

Lecture Notes in Artificial Intelligence 6904

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

Martin Atzmueller Andreas Hotho
Markus Strohmaier Alvin Chin (Eds.)

Analysis of Social Media and Ubiquitous Data

International Workshops
MSM 2010, Toronto, ON, Canada, June 13, 2010, and
MUSE 2010, Barcelona, Spain, September 20, 2010
Revised Selected Papers

Series Editors

Randy Goebel, University of Alberta, Edmonton, Canada
Jörg Siekmann, University of Saarland, Saarbrücken, Germany
Wolfgang Wahlster, DFKI and University of Saarland, Saarbrücken, Germany

Volume Editors

Martin Atzmueller
University of Kassel, Germany
E-mail: atzmueller@cs.uni-kassel.de

Andreas Hotho
University of Würzburg, Germany
E-mail: hotho@informatik.uni-wuerzburg.de

Markus Strohmaier
Graz University of Technology, Austria
E-mail: markus.strohmaier@tugraz.at

Alvin Chin
Nokia Research Center, Beijing, China
E-mail: alvin.chin@nokia.com

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-23598-6 e-ISBN 978-3-642-23599-3
DOI 10.1007/978-3-642-23599-3
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011934915

CR Subject Classification (1998): I.2, H.5.3, H.5, C.2, K.4, K.5, K.6

LNCS Sublibrary: SL 7 – Artificial Intelligence

© Springer-Verlag Berlin Heidelberg 2011

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

In the last decade, the reach of computational systems has dramatically expanded both in breadth and depth. This development has led computational devices and applications to permeate societal and social systems in an unprecedented manner.

Today, an increasing entwinement of social phenomena, ubiquitous data, and computational processes can be observed in many domains and contexts, including social media, online social networking, and mobile computing. Such systems, in which social, ubiquitous, and computational processes are interdependent and tightly interwoven, can be characterized as distributed social – computational systems, i.e., integrated systems of people, sensors and computers. Typically, the properties of such systems can be considered to be emergent, which means (a) they are influenced by a combination of social phenomena, algorithmic computation, and ubiquitous data and (b) they are usually beyond the direct control of system engineers. In such systems, potentially critical system properties emerge through social adoption and usage.

Therefore, understanding and engineering social – computational systems requires novel approaches and new techniques to system analysis and engineering. This book sets out to explore this emerging space by presenting a number of current approaches and early important work addressing selected aspects of this problem. The individual contributions of this book represent the first steps in this direction, focusing on problems related to the modeling and mining of social and ubiquitous computational systems. Methods for mining, modeling, and development can help to advance our understanding of the dynamics and structures inherent to these systems, and can help to make social – computational systems and ubiquitous data amenable to deeper quantitative analysis.

The papers presented in this book are revised and significantly extended versions of papers submitted to two related workshops: The Modeling Social Media Workshop (MSM 2010) that was held on June 13, 2010 in conjunction with the 21st ACM Conference on Hypertext and Hypermedia (Hypertext 2010), in Toronto, Canada, and the Mining Ubiquitous and Social Environments (MUSE 2010) International Workshop, which was held on September 20, 2010 in conjunction with the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD 2010) in Barcelona, Spain. In the following, we briefly discuss the themes of those two workshops in more detail.

Social Media: Social media applications such as blogs, microblogs, wikis, news aggregation sites, and social tagging systems have pervaded the Web and have transformed the way people communicate and interact with each other online. In order to understand and effectively design social media systems, we need to develop models that are capable of reflecting their complex, multi-faceted

socio-technological nature. Modeling social media applications enables us to understand and predict their evolution, explain their dynamics, or to describe their underlying social – computational mechanics.

Ubiquitous Data: Ubiquitous data require novel analysis methods including new methods for data mining and machine learning. Unlike in traditional data mining scenarios, data do not emerge from a small number of (heterogeneous) data sources, but potentially from hundreds to millions of different sources. As there is only minimal coordination, these sources can overlap or diverge in any possible way. In typical ubiquitous settings, the mining system can be implemented inside the small devices and sometimes on central servers, for real-time applications, similar to common mining approaches. Steps into this new and exciting application area are the analysis of the collected new data, the adaptation of well-known data mining and machine learning algorithms, and finally the development of new algorithms. The advancement of such algorithms and their application in social and ubiquitous settings is one of the core contributions of this book.

Considering these two workshop themes, the papers contained in this volume form a starting point for bridging the gap between both worlds: Both social media applications and ubiquitous systems benefit from modeling aspects, either at the system level, or for providing a sound data basis for further analysis and mining options. On the other hand, data analysis and data mining can provide novel insights and thus similarly enhance and support modeling prospects. In “A Framework for Mobile User Activity Logging,” Wolfgang Woerndl, Alexander Manhardt, Florian Schulze, and Vivian Prinz provide a unified approach for collecting user activity data on mobile devices for user modeling in social computational and ubiquitous systems. In “Intentional Modeling of Social Media Design Knowledge for Government – Citizen Communication,” Andrew Hilts and Eric Yu present how the agent-oriented modeling framework *i** can be applied to analyze the impact of different social media configurations on the goals and relationships of the actors involved. In “Grooming Analysis—Modeling the Social Dynamics of Online Discussion Groups,” Else Nygren presents results from an empirical study of social interactions (in particular: grooming) in a social – computational system.

Next, the chapter “Exploring Gender Differences in Member Profiles of an Online Dating Site Across 35 Countries,” Slava Kisilevich and Mark Last describe the construction of classification models for characterizing gender differences in social networking sites, specifically online dating sites for different countries stressing both modeling and mining aspects. In “Community Assessment Using Evidence Networks,” Folke Mitzlaff, Martin Atzmueeller, Dominik Benz, Andreas Hotho, and Gerd Stumme present a community assessment approach using evidence networks of user activities; the experiments show that (implicit) evidence networks are well suited for consistent relative community ratings, for evaluation and comparison of mined community structures.

The work “Towards Adjusting Mobile Devices to User’s Behaviour” by Peter Fricke, Felix Jungermann, Katharina Morik, Nico Piatkowski, Olaf Spinczyk,

Marco Stolpe, and Jochen Streicher discusses the optimization of mobile (and ubiquitous) devices with respect to the behavior of users. The paper “Bayesian Networks to Predict Data Mining Algorithm Behavior in Ubiquitous Environments” by Aysegul Cayci, Santiago Eibe, Yucel Saygin, and Ernestina Menasalvas describes an approach for parameter estimation and method adaptation in the context of ubiquitous environments with limited resources. Finally, the paper “Online and Offline Trend Cluster Discovery in Spatially Distributed Data Streams” by Anna Ciampi, Annalisa Appice, and Donato Malerba discusses an algorithm for interleaving spatial clustering and trend discovery, with a broad application scope.

It is the hope of the editors that this book (a) catches the attention of an audience interested in recent problems and advancements in the fields of social media, online social networks, and ubiquitous data and (b) helps to spark a conversation on new problems related to the design and analysis of ubiquitous social – computational systems.

We want to thank our reviewers for their careful help in selecting and improving the provided submissions.

June 2011

Martin Atzmueller
Andreas Hotho
Markus Strohmaier
Alvin Chin

Organization

Program Committee

Martin Atzmueller	University of Kassel, Germany
Ulf Brefeld	Yahoo! Research, Spain
Jordi Cabot	INRIA-École des Mines de Nantes, France
Alvin Chin	Nokia Research Center, China
Marco De Gemmis	University of Bari, Italy
Wai-Tat Fu	University of Illinois, USA
Daniel Gayo Avello	University of Oviedo, Spain
Tad Hogg	Hewlett Packard Laboratories, USA
Andreas Hotho	University of Würzburg, Germany
Thomas Kannampallil	University of Texas, USA
Katharina Morik	TU Dortmund, Germany
Ion Muslea	Language Weaver, Inc., USA
Harald Sack	Hasso-Plattner-Institut, Universität Potsdam, Germany
Sergej Sizov	University of Koblenz-Landau, Germany
Marc Smith	Connected Action Consulting Group, USA
Markus Strohmaier	Graz University of Technology, Austria

Additional Reviewer

Hercher, Johannes

Table of Contents

Logging User Activities and Sensor Data on Mobile Devices	1
<i>Wolfgang Woerndl, Alexander Manhardt, Florian Schulze, and Vivian Prinz</i>	
Intentional Modeling of Social Media Design Knowledge for Government-Citizen Communication	20
<i>Andrew Hiltz and Eric Yu</i>	
Grooming Analysis Modeling the Social Interactions of Online Discussion Groups	37
<i>Else Nygren</i>	
Exploring Gender Differences in Member Profiles of an Online Dating Site Across 35 Countries	57
<i>Slava Kisilevich and Mark Last</i>	
Community Assessment Using Evidence Networks	79
<i>Folke Mitzlaff, Martin Atzmueller, Dominik Benz, Andreas Hotho, and Gerd Stumme</i>	
Towards Adjusting Mobile Devices to User's Behaviour	99
<i>Peter Fricke, Felix Jungermann, Katharina Morik, Nico Piatkowski, Olaf Spinczyk, Marco Stolpe, and Jochen Streicher</i>	
Bayesian Networks to Predict Data Mining Algorithm Behavior in Ubiquitous Computing Environments	119
<i>Aysegul Cayci, Santiago Eibe, Ernestina Menasalvas, and Yucel Saygin</i>	
Online and Offline Trend Cluster Discovery in Spatially Distributed Data Streams	142
<i>Anna Ciampi, Annalisa Appice, and Donato Malerba</i>	
Author Index	163