# Lecture Notes in Artificial Intelligence

#### 8502

# Subseries of Lecture Notes in Computer Science

#### **LNAI Series Editors**

Randy Goebel
University of Alberta, Edmonton, Canada
Yuzuru Tanaka
Hokkaido University, Sapporo, Japan
Wolfgang Wahlster
DFKI and Saarland University, Saarbrücken, Germany

#### **LNAI Founding Series Editor**

Joerg Siekmann

DFKI and Saarland University, Saarbrücken, Germany

Troels Andreasen Henning Christiansen Juan-Carlos Cubero Zbigniew W. Raś (Eds.)

# Foundations of Intelligent Systems

21st International Symposium, ISMIS 2014 Roskilde, Denmark, June 25-27, 2014 Proceedings



#### Volume Editors

Troels Andreasen Roskilde University, Denmark

E-mail: troels@ruc.dk

Henning Christiansen Roskilde University, Denmark E-mail: henning@ruc.dk

Juan-Carlos Cubero University of Granada, Spain E-mail: jc.cubero@decsai.ugr.es

Zbigniew W. Raś University of North Carolina, Charlotte, NC, USA and Warsaw University of Technology, Poland E-mail: ras@uncc.edu

ISSN 0302-9743 ISBN 978-3-319-08325-4 DOI 10.1007/978-3-319-08326-1 e-ISSN 1611-3349 e-ISBN 978-3-319-08326-1

Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014941587

LNCS Sublibrary: SL 7 – Artificial Intelligence

#### © Springer International Publishing Switzerland 2014

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 ist 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.

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**

This volume contains the papers presented at ISMIS 2014: 21st International Symposium on Methodologies for Intelligent Systems held during June 25–27, 2014, in Roskilde, Denmark. The symposium was organized by members of the PLIS research group: Programming, Logic and Intelligent Systems, of the Department of Communication, Business and Information Technologies at Roskilde University.

ISMIS is a conference series that started in 1986 and has developed into an established and prestigious conference for exchanging the latest research results in building intelligent systems. The scope of ISMIS represents a wide range of topics on applying artificial intelligence techniques to areas as diverse as decision support, automated deduction, reasoning, knowledge-based systems, machine learning, computer vision, robotics, planning, databases, information retrieval, and so on. ISMIS provides a forum and a means for exchanging information for those interested purely in theory, those interested primarily in implementation, and those interested in specific research and industrial applications.

We want to express our special thanks to the Program Committee members and everyone who contributed at any level to the organization of ISMIS 2014. Also, special thanks to our invited speakers, Matthias Jarke, Xavier Serra and Steffen Staab. We would like to thank every author who submitted a paper to ISMIS 2014 and finally the team of EasyChair, without whose free software the handling of submissions and editing of the proceedings could not have been managed so smoothly by a small group of people. Last but not the least, we thank Alfred Hofmann of Springer for his continuous support.

April 2014

Troels Andreasen Henning Christiansen Juan-Carlos Cubero Zbigniew W. Raś

## Organization

The symposium was organized by members of the PLIS research group: Programming, Logic and Intelligent Systems, of the Department of Communication, Business and Information Technologies at Roskilde University.

#### General Chair

Zbigniew W. Raś University of North Carolina, Charlotte, USA

and Warsaw University of Technology,

Poland

#### Symposium Chair

Troels Andreasen Roskilde University, Denmark

#### **Program Co-chairs**

Henning Christiansen Roskilde University, Denmark Juan Carlos Cubero University of Granada, Spain

#### Steering Committee

Aijun An York University, Canada

Alexander Felfernig Graz University of Technology, Austria

Andrzej Skowron University of Warsaw, Poland

Dominik Ślęzak Infobright Inc., Canada; and University of

Warsaw, Poland

Henryk Rybinski Warsaw University of Technology, Poland

Jaime Carbonell CMU, USA

Jan Rauch University of Economics, Prague,

Czech Republic

Jiming Liu Hong Kong Baptist University, Hong Kong,

SAR China

Li Chen Hong Kong Baptist University, Hong Kong,

SAR China

Lorenza Saitta University of Piemonte Orientale, Italy

Maria Zemankova NSF, USA

Marzena Kryszkiewicz Warsaw University of Technology, Poland

Nick Cercone York University, Canada

#### VIII Organization

Petr Berka University of Economics, Prague,

Czech Republic

Stan Matwin University of Ottawa, Canada

Tapio Elomaa Tampere University of Technology, Finland Zbigniew W. Raś UNC-Charlotte, USA; and Warsaw University

of Technology, Poland

#### **Program Committees**

#### ISMIS Regular Papers and Posters

Luigia Carlucci Aiello Sapienza Università di Roma, Italy

Aijun An York University, Canada
Troels Andreasen Roskilde University, Denmark
Annalisa Appice University of Bari, Italy
Salima Benbernou University of Paris V, France

Marenglen Biba University of New York Tirana, Albania Maria Bielikova Slovak University of Technology, Slovakia

Ivan Bratko University of Ljubljana, Slovenia Francois Bry University of Munich, Germany Henrik Bulskov Roskilde University, Denmark Sandra Carberry University of Delaware, USA Michelangelo Ceci University of Bari, Italy

Jianhua Chen Louisiana State University, USA

Henning Christiansen Roskilde University (Co-chair), Denmark William J. Clancey Florida Institute for Human and Machine

Cognition, USA

Luca Console Università di Torino, Italy

Bruno Cremilleux Université de Caen Basse-Normandie, France Juan Carlos Cubero University of Granada (Co-chair), Spain ICAR-CNR, University of Calabria, Italy Ramon Lopez de Mantaras Spanish National Research Council, Spain Università degli Studi di Bari "Aldo Moro",

Ital

Jørgen Fischer NilssonTechnical University of DenmarkVladimir A. FomichovNational Research University, FranceLaura GiordanoUniversity of Piemonte Orientale, ItalyJacek GrekowBialystok University of Technology, Poland

Jerzy Grzymala-Busse University of Kansas, USA Hakim Hacid Alcatel-Lucent Bell Labs Allel Hadiali Université de Rennes 1, France Shoji Hirano Shimane Medical University, Japan Lothar Hotz University of Hamburg, Germany Manfred Jaeger Aalborg University, Denmark Nathalie Japkowicz University of Ottawa, Canada Polish Academy of Sciences, Poland Mieczyslaw Klopotek

Bozena Kostek Patrick Lambrix Rory Lewis

Michael Lowry Donato Malerba Giuseppe Manco

Krzysztof Marasek

Maria José Martín Bautista

Nicolas Marín Ruiz Elio Masciari Paola Mello

Ernestina Menasalvas Ruiz

Neil Murray

Agnieszka Mykowiecka John Mylopoulos Thomas D. Nielsen

Jean-Marc Petit Olivier Pivert

Henri Prade Vijay Raghavan Hiroshi Sakai

Daniel Sanchez Fernandez Dominik Slezak

Jerzy Stefanowski Jaroslaw Stepaniuk Marcin Sydow

Erich Teppan K. Thirunarayan Christel Vrain

Alicja Wieczorkowska

Franz Wotawa

Yivu Yao

Slawomir Zadrozny Włodek Zadrozny

Ning Zhong

Djamel Zighed

Gdansk University of Technology, Poland

Linkoping University, Sweden

University of Colorado at Colorado Springs,

USA

NASA Ames, USA

University of Bari, Italy

Universitá della Calabria, Italy

Polish-Japanese Institute of Information

Technology, Poland

University of Granada, Spain University of Granada, Spain Università della Calabria, Italy University of Bologna, Italy

Universidad Politécnica de Madrid, Spain University at Albany, State University of

New York, USA

Polish Academy of Sciences, Poland University of Toronto, Canada Aalborg University, Denmark

Université de Lyon, France University of Rennes 1, France

National Center for Scientific Research, France University of Louisiana at Lafayette, USA Kyushu Institute of Technology, Japan

University of Granada, Spain University of Warsaw, Poland Poznan Technical University, Poland

Bialystok University of Technology, Poland Polish-Japanese Institute of Information

Technology, Poland

University of Klagenfurt, Germany Wright State University, USA Orleans University, France

Polish-Japanese Institute of Information

Technology, Poland

Graz University of Technology, Poland

University of Regina, Canada

University of North Carolina at Charlotte, USA University of North Carolina at Charlotte, USA

Maebashi Institute of Technology, Japan

University of Lyon 2, France

#### Special Session: Challenges in Text Mining and Semantic Information Retrieval

Piotr Andruszkiewicz Warsaw University of Technology, Poland Robert Bembenik Warsaw University of Technology, Poland

#### X Organization

Philipp Cimiano Bielefeld University, Germany

Tomasz Gambin Warsaw University of Technology, Poland Piotr Gawrysiak Warsaw University of Technology, Poland

Marzena Kryszkiewicz Warsaw University of Technology

(Co-chair), Poland)

Evangelos Milios Dalhousie University, Canada Mikołaj Morzy Poznań University of Technology, Poland

virkolaj violey i oznan Chiversity of Technology, i ola

Robert Nowak University of Technology

Grzegorz Protaziuk Warsaw University of Technology, Poland Henryk Rybinski University of Technology (Co-chair), Poland Lukasz Skonieczny Warsaw University of Technology (Co-chair),

Poland

Jerzy Stefanowski Poznań Technical University, Poland Julian Szymanski Gdansk University of Technology, Poland Krzysztof Walczak Warsaw University of Technology, Poland Wlodek Zadrozny University of North Carolina at Charlotte, USA

# Special Session: Warehousing and OLAPing Complex, Spatial and Spatio-Temporal Data

Michelangelo Ceci University of Bari, Italy

Alfredo Cuzzocrea ICAR-CNR & University of Calabria (Chair),

Italy

Sergio Flesca University of Calabria, Italy
Filippo Furfaro University of Calabria, Italy
Carson Leung University of Manitoba, Canada

#### Additional Reviewers

Agrawal, Ameeta Kapanipathi, Pavan Aiello, Marco Klec, Mariusz

Boulkrinat, Samia Lanotte, Pasqua Fabiana

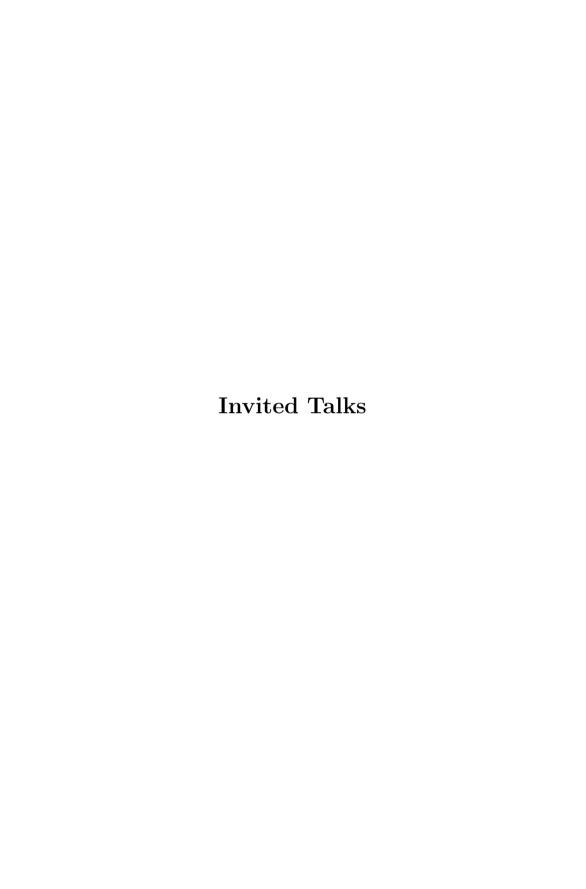
Béchet, Nicolas Liberatore, Paolo

Cancelliere, Rossella Lyu, Siwei Chesani, Federico Pio, Gianvito

Choiref, Zahira
Corby, Olivier
Pusala, Murali
Dragisic, Zlatan
Fumarola, Fabio
Sharif, Mohammad
Control of Con

Fumarola, Fabio Sharif, Mohammad Grisetti, Giorgio Spillane, Sean Guarascio, Massimo Thion, Virginie

Hose, Katja Wang, Wenbo Ivanova, Valentina Zihayat, Morteza



# Exploiting Cultural Specificity in Music Information Research

Xavier Serra

Music Technology Group Universitat Pompeu Fabra, Barcelona xavier.serra@upf.edu

Music Information Research (MIR) is a discipline that aims to understand and model music from an information processing perspective and one of its major challenges relates to the automatic generation of musically meaningful information with which to better describe and exploit audio music recordings. The goal is to integrate and process a variety of data sources, like the actual audio recordings, plus editorial metadata and contextual information, to obtain structured information that is semantically and musically meaningful and that is of use in search, retrieval and discovery tasks [1].

A piece of music is an information entity that makes sense specially within a particular social and cultural context. Its analysis and description has to take that into account and thus the data-driven approaches have to incorporate domain knowledge from that particular context in order to make sense of the available information on that piece of music.

In this presentation I will introduce the research currently being done in CompMusic (http://compmusic.upf.edu), a project funded by the European Research Council that focuses on a number of MIR problems through the study of five music cultures: Hindustani (North India), Carnatic (South India), Turkishmakam (Turkey), Arab-Andalusian (Maghreb), and Beijing Opera (China). We work on the extraction of musically relevant features from audio music recordings related to melody and rhythm, and on the semantic analysis of the contextual information of those recordings [2].

Given that most of the research in MIR has been based on studying the western commercial music of the last few decades, our claim is that the technologies developed have a strong bias towards that music, thus not being appropriate for other music repertories. We want to identify the current limitations and propose information processing approaches that can go beyond those boundaries. For that we selected a few music cultures that had personalities contrasting with the popular western music, that had alive performance practices and strong social and cultural relevance, for which there were musicological and cultural studies, and for which it was feasible to collect sufficient and coherent machine-readable music data. At the same time we wanted to have a diverse set of music repertoires with which to study a variety of new and diverse MIR problems.

A major effort in CompMusic has been the creation of research corpora. The types of data that we have gathered are mainly audio recordings and editorial metadata, which are then complemented with descriptive information about the items we have, and in some cases with music scores and/or lyrics. In order to evaluate our research results we have defined a user scenario and have developed a complete system-level application with which users can interact and with which we can evaluate most of the research results from a user perspective. The system, Dunya (http://dunya.compmusic.upf.edu), is a web-based application to explore music collections aimed at music connoisseurs of the particular music traditions. It uses the technologies developed for melodic and rhythmic description and semantic analysis to navigate through the audio recordings and the information items available. This navigation promotes the discovery of relationships between the different information items.

#### References

- Serra, X., Magas, M., Benetos, E., Chudy, M., Dixon, S., Flexer, A., Gómez, E., Gouyon, F., Herrera, P., Jordà, S., Paytuvi, O., Peeters, G., Schlüter, J., Vinet, H., Widmer, G.: Roadmap for Music Information ReSearch (2013) ISBN: 978-2-9540351-1-6
- Serra, X.: A Multicultural Approach in Music Information Research. In: Int. Soc. for Music Information Retrieval Conf. (ISMIR), pp. 151–156 (2011)

## Big Data Workflows: Issues and Challenges

#### Matthias Jarke

RWTH Aachen University, Informatik 5 & Fraunhofer FIT Ahornstr. 55m 52074 Aachen, Germany jarke@cs.rwth-aachen.de

Abstract. Big Data is often seen as a rather uniform, if not well understood conglomerate of research and practice issues related to massively increased Volume, Velocity, and Variety of data. In reality, there is an enormous diversity of requirements, architectural and algorithmic settings, in which key success factors can range from efficient sensor fusion to rapid query processing and stream mining, to careful semantic-preserving data integration, to aspect such as data protection, provenance maintenance, novel business models, and digital rights management. In the Fraunhofer Big Data [1], twenty-four research institutes from different disciplines in science and engineering have joined forces to explore big data in the six domains of production, logistics, life sciences/healthcare, energy management, security, business and finance.

This presentation will illustrate this variety from the perspective of intelligent, automated workflow assistance in different big data settings. We start with an overview of attempts for automating the integration of heterogeneous structured and semi-structured data, e.g. in corporate as well as cross-organizational multi-database settings [2]. One set of application projects we are currently engaged in comprises different aspects of research data management, focusing on prevention of scientific fraud and traceability, while ensuring correct data ownership and shared understanding. Another important aspect in this context is the evolution analysis of contributions to scientific communities [3], be it open source communities or simply publication and citation networks. As another extreme, we look at highly scalable digital rights policies and management workflows within very large-scale video databases, contrasting automated techniques and "wisdom of the crowd" [4].

#### References

- 1. http://www.bigdata.fraunhofer.de
- Jarke, M., Jeusfeld, M.A., Quix, C.: Data-centric intelligent information integration

   from concepts to automation. Journal of Intelligent Information Systems (to appear, 2014)
- 3. Pham, M.C., Klamma, R., Jarke, M.: Development of computer science disciplines a social network analysis approach. Social Network Analysis and Mining (SNAM) 1(4), 321–340 (2011)

#### XVI M. Jarke

4. Rashed, K., Renzel, D., Klamma, R., Jarke, M.: Community and trust-aware fake media detection. Multimedia Tools and Applications (2012), doi:10.1007/s11042-012-1103-3; Special issue on Multimedia on the Web 2012

# Table of Contents

Complex Networks and Data Stream Mining	
Community Detection by an Efficient Ant Colony Approach Lúcio Pereira de Andrade, Rogério Pinto Espíndola, and Nelson Francisco Favilla Ebecken	1
Adaptive XML Stream Classification Using Partial Tree-Edit Distance	10
RILL: Algorithm for Learning Rules from Streaming Data with Concept Drift	20
Community Detection for Multiplex Social Networks Based on Relational Bayesian Networks	30
Mining Dense Regions from Vehicular Mobility in Streaming Setting  Corrado Loglisci and Donato Malerba	40
Mining Temporal Evolution of Entities in a Stream of Textual  Documents	50
An Efficient Method for Community Detection Based on Formal Concept Analysis	61
Data Mining Methods	
On Interpreting Three-Way Decisions through Two-Way Decisions Xiaofei Deng, Yiyu Yao, and JingTao Yao	73
FHM: Faster High-Utility Itemset Mining Using Estimated Utility Co-occurrence Pruning	83
Automatic Subclasses Estimation for a Better Classification with HNNP	93

A Large-Scale, Hybrid Approach for Recommending Pages Based on Previous User Click Pattern and Content	103
EverMiner Prototype Using LISp-Miner Control Language	113
Local Characteristics of Minority Examples in Pre-processing of Imbalanced Data	123
Visual-Based Detection of Properties of Confirmation Measures Robert Susmaga and Izabela Szczęch	133
Intelligent Systems Applications	
A Recursive Algorithm for Building Renovation in Smart Cities	144
Spike Sorting Based upon PCA over DWT Frequency Band Selection  Konrad Ciecierski, Zbigniew W. Raś, and Andrzej W. Przybyszewski	154
Neural Network Implementation of a Mesoscale Meteorological Model	164
Spectral Machine Learning for Predicting Power Wheelchair Exercise Compliance	174
Mood Tracking of Radio Station Broadcasts	184
Evidential Combination Operators for Entrapment Prediction in Advanced Driver Assistance Systems	194
Influence of Feature Sets on Precision, Recall, and Accuracy of Identification of Musical Instruments in Audio Recordings  Elżbieta Kubera, Alicja A. Wieczorkowska, and Magdalena Skrzypiec	204
Multi-label Ferns for Efficient Recognition of Musical Instruments in Recordings	214

Methodology and Case Study.....

Alexander Smirnov, Tatiana Levashova, Alexey Kashevnik, and

Nikolay Shilov

325

# Textual Data Analysis and Mining

Unsupervised Aggregation of Categories for Document Labelling Piotr Borkowski, Krzysztof Ciesielski, and Mieczysław A. Kłopotek	335
Classification of Small Datasets: Why Using Class-Based Weighting Measures?	345
Improved Factorization of a Connectionist Language Model for Single-Pass Real-Time Speech Recognition	355
Automatic Extraction of Logical Web Lists	365
Combining Formal Logic and Machine Learning for Sentiment Analysis	375
Clustering View-Segmented Documents via Tensor Modeling	385
Searching XML Element Using Terms Propagation Method	395
Special Session: Challenges in Text Mining and Semantic Information Retrieval	
AI Platform for Building University Research Knowledge Base Jakub Koperwas, Łukasz Skonieczny, Marek Kozłowski, Piotr Andruszkiewicz, Henryk Rybiński, and Wacław Struk	405
A Seed Based Method for Dictionary Translation	415
SAUText — A System for Analysis of Unstructured Textual Data Grzegorz Protaziuk, Jacek Lewandowski, and Robert Bembenik	425
Evaluation of Path Based Methods for Conceptual Representation of the Text	435
Łukasz Kucharczyk and Julian Szymański	

XXI

Special Session: Warehousing and OLAPing Complex, Spatial and Spatio-Temporal Data	
Restructuring Dynamically Analytical Dashboards Based on Usage	445
Profiles	445
Enhancing Traditional Data Warehousing Architectures with Real-Time Capabilities	456
Inference on Semantic Trajectory Data Warehouse Using an Ontological	100
Approach	466
Combining Stream Processing Engines and Big Data Storages for Data Analysis	476
ISMIS Posters	
Representation and Evolution of User Profile in Information Retrieval Based on Bayesian Approach	486
Creating Polygon Models for Spatial Clusters	493
Skeleton Clustering by Autonomous Mobile Robots for Subtle Fall Risk Discovery	500
Sonar Method of Distinguishing Objects Based on Reflected Signal Specifics	506
Endowing Semantic Query Languages with Advanced Relaxation Capabilities	512
A Business Intelligence Solution for Monitoring Efficiency of Photovoltaic Power Plants	518

#### XXII Table of Contents

WBPL: An Open-Source Library for Predicting Web Surfing Behaviors	524
Ted Gueniche, Philippe Fournier-Viger, Roger Nkambou, and Vincent S. Tseng	924
Data-Quality-Aware Skyline Queries	530
Neuroscience Rough Set Approach for Credit Analysis of Branchless Banking	536
Collective Inference for Handling Autocorrelation in Network Regression	542
On Predicting a Call Center's Workload: A Discretization-Based Approach	548
Improved Approximation Guarantee for Max Sum Diversification with Parameterised Triangle Inequality	554
Learning Diagnostic Diagrams in Transport-Based Data-Collection  Systems	560
Author Index	567