

Communications in Computer and Information Science

1054

Commenced Publication in 2007

Founding and Former Series Editors:

Phoebe Chen, Alfredo Cuzzocrea, Xiaoyong Du, Orhun Kara, Ting Liu,
Krishna M. Sivalingam, Dominik Ślęzak, Takashi Washio, and Xiaokang Yang

Editorial Board Members

Simone Diniz Junqueira Barbosa

*Pontifical Catholic University of Rio de Janeiro (PUC-Rio),
Rio de Janeiro, Brazil*

Joaquim Filipe

Polytechnic Institute of Setúbal, Setúbal, Portugal

Ashish Ghosh

Indian Statistical Institute, Kolkata, India

Igor Kotenko

*St. Petersburg Institute for Informatics and Automation of the Russian
Academy of Sciences, St. Petersburg, Russia*

Junsong Yuan

University at Buffalo, The State University of New York, Buffalo, NY, USA

Lizhu Zhou

Tsinghua University, Beijing, China

More information about this series at <http://www.springer.com/series/7899>

Muhammad Younas · Irfan Awan ·
Salima Benbernou (Eds.)

Big Data Innovations and Applications

5th International Conference, Innovate-Data 2019
Istanbul, Turkey, August 26–28, 2019
Proceedings

Editors

Muhammad Younas
School of ECM
Oxford Brookes University
Oxford, UK

Irfan Awan
Department of Informatics
University of Bradford
Bradford, UK

Salima Benbernou
Universite Paris Descartes
Paris, France

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-030-27354-5 ISBN 978-3-030-27355-2 (eBook)
<https://doi.org/10.1007/978-3-030-27355-2>

© Springer Nature Switzerland AG 2019

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Welcome to the proceedings of the 5th International Conference on Big Data Innovations and Applications (Innovate-Data 2019), which was held during August 26–28, 2019, in Istanbul, Turkey. The Innovate-Data 2019 conference was co-located with the International Conference on Future Internet of Things and Cloud (FiCloud-2019), the International Conference on Mobile Web and Intelligent Information Systems (MobiWis-2019), the International Conference on Deep Learning and Machine Learning in Emerging Applications (DEEP-ML 2019), and different workshops and symposia.

Big data has become a new trend in modern information technology and has been transforming various domains and disciplines including business, finance, science, education, the public sector, health care, life style and society as a whole. For instance, businesses can exploit big data in order to improve customer relationships and to serve customers in an efficient and effective way. Big data tools can also help in cost and time savings as they can analyze data efficiently and make quick decisions.

The concept of big data revolves not only around the size but how the data can be efficiently collected, processed, and analyzed in order to achieve the desired goals. The aim of the Innovate-Data conference is to solicit research and development work in order to further improve the underlying models, methods, and tools of big data and to find new ways of using big data in current and future applications. The Innovate-Data 2019 conference included interesting and timely topics such as: big data storage, representation, and processing; data engineering and design; big data security, privacy, and trust; big data models, infrastructure, and platforms; visualization of big data; big data analytics and metrics; and the overall applications and innovations of big data and related technologies.

In response to the call for papers, Innovate-Data 2019 received submissions from different countries across the world. Each paper was reviewed by multiple reviewers. Based on the reviews, 16 papers were accepted for the technical program, that is, an acceptance rate of 33%. The accepted papers address interesting research and practical issues related to big data processing and applications, big data analytics, big data security, privacy and trust, machine/deep learning and big data, and the use of big data in innovative applications.

We express our thanks to the members of the Organizing and Program Committees for helping in the organization and the review process of the research papers. Their timely and useful feedback to authors of the submitted papers is highly appreciated. We would also like to thank all authors for their contributions to Innovate-Data 2019. Putting together such an interesting technical program would not have been possible without the support from the technical committees and the authors.

We sincerely thank Prof. William Knottenbelt (General Chair) for his help and support. We would like to thank Dr. Filipe Portela (Workshop Coordinator), Dr. Fang-Fang Chua (Publicity Chair), Dr. Lin Guan (Journal Special Issue Coordinator), and Prof. Irfan Awan (Publication Chair). We sincerely appreciate the contributions of the local organizing chairs, Dr. Perin Ünal, Teknopar, Dr. Tacha Serif, and Dr. Sezer Gören Uğurdağ, for their help and support in the organization of the conference.

Innovate-Data 2019 and the co-located conferences had joint keynote and invited sessions in which interesting talks were delivered. We sincerely thank the speakers Prof. Pierangela Samarati (Università degli Studi di Milano, Italy), Mr. Gökhan Büyükdığan (Arçelik A.S., Turkey), and Dr. Soumya Kanti Datta (EURECOM, France).

Our sincere thanks also go to the Springer CCIS team for their valuable support in the production of the conference proceedings.

August 2019

Salima Benbernou
Muhammad Younas

Organization

Innovate-Data 2019 Organizing Committee

General Chair

William Knottenbelt Imperial College London, UK

Program Co-chairs

Salima Benbernou Université Paris Descartes, France
Muhammad Younas Oxford Brookes University, UK

Local Organizing Co-chairs

Perin Ünal Teknopar, Turkey
Sezer Gören Uğurdağ Yeditepe University, Turkey
Tacha Serif Yeditepe University, Turkey

Publication Chair

Irfan Awan University of Bradford, UK

Workshop Coordinator

Filipe Portela University of Minho, Portugal

Publicity Chair

Fang-Fang Chua Multimedia University, Malaysia

Journal Special Issue Coordinator

Lin Guan Loughborough University, UK

Program Committee

Abdelmounaam Rezgui Illinois State University, USA
Abdulsalam Yassine Lakehead University, Canada
Afonso Ferreira CNRS - Institut de Recherches en Informatique
 de Toulouse, France
Ahmed Awad Cairo University, Egypt

Ahmed Zouinkhi	National Engineering School of Gabes, Tunisia
Akiyo Nadamoto	Konan University, Japan
Allaoua Chaoui	University Mentouri Constantine, Algeria
Amin Beheshti	Macquarie University, Australia
Armin Lawi	Hasanuddin University, Indonesia
Amr Magdy	University of California, Riverside, USA
Aris Gkoulalas-Divanis	IBM Watson Health, Cambridge, MA, USA
Athena Vakali	Aristotle University of Thessaloniki, Greece
Ayman Alahmar	Lakehead University, Canada
Cándido Caballero-Gil	Universidad de La Laguna, Spain
Dimitri Plexousakis	University of Crete, Greece
Dimka Karastoyanova	University of Groningen, The Netherlands
Domenico Talia	University of Calabria, Italy
Emad Mohammed	University of Calgary, Canada
Ernesto Damiani	University of Milan, Italy
Fanny Klett	German Workforce ADL Partnership Laboratory, Germany
Faouzi Alaya Cheikh	Norwegian University of Science and Technology, Norway
Federica Paci	University of Trento, Italy
Giuseppe Di Modica	University of Catania, Italy
George Pallis	University of Cyprus, Cyprus
Guy De Tré	Ghent University, Belgium
Hesham Hallal	Fahad Fahad Bin Sultan University, Saudi Arabia
Hiroaki Higaki	Tokyo Denki University, Japan
Ibrahim Korpeoglu	Bilkent University, Turkey
Irena Holubova	Charles University in Prague, Czech Republic
Iván Santos-González	Universidad de La Laguna, Spain
Jezabel M. Molina-Gil	Universidad de La Laguna, Spain
Jorge Bernardino	Polytechnic Institute of Coimbra - ISEC, Portugal
Kazuaki Tanaka	Kyushu Institute of Technology, Japan
Khouloud Boukadi	University of Sfax, Tunisia
Mohamed Boukhebouze	CETIC Research Center, Belgium
Morad Benyoucef	University of Ottawa, Canada
Mourad Khayati	University of Fribourg, Switzerland
Mourad Ouziri	Université Paris Descartes - LIPADE, France
Muhammad Rizwan Asghar	The University of Auckland, New Zealand
Orazio Tomarchio	University of Catania, Italy
Pino Caballero-Gil	DEIOC, University of La Laguna, Spain
Rachid Benlamri	Lakehead University, Canada
Radwa El Shawi	University of Tartu, Estonia
Rafael Santos	Brazilian National Institute for Space Research, Brazil
Raghava Mutharaju	IIT-Delhi, India
Raj Sunderraman	Georgia State University, USA
Ridha Hamila	Qatar University, Qatar
Saad Harous	UAE University, UAE

Sadiki Tayeb	International University Rabat, Morocco
Sebastian Link	The University of Auckland, New Zealand
Shantanu Sharma	University of California, Irvine, USA
Stephane Bressan	National University of Singapore, Singapore
Toshihiro Yamauchi	Okayama University, Japan
Tengku Adil	University Technology of MARA, Malaysia
Vasilios Andrikopoulos	University of Groningen, The Netherlands
Yacine Atif	Skövde University, Sweden
Zaki Malik	Eastern Michigan University, USA
Zuhair Khayyat	KAUST, Saudi Arabia

Contents

Advances in Big Data Systems

Enabling Joins over Cassandra NoSQL Databases	3
<i>Haridimos Kondylakis, Antonis Fountouris, Apostolos Planas, Georgia Troullinou, and Dimitris Plexousakis</i>	
MapReduce Join Across Geo-Distributed Data Centers	18
<i>Giuseppe Di Modica and Orazio Tomarchio</i>	
Mobile Device Identification via User Behavior Analysis	32
<i>Kadriye Dogan and Ozlem Durmaz Incel</i>	
A SAT-Based Formal Approach for Verifying Business Process Configuration	47
<i>Abderrahim Ait Wakrime, Souha Boubaker, Slim Kallel, and Walid Gaaloul</i>	

Machine Learning and Data Analytics

TSWNN+: Check-in Prediction Based on Deep Learning and Factorization Machine	65
<i>Chang Su, Ningning Liu, Xianzhong Xie, and Shaowen Peng</i>	
Deep Learning Based Sentiment Analysis on Product Reviews on Twitter . . .	80
<i>Aytuğ Onan</i>	
A Cluster-Based Machine Learning Model for Large Healthcare Data Analysis	92
<i>Fatemeh Sharifi, Emad Mohammed, Trafford Crump, and Behrouz H. Far</i>	
Satire Detection in Turkish News Articles: A Machine Learning Approach	107
<i>Mansur Alp Toçoğlu and Aytuğ Onan</i>	

Big Data Innovation and Applications

Committee of the SGTM Neural-Like Structures with Extended Inputs for Predictive Analytics in Insurance	121
<i>Roman Tkachenko, Ivan Izonin, Michal Greguš ml., Pavlo Tkachenko, and Ivanna Dronyuk</i>	

Game Analytics on Free to Play	133
<i>Robert Flunger, Andreas Mladenow, and Christine Strauss</i>	
A Decentralized File Sharing Framework for Sensitive Data	142
<i>Onur Demir and Berkay Kocak</i>	
The Information System for the Research in Carotid Atherosclerosis	150
<i>Jiri Blahuta and Tomas Soukup</i>	
Security and Risk Analysis	
Big Data Analytics for Financial Crime Typologies	165
<i>Kirill Plaksiy, Andrey Nikiforov, and Natalia Miloslavskaya</i>	
Development of a Model for Identifying High-Risk Operations for AML/CFT Purposes	179
<i>Pavel Y. Leonov, Viktor P. Suyts, Oksana S. Kotelyanets, and Nikolai V. Ivanov</i>	
Monitoring System for the Housing and Utility Services Based on the Digital Technologies IIoT, Big Data, Data Mining, Edge and Cloud Computing	193
<i>Vasiliy S. Kireev, Pyotr V. Bochkaryov, Anna I. Guseva, Igor A. Kuznetsov, and Stanislav A. Filippov</i>	
K-Means Method as a Tool of Big Data Analysis in Risk-Oriented Audit . . .	206
<i>Pavel Y. Leonov, Viktor P. Suyts, Oksana S. Kotelyanets, and Nikolai V. Ivanov</i>	
Author Index	217