

Lecture Notes
in Business Information Processing

138

Series Editors

Wil van der Aalst

Eindhoven Technical University, The Netherlands

John Mylopoulos

University of Trento, Italy

Michael Rosemann

Queensland University of Technology, Brisbane, Qld, Australia

Michael J. Shaw

University of Illinois, Urbana-Champaign, IL, USA

Clemens Szyperski

Microsoft Research, Redmond, WA, USA

Marie-Aude Aufaure
Esteban Zimányi (Eds.)

Business Intelligence

Second European Summer School, eBISS 2012
Brussels, Belgium, July 15-21, 2012
Tutorial Lectures



Springer

Volume Editors

Marie-Aude Aufaure
Ecole Centrale Paris
MAS Laboratory
Châtenay-Malabry, France
E-mail: marie-aude.aufaure@ecp.fr

Esteban Zimányi
Université Libre de Bruxelles
Department of Computer and Decision Engineering (CoDE)
Brussels, Belgium
E-mail: ezimanyi@ulb.ac.be

ISSN 1865-1348 e-ISSN 1865-1356
ISBN 978-3-642-36317-7 e-ISBN 978-3-642-36318-4
DOI 10.1007/978-3-642-36318-4
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012956061

ACM Computing Classification (1998): J.1, H.2, H.3, D.1, G.3, G.2

© Springer-Verlag Berlin Heidelberg 2013

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, 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.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The Second European Business Intelligence Summer School (eBISS 2012) took place in Brussels, Belgium, in July 2012. Tutorials were given by experts in business intelligence and covered various hot topics. This volume contains the lecture notes of the summer school.

The first chapter shows how to support multidimensional business intelligence for complex application domains such as medical data, location-based services, music data, web data, and text data. In particular, complex dimension hierarchies, complex measures, and integration of multidimensional data with complex external data are detailed.

The second chapter introduces business process modeling. Business process management can be seen as an extension of classical workflow management systems. This chapter discusses the similarities and differences between these two notions, and then introduces the current OMG standard BPMN 2.0 (Business Process Modeling and Notation) and discusses BPEL (Business Process Execution Language). Finally, hot research topics in this field such as business process mining are depicted.

The third chapter provides an overview of machine learning techniques in time series forecasting, and outlines the challenging issues related to predicting the future in applied sciences. Three aspects are discussed in this chapter: the formalization of one-step forecasting problems as supervised learning tasks, the discussion of local learning techniques as an effective tool for dealing with temporal data, and the role of the forecasting strategy when we move from one-step to multiple-step forecasting.

The fourth chapter gives an overview of Markov logic networks from a theoretical and a practical viewpoint. Statistical relational learning approaches, more specifically Bayesian logic networks, are then detailed. Inferencing and learning processes are then explained together with the best scaling algorithms known today. This chapter concludes with an overview of application areas.

The fifth chapter focuses on recent developments, challenges, and potential solutions for mining large graphs. The main challenge of new tools and frameworks lies in the development of new paradigms that are scalable, efficient, and flexible. Distributed computing is depicted through the MapReduce paradigm. Finally, a new field of research, graph data warehousing, which is deeply linked with large graph mining, is introduced.

The sixth chapter addresses the challenges induced by big data analytics on modern architectures. Massively parallel analysis systems and their programming models are discussed, as well as the application of these modern architectures on database processing.

The seventh chapter introduces decision aid, and more specifically multicriteria decision aid, with a focus on two methods: PROMETHEE and GAIA. An illustrative example, highlighting the added value of using interactive and visual tools in complex decision processes, is analyzed with the D-Sight software.

The eighth chapter explores the importance of semantic technologies (ontologies) and knowledge extraction techniques for knowledge management, search, and capture in e-business processes. Semantic technologies and ontology learning from web data are detailed, and the use of ontologies for business intelligence is discussed through use cases described in several fields.

Finally, the ninth chapter presents the Business Semantics Management (BSM) method, a fact-oriented approach to knowledge modeling grounded in natural language. This method constitutes an interface between Enterprise Information Management and the Web of Data. BSM was implemented in the Flemish Public Administration for building the Flanders Research Information Space (FRIS) program.

We would like to thank the attendants of the summer school for their active participation, as well as the speakers and their co-authors for the high quality of their contributions in a constantly evolving and highly competitive domain. Finally, the lectures in this volume greatly benefited from the comments of the external reviewers.

November 2012

Marie-Aude Aufaure
Esteban Zimányi
eBISS 2012 Co-chairs

Organization

The Second European Business Intelligence Summer School (eBISS 2012) was organized by the MAS Laboratory of the Ecole Centrale de Paris and the Department of Computer and Decision Engineering (CoDE) of the Université Libre de Bruxelles.

Program Committee

Alberto Abelló	Universitat Politècnica de Catalunya, Spain
Marie-Aude Aufaure	Ecole Centrale de Paris, France
Patrick Marcel	Université François Rabelais de Tours, France
Alexander Löser	Technische Universität Berlin, Germany
Esteban Zimányi	Université Libre de Bruxelles, Belgium

Local Organizers

Angélique Dufresnes	Université Libre de Bruxelles, Belgium
Alejandro Vaisman (Organization Chair)	Université Libre de Bruxelles, Belgium
Stijn Vansummeren	Université Libre de Bruxelles, Belgium
Boris Verhaegen	Université Libre de Bruxelles, Belgium
Gary Verhaegen	Université Libre de Bruxelles, Belgium
Esteban Zimányi	Université Libre de Bruxelles, Belgium

External Referees

Mohammad Al Hasan	Indiana University-Purdue University Indianapolis, USA
Omar Boussaid	ERIC Laboratory, University of Lyon 2, France
Paola Cerchiello	University of Pavia, Italy
Giorgio Corani	IDSIA (Istituto Dalle Molle di Studi sull'Intelligenza Artificiale), Manno, Switzerland
Etienne Cuvelier	Ecole Centrale Paris, France
Daniel Deutch	Ben Gurion University, Israel
Luis Dias	University of Coimbra, Portugal
Marlon Dumas	University of Tartu, Estonia

VIII Organization

Bingsheng He

Mustafa Jarrar

Gabriele Kern-Isberner

Jens Lechtenböcker

Fionn Murtagh

Wim Peters

Hans Weigand

Nanyang Technological University, Singapore

Birzeit University, Palestine

Technische Universität Dortmund, Germany

Westfälische Wilhelms-Universität, Münster,
Germany

Science Foundation, Ireland

University of Sheffield, UK

Tilburg University, The Netherlands

Table of Contents

Managing Complex Multidimensional Data	1
<i>Torben Bach Pedersen</i>	
An Introduction to Business Process Modeling	29
<i>Alejandro Vaisman</i>	
Machine Learning Strategies for Time Series Forecasting	62
<i>Gianluca Bontempi, Souhaib Ben Taieb, and Yann-Aël Le Borgne</i>	
Knowledge Discovery from Constrained Relational Data: A Tutorial on Markov Logic Networks	78
<i>Marcus Spies</i>	
Large Graph Mining: Recent Developments, Challenges and Potential Solutions	103
<i>Sabri Skhiri and Salim Jouili</i>	
Big Data Analytics on Modern Hardware Architectures: A Technology Survey	125
<i>Michael Saecker and Volker Markl</i>	
An Introduction to Multicriteria Decision Aid: The PROMETHEE and GAIA Methods	150
<i>Yves De Smet and Karim Lidouh</i>	
Knowledge Harvesting for Business Intelligence	177
<i>Nesrine Ben Mustapha and Marie-Aude Aufaure</i>	
Business Semantics as an Interface between Enterprise Information Management and the Web of Data: A Case Study in the Flemish Public Administration	208
<i>Christophe Debruyne and Pieter De Leenheer</i>	
Author Index	235