

Lecture Notes in Artificial Intelligence 7472

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 Alvin Chin
Denis Helic Andreas Hotho (Eds.)

Modeling and Mining Ubiquitous Social Media

International Workshops
MSM 2011, Boston, MA, USA, October 9, 2011
and MUSE 2011, Athens, Greece, September 5, 2011
Revised Selected Papers



Springer

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, Knowledge and Data Engineering Group
Wilhelmshöher Allee 73, 34121 Kassel, Germany
E-mail: atzmueller@cs.uni-kassel.de

Alvin Chin
Nokia Research Center, Mobile Social Networking Group
Beijing 100176, China
E-mail: alvin.chin@nokia.com

Denis Helic
Graz University of Technology, Faculty of Computer Science
Inffeldgasse 21a, 8010 Graz, Austria
E-mail: dhelic@tugraz.at

Andreas Hotho
University of Würzburg, Data Mining and Information Retrieval Group
Am Hubland, 97074 Würzburg, Germany
E-mail: hotho@informatik.uni-wuerzburg.de

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-33683-6 e-ISBN 978-3-642-33684-3
DOI 10.1007/978-3-642-33684-3
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012947587

CR Subject Classification (1998): I.2.6, H.3.3-5, H.3.7, H.5.3-4, H.2.8, H.2.4, C.5.3, I.2.9, K.4.4, K.6.5, C.2.4

LNCS Sublibrary: SL 7 – Artificial Intelligence

© Springer-Verlag Berlin Heidelberg 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 emergence of ubiquitous computing has started to create new environments consisting of small, heterogeneous, and distributed devices that foster the social interaction of users in several dimensions. Similarly, the upcoming social web and social media technologies also integrate the user interactions in social networking environments.

Social media and ubiquitous data are thus expanding their breadth and depth. On the one hand, there are more and more social media applications; on the other hand, ubiquitous sensors are becoming part of personal and societal life at larger scales, as well as the accompanying social computational devices and applications. This can be observed in many domains and contexts, including events and activities in business and personal life.

Understanding and modeling ubiquitous (and) social systems require novel approaches and new techniques for their analysis. This book sets out to explore this emerging space by presenting a number of current approaches and important work addressing selected aspects of this problem. The individual contributions of this book focus on problems related to the modeling and mining of ubiquitous social media. Methods for mining, modeling, and engineering can help to advance our understanding of the dynamics and structures inherent to systems integrating and applying ubiquitous social media. Specifically, we advance on the analysis of dynamics and behavior on social media and ubiquitous data, e.g., concerning human contact networks or anomalous behavior. Furthermore, we consider helpful preprocessing methods for ubiquitous data and also focus on user modeling, privacy, and security aspects in order to address the user perspective of the respective systems.

The papers presented in this book are revised and significantly extended versions of papers submitted to two related workshops: The Second International Workshop on Mining Ubiquitous and Social Environments (MUSE 2011), which was held on September 5, 2011, in conjunction with the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD 2011) in Athens, Greece, and the Second International Workshop on Modeling Social Media (MSM'2011) that was held on October 9, 2011, in conjunction with IEEE SocialCom 2011 in Boston, USA. With respect to these two complementing workshop themes, the papers contained in this volume form a starting point for bridging the gap between the social and ubiquitous 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 into the user behavior in social media systems and thus similarly enhance and support modeling prospects. In the following, we briefly discuss the themes of these two workshops in more detail.

Social Media Modeling: 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, multifaceted 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. Much of the user interfaces used to access social media are difficult to use and do not translate well when shown on a mobile device such as a mobile phone. Moreover, design of social media on mobile devices is significantly different to creating PC-based or Web-based social media applications. Also, social media applications are typically modeled in an application-specific way and therefore there is no standardized method of creating a user interface. User interface modeling in social media can use a wide range of modeling perspectives such as justificative, explanative, descriptive, formative, predictive models and approaches (statistical modeling, conceptual modeling, temporal modeling, etc).

Ubiquitous Data Mining: 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 with their application in social and ubiquitous settings is one of the core contributions of this book.

Concerning the range of topics, we broadly consider three main themes: communities and networks in ubiquitous social media, mining approaches, and issues of user modeling, privacy, and security.

For the first main theme, we focus on the dynamics and the behavior concerning communities, anomalies, and human contact networks: We consider advanced community mining and analysis methods with respect to both social media and insights into ubiquitous human contact networks. Furthermore, network evaluation and anomalous behavior are discussed. “Integrating Social Media Data for Community Detection” by Jiliang Tang, Xufei Wang, and Huan Liu presents an approach for combining different data sources available in social media (networks) for enhanced community detection. Martin Atzmueller, Stephan Doerfel, Andreas Hotho, Folke Mitzlaff, and Gerd Stumme present related analysis and approaches in “Face-to-Face Contacts at a Conference: Dynamics of Communities and Roles” in a conferencing scenario. Philipp Singer, Claudia Wagner, and Markus Strohmaier discuss the evolution of content and social media in

“Factors Influencing the Co-evolution of Social and Content Networks in Online Social Media.” In “Mining Dense Structures to Uncover Anomalous Behavior in Financial Network Data,” Ursula Redmond, Martin Harrigan, and Padraig Cunningham discuss the discovery of anomalies in social networks and social media.

Concerning the mining approaches in ubiquitous social media, we distinguish explorative approaches for characterization and description, and mining methods for improving models, e.g., for interpolating ubiquitous data and for modeling temporal effects. For the first dimension, “Describing Locations Using Tags and Images: Explorative Pattern Mining in Social Media” by Florian Lemmerich and Martin Atzmueller presents an explorative pattern-mining approach for characterizing locations using tagging and image data. For the model-oriented techniques, “Learning and Transferring Geographically Weighted Regression Trees Across Time” by Annalisa Appice, Michelangelo Ceci, Donato Malerba, and Antonietta Lanza proposes a spatial data-mining method for geographically weighted regression. After that the paper “Trend Cluster-Based Kriging Interpolation in Sensor Data Networks” by Pietro Guccione, Annalisa Appice, Anna Ciampi, and Donato Malerba describes an interpolation method for ubiquitous data acquired, for example, in the context of pervasive sensor networks.

For user modeling, privacy, and security, we include a simulation approach and the modeling of user interfaces targeting privacy and security aspects. Else Nygren presents a simulation-based approach in “Simulation of User Participation and Interaction in Online Discussion Groups.” Next, Ricardo Tesoriero, Mohamed Bourimi, Fatih Karatas, Pedro Villanueva, Thomas Barth, and Philipp Schwarte discuss privacy and security issues in “Privacy and Security in Multi-Modal UIs with Model-Driven Modeling.”

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 engineering, modeling, mining, and analysis of ubiquitous social media and systems integrating these areas.

We want to thank the workshop and proceedings reviewers for their careful help in selecting and the authors for improving the provided submissions. We also thank all the authors for their contributions and the presenters for the interesting talks and the lively discussion at both workshops. Without these individuals we would not have been able to produce such a book.

May 2012

Martin Atzmueller
Alvin Chin
Denis Helic
Andreas Hotho

Organization

Program Committee

Martin Atzmueller	University of Kassel, Germany
Jordi Cabot	INRIA-École des Mines de Nantes, France
Ciro Cattuto	Institute for Scientific Interchange (ISI) Foundation, Italy
Michelangelo Ceci	Università degli Studi di Bari, Italy
Alvin Chin	Nokia Research Center
Marco De Gemmis	University of Bari, Italy
Wai-Tat Fu	University of Illinois, USA
Denis Helic	KMI, TU-Graz, Austria
Andreas Hotho	University of Würzburg, Germany
Huan Liu	Arizona State University, USA
Ion Muslea	Language Weaver, Inc.
Claudia Müller-Birn	FU Berlin, Germany
Else Nygren	Uppsala University, Sweden
Giovanni Semeraro	University of Bari, Italy
Marc Smith	Connected Action Consulting Group
Myra Spiliopoulou	University of Magdeburg, Germany
Gerd Stumme	University of Kassel, Germany
Christoph Trattner	IICM, TU-Graz, Austria

Additional Reviewers

Abbasi, Mohammad Ali
Aiello, Luca Maria
Canovas Izquierdo, Javier Luis
Martínez, Salvador
Panisson, André
Pio, Gianvito

Table of Contents

Integrating Social Media Data for Community Detection	1
<i>Jiliang Tang, Xufei Wang, and Huan Liu</i>	
Face-to-Face Contacts at a Conference: Dynamics of Communities and Roles	21
<i>Martin Atzmueller, Stephan Doerfel, Andreas Hotho, Folke Mitzlaff, and Gerd Stumme</i>	
Factors Influencing the Co-evolution of Social and Content Networks in Online Social Media	40
<i>Philipp Singer, Claudia Wagner, and Markus Strohmaier</i>	
Mining Dense Structures to Uncover Anomalous Behaviour in Financial Network Data	60
<i>Ursula Redmond, Martin Harrigan, and Pádraig Cunningham</i>	
Describing Locations Using Tags and Images: Explorative Pattern Mining in Social Media	77
<i>Florian Lemmerich and Martin Atzmueller</i>	
Learning and Transferring Geographically Weighted Regression Trees across Time	97
<i>Annalisa Appice, Michelangelo Ceci, Donato Malerba, and Antonietta Lanza</i>	
Trend Cluster Based Kriging Interpolation in Sensor Data Networks	118
<i>Pietro Guccione, Annalisa Appice, Anna Ciampi, and Donato Malerba</i>	
Simulation of User Participation and Interaction in Online Discussion Groups	138
<i>Else Nygren</i>	
Model-Driven Privacy and Security in Multi-modal Social Media UIs . . .	158
<i>Ricardo Tesoriero, Mohamed Bourimi, Fatih Karatas, Thomas Barth, Pedro G. Villanueva, and Philipp Schwarte</i>	
Author Index	183