

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Alfred Kobsa

University of California, Irvine, CA, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

University of Dortmund, Germany

Madhu Sudan

Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max-Planck Institute of Computer Science, Saarbruecken, Germany

Konstantin Avrachenkov Debora Donato
Nelly Litvak (Eds.)

Algorithms and Models for the Web-Graph

6th International Workshop, WAW 2009
Barcelona, Spain, February 12-13, 2009
Proceedings

Volume Editors

Konstantin Avrachenkov
INRIA Sophia Antipolis
2004 Route des Lucioles, 06902 Sophia Antipolis, France
E-mail: k.avrachenkov@sophia.inria.fr

Debora Donato
Yahoo! Research, Barcelona
Ocata 1, 1st floor, 08003 Barcelona, Spain
E-mail: donato@yahoo-inc.com

Nelly Litvak
University of Twente
Faculty of Electrical Engineering, Mathematics and Computer Science
P.O. Box 217, 7500 AE Enschede, The Netherlands
E-mail: n.litvak@ewi.utwente.nl

Library of Congress Control Number: 2008943853

CR Subject Classification (1998): F.2, G.2, H.4, H.3, C.2, H.2.8, E.1

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

ISSN 0302-9743
ISBN-10 3-540-95994-7 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-95994-6 Springer Berlin Heidelberg New York

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.

springer.com

© Springer-Verlag Berlin Heidelberg 2009
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12611911 06/3180 5 4 3 2 1 0

Preface

This volume constitutes the refereed proceedings of the 6th Workshop on Algorithms and Models for the Web Graph, WAW 2009, held in Barcelona in February 2009.

The World Wide Web has become part of our everyday life, and information retrieval and data mining on the Web are now of enormous practical interest. The algorithms supporting these activities combine the view of the Web as a text repository and as a graph, induced in various ways by links among pages, links among hosts, or other similar networks. We also witness an increasing role of the second-generation Web-based applications Web 2.0, such as social networking sites and wiki sites.

The workshop program consisted of 14 regular papers and two invited talks. The invited talks were given by Ravi Kumar (Yahoo! Research, USA) and José Fernando Mendes (University of Aveiro, Portugal). The regular papers went through a thorough review process. The workshop papers were naturally clustered in three sections: “Graph Models for Complex Networks,” “PageRank and Web Graph” and “Social Networks and Search.” The first section lays a foundation for theoretical and empirical analysis of the Web graph and Web 2.0 graphs. The second section analyzes random walks on the Web and Web 2.0 graphs and their applications. It is interesting to observe that the PageRank algorithm finds new exciting applications beyond now classical Web page ranking. Nowadays social networks are among the most popular applications on the Web. The third section of the workshop program is devoted to the design and performance evaluation of the algorithms for social networks. We hope that the papers presented at WAW 2009 will stimulate further development of Web 2.0 applications and deepen our understanding of the World Wide Web evolution.

We would like to thank the General Chair, Andrei Broder, and the Program Committee members for their time and effort which resulted in a very high quality program. We would like to thank the organizers of WAW 2007, Fan Chung and Anthony Bonato, for their valuable advice. We also would like to thank the Springer LNCS editorial team, particularly Alfred Hofmann, Ursula Barth and Anna Kramer, for their advice and prompt help. We thank Yana Volkovich very much for the design and continuing support of the WAW 2009 website. We would like to thank cordially our industrial sponsors Yahoo! Inc. and Google Inc. and our technical sponsors INRIA, University of Twente, Pompeu Fabra University, and Barcelona Media - Innovation Centre. Last but not least, we would like to extend our thanks to all the authors for their high-quality scientific contribution.

February 2009

Konstantin Avrachenkov
Debora Donato
Nelly Litvak

Organization

Executive Committee

Conference Chair	Andrei Broder (Yahoo! Research, USA)
Program Committee	
Co-chair	Konstantin Avrachenkov (INRIA, France)
Program Committee	
Co-chair	Debora Donato (Yahoo! Research, Spain)
Program Committee	
Co-chair	Nelly Litvak (University of Twente, The Netherlands)

Organizing Committee

Ricardo Baeza-Yates	Yahoo! Research, Spain
Debora Donato	Yahoo! Research, Spain
Mari Carmen Marcos	Pompeu Fabra University, Spain
Yana Volkovich	University of Twente, The Netherlands

Program Committee

Paolo Boldi	University of Milan, Italy
Anthony Bonato	Ryerson University, Canada
Guido Caldarelli	Centre Statistical Mechanics and Complexity CNR-INFN, Italy
Fan Chung Graham	University of California, San Diego, USA
Vladimir Dobrynin	St. Petersburg State University, Russia
Jeannette Janssen	Dalhousie University, Canada
Ravi Kumar	Yahoo! Research, USA
Amy N. Langville	College of Charleston, USA
Stefano Leonardi	Sapienza University of Rome, Italy
David Liben-Nowell	Carleton College, USA
Mark Manasse	Microsoft Research, USA
Kevin McCurley	Google Inc., USA
Igor Nekrestyanov	St. Petersburg State University, Russia
Remco van der Hofstad	TU Eindhoven, The Netherlands
Laurent Viennot	INRIA, France
Sebastiano Vigna	University of Milan, Italy
Dorothea Wagner	Karlsruhe University, Germany
Walter Willinger	AT&T Research, USA
Alexander Zelikovsky	Georgia State University, USA

Sponsoring Institutions

Pompeu Fabra University

Google Inc.

Yahoo! Research

Table of Contents

Graph Models for Complex Networks

Information Theoretic Comparison of Stochastic Graph Models: Some Experiments.....	1
<i>Kevin J. Lang</i>	
Approximating the Number of Network Motifs.....	13
<i>Mira Gonen and Yuval Shavitt</i>	
Finding Dense Subgraphs with Size Bounds	25
<i>Reid Andersen and Kumar Chellapilla</i>	
The Giant Component in a Random Subgraph of a Given Graph	38
<i>Fan Chung, Paul Horn, and Linyuan Lu</i>	
Quantifying the Impact of Information Aggregation on Complex Networks: A Temporal Perspective	50
<i>Fernando Mourão, Leonardo Rocha, Lucas Miranda, Virgílio Almeida, and Wagner Meira Jr.</i>	

PageRank and Web Graph

A Local Graph Partitioning Algorithm Using Heat Kernel Pagerank....	62
<i>Fan Chung</i>	
Choose the Damping, Choose the Ranking?	76
<i>Marco Bressan and Enoch Peserico</i>	
Characterization of Tail Dependence for In-Degree and PageRank	90
<i>Nelly Litvak, Werner Scheinhardt, Yana Volkovich, and Bert Zwart</i>	
Web Page Rank Prediction with PCA and EM Clustering	104
<i>Polyxeni Zacharouli, Michalis Titsias, and Michalis Vazirgiannis</i>	
Permuting Web Graphs	116
<i>Paolo Boldi, Massimo Santini, and Sebastiano Vigna</i>	

Social Networks and Search

A Dynamic Model for On-Line Social Networks	127
<i>Anthony Bonato, Noor Hadi, Paul Horn, Paweł Pralat, and Changping Wang</i>	

TC-SocialRank: Ranking the Social Web.....	143
<i>Antonio Gulli, Stefano Cataudella, and Luca Foschini</i>	
Exploiting Positive and Negative Graded Relevance Assessments for Content Recommendation.....	155
<i>Maarten Clements, Arjen P. de Vries, and Marcel J.T. Reinders</i>	
Cluster Based Personalized Search	167
<i>Hyun Chul Lee and Allan Borodin</i>	
Author Index	185