Communications
in Computer and Information Science

Editorial Board

Simone Diniz Junqueira Barbosa
  Pontifical Catholic University of Rio de Janeiro (PUC-Rio),
  Rio de Janeiro, Brazil
Phoebe Chen
  La Trobe University, Melbourne, Australia
Alfredo Cuzzocrea
  ICAR-CNR and University of Calabria, Cosenza, Italy
Xiaoyong Du
  Renmin University of China, Beijing, China
Joaquim Filipe
  Polytechnic Institute of Setúbal, Setúbal, Portugal
Orhun Kara
  TÜBİTAK BİLGEM and Middle East Technical University, Ankara, Turkey
Igor Kotenko
  St. Petersburg Institute for Informatics and Automation of the Russian
  Academy of Sciences, St. Petersburg, Russia
Krishna M. Sivalingam
  Indian Institute of Technology Madras, Chennai, India
Dominik Śleżak
  University of Warsaw and Infobright, Warsaw, Poland
Takashi Washio
  Osaka University, Osaka, Japan
Xiaokang Yang
  Shanghai Jiao Tong University, Shanghai, China
Preface

Collecting, processing, and analyzing data became important branches of computer science. Many areas of our existence generate a wealth of information that must be stored in a structured manner and processed appropriately in order to gain the knowledge from the inside. Databases have become a ubiquitous way of collecting and storing data. They are used to hold data describing many areas of human life and activity, and as a consequence, they are also present in almost every IT system. Today’s databases have to face the problem of data proliferation and growing variety. More efficient methods for data processing are needed more than ever. New areas of interests that deliver data require innovative algorithms for data analysis.

Beyond Databases, Architectures and Structures (BDAS) is a series of conferences that intends to give the state of the art of the research that satisfies the needs of modern, widely understood database systems, architectures, models, structures, and algorithms focused on processing various types of data. The aim of the conference is to reflect the most recent developments of databases and allied techniques used for solving problems in a variety of areas related to database systems, or even go one step forward – beyond the horizon of existing databases, architectures, and data structures. The 11th International BDAS Scientific Conference was a continuation of the highly successful BDAS conference series started in 2005 in Ustroń, Poland. For many years BDAS has been attracting hundreds or even thousands of researchers and professionals working in the field of databases. Among attendees of our conference were scientists and representatives of IT companies. Several editions of BDAS were supported by our commercial, world-renowned partners, developing solutions for the database domain, such as IBM, Microsoft, Sybase, Oracle, and others. BDAS annual meetings have become an arena for exchanging information on the widely understood database systems and data processing algorithms.

BDAS 2015 was the 11th edition of the conference, organized under the technical cosponsorship of the IEEE Poland Section. We also continued our successful cooperation with Springer, which resulted in the publication of this book. The conference attracted more than a hundred participants from 15 countries, who made this conference a successful and memorable event. There were five keynote talks given by leading scientists: Prof. Bora İ. Kumova from Department of Computer Engineering, İzmir Institute of Technology (İYTE), İzmir, Turkey spoke on ‘Fuzzy syllogistic reasoning over relational data.’ Prof. Dirk Labudde from Bioinformatics group Mittweida (bigM) and Forensic Science Investigation Lab (FoSIL), University of Applied Sciences, Mittweida, Germany gave an excellent talk entitled ‘Bioinformatics and Forensics - How today’s Life Science technologies can shape the Crime Sciences of tomorrow.’ Prof. Jean-Charles Lamirel from SYNALP team, LORIA, France gave a very enlightening speech on ‘New metrics and related statistical approaches for efficient mining in very large and highly multidimensional databases,’ Prof. Mikhail Moshkov from King Abdullah University for Science and Technology (KAUST), Saudi Arabia honored us with
a presentation on ‘Extensions of dynamic programming for design and analysis of decision trees,’ and Dr. Riccardo Rasconi from Institute of Cognitive Science and Technology, National Research Council, Rome, Italy spoke on ‘Surveying the versatility of constraint-based large neighborhood search for scheduling problems.’ The keynote speeches and plenary sessions gained insight into new areas.

BDAS is focused on all aspects of databases. It is intended to have a broad scope, including different kinds of data acquisition, processing, and storing, and this book reflects fairly well the large span of research presented at BDAS 2015. This volume consists of 53 carefully selected papers. The first three papers accompany the stunning keynote talks. The remainder of the papers are assigned to eight thematic groups:

- Database architectures and performance
- Data integration, storage, and data warehousing
- Ontologies and Semantic Web
- Artificial intelligence, data mining, and knowledge discovery
- Image analysis and multimedia mining
- Spatial data analysis
- Database systems development
- Applications of database systems

The first group is related to various database architectures, query optimization, and database performance. Papers gathered in this group discuss hot topics of query selectivity estimation, testing performance of various database systems, NoSQL and data consistency, temporal and probabilistic databases. The next group of papers concern issues related to data integration, data storage, and data warehousing. The group consists of seven papers presenting research devoted to the data mapping semantics while sharing and exchanging data, novel data integration architectures, efficiency of storage space configuration, new ETL concepts, and data warehouse modeling.

The third group consists of three papers devoted to ontologies and the Semantic Web. These papers discuss problems of automatic approaches for building ontology from relational data, data integration with ontology, and RDF graph partitioning. The research devoted to artificial intelligence and data mining is presented in eight papers gathered in the fourth group. These papers show a wide spectrum of applications of various exploration methods, like decision rules, knowledge-based systems, clustering, artificial immune systems and memetic algorithms, Dynamic Gaussian Bayesian Network models, to solve many real problems.

The next group of papers is focused on image analysis and multimedia mining. This group consists of six papers devoted to lossless compression of images, querying multimedia databases, real-time object detection from depth images, analysis of facial expressions, emotions, and medical images.

Some aspects of the spatial data collecting and processing are discussed in three successive papers. The next three papers show various aspects of database systems development. Finally, the last 10 papers present different usage of databases starting from mining and metallurgical industries, through different fuel and energy consumption related problems and ERP systems, ending with bioinformatics knowledgebase and databases storing affect-annotated data and faces.
We would like to thank all Program Committee members and additional reviewers for their effort in reviewing the papers. Special thanks to Piotr Kuźniacki - builder and for ten years administrator of our website www.bdas.pl. The conference organization would not have been possible without the technical staff: Dorota Huget and Jacek Pietraszuk.

We hope that the broad scope of topics related to databases covered in this proceedings volume will help the reader to understand that databases have become an important element of nearly every branch of computer science.

April 2015

Stanisław Kozielski
Dariusz Mrozek
Paweł Kasprowski
Bożena Małysiak-Mrozek
Daniel Kostrzewa
Organization

BDAS 2015 was organized by Institute of Informatics, Silesian University of Technology, Poland.

BDAS 2015 Program Committee

Honorary Member
Lotfi A. Zadeh University of California, Berkeley, USA

Chair
Stanisław Kozielski Silesian University of Technology, Poland

Members
Sansanee Auephanwiriyakul Chiang Mai University, Thailand
Werner Backes Sirrix AG Security Technologies, Bochum, Germany
Susmit Bagchi Gyeongsang National University, South Korea
Patrick Bours Gjøvik University College, Norway
George D.C. Cavalcanti Universidade Federal de Pernambuco, Brazil
Po-Yuan Chen China Medical University, Taichung, Taiwan, University of British Columbia, BC, Canada
Yixiang Chen East China Normal University, Shanghai
Tadeusz Czachórski IITiS, Polish Academy of Sciences, Poland
Andrzej Chydzieński Silesian University of Technology, Poland
Sebastian Deorowicz Silesian University of Technology, Poland
Jack Dongarra University of Tennessee, Knoxville, USA
Andrzej Drygajło École Polytechnique Fédérale de Lausanne, Switzerland
Moawia Elfaki Yahia King Faisal University, Saudi Arabia
Rudolf Fleischer German University of Technology, Oman
Hamido Fujita Iwate Prefectural University, Japan
Krzysztof Goczyła Gdańsk University of Technology, Poland
Marcin Gorawski Silesian University of Technology, Poland
Jarek Gryz York University, Ontario, Canada
Andrzej Grzywak Silesian University of Technology, Poland
Brahim Hnich Izmir University of Economics, Izmir, Turkey
Edward Hryniewicz
Jiewen Huang
Xiaohua Tony Hu
Zbigniew Huzar
Tomasz Imieliński
Paweł Kasprowski
Przemysław Kazienko
Jerzy Klamka
Bora İ Kumova
Andrzej Kwiecień
Jean-Charles Lamirel
Sérgio Lifschitz
Antoni Ligęza
Bożena Małysiak-Mrozek
Marco Masseroli
Zygmunt Mazur
Yasser F. O. Mohammed
Tadeusz Morzy
Mikhail Moshkov
Dariusz Mrozek
Mieczysław Muraszkiewicz
Sergio Nesmachnow
Tadeusz Pankowski
Witold Pedrycz
Adam Pelikan
Ewa Piętka
Bolesław Pochopień
Andrzej Polański
Hugo Proença
Riccardo Rasconi
Marek Rejman-Greene
Jerzy Rutkowski
Henryk Rybiński
Galina Setlak
Marek Sikora
Przemysław Stpiczyński

Silesian University of Technology, Poland
Google Inc., Mountain View, CA, USA
Drexel University, Philadelphia, PA, USA
Wrocław University of Technology, Poland
Rutgers University, New Brunswick, New Jersey, USA
Silesian University of Technology, Poland
Wrocław University of Technology, Poland
IITiS, Polish Academy of Sciences, Poland
Izmir Institute of Technology, Turkey
Silesian University of Technology, Poland
LORIA, Nancy, France, University of Strasbourg, France
Pontificia Universidade Católica do Rio de Janeiro, Brazil
AGH University of Science and Technology, Poland
Silesian University of Technology, Poland
Politecnico di Milano, Italy
Wrocław University of Technology, Poland
Assiut University, Egypt
Poznań University of Technology, Poland
King Abdullah University of Science and Technology, Saudi Arabia
Silesian University of Technology, Poland
Warsaw University of Technology, Poland
Universidad de la República, Uruguay
Poznań University of Technology, Poland
University of Alberta, Canada
Lodz University of Technology, Poland
Silesian University of Technology, Poland
Silesian University of Technology, Poland
University of Beira Interior, Portugal
Institute for Cognitive Sciences and Technologies, Italian National Research Council, Italy
Centre for Applied Science and Technology in Home Office Science, UK
Silesian University of Technology, Poland
Warsaw University of Technology, Poland
Rzeszow University of Technology, Poland
Silesian University of Technology and Institute of Innovative Technologies EMAG, Poland
Maria Curie-Skłodowska University, Poland

Organization XI

Dominik Ślęzak University of Warsaw and Infobright Inc., Poland
Andrzej Świerniak Silesian University of Technology, Poland
Adam Świtoński Silesian University of Technology, Poland
Karin Verspoor University of Melbourne, Australia
Alicja Wakulicz-Deja University of Silesia in Katowice, Poland
Sylwester Warecki Intel Corporation, San Diego, California, USA
Tadeusz Wieczorek Silesian University of Technology, Poland
Konrad Wojciechowski Silesian University of Technology, Poland
Robert Wrembel Poznań University of Technology, Poland
Stanisław Wrycza University of Gdańsk, Poland
Miroslaw Zaborowski IITiS, Polish Academy of Sciences, Poland
Grzegorz Zaręba University of Arizona, Tucson, USA
Krzysztof Zieliński AGH University of Science and Technology, Poland
Quan Zou Xiamen University, People’s Republic of China

Organizing Committee
Bożena Małysiak-Mrozek Silesian University of Technology, Poland
Dariusz Mrozek Silesian University of Technology, Poland
Paweł Kasprowski Silesian University of Technology, Poland
Daniel Kostrzewa Silesian University of Technology, Poland
Piotr Kuźniacki Silesian University of Technology, Poland

Additional Reviewers
Augustyn Dariusz Rafał Piórkowski Adam
Bach Małgorzata Pluciennik-Psota Ewa
Bajerski Piotr Respondek Jerzy
Brzeski Robert Romuk Ewa
Duszenko Adam Sitek Paweł
Frączek Jacek Świderski Michał
Harężlak Katarzyna Szwoch Wioleta
Josński Henryk Traczyk Tomasz
Kawulok Michał Tutajewicz Robert
Kozielski Michał Walosz Wójciech
Michalak Marcin Werner Aleksandra
Momot Alina Wycislik Łukasz
Niedbała Sławomir Zghidi Hafed
Nowak-Brzezińska Agnieszka Zielosko Beata
Nurzyńska Karolina

Sponsoring Institutions
Technical cosponsorship of the IEEE Poland Section
Contents

Invited Papers

New Metrics and Related Statistical Approaches for Efficient Mining in Very Large and Highly Multidimensional Databases .......................... 3
   Jean-Charles Lamirel

Generating Ontologies from Relational Data with Fuzzy-Syllogistic Reasoning ............................................................... 21
   Bora ˙I. Kumova

Surveying the Versatility of Constraint-Based Large Neighborhood Search for Scheduling Problems ........................................ 33
   Riccardo Rasconi, Angelo Oddi, and Amedeo Cesta

Database Architectures and Performance

Query Workload Aware Multi-histogram Based on Equi-width Sub-histograms for Selectivity Estimations of Range Queries .............. 47
   Dariusz Rafał Augustyn

Analysis of the Effect of Chosen Initialization Parameters on Database Performance .......................................................... 60
   Wanda Gryglewicz-Kacerka and Jarosław Kacerka

Database Under Pressure - Scaling Database Performance Tests in Microsoft Azure Public Cloud ............................................. 69
   Dariusz Mrozek, Anna Paliga, Bożena Małysiak-Mrozek, and Stanisław Kozielski

Comparison Between Performance of Various Database Systems for Implementing a Language Corpus ........................................ 82
   Dimuthu Upeksha, Chamila Wijayarathna, Maduranga Siriwardena, Lahiri Lasandun, Chinthana Wimalasuriya, N.H.N.D. de Silva, and Gihan Dias

A Comparison of Different Forms of Temporal Data Management ...... 92
   Florian Künzner and Dušan Petković

Performance Aspects of Migrating a Web Application from a Relational to a NoSQL Database ...................................................... 107
   Katarzyna Harezlak and Robert Skowron

A Consensus Quorum Algorithm for Replicated NoSQL Data ............ 116
   Tadeusz Pankowski
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preserving Data Consistency in Scalable Distributed Two Layer Data</td>
<td>126</td>
</tr>
<tr>
<td>Structures</td>
<td></td>
</tr>
<tr>
<td><em>Adam Krechowicz, Stanislaw Denizak, Grzegorz Lukawski,</em></td>
<td></td>
</tr>
<tr>
<td><em>and Mariusz Bedla</em></td>
<td></td>
</tr>
<tr>
<td>Modern Temporal Data Models: Strengths and Weaknesses</td>
<td>136</td>
</tr>
<tr>
<td><em>Dušan Petković</em></td>
<td></td>
</tr>
<tr>
<td>Multiargument Relationships in Possibilistic Databases</td>
<td>147</td>
</tr>
<tr>
<td><em>Krzysztof Myszkorowski</em></td>
<td></td>
</tr>
<tr>
<td><strong>Data Integration, Storage and Data Warehousing</strong></td>
<td></td>
</tr>
<tr>
<td>Data Sharing and Exchange: General Data-Mapping Semantics</td>
<td>159</td>
</tr>
<tr>
<td><em>Rana Awada and Iluju Kiringa</em></td>
<td></td>
</tr>
<tr>
<td>A Universal Cuboid-Based Integration Architecture for Polyglotic</td>
<td>170</td>
</tr>
<tr>
<td>Querying of Heterogeneous Datasources</td>
<td></td>
</tr>
<tr>
<td><em>Michał Chromiak, Piotr Wiśniewski, and Krzysztof Stencel</em></td>
<td></td>
</tr>
<tr>
<td>Efficient Multidisk Database Storage Configuration</td>
<td>180</td>
</tr>
<tr>
<td><em>Mateusz Smolinski</em></td>
<td></td>
</tr>
<tr>
<td>E–LT Concept in a Light of New Features of Oracle Data Integrator</td>
<td>190</td>
</tr>
<tr>
<td>12c Based on Data Migration within a Hospital Information System</td>
<td></td>
</tr>
<tr>
<td><em>Lukasz Wycislik, Dariusz Rafał Augustyn, Dariusz Mrozek,</em></td>
<td></td>
</tr>
<tr>
<td><em>Ewa Pluciennik, Hafed Zghidi, and Robert Brzeski</em></td>
<td></td>
</tr>
<tr>
<td>Automating Schema Integration Technique Case Study: Generating</td>
<td>200</td>
</tr>
<tr>
<td>Data Warehouse Schema from Data Mart Schemas</td>
<td></td>
</tr>
<tr>
<td><em>Nouha Arfaoui and Jalel Akachi</em></td>
<td></td>
</tr>
<tr>
<td>Proposal of a New Data Warehouse Architecture Reference Model</td>
<td>210</td>
</tr>
<tr>
<td><em>Dariusz Dymek, Wojciech Komnata, and Piotr Szwed</em></td>
<td></td>
</tr>
<tr>
<td>DWARM: An Ontology of Data Warehouse Architecture Reference Model</td>
<td>222</td>
</tr>
<tr>
<td><em>Piotr Szwed, Wojciech Komnata, and Dariusz Dymek</em></td>
<td></td>
</tr>
<tr>
<td><strong>Ontologies and Semantic Web</strong></td>
<td></td>
</tr>
<tr>
<td>Ontology Learning from Relational Database: How to Label the</td>
<td>235</td>
</tr>
<tr>
<td>Relationships Between Concepts?</td>
<td></td>
</tr>
<tr>
<td><em>Bouchra El Idrissi, Salah Baïna, and Karim Baïna</em></td>
<td></td>
</tr>
<tr>
<td>Integration of Facebook Online Social Network User Profiles into a</td>
<td>245</td>
</tr>
<tr>
<td>Knowledgebase</td>
<td></td>
</tr>
<tr>
<td><em>Wojciech Kijas and Michał Kozielski</em></td>
<td></td>
</tr>
</tbody>
</table>
RDF Graph Partitions: A Brief Survey .......................... 256
Dominik Tomaszuk, Łukasz Skonieczny, and David Wood

Artificial Intelligence, Data Mining and Knowledge Discovery

Optimization of Inhibitory Decision Rules Relative to Coverage – Comparative Study ........................................... 267
Beata Zielosko

Application of the Shapley-Shubik Power Index in the Process of Decision Making on the Basis of Dispersed Medical Data ........ 277
Małgorzata Przybyła-Kasperek

Inference in Expert Systems Using Natural Language Processing ...... 288
Tomasz Jach and Tomasz Xięski

Impact of Parallel Memetic Algorithm Parameters on Its Efficacy ...... 299
Miroslaw Blocho and Jakub Nalepa

Data Processing in Immune Optimization of the Structure .......... 309
Arkadiusz Poteralski

A Prudent Based Approach for Customer Churn Prediction .......... 320
Adnan Amin, Faisal Rahim, Muhammad Ramzan, and Sajid Anwar

Forecasting Daily Urban Water Demand Using Dynamic Gaussian Bayesian Network ................................................. 333
Wojciech Froelich

Parallel Density-Based Stream Clustering Using a Multi-user GPU Scheduler ................................................................. 343
Ayman Tarakji, Marwan Hassani, Lyubomir Georgiev, Thomas Seidl, and Rainer Leupers

Image Analysis and Multimedia Mining

Lossless Compression of Medical and Natural High Bit Depth Sparse Histogram Images .................................................. 363
Roman Starosolski

Query by Shape for Image Retrieval from Multimedia Databases .... 377
Stanisław Deniziak and Tomasz Michno

Real-Time People Counting from Depth Images .......................... 387
Jakub Nalepa, Janusz Szymanek, and Michal Kawulok
PCA Application in Classification of Smiling and Neutral Facial Displays ................................................................. 398  
Karolina Nurzynska and Bogdan Smolka

Detection of Tuberculosis Bacteria in Sputum Slide Image Using Morphological Features ........................................... 408  
Zahoor Jan, Muhammad Rafiq, Hayat Muhammad, and Noor Zada

Automatic Medical Objects Classification Based on Data Sets and Domain Knowledge .............................................. 415  
Przemysław Wiktor Pardel, Jan G. Bazan, Jacek Zarychta, and Stanisława Bazan-Socha

Spatial Data Analysis

Interpolation as a Bridge Between Table and Array Representations of Geofields in Spatial Databases................................. 427  
Piotr Bajerski

A Proposal of Hybrid Spatial Indexing for Addressing the Measurement Points in Monitoring Sensor Networks .......................................... 437  
Michał Lupa, Monika Chuchro, Adam Piórkowski, Anna Pięta, and Andrzej Leśniak

An Attempt to Automate the Simplification of Building Objects in Multiresolution Databases ........................................ 448  
Michał Lupa, Krystian Koziol, and Andrzej Leśniak

Database Systems Development

Motivation Modeling and Metrics Evaluation of IT Architect Certification Programs......................................................... 463  
Michał Turek and Jan Werewka

Supporting Code Review by Automatic Detection of Potentially Buggy Changes ............................................................. 473  
Mikołaj Fejzer, Michał Wojtyna, Marta Burzańska, Piotr Wiśniewski, and Krzysztof Stencel

Project Management in the Scrum Methodology ................................. 483  
Maria Tytkowska, Aleksandra Werner, and Małgorzata Bach

Applications of Database Systems

Regression Rule Learning for Methane Forecasting in Coal Mines ........ 495  
Michał Kozieński, Adam Skowron, Łukasz Wróbel, and Marek Sikora
On Design of Domain-Specific Query Language for the Metallurgical Industry ........................................................ 505  
   Andrey Borodin, Yuri Kiselev, Sergey Mirvoda, and Sergey Porshnev

Approach to the Monitoring of Energy Consumption in Eco-grinder Based on ABC Optimization ...................................... 516  
   Jacek Czerniak, Dawid Ewald, Marek Macko, Grzegorz Śmigielski, and Krzysztof Tyszczuk

MercuryDb - A Database System Supporting Management and Limitation of Mercury Content in Fossil Fuels .................... 530  
   Sebastian Iwaszenko, Karolina Nurzynska, and Barbara Bialecka

Liquefied Petroleum Storage and Distribution Problems and Research Thesis .......................................................... 540  
   Marcin Gorawski, Anna Gorawska, and Krzysztof Pasterak

Idea of Impact of ERP-APS-MES Systems Integration on the Effectiveness of Decision Making Process in Manufacturing Companies .......................................................... 551  
   Edyta Kucharska, Katarzyna Grobler-Dębska, Jarosław Gracel, and Mieczysław Jagodziński

A Concept of Decision Support in Supply Chain Management – A Hybrid Approach ..................................................... 565  
   Paweł Sitek

eProS – A Bioinformatics Knowledgebase, Toolbox and Database for Characterizing Protein Function ................................ 575  
   Florian Heinke, Daniel Stockmann, Stefan Schildbach, Mathias Langer, and Dirk Labudde

Evaluation Criteria for Affect-Annotation Databases .................... 585  
   Agata Kolakowska, Agnieszka Landowska, Mariusz Szwoch, Wioleta Szwoch, and Michal R. Wrobel

The Mask: A Face Network System for Bell’s Palsy Recovery Surveillance .............................................................. 598  
   Hanen Bouali and Jalel Akaichi

Author Index .................................................. 611