Lecture Notes in Computer Science **9093**

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

Mohamed A. Sharaf · Muhammad Aamir Cheema
Jianzhong Qi (Eds.)

Databases Theory and Applications

26th Australasian Database Conference, ADC 2015
Melbourne, VIC, Australia, June 4–7, 2015
Proceedings

Springer

Editors
Mohamed A. Sharaf
University of Queensland
Brisbane
Queensland
Australia

Jianzhong Qi
The University of Melbourne
Melbourne
Australia

Muhammad Aamir Cheema
Monash University
Clayton
Australia

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISBN 978-3-319-19547-6 ISBN 978-3-319-19548-3 (eBook)
DOI 10.1007/978-3-319-19548-3

Library of Congress Control Number: 2015939808

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

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Preface

It is our pleasure to present to you the proceedings of the 26th Australasian Database Conference (ADC2015), which took place in Melbourne, Australia. The Australasian Database Conference is an annual international forum for sharing the latest research advancements and novel applications of database systems, data-driven applications, and data analytics between researchers and practitioners from around the globe, particularly Australia and New Zealand. The mission of ADC is to share novel research solutions to problems of today's information society that fulfill the needs of heterogeneous applications and environments and to identify new issues and directions for future research. ADC seeks papers from academia and industry presenting research on all practical and theoretical aspects of advanced database theory and applications, as well as case studies and implementation experiences. All topics related to database are of interest and within the scope of the conference. ADC gives researchers and practitioners a unique opportunity to share their perspectives with others interested in the various aspects of database systems.

ADC 2015 was co-located with ACM SIGMOD 2015 (May 31-June 4, 2015, Melbourne, Australia) and held immediately after it. As in previous years, ADC 2015 accepted all papers that the Program Committee considered as being of ADC quality without setting any predefined quota. The conference received 43 submissions and accepted 29 papers, including 24 full research papers and five demo papers. The Program Committee that selected the papers consisted of 49 members from around the globe, including Australia, China, Germany, Hong Kong, Japan, New Zealand, Switzerland, Taiwan, the UK, and the USA, who were thorough and dedicated to the reviewing process. Each paper was peer reviewed in full by at least three independent reviewers, and in some cases four referees produced independent reviews. A conscious decision was made to select the papers for which all reviews were positive and favorable.

We would like to thank all our colleagues who served on the Program Committee or acted as external reviewers. We would also like to thank all the authors who submitted their papers, and the attendees. We hope that with these proceedings, you can have an overview of this vibrant research community and its activities. We encourage you to make submissions to the next ADC conference and contribute to this community.

June 2015

Mohamed A. Sharaf
Muhammad A. Cheema
Jianzhong Qi

General Chair's Welcome Message

Welcome to the proceedings of the 26th Australasian Database Conference (ADC2015)! ADC is a leading Australia- and New Zealand-based international conference on research and applications of database systems, data-driven applications, and data analytics. In the past 10 years, ADC has been held in Brisbane (2014), Adelaide (2013), Melbourne (2012), Perth (2011), Brisbane (2010), Wellington (2009), Wollongong (2008), Ballarat (2007), Hobart (2006), and Newcastle (2005). This year, the ADC conference came back to Melbourne.

In the past, the ADC conference series was held as part of the Australasian Computer Science Week (ACSW). Starting from 2014, the ADC conferences have departed from ACSW as the database research community in Australasia has grown significantly larger. Now the new ADC conference has an expanded research program and focuses on community-building through a PhD School. ADC 2015 was the second of this new ADC conference series.

The conference this year had three eminent speakers to give keynote speeches: Tamer Özsu from the University of Waterloo, Canada, Gerhard Weikum from the Max Planck Institute for Informatics, Germany, and Bingsheng He from the Nanyang Technological University, Singapore. In addition to 24 full research papers and five demo papers carefully selected by the Program Committee, we were also very fortunate to have two invited talks presented by world-leading researchers: Cyrus Shahabi from the University of Southern California, USA, and Lu Qin from the University of Technology, Sydney, Australia. We had a three-day PhD School program as part of this year's ADC.

We wish to take this opportunity to thank all speakers, authors, and organizers. I would also specially thank our Organizing Committee members: Program Committee Co-chairs Mohamed A. Sharaf and Muhammad A. Cheema, for their dedication in ensuring a high-quality program, Proceedings Chair Jianzhong Qi, for his effort in delivering the conference proceedings timely, and Local Co-chairs Jianxin Li and Jeffrey Chan, for their consideration in covering every detail of the conference logistics. Without them, this year's ADC would not have been a success.

Melbourne is a multicultural city and ADC 2015 was held on the campus of the RMIT University residing at the heart of the City of Melbourne. We trust all ADC2015 participants had a wonderful experience with the conference, the campus, and the city.

Rui Zhang

Organization

General Chair

Rui Zhang	University of Melbourne, Australia

Program Committee Co-chairs

Mohamed A. Sharaf	University of Queensland Australia
Muhammad A. Cheema	Monash University, Australia

Local Co-chairs

Jianxin Li	RMIT University, Australia
Jeffrey Chan	RMIT University, Australia

Proceedings Chair

Jianzhong Qi	University of Melbourne, Australia

Steering Committee

Rao Kotagiri	University of Melbourne, Australia
Timos Sellis	RMIT University, Australia
Gill Dobbie	University of Auckland, New Zealand
Alan Fekete	University of Sydney, Australia
Xuemin Lin	University of New South Wales, Australia
Yanchun Zhang	Victoria University, Australia
Xiaofang Zhou	University of Queensland

Program Committee

Zhifeng Bao	University of Tasmania, Australia
Ljiljana Brankovic	University of Newcastle, Australia
Lijun Chang	University of New South Wales, Australia
Byron Choi	Hong Kong Baptist University, Hong Kong, SAR China

Mohamed Eltabakh	WPI, USA
Markus Endres	University of Augsburg, Germany
Junbin Gao	Charles Sturt University, Australia
Dimitrios Georgakopoulos	RMIT University, Australia
Janusz Getta	University of Wollongong, Australia
Yu Gu	Northeastern University, China
Michael Houle	National Institute of Informatics, Japan
Guangyan Huang	Deakin University, Australia
Yoshiharu Ishikawa	Nagoya University, Japan
Jiuyong Li	University of South Australia, , Australia
Ruixuan Li	Huazhong University of Science and Technology, China
Xue Li	University of Queensland, Australia
Zhixu Li	Soochow University, China
Muhammad Asif Naeem	University of Auckland, New Zealand
Lu Qin	University of Technology, Sydney, Australia
Goce Ristanoski	University of Melbourne, Australia
Uwe Roehm	University of Sydney, Australia
Shazia Sadiq	University of Queensland, Australia
Sherif Sakr	NICTA, Australia
Mohamed Sarwat	Arizona State University, USA
Saket Sathe	EPFL, Switzerland
Shuo Shang	China University of Petroleum, China
Quan Z. Sheng	University of Adelaide, Australia
Laurianne Sitbon	Queensland University of Technology, Australia
Bela Stantic	Griffith University, Australia
Millist Vincent	University of South Australia, Australia
Junhu Wang	Griffith University, Australia
Lei Wang	University of Wollongong, Australia
Yan Wang	Macquarie University, Australia
Chuan Xiao	Nagoya University, Japan
Guandong Xu	University of Technology, Australia
Jian Yang	Macquarie University, Australia
Mi-Yen Yeh	Academia Sinica, Taiwan
Daisaku Yokoyama	University of Tokyo, Japan
Weiren Yu	Imperial College, UK
Nayyar Zaidi	Monash University, Australia
Wenjie Zhang	University of New South Wales, Australia
Xianchao Zhang	Dalian University of Technology, China
Xiuzhen Zhang	RMIT University, Australia
Ying Zhang	University of Technology, Sydney, Australia
Xiang Zhao	National University of Defence Technology, China
Kai Zheng	University of Queensland, Australia
Xiangmin Zhou	Victoria University, Australia
Guido Zuccon	Queensland University of Technology, Australia
Andreas Zufle	Ludwig-Maximilians-Universität, Germany

External Reviewers

Zhan Bu	Nanjing University of Finance and Economics, China
Po-Yu Chen	Imperial College London, UK
Hamidu Abdel-Fatao	University of South Australia
Tao Hoang	University of South Australia
Lili Jiang	Max Planck Institute, Germany
Selasi Kwashie	University of South Australia
Yongrui Qin	The University of Adelaide, Australia
Lie Qu	Macquarie University, Australia
Jie Shao	University of Electronic Science and Technology of China
Xianzhi Wang	The University of Adelaide, Australia
Dong Wen	University of Technology, Sydney, Australia
Zhengyuan Xue	Huazhong University of Science and Technology, China
Diyi Yang	Carnegie Mellon University, USA
Haibin Zhang	Macquarie University, Australia
Cong Zhao	Imperial College London, UK
Xiaoming Zheng	Macquarie University, Australia

Tutorials

An Overview of Graph Data Management and Analysis

M. Tamer Özsu

University of Waterloo

Abstract. Graphs have always been important data types for database researchers. With the recent growth of social networks, Wikipedia, Linked Data, RDF, and other networks, the interest in managing very large graphs have again gained momentum. In this talk I will first present a taxonomy of graph processing systems and then summarize research on querying and analytics over property graphs and management and querying of RDF graphs.

Short Biography. M. Tamer Özsu is Professor of Computer Science at the David R. Cheriton School of Computer Science, and Associate Dean (Research) of the Faculty of Mathematics at the University of Waterloo. His research is in data management focusing on large-scale data distribution and management of non-traditional data. He is a Fellow of the Association for Computing Machinery (ACM), and of the Institute of Electrical and Electronics Engineers (IEEE), an elected member of the Science Academy of Turkey, and member of Sigma Xi and American Association for the Advancement of Science (AAAS). He currently holds a Cheriton Faculty Fellowship at the University of Waterloo.

Knowledge Graphs: From a Fistful of Triples to Deep Data and Deep Text

Gerhard Weikum

Max Planck Institute for Informatics

Abstract. Knowledge graphs (KG's), aka. knowledge bases, are huge repositories of entities, their types, properties, and relationships between entities. KG's have become a key asset for search, analytics, recommendations, and data integration on the Web and in enterprises. Rooted in academic research and community projects such as DBpedia, Freebase, and Yago, KG's are now intensively used at big industrial stakeholders such as Google, Microsoft, Yahoo, Alibaba, Bloomberg, Walmart, and many others.

This talk reviews the knowledge graph technology, discussing strengths and limitations and pointing out opportunities for further research. The talk spans a spectrum of issues that arise in the life-cycle and use-cases of a KG: construction from data and text sources, maintaining over time, extension with common sense knowledge, querying and mining, boosting language understanding and text analytics, and usability issues in interactive exploration.

Short Biography. Gerhard Weikum is a Scientific Director at the Max Planck Institute for Informatics in Saarbruecken, Germany, and also an Adjunct Professor at Saarland University. He graduated from the University of Darmstadt, Germany.

Weikum's research spans transactional and distributed systems, self-tuning database systems, DB&IR integration, and the automatic construction of knowledge bases from Web and text sources. He co-authored a comprehensive textbook on transactional systems, received the VLDB 10-Year Award for his work on automatic DB tuning, and is one of the creators of the YAGO knowledge base.

Gerhard Weikum is an ACM Fellow, a member of the Academia Europaea, and a member of several academies in Germany. He has served on various editorial boards, including Communications of the ACM, ACM TODS and ACM TWEB, and as PC chair of conferences like ACM SIGMOD, Data Engineering, and CIDR. From 2003 through 2009 he was president of the VLDB Endowment. He received a Google Focused Research Award in 2010, the ACM SIGMOD Contributions Award in 2011, and an ERC Synergy Grant in 2013.

Emerging HPC Technologies for Real-Time (Big) Data Analytics: A Tutorial

Bingsheng He

Nanyang Technological University

Abstract. Big data has become a buzz word. Among various big-data challenges, real-time data analytics has been identified as one of the most exciting and promising areas for both academia and industry. We are facing the challenges at all levels ranging from sophisticated algorithms and procedures to mine the gold from massive data to high-performance computing (HPC) techniques and systems to get the useful data in time. In this tutorial, we review the system design and implementation of HPC technologies (including GPUs, FPGAs and RDMA etc) as weapons to address the performance requirement of real-time data analytics. Particularly, we focus on the interplay between HPC and real-time data analytics, where real-time data analytics also poses significant challenges to the design and implementation of HPC technologies. I will also present our recent research efforts in developing real-time data analytics systems by Xtra Computing Group at NTU Singapore (http://pdcc.ntu.edu.sg/xtra/) as well as related research from other groups. Finally, I will outline some open problems in this field.

Short Biography. Dr. Bingsheng He is currently an Assistant Professor at Division of Networks and Distributed Systems, School of Computer Engineering, Nanyang Technological University. Before that, he held a research position in the System Research group of Microsoft Research Asia (2008 - 2010), where his major research was building high performance cloud computing systems for Microsoft. He got the Bachelor degree in Shanghai Jiao Tong University (1999 - 2003), and the Ph.D. degree in Hong Kong University of Science & Technology (2003 - 2008). His current research interests include cloud computing, database systems and high performance computing. His papers are published in prestigious international journals (such as ACM TODS, IEEE TKDE/TPDS/TC) and proceedings (such as ACM SIGMOD, VLDB/PVLDB, ACM/IEEE SuperComputing, PACT, HPDC, ACM SoCC, and CIDR). He has been awarded with the IBM Ph.D. fellowship (2007 - 2008) and with NVIDIA Academic Partnership (2010 - 2011).

Invited Talks

Graph Processing in the Era of Big Data

Lu Qin

University of Technology, Sydney

Abstract. With the emergence and rapid proliferation of applications that deal with big graphs, such as web graphs (Google, Yahoo), social networks (Facebook, Twitter), e-commerce networks (Amazon, Ebay), and road networks, graph processing has become increasingly prevalent and important in recent years. However, in the era of big data, the explosion and profusion of available graph data in a wide range of application domains rise up new challenges and opportunities in graph processing. In this talk, I will first investigate these new challenges. To tickle these challenges, I will then introduce our recent research on big graph processing in terms of new graph query semantics, new query processing algorithms, new graph indexing techniques, and new computing paradigms. Finally, I will discuss our potential future research directions for graph processing.

Short Biography. Dr. Lu Qin received his PhD degree in 2010 from the department of Systems Engineering and Engineering Management (SEEM) in the Chinese University of Hong Kong (CUHK). Dr. Qin is currently a core member in the Centre of Quantum Computation and Intelligent Systems (QCIS) at the University of Technology, Sydney (UTS). Dr. Qin's research interests include algorithm design and analysis on big data, big graph processing in the cloud, and big graph searching and mining. He has published 50+ top-tier conference/journal papers including 7 SIGMOD papers, 9 VLDB papers, and 7 ICDE papers in the top-3 database conferences, and 7 VLDB journal papers, 1 Algorithmica paper, and 2 TKDE papers in the top-ranked database and algorithm journals. His book entitled "Keyword Search in Databases" is the first monograph on keyword search in databases. Dr. Qin served as a program committee member of a lot of top database and data mining conferences. He has received several research funds from Australian government and UTS.

Spatial Indexing and Spatial Crowdsourcing of User-Generated-Video

Cyrus Shahabi

University of Southern California

Abstract. I will start by showing a demo of our MediaQ system prototype. MediaQ is a novel online media management system to collect, organize, share, and search user-generated mobile videos (UGV) from the public. Subsequently, I focus on two of the research challenges underlying MediaQ. First, I discuss our approach to index and search UGVs more effectively by utilizing the smartphone sensors (e.g., GPS locations, compass directions) to geo-tag each video frame by the spatial extent of its coverage area. Next, I introduce our generic framework for spatial crowdsourcing and discuss various techniques for optimal assignment of spatiotemporal tasks (e.g., UGV data collection) to human workers. Finally, I conclude by summarizing our ongoing efforts in spatial crowdsourcing.

Short Biography. Cyrus Shahabi is a Professor of Computer Science and Electrical Engineering and the Director of the Information Laboratory (InfoLAB) at the Computer Science Department and also the Director of the NSF's Integrated Media Systems Center (IMSC) at the University of Southern California (USC). He is also the director of the Informatics Program at USC's Viterbi School of Engineering. He was the CTO and co-founder of a USC spin-off, Geosemble Technologies, which was acquired in July 2012. Since then, he founded another company, ClearPath, focusing on predictive path-planning for car navigation systems. He received his B.S. in Computer Engineering from Sharif University of Technology in 1989 and then his M.S. and Ph.D. Degrees in Computer Science from the University of Southern California in May 1993 and August 1996, respectively. He authored two books and more than two hundred research papers in the areas of databases, GIS and multimedia with an h-index of 41. He also holds more than 12 US Patents.

Dr. Shahabi has received funding from agencies including NSF, NIJ, NASA, NIH, DARPA, AFRL, and DHS as well as industries such as Chevron, Google, HP, Intel, Microsoft, NCR, NGC and Oracle. He was an Associate Editor of IEEE Transactions on Parallel and Distributed Systems (TPDS) from 2004 to 2009 and IEEE Transactions on Knowledge and Data Engineering (TKDE) from 2010 to 2013. He is currently on the editorial board of the VLDB Journal, ACM Transactions on Spatial Algorithms and Systems (TSAS), and ACM Computers in Entertainment. He is the founding chair of IEEE NetDB workshop and also the general co-chair of ACM GIS 2007 - 2009. He chaired the nomination committee of ACM SIGSPATIAL 2011 - 2014. He was a PC co-Chair of DASFAA 2015, IEEE MDM 2013 and IEEE BigData 2013, and regularly serves on the program committee of major conferences such as VLDB, ACM SIGMOD, IEEE ICDE, ACM SIGKDD, and ACM Multimedia.

Dr. Shahabi is a fellow of IEEE, and a recipient of the ACM Distinguished Scientist award in 2009, the 2003 U.S. Presidential Early Career Awards for Scientists and Engineers (PECASE), the NSF CAREER award in 2002, and the 2001 Okawa Foundation Research Grant for Information and Telecommunications.

Contents

Research Papers

Demo Papers