

*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, Chennai, 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>

Sven Hartmann · Hui Ma  
Abdelkader Hameurlain · Günther Pernul  
Roland R. Wagner (Eds.)

# Database and Expert Systems Applications

29th International Conference, DEXA 2018  
Regensburg, Germany, September 3–6, 2018  
Proceedings, Part I

*Editors*

Sven Hartmann  
Clausthal University of Technology  
Clausthal-Zellerfeld  
Germany

Hui Ma  
Victoria University of Wellington  
Wellington  
New Zealand

Abdelkader Hameurlain  
Paul Sabatier University  
Toulouse  
France

Günther Pernul  
University of Regensburg  
Regensburg  
Germany

Roland R. Wagner  
Johannes Kepler University  
Linz  
Austria

ISSN 0302-9743                      ISSN 1611-3349 (electronic)  
Lecture Notes in Computer Science  
ISBN 978-3-319-98808-5              ISBN 978-3-319-98809-2 (eBook)  
<https://doi.org/10.1007/978-3-319-98809-2>

Library of Congress Control Number: 2018950662

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

© Springer Nature Switzerland AG 2018

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. 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

This volume contains the papers presented at the 29th International Conference on Database and Expert Systems Applications (DEXA 2018), which was held in Regensburg, Germany, during September 3–6, 2018. On behalf of the Program Committee, we commend these papers to you and hope you find them useful.

Database, information, and knowledge systems have always been a core subject of computer science. The ever-increasing need to distribute, exchange, and integrate data, information, and knowledge has added further importance to this subject. Advances in the field will help facilitate new avenues of communication, to proliferate interdisciplinary discovery, and to drive innovation and commercial opportunity.

DEXA is an international conference series that showcases state-of-the-art research activities in database, information, and knowledge systems. The conference and its associated workshops provide a premier annual forum to present original research results and to examine advanced applications in the field. The goal is to bring together developers, scientists, and users to extensively discuss requirements, challenges, and solutions in database, information, and knowledge systems.

DEXA 2018 solicited original contributions dealing with any aspect of database, information, and knowledge systems. Suggested topics included, but were not limited to:

- Acquisition, Modeling, Management, and Processing of Knowledge
- Authenticity, Privacy, Security, and Trust
- Availability, Reliability, and Fault Tolerance
- Big Data Management and Analytics
- Consistency, Integrity, Quality of Data
- Constraint Modeling and Processing
- Cloud Computing and Database-as-a-Service
- Database Federation and Integration, Interoperability, Multi-Databases
- Data and Information Networks
- Data and Information Semantics
- Data Integration, Metadata Management, and Interoperability
- Data Structures and Data Management Algorithms
- Database and Information System Architecture and Performance
- Data Streams and Sensor Data
- Data Warehousing
- Decision Support Systems and Their Applications
- Dependability, Reliability, and Fault Tolerance
- Digital Libraries and Multimedia Databases
- Distributed, Parallel, P2P, Grid, and Cloud Databases
- Graph Databases
- Incomplete and Uncertain Data
- Information Retrieval

- Information and Database Systems and Their Applications
- Mobile, Pervasive, and Ubiquitous Data
- Modeling, Automation, and Optimization of Processes
- NoSQL and NewSQL Databases
- Object, Object-Relational, and Deductive Databases
- Provenance of Data and Information
- Semantic Web and Ontologies
- Social Networks, Social Web, Graph, and Personal Information Management
- Statistical and Scientific Databases
- Temporal, Spatial, and High-Dimensional Databases
- Query Processing and Transaction Management
- User Interfaces to Databases and Information Systems
- Visual Data Analytics, Data Mining, and Knowledge Discovery
- WWW and Databases, Web Services
- Workflow Management and Databases
- XML and Semi-structured Data

Following the call for papers, which yielded 160 submissions, there was a rigorous review process that saw each submission refereed by three to six international experts. The 35 submissions judged best by the Program Committee were accepted as full research papers, yielding an acceptance rate of 22%. A further 40 submissions were accepted as short research papers.

As is the tradition of DEXA, all accepted papers are published by Springer. Authors of selected papers presented at the conference were invited to submit substantially extended versions of their conference papers for publication in the Springer journal *Transactions on Large-Scale Data- and Knowledge-Centered Systems (TLDKS)*. The submitted extended versions underwent a further review process.

The success of DEXA 2018 was the result of collegial teamwork from many individuals. We wish to thank all authors who submitted papers and all conference participants for the fruitful discussions.

We are grateful to Xiaofang Zhou (The University of Queensland) for his keynote talk on “Spatial Trajectory Analytics: Past, Present, and Future” and to Tok Wang Ling (National University of Singapore) for his keynote talk on “Data Models Revisited: Improving the Quality of Database Schema Design, Integration and Keyword Search with ORA-Semantics.”

This edition of DEXA also featured three international workshops covering a variety of specialized topics:

- BDMICS 2018: Third International Workshop on Big Data Management in Cloud Systems
- BIODDD 2018: 9th International Workshop on Biological Knowledge Discovery from Data
- TIR 2018: 15th International Workshop on Technologies for Information Retrieval

We would like to thank the members of the Program Committee and the external reviewers for their timely expertise in carefully reviewing the submissions. We are grateful to our general chairs, Abdelkader Hameurlain, Günther Pernul, and

Roland R. Wagner, to our publication chair, Vladimir Marik, and to our workshop chairs, A Min Tjoa and Roland R. Wagner.

We wish to express our deep appreciation to Gabriela Wagner of the DEXA conference organization office. Without her outstanding work and excellent support, this volume would not have seen the light of day.

Finally, we like to thank Günther Pernul and his team for being our hosts during the wonderful days in Regensburg.

July 2018

Sven Hartmann  
Hui Ma

# Organization

## General Chairs

Abdelkader Hameurlain	IRIT, Paul Sabatier University, Toulouse, France
Günther Pernul	University of Regensburg, Germany
Roland R. Wagner	Johannes Kepler University Linz, Austria

## Program Committee Chairs

Hui Ma	Victoria University of Wellington, New Zealand
Sven Hartmann	Clausthal University of Technology, Germany

## Publication Chair

Vladimir Marik	Czech Technical University, Czech Republic
----------------	--

## Program Committee

Slim Abdennadher	German University, Cairo, Egypt
Hamideh Afsarmanesh	University of Amsterdam, The Netherlands
Riccardo Albertoni	Institute of Applied Mathematics and Information Technologies - Italian National Council of Research, Italy
Idir Amine Amarouche	University Houari Boumediene, Algeria
Rachid Anane	Coventry University, UK
Annalisa Appice	Università degli Studi di Bari, Italy
Mustafa Atay	Winston-Salem State University, USA
Faten Atigui	CNAM, France
Spiridon Bakiras	Hamad bin Khalifa University, Qatar
Zhifeng Bao	National University of Singapore, Singapore
Ladjel Bellatreche	ENSMA, France
Nadia Bennani	INSA Lyon, France
Karim Benouaret	Université Claude Bernard Lyon 1, France
Benslimane Djamal	Lyon 1 University, France
Morad Benyoucef	University of Ottawa, Canada
Catherine Berrut	Grenoble University, France
Athman Bouguettaya	University of Sydney, Australia
Omar Boussaid	University of Lyon/Lyon 2, France
Stephane Bressan	National University of Singapore, Singapore
Barbara Catania	DISI, University of Genoa, Italy
Michelangelo Ceci	University of Bari, Italy
Richard Chbeir	UPPA University, France



Cindy Chen	University of Massachusetts Lowell, USA
Phoebe Chen	La Trobe University, Australia
Max Chevalier	IRIT - SIG, Université de Toulouse, France
Byron Choi	Hong Kong Baptist University, Hong Kong, SAR China
Soon Ae Chun	City University of New York, USA
Deborah Dahl	Conversational Technologies, USA
Jérôme Darmont	Université de Lyon (ERIC Lyon 2), France
Roberto De Virgilio	Università Roma Tre, Italy
Vincenzo Deufemia	Università degli Studi di Salerno, Italy
Gayo Diallo	Bordeaux University, France
Juliette Dibie-Barthélemy	AgroParisTech, France
Dejing Dou	University of Oregon, USA
Cedric du Mouza	CNAM, France
Johann Eder	University of Klagenfurt, Austria
Suzanne Embury	The University of Manchester, UK
Markus Endres	University of Augsburg, Germany
Noura Faci	Lyon 1 University, France
Bettina Fazzinga	ICAR-CNR, Italy
Leonidas Fegaras	The University of Texas at Arlington, USA
Stefano Ferilli	University of Bari, Italy
Flavio Ferrarotti	Software Competence Center Hagenberg, Austria
Vladimir Fomichov	School of Business Informatics, National Research University Higher School of Economics, Moscow, Russian Federation
Flavius Frasincar	Erasmus University Rotterdam, The Netherlands
Bernhard Freudenthaler	Software Competence Center Hagenberg, Austria
Hiroaki Fukuda	Shibaura Institute of Technology, Japan
Steven Furnell	Plymouth University, UK
Joy Garfield	University of Worcester, UK
Claudio Gennaro	ISTI-CNR, Italy
Manolis Gergatsoulis	Ionian University, Greece
Javad Ghofrani	Leibniz Universität Hannover, Germany
Fabio Grandi	University of Bologna, Italy
Carmine Gravino	University of Salerno, Italy
Sven Groppe	Lübeck University, Germany
Jerzy Grzymala-Busse	University of Kansas, USA
Francesco Guerra	Università degli Studi di Modena e Reggio Emilia, Italy
Giovanna Guerrini	University of Genoa, Italy
Allel Hadjali	ENSMA, Poitiers, France
Abdelkader Hameurlain	Paul Sabatier University, France
Ibrahim Hamidah	Universiti Putra Malaysia, Malaysia
Takahiro Hara	Osaka University, Japan
Sven Hartmann	Clausthal University of Technology, Germany
Wynne Hsu	National University of Singapore, Singapore

Yu Hua	Huazhong University of Science and Technology, China
San-Yih Hwang	National Sun Yat-Sen University, Taiwan
Theo Härder	TU Kaiserslautern, Germany
Ionut Emil Iacob	Georgia Southern University, USA
Sergio Ilarri	University of Zaragoza, Spain
Abdessamad Imine	Inria Grand Nancy, France
Yasunori Ishihara	Nanzan University, Japan
Peiquan Jin	University of Science and Technology of China, China
Anne Kao	Boeing, USA
Dimitris Karagiannis	University of Vienna, Austria
Stefan Katzenbeisser	Technische Universität Darmstadt, Germany
Anne Kayem	Hasso Plattner Institute, Germany
Carsten Kleiner	University of Applied Sciences and Arts Hannover, Germany
Henning Koehler	Massey University, New Zealand
Harald Kosch	University of Passau, Germany
Michal Krátký	Technical University of Ostrava, Czech Republic
Petr Kremen	Czech Technical University in Prague, Czech Republic
Sachin Kulkarni	Macquarie Global Services, USA
Josef Küng	University of Linz, Austria
Gianfranco Lamperti	University of Brescia, Italy
Anne Laurent	LIRMM, University of Montpellier 2, France
Lenka Lhotska	Czech Technical University, Czech Republic
Yuchen Li	Singapore Management University, Singapore
Wenxin Liang	Dalian University of Technology, China
Tok Wang Ling	National University of Singapore, Singapore
Sebastian Link	The University of Auckland, New Zealand
Chuan-Ming Liu	National Taipei University of Technology, Taiwan
Hong-Cheu Liu	University of South Australia, Australia
Jorge Lloret Gazo	University of Zaragoza, Spain
Alessandra Lumini	University of Bologna, Italy
Hui Ma	Victoria University of Wellington, New Zealand
Qiang Ma	Kyoto University, Japan
Stephane Maag	TELECOM SudParis, France
Zakaria Maamar	Zayed University, United Arab Emirates
Elio Masciari	ICAR-CNR, Università della Calabria, Italy
Brahim Medjahed	University of Michigan - Dearborn, USA
Harekrishna Mishra	Institute of Rural Management Anand, India
Lars Moench	University of Hagen, Germany
Riad Mokadem	IRIT, Paul Sabatier University, France
Yang-Sae Moon	Kangwon National University, South Korea
Franck Morvan	IRIT, Paul Sabatier University, France
Dariusz Mrozek	Silesian University of Technology, Poland
Francesc Munoz-Escoi	Universitat Politècnica de Valencia, Spain
Ismael Navas-Delgado	University of Málaga, Spain

Wilfred Ng	Hong Kong University of Science and Technology, Hong Kong, SAR China
Javier Nieves Acedo	IK4-Azterlan, Spain
Mourad Oussalah	University of Nantes, France
George Pallis	University of Cyprus, Cyprus
Ingrid Pappel	Tallinn University of Technology, Estonia
Marcin Paprzycki	Polish Academy of Sciences, Warsaw Management Academy, Poland
Oscar Pastor Lopez	Universitat Politecnica de Valencia, Spain
Francesco Piccialli	University of Naples Federico II, Italy
Clara Pizzuti	Institute for High Performance Computing and Networking (ICAR)-National Research Council (CNR), Italy
Pascal Poncelet	LIRMM, France
Elaheh Pourabbas	National Research Council, Italy
Claudia Raibulet	Università degli Studi di Milano-Bicocca, Italy
Praveen Rao	University of Missouri-Kansas City, USA
Rodolfo Resende	Federal University of Minas Gerais, Brazil
Claudia Roncancio	Grenoble University/LIG, France
Massimo Ruffolo	ICAR-CNR, Italy
Simonas Saltenis	Aalborg University, Denmark
N. L. Sarda	I.I.T. Bombay, India
Marinette Savonnet	University of Burgundy, France
Florence Sedes	IRIT, Paul Sabatier University, Toulouse, France
Nazha Selmaoui	University of New Caledonia, New Caledonia
Michael Sheng	Macquarie University, Australia
Patrick Siarry	Université Paris 12 (LiSSi), France
Gheorghe Cosmin Silaghi	Babes-Bolyai University of Cluj-Napoca, Romania
Hala Skaf-Molli	Nantes University, France
Bala Srinivasan	Retried, Monash University, Australia
Umberto Straccia	ISTI - CNR, Italy
Maguelonne Teisseire	Irstea - TETIS, France
Sergio Tessaris	Free University of Bozen-Bolzano, Italy
Olivier Teste	IRIT, University of Toulouse, France
Stephanie Teufel	University of Fribourg, Switzerland
Jukka Teuhola	University of Turku, Finland
Jean-Marc Thevenin	University of Toulouse 1 Capitole, France
A Min Tjoa	Vienna University of Technology, Austria
Vicenc Torra	University of Skövde, Sweden
Traian Marius Truta	Northern Kentucky University, USA
Theodoros Tzouramanis	University of the Aegean, Greece
Lucia Vaira	University of Salento, Italy
Ismini Vasileiou	University of Plymouth, UK
Krishnamurthy Vidyasankar	Memorial University of Newfoundland, Canada
Marco Vieira	University of Coimbra, Portugal
Junhu Wang	Griffith University, Brisbane, Australia

Wendy Hui Wang  
 Piotr Wisniewski  
 Ming Hour Yang  
 Yang, Xiaochun  
 Yanchang Zhao  
 Qiang Zhu  
 Marcin Zimniak  
 Ester Zumpano

Stevens Institute of Technology, USA  
 Nicolaus Copernicus University, Poland  
 Chung Yuan Christian University, Taiwan  
 Northeastern University, China  
 CSIRO, Australia  
 The University of Michigan, USA  
 Leipzig University, Germany  
 University of Calabria, Italy

## Additional Reviewers

Valentyna Tsap  
 Liliana Ibanescu  
 Cyril Labbé  
 Zouhaier Brahmia  
 Dunren Che  
 Feng George Yu  
 Gang Qian  
 Lubomir Stanchev  
 Jorge Martinez-Gil  
 Loredana Caruccio  
 Valentina Indelli Pisano  
 Jorge Bernardino  
 Bruno Cabral  
 Paulo Nunes  
 William Ferng  
 Amin Mesmoudi  
 Sabeur Aridhi

Julius Köpke  
 Marco Franceschetti  
 Meriem Laifa  
 Sheik Mohammad  
 Mostakim Fattah  
 Mohammed Nasser  
 Mohammed Ba-hutair  
 Ali Hamdi Fergani Ali  
 Masoud Salehpour  
 Adnan Mahmood  
 Wei Emma Zhang  
 Zawar Hussain  
 Hui Luo  
 Sheng Wang  
 Lucile Sautot  
 Jacques Fize

Tallinn University of Technology, Estonia  
 AgroParisTech, France  
 Université Grenoble-Alpes, France  
 University of Sfax, Tunisia  
 Southern Illinois University, USA  
 Youngstown State University, USA  
 University of Central Oklahoma, USA  
 Cal Poly, USA  
 Software Competence Center Hagenberg, Austria  
 University of Salerno, Italy  
 University of Salerno, Italy  
 Polytechnic Institute of Coimbra, Portugal  
 University of Coimbra, Portugal  
 Polytechnic Institute of Guarda, Portugal  
 Boeing, USA  
 LIAS/University of Poitiers, France  
 LORIA, University of Lorraine - TELECOM Nancy, France  
 Alpen Adria Universität Klagenfurt, Austria  
 Alpen Adria Universität Klagenfurt, Austria  
 Bordj-Bouarreridj University, Algeria  
 University of Sydney, Australia  
 University of Sydney, Australia  
 University of Sydney, Australia  
 Macquarie University, Australia  
 Macquarie University, Australia  
 Macquarie University, Australia  
 RMIT University, Australia  
 RMIT University, Australia  
 AgroParisTech, France  
 Cirad, Irstea, France

María del Carmen Rodríguez-Hernández	Technological Institute of Aragón, Spain
Ramón Hermoso	University of Zaragoza, Spain
Senen Gonzalez	Software Competence Center Hagenberg, Austria
Ermelinda Oro	High Performance and Computing Institute of the National Research Council (ICAR-CNR), Italy
Shaoyi Yin	Paul Sabatier University, France
Jannai Tokotoko	ISEA University of New Caledonia, New Caledonia
Xiaotian Hao	Hong Kong University of Science and Technology, Hong Kong, SAR China
Ji Cheng	Hong Kong University of Science & Technology, Hong Kong, China
Radim Bača	Technical University of Ostrava, Czech Republic
Petr Lukáš	Technical University of Ostrava, Czech Republic
Peter Chovanec	Technical University of Ostrava, Czech Republic
Galicia Auyon Jorge	ISAE-ENSMA, Poitiers, France
Armando	
Nabila Berkani	ESI, Algiers, Algeria
Amine Roukh	Mostaganem University, Algeria
Chourouk Belheouane	USTHB, Algiers, Algeria
Angelo Impedovo	University of Bari, Italy
Emanuele Pio Barracchia	University of Bari, Italy
Arpita Chatterjee	Georgia Southern University, USA
Stephen Carden	Georgia Southern University, USA
Tharanga Wickramarachchi	U.S. Bank, USA
Divine Wanduku	Georgia Southern University, USA
Lama Saeeda	Czech Technical University in Prague, Czech Republic
Michal Med	Czech Technical University in Prague, Czech Republic
Franck Ravat	Université Toulouse 1 Capitole - IRIT, France
Julien Aligon	Université Toulouse 1 Capitole - IRIT, France
Matthew Damigos	Ionian University, Greece
Eleftherios Kalogeros	Ionian University, Greece
Srini Bhagavan	IBM, USA
Monica Senapati	University of Missouri-Kansas City, USA
Khulud Alsultan	University of Missouri-Kansas City, USA
Anas Katib	University of Missouri-Kansas City, USA
Jose Alvarez	Telecom SudParis, France
Sarah Dahab	Telecom SudParis, France
Dietrich Steinmetz	Clausthal University of Technology, Germany

## **Abstracts of Keynote Speakers**

# Data Models Revisited: Improving the Quality of Database Schema Design, Integration and Keyword Search with ORA-Semantics (Extended Abstract)

Tok Wang Ling<sup>1</sup>, Mong Li Lee<sup>1</sup>, Thuy Ngoc Le<sup>2</sup>, and Zhong Zeng<sup>3</sup>

<sup>1</sup> Department of Computer Science, School of Computing, National University of Singapore

{lingtw, leeml}@comp.nus.edu.sg

<sup>2</sup> Google Singapore

le.thuy.ngoc@gmail.com

<sup>3</sup> Data Center Technology Lab, Huawei

zengzhong4@huawei.com

## Introduction

Object class, relationship type, and attribute of object class and relationship type, are three basic concepts in the Entity Relationship Model. We call them ORA-semantics. In this talk, we highlight the limitations of the common database models such as the relational and XML data model. One serious common limitation of these database models is their inability to capture and explicitly represent object classes and relationship types together with their attributes in their schema languages. In fact, these data models have no concepts of object class, relationship type, and their attribute.

Without using ORA-semantics in databases, the quality of important database tasks such as relational and XML database schema design, data and schema integration, and relational and XML keyword query processing are low, and serious problems may arise. We show the reasons that lead to these problems, and demonstrate how ORA-semantics can be used to improve the result quality of these database tasks significantly.

## Limitations of Relational Model

In the *relational model*, functional dependencies (FDs) and multivalued dependencies (MVDs) are integrity constraints; many of which are artificially imposed by organization or database designers. These constraints have no semantics, and cannot be automatically discovered by data mining techniques.

FDs and MVDs are used to remove redundancy and obtain normal form relations in database schema design. During normalization, we must cover the given set of FDs

(i.e., the closure of the set of FDs remain unchanged), and we want to remove all MVDs. However, MVDs are relation sensitive, and it is very difficult to detect them. The existence of MVDs in a relation is because some unrelated multivalued attributes (of an object class or a relationship type) are wrongly grouped in the relation [10]. Key in relation is not the same as OID of object class. There is no concept of ORA-semantics in the relational model.

## ORA Semantics in Database Schema Design

There are three common approaches for relational database schema design:

- a. **Decomposition.** This approach is based on the Universal Relation Assumption (URA) that a database can be represented by a universal relation which contains all the attributes of the database and this relation is then decomposed into smaller relations in some good normal forms such as 3NF, BCNF, 4NF, etc. in order to remove redundant data using the given FDs and MVDs. The process is non-deterministic, and the relations obtained depend on the order of FDs and MVDs chosen for decomposition, which may not cover the given set of FDs.
- b. **Synthesis [1].** This approach is based on the assumption that a database can be described by a given set of attributes and a given set of functional dependencies. It also assumes URA, and a set of 3NF and BCNF relations is synthesized based on the given set of dependencies. The process is non-deterministic, and depends on the order of the redundant FDs found to generate 3NF relations. It does not consider MVDs and does not guarantee reconstructibility.
- c. **ER Approach.** An ER diagram (ERD) is first constructed based on the database specification and requirements, and then normalized to a normal form ERD. The normal form ERD is then translated to a set of normal form relations together with a set of additional constraints that exist in the ERD but cannot be represented in the relational schema [11]. Multivalued attributes of object classes and relationships will be in separated relations. Users do not need to consider MVDs which are relation sensitive. ERD can use relaxed URA, i.e. only object identifier names must be unique, which is much more convenient than using URA.

Both the decomposition and synthesis approaches cannot handle complex relationship types such as recursive relationship type, ISA relationship, and multiple relationship types defined among 2 or more object classes. They also do not have the concept of ORA-semantics and have many problems and short comings. Other problems and issues that arise when using decomposition and synthesis methods to design a database include

- (i) How to find a given set of FDs in a relational database? Can we use some data mining techniques to find FDs and MVDs in a relational database?
- (ii) If a relation is not in BCNF, can we always normalize it to a set of BCNF relations?



- (iii) If a relation is not in 4NF, is there a non-loss decomposition of the relation into a set of 4NF relations which cover all the given FDs?
- (iv) 3NF and BCNF relations are defined on individual relations, rather than on the whole database. Hence, they cannot detect redundancy among relations of the database and may contain global redundant attributes [13].

In contrast, the ER approach captures ORA-semantics and avoids the problems of the decomposition method and synthesis method.

## ORA Semantics in Data and Schema Integration

In data and schema integration, entity resolution (or object identification) is widely studied. However, this problem is still not well solved and cannot be handled fully automatically, e.g., we cannot automatically identify authors of papers completely in DBLP.

Besides entity resolution, we need to consider relationship resolution which aims to identify different relationship types between/among same object classes. We also need to differentiate between primary key vs object identifier (OID), local OID vs global OID, system generated OID vs manually designed OID, local FD vs global FD, semantic dependency vs FD/MVD constraint, structural conflicts [9], as well as schematic discrepancy [3] among schemas. All these concepts are related to ORA-semantics and they have a big impact on the quality of the integrated database and schema.

The challenge to achieve a good quality integration remains. Since the ER model can capture ORA-semantics, it is more promising to use the ER approach for data and schema integration.

## ORA Semantics in Relational Keyword Search

Methods for *relational keyword search* [4, 5] can be broadly classified into two categories: data graph approach and schema graph approach. In the data graph approach, the relational database is modeled as a graph where each node represents a tuple and each edge represents a foreign key-key reference. An answer to a keyword query is typically defined as a minimal connected subgraph which contains all the keywords. This graph search is equivalent to the Steiner tree problem, which is NP-complete.

In schema graph approach, the database schema is modeled as a schema graph where each node represents a relation and each edge represents a foreign key-key constraint between two relations. Based on the schema graph, a keyword query is translated into a set of SQL statements that join the relations with tuples matching the keywords.

We identify the serious limitations of existing relational keyword search, which include incomplete answers, meaningless answers, inconsistent answers, and user difficulty in understanding the answers when they are represented as Steiner trees, etc.

In addition, the answers returned depend on the normal form of the relational database, i.e., database schema dependence. We can improve the correctness and completeness of relational keyword search by exploiting ORA-semantics because these semantics enable us to detect duplication of objects and relationships and address the above mentioned limitations [16].

We extend keyword queries by allowing keywords that match the metadata, i.e., relation name and attribute name. We also extend keyword queries with group-by and aggregate functions including sum, max, min, avg, count, etc. In order to process these extended keyword queries correctly, we use ORA-semantics to detect duplication of objects and relationships. Without using ORA-semantics, the results of aggregate functions may be computed wrongly. For more details, see [15, 17].

## Limitations of XML Data Model

The *XML data model* also cannot capture ORA-Semantics [2, 12]. The constraints on the structure and content of XML can be described by DTD or XML Schema. The ID in DTD is not the same as the object identifier, ID attribute is OID of the object class, but OID of an object class may not be able to declare as ID, and a multivalued attribute of object class cannot be represented directly as an attribute in DTD/XML Schema. IDREF is not the same as foreign key to key reference in RDB. IDREF has no type.

DTD/XML Schema can only represent the hierarchical structures with simple constraints; they have no concept on ORA-semantics. The parent-child relationship in XML may not represent relationship type; relationship type (especially n-ary) is not explicitly captured in DTD/XML Schema. They cannot distinguish between attribute of object class vs attribute of relationship type.

## ORA Semantics in XML Keyword Search

Existing approaches to *XML keyword search* are structure-based because they mainly rely on the exploration of the structure of XML data. These approaches can be classified as tree-based and graph-based search. Tree-based search is used when an XML document is modeled as a tree, i.e. without ID references (IDREFs), while graph-based search is used for XML documents with IDREFs.

Almost all tree-based approaches are based on some variations of LCA (Least Common Ancestor) semantics such as SLCA, MLCA, VLCA, and ELCA [14]. Given the lack of awareness of semantics in XML data, LCA-based methods do not exploit hidden ORA-semantics in data-centric XML document. This causes serious problems in processing LCA-based XML keyword queries, such as returning meaningless answers, duplicated answers, incomplete answers, missing answers, and inconsistent answers.

We can use ORA-semantics to improve the correctness and completeness of XML keyword search by detecting duplication of objects and relationships. We introduce the

concepts of object tree, reversed object tree, and relative of objects to address the above mentioned problems of XML keyword search [6, 8]. We also extend XML keyword queries by considering keywords that match the metadata, i.e., tag names of XML data, and with group-by and aggregate functions [7].

## Conclusion

In summary, the schemas of relational model and XML data model cannot capture the ORA-semantics which exist in the ER model. We highlight the serious problems on the quality of some database tasks due to the lack of knowledge on ORA-semantics in the relational model and XML data model. However, programmers must know the ORA-semantics of the database in order to write SQL and XQuery programs correctly. ORA-SS data model [2, 12] is designed to capture ORA-semantics in XML data.

We conclude this talk with suggestions on further research on data and schema integration, keyword query search in relational databases and XML databases such as data model independent keyword query search, and the use of ORA-semantics in NoSQL and big data applications.

## References

1. Bernstein, P.A.: Synthesizing third normal form relations from functional dependencies. *Trans. Database Syst.* (1976)
2. Dobbie, G., Wu, X., Ling, T.W., Lee, M.L.: Ora-ss: an object-relationship-attribute model for semistructured data. Technical report, National University of Singapore (2000)
3. He, Q., Ling, T.W.: Extending and inferring functional dependencies in schema transformation. In: *ACM CIKM* (2004)
4. Hristidis, V., Papakonstantinou, Y.: Discover: keyword search in relational databases. In: *VLDB* (2002)
5. Hulgeri, A., Nakhe, C.: Keyword searching and browsing in databases using banks. In: *IEEE ICDE* (2002)
6. Le, T.N., Bao, Z., Ling, T.W.: Schema-independence in xml keyword search. In: Yu, E., Dobbie, G., Jarke, M., Purao, S. (eds.) *ER 2014*. LNCS, vol. 8824. Springer, Cham (2014)
7. Le, T.N., Bao, Z., Ling, T.W., Dobbie, G.: Group-by and aggregate functions in XML keyword search. In: *DEXA* (2014)
8. Le, T.N., Wu, H., Ling, T.W., Li, L., Lu, J.: From structure-based to semantics-based: towards effective XML keyword search. In: Ng, W., Storey, V.C., Trujillo J.C. (eds.) *ER 2013*. LNCS, vol. 8217. Springer, Heidelberg (2013)
9. Lee, M.L., Ling, T.W.: Resolving structural conicts in the integration of entity relationship schemas. In: *ER* (1995)
10. Ling, T.W.: An analysis of multivalued and join dependencies based on the entity-relationship approach. *Data Knowl. Eng.* (1985)
11. Ling, T.W.: A normal form for entity-relationship diagrams. In: *ER* (1985)
12. Ling, T.W., Lee, M.L., Dobbie, G.: *Semistructured Database Design*. Springer, New York (2005)

13. Ling, T.W., Tompa, F.W., Kameda, T.: An improved third normal form for relational databases. *Trans. Database Syst.* (1981)
14. Xu, Y., Papakonstantinou, Y.: Efficient keyword search for smallest LCAs in XML databases. In: *ACM SIGMOD* (2005)
15. Zeng, Z., Bao, Z., Le, T.N., Lee, M.L., Ling, T.W.: Expressq: identifying keyword context and search target in relational keyword queries. In: *ACM CIKM* (2014)
16. Zeng, Z., Bao, Z., Lee, M.L., Ling, T.W.: A semantic approach to keyword search over relational databases. In: Ng, W., Storey, V.C., Trujillo, J.C. (eds.) *ER 2013. LNCS*, vol. 8217. Springer, Heidelberg (2013)
17. Zeng, Z., Lee, M.L., Ling, T.W.: Answering keyword queries involving aggregates and group by on relational databases. In: *EDBT* (2016)

# **Spatial Trajectory Analytics: Past, Present and Future (Extended Abstract)**

Xiaofang Zhou

School of Information Technology and Electrical Engineering,  
The University of Queensland, Australia  
zxf@itee.uq.edu.au

Trajectory computing involves a wide range of research topics centered around spatiotemporal data, including data management, query processing, data mining and recommendation systems, and more recently, data privacy and machine learning. It can find many applications in intelligent transport systems, location-based systems, urban planning and smart city.

Spatial trajectory computing research has attracted an extensive amount of effort from researchers in database and data mining communities. In 2011 we edited a book to introduce the basic concepts and main research topics and progresses at that time in spatial trajectory computing [6]. This area has been developed at a very rapid and still accelerating speed, driven by the availability of massive volumes of both historical and real-time streaming trajectory data from many sources such as GPS devices, smart phones and social media applications. Major businesses also start to treat spatial trajectory data as enterprise data to support all business units that require location and movement intelligence. Trajectory data have now been embedded into traffic navigation and car sharing services, mobile apps and online social network applications, leading to more sophisticated time-dependent queries [3] and millions of concurrent queries that have not been considered in previous spatial query processing research. New computing platforms and new computational and analytics tools such as machine learning [4] have also contributed the current surge of research effort in this area. As trajectory data can reveal highly unique information about individuals [1], there are new research opportunities to address the both sides of the problem: to protect user's location and movement privacy and to link users from different trajectory datasets.

There are strong industry demands to manage and process extremely large amount of trajectory data for a diversified range of applications. Our community has developed a quite comprehensive spectrum of solutions in the past to address different aspects of trajectory analytics problems. There is an urgent need now to develop flexible and powerful trajectory data management systems with proper support from data acquisition, management to analytics. Such a system should cater for the hierarchical nature of spatial data [2] such that analytics can be applied at the right level to generate meaningful results (for example, trajectory similarity analysis can only be done using calibrated data [5]). This is the future direction of spatial trajectory computing research.

## References

1. de Montjoye, Y.-A., Hidalgo, C.A., Verleysen, M., Blondel, V.D.: Unique in the crowd: the privacy bounds of human mobility. *Sci. Rep.* **3**, 1376 (2013). EP 03
2. Kuipers, B.: The spatial semantic hierarchy. *Artif. Intell.* **119**(1–2), 191–233 (2000)
3. Li, L., Hua, W., Du, X., Zhou, X.: Minimal on-road time route scheduling on time-dependent graphs. *PVLDB* **10**(11), 1274–1285 (2017)
4. Lv, Z., Xu, J., Zheng, K., Yin, H., Zhao, P., Zhou, X.: LC-RNN: a deep learning model for traffic speed prediction. In: *IJCAI* (2018)
5. Su, H., Zheng, K., Huang, J., Wang, H., Zhou, X.: Calibrating trajectory data for spatio-temporal similarity analysis. *VLDB J.* **24**(1), 93–116 (2015)
6. Zheng, Y., Zhou, X.: *Computing with Spatial Trajectories*. Springer, New York (2011)

# Contents – Part I

## Big Data Analytics

Scalable Vertical Mining for Big Data Analytics of Frequent Itemsets . . . . .	3
<i>Carson K. Leung, Hao Zhang, Joglas Souza, and Wookey Lee</i>	
ScaleSCAN: Scalable Density-Based Graph Clustering . . . . .	18
<i>Hiroaki Shiokawa, Tomokatsu Takahashi, and Hiroyuki Kitagawa</i>	
Sequence-Based Approaches to Course Recommender Systems. . . . .	35
<i>Ren Wang and Osmar R. Zaiane</i>	

## Data Integrity and Privacy

BFASTDC: A Bitwise Algorithm for Mining Denial Constraints. . . . .	53
<i>Eduardo H. M. Pena and Eduardo Cunha de Almeida</i>	
BOUNCER: Privacy-Aware Query Processing over Federations of RDF Datasets . . . . .	69
<i>Kemele M. Endris, Zuhair Almhithawi, Ioanna Lytra, Maria-Esther Vidal, and Sören Auer</i>	
Minimising Information Loss on Anonymised High Dimensional Data with Greedy In-Memory Processing. . . . .	85
<i>Nikolai J. Podlesny, Anne V. D. M. Kayem, Stephan von Schorlemer, and Matthias Uflacker</i>	

## Decision Support Systems

A Diversification-Aware Itemset Placement Framework for Long-Term Sustainability of Retail Businesses. . . . .	103
<i>Parul Chaudhary, Anirban Mondal, and Polepalli Krishna Reddy</i>	
Global Analysis of Factors by Considering Trends to Investment Support . . .	119
<i>Makoto Kirihata and Qiang Ma</i>	
Efficient Aggregation Query Processing for Large-Scale Multidimensional Data by Combining RDB and KVS. . . . .	134
<i>Yuya Watari, Atsushi Keyaki, Jun Miyazaki, and Masahide Nakamura</i>	

**Data Semantics**

Learning Interpretable Entity Representation in Linked Data . . . . .	153
<i>Takahiro Komamizu</i>	
GARUM: A Semantic Similarity Measure Based on Machine Learning and Entity Characteristics . . . . .	169
<i>Ignacio Traverso-Ribón and Maria-Esther Vidal</i>	
Knowledge Graphs for Semantically Integrating Cyber-Physical Systems . . . .	184
<i>Irlán Grangel-González, Lavdim Halilaj, Maria-Esther Vidal, Omar Rana, Steffen Lohmann, Sören Auer, and Andreas W. Müller</i>	

**Cloud Data Processing**

Efficient Top- $k$ Cloud Services Query Processing Using Trust and QoS . . . .	203
<i>Karim Benouaret, Idir Benouaret, Mahmoud Barhamgi, and Djamel Benslimane</i>	
Answering Top- $k$ Queries over Outsourced Sensitive Data in the Cloud . . . .	218
<i>Sakina Mahboubi, Reza Akbarinia, and Patrick Valduriez</i>	
$R^2$ -Tree: An Efficient Indexing Scheme for Server-Centric Data Center Networks . . . . .	232
<i>Yin Lin, Xinyi Chen, Xiaofeng Gao, Bin Yao, and Guihai Chen</i>	

**Time Series Data**

Monitoring Range Motif on Streaming Time-Series . . . . .	251
<i>Shinya Kato, Daichi Amagata, Shunya Nishio, and Takahiro Hara</i>	
MTSC: An Effective Multiple Time Series Compressing Approach . . . . .	267
<i>Ningting Pan, Peng Wang, Jiaye Wu, and Wei Wang</i>	
DANCINGLINES: An Analytical Scheme to Depict Cross-Platform Event Popularity . . . . .	283
<i>Tianxiang Gao, Weiming Bao, Jinning Li, Xiaofeng Gao, Boyuan Kong, Yan Tang, Guihai Chen, and Xuan Li</i>	

**Social Networks**

Community Structure Based Shortest Path Finding for Social Networks . . . .	303
<i>Yale Chai, Chunyao Song, Peng Nie, Xiaojie Yuan, and Yao Ge</i>	



On Link Stability Detection for Online Social Networks . . . . .	320
<i>Ji Zhang, Xiaohui Tao, Leonard Tan, Jerry Chun-Wei Lin, Hongzhou Li, and Liang Chang</i>	

EPOC: A Survival Perspective Early Pattern Detection Model for Outbreak Cascades . . . . .	336
<i>Chaoqi Yang, Qitian Wu, Xiaofeng Gao, and Guihai Chen</i>	

## **Temporal and Spatial Databases**

Analyzing Temporal Keyword Queries for Interactive Search over Temporal Databases . . . . .	355
<i>Qiao Gao, Mong Li Lee, Tok Wang Ling, Gillian Dobbie, and Zhong Zeng</i>	

Implicit Representation of Bigrular Rules for Multigranular Data . . . . .	372
<i>Stephen J. Hegner and M. Andrea Rodríguez</i>	

QDR-Tree: An Efficient Index Scheme for Complex Spatial Keyword Query . . . . .	390
<i>Xinshi Zang, Peiwen Hao, Xiaofeng Gao, Bin Yao, and Guihai Chen</i>	

## **Graph Data and Road Networks**

Approximating Diversified Top- $k$ Graph Pattern Matching . . . . .	407
<i>Xin Wang and Huayi Zhan</i>	

Boosting PageRank Scores by Optimizing Internal Link Structure . . . . .	424
<i>Naoto Ohsaka, Tomohiro Sonobe, Naonori Kakimura, Takuro Fukunaga, Sumio Fujita, and Ken-ichi Kawarabayashi</i>	

Finding the Most Navigable Path in Road Networks: A Summary of Results. . . . .	440
<i>Ramneek Kaur, Vikram Goyal, and Venkata M. V. Gunturi</i>	

Load Balancing in Network Voronoi Diagrams Under Overload Penalties . . .	457
<i>Ankita Mehta, Kapish Malik, Venkata M. V. Gunturi, Anurag Goel, Pooja Sethia, and Aditi Aggarwal</i>	

<b>Author Index</b> . . . . .	477
-------------------------------	-----

## Contents – Part II

### Information Retrieval

Template Trees: Extracting Actionable Information from Machine Generated Emails . . . . .	3
<i>Manoj K. Agarwal and Jitendra Singh</i>	
Parameter Free Mixed-Type Density-Based Clustering . . . . .	19
<i>Sahar Behzadi, Mahmoud Abdelmottaleb Ibrahim, and Claudia Plant</i>	
CROP: An Efficient Cross-Platform Event Popularity Prediction Model for Online Media . . . . .	35
<i>Mingding Liao, Xiaofeng Gao, Xuezheng Peng, and Guihai Chen</i>	
Probabilistic Classification of Skeleton Sequences . . . . .	50
<i>Jan Sedmidubsky and Pavel Zezula</i>	

### Uncertain Information

A Fuzzy Unified Framework for Imprecise Knowledge . . . . .	69
<i>Soumaya Moussa and Saoussen Bel Hadj Kacem</i>	
Frequent Itemset Mining on Correlated Probabilistic Databases . . . . .	84
<i>Yasemin Asan Kalaz and Rajeev Raman</i>	
Leveraging Data Relationships to Resolve Conflicts from Disparate Data Sources . . . . .	99
<i>Romila Pradhan, Walid G. Aref, and Sunil Prabhakar</i>	

### Data Warehouses and Recommender Systems

Direct Conversion of Early Information to Multi-dimensional Model . . . . .	119
<i>Deepika Prakash</i>	
OLAP Queries Context-Aware Recommender System . . . . .	127
<i>Elsa Negre, Franck Ravat, and Olivier Teste</i>	
Combining Web and Enterprise Data for Lightweight Data Mart Construction . . . . .	138
<i>Suzanne McCarthy, Andrew McCarren, and Mark Roantree</i>	

FairGRecs: Fair Group Recommendations by Exploiting Personal Health Information . . . . .	147
<i>Maria Stratigi, Haridimos Kondylakis, and Kostas Stefanidis</i>	

### Data Streams

Big Log Data Stream Processing: Adapting an Anomaly Detection Technique . . . . .	159
<i>Marietheres Dietz and Günther Pernul</i>	

Information Filtering Method for Twitter Streaming Data Using Human-in-the-Loop Machine Learning. . . . .	167
<i>Yu Suzuki and Satoshi Nakamura</i>	

Parallel $n$ -of- $N$ Skyline Queries over Uncertain Data Streams . . . . .	176
<i>Jun Liu, Xiaoyong Li, Kaijun Ren, Junqiang Song, and Zongshuo Zhang</i>	

A Recommender System with Advanced Time Series Medical Data Analysis for Diabetes Patients in a Telehealth Environment . . . . .	185
<i>Raid Lafta, Ji Zhang, Xiaohui Tao, Jerry Chun-Wei Lin, Fulong Chen, Yonglong Luo, and Xiaoyao Zheng</i>	

### Information Networks and Algorithms

Edit Distance Based Similarity Search of Heterogeneous Information Networks . . . . .	195
<i>Jianhua Lu, Ningyun Lu, Sipei Ma, and Baili Zhang</i>	

An Approximate Nearest Neighbor Search Algorithm Using Distance-Based Hashing. . . . .	203
<i>Yuri Itotani, Shin'ichi Wakabayashi, Shinobu Nagayama, and Masato Inagi</i>	

Approximate Set Similarity Join Using Many-Core Processors . . . . .	214
<i>Kenta Sugano, Toshiyuki Amagasa, and Hiroyuki Kitagawa</i>	

Mining Graph Pattern Association Rules . . . . .	223
<i>Xin Wang and Yang Xu</i>	

### Database System Architecture and Performance

Cost Effective Load-Balancing Approach for Range-Partitioned Main-Memory Resident Data . . . . .	239
<i>Djahida Belayadi, Khaled-Walid Hidouci, Ladjel Bellatreche, and Carlos Ordonez</i>	

Adaptive Workload-Based Partitioning and Replication for RDF Graphs . . . .	250
<i>Ahmed Al-Ghezi and Lena Wiese</i>	

<i>QUIOW</i> : A Keyword-Based Query Processing Tool for RDF Datasets and Relational Databases . . . . .	259
<i>Yenier T. Izquierdo, Grettel M. García, Elisa S. Menendez, Marco A. Casanova, Frederic Dartayre, and Carlos H. Levy</i>	

An Abstract Machine for Push Bottom-Up Evaluation of Datalog . . . . .	270
<i>Stefan Brass and Mario Wenzel</i>	

## Novel Database Solutions

What Lies Beyond Structured Data? A Comparison Study for Metric Data Storage. . . . .	283
<i>Pedro H. B. Siqueira, Paulo H. Oliveira, Marcos V. N. Bedo, and Daniel S. Kaster</i>	

A Native Operator for Process Discovery . . . . .	292
<i>Alifah Syamsiyah, Boudewijn F. van Dongen, and Remco M. Dijkman</i>	

Implementation of the Aggregated R-Tree for Phase Change Memory . . . . .	301
<i>Maciej Jurga and Wojciech Macyna</i>	

Modeling Query Energy Costs in Analytical Database Systems with Processor Speed Scaling. . . . .	310
<i>Boming Luo, Yuto Hayamizu, Kazuo Goda, and Masaru Kitsuregawa</i>	

## Graph Querying and Databases

Sprouter: Dynamic Graph Processing over Data Streams at Scale . . . . .	321
<i>Tariq Abughofa and Farhana Zulkernine</i>	

A Hybrid Approach of Subgraph Isomorphism and Graph Simulation for Graph Pattern Matching . . . . .	329
<i>Kazunori Sugawara and Nobutaka Suzuki</i>	

Time Complexity and Parallel Speedup of Relational Queries to Solve Graph Problems . . . . .	339
<i>Carlos Ordóñez and Predrag T. Tosić</i>	

Using Functional Dependencies in Conversion of Relational Databases to Graph Databases . . . . .	350
<i>Younna A. Megid, Neamat El-Tazi, and Aly Fahmy</i>	

## Learning

A Two-Level Attentive Pooling Based Hybrid Network for Question Answer Matching Task . . . . .	361
<i>Zhenhua Huang, Guangxu Shan, Jiuju Cheng, and Juan Ni</i>	
Features' Associations in Fuzzy Ensemble Classifiers . . . . .	369
<i>Ilef Ben Slima and Amel Borgi</i>	
Learning Ranking Functions by Genetic Programming Revisited. . . . .	378
<i>Ricardo Baeza-Yates, Alfredo Cuzzocrea, Domenico Crea, and Giovanni Lo Bianco</i>	
A Comparative Study of Synthetic Dataset Generation Techniques . . . . .	387
<i>Ashish Dandekar, Remmy A. M. Zen, and Stéphane Bressan</i>	

## Emerging Applications

The Impact of Rainfall and Temperature on Peak and Off-Peak Urban Traffic . . . . .	399
<i>Aniekan Essien, Ilias Petrounias, Pedro Sampaio, and Sandra Sampaio</i>	
Fast Identification of Interesting Spatial Regions with Applications in Human Development Research . . . . .	408
<i>Carl Duffy, Deepak P., Cheng Long, M. Satish Kumar, Amit Thorat, and Amaresh Dubey</i>	
Creating Time Series-Based Metadata for Semantic IoT Web Services. . . . .	417
<i>Kasper Apajalahti</i>	
Topic Detection with Danmaku: A Time-Sync Joint NMF Approach. . . . .	428
<i>Qingchun Bai, Qinmin Hu, Faming Fang, and Liang He</i>	

## Data Mining

Combine Value Clustering and Weighted Value Coupling Learning for Outlier Detection in Categorical Data . . . . .	439
<i>Hongzuo Xu, Yongjun Wang, Zhiyue Wu, Xingkong Ma, and Zhiquan Qin</i>	
Mining Local High Utility Itemsets . . . . .	450
<i>Philippe Fournier-Viger, Yimin Zhang, Jerry Chun-Wei Lin, Hamido Fujita, and Yun Sing Koh</i>	
Mining Trending High Utility Itemsets from Temporal Transaction Databases. . . . .	461
<i>Acquah Hackman, Yu Huang, and Vincent S. Tseng</i>	

Social Media vs. News Media: Analyzing Real-World Events from Different Perspectives . . . . .	471
<i>Liqiang Wang, Ziyu Guo, Yafang Wang, Zeyuan Cui, Shijun Liu, and Gerard de Melo</i>	
<b>Privacy</b>	
Differential Privacy for Regularised Linear Regression. . . . .	483
<i>Ashish Dandekar, Debabrota Basu, and Stéphane Bressan</i>	
A Metaheuristic Algorithm for Hiding Sensitive Itemsets . . . . .	492
<i>Jerry Chun-Wei Lin, Yuyu Zhang, Philippe Fournier-Viger, Youcef Djenouri, and Ji Zhang</i>	
<b>Text Processing</b>	
Constructing Multiple Domain Taxonomy for Text Processing Tasks. . . . .	501
<i>Yihong Zhang, Yongrui Qin, and Longkun Guo</i>	
Combining Bilingual Lexicons Extracted from Comparable Corpora: The Complementary Approach Between Word Embedding and Text Mining . . . . .	510
<i>Sourour Belhaj Rhouma, Chiraz Latiri, and Catherine Berrut</i>	
<b>Author Index</b> . . . . .	519