Lecture Notes in Business Information Processing 237

Series Editors

Wil van der Aalst
Eindhoven Technical University, Eindhoven, The Netherlands

John Mylopoulos
University of Trento, Povo, 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

More information about this series at http://www.springer.com/series/7911

Paolo Ceravolo · Barbara Russo
Rafael Accorsi (Eds.)

Data-Driven Process Discovery and Analysis

4th International Symposium, SIMPDA 2014
Milan, Italy, November 19–21, 2014
Revised Selected Papers

Springer

Editors
Paolo Ceravolo
Università degli Studi di Milano
Crema
Italy

Rafael Accorsi
Universität Freiburg
Freiburg
Germany

Barbara Russo
Free University of Bozen/Bolzano
Bolzano
Italy

ISSN 1865-1348 ISSN 1865-1356 (electronic)
Lecture Notes in Business Information Processing
ISBN 978-3-319-27242-9 ISBN 978-3-319-27243-6 (eBook)
DOI 10.1007/978-3-319-27243-6

Library of Congress Control Number: 2015955854

Springer Cham Heidelberg New York Dordrecht London
© IFIP International Federation for Information Processing 2015
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.
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.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Preface

The rapid growth of organizational and business process managed via information systems made available a big variety of data that as a consequence created a high demand for making available data analysis techniques more effective and valuable. The fourth edition of the International Symposium on Data-Driven Process Discovery and Analysis (SIMPDA 2014) was conceived to offer a forum where researchers from different communities and the industry can share their insights in this hot new field. The symposium featured a number of advanced keynotes illustrating new approaches as well as presentations on recent research. The goal is to foster exchanges between academic researchers, industry, and a wider audience interested in process discovery and analysis. The event is organized by the IFIP WG 2.6. This year, the symposium was held in Milan, the city of Expo 2015.

The submissions cover theoretical issues related to process representation, discovery, and analysis or provide practical and operational experiences in process discovery and analysis. To improve the quality of the contributions, the symposium fostered discussion during the presentations, giving authors the opportunity to improve their work extending the results presented. For this reason, authors of accepted papers and keynote speakers were invited to submit extended articles to this proceedings volume in the LNBIP series. There were 21 submissions and five papers were accepted for publication.

During this edition the presentations and the discussions frequently focused on the implementation of process-mining algorithms in contexts where the analytical process is fed by data streams. The current selection of papers underlines the most relevant challenges that were identified and proposes novel solutions and approaches facing these challenges.

In the first paper "Discovery of Frequent Episodes in Event Logs," Maikel Leemans and Wil M.P. van der Aalst present an approach to detect frequently occurring episodes, i.e., a partially ordered collection of events, in an event log. Moreover, this work uses comparison with existing discovery algorithms to demonstrate that episode mining benefits from exploiting parameters that are encoding the process behavior.

The second paper by Bart Hompes et al., "Finding Suitable Activity Clusters for Decomposed Process Discovery," focused on decomposition as a strategy for the parallelization of process-mining algorithms. Analysis shows that although the decomposition step takes a relatively small amount of time, it is of key importance in finding a high-quality process model and for the computation time required to discover the individual parts. Moreover, the authors propose three metrics that can be used to assess the quality of a decomposition, before using it to discover a model or check conformance.

The third paper by Mahdi Alizadeh, Massimiliano de Leoni, and Nicola Zannone, "History-Based Construction of Alignments for Conformance Checking: Formalization and Implementation?", proposes an approach to automatically define the cost function for alignment-based conformance-checking techniques. Based on the information

extracted from the past process executions, the cost function is derived relying on objective factors and thus enabling the construction of probable alignments, i.e., alignments that provide probable explanations of nonconformity.

The fourth paper, by David Redlich et al., "Dynamic Constructs Competition Miner – Occurrence vs. Time-Based Ageing," extends a divide-and-conquer algorithm for discovering block-structured processes from event logs possibly consisting of exceptional behavior. In particular, this paper proposes a set of modifications to enable dynamic business process discovery at run-time from a stream of events.

The fifth paper by Marco Anisetti et al., "Trustworthy Cloud Certification: A Model-Based Approach," discusses the problem of tracking and assessing the behavior of cloud services/processes. One of the main limitations of existing approaches is the uncertainty introduced by the cloud on the validity and correctness of existing certificates. The authors present a trustworthy cloud certification approach by continuously verifying the correctness of the service model at the basis of certification activities against real and synthetic service execution traces.

We gratefully acknowledge the strong research community that gathered around the research problems related to process data analysis and the high quality of their research work, which is hopefully reflected in the papers of this volume. We would also like to express our deep appreciation for the reviewers hard work and dedication. Above all, thanks are due the authors for submitting the best results of their work to the Symposium on Data-Driven Process Discovery and Analysis.

We are very grateful to the Università degli Studi di Milano and to IFIP for their financial support, and to the University of Freiburg and the Free University of Bozen/Bolzano.

October 2015

Paolo Ceravolo
Rafael Accorsi
Barbara Russo

Organization

Conference Co-chairs

Paolo Ceravolo	Università degli Studi di Milano, Italy
Rafael Accorsi	Universität Freiburg, Germany
Barbara Russo	Free University of Bozen/Bolzano, Italy

Advisory Board

Karl Aberer	EPFL, Switzerland
Ernesto Damiani	Università degli Studi di Milano, Italy
Tharam Dillon	La Trobe University, Australia
Dragan Gasevic	Athabasca University, Canada
Marcello Leida	EBTIC (Etisalat BT Innovation Centre), UAE
Erich Neuhold	University of Vienna, Austria
Maurice van Keulen	University of Twente, The Netherlands

PhD Award Committee

Paolo Ceravolo	Università degli Studi di Milano, Italy
Marcello Leida	EBTIC (Etisalat BT Innovation Centre), UAE
Gregorio Piccoli	Zucchetti spa, Italy

Web Chair and Publicity Chair

Fulvio Frati	Università degli Studi di Milano, Italy

Program Committee

Irene Vanderfeesten	Eindhoven University of Technology, The Netherlands
Maurice Van Keulen	University of Twente, The Netherlands
Manfred Reichert	University of Ulm, Germany
Schahram Dustdar	Vienna University of Technology, Austria
Mohamed Mosbah	University of Bordeaux, France
Meiko Jensen	Ruhr-University Bochum, Germany
Helen Balinsky	Hewlett-Packard Laboratories, UK
Valentina Emilia Balas	University of Arad, Romania
Karima Boudaoud	Ecole Polytechnique de Nice Sophia Antipolis, France
George Spanoudakis	City University London, UK
Richard Chbeir	University of Bourgogne, France
Gregorio Martinez Perez	University of Murcia, Spain

Ebrahim Bagheri	Ryerson University, Canada
Jan Mendling	Vienna University of Economics and Business, Austria
Farookh Hussain	University of Technology Sydney, Australia
Marcello Leida	Ebtic (Etisalat BT Innovation Centre), UAE
Wil Van Der Aalst	Technische Universiteit Eindhoven, The Netherlands
Ronald Maier	University of Innsbruck, Austria
Chintan Amrit	University of Twente, The Netherlands
Marco Montali	Free Unviersity of Bozen - Bolzano, Italy
Elizabeth Chang	University New South Wales, Australia
Peter Spyns	Flemish Government, Belgium
Angelo Corallo	University of Salento, Italy
Antonio Maña Gómez	University of Málaga, Spain
Mustafa Jarrar	Birzeit University, Palestinian Territory
Isabella Seeber	University of Innsbruck, Austria
Chi Hung	Tsinghua University, China
Alessandra Toninelli	Engineering Group, Italy
Haris Mouratidis	University of Brighton, UK
Abder Koukam	University of Technology, UTBM, France
Fabrizio Maria Maggi	University of Tartu, Estonia
Massimiliano De Leoni	Eindhoven TU, Netherlands
Edgar Weippl	TU Vienna, Austria
Pnina Soffer	University of Haifa, Israel
Jianmin Wang	Tsinghua University Beijing, China
Minseok Song	UNIST, South Korea
Roland Rieke	Fraunhofer SIT, Germany
Josep Carmona	UPC - Barcelona, Spain
Mark Strembeck	WU Vienna, Austria
Matthias Weidlich	Imperial College, UK
Mohamed Mosbah	University of Bordeaux
Maria Leitner	University of Vienna, Austria
Benoit Depaire	University of Hasselt, Belgium
Barbara Weber	University of Innsbruck, Austria
Babiga Birregah	University of Technology of Troyes, France

Contents

Discovery of Frequent Episodes in Event Logs 1
Maikel Leemans and Wil M.P. van der Aalst

Finding Suitable Activity Clusters for Decomposed Process Discovery 32
B.F.A. Hompes, H.M.W. Verbeek, and W.M.P. van der Aalst

History-Based Construction of Alignments for Conformance Checking: Formalization and Implementation 58
Mahdi Alizadeh, Massimiliano de Leoni, and Nicola Zannone

Dynamic Constructs Competition Miner - Occurrence- vs. Time-Based Ageing 79
David Redlich, Thomas Molka, Wasif Gilani, Gordon Blair, and Awais Rashid

Trustworthy Cloud Certification: A Model-Based Approach 107
Marco Anisetti, Claudio A. Ardagna, Ernesto Damiani, and Nabil El Ioini

Author Index 123