

Lecture Notes
in Business Information Processing

116

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

Karl Aberer
Ernesto Damiani
Tharam Dillon (Eds.)

Data-Driven Process Discovery and Analysis

First International Symposium, SIMPDA 2011
Campione d'Italia, Italy, June 29 – July 1, 2011
Revised Selected Papers

 Springer

Volume Editors

Karl Aberer
EPFL Lausanne
Lausanne, Switzerland
E-mail: karl.aberer@epfl.ch

Ernesto Damiani
Università degli Studi di Milano
Crema CR, Italy
E-mail: ernesto.damiani@unimi.it

Tharam Dillon
University of Technology
Sydney, NSW, Australia
E-mail: tharam@it.uts.edu.au

ISSN 1865-1348
ISBN 978-3-642-34043-7
DOI 10.1007/978-3-642-34044-4
Springer Heidelberg Dordrecht London New York

e-ISSN 1865-1356
e-ISBN 978-3-642-34044-4

Library of Congress Control Number: 2012948778

ACM Computing Classification (1998): H.4, J.1, H.3

© IFIP International Federation for Information Processing 2012

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 rapid growth of organizational and business processes supported by ICT is making available a huge amount of process data. As a consequence, there is a growing demand for effective and efficient process analysis techniques. The International Symposium on Data-Driven Process Discovery and Analysis (SIMPDA), jointly organized by IFIP WG 2.6 and W.G 2.12, was conceived as a forum where researchers from different scientific communities – to name but a few, process representation and modelling, data warehousing and analysis and data mining – could freely share their insights and discuss them with potential industrial partners in open sessions. The first symposium took place in June 2011 in the unique location of Campione d’Italia, between Italy and Switzerland; it featured a number of advanced keynotes, presentations on recent research, a competitive PhD seminar, and a small number of selected research and industrial demonstrations.

After the symposium, a call for papers was circulated to solicit submissions to this LNBIP volume. The call was open to selected SIMPDA participants – both paper presenters and keynotes – as well as to the research community at large. There were 31 submissions and 11 papers were accepted for publication, covering a wide range of topics, from theoretical issues related to process representation to practical experience in process discovery and analysis.

In the paper “Towards Distributed Collaborative Workflow Management for Mobile Devices”, Anna Kocurova et al. present their approach to implementing a workflow management system supporting content management. The main goal of this work is to study the implications related to scenarios where people can cooperate via mobile systems. The design of a conceptual reference framework for the definition, management and execution of mobile context-aware content-centric workflows is introduced.

The paper by Francesco Arigliano et al., “Monitoring Business Processes in the Networked Enterprise”, addresses the problem of defining frameworks and languages for measuring processes. The paper proposes a framework that enables the specification and measurement of metrics, supporting the definition of relationships among metrics at different organizational levels, providing a way to bind strategic objectives to organizational and technological performance.

The paper by Wil van der Aalst et al., “Towards Improving the Representational Bias of Process Mining”, describes a technique to reduce internal inconsistencies of process models created by process mining techniques. The authors propose a new generation of genetic process discovery algorithms that limit the search space to improve the quality of the results discovered and speed up the search process.

The paper by Claudio Di Ciccio et al., “MailOfMine: Analyzing Mail Messages for Mining Artful Collaborative Processes”, studies the problem of an efficient approach to automatically building, on top of a collection of email messages, a set of workflow models that represent the processes underlying the knowledge workers’ activities.

The paper by Gianluca Demartin et al., “Benchmarking RDF Analytics”, presents a novel benchmark for evaluating and comparing the efficiency of Semantic Web data management systems on analytic queries. The benchmark models a real-world setting derived from the Bologna process and includes a broad set of queries reflecting a large panel of concrete, data-intensive user needs.

The paper by Ioana Ciuciu et al., “Towards Evaluating an Ontology-Based Data Matching Strategy for Retrieval and Recommendation of Security Annotations for Business Process Models”, develops a method to provide semantic support to process modellers during the design of secure business process models.

The paper by Jerzy Korczak et al., “FP-Growth in Discovery of Customer Patterns” describes an algorithm for finding association rules and develops an experiment for evaluating the performances of the algorithm.

The paper by Marcello Leida et al., “Case Study in Process Mining in a Multinational Enterprise” deals with one such case study and describes the issues that were encountered and overcome.

The paper by Rafael Accorsi et al., “Discovering Workflow Changes with Time-Based Trace Clustering”, proposes an algorithm for trace clustering, a vital preliminary step to process mining. The approach includes a time-oriented clustering of process logs that directly reflects the process dynamics.

The paper by Vera Kunzle et al., “Striving for Object-Aware Process Support: How Existing Approaches Fit Together”, introduces the fundamental features of object-aware processes. The proposed approach supports a tighter integration of processes and data along these characteristics.

Finally the paper by Yishuai Lin et al., “Scrum Conceptualization Using K-CRIO Ontology”, presents an ontology that allows the description of specific types of business processes: those that are dedicated to the design of a product. The ontology provides a means for annotating resources, reasoning, monitoring design processes, enabling searches and proactively proposing tips and proper contents.

Putting together a volume like this is always a team effort. We are grateful to the research community that gathered around the symposium and acknowledge the high quality of their research work, which is hopefully reflected in the papers of this volume. Hopefully, the lively discussions that took place at SIMPDA helped authors to improve their work and fostered interesting extensions. We also would like to express our deep appreciation for the referees’ hard work and dedication. Above all, however, heartfelt thanks are due to the authors for submitting the best results of their work to this publication.

August 2012

Karl Aberer
Ernesto Damiani
Tharam Dillon

Symposium Committee

Co-chairs

Karl Aberer	EPFL Lausanne, Switzerland
Ernesto Damiani	Università degli Studi di Milano, Italy
Tharam Dillon	University of Technology Sydney, Australia

Steering Committee

Erich Neuhold	University of Vienna, Austria
Lionel Brunie	Institut National des Sciences Appliquées (INSA) de Lyon, France
Paolo Ceravolo	Università degli Studi di Milano, Italy
Tharam Dillon	Curtin University, Australia
Elizabeth Chang	Curtin University, Australia
Giuseppina Passiante	Università del Salento, Italy

Conference Program Committee

Peter Spyns	Vrije Universiteit Brussel - STAR Lab, Belgium
Daniele Bonetta	Università della Svizzera Italiana, Switzerland
Philippe Cudre-Mauroux	MIT-CSAIL, USA
Maurice Van Keulen	University of Twente, The Netherlands
Avigdor Gal	Israel Institute of Technology, Israel
Mohand-Said Hacid	Université Claude Bernard Lyon 1, France
Angelo Corallo	Università del Salento, Italy
Irene Vanderfeesten	Eindhoven University of Technology, The Netherlands
Rafael Accorsi	University of Freiburg, Germany
Farookh Khadeer Hussain	Curtin University, Australia
Thomas Risse	L3S Research Center, Germany
Wolfgang Klas	University of Vienna, Austria
Davide Storelli	Università del Salento, Italy
Marcello Leida	EBTIC (Etisalat BT Innovation Centre), UAE
Gabriele Ruffatti	Engineering Group, Italy
Jerzy Korczak	Wroclaw University of Economics, Poland
Abder Koukam	University of Technology UTBM, France
Renato Iannella	Semantic Identity, Australia
Manfred Reichert	University of Ulm, Germany
Wei-Chiang Hong	Oriental Institute of Technology, Taiwan (China)
Mustafa Jarrar	Birzeit University, Palestinian Territories

VIII Symposium Committee

Schahram Dustdar	Vienna University of Technology, Austria
Mohamed Achemlal	Orange Labs, France
Jose M. Alcaraz Calero	Hewlett-Packard, UK
Mohamed Mosbah	University of Bordeaux, France
Eduardo Fernández-Medina	University of Castilla-La Mancha, Spain
Meiko Jensen	University of Bochum, Germany
Haris Mouratidis	University of East London, UK
Manfred Reichert	University of Ulm, Germany
Debasis Giri	Haldia Institute of Technology, India
Helen Balinsky	Hewlett-Packard Laboratories, UK
Valentina Emilia Balas	University of Arad, Romania
Antonio Mana Gomez	University of Malaga, Spain
Frédéric Cuppens	Télécom Bretagne, France
Nora Cuppens	Télécom Bretagne, France
Mihaela Cardei	Florida Atlantic University, USA
Eduardo Fernandez	Florida Atlantic University, USA
Andreas Wombacher	University of Twente, The Netherlands
Karima Boudaoud	Ecole Polytechnique de Nice Sophia Antipolis, France
George Spanoudakis	City University of London, UK
Artur Hecker	Télécom ParisTech, France
Etienne Riviere	Université de Neuchâtel, Switzerland
Richard Chbeir	University of Bourgogne, France
Chi Hung	Tsinghua University, China
Anas Abouelkalam	Institut National Polytechnique de Toulouse, France
Luis Soares	Barbosa Universidade do Minho, Portugal
Adrian Pasarariu	Florida Atlantic University, USA
Gregorio Martinez Perez	University of Murcia, Spain
Ioana Georgiana Ciuciu	Free University of Brussels, Belgium
Alfredo Cuzzocrea	Università della Calabria, Italy

Keynote Speakers

Wil van der Aalst	Technische Universiteit Eindhoven, The Netherlands
Florian Kerschbaum	SAP, Germany
Dragan Gasevic	Athabasca University, Canada
Barbara Pernici	Politecnico di Milano, Italy

Table of Contents

Towards Distributed Collaborative Workflow Management for Mobile Devices	1
<i>Anna Kocurova, Samia Oussena, Peter Komisarczuk, Tony Clark, and Dean Kramer</i>	
Monitoring Business Processes in the Networked Enterprise	21
<i>Francesco Arigliano, Devis Bianchini, Cinzia Cappiello, Angelo Corallo, Paolo Ceravolo, Ernesto Damiani, Valeria De Antonellis, Barbara Pernici, Pierluigi Plebani, Davide Storelli, and Claudia Vicari</i>	
Towards Improving the Representational Bias of Process Mining	39
<i>Wil van der Aalst, Joos Buijs, and Boudewijn van Dongen</i>	
MailOfMine – Analyzing Mail Messages for Mining Artful Collaborative Processes	55
<i>Claudio Di Ciccio, Massimo Mecella, Monica Scannapieco, Diego Zardetto, and Tiziana Catarci</i>	
BowlognaBench—Benchmarking RDF Analytics	82
<i>Gianluca Demartini, Iliya Enchev, Marcin Wylot, Joël Gapany, and Philippe Cudré-Mauroux</i>	
Towards Evaluating an Ontology-Based Data Matching Strategy for Retrieval and Recommendation of Security Annotations for Business Process Models	103
<i>Ioana Ciuciu, Yan Tang, and Robert Meersman</i>	
FP-Growth in Discovery of Customer Patterns	120
<i>Jerzy Korczak and Piotr Skrzypczak</i>	
Case Study in Process Mining in a Multinational Enterprise	134
<i>Paul Taylor, Marcello Leida, and Basim Majeed</i>	
Discovering Workflow Changes with Time-Based Trace Clustering	154
<i>Rafael Accorsi and Thomas Stocker</i>	
Striving for Object-Aware Process Support: How Existing Approaches Fit Together	169
<i>Vera Künzle and Manfred Reichert</i>	
Scrum Conceptualization Using K-CRIO Ontology	189
<i>Yishuai Lin, Vincent Hilaire, Nicolas Gaud, and Abderrafaa Koukam</i>	
Author Index	213