashid Ahmad, Pawan Kumar, Ashutosh Kumar, Mukul K. Sinha, and B. D. Chaudhary

A Study on BPaaS with TCO Model ......................................................................................................................249
    Thi My Hanh Le, Luis Alfaro Alfredo, Hyung Rim Choi, Min Je Cho, and Chae Soo Kim

A Framework for Tracking Reliable Data in the Cloud for Port Logistics .................................................................257
    Luis Alfredo Alfaro, Thi My Hanh Le, Hyung Rim Choi, Min Je Cho, and Chae Soo Kim

A Multi-dimensional Weighting Method for Historical Records in Cloud Service Evaluation .................................................................265
    Lianyong Qi, Jiancheng Ni, Xiaona Xia, Hua Wang, and Chao Yan

Distributed Data Stream Processing with Onix ........................................................................................................267
    Roman Y. Shtykh and Toshihiro Suzuki

Interoperating Cloud Services for Enhanced Data Management ..................................................................................269
    Attila Kertesz

UDaaS: A Cloud-Based URL-Deduplication-as-a-Service for Big Datasets .................................................................271
    Shams Zawoad, Ragib Hasan, Gary Warner, and Anthony Skjellum

Hybrid Cache Architecture Using Big Data Analysis for Content Delivery Network .................................................................273
    Tai-Yeon Ku, John D. Shin, Young-Sik Chung, and Hoon Choi

Assessment of DM-cache Running on Virtual Linux ........................................................................................................275
    Jaemyoun Lee and Kyungtae Kang

Detection of Web Spambot in the Presence of Decoy Actions ..................................................................................277
    Vida Ghanaei, Costas S. Iliopoulos, and Solon P. Pissis

Modeling, Optimization and Performance Evaluation of Scientific Workflows in Clouds .................................................................280
    Kamil Figiela and Maciej Malawski

Filess – File-Less Architecture for Future Information Systems ..................................................................................281
    Bartosz Kryza and Jacek Kitowski

Towards Adaptable Data Farming in Clouds ...............................................................................................................283
    Dariusz Król and Jacek Kitowski
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using Accumulo for Graph Twiddling</td>
<td>285</td>
</tr>
<tr>
<td>Darren Webb</td>
<td></td>
</tr>
<tr>
<td>The Seventh IEEE International Conference on Social Computing and Networking (SocialCom 2014)</td>
<td></td>
</tr>
<tr>
<td>Item Recommendation Using Collaborative Filtering in Mobile Social Games:</td>
<td>293</td>
</tr>
<tr>
<td>Zhaojie Tao, Ming Cheung, James She, and Ringo Lam</td>
<td></td>
</tr>
<tr>
<td>Empirical Analysis of Workflow Patterns for Use in Knowledge Advantage Machines</td>
<td>298</td>
</tr>
<tr>
<td>Dan Sloan, Ramana Reddy, and Sumitra Reddy</td>
<td></td>
</tr>
<tr>
<td>Utilizing Favorites Lists for Better Recommendations</td>
<td>303</td>
</tr>
<tr>
<td>Mustafa Abualsaud and Alex Thomo</td>
<td></td>
</tr>
<tr>
<td>Sentiment Enhanced Hybrid TF-IDF for Microblogs</td>
<td>311</td>
</tr>
<tr>
<td>Atakan Simsek and Pinar Karagoz</td>
<td></td>
</tr>
<tr>
<td>A Social and Popularity-Based Tag Recommender</td>
<td>318</td>
</tr>
<tr>
<td>Modou Gueye, Talel Abdessalem, and Hubert Nacke</td>
<td></td>
</tr>
<tr>
<td>Personalized Recommender System on Whom to Follow in Twitter</td>
<td>326</td>
</tr>
<tr>
<td>Masudul Islam, Chen Ding, and Chi-Hung Chi</td>
<td></td>
</tr>
<tr>
<td>Engagement in Motion: Exploring Short Term Dynamics in Page-Level Social Media Metrics</td>
<td>334</td>
</tr>
<tr>
<td>Benjamin Lucas, Ahmed Shamsul Arefin, Natalie Jane de Vries, Regina Berretta, Jamie Carlson, and Pablo Moscato</td>
<td></td>
</tr>
<tr>
<td>Discovery of Really Popular Friends from Social Networks</td>
<td>342</td>
</tr>
<tr>
<td>Fan Jiang, Carson Kai-Sang Leung, Dacheng Liu, and Aaron M. Peddle</td>
<td></td>
</tr>
<tr>
<td>The Potential Use of Multi-agent and Hybrid Data Mining Approaches in Social Informatics for Improving e-Health Services</td>
<td>350</td>
</tr>
<tr>
<td>Dharmendra Sharma and Fariba Shadabi</td>
<td></td>
</tr>
<tr>
<td>A Comparison of Common Users across Instagram and Ask.fm to Better Understand Cyberbullying</td>
<td>355</td>
</tr>
<tr>
<td>Homa Hosseinmardi Shaosong Li, Zhiyi Yang, Qin Lv, Rafiq Richard Han, and Shivakant Mishra</td>
<td></td>
</tr>
<tr>
<td>CORE Analysis for Efficient Shortest Path Traversal Queries in Social Graphs</td>
<td>363</td>
</tr>
<tr>
<td>Waqas Nawaz, Kifayat-Ullah Khan, and Young-Koo Lee</td>
<td></td>
</tr>
<tr>
<td>Proposal of Alleviative Method of Community Analysis with Overlapping Nodes</td>
<td>371</td>
</tr>
<tr>
<td>Atsushi Tanaka</td>
<td></td>
</tr>
<tr>
<td>Set-Based Unified Approach for Attributed Graph Summarization</td>
<td>378</td>
</tr>
<tr>
<td>Kifayat Ullah Khan, Waqas Nawaz, and Young-Koo Lee</td>
<td></td>
</tr>
</tbody>
</table>
Participatory Sensor Networks as Sensing Layers .................................................................386

Thiago H. Silva, Pedro O. S. Vaz de Melo, Jussara M. Almeida, Aline C. Viana,
Juliana Salles, and Antonio A. F. Loureiro

Socio Path: Protecting Privacy by Self-Sufficient Data Distribution
in User-Centric Networks .................................................................................................394

Fabian Hartmann and Ingmar Baumgart

Social-Based Multi-label Routing in Delay Tolerant Networks .....................................402

Linmao Song, Yang Li, Shuija Fan, and Xiumei Fan

Efficient Event Detection for the Blogosphere ...............................................................408

Patrick Hennig, Philipp Berger, Daniel Kurzynski, Hannes Rantzsch,
and Christoph Meinel

Cluster Labeling for the Blogosphere ........................................................................416

Patrick Hennig, Philipp Berger, Claus Steuer, Christia Wuerz,
and Christoph Meinel

An Enhanced Content-Based Recommender System for Academic Social
Networks ..........................................................................................................................424

Vala Ali Rohani, Zarinah Mohd Kasirun, and Kuru Ratnavelu

Friend News System: A Modern Implementation of Usenet over Social VPNs ..........432

Yasushi Shinjo, Xiao Kunyao, Naoki Kainuma, Daiyuu Nobori, and Akira Sato

Social Network Observatory for Innovation in Enterprise-Employee
Engagement ......................................................................................................................441

Jayashree Ramanathan, Zhe Xu, Ramiya Venkatchalalm, Rajiv Ramnath,
Bruce Craig, and Tara Paider

Characterization of the Use of Social Media in Natural Disasters: A Systematic
Review .............................................................................................................................449

Babak Abedin, Abdul Babar, and Alireza Abbasi

Incorporating User Reviews as Implicit Feedback for Improving Recommender
Systems ..........................................................................................................................455

Yasamin Heshmat Dehkordi, Alex Thomo, and Sudhakar Ganti

A Community-Structure Based Adaptively Optimized Link Prediction Algorithm .................................................................463

Zhaojun Yang, Jiayu Song, Zhaolong Huang, Xuzhen Zhu, and Hui Tian

An Improved Latent Dirichlet Allocation Model for Hot Topic Extraction ..................470

Guolong Liu, Xiaofei Xu, Ying Zhu, and Li Li

"Hey #311, Come Clean My Street!": A Spatio-temporal Sentiment Analysis
of Twitter Data and 311 Civil Complaints .....................................................................477

Ryan Eshleman and Hui Yang

Gauging Heterogeneity in Online Consumer Behaviour Data: A Proximity
Graph Approach .............................................................................................................485

Natalie Jane de Vries, Ahmed Shamsul Arefin, and Pablo Moscato
Centralized Payment System Using Social Networks Account ..........................................................493
   Alireza Beikverdi, InHwan Kim, and JooSeok Song

Identifying Communities of Trust and Confidence in the Charity and Not-for-Profit Sector: A Memetic Algorithm Approach ..........................................................500
   Leila Moslemi Naeni, Natalie Jane de Vries, Rodrigo Reis,
   Ahmed Shamsul Arefin, Regina Berretta, and Pablo Moscato

Foto2Events: From Photos to Event Discovery and Linking in Online Social Networks ..........................................................................................................................508
   Eliana J. Raad and Richard Chbeir

The Use of a Social Networking Site in the Facilitation of Internationalization in Higher Education: A Case Study Using the Actor Network Theory Perspective ..........................................................................................................................516
   Kim Keith and Jean-Paul Van Belle

Influence Level-Based Sybil Attack Resistant Recommender Systems ............................................524
   Giseop Noh and Hayoung Oh

Video Annotation with Aggregate Social Network Data ..........................................................................................................................532
   Georgios Palaiokrassas, Kleopatra Konstanteli, Athanasios Voulodimos,
   Konstantinos Psychas, Theodora Varvarigou, David Salama Osborne,
   and Efstathia Chatzi

Cognition-Based Semantic Annotation for Web Images ..........................................................................................................................540
   Jinbiao Jing, Xiangfeng Luo, Junyu Xuan, and Weidong Liu

Perception-Based Resilience: Accounting for the Impact of Human Perception on Resilience Thinking ..........................................................................................................................547
   Roberto Legaspi, Hiroshi Maruyama, Rungsiman Nararatwong, and Hitoshi Okada

Behavioral Strategies in Online Forums with Different Feedback Types ............................................555
   Sanja Tanasijevic, Klemens Böhm, and Karl-Martin Ehrhart

Uncovering Diffusion in Academic Publications Using Model-Driven and Model-Free Approaches ..........................................................................................................................564
   Minkyung Kim, David Newth, and Peter Christen

How Social Identity May Matter Most in Brand Crisis Management ..................................................572
   Rungsiman Nararatwong, Kotaro Okazaki, Hitoshi Okada, and Katsumi Inoue

Reducing Noises for Recall-Oriented Patent Retrieval ........................................................................579
   Wookey Lee, Carson Kai-Sang Leung, and Justin JongSu Song

A Dynamic Social Network Experiment with Multi-team Systems ..................................................587
   Andrew Pilny, Alex Yahja, Marshall Scott Poole, and Melissa Dobosh
The Fourth International Conference on Sustainable Computing and Communications (SustainCom 2014)

A High Efficient Real Time Data Aggregation Algorithm for WSNs .......................................................... 594
Tao Du, Shouning Qu, Qingbei Guo, and Zhe Qu

New Progress in Wind Prediction Based on Nonlinear Amendment ......................................................... 599
Yagang Zhang, Jingyun Yang, Kangcheng Wang, Zheng Zhao, Jinkang Liu, and Yinding Wang

Comparison of the Robustness of RNN, MPC and ANN Controller for Residential Heating System ..................... 604
Abbas Javed, Hadi Larijani, Ali Ahmdinia, and Rohinton Emmanuel

Sustainable Software System Engineering ........................................................................................................ 612
Stefanie Betz and Timm Caporale

Vehicular Travel Initiated Sustainable USB Mobile Charging and Travel Analytics System ........................................ 620
Akshay Atul Potnis, Harshavardhan Vivek Pandit, and Siddharth Sunil Deshpande

Intelligent Mobility for Smart Cities: Driver Behaviour Models for Assessment of Sustainable Transport ................. 625
Hussein Dia and Sakda Panwai

Energy Modeling of Virtual Machine Replication Schemes with Checkpointing in Data Centers ........................................ 633
Subrota K. Mondal and Jogesh K. Muppala

Defining Energy Consumption Plans for Data Querying Processes ........................................................................ 641
Ricardo Gonçalves, João Saraiva, and Orlando Belo

Energy-Efficient User Association in HetNets: An Evolutionary Game Approach .................................................. 648
Kaifeng Han, Dantong Liu, Yue Chen, and Kok Keong Chai

A Simplified Method of Measurement of Energy Consumption in Cloud and Virtualized Environment ...................... 654
I Made Murwantara and Behzad Bordbar

Power Consumption of Erbium Doped Fibre Amplified Links ........................................................................ 662
Kerry Hinton, Peng Wang, Peter Farrell, and Bipin Pilai

Blind Frequency Hopping Spectrum Estimation: A Bayesian Approach ............................................................ 669
Lifan Zhao, Lu Wang, Guoan Bi, and Haijian Zhang

Dependability and Resource Optimization Analysis for Smart Grid Communication Networks .................................. 676
Ming Xiang, Sotharith Tauch, and William Liu
A User Profile-Aware Policy-Based Management Framework for Greening the Cloud ..................................................................................................................................................682
  Fadi Alhaddadin, William Liu, and Jairo A. Gutiérrez

A Low Cost Implementation of Home Area Networks for Home Energy Management Systems ..................................................................................................................................................688
  Saira Hussain and Muhammad Jahangir Ikram

A Generic and Extensible Framework for Monitoring Energy Consumption of OpenStack Clouds .............................................................................................................................................................................696
  François Rossigneux, Laurent Lefevre, Jean-Patrick Gelas, and Marcos Dias de Assuncao

Towards Generalizing "Big Little" for Energy Proportional HPC and Cloud Infrastructures ..................................................................................................................................................703
  Violaine Villebonnet, Georges da Costa, Laurent Lefevre, Jean-Marc Pierson, and Patricia Stolf

Risk Prediction System Based on Risk Prediction Model with Complex Event Processing: Risk Prediction in Real Time on Complex Event Processing Engine ..................................................................................................................................................711
  Yoon-Ki Kim and Chang-Sung Jeong

A Scheme for Software Defined ORS Satellite Networking ............................................................................................................................................................716
  Jing Feng, Lei Jiang, Ye Shen, WeiJun Ma, and Min Yin

Energy-Efficiency Based Resource Allocation for D2D Communication and Cellular Networks ............................................................................................................................................................722
  Layanah AlWreikat, Rong Chai, and Osama M.F. Abu-Sharkh

Smart Cultural Heritage and Open Source as Regeneration of Historical Centers: Fruition, Conservation and Preservation in Current Methods of 2D/3D Digitization and 3D Modelling ............................................................................................................................................................729
  Alessandro de Masi

The Third International Symposium on Privacy and Security in Cloud and Big Data

Fine Grain Cross-VM Attacks on Xen and VMware ............................................................................................................................................................737
  Gorka Irazoqui, Mehmet Sinan Inci, Thomas Eisenbarth, and Berk Sunar

On the Difficulty of Securing Web Applications Using CryptDB ............................................................................................................................................................745
  Ihsan H. Akin and Berk Sunar

Realizing Purpose-Based Privacy Policies Succinctly via Information-Flow Labels ............................................................................................................................................................753
  N.V. Narendra Kumar and R.K. Shyamasundar
The Third International Workshop on Graph Databases and Social Networking

A Multi-dimensional Analysis and Data Cube for Unstructured Text and Social Media ..................................................................................................................761

  Suan Lee, Namsoo Kim, and Jinho Kim

A Person Identification Method in CUG Using Voice Pitch Analysis ............................................................765

  So-Hyun Park, Young-Ho Park, Aziz Nasridinov, and Joo-Yeoun Lee

Path Pattern Query Processing on Large Graphs .........................................................................................767

  Yiyuan Bai, Chaokun Wang, Xiang Ying, Meng Wang, and Yunqing Gong

A GPU-accelerated Density-Based Clustering Algorithm ........................................................................775

  Woong-Kee Loh and Young-Kuk Kim

On the Graph Decomposition .......................................................................................................................777

  Yangjun Chen and Yibin Chen

Author Index ..................................................................................................................................................785