

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, Lancaster, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Friedemann Mattern

ETH Zurich, Zurich, Switzerland

John C. Mitchell

Stanford University, Stanford, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

TU Dortmund University, Dortmund, Germany

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max Planck Institute for Informatics, Saarbrücken, Germany

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

Muhammad Aamir Cheema · Wenjie Zhang
Lijun Chang (Eds.)

Databases Theory and Applications

27th Australasian Database Conference, ADC 2016
Sydney, NSW, September 28–29, 2016
Proceedings

Editors

Muhammad Aamir Cheema
Monash University
Clayton
Australia

Lijun Chang
University of New South Wales
Sydney, NSW
Australia

Wenjie Zhang
School of Computer Science and
Engineering
University of New South Wales
Sydney
Australia

ISSN 0302-9743

ISSN 1611-3349 (electronic)

Lecture Notes in Computer Science

ISBN 978-3-319-46921-8

ISBN 978-3-319-46922-5 (eBook)

DOI 10.1007/978-3-319-46922-5

Library of Congress Control Number: 2016952520

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

© Springer International Publishing AG 2016

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, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

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

It is our pleasure to present to you the proceedings of the 27th Australasian Database Conference (ADC2016), which took place in Sydney, 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 2016 was held during September 26–29, 2016 in Sydney, Australia. As in previous years, ADC 2016 accepted all papers that the Program Committee considered as being of ADC quality without setting any predefined quota. The conference accepted 33 research papers including some invited papers, and 11 demo papers. Each paper including every invited submission 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. The Program Committee that selected the papers comprised 53 members from around the globe including Australia, Bangladesh, Canada, China, Finland, Germany, Japan, New Zealand, Qatar, Singapore, Taiwan, and the UK, who were thorough and dedicated to the reviewing process.

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.

August 2016

Muhammad A. Cheema
Wenjie Zhang
Lijun Chang

General Chair's Welcome Message

Welcome to the proceedings of the 27th Australasian Database Conference (ADC2016)! 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 Melbourne (2015), Brisbane (2014), Adelaide (2013), Melbourne (2012), Perth (2011), Brisbane (2010), Wellington (2009), Wollongong (2008), Ballarat (2007), and Hobart (2006). This year, the ADC conference came to Sydney.

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 2016 was the third of this new ADC conference series.

The conference this year had four eminent speakers to give keynote speeches: Michael Stonebraker from Massachusetts Institute of Technology, USA, Mark Sanderson from RMIT University, Australia, Stephan Winter from the University of Melbourne, Australia, and Mohamed F. Mokbel from the University of Minnesota, USA. In addition to 33 full research papers and 11 demo papers carefully selected by the Program Committee, we were also very fortunate to have two invited talks presented by world-leading researchers: Zi (Helen) Huang from the University of Queensland, Australia, and Ying Zhang from the University of Technology Sydney, Australia. We had a two-day PhD School program as part of this year's ADC with great support from invited speakers: Alan David Fekete from the University of Sydney, Australia, Jeffrey Xu Yu from the Chinese University of Hong Kong, China, Zi (Helen) Huang from the University of Queensland, Australia, Ying Zhang from the University of Technology Sydney, Australia, and Kai (Alex) Qin from RMIT University, Australia.

We wish to take this opportunity to thank all speakers, authors, and organizers. I would especially like to thank our Organizing Committee members: Program Committee co-chairs Muhammad A. Cheema and Wenjie Zhang, for their dedication in ensuring a high-quality program, proceedings chair Lijun Chang, for his effort in delivering the conference proceedings timely, local co-chairs Jianbin Qin, Shiyu Yang, and Xiaoyang Wang, for their efforts in covering every detail of the conference logistics, and webmaster Xiang Wang for his effort in maintaining the conference website. I would also like to thank the University of New South Wales (UNSW) for the generous financial support. Without them, this year's ADC would not have been a success.

Sydney is a multi-cultural city and ADC 2016 was held on the main campus of the University of New South Wales located on a 38-hectare site in the suburb of Kensington. We trust all ADC 2016 participants had wonderful experience with the conference, the campus, and the city.

Xuemin Lin

Organization

General Chair

Xuemin Lin University of New South Wales, Australia

Program Co-chairs

Muhammad A. Cheema Monash University, Australia
Wenjie Zhang University of New South Wales, Australia

Local Co-chairs

Jianbin Qin University of New South Wales, Australia
Shiyu Yang University of New South Wales, Australia
Xiaoyang Wang University of Technology Sydney, Australia

Proceedings Chair

Lijun Chang University of New South Wales, Australia

Steering Committee

Rao Kotagiri University of Melbourne, Australia
Timos Sellis Swinburne University of Technology, 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, Australia

Program Committee

Muhammad Ahmad McGill University, Canada
Shafiq Alam The University of Auckland, New Zealand
Mohammed Ali Bangladesh University of Engineering and Technology,
 Bangladesh
Zhifeng Bao RMIT University, Australia
Ljiljana Brankovic University of Newcastle, Australia
Junbin Gao Charles Sturt University, Australia
Dimitrios RMIT University, Australia
 Georgakopoulos

Janusz Getta	University of Wollongong, Australia
Yusuke Gotoh	Okayama University, Japan
Yu Gu	Northeastern University, China
Tanzima Hashem	Bangladesh University of Engineering and Technology, Bangladesh
Michael Houle	National Institute of Informatics, Japan
Guangyan Huang	Deakin University, Australia
Arijit Khan	Nanyang Technological University, Singapore
Khalid Latif	Aalto University, Finland
Jianxin Li	University of Western Australia, Australia
Jianzhong Li	Harbin Institute of Technology, China
Jiuyong Li	University of South Australia, Australia
Xue Li	University of Queensland, Australia
Zhixu Li	Soochow University, China
Ghulam Murtuza	Intersect, Australia
Muhammad Naem	University of Auckland, New Zealand
Mourad Ouzzani	Qatar Computing Research Institute, Qatar
Weining Qian	East China Normal University, China
Jianbin Qin	University of New South Wales, Australia
Lu Qin	University of Technology, Sydney, Australia
Goce Ristanoski	National ICT Australia, Australia
Shazia Sadiq	University of Queensland, Australia
Sherif Sakr	University of New South Wales, Australia
Flora Salim	RMIT University, Australia
Haichuan Shang	National Institute of Information and Communications, Japan
Quan Sheng	University of Adelaide, Australia
Bela Stantic	Griffith University, Australia
Farhan Tauheed	Oracle
Anwaar Ulhaq	Victoria University, Australia
Junhu Wang	Griffith University, Australia
Lei Wang	University of Wollongong, Australia
Yan Wang	Macquarie University, Australia
Yang Wang	University of New South Wales, Australia
Lin Wu	University of Adelaide, Australia
Chuan Xiao	Nagoya University, Japan
Guandong Xu	University of Technology, Australia
Mi-Yen Yeh	Academia Sinica, Taiwan
Daisaku Yokoyama	University of Tokyo, Japan
Weiren Yu	Imperial College London, UK
Nayyar Zaidi	Monash University, Australia
Xianchao Zhang	Dalian University of Technology, China
Xiuzhen Zhang	RMIT University, Australia
Ying Zhang	University of Technology Sydney, Australia
Kai Zheng	University of Queensland, Australia
Aoying Zhou	East China Normal University, China

Xiangmin Zhou	Victoria University, Australia
Guido Zuccon	Queensland University of Technology, Australia
Andreas Zufle	Ludwig-Maximilians-Universität, Germany

External Reviewers

Maitraye Das	Bangladesh University of Engineering and Technology, Bangladesh
Wei Kang	University of South Australia, Australia
Ali Shemshadi	University of Adelaide, Australia
Mehnaz Tabassum	Bangladesh University of Engineering and Technology, Bangladesh
Xin Wang	Tianjin University, China
Wei Zhang	University of Adelaide, Australia
Yihong Zhang	University of Adelaide, Australia

ADC Keynotes

Big Data, Technological Disruption and the 800 Pound Gorilla in the Corner

Michael Stonebraker

Massachusetts Institute of Technology, Cambridge, USA

Abstract. This talk will focus on the current market for “Big Data” products, specifically those that deal with one or more of “the 3 V’s”. I will suggest that the **volume** problem for business intelligence applications is pretty well solved by the data warehouse vendors; however upcoming “data science” tasks are poorly supported at present. On the other hand, there is rapid technological progress, so “stay tuned”. In the **velocity** arena recent “new SQL” and stream processing products are doing a good job, but there are a few storm clouds on the horizon. The **variety** space has a collection of mature products, along with considerable innovation from startups. I will discuss opportunities in this space, especially those enabled by possible disruption from new technology. Also discussed will be the pain levels I observe in current enterprises, culminating in my presentation of “the 800 pound gorilla in the corner”.

Short Biography. Professor Stonebraker has been a pioneer of data base research and technology for more than forty years. He was the main architect of the INGRES relational DBMS, and the object-relational DBMS, POSTGRES. These prototypes were developed at the University of California at Berkeley where Stonebraker was a Professor of Computer Science for twenty five years. More recently at M.I.T. he was a co-architect of the Aurora/Borealis stream processing engine, the C-Store column-oriented DBMS, the H-Store transaction processing engine, the SciDB array DBMS, and the Data Tamer data curation system. Presently he serves as Chief Technology Officer of Paradigm4 and Tamr, Inc. Professor Stonebraker was awarded the ACM System Software Award in 1992 for his work on INGRES. Additionally, he was awarded the first annual SIGMOD Innovation award in 1994, and was elected to the National Academy of Engineering in 1997. He was awarded the IEEE John Von Neumann award in 2005 and the 2014 Turing Award, and is presently an Adjunct Professor of Computer Science at M.I.T, where he is co-director of the Intel Science and Technology Center focused on big data.

Getting Rid of the Ten Blue Links

Mark Sanderson

RMIT University, Melbourne, Australia

Abstract. In this talk, I will first give a brief overview of the IR group at RMIT. Then I will describe the work we are doing at RMIT to change one of the commonest web pages we all look at: the Search Result Page (SERP). In our work we are looking to replace the SERP with a set of answer passages that address the user's query. In the context of general web search, the problem of finding answer passages has not been explored extensively. Previous studies have found that many informational queries can be answered by a passage of text extracted from a retrieved document, relieving the user from having to read the actual document. While current passage retrieval methods that focus on topical relevance have been shown to be not effective at finding answers, the result shows that more knowledge is required to identify answers in the document. We have been formulating the answer passage extraction problem as a summarization task. We initially used term distributions extracted from a Community Question Answering (CQA) service to generate more effective summaries of retrieved web pages. An experiment was conducted to see the benefit of using the CQA data in finding answer passages. We analyze the fraction of answers covering a set of queries, the quality of the corresponding result from the answering service, and their impact on the generated summaries. I will also talk about recent work where we re-rank retrieved passages according to the summary quality and incorporating document summarizability into the ranking function.

Short Biography. Prof. Mark Sanderson is the deputy head of the School of Computer Science and IT at RMIT University in Melbourne, Australia. According to a range of international ranking systems RMIT is in the top 8 of CS schools in Australia. Prof Sanderson is head of the RMIT Information Retrieval (IR) group, which is regarded as the leading IR group in Australia. He is co-editor of Foundations and Trends in Information Retrieval, which is currently the highest impact rated IR journal. He is also an associate editor of IEEE TKDE and of ACM TWeb. Prof. Sanderson was co-PC chair of ACM SIGIR in 2009 and 2012, and general chair of the conference in 2004. Prof Sanderson is also a visiting professor at NII in Tokyo.

ADC Invited Talks

Location Understanding in Social Media

Zi (Helen) Huang

The University of Queensland, St Lucia, Australia

Abstract. Location data has been playing an important role in many social media applications, particularly the location-based services. Unfortunately, location information is often missing in social media data, such as online images. In this talk, we introduce novel methods to estimate missing locations for social images by effectively fusing multi-modalities of social media data. Interestingly, by integrating visual data and location data, such as geo-tagged images and check-ins, important location proximities can be better understood and discovered to users.

Short Biography. Dr Zi Huang received her BSc degree from Tsinghua University, China, in 2001, and her PhD in Computer Science from the University of Queensland, Australia, in 2007. She is currently an ARC Future Fellow with the School of Information Technology and Electrical Engineering, University of Queensland. Her research interests include multimedia indexing and search, social data analysis and knowledge discovery.

Continuous Spatial-Keyword Queries over Streaming Data

Ying Zhang

University of Technology Sydney, Ultimo, Australia

Abstract. As the prevalence of social media and GPS-enabled devices, a massive amount of geo-textual data has been generated in a stream fashion, leading to a variety of applications such as location-based recommendation and information dissemination. For example, a location-based e-coupon system may allow potentially millions of users to register their continuous spatial-keyword queries (e.g., interests in nearby sales) by specifying a set of keywords and a spatial region; the system then delivers each incoming spatial-textual object (e.g., a geo-tagged e-coupon) to all the matched queries (i.e., users) whose spatial and textual requirements are satisfied. In this talk, I will introduce our recent work on continuous spatial-keyword queries over streaming data. Novel indexing structures, which seamlessly and effectively integrate keyword and spatial information, will be presented to support various continuous spatial-keyword queries.

Short Biography. Ying Zhang is a senior lecturer and ARC DECRA research fellow (2014–2016) at QCIS, the University of Technology Sydney (UTS). He received his BSc and MSc degrees in Computer Science from Peking University, China, and PhD in Computer Science from the University of New South Wales, Australia. His research interests include query processing on spatial data, spatial-textual data, streaming data, uncertain data and graphs. He has published 40+ papers on prestigious conferences and journals such as SIGMOD, SIGIR, VLDB, ICDE, TODS, VLDBJ, and TKDE. He was an Australian Research Council Australian Postdoctoral Fellowship (ARC APD) holder during 2010 and 2013.

PhD School Keynotes

Where am I? Where Do I Want to Go?

Stephan Winter

The University of Melbourne, Parkville, Australia

Abstract. Sensors and the related information and communication technology get ever smarter in localizing people, vehicles, events or goods. And yet, at the end always a person is consuming this information, or even producing it (sometimes called “people as sensors”). In this talk I will focus on the gap between the concepts of people about their environment, and the concepts of sensors, spatial databases and geographic information systems. This gap is a major impediment for communication between people and systems; some examples are emergency calls, tracking bushfires, guiding an autonomously driving car, planning a trip through a city by public transport, helping people to evacuate, or simply the general search in a search engine (Ed Parsons, Geospatial Chief Technologist of Google, indicated that “about 1 in 3 of queries that people just type into a standard Google search bar are about places”). The talk will identify the issues with this gap, and show some steps to overcome this gap, including novel, complementary ways of representing spatial information in databases.

Short Biography. Stephan Winter is Professor in Spatial Information Science at the Department of Infrastructure Engineering, The University of Melbourne. He holds a PhD (Dr.-Ing.) from the University of Bonn (1997), and a habilitation from the Technical University Vienna (2001). Within spatial information science Stephan Winter is specializing on human wayfinding and navigation, with a vision of developing intelligent spatial machines. He has contributed to topics such as spatial human-computer interaction, network analysis, routing heuristics, and collaborative transportation and evacuation.

The Era of Big Spatial Data

Mohamed F. Mokbel

University of Minnesota, Minneapolis, USA

Abstract. In recent years, there has been an explosion in the amounts of spatial and spatio-temporal data produced from several devices including smart phones, space telescopes, medical devices. Unfortunately, managing and analyzing such big spatial data is hampered by the lack of specialized systems, techniques, and algorithms. While big data is well supported with a variety of distributed systems and cloud infrastructure, none of these systems or infrastructure provide any special support for spatial or spatio-temporal data. This talk presents our efforts in indexing, querying, and visualizing big spatial and spatio-temporal data. We will describe our efforts within SpatialHadoop; our full-fledged MapReduce framework with native support for spatial data, including support for basic spatial operations, computational geometry, and spatial visualization.

Short Biography. Mohamed F. Mokbel (Ph.D., Purdue University, MS, B.Sc., Alexandria University) is Associate Professor in the Department of Computer Science and Engineering, University of Minnesota. His research interests include the interaction of GIS and location-based services with database systems and cloud computing. His research work has been recognized by five Best Paper Awards and by the NSF CAREER award. Mohamed was the program co-chair for the ACM SIGSPATIAL GIS conference from 2008 to 2010, IEEE MDM Conference 2011 and 2014, and the General Chair for SSTD 2011. He is an Associate Editor for ACM TODS, ACM TSAS, VLDB journal, and GeoInformatica. Mohamed is an elected Chair of ACM SIGSPATIAL 2014-2017. For more information, please visit: www.cs.umn.edu/~mokbel.

PhD School Tutorials

Consistency Properties for Distributed Storage Platforms

Alan David Fekete

University of Sydney, Sydney, Australia

Abstract. A scalable and fault-tolerant data storage layer is extremely useful when constructing scalable fault-tolerant application software. The application developer is a consumer of a service provided by the storage layer, and the interface between these parties needs to be precise. This tutorial reflects on several bodies of research that relate to understanding the implications for the consumer, of the consistency aspects of that interface. We cover in turn how consistency properties can be defined, how the consumer can measure consistency, and how to reason about applications when they must run over storage with consistency that is weaker-than-ideal.

Short Biography. Alan Fekete is Professor of Enterprise Software Systems within the School of Information Technologies at the University of Sydney. His undergraduate education was at the University of Sydney, and his doctorate was earned in the mathematics department of Harvard University. He has been recognized as a Distinguished Scientist by ACM, and he serves as Trustee for the VLDB Endowment. He is particularly known for a body of research on transaction management. He is also active in CS Education.

Large Graph Processing: Algorithms and Systems

Jeffrey Xu Yu

The Chinese University of Hong Kong, Hong Kong, China

Abstract. The real applications that need graph processing techniques to handle a large graph can be found from many real applications including online social networks, biological networks, ontology, transportation networks, etc. In this talk, we will discuss some selected research topics on graph processing over large graphs from the algorithm perspectives and the systems perspectives.

Short Biography. Dr Jeffrey Xu Yu is a Professor in the Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong. His current main research interests include graph mining, graph query processing, graph pattern matching, keywords search in databases, and online social networks. Dr. Yu served as an Information Director and a member in ACM SIGMOD executive committee (2007–2011), an associate editor of IEEE Transactions on Knowledge and Data Engineering (2004–2008), and an associate editor in VLDB Journal (2007-2013). Currently he serves as an associate editor of WWW Journal, the International Journal of Cooperative Information Systems, the Journal on Health Information Science and Systems (HISS), and Journal of Information Processing. Dr. Yu served/serves in many organization committees and program committees in international conferences/workshops including PC Co-chair of APWeb'04, WAIM'06, APWeb/WAIM'07, WISE'09, PAKDD'10, DASFAA'11, ICDM'12, NDBC'13, ADMA'14, and CIKM'15.

Social Event and Behavior Modelling

Zi (Helen) Huang

The University of Queensland, St Lucia, Australia

Abstract. Social media data has provided great opportunities for many challenging data mining tasks. Its value has been widely exhibited in real-world applications. In this tutorial, we are focused on the impact of social media on public social event and individual online user behavior using heterogeneous social media data. We will review recent research advances in social event detection and prediction, and online user behavior modelling and prediction. As a step further, we will also discuss the effect of public events on individual behavior and explore their potential sequential correlations for new research opportunities.

Short Biography. Dr Zi Huang received her BSc degree from Tsinghua University, China, in 2001, and her PhD in Computer Science from the University of Queensland, Australia, in 2007. She is currently an ARC Future Fellow with the School of Information Technology and Electrical Engineering, University of Queensland. Her research interests include multimedia indexing and search, social data analysis and knowledge discovery.

Querying and Mining of Geo-Textual Data

Ying Zhang

University of Technology Sydney, Sydney, Australia

Abstract. Proliferation of geo-position technologies (e.g., smart phones, general mobile devices and sensor networks) and online social media (e.g., Twitter, Foursquare and Facebook) has resulted in a huge flood of location data being integrated with various textual data (e.g., tweets and news), leading to the “geo-textual” data. The ever increasing amounts of geo-textual data have tremendous potential for the discovery of new and useful knowledge in many key applications such as location-based services (LBS), e-marketing and social networks. In this tutorial, we first highlight the importance of geo-textual data management and the unique challenges that need to be addressed. Subsequently, we provide an overview of the existing research on geo-textual data, covering modelling, ad-hoc spatial-keyword queries, continuous spatial-keyword queries, mining of geo-textual data and other relevant topics. Finally, we discuss the future research directions in this important and growing research area.

Short Biography. Ying Zhang is a senior lecturer and ARC DECRA research fellow (2014–2016) at QCIS, the University of Technology Sydney (UTS). He received his BSc and MSc degrees in Computer Science from Peking University, China, and PhD in Computer Science from the University of New South Wales, Australia. His research interests include query processing on spatial data, spatial-textual data, streaming data, uncertain data and graphs. He has published 40+ papers on prestigious conferences and journals such as SIGMOD, SIGIR, VLDB, ICDE, TODS, VLDBJ, and TKDE. He was an Australian Research Council Australian Postdoctoral Fellowship (ARC APD) holder during 2010 and 2013.

Data-Driven Evolutionary Optimisation

Kai (Alex) Qin

RMIT University, Melbourne, Australia

Abstract. Optimisation aims at finding the best solution from numerous feasible ones, which is demanded in nearly every field when resolving various problems arising therein. Evolutionary optimisation represents a family of optimisation techniques based on Darwinian principles, characterized by a population of candidate solutions which will be evolved via nature-inspired operators to search for the optimum. Intrinsically, it belongs to a generate-and-test problem solver which incrementally produces a large volume of “data” (i.e. candidate solutions) as search progresses with search experience encoded by such “data”. In the past few decades, a lot of efforts had been made to enhance evolutionary optimisation techniques via exploiting (e.g. using data analytics techniques) the “data” generated in the course of search. However, modern optimisation problems, featured with the fast-growing scale, complexity and uncertainty, can seldom be tackled by simply hybridizing evolutionary optimisation with some off-the-shelf data analytics techniques, and therefore call for an in-depth investigation on how to leverage the “data” generated during search to facilitate optimisation. This tutorial aims to introduce a unified perspective on evolutionary optimisation techniques that adopts data analytics as an indispensable component, describe how to identify and address various data analytics tasks during the search process, and discuss an emerging research trend which makes use of search experience gained by solving some problems to facilitate solving other problems via knowledge transfer. The audience is expected to get to know the fundamentals and recent developments in data-driven evolutionary optimisation, and be inspired to employ such techniques to deal with their encountered optimisation problems.

Short Biography. Dr Kai Qin received his BEng degree from Southeast University, China, in 2001, and his PhD from Nanyang Technological University, Singapore, in 2007. He is now a lecturer in Computer Science and Information Technology at RMIT University. His research interests include evolutionary computation, machine learning, image processing, GPU computing and service computing. He has published 60+ papers and received two best paper awards. Two of his co-authored papers are the 1st and 3rd most cited papers (Thomson Reuters) in IEEE Transactions on Evolutionary Computation (ERA A*) over the last 10 years. He is currently chairing the IEEE Computational Intelligence Society task force on collaborative learning and optimisation, promoting research on synergizing machine learning and intelligent optimisation techniques to resolve challenging real-world problems which involve learning and optimisation as indispensable and interwoven tasks.

Contents

Social Network and Graphs

Finding Influencers in Temporal Social Networks Using Intervention Analysis.	3
<i>Maximilian Franzke, Janina Bleicher, and Andreas Züfle</i>	
Retrieving Top-k Famous Places in Location-Based Social Networks.	17
<i>Ammar Sohail, Ghulam Murtaza, and David Taniar</i>	
Comprehensive Graph and Content Feature Based User Profiling	31
<i>Peihao Tong, Junjie Yao, Liping Wang, and Shiyu Yang</i>	
Efficient Maximum Closeness Centrality Group Identification.	43
<i>Chen Chen, Wei Wang, and Xiaoyang Wang</i>	
Analyzing EEG Signal Data for Detection of Epileptic Seizure: Introducing Weight on Visibility Graph with Complex Network Feature	56
<i>Supriya, Siuly, Hua Wang, Guangping Zhuo, and Yanchun Zhang</i>	

Spatial Database

Spatial Textual Top-k Search in Mobile Peer-to-Peer Networks.	69
<i>Thao P. Nghiem, Cong Ma, J. Shane Culpepper, and Timos Sellis</i>	
Continuous Maximum Visibility Query for a Moving Target	82
<i>Ch. Md. Rakin Haider, Arif Arman, Mohammed Eunus Ali, and Farhana Murtaza Choudhury</i>	
An Index-Based Method for Efficient Maximizing Range Sum Queries in Road Network	95
<i>Xiaoling Zhou and Wei Wang</i>	
Trip Planning Queries for Subgroups in Spatial Databases	110
<i>Tanzima Hashem, Tahrima Hashem, Mohammed Eunus Ali, Lars Kulik, and Egemen Tanin</i>	
A Sketch-First Approach for Finding TSP	123
<i>Weihuang Huang, Jeffrey Xu Yu, and Zechao Shang</i>	
Finding Least On-Road Travel Time on Road Network	137
<i>Lei Li, Xiaofang Zhou, and Kevin Zheng</i>	

A Weighted K-AP Query Method for RSSI Based Indoor Positioning	150
<i>Huan Huo, Xiufeng Liu, Jifeng Li, Huhu Yang, Dunlu Peng, and Qingkui Chen</i>	
AutoCadroid: An Automated Tool for Building Smartphone Indoor Navigation System from AutoCad Files	164
<i>Shamim Hasnath, Mohammed Eunus Ali, M. Kaysar Abdullah, and S.M. Farhad</i>	
Uncertain Data and Trajectory	
Integration of Probabilistic Information	179
<i>Fereidoon Sadri and Gayatri Tallur</i>	
Top-k Dominance Range-Based Uncertain Queries	191
<i>Ha Thanh Huynh Nguyen and Jinli Cao</i>	
Classification Based on Compressive Multivariate Time Series	204
<i>Chandra Utomo, Xue Li, and Sen Wang</i>	
Adapting ELM to Time Series Classification: A Novel Diversified Top-k Shapelets Extraction Method	215
<i>Qiuyan Yan, Qifa Sun, and Xinming Yan</i>	
Scalable and Fast Top-k Most Similar Trajectories Search Using MapReduce In-Memory	228
<i>Douglas Alves Peixoto and Nguyen Quoc Viet Hung</i>	
Data Mining and Analytics	
Prescriptive Analytics for Big Data	245
<i>Reza Soltanpoor and Timos Sellis</i>	
Topical Event Detection on Twitter	257
<i>Lishan Cui, Xiuzhen Zhang, Xiangmin Zhou, and Flora Salim</i>	
A Time and Opinion Quality-Weighted Model for Aggregating Online Reviews	269
<i>Yassien Shaalan and Xiuzhen Zhang</i>	
An Effective Spatio-Temporal Approach for Predicting Future Semantic Locations	283
<i>Hamidu Abdel-Fatao, Jiuyong Li, Jixue Liu, and Rahman Ashfaqur</i>	
Automatic Labelling of Topics via Analysis of User Summaries	295
<i>Lishan Cui, Xiuzhen Zhang, Amanda Kimpton, and Daryl D'Souza</i>	

Content-Based Top- N Recommendation Using Heterogeneous Relations	308
<i>Yifan Chen, Xiang Zhao, Junjiao Gan, Junkai Ren, and Yanli Hu</i>	
Exploring Data Mining Techniques in Medical Data Streams	321
<i>Le Sun, Jiangang Ma, Yanchun Zhang, and Hua Wang</i>	
Using Detected Visual Objects to Index Video Database	333
<i>Xingzhong Du, Hongzhi Yin, Zi Huang, Yi Yang, and Xiaofang Zhou</i>	
Virtual Samples Construction Using Image-Block-Stretching for Face Recognition	346
<i>Yingnan Zhao, Xiangjian He, and Beijing Chen</i>	

Miscellaneous

XB+-Tree: A Novel Index for PCM/DRAM-Based Hybrid Memory	357
<i>Lu Li, Peiquan Jin, Chengcheng Yang, Shouhong Wan, and Lihua Yue</i>	
Effective Order Preserving Estimation Method	369
<i>Chen Chen, Wei Wang, Xiaoyang Wang, and Shiyu Yang</i>	
Joint Top- k Subscription Query Processing over Microblog Threads	381
<i>Liangjun Song, Zhifeng Bao, Farhana Choudhury, and Timos Sellis</i>	
A Pattern-Based Framework for Addressing Data Representational Inconsistency	395
<i>Bingyu Yi, Wen Hua, and Shazia Sadiq</i>	
Optimising Queue-Based Semi-stream Joins by Introducing a Queue of Frequent Pages	407
<i>M. Asif Naeem, Gerald Weber, and Christof Lutteroth</i>	
Exploiting Hierarchies for Efficient Detection of Completeness in Stream Data	419
<i>Simon Razniewski, Shazia Sadiq, and Xiaofang Zhou</i>	

Demo Papers

Visualization-Aided Exploration of the Real Estate Data	435
<i>Mingzhao Li, Zhifeng Bao, Timos Sellis, and Shi Yan</i>	
A Real Time System for Indoor Shortest Path Query with Indexed Indoor Datasets	440
<i>Zhou Shao, Joon Bum Lee, David Taniar, and Yang Bo</i>	
SKPS: Towards Efficient Processing of Spatial-Keyword Publish/Subscribe System.	444
<i>Xiang Wang, Shiyu Yang, and Ying Zhang</i>	

A Peer to Peer Housing Rental System with Continuous Spatial-Keyword Searching.	448
<i>Yingqian Hou, Xinyin Wang, Liping Wang, and Junjie Yao</i>	
SCHOLAT: An Innovative Academic Information Service Platform	453
<i>Feiyi Tang, Jia Zhu, Chaobo He, Chengzhou Fu, Jing He, and Yong Tang</i>	
Data-Driven Prediction and Visualisation of Dynamic Bushfire Risks	457
<i>Laura Rusu, Hoang Tam Vo, Ziyuan Wang, Mahsa Salehi, and Anna Phan</i>	
EDMS: A System for Efficient Processing Distance-Aware Influence Maximization	462
<i>Xiaoyang Wang, Chen Chen, and Ying Zhang</i>	
Linking News and Tweets	467
<i>Xiaojie Lin, Ye Gu, Rui Zhang, and Ju Fan</i>	
A Data-Packets-Saving Multi-agent Scheduling of Large Online Meetings . . .	471
<i>Usman Ali and Guangyan Huang</i>	
A Classifier Hub for Imbalanced Financial Data	476
<i>Chirath Abeyasinghe, Jianguo Li, and Jing He</i>	
ProvRPQ: An Interactive Tool for Provenance-Aware Regular Path Queries on RDF Graphs	480
<i>Xin Wang and Junhu Wang</i>	
Author Index	485