

Editor-in-Chief

Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspan.waw.pl

Mykola Pechenizkiy and Marek Wojciechowski (Eds.)

New Trends in Databases and Information Systems

 Springer

Editors

Mykola Pechenizkiy
Department of Computer Science
Eindhoven University of Technology
Eindhoven
The Netherlands

Marek Wojciechowski
Institute of Computing Science
Poznan University of Technology
Poznan
Poland

ISSN 2194-5357

ISBN 978-3-642-32517-5

DOI 10.1007/978-3-642-32518-2

Springer Heidelberg New York Dordrecht London

e-ISSN 2194-5365

e-ISBN 978-3-642-32518-2

Library of Congress Control Number: 2012943846

© Springer-Verlag Berlin Heidelberg 2013

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. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

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.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

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

Preface

Database and information systems technologies have been rapidly evolving in several directions over the past years. New types and kinds of data, new types of applications and information systems to support them raise diverse challenges to be addressed. The so-called big data challenge, streaming data management and processing, social networks and other complex data analysis, including semantic reasoning into information systems supporting for instance trading, negotiations, and bidding mechanisms are just some of the emerging research topics. This volume contains papers contributed by six workshops and the Ph.D. Consortium associated with the ADBIS 2012 conference that report on the recent developments and an ongoing research in the aforementioned areas.

The 16th East-European Conference on Advances in Databases and Information Systems (ADBIS 2012) was held on September 17–21, 2012 in Poznan, Poland. The main objective of the ADBIS series of conferences is to provide a forum for the dissemination of research accomplishments and to promote interaction and collaboration between the database and information systems research communities from Central and East European countries and the rest of the world. The general idea behind the workshops and the Ph.D. Consortium was to collect contributions from various subdomains of the broad research areas of databases and information systems, representing new trends in these two important areas. Each of the events complementing the main ADBIS conference had its own international program committee, whose members served as the reviewers of papers contributed to the corresponding part of this volume.

ADBIS Workshop on GPUs in Databases (GID 2012) was chaired by Witold Andrzejewski (Poznan University of Technology, Poland), Krzysztof Kaczmarski (Warsaw University of Technology, Poland), and Tobias Lauer (Jedox AG, Germany). The motivation behind the workshop was an observation that while other application domains strongly benefit from utilizing the Graphics Processing Units (GPUs) to increase the performance of processing, database-related applications and algorithms do not get enough attention. The main goal of GID 2012 was to fill this gap. Members of the workshop's Program Committee were: Amitava Datta (University of Western Australia, Commonwealth of Australia), Jarosław Gramacki

(University of Zielona Góra, Poland), Bingsheng He (Nanyang Technological University, Singapore), Ming Ouyang (University of Louisville, USA), John D. Owens (University of California, Davis, USA), Krzysztof Stencel (Warsaw University of Technology, Poland), and Paweł Wojciechowski (Poznan University of Technology, Poland).

Mining Complex and Stream Data (MCSD'12) was chaired by Jerzy Stefanowski (Poznan University of Technology, Poland) and Dominik Ślęzak (University of Warsaw & Infobright Inc., Poland). The aim of this workshop was introducing new algorithmic foundations and application aspects of mining real world difficult data with a particular focus on the emerging domain of stream data mining and mining large volumes of data having a complex structure. The Program Committee of MCSD'12 consisted of the following members: Jan Bazan (University of Rzeszów, Poland), Petr Berka (University of Economics, Prague, Czech Republic), Albert Bifet (University of Waikato, New Zealand), Michelangelo Ceci (University of Bari, Italy), Alfredo Cuzzocrea (ICAR-CNR & University of Calabria, Italy), Simon Fischer (Rapid-I GmbH, Dortmund, Germany), Mohamed Gaber (University of Portsmouth, UK), Jerzy Grzymała-Busse (University of Kansas, USA), Rudolf Kruse (Otto-von-Guericke University, Magdeburg, Germany), Marzena Kryszkiewicz (Warsaw University of Technology, Poland), Stan Matwin (University of Ottawa, Canada), Ernestina Menasalvas (Technical University of Madrid, Spain), Mikołaj Morzy (Poznan University of Technology, Poland), Hung Son Nguyen (The University of Warsaw, Poland), Zbigniew Raś (University of North Carolina, Charlotte, USA), Alexey Tsymbal (Siemens AG, Erlangen, Germany), Michał Woźniak (Wrocław University of Technology, Poland), and Indre Zliobaite (Bournemouth University, UK). The reviewing process was also supported by Dariusz Czerski (IPI PAN at Warsaw, Poland), Piotr Sobolewski (Wrocław University of Technology, Poland), and Gianvito Pio (University of Bari, Italy).

International Workshop on Ontologies meet Advanced Information Systems (OAIS'2012) was chaired by Ladjel Bellatreche (LIAS/ISAE-ENSMA, France) and Yamine Ait Ameer (IRIT-ENSEEIH, France). The workshop had two main objectives. The first one was to present new and challenging issues in the contribution of ontologies for designing high quality information systems. The second one was to present new research and technological developments that use ontologies all over the life cycle of information systems. The Program Committee members of OAIS'2012 were: Brahim Medjahed (Michigan University, USA), David Taniar (Monash University, Australia), Oscar Romero Moral (Universitat Politècnica de Catalunya), Pascal Hitzler (Wright State University, USA), Carlos Ordóñez (Houston University USA), Dickson K.W. Chiu (University of Hong Kong, China), Francesco Guerra (Università di Modena e Reggio Emilia, Italy), Fernando Silva Parreiras (FUMEC University, Brazil), Dimitris Plexousakis (Crete University, Greece), Leandro Krug Wives (Federal University of Rio Grande do Sul, Brazil), Haridimos Kondylakis (FORTH-ICS and University of Crete, Greece), Reza Akbarinia (INRIA, Montpellier, France), Manolis Koubarakis (National and Kapodistrian University of Athens, Greece), Filipe Mota Pinto (Polytechnic Institute of Leiria, Portugal), Stéphane Jean (LIAS/ISAE-ENSMA, France), Mimoun Malki (Sidi Bel Abbès University,

Algeria), Boufaïda Zizette (Constantine University, Algeria), Abdelkamel Tari (Béjaïa University, Algeria), Farouk Toumani (Clermont Ferrand University, France), Chantal Reynaud (LRI, Paris, France), Alfredo Cuzzocrea (ICAR-CNR and University of Calabria, Italy), Daniela Grigori (Prism, Versailles University, France), Juan C. Trujillo (University of Alicante, Spain), Zohra Bellahsene (LIRMM, Montpellier, France), Shonali Krishnaswamy (Monash University, Australia), Selma Khouri (LIAS/ISAE-ENSMA, France), Idir Ait Sadoune (Supelec, France), and Simitsis Alkis (HP, USA).

Second Workshop on Modeling Multi-commodity Trade: Data models and processing (MMT'12) was chaired by Eugeniusz Toczyłowski (Warsaw University of Technology, Poland) and Mariusz Kaleta (Warsaw University of Technology, Poland). Its goal was to address the current scientific and technological challenges in information systems supporting the market mechanism, including: semantics issues in trading, architectures of information systems for market mechanisms, incorporating social networks into trade processes, data modeling for negotiations and bidding, as well as market processes modeling and management. The Program Committee of MMT'12 consisted of the following members: Stanisław Ambroszkiewicz (Polish Academy of Sciences, Poland), Costin Badica (University of Craiova, Romania), Janusz Granat (National Institute of Telecommunications, Poland), Przemysław Kazienko (Wrocław University of Technology, Poland), Zbigniew Nahorski (Polish Academy of Science, Poland), Marcin Paprzycki (Polish Academy of Science, Poland), and Adam Wierzbicki (Polish-Japanese Institute of Information Technology, Poland).

1st ADBIS Workshop on Social Data Processing (SDP'12) was chaired by Jaroslav Pokorný (Charles University in Prague, Czech Republic), Athena Vakali (Aristotle University of Thessaloniki, Greece), and Václav Snášel (VSB - Technical University of Ostrava, Czech Republic). The workshop aimed at addressing the research issues associated with online social networks, including the topics of distributed computing, databases, and storage systems as well as modeling professional profiles of objects like workers, specialists, projects, supervisors, etc. The members of the Program Committee of SDP'12 were: Peter Vojtáš (Charles University in Prague, Czech Republic), Kamil Matoušek (Czech Technical University, Czech Republic), Jiří Kubalík (Czech Technical University, Czech Republic), Petr Křemen (Czech Technical University, Czech Republic), Hakim Hacid (Bell Labs, France), Nick Papanikolaou (HP Labs, UK), Myra Spiliopoulou (University of Magdeburg, Germany), Ernestina Menasalvas (Technical University of Madrid, Spain), Maria Augusta Nunes (Federal University of Sergipe, Brazil), Miloš Kudělka (VSB - Technical University of Ostrava, Czech Republic), and Jan Martinovič (VSB - Technical University of Ostrava, Czech Republic).

1st ADBIS Workshop on Social and Algorithmic Issues in Business Support (SAIBS) was chaired by Adam Wojciechowski (Poznan University of Technology, Poland) and Alok Mishra (Atılım University, Turkey). The focus of the workshop was on computational and optimization issues that can be supported by crowd input or social intelligence. Specific goals included addressing the following problems: how far and on which fields business may benefit from utilizing social contribution,

and how computer systems may understand social behavior and support humans in making decisions. The Program Committee of SAIBS consisted of the following members: Ricardo Colomo Palacios (Carlos III University of Madrid, Spain), Arianna D'Ulizia (IRPPS, National Research Council, Rome, Italy), Fernando Ferri (IRPPS, National Research Council, Rome, Italy), Patrizia Grifoni (IRPPS, National Research Council, Rome, Italy), Kyoung Jun Lee (Kyung Hee University, Korea), Mirosław Ochodek (Poznan University of Technology, Poland), Rory O'Connor (Dublin City University, Ireland), Robert Susmaga (Poznan University of Technology, Poland), and Agnieszka Węgrzyn (University of Zielona Góra, Poland).

ADBIS Ph.D. Consortium was established as a forum for Ph.D. students to present their research ideas, confront them with the scientific community, and receive feedback from senior mentors. Ph.D. students at an advanced stage of research were given an opportunity to prepare a paper devoted to their research area with a possibility of sharing their achieved preliminary results. The chairs of the Ph.D. Consortium, responsible for selecting the papers from this category, were Mikołaj Morzy (Poznan University of Technology, Poland) and Alexandros Nanopoulos (Catholic University of Eichstätt-Ingolstadt, Germany).

We would like to thank the authors, the reviewers, and the chairs of the ADBIS 2012 workshops for their work and effort without which assembling this volume would not be possible.

September 2012

Mykola Pechenizkiy
Marek Wojciechowski

allegro group

POZnań*
*Miasto know-how

IBM

Microsoft®

Roche

T A R G I T  **®**
business intelligence

SAMSUNG

EDGE  **SOLUTIONS**

itelligence

Contents

Part I: GPUs in Databases

Applying CUDA Technology in DCT-Based Method of Query Selectivity Estimation	3
<i>Dariusz Rafal Augustyn, Sebastian Zederowski</i>	
Processing of Range Query Using SIMD and GPU	13
<i>Pavel Bednář, Petr Gajdoš, Michal Krátký, Peter Chovanec</i>	
Towards Optimization of Hybrid CPU/GPU Query Plans in Database Systems	27
<i>Sebastian Breß, Eike Schallehn, Ingolf Geist</i>	
Thrust and CUDA in Data Intensive Algorithms	37
<i>Krzysztof Kaczmarski, Paweł Rzążewski</i>	

Part II: Mining Complex and Stream Data

A Detection of the Most Influential Documents	49
<i>Dariusz Ceglarek, Konstanty Haniewicz</i>	
Approximation Algorithms for Massive High-Rate Data Streams	59
<i>Alfredo Cuzzocrea</i>	
Comparing Block Ensembles for Data Streams with Concept Drift	69
<i>Magdalena Deckert, Jerzy Stefanowski</i>	
Adapting Travel Time Estimates to Current Traffic Conditions	79
<i>Przemysław Gawęł, Krzysztof Dembczyński, Robert Susmaga, Przemysław Wesolek, Piotr Zielniewicz, Andrzej Jaszkievicz</i>	

SONCA: Scalable Semantic Processing of Rapidly Growing Document Stores	89
<i>Marek Grzegorowski, Przemysław Wiktor Pardel, Sebastian Stawicki, Krzysztof Stencel</i>	
Collective Classification Techniques: An Experimental Study	99
<i>Tomasz Kajdanowicz, Przemysław Kazienko, Marcin Janczak</i>	
Granular Knowledge Discovery Framework: A Case Study of Incident Data Reporting System	109
<i>Adam Krasuski, Dominik Ślęzak, Karol Kreński, Stanisław Łazowy</i>	
Diversity in Ensembles for One-Class Classification	119
<i>Bartosz Krawczyk</i>	
Evaluation of Stream Data by Formal Concept Analysis	131
<i>Martin Radvanský, Vladimír Sklenář, Václav Snášel</i>	
Soft Competitive Learning for Large Data Sets	141
<i>Frank-Michael Schleich, Xibin Zhu, Barbara Hammer</i>	
Enhancing Concept Drift Detection with Simulated Recurrence	153
<i>Piotr Sobolewski, Michał Woźniak</i>	
DeltaDens – Incremental Algorithm for On-Line Density-Based Clustering	163
<i>Radostaw Z. Ziemiński</i>	
Part III: Ontologies Meet Advanced Information Systems	
Introducing Artificial Neural Network in Ontologies Alignment Process	175
<i>Warith Eddine Djeddi, Mohamed Tarek Khadir</i>	
Time Integration in Semantic Trajectories Using an Ontological Modelling Approach: A Case Study with Experiments, Optimization and Evaluation of an Integration Approach	187
<i>Rouaa Wannous, Jamal Malki, Alain Bouju, Cécile Vincent</i>	
WebOMSIE: An Ontology-Based Multi Source Web Information Extraction	199
<i>Zineb Younsi, Mohamed Quafafou, Redouane Ouzegane, Abdelkamel Tari</i>	
Part IV: Modeling Multi-commodity Trade: Data Models and Processing	
Bidding Languages for Continuous Auctions	211
<i>Mariusz Kaleta</i>	

Auction of Time as a Tool for Solving Multiagent Scheduling Problems	221
<i>Piotr Modliński</i>	
Application of an Auction Algorithm in an Agent-Based Power Balancing System	231
<i>Piotr Pałka, Weronika Radziszewska, Zbigniew Nahorski</i>	
Multi-commodity Trade Application to the Routing Algorithm for the Delay and Disruptive Tolerant Networks	241
<i>Piotr Pałka, Radosław Schoeneich</i>	
Offers Discovery and Identifying User Requirements for Multi-commodity Trade in Open Markets	251
<i>Dominik Ryzko, Anna Wróblewska</i>	
Fair Resource Allocation in Multi-commodity Networks	261
<i>Tomasz Śliwiński</i>	
Part V: Social Data Processing	
Heuristic Approach to Automatic Wrapper Generation for Social Media Websites	273
<i>Bartosz Baziński, Michał Brzezicki</i>	
Spectral Clustering: Left-Right-Oscillate Algorithm for Detecting Communities	285
<i>Pavla Dráždilová, Jan Martinovič, Kateřina Slaninová</i>	
Exploiting Potential of the Professional Social Network Portal “SitIT”	295
<i>Kamil Matoušek, Jiří Kubalík, Martin Nečaský, Peter Vojtáš</i>	
Modeling and Storing Complex Network with <i>Graph-Tree</i>	305
<i>Adan Lucio Pereira, Ana Paula Appel</i>	
Evolution of Author’s Profiles Based on Analysis of DBLP Data	317
<i>Martin Radvanský, Zdeněk Horák, Miloš Kudělka, Václav Snášel</i>	
Towards Effective Social Network System Implementation	327
<i>Jaroslav Škrabálek, Petr Kunc, Filip Nguyen, Tomáš Pitner</i>	
Part VI: Social and Algorithmic Issues in Business Support	
Community Traffic: A Technology for the Next Generation Car Navigation	339
<i>Przemysław Gaweł, Krzysztof Dembczyński, Wojciech Kotłowski, Marek Kubiak, Robert Susmaga, Przemysław Wesolek, Piotr Zielniewicz, Andrzej Jaszkievicz</i>	

Situational Requirement Method System: Knowledge Management in Business Support	349
<i>Deepti Mishra, Secil Aydin, Alok Mishra</i>	
Effectiveness Analysis of Promotional Features Used in Internet Auctions: Empirical Study	361
<i>Adam Wojciechowski, Paweł Warczynski</i>	
Part VII: Ph.D. Consortium	
Data Mining Approach to Digital Image Processing in Old Painting Restoration	373
<i>Joanna Gancarczyk</i>	
Determining Document's Semantic Orientation Using kNN Algorithm	383
<i>Krzysztof Jędrzejewski, Maurycy Zamorski</i>	
Designing a Software Transactional Memory for Peer-to-Peer Systems	395
<i>Aurel Paulovič, Pavol Návrat</i>	
Traceability in Software Architecture Decisions Based on Notes about Documents	403
<i>Gilberto Pedraza-Garcia, Dario Correal</i>	
OLAP Models for Sequential Data – Current State of Research and Open Problems	415
<i>Łukasz Nienartowicz</i>	
Data Management for Fingerprint Recognition Algorithm Based on Characteristic Points' Groups	425
<i>Michał Szczepanik, Ireneusz Józwiak</i>	
Data Prefetching Based on Long-Term Periodic Access Patterns	433
<i>Dmitri Vasilik</i>	
E-ETL: Framework for Managing Evolving ETL Processes	441
<i>Artur Wojciechowski</i>	
Author Index	451