

2014 IEEE/ACM International Symposium on Big Data Computing (BDC 2014)

**London, United Kingdom
8 – 11 December 2014**



**IEEE Catalog Number: CFP14B46-POD
ISBN: 978-1-4799-1898-0**

2014 International Symposium on Big Data Computing

BDC 2014

Table of Contents

Conference Organization.....	vii
Program Committee.....	viii

Full Papers

Soccer Video Summarization Using Video Content Analysis and Social Media Streams	1
<i>Said Jai-Andaloussi, Aboulaite Mohamed, Nabil Madrane, and Abderrahim Sekkaki</i>	
Progressive Data Stream Mining and Transaction Classification for Workload-Aware Incremental Database Repartitioning	8
<i>Joarder Mohammad Mustafa Kamal, Manzur Murshed, and Mohamed Medhat Gaber</i>	
A Scalable Data Science Workflow Approach for Big Data Bayesian Network Learning	16
<i>Jianwu Wang, Yan Tang, Mai Nguyen, and Ilkay Altintas</i>	
Big Data Staging with MPI-IO for Interactive X-ray Science	26
<i>Justin M. Wozniak, Hemant Sharma, Timothy G. Armstrong, Michael Wilde, Jonathan D. Almer, and Ian Foster</i>	
GeoLens: Enabling Interactive Visual Analytics over Large-Scale, Multidimensional Geospatial Datasets	35
<i>Jared Koontz, Matthew Malensek, and Sangmi Pallickara</i>	
A Covering Classification Rule Induction Approach for Big Datasets	45
<i>Vasilis Koliass, Ioannis Anagnostopoulos, and Eleftherios Kayafas</i>	
Are Clouds Ready to Accelerate Ad Hoc Financial Simulations?	54
<i>Blesson Varghese and Adam Barker</i>	
A Practical Approach to Scalable Big Data Computing for the Personalization of Services at Samsung	64
<i>Jong Hoon Ahn</i>	

Short Papers

Toward Optimal Resource Provisioning for Cloud MapReduce and Hybrid Cloud Applications	74
<i>Arkaitz Ruiz-Alvarez and Marty Humphrey</i>	
Characterizing the Communication Demands of the Graph500 Benchmark on a Commodity Cluster	83
<i>Pablo Fuentes, José Luis Bosque, Ramón Beivide, Mateo Valero, and Cyriel Minkenberg</i>	
Monitoring Data Streams at Process Level in Scientific Big Data Batch Clusters	90
<i>Eileen Kuehn, Max Fischer, Christopher Jung, Andreas Petzold, and Achim Streit</i>	
Genetic Algorithm Based Data-Aware Group Scheduling for Big Data Clouds	96
<i>Raghavendra Kune, Pramod Kumar Konugurthi, Arun Agarwal, Raghavendra Rao Chillarige, and Rajkumar Buyya</i>	
The Impact of Basic Matrix Factorization Refinements on Recommendation Accuracy	105
<i>Parisa Lak, Bora Caglayan, and Ayse Basar Bener</i>	
A Biologically-Inspired Approach to Network Traffic Classification for Resource-Constrained Systems	113
<i>Brian Schmidt, Dionysios Kountanis, and Ala Al-Fuqaha</i>	
Author Index	119