PROCEEDINGS OF
THE 2016 INTERNATIONAL CONFERENCE ON
ADVANCES IN BIG DATA ANALYTICS

ABDA 2016

Editors

Hamid R. Arabnia
Fernando G. Tinetti, Mary Yang

Associate Editors

George Jandieri, James J. (Jong Hyuk) Park
Ashu M. G. Solo

WORLDCOMP'16
July 25-28, 2016
Las Vegas Nevada, USA
www.worldcomp.org

©CSREA Press

This volume contains papers presented at The 2016 International Conference on Advances in Big Data Analytics (ABDA'16). Their inclusion in this publication does not necessarily constitute endorsements by editors or by the publisher.

Copyright and Reprint Permission

Copying without a fee is permitted provided that the copies are not made or distributed for direct commercial advantage, and credit to source is given. Abstracting is permitted with credit to the source. Please contact the publisher for other copying, reprint, or republication permission.

Copyright © 2016 CSREA Press
ISBN: 1-60132-427-8
Printed in the United States of America

CSREA Press
U. S. A.

Foreword

It gives us great pleasure to introduce this collection of papers to be presented at the 2016 International Conference on Advances in Big Data Analytics (ABDA'16), July 25-28, 2016, at Monte Carlo Resort, Las Vegas, USA.

An important mission of the World Congress in Computer Science, Computer Engineering, and Applied Computing (a federated congress to which this conference is affiliated with) includes *"Providing a unique platform for a diverse community of constituents composed of scholars, researchers, developers, educators, and practitioners. The Congress makes concerted effort to reach out to participants affiliated with diverse entities (such as: universities, institutions, corporations, government agencies, and research centers/labs) from all over the world. The congress also attempts to connect participants from institutions that have* ***teaching*** *as their main mission with those who are affiliated with institutions that have* ***research*** *as their main mission. The congress uses a quota system to achieve its institution and geography diversity objectives."* By any definition of diversity, this congress is among the most diverse scientific meeting in USA. We are proud to report that this federated congress has authors and participants from 74 different nations representing variety of personal and scientific experiences that arise from differences in culture and values. As can be seen (see below), the program committee of this conference as well as the program committee of all other tracks of the federated congress are as diverse as its authors and participants.

The program committee would like to thank all those who submitted papers for consideration. About 60% of the submissions were from outside the United States. Each submitted paper was peer-reviewed by two experts in the field for originality, significance, clarity, impact, and soundness. In cases of contradictory recommendations, a member of the conference program committee was charged to make the final decision; often, this involved seeking help from additional referees. In addition, papers whose authors included a member of the conference program committee were evaluated using the double-blinded review process. One exception to the above evaluation process was for papers that were submitted directly to chairs/organizers of pre-approved sessions/workshops; in these cases, the chairs/organizers were responsible for the evaluation of such submissions. The overall paper acceptance rate for regular papers was 26%; 12% of the remaining papers were accepted as poster papers (at the time of this writing, we had not yet received the acceptance rate for a couple of individual tracks.)

We are very grateful to the many colleagues who offered their services in organizing the conference. In particular, we would like to thank the members of Program Committee of ABDA'16, members of the congress Steering Committee, and members of the committees of federated congress tracks that have topics within the scope of ABDA. Many individuals listed below, will be requested after the conference to provide their expertise and services for selecting papers for publication (extended versions) in journal special issues as well as for publication in a set of research books (to be prepared for publishers including: Springer, Elsevier, BMC journals, and others).

- *Prof. Afrand Agah; Department of Computer Science, West Chester University of Pennsylvania, West Chester, PA, USA*
- *Prof. Abbas M. Al-Bakry (Congress Steering Committee); University President, University of IT and Communications, Baghdad, Iraq*
- *Prof. Nizar Al-Holou (Congress Steering Committee); Professor and Chair, Electrical and Computer Engineering Department; Vice Chair, IEEE/SEM-Computer Chapter; University of Detroit Mercy, Detroit, Michigan, USA*
- *Prof. Hamid R. Arabnia (Congress Steering Committee & Coordinator); Graduate Program Director (PhD, MS, MAMS); The University of Georgia, USA; Editor-in-Chief, Journal of Supercomputing (Springer); Editor-in-Chief, Transactions of Computational Science & Computational Intelligence (Springer); Fellow, Center of Excellence in Terrorism, Resilience, Intelligence & Organized Crime Research (CENTRIC).*
- *Prof. Mehran Asadi; Interim Chair, Department of Business and Entrepreneurial Studies, The Lincoln University, Pennsylvania, USA*
- *Prof. Juan Jose Martinez Castillo; Director, The Acantelys Alan Turing Nikola Tesla Research Group and GIPEB, Universidad Nacional Abierta, Venezuela*

- *Dr. Daniel Bo-Wei Chen; Chair, IEEE Signal Processing Chapter, IEEE Harbin Section; Guest Editor in ACM Transactions in Embedded Computing; School of Information Technology, Monash University Sunway Campus, Australia*
- *Prof. Kevin Daimi (Congress Steering Committee); Director, Computer Science and Software Engineering Programs, Department of Mathematics, Computer Science and Software Engineering, University of Detroit Mercy, Detroit, Michigan, USA*
- *Dr. Lamia Atma Djoudi (Chair, Doctoral Colloquium & Demos Sessions); Synchrone Technologies, France*
- *Prof. Mary Mehrnoosh Eshaghian-Wilner (Congress Steering Committee); Professor of Engineering Practice, University of Southern California, California, USA; Adjunct Professor, Electrical Engineering, University of California Los Angeles, Los Angeles (UCLA), California, USA*
- *Prof. George A. Gravvanis (Congress Steering Committee); Director, Physics Laboratory & Head of Advanced Scientific Computing, Applied Math & Applications Research Group; Professor of Applied Mathematics and Numerical Computing and Department of ECE, School of Engineering, Democritus University of Thrace, Xanthi, Greece; former President of the Technical Commission on Data Processing, Social Security for the Migrant Workers, European Commission, Hellenic Presidency, Greece*
- *Dr. Ruizhu Huang; Texas Advanced Computing Center, University of Texas, Austin, Texas, USA*
- *Prof. George Jandieri (Congress Steering Committee); Georgian Technical University, Tbilisi, Georgia; Chief Scientist, The Institute of Cybernetics, Georgian Academy of Science, Georgia; Ed. Member, International Journal of Microwaves and Optical Technology, The Open Atmospheric Science Journal, American Journal of Remote Sensing, Georgia*
- *Prof. Byung-Gyu Kim (Congress Steering Committee); Multimedia Processing Communications Lab.(MPCL), Department of Computer Science and Engineering, College of Engineering, SunMoon University, South Korea*
- *Prof. Tai-hoon Kim; School of Information and Computing Science, University of Tasmania, Australia*
- *Prof. D. V. Kodavade; Professor & Head, Computer Science & Engineering Department, D.K.T.E Society's Textile & Engineering Institute, Ichalkaranji, Maharashtra State, India*
- *Prof. Dr. Guoming Lai; Computer Science and Technology, Sun Yat-Sen University, Guangzhou, P. R. China*
- *Dr. Muhammad Naufal Bin Mansor; Faculty of Engineering Technology, Kampus Uniciti Alam, Universiti Malaysia Perlis, UniMAP, Malaysia*
- *Dr. Andrew Marsh (Congress Steering Committee); CEO, HoIP Telecom Ltd (Healthcare over Internet Protocol), UK; Secretary General of World Academy of BioMedical Sciences and Technologies (WABT) a UNESCO NGO, The United Nations*
- *Prof. Ali Mostafaeipour; Industrial Engineering Department, Yazd University, Yazd, Iran*
- *Prof. James J. (Jong Hyuk) Park (Congress Steering Committee); Department of Computer Science and Engineering (DCSE), SeoulTech, Korea; President, FTRA, EiC, HCIS Springer, JoC, IJITCC; Head of DCSE, SeoulTech, Korea*
- *Prof. Shashikant Patil; Electronics & Telecommunication Engineering Department, Head of SVKMs NMiMS Bosch Rexroth Center of Excellence in Automation Technologies, Shirpur Campus, India*
- *Dr. Akash Singh (Congress Steering Committee); IBM Corporation, Sacramento, California, USA; Chartered Scientist, Science Council, UK; Fellow, British Computer Society; Member, Senior IEEE, AACR, AAAS, and AAAI; IBM Corporation, USA*
- *Ashu M. G. Solo, (Publicity Chair), Fellow of British Computer Society, Principal/R&D Engineer, Maverick Technologies America Inc.*
- *Dr. Jaya Thomas; Department of Computer Science, State University of New York, Korea (SUNY Korea) and Department of Computer Science, Stony Brook University, USA*
- *Prof. Fernando G. Tinetti (Congress Steering Committee); School of Computer Science, Universidad Nacional de La Plata, La Plata, Argentina; Co-editor, Journal of Computer Science and Technology (JCS&T).*
- *Prof. Hahanov Vladimir (Congress Steering Committee); Vice Rector, and Dean of the Computer Engineering Faculty, Kharkov National University of Radio Electronics, Ukraine and Professor of Design Automation Department, Computer Engineering Faculty, Kharkov; IEEE Computer Society Golden Core Member; National University of Radio Electronics, Ukraine*
- *Dr. Mohd Helmy Abd Wahab; Former Head of Intelligent System Lab., Senior Lecturer and Academic Advisor, Department of Computer Engineering, University Tun Hussein Onn Malaysia, Malaysia*
- *Prof. Shiuh-Jeng Wang (Congress Steering Committee); Director of Information Cryptology and Construction Laboratory (ICCL) and Director of Chinese Cryptology and Information Security Association (CCISA); Department of Information Management, Central Police University, Taoyuan, Taiwan; Guest Ed., IEEE Journal on Selected Areas in Communications.*
- *Prof. Dr. Bernd E. Wolfinger; Telecommunications and Computer Networks Division, Computer Science Department, University of Hamburg, Hamburg (Stellingen), Germany*

- *Prof. Mary Yang (Co-Editor); Director of MidSouth Bioinformatics Center and Director of Joint Bioinformatics Ph.D. Program, University of Arkansas Little Rock George Washington Donaghey College of Engineering & Information Technology and University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA*
- *Prof. Jane You (Congress Steering Committee & Vice-Chair of IPCV'16); Associate Head, Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong*

We would like to extend our appreciation to the referees, the members of the program committees of individual sessions, tracks, and workshops; their names do not appear in this document; they are listed on the web sites of individual tracks.

As Sponsors-at-large, partners, and/or organizers each of the followings (separated by semicolons) provided help for at least one track of the Congress: Computer Science Research, Education, and Applications Press (CSREA); US Chapter of World Academy of Science (http://www.worldcomp.org/) ; American Council on Science & Education & Federated Research Council (http://www.americancse.org/); HoIP, Health Without Boundaries, Healthcare over Internet Protocol, UK (http://www.hoip.eu); HoIP Telecom, UK (http://www.hoip-telecom.co.uk); and WABT, Human Health Medicine, UNESCO NGOs, Paris, France (http://www.thewabt.com/). In addition, a number of university faculty members and their staff (names appear on the cover of the set of proceedings), several publishers of computer science and computer engineering books and journals, chapters and/or task forces of computer science associations/organizations from 3 regions, and developers of high-performance machines and systems provided significant help in organizing the conference as well as providing some resources. We are grateful to them all.

We express our gratitude to keynote, invited, and individual conference/tracks and tutorial speakers - the list of speakers appears on the conference web site. We would also like to thank the followings: UCMSS (Universal Conference Management Systems & Support, California, USA) for managing all aspects of the conference; Dr. Tim Field of APC for coordinating and managing the printing of the proceedings; and the staff of Monte Carlo Resort (Convention department) in Las Vegas for the professional service they provided. Last but not least, we would like to thank the Co-Editors and Associate Co-Editors of ABDA'16: Prof. Hamid R. Arabnia, Prof. George Jandieri, Prof. Mary Yang, Prof. James J. (Jong Hyuk) Park, Ashu M. G. Solo, and Prof. Fernando G. Tinetti.

We present the proceedings of ABDA'16.

Steering Committee, 2016
http://www.worldcomp.org/

Contents

SESSION: PROTEIN CLASSIFICATION, STRUCTURE PREDICTION, AND COMPUTATIONAL STRUCTURAL BIOLOGY

Bidirectional Representation and Backpropagation Learning **3**
Olaoluwa Adigun, Bart Kosko

Processing and Analysis of Large Data Sets Using High Performance Computing: Beyond Experimental Data **10**
Brian Panneton, Brian Henz, Pritesh Patel, James Adametz

Event Sequence Detection over Interval-Based Event Streams **17**
Salah Ahmed, Olga Poppe, Elke A. Rundensteiner

Efficient Lossless Compression of 4D Hyperspectral Image Data **24**
Hongda Shen, W. David Pan, Yuhang Dong

String Vector Based KNN for Index Optimization **31**
Taeho Jo

Data Quality Evaluation Previous to Big Data Analytics **37**
M. Mejia-Lavalle, J. Perez-Ortega, A. Magadan-Salazar, E. Perez-Luna, G. Reyes Salgado, D. Mujica Vargas

SESSION: GENE EXPRESSION, REGULATORY NETWORKS, MICROARRAY, SEQUENCING, ALIGNMENT, AND RELATED STUDIES

A Novel Control-flow based Intrusion Detection Technique for Big Data Systems **45**
Santosh Aditham, Nagarajan Ranganathan

Big Data to Optimise Product Strategy in the Electronics Industry **52**
Nawaz Khan, Vijayalakshmi Subbiah, Elli Georgiadou, Angela Repanovich

Effective Detecting MicroblogSpammers Using Big Data Fusion Algorithm **59**
Yang Qiao, Huaping Zhang, Yanping Zhao, Yu Zhang, Yu Min

A Preliminary Report on Infusing Data-Enabled Active Learning in Undergraduate CS Mathematics and Statistics Courses **66**
Carl Pettis, Rajendran Swamidurai, Ash Abebe

The Impact of Macroeconomic Factors on Saudi Stock Market (Tadawul) Prices **71**
Mu'tasem Jarrah, Naomie Salim

Bigdata Platform based Approach for Defending against DDoS **77**
Yoon Joo Chae, Nikitha Johnsirani Venkatesan, Dong Ryeol Shin

Application of Deep Convolutional Neural Networks for Detecting Extreme Weather in Climate Datasets **81**
Yunjie Liu, Evan Racah, . Prabhat, Joaquin Correa, Amir Khosrowshahi, David Lavers, Kenneth Kunkel, Michael Wehner, William Collins

SESSION: SYSTEMS BIOLOGY AND RNA SEQUENCE DATA PROCESSING + HIGH PERFORMANCE COMPUTING

Star Plot Visualization of Ultrahigh Dimensional Multivariate Data **91**
Shabana Sangli, Gurminder Kaur, Bijaya B. Karki

Simulating Spatial Correlation for Catastrophic Events **98**
Georg Hofmann

Detecting Sarcastic Tweets: A SentiStrength Modeling Approach **102**
samaneh Nadali, Masrah Azrifah Azmi Murad , Nurfadhlina Mohamad Sharef

Why We Need Big Data? **107**
Eizo Kinoshita, Takafumi Mizuno

SESSION: POSTER PAPERS

Data-Driven Algorithms for Fault Detection and Diagnosis in Industrial Process **115**
Mohamed El Koujok, Mouloud Amazouz

Data and Parity Block Placement Policy to Enhance Storage Eficiency and Utilization **117**
Dayeon Kim, Dongryul Shin

Dark Side of Social Networking and the Data it Accumulate **119**
Muhammad Fahim Uddin

SESSION: LATE BREAKING PAPERS

IDEAS: An Online Tool to Identify Differential Expression of Genes for Applications in Genome-wide Studies **123**
William Yang, Kenji Yoshigoe, Xiaosheng Wang, Dan Li, Yifan Zhang, Patrycja Krakowiak, Wenbing Zhao, Hong Zhou, Zuojie Luo, Xiang Qin, Guo-Zheng Li, Hamid R. Arabnia, Weida Tong, Jun S. Liu, Mary Qu Yang

Dimensionality Reduction via the Johnson-Lindenstrauss Lemma **134**
John Fedoruk, Byron Schmuland, Julia Johnson, Giseon Heo

Correlating Algorithm and Model for Personality Features with Academic Relevance in Big Data **140**
Muhammad Fahim Uddin, Jeongkyu Lee

Gremlin By Example **144**
Kornelije Rabuzin, Mirko Malekovic, Martina Sestak

Proposing Good Fit Student Algorithm (GFS-A) to Utilize Big Data and Academic Data **150**
Muhammad Uddin, Jeongkyu Lee

Tutorial Lecture: Resonance of Big-Data Analytics and Precision Medicine Research is Producing a Profound Impact on Optimized Individual Healthcare **154**
William Yang

An Approach to Developing and Maintaining the Information Security Systems in Design of Big data Systems **156**
Venkateswarlu Sunkari, V Surendra Amareswarapu, Kranthi Kumar Singamaneni, Ranga Chaitanya Singamaneni