

Contents

Second International DASFAA Workshop on Big Data Management and Analytics (BDMA)

Meme Media and Knowledge Federation for Exploratory Visual Analytics of Big Data	3
<i>Yuzuru Tanaka</i>	
Online Data Clustering Using Variational Learning of a Hierarchical Dirichlet Process Mixture of Dirichlet Distributions	18
<i>Wentao Fan and Nizar Bouguila</i>	
Distributed Skyline Computation of Vertically Splitted Databases by Using MapReduce	33
<i>Md. Anisuzzaman Siddique, Hao Tian, and Yasuhiko Morimoto</i>	
Short-Term Speed Prediction on Urban Highways by Ensemble Learning with Feature Subset Selection.	46
<i>Mohammad Arif Rasyidi and Kwang Ryel Ryu</i>	
Graph Summarization Using Word Correlation Analysis on Large Set of Documents.	61
<i>Putu Y. Kusmawan and Joonho Kwon</i>	
Distributed K-Distance Indexing Approach for Efficient Shortest Path Discovery on Large Graphs	75
<i>Jihye Hong, Hyunwook Kim, Waqas Nawaz, Kisung Park, Byeong-Soo Jeong, and Young-Koo Lee</i>	
Customized Information Interface with Web Applications	89
<i>Wookey Lee, Suan Lee, and Jinho Kim</i>	
Leveraging Enterprise Application Characteristics to Optimize Incremental Aggregate Maintenance in a Columnar In-Memory Database	102
<i>Stephan Müller, Paul Möller, and Hasso Plattner</i>	
MaiterStore: A Hot-Aware, High-Performance Key-Value Store for Graph Processing.	117
<i>Dong Chang, Yanfeng Zhang, and Ge Yu</i>	
Vertical Bit-Packing: Optimizing Operations on Bit-Packed Vectors Leveraging SIMD Instructions	132
<i>Martin Faust, Martin Grund, Tim Berning, David Schwalb, and Hasso Plattner</i>	

Efficient Streaming Detection of Hidden Clusters in Big Data Using Subspace Stream Clustering	146
<i>Marwan Hassani and Thomas Seidl</i>	
A Comparison of Systems to Large-Scale Data Access	161
<i>Amin Mesmoudi and Mohand-Saïd Hacıd</i>	
Third International Workshop on Data Management for Emerging Network Infrastructure (DaMEN)	
A Framework to Measure Storage Utilization in Cloud Storage Systems	179
<i>Xiao Zhang, Wan Guo, Zhanhuai Li, Xiaonan Zhao, and Xiao Qin</i>	
Personalized Recommendation via Relevance Propagation on Social Tagging Graph	192
<i>Huiming Li, Hao Li, Zimu Zhang, and Hao Wu</i>	
Optimizing Pipelined Execution for Distributed In-Memory OLAP System . . .	204
<i>Li Wang, Lei Zhang, Chengcheng Yu, and Aoying Zhou</i>	
Hash ^{ed} -Join: Approximate String Similarity Join with Hashing.	217
<i>Peisen Yuan, Chaofeng Sha, and Yi Sun</i>	
Minimizing Explanations of Why-Not Questions	230
<i>Chuanyu Zong, Bin Wang, Jing Sun, and Xiaochun Yang</i>	
HadoopM: A Message-Enabled Data Processing System on Large Clusters . . .	243
<i>Wei Pan, Zhanhuai Li, Bo Suo, and Zhuo Wang</i>	
AntiqueData: A Proxy to Maintain Computational Transparency in Cloud. . .	256
<i>Himel Dev, Mohammed Eunus Ali, Tanmoy Sen, and Madhusudan Basak</i>	
Third International Workshop on Spatial Information Modeling, Management and Mining (SIM³)	
Monitoring Query Processing in Mobile Robot Databases	271
<i>Kento Sugiura, Arata Hayashi, Tingting Dong, and Yoshiharu Ishikawa</i>	
Efficiently Evaluating Range-Constrained Spatial Keyword Query on Road Networks	283
<i>Wengen Li, Jihong Guan, and Shuigeng Zhou</i>	
A Spatial-Temporal Analysis of Users' Geographical Patterns in Social Media: A Case Study on Microblogs	296
<i>Chao Li, Zhongying Zhao, Jun Luo, Ling Yin, and Qiming Zhou</i>	

Solving Multiple Bichromatic Mutual Nearest Neighbor Queries with the GPU	308
<i>Marta Fort and J. Antoni Sellarès</i>	
A Kernel Density Method for Aggregating Boundary Collision Data into Areal Units	317
<i>Ge Cui, Xin Wang, and Dae-Won Kwon</i>	
Integrated Indoor Positioning with Mobile Devices for Location-Based Service Applications	329
<i>Bei Huang and Yang Gao</i>	
A Hybrid Scale-Out Cloud-Based Data Service for Worldwide Sensors	342
<i>Tania Khalafbeigi, Chih-Yuan Huang, Steve Liang, and Mea Wang</i>	
DASFAA Workshop on Uncertain and Crowdsourced Data (UnCrowd)	
Uncertainty in Crowd Data Sourcing Under Structural Constraints	351
<i>Antoine Amarilli, Yael Amsterdamer, and Tova Milo</i>	
Integration of Web Sources Under Uncertainty and Dependencies Using Probabilistic XML	360
<i>M. Lamine Ba, Sebastien Montenez, Ruiming Tang, and Talel Abdessalem</i>	
Skill Ontology-Based Model for Quality Assurance in Crowdsourcing	376
<i>Kinda El Maarry, Wolf-Tilo Balke, Hyunsouk Cho, Seung-won Hwang, and Yukino Baba</i>	
ProbKS: Keyword Search on Probabilistic Spatial Data	388
<i>Feng Gao, Rohit Jain, Sunil Prabhakar, and Luo Si</i>	
Towards Mobile Sensor-Aware Crowdsourcing: Architecture, Opportunities and Challenges	403
<i>Jiyin He, Kai Kunze, Christoph Lofi, Sanjay K. Madria, and Stephan Sigg</i>	
Conditioning Probabilistic Relational Data with Referential Constraints	413
<i>Ruiming Tang, Dongxu Shao, M. Lamine Ba, and Huayu Wu</i>	
Author Index	429