

**2015 31st IEEE International
Conference on Data
Engineering Workshops
(ICDEW 2015)**

**Seoul, South Korea
13-17 April 2015**



**IEEE Catalog Number: CFP1545A-POD
ISBN: 978-1-4799-8443-5**

CloudDM 2015 — First Workshop on Cloud Data Management - Joint Workshop of Data Management in the Cloud (DMC) and CloudDB

- 1 **CloudDM 2015 Preface Pages**
- 4 **A Biased Guide to Building a Web-Scale OLTP Service**
(Mark Callaghan)
- 5 **The Service Contract Between Cloud Provider and Consumer: Consistency Properties**
(Alan Fekete)
- 6 **Metadata-as-a-Service**
(Akon Dey, Gajanan Chinchwadkar, Alan Fekete, Krishna Ramachandran)
- 10 **Dynamic Resource Management in a MapReduce-Style Platform for Fast Data Processing**
(Kasper Grud Skat Madsen, Yongluan Zhou)
- 14 **Spotgres — Parallel Data Analytics on Spot Instances**
(Carsten Binnig, Abdallah Salama, Erfan Zamanian, Muhammad El-Hindi, Sebastian Feil, Tobias Ziegler)
- 22 **AIR: Adaptive Index Replacement in Hadoop**
(Stefan Schuh, Jens Dittrich)
- 30 **Managing Information for Personal Goals (Vision)**
(Alan Fekete, Judy Kay, Michael Franklin, Debjane Barua, Bob Kummerfeld)
- 34 **Large-Scale Spatial Join Query Processing in Cloud**
(Simin You, Jianting Zhang, Le Gruenwald)
- 42 **The Era of Big Spatial Data**
(Ahmed Eldawy, Mohamed F. Mokbel)

DESWeb 2015 — Data Engineering Meets the Semantic Web

- 50 **DESWeb 2015 Preface Pages**
- 53 **Graph Mining Meets the Semantic Web**
(Sangkeun Lee, Sreenivas R. Sukumar, Seung-Hwan Lim)
- 59 **Challenges in Chinese Knowledge Graph Construction**
(Chengyu Wang, Ming Gao, Xiaofeng He, Rong Zhang)
- 62 **SemTree: An Index for Supporting Semantic Retrieval of Documents**
(Flora Amato, Aniello De Santo, Francesco Gargiulo, Vincenzo Moscato, Fabio Persia, Antonio Picariello, Silvestro Roberto Poccia)
- 68 **Towards Web-Scale How-Provenance**
(Daniel Deutch, Amir Gilad, Yuval Moskovitch)
- 71 **Efficient OLAP Operations for RDF Analytics**
(Elham Akbari Azirani, François Goasdoué, Ioana Manolescu, Alexandra Roatis)
- 77 **Big RDF Data Cleaning**
(Nan Tang)

BioBAD 2015 — Big Data Analytics for Biosecurity

- 80 **BioBAD 2015 Preface Pages**
- 83 **Multi-Scale Population and Mobility Estimation with Geo-Tagged Tweets**
(Jiajun Liu, Kun Zhao, Saeed Khan, Mark Cameron, Raja Jurdak)
- 87 **United, We Stand: Combining Cross-Governmental Data Resources to Refine Border Activities**
(Sandy Clarke, Alan Küffer, Greg Hood, Andrew Robinson)
- 92 **Ant Based Defense Mechanism for Selective Forwarding Attack in MANET**
(S. Vadhana Kumari, B. Paramasivan)
- 98 **Detection of Oil Pollution in Seawater: Biosecurity Prevention Using Electronic Nose Technology**
(Rob Chandler, Aruneema Das, Tim Gibson, Ritaban Dutta)
- 101 **Farm Biosecurity Hot Spots Prediction Using Big Data Analytics**
(Cecil Li, Ritaban Dutta, Daniel Smith, Aruneema Das, Jagannath Aryal)

DAMASCA 2015 — Data Mining and Smart Cities Applications Workshop

- 105 **DAMASCA 2015 Preface Pages**
- 108 **Smart City and Geospatiality: Hobart Deeply Learned**
(Jagannath Aryal, Ritaban Dutta)
- 110 **Managing Travels with PETRA: The Rome Use Case**
(Adi Botea, Stefano Braghin, Nuno Lopes, Riccardo Guidotti, Francesca Pratesi)
- 112 **Monitoring the Citizens' Perception on Urban Security in Smart City Environments**
(Luca Cagliero, Tania Cerquitelli, Silvia Chiusano, Pierangelo Garino, Marco Nardone, Barbara Pralio, Luca Venturini)
- 117 **POP: A Passenger-Oriented Partners Matching System**
(Xiaoyi Duan, Cheqing Jin, Xiaoling Wang)
- 119 **Data Mining for Safety Transportation by Means of Using Internet Survey**
(Masahiro Miyaji)
- 124 **On Crowdsensed Data Acquisition Using Multi-Dimensional Point Processes**
(Saket Sathe, Timos Sellis, Karl Aberer)
- 129 **Discovering Underpasses from Walking Trajectories**
(Qiuge Song, Cheqing Jin, Xiaoling Wang, Ming Gao, Aoying Zhou)
- 134 **Hotel Recommendation Based on User Preference Analysis**
(Kai Zhang, Keqiang Wang, Xiaoling Wang, Cheqing Jin, Aoying Zhou)

HardBD 2015 — International Workshop on Big Data Management on Emerging Hardware

- 139 **HardBD 2015 Preface Pages**
- 142 **Recent Advances in Flash Solutions**
(Sangyeun Cho)
- 143 **Scalable and Efficient Spatial Data Management on Multi-Core CPU and GPU Clusters: A Preliminary Implementation Based on Impala**
(Simin You, Jianting Zhang, Le Gruenwald)
- 149 **Optimizing CPU Cache Performance for Pregel-Like Graph Computation**
(Songjie Niu, Shimin Chen)
- 155 **Query Processing on Low-Energy Many-Core Processors**
(Annett Ungethüm, Dirk Habich, Tomas Karnagel, Wolfgang Lehner, Nils Asmussen, Marcus Völp, Benedikt Nöthen, Gerhard Fettweis)
- 161 **Flash-Aware Index Scan in PostgreSQL**
(Da-som Hwang, Woon-hak Kang, Gihwan Oh, Sang-won Lee)

SSEPM 2015 — 1st International Workshop on Scalable Social Event Processing and Management

- 167 **SSEPM 2015 Preface Pages**
- 170 **Analyzing Online News Dissemination via Structure Learning: An Experimental View**
(Ruiqi Li, Yanli Hu, Jiuyang Tang, Weidong Xiao)
- 174 **Distinguishing Re-Sharing Behaviors from Re-Creating Behaviors in Information Diffusion**
(Yiran Xie, Hongzhi Yin, Bin Cui, Junjie Yao, Quanqing Xu)
- 182 **Scalable Collective Spatial Keyword Query**
(Peijun He, Hao Xu, Xiang Zhao, Zhitao Shen)
- 190 **News Event Summarization Complemented by Micropoints**
(Pingping Lin, Rong Xiao, Yan Zhang)
- 198 **Sampling Social Streams for Hot Social Events Analytics**
(Ye Li, Fan Xia, Chengcheng Yu, Weining Qian, Aoying Zhou)
- 202 **An SVD-Based Multimodal Clustering Method for Social Event Detection**
(Yun Ma, Qing Li, Zhenguo Yang, Zheng Lu, Haiwei Pan, Antoni B. Chan)
- 210 **SINCA: Scalable In-Memory Event Aggregation Using Clustered Operators**
(Mahesh Kumar Behera, Kalyan S., Prasanna Venkatesh, Antoni Wolski)
- 216 **On Bottleneck-Aware Arrangement for Event-Based Social Networks**
(Yongxin Tong, Rui Meng, Jieying She)
- 224 **On the Rise and Fall of Sina Weibo: Analysis Based on a Fixed User Group**
(Fan Xia, Qunyan Zhang, Chengyu Wang, Weining Qian, Aoying Zhou)

ICDE 2015 PhD Symposium

- 232 **ICDE PhD Symposium 2015 Preface Pages**
- 235 **Friend Recommendation with a Target User in Social Networking Services**
(Sundong Kim)
- 240 **Integrating Query Processing with Parallel Languages**
(Brandon Myers)
- 245 **Improving the Quality of Large-Scale Database-Centric Software Systems by Analyzing Database Access Code**
(Tse-Hsun Chen)
- 250 **Pregel Meets UnCAL: A Systematic Framework for Transforming Big Graphs**
(Le-Duc Tung)
- 255 **Set-Based Approach for Lossless Graph Summarization Using Locality Sensitive Hashing**
(Kifayat Ullah Khan)
- 260 **Traversal Optimizations and Analysis for Large Graph Clustering**
(Waqas Nawaz)