

# **INNS Conference on Big Data 2015**

Procedia Computer Science Volume 53

San Francisco, California, USA  
8 - 10 August 2015

## **Editors:**

**Asim Roy  
Plamen Angelov  
Adel Alimi**

**Kumar Venayagamoorthy  
Theodore Trafalis**

ISBN: 978-1-5108-2328-0

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© by Elsevier B.V.  
All rights reserved.

Printed by Curran Associates, Inc. (2016)

For permission requests, please contact Elsevier B.V.  
at the address below.

Elsevier B.V.  
Radarweg 29  
Amsterdam 1043 NX  
The Netherlands

Phone: +31 20 485 3911  
Fax: +31 20 485 2457

<http://www.elsevierpublishingsolutions.com/contact.asp>

**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  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)



## Contents

### Oral Sessions

#### *Challenges, Structuring and Handling*

A Nested Hierarchy of Dynamically Evolving Clouds for Big Data Structuring and Searching P. Angelov, P. Sadeghi-Tehran	1
Evolving Classifier TEDAClass for Big Data D. Kangin, P. Angelov, J.A. Iglesias, A. Sanchis	9
Managing Vagueness with Fuzzy in Hierarchical Big Data D.J. Lewis, T.P. Martin	19
Discriminating Variable Star Candidates in Large Image Databases from the HiTS Survey Using NMF P. Huijse, P.A. Estévez, F. Föörster, E. Berrocal	29
Segmentation of Geophysical Data: A Big Data Friendly Approach D.R.B. Stockwell, L. Zhang, B. Verma	39
Connectomics to Semantomics: Addressing the Brain's Big Data Challenge X.D. Arsiwalla, D. Dalmazzo, R. Zucca, A. Betella, S. Brandi, E. Martinez, P. Omedas, P. Verschure	48

#### *Neural and Statistical Optimization Tools*

The Art of Encephalography to Understand and Discriminate Higher Cognitive Functions Visualizing Big Data on Brain Imaging Using Brain Dynamics Movies J.J. Joshua Davis, R. Kozma, W.J. Freeman	56
Fast $\ell_1$ -norm Nearest Neighbor Search Using A Simple Variant of Randomized Partition Tree K. Sinha	64
Quantum-Inspired Features and Parameter Optimization of Spiking Neural Networks for a Case Study from Atmospheric M.C. Cardoso, M. Silva, M.M.B.R. Vellasco, E. Cataldo	74

#### *Computational Intelligence and Applications*

A Computational Framework for Implementation of Neural Networks on Multi-Core Machine W. Wang, Y. Murphey, P. Watta	82
Comparison of Adaptive Neuro-Fuzzy Inference Systems and Echo State Networks for PV Power Prediction I. Jayawardene, G.K. Venayagamoorthy	92
Don't Pay for Validation: Detecting Drifts from Unlabeled Data Using Margin Density T.S. Sethi, M. Kantardzic	103

#### *Analytics*

Using Noise to Speed up Markov Chain Monte Carlo Estimation B. Franzke, B. Kosko	113
Big Data Analytics in the Cloud: Spark on Hadoop vs MPI/OpenMP on Beowulf J.L. Reyes-Ortiz, L. Oneto, D. Anguita	121
An Investigation of Parallel Road Map Inference from Big GPS Traces Data W. Elleuch, A. Wali, A.M. Alimi	131
Big Data Analytics as a Service for Affective Humanoid Service Robots M. Jiang, L. Zhang	141

#### *Big Data Clustering*

HiClus: Highly Scalable Density-based Clustering with Heterogeneous Cloud C.-C. Chen, M.-S. Chen	149
Micro-Batching Growing Neural Gas for Clustering Data Streams Using Spark Streaming M. Ghesmoune, M. Lebbah, H. Azzag	158
On the Distributed Implementation of Unsupervised Extreme Learning Machines for Big Data Y. Rizk, M. Awad	167

Large-scale Monitoring for Cyber Attacks by Using Cluster Information on Darknet Traffic Features H. Nishikaze, S. Ozawa, J. Kitazono, T. Ban, J. Nakazato, J. Shimamura	175
A Time Series Clustering Technique Based on Community Detection in Networks L.N. Ferreira, L. Zhao	183
Evolving Gustafson-Kessel Possibilistic c-Means Clustering I. Škrjanc, D. Dovžan	191
<i>Big Data in Social Media and Real Applications</i>	
Ontology-based Sentiment Analysis Process for Social Media Content P. Thakor, S. Sasi	199
Active Manifold Learning with Twitter Big Data C. Silva, M. Antunes, J. Costa, B. Ribeiro	208
MapReduce Based Text Detection in Big Data Natural Scene Videos Abdelkarim Ben Ayed, Mohamed Ben Halima, Adel M. Alimi	216
Automated Detection of Human Users in Twitter M.A. Fernandes, P. Patel, T. Marwala	224
Toward a Framework for Improving the Execution of the Big Data Applications W. Abbes, Z. Kechaou, A.M. Alimi	232
Low-Rank Tensor Recovery for Geo-Demand Estimation in Online Retailing L. He, Z. (Tony) Qin, J. Bewli	239
<i>Algorithms and Modeling</i>	
Instance Selection with Ant Colony Optimization I.M. Anwar, K.M. Salama, A.M. Abdelbar	248
Formulation of the Tree Approximation Problem as a Detection Problem and Relation between the AUC and Information Theory Divergences N.T. Khajavi, A. Kuh	257
Leveraging <i>Big Data</i> to Model the Likelihood of Developing Psychological Conditions After a Concussion F. Dabek, J.J. Caban	265
<i>Big Data Methodologies</i>	
Extracting Clinical Relations in Electronic Health Records Using Enriched Parse Trees J. Kim, Y. Choe, K. Mueller	274
Building Efficient Probability Transition Matrix Using Machine Learning from Big Data for Personalized Route Prediction X. Wang, Y. Ma, J. Di, Y.L. Murphey, S. Qiu, J. Kristinsson, J. Meyer, F. Tseng, T. Feldkamp	284
Towards an offloading Framework Based on Big Data Analytics in Mobile Cloud Computing Environments H. Kchaou, Z. Kechaou, A.M. Alimi	292
<b>Special sessions</b>	
<i>Special Session on Scalable Machine Learning</i>	
MapReduce SVM Game C.M. Vineyard, S.J. Verzi, C.D. James, J.B. Aimone, G.L. Heileman	298
Scalability of Stochastic Gradient Descent Based on “Smart” Sampling Techniques S. Clémentinçon, A. Bellet, O. Jelassi, G. Papa	308
WWN-8: Incremental Online Stereo with Shape-from-X Using Life-Long Big Data from Multiple Modalities M. Solgi, J. Weng	316
A MapReduce Cortical Algorithms Implementation for Unsupervised Learning of Big Data N. Hajj, Y. Rizk, M. Awad	327
Learning Morphological Transformations with Recurrent Neural Networks S. Biswas, T. Breuel	335
A Classification Algorithm for High-dimensional Data A. Roy	345
<i>Special Session on Brain-inspired Computing: Theories and Applications</i>	
A Localist Paradigm for Big Data T. Achler	356
Robust Discovery of Temporal Structure in Multi-neuron Recordings Using Hopfield Networks C. Hillar, F. Effenberger	365
EEG-based Subject Independent Affective Computing Models L. Bozhkov, P. Georgieva, I. Santos, A. Pereira, C. Silva	375
A New Multi-Channels Sequence Recognition Framework Using Deep Convolutional Neural Network R. Zhang, C. Li, D. Jia	383

*Special Session on Soft computing methods in large-scale content-based image retrieval*

Generic Object Recognition with Local Receptive Fields Based Extreme Learning Machine Z. Bai, L.L.C. Kasun, G.-B. Huang .....	391
Efficient Skin Segmentation via Neural Networks: HP-ELM and BD-SOM C. Swaney, A. Akusok, K.-M. Björk, Y. Miche, A. Lendasse .....	400
Dynamic Phase and Group Detection in Pedestrian Crowd Data Using Multiplex Visibility Graphs C. Stephen .....	410
Regularized Extreme Learning Machine for Large-scale Media Content Analysis A. Iosifidis, A. Tefas, I. Pitas .....	420

*Special Session on Advances and Perspectives of Big Data Analytics in Transportation*

Approaching Camera-based Real-World Navigation Using Object Recognition Z. Zheng, X. He, J. Weng .....	428
Condition Based Maintenance in Railway Transportation Systems Based on Big Data Streaming Analysis E. Fumeo, L. Oneto, D. Anguita .....	437
Evaluation of Self-Organized Approach for Predicting Compressor Faults in a City Bus Fleet Y. Fan, S. Nowaczyk, T. Rögnvaldsson .....	447
Railway Assets: A Potential Domain for Big Data Analytics A. Thaduri, D. Galar, U. Kumar .....	457
Learning from Distributed Data Sources Using Random Vector Functional-Link Networks S. Scardapane, M. Panella, D. Comminiello, A. Uncini .....	468
Navigating a Robot through Big Visual Sensory Data A. Altahhan .....	478