Lecture Notes in Computer Science 7608

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Liliana Calderón-Benavides
Cristina González-Caro
Edgar Chávez Nivio Ziviani (Eds.)

String Processing and Information Retrieval

19th International Symposium, SPIRE 2012
Cartagena de Indias, Colombia, October 21-25, 2012
Proceedings

Springer

Volume Editors

Liliana Calderón-Benavides
Cristina González-Caro
Universidad Autónoma de Bucaramanga
Information Technologies Research Group
Bucaramanga, Colombia
E-mail: {mcalderon, cgonzalc}@unab.edu.co

Edgar Chávez
Universidad Michoacana
School of Physics and Mathematics
Edificio B, Ciudad Universitaria
Morelia, México 58000, Mexico
E-mail: elchavez@umich.mx

Nivio Ziviani
Universidade Federal de Minas Gerais
Department of Computer Science
Av. Antonio Carlos 6627, Belo Horizonte 31270-010, MG, Brazil
E-mail: nivio@dcc.ufmg.br

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-34108-3 e-ISBN 978-3-642-34109-0
DOI 10.1007/978-3-642-34109-0
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: Applied for

CR Subject Classification (1998): H.3, J.3, H.2.8, I.5, I.2.7, H.4

LNCS Sublibrary: SL 1 – Theoretical Computer Science and General Issues

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 the 19th International Symposium on String Processing and Information Retrieval (SPIRE), held in Cartagena de Indias, Colombia, during October 21–25, 2012. SPIRE 2012 was organized in tandem with the 8th Latin American Web Congress (LA-WEB), with both conferences sharing a common day in Web Retrieval.

In the 19 years since its first edition back in 1993, SPIRE has become the reference meeting for the interdisciplinary community of researchers whose activity lies at the crossroads of string processing and information retrieval. The first four editions of this event concentrated mainly on string processing, and were held in South America under the title "South American Workshop on String Processing" (WSP) in 1993 (Belo Horizonte, Brazil), 1995 (Valparaiso, Chile), 1996 (Recife, Brazil), and 1997 (Valparaiso, Chile). WSP was renamed as SPIRE in 1998 (Santa Cruz, Bolivia) when the scope of the event was broadened to include also information retrieval. The change was motivated by the increasing relevance of information retrieval and its close interrelationship with the general area of string processing. From 1999 to 2007, the venue of SPIRE alternated between South / Latin America (odd years) and Europe (even years), with Cancun, Mexico, in 1999, A Coruña, Spain, in 2000, Laguna de San Rafael, Chile, in 2001, Lisbon, Portugal, in 2002, Manaus, Brazil, in 2003, Padova, Italy, in 2004, Buenos Aires, Argentina, in 2005, Glasgow, UK, in 2006, and Santiago, Chile, in 2007. This pattern was broken when SPIRE 2008 was held in Melbourne, Australia, but it was restarted in 2009 when the venue was Saariselk, Finland, Los Cabos, Mexico, in 2010, followed by Pisa, Italy, in 2011.

The SPIRE 2012 call for papers resulted in the submission of 81 papers. Each submitted paper was reviewed by at least three of the 40 members of the Program Committee (PC), which eventually engaged in discussions coordinated by the two PC Chairmen in case of lack of consensus. We believe this resulted in a very accurate selection of the truly best submitted papers. As a result, 26 long papers and 13 short papers were accepted, and are published in these proceedings.

The program of SPIRE 2012 started on October 21 with three tutorials providing in-depth coverage of topics in string processing ("Space-Efficient Data Structures," by Francisco Claude and Gonzalo Navarro) and information retrieval ("Evaluation Metrics for Information Access," by Enrique Amig and Julio Gonzalo and "Information Dissemination in Social Networks," by Aristides Gionis). On October 22, SPIRE 2012 hosted two workshops, i.e., the 7th Workshop on Compression, Text, and Algorithms (WCTA) and the Workshop on Algorithmic Analysis of Biological Data (WAABD). On the following three days the main conference featured keynote speeches by Amihood Amir, Ricardo Baeza-Yates, and Ian H. Witten, plus the presentations of the 26 full papers and

13 short papers. A Best Paper Award and a Best Student Paper Award were also assigned.

We would like to take the opportunity to thank Google, Yahoo! Research, and Universidad Autónoma de Bucaramanga. We would also like to thank everybody involved in making SPIRE 2012 such an exciting event. Specifically, we would like to thank all conference, tutorial, and workshop participants and presenters, who provided a fascinating one-week program of high-quality presentations and intensive discussions. Thanks also to all the members of the PC and to the additional reviewers, who went to great lengths to ensure the high quality of this conference. We are specially grateful to EasyChair for saving us a lot of work and for providing timely support.

Furthermore, we would like to thank all the members of the local organizing team at the Universidad Autónoma de Bucaramanga (UNAB). Particularly, we would like to thank Fabrizio Silvestri, who acted as Tutorials Chair, Gonzalo Navarro, who acted as Workshops Chair, Luz Emilia Jimenez, who gave us support in local arrangements, the IT team from UNAB who designed the official image and website of the symposium, and to all the student volunteers. They all made tremendous efforts to make sure that this event became an exciting and enjoyable one. It is due to them that the organization of SPIRE 2012 was a pleasure.

October 2012

Liliana Calderón-Benavides
Edgar Chávez
Cristina González-Caro
Nivio Ziviani

Organization

Program Committee

Giambattista Amati	Fondazione Ugo Bordoni, Italy
Amihood Amir	Bar Ilan University and Johns Hopkins University, Israel/USA
Ricardo Baeza-Yates	Yahoo! Research
Paolo Boldi	Università degli Studi di Milano, Italy
Liliana Calderón-Benavides	Universidad Autónoma de Bucaramanga, Colombia
Jamie Callan	Carnegie Mellon University, USA
Edgar Chavez	Universidad Michoacana, Mexico
Francisco Claude	University of Waterloo, Canada
Cesar A. Collazos	Universidad del Cauca, Colombia
Fabio Crestani	University of Lugano, Switzerland
Marco Cristo	Universidade Federal do Amazonas, Brazil
Maxime Crochemore	Kings College London and Université Paris-Est, UK/France
Bruce Croft	University of Massachusetts Amherst, USA
Edleno Silva De Moura	Universidade Federal do Amazonas, Brazil
Marcos Goncalves	Universidade Federal do Minas Gerais, Brazil
Cristina González-Caro	Universidad Autónoma de Bucaramanga, Colombia
Concettina Guerra	University of Padova and Georgia Tech, Italy/USA
Jan Holub	Czech Technical University in Prague, Czech Republic
Lucian Ilie	University of Western Ontario, Canada
Costas Iliopoulos	King's College London, Uk
Shen Jialie	Singapore Management University, Singapore
Gregory Kucherov	CNRS/LIGM, France
Alberto Laender	Universidade Federal de Minas Gerais, Brazil
Mounia Lalmas	Yahoo Research
Moshe Lewenstein	Bar Ilan University, Israel
Alistair Moffat	University of Melbourne, Australia
Veli Mäkinen	University of Helsinki, Finland
Gonzalo Navarro	University of Chile, Chile
Laxmi Parida	IBM T.J. Watson Research Center, USA

Kunsoo Park	Seoul National University, Korea
Marco Pellegrini	Institute for Informatics and Telematics of C.N.R., Italy
Yoan Pinzon	National University of Colombia - ALGOS UN, Colombia
Simon Puglisi	Royal Melbourne Institute of Technology, Australia
Berthier Ribeiro-Neto	Google and UFMG, Brazil
Luis M. S. Russo	IST / INESC-ID, Portugal
Rahul Shah	Louisiana State University, USA
Torsten Suel	Yahoo! Research
Esko Ukkonen	University of Helsinki, Finland
Adriano Veloso	UFMG, Brazil
Jeff Vitter	University of Kansas, USA
Nivio Ziviani	Universidade Federal de Minas Gerais, Brazil

Additional Reviewers

Alatabi, Ali
Badkobeh, Golnaz
Baier, Jan
Barros, Evandrino
Barton, Carl
Berlt, Klessius
Brandão, Michele
Carvalho, André
Cheng, Zhiyong
Christou, Michalis
Cortez, Eli
Dimopoulos, Constantinos
Fariña, Antonio
Fernandes, David
Fernandes, Francisco
Ferreira, Anderson
Francisco, Alexandre P.
Fredriksson, Kimmo
Frousios, Kimon
Gagie, Travis
Gerani, Shima
Gottesman-Gelley, Bluma
Gupta, Varun
Haiminen, Niina
He, Dan
Inches, Giacomo
Jiang, Wei
Kazi, Serizhan
Kim, Yubin
Kolpakov, Roman
Konow, Roberto
Kubica, Marcin
Kulkarni, Anagha
Ladra, Susana
Landau, Gad
Liptak, Zsuzsanna
Lonardi, Stefano
Mahdabi, Parvaz
Markov, Ilya
Miranda, Eulanda dos Santos
Mostafa, Keikha
Nekrich, Yakov
Nepomnyachiy, Sergey
Nicolas, Francois
Parama, Jose R.
Petri, Matthias
Pisanti, Nadia
Pissis, Solon
Polishchuk, Valentin
Porat, Ely
Rodrigues, Kaio
Roma, Nuno

Salles, Thiago
Salmela, Leena
Sheng, Cheng
Sirén, Jouni
Souza, Jucimar
Thankachan, Sharma
Tsur, Dekel
Tyczynski, Wojciech
Utro, Filippo
Vahabi, Hossein
Välimäki, Niko
Wagner Rodrigues, Kaio
Will, Sebastian
Zhao, Le

Table of Contents