Lecture Notes in Computer Science **10440**

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

More information about this series at http://www.springer.com/series/7409

Ladjel Bellatreche · Sharma Chakravarthy (Eds.)

Big Data Analytics and Knowledge Discovery

19th International Conference, DaWaK 2017
Lyon, France, August 28–31, 2017
Proceedings

Springer

Editors
Ladjel Bellatreche
LIAS/ISAE-ENSMA
Chasseneuil
France

Sharma Chakravarthy
University of Texas at Arlington
Arlington, TX
USA

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISBN 978-3-319-64282-6 ISBN 978-3-319-64283-3 (eBook)
DOI 10.1007/978-3-319-64283-3

Library of Congress Control Number: 2017947764

LNCS Sublibrary: SL3 – Information Systems and Applications, incl. Internet/Web, and HCI

Printed on acid-free paper

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

Preface

Big data analytics and knowledge discovery technologies have been growing over time and they are now a practical need in every major small or large company. The adoption of these technologies by researchers and industries encourages these companies to integrate heterogeneous, distributed, autonomous, and evolving data with high variety coming from traditional (databases) and advanced sources—such as sensors and social networks, IoT— into a single large database to enable advanced querying, analysis, and recommendation. With the explosion of the diversity of deployment platforms motivated by high-performance computing (HPC) and advanced programming paradigms, offering the process of retrieval, knowledge discovery from this huge amount of heterogeneous complex data and sentiment analysis forms the litmus test for research in this area. Faced with this volume of data managed by the new generation of data warehouses, companies are very sensitive to non-functional requirement satisfaction, by proposing mathematical cost models estimating different metrics such as query performance, elasticity, etc.

During the past few years, the International Conference on Big Data Analytics and Knowledge Discovery (DaWaK) has become one of the most important international scientific events bringing together researchers, developers, and practitioners to discuss the latest research issues and experiences in developing and deploying a new generation of data warehouses and knowledge discovery systems, applications, and solutions. This year's conference (DaWaK 2017) built on this tradition of facilitating the cross-disciplinary exchange of ideas, experience, and potential research directions. DaWaK 2017 sought to introduce innovative principles, methods, models, algorithms and solutions, industrial products, and experiences to challenging problems faced in the development of a new generation of data warehouses in the big data era, knowledge discovery, data-mining applications, and the emerging area of HPC.

This year we received 97 papers and the Program Committee finally selected 24 full papers and 11 short papers, making an acceptance rate of 25%. The accepted papers cover a number of broad research areas on both theoretical and practical aspects of new generations of data warehouses and knowledge discovery. In the area of big data, the topics covered included the modeling and designing of a new generation of data warehouses by considering the big data aspects, data flows, NoSQL databases, cloud computing, cost models, data streaming, advanced programming paradigms (map-reduce, Hadoop), query optimization, data privacy, multidimensional analysis of text documents, and data warehousing and big data for real-world applications such as agricultural trade data, air quality measurement, electrical vehicle, etc. In the areas of data mining and knowledge discovery, the topics included traditional data-mining topics such as frequent item sets, association, etc., and machine-learning techniques. It was especially notable that some papers covered emerging real-world applications with a special focus on social networks.

Due to the maturity of DaWaK, several editors of well-known journals support our conference. This year, we had three special issues of well-known journals: *Distributed and Parallel Databases*, Springer, *Journal of Concurrency and Computation: Practice and Experience*, Wiley, and *Transactions on Large-Scale Data- and Knowledge-Centered Systems*, TLDKS, Springer.

We would like to thank all authors for submitting their research papers to DaWaK 2017. We express our gratitude to all the Program Committee members and to the external reviewers, who reviewed the papers thoroughly and in a timely manner. Finally, we would like to thank Gabriela Wagner for her endless help and support.

Hope you enjoy the proceedings.

June 2017

Ladjel Bellatreche
Sharma Chakravarthy

Organization

Program Committee Co-chairs

Ladjel Bellatreche	LIAS/ISAE-ENSMA, Poitiers, France
Sharma Chakravarthy	The University of Texas at Arlington, USA

Program Committee

Alberto Abelló	Universitat Politecnica de Catalunya, Spain
Sonali Agarwal	Indian Institute of information Technology, Allahabad, India
Mohammed Al-Kateb	Teradata Labs, USA
Toshiyuki Amagasa	University of Tsukuba, Japan
Torben Bach Pedersen	Aalborg University, Denmark
Elena Baralis	Politecnico di Torino, Italy
Ladjel Bellatreche	ENSMA, France
Sadok Ben Yahia	Faculty of Sciences of Tunis, Tunisia
Jorge Bernardino	ISEC - Polytechnic Institute of Coimbra, Portugal
Mikael Berndtsson	University of Skovde, Sweden
Vasudha Bhatnagar	Delhi University, India
Omar Boussaid	University of Lyon, France
Stephane Bressan	National University of Singapore, Singapore
Sharma Chakravarthy	The University of Texas at Arlington, USA
Isabelle Comyn-Wattiau	ESSEC Business School, Paris, France
Bruno Cremilleux	Université de Caen, France
Alfredo Cuzzocrea	University of Trieste, Italy
Laurent d'Orazio	University of Rennes 1, France
Karen Davis	University of Cincinnati, USA
Claudia Diamantini	Università Politecnica delle Marche, Italy
Alin Dobra	University of Florida, USA
Josep Domingo-Ferrer	Universitat Rovira i Virgili, Spain
Dejing Dou	University of Oregon, USA
Curtis Dyreson	Utah State University, USA
Markus Endres	University of Augsburg, Germany
Leonidas Fegaras	The University of Texas at Arlington, USA
Filippo Furfaro	DIMES - University of Calabria, Italy
Pedro Furtado	Universidade de Coimbra, Portugal, Portugal
Carlos Garcia-Alvarado	Amazon
Kazuo Goda	University of Tokyo, Japan
Matteo Golfarelli	University of Bologna, Italy
Sergio Greco	University of Calabria, Italy
Takahiro Hara	Osaka University, Japan

Frank Hoppner	Ostfalia University of Applied Sciences, Germany
Yoshiharu Ishikawa	Nagoya University, Japan
Stéphane Jean	LIAS/ISAE-ENSMA and University of Poitiers, France
Lili Jiang	Umeå University, Sweden
Vana Kalogeraki	Athens University of Economics and Business, Greece
Selma Khouri	LCSI/ESI, Algeria and LIAS/ISAE-ENSMA, France
Uday Kiran	University of Tokyo, Japan
Nhan Le-Thanh	Nice Sophia Antipolis University, France
Jens Lechtenboerger	Westfalische Wilhelms - Universität Münster, Germany
Wookey Lee	Inha University, South Korea
Carson K. Leung	University of Manitoba, Canada
Sofian Maabout	University of Bordeaux, France
Sanjay Kumar Madria	Missouri University of Science and Technology, USA
Yannis Manolopoulos	Aristotle University, Greece
Patrick Marcel	Université François Rabelais Tours, France
Amin Mesmoudi	Poitiers University, France
Jun Miyazaki	Tokyo Institute of Technology, Japan
Anirban Mondal	Shiv Nadar University, India
Yasuhiko Morimoto	Hiroshima University, Japan
Makoto Onizuka	Osaka University, Japan
Carlos Ordonez	University of Houston, USA
Alex Poulovassilis	Birkbek, University of London, UK
Praveen Rao	University of Missouri-Kansas City, USA
Goce Ristanoski	Data61, CSIRO, Australia
Laura Rusu	IBM, Australia
Alkis Simitsis	HP Labs, USA
David Taniar	Monash University, Australia
Olivier Teste	IRIT, University of Toulouse, France
Dimitri Theodoratos	New Jersey Institute of Technology, USA
Predrag Tosic	Washington State University, USA
Panos Vassiliadis	University of Ioannina, Greece
Guangtao Wang	NTU, Singapore
Robert Wrembel	Poznan University of Technology, Poland
Haruo Yokota	Tokyo Institute of Technology, Japan
Osmar Zaiane	University of Alberta, Canada

Additional Reviewers

Bijay Neupane	Aalborg University, Denmark
Christian Thomsen	Aalborg University, Denmark
Muhammad Aamir Saleem	Aalborg University, Denmark
Yuya Sasaki	Osaka University, Japan
Hieu Hanh Le	Tokyo Institute of Technology, Japan
Dominique Li	University of Tours, France
Yuto Hayamizu	University of Tokyo, Japan

Hiroyuki Yamda	University of Tokyo, Japan
Sharanjit Kaur	University of Delhi, India
Rakhi Saxena	University of Delhi, India
Swagata Duari	University of Delhi, India
Fan Jiang	University of Manitoba, Canada
Adam Pazdor	University of Manitoba, Canada
Syed Tanbeer	University of Manitoba, Canada
Xiaoying Wu	Wuhan University, China
Aggeliki Dimitriou	National Technical University of Athens, Greece
Souvik Sinha	New Jersey Institute of Technology, USA
Antonio Corral	University of Almeria, Spain
Anastasios Gounaris	Aristotle University of Thessaloniki, Greece
Johannes Kastner	University of Augsburg, Germany
Lena Rudenko	University of Augsburg, Germany
Rohit Kumar	Université Libre de Bruxelles, Belgium
Rana Faisal Munir	Universitat Politècnica de Catalunya, Spain
Luca Cagliero	Politecnico di Torino, Italy
Evelina Di Corso	Politecnico di Torino, Italy
Paolo Garza	Politecnico di Torino, Italy
Trung Dung LE	Université de Rennes 1, France
Anas Katib	University of Missouri-Kansas City, USA
Monica Senapati	University of Missouri-Kansas City, USA
Dig Vijay Kumar Yarlagadda	University of Missouri-Kansas City, USA
Rodrigo Rocha Silva	São Paulo State Technological College, FATEC-MC, Brazil
Emanuele Storti	Università Politecnica delle Marche, Italy
Alex Mircoli	Università Politecnica delle Marche, Italy

Contents