Computer Communications and Networks

For other titles published in this series, go to
http://www.springer.com/series/4198

The **Computer Communications and Networks** series is a range of textbooks, monographs and handbooks. It sets out to provide students, researchers and nonspecialists alike with a sure grounding in current knowledge, together with comprehensible access to the latest developments in computer communications and networking.

Emphasis is placed on clear and explanatory styles that support a tutorial approach, so that even the most complex of topics is presented in a lucid and intelligible manner.

Ajith Abraham • Aboul-Ella Hassanien
Václav Snášel
Editors

Computational Social Network Analysis

Trends, Tools and Research Advances

Springer

Editors
Prof. Dr. Ajith Abraham
Machine Intelligence Research
Labs (MIR)
Scientific Network for Innovation
& Research Excellence
P.O.Box 2259
Auburn WA 98071-2259
USA
ajith.abraham@ieee.org

Prof. Aboul-Ella Hassanien
Cairo University
Fac. Computers & Information
Dept. Information Technology
5 Ahmed Zewal Street
Orman, Giza 12613
Egypt
a.hassanien@fci-cu.edu.eg

Václav Snášel
Technical University Ostrava
Dept. Computer Science
Tr. 17. Listopadu 15
708 33 Ostrava
Czech Republic
vaclav.snasel@vsb.cz

Series Editor
Professor A.J. Sammes, BSc, MPhil, PhD, FBCS, CEng
Centre for Forensic Computing
Cranfield University
DCMT, Shrivenham
Swindon SN6 8LA
UK

ISSN 1617-7975
ISBN 978-1-84882-228-3 e-ISBN 978-1-84882-229-0
DOI 10.1007/978-1-84882-229-0
Springer Dordrecht Heidelberg London New York

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Control Number: 2009942893

© Springer-Verlag London Limited 2010
Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms of licenses issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.
The use of 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 laws and regulations and therefore free for general use.
The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this book and cannot accept any legal responsibility or liability for any errors or omissions that may be made.

Cover design: SPi Publisher Services

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Social networks provide a powerful abstraction of the structure and dynamics of diverse kinds of people or people-to-technology interaction. Web 2.0 has enabled a new generation of Web-based communities, social networks, and folksonomies to facilitate collaboration among different communities. Social network analysis is a rapidly growing field within the Web intelligence domain. The recent developments in Web 2.0 have provided more opportunities to investigate the dynamics and structure of Web-based social networks. Recent trends also indicate the usage of social networks as a key feature for next-generation usage and exploitation of the Web. This book provides an opportunity to compare and contrast the ethological approach to social behavior in animals (including the study of animal tracks and learning by members of the same species) with web-based evidence of social interaction, perceptual learning, information granulation, the behavior of humans, and affinities between web-based social networks. The main topics cover the design and use of various computational intelligence tools and software, simulations of social networks, representation and analysis of social networks, use of semantic networks in the design, and community-based research issues such as knowledge discovery, privacy and protection, and visualization. This book presents some of the latest advances of various topics in intelligence social networks and illustrates how organizations can gain competitive advantages by applying the different emergent techniques in the real-world scenarios. Experience reports, survey articles, and intelligence techniques and theories with specific networks technology problems are depicted. We hope that this book will be useful for researchers, scholars, postgraduate students, and developers who are interested in intelligence social networks research and related issues. In particular, the book will be a valuable companion and comprehensive reference for both postgraduate and senior undergraduate students who are taking a course in Social Intelligence. The book contains 18 chapters, which are organized into two parts, and all articles are self-contained to provide greatest reading flexibility.

This volume is divided into three parts as detailed below. Part I deals with Social Network Mining Tools and consists of seven chapters.

The increasing achievement of the Web has led people to exploit collaborative technologies in order to encourage partnerships among different groups. The interest in analyzing Virtual Social Networks has grown massively in recent years

and it has involved researches from different fields. This led to the development of different methods to study relationships between people, groups, organizations, and other knowledge-processing entities on the Web. In Chapter 1, "An Overview of Methods for Virtual Social Networks Analysis," D'Andrea et al. classify these methods in two categories. The first category concerns methods used for the network data collection, while the second category deals with methods used for the network data visualization. The chapter gives an example of application of these methods to analyze the Virtual Social Network LinkedIn.

The discovery of single key players in social networks is commonly done using some of the centrality measures employed in social network analysis. In Chapter 2, "Discovering Sets of Key Players in Social Networks," Ortiz-Arroyo presents a brief survey of such methods. The methods described include a variety of techniques ranging from those based on traditional centrality measures using optimizing criteria to those based on measuring the efficiency of a network. Additionally, the author describes and evaluates a new approach to discover sets of key players based on entropy measures.

Barton et al. in Chapter 3, "Towards Self-Organizing Search Systems," propose a general three-layer model for designing and implementing a self-organizing system that aims at searching in multimedia data. This model gives developer guidelines about what component must be implemented and how they should behave. The usability of this model is illustrated on a system called Metric Social Network.

Chapter 4, "DISSECT: Data-Intensive Socially Similar Evolving Community Tracker" by Chin and Chignell, examines the problem of tracking community in social networks inferred from online interactions by tracking evolution of known subgroups over time. A variety of approaches have been suggested to address this problem, and the corresponding research literature is reviewed, which include centrality and clustering for identifying subgroup members, and optimization. The main contribution of this chapter is the DISSECT method wherein multiple known subgroups within a social network are tracked in terms of similarity-based cohesiveness over time. The DISSECT method relies on cluster analysis of snapshots of network activity at different points in time followed by similarity analysis of subgroup evolution over successive time periods.

The Blogosphere is expanding at an unprecedented speed, and a better understanding of the blogosphere can greatly facilitate the development of the Social Web to serve the needs of users, service providers, and advertisers. One important task in this process is the clustering of blog sites. In Chapter 5, "Clustering of Blog Sites Using Collective Wisdom," Agarwal et al. illustrate how clustering with collective wisdom can be achieved and compare its performance with respect to representative traditional clustering methods.

In Chapter 6, "Exploratory Analysis of the Social Network of Researchers in Inductive Logic Programming," Lavrač et al. present selected techniques for social network analysis and text mining, and interpret the results of exploratory analysis of the social network of researchers in inductive logic programming (ILP), on the basis of the ILP scientific publications database collected within the ILPnet2 project. In addition the chapter also presents a novel methodology for topic

ontology learning from text documents. The proposed methodology, named OntoTermExtraction (Term Extraction for Ontology learning), is based on OntoGen, a semiautomated tool for topic ontology construction, upgraded by using an advanced terminology extraction tool in an iterative, semiautomated ontology construction process.

In Chapter 7, "Information Flow in Systems of Interacting Agents as a Function of Local and Global Topological Features," Andre S. Ribeiro presents the latest results on how a system's structure, namely, its topological features, at global and local levels, affects the flow of information among its elements. It is shown how the topology determines the amount and diversity of information flow, and the ability to cope with noise and uncertainty in the information transmitted. The author uses Boolean Networks as models of dynamical systems of interacting elements whose structure is characterized at a global level by size and average connectivity and at a local level by a generalized clustering coefficient. The dynamics is characterized by the pair-wise mutual information of the time series of the elements states, which quantifies the amount of information flow, and by the Lempel-Ziv complexity, which characterizes the complexity, i.e., diversity of the messages flowing in the system.

Part II consists of eight chapters and deals with Social Network Evolution.

In Chapter 8, "Network Evolution: Theory and Mechanisms," Omidi and Masoudi-Nejad review some of the important network models that are introduced in recent years. The aim of all of these models is imitating the real-world network properties. Real-world networks exhibit behaviors such as small-world, scale-free, and high clustering coefficient.

In Chapter 9, "Vmap-Layout, a Layout Algorithm for Drawing Scientograms," Quirin and Cordón present a drawing algorithm to represent graphically co-citation networks. The proposed method can print real-world networks in an aesthetic way, highlighting the backbone and pushing the less important links to the boundaries. The algorithm is detailed and compared with the classical Kamada-Kawai drawing algorithm on several scientograms.

In Chapter 10, "Nature-Inspired Dissemination of Information in P2P Networks," by Christophe illustrates the dissemination of information within groups of people and aim at answering one question: can we find an effort-less way of sharing information on the web? A nature-inspired framework is introduced as an answer to this question. This framework features artificial ants taking care of the dissemination of information items within the network.

In Chapter 11, "Analysis and Visualization of Relations in e-Learning," Pavla et al. focus on searching of latent social networks in e-Learning systems data. This data consists of student's activity records where latent ties among actors are embedded. The social network studied in this chapter is represented by groups of students who have similar contacts and interact in similar social circles. Different methods of data clustering analysis are applied to these groups and the findings illustrate the existence of latent ties among the group members. The second part of this chapter focuses on social network visualization for monitoring the study activities of individuals or groups, as well as the planning of educational curriculum, the evaluation of study processes, etc.

In Chapter 12, "Interdisciplinary Matchmaking: Choosing Collaborators by Skill, Acquaintance and Trust," Hupa et al. present fundamental concepts on how to score a team based on members' social context and their suitability for a particular project. The chapter represents the social context of an individual as a three-dimensional social network (3DSN), composed of a knowledge dimension expressing skills, a trust dimension, and an acquaintance dimension. Dimensions of a 3DSN are used to mathematically formalize the criteria for prediction of team's performance.

In Chapter 13, "Web Communities Defined by Web Page Content," Kudělka et al. explore the relationship between the intent of web pages, their architecture, and the communities who take part in their usage and creation. The chapter describes techniques, which can be used to extract the mentioned information as well as tools usable in the analysis of this information. Authors define the MicroGenre as a building block of web pages, which is based on social interaction.

In Chapter 14, "Extended Generalized Block Modeling for Compound Communities and External Actors," Radoslaw and Krawczyk consider social communities composed of several cohesive subgroups, which they call compound communities. For such communities, an extended generalized block modeling is proposed in this chapter taking into account the structure of compound communities and relations with external actors. Using the extension, the community protection approach is proposed and used in detection of spam directed toward an e-mail local society.

In Chapter 15, "Analyzing Collaborations Through Content-Based Social Networks," Cucchiarelli et al. present a methodology and a software application to support the analysis of collaborations and collaboration content in scientific communities. High quality terminology extraction, semantic graphs, and clustering techniques are used to identify the relevant research topics. Traditional and novel social analysis tools are then used to study the emergence of interests around certain topics, the evolution of collaborations around these themes, and to identify potential for better cooperation.

Part III deals with Social Network Applications and consists of three chapters.

Grzegorz et al. in Chapter 16 "IA-Regional-Radio – Social Network for Radio Recommendation," present a system based on social network for radio application. This system carries out automatic collection, evaluation, and rating of music reviewers, the possibility for listeners to rate musical hits and recommendations deduced from auditor's profiles in the form of regional Internet radio. First, the system searches and retrieves probable music reviews from the Internet. Subsequently, the system carries out an evaluation and rating of those reviews. From this list of music hits, the system directly allows notation from our application. Finally, the system automatically creates the record list diffused each day depending on the region, the year season, day hours, and age of listeners.

In Chapter 17, "On the Use of Social Networks in Web Services – Application to the Discovery Stage," Maamar et al. discusses the use of social networks in web services with focus on the discovery stage that characterizes the life cycle of these Web services. Traditional discovery techniques are based on registries such as Universal Description, Discovery and Integration (UDDI), and Electronic Business using eXtensible Markup Language (ebXML). Unfortunately, despite the

different improvements that these techniques have been subject to, they still suffer from various limitations that could slow down the acceptance trend of web services by the IT community. Social networks seem to offer solutions for some of these limitations, but raise at the same time some issues that are discussed within this chapter.

The friendship relation, a social relation among individuals, is one of the primary relations modeled in some of the world's largest online social networking sites, such as FaceBook. On the other hand, the co-occurrence relation, as a relation among faces appearing in pictures, is one that is easily detectable using modern face detection techniques. These two relations, though appearing in different realms (social vs. visual sensory), have a strong correlation: faces that co-occur in photos often belong to individuals that are friends. Using real-world data gathered from Facebook, which were gathered as part of the FaceBots project, the world's first physical face recognizing and conversing robot can utilize and publish information on Facebook. Mavridis et al. in Chapter 18, "Friends with Faces: How Social Networks Can Enhance Face Recognition and Vice Versa," present methods as well as results for utilizing this correlation in both directions. Both algorithms for utilizing knowledge of the social context for faster and better face recognition are given, as well as algorithms for estimating the friendship network of a number of individuals given photos containing their faces.

We are very much grateful to the authors of this volume and to the reviewers for their tremendous service in critically reviewing the chapters. Most of the authors of chapters included in this book also served as referees for chapters written by other authors. Thanks go to all those who provided constructive and comprehensive reviews. The editors would like to thank Wayne Wheeler and Simon Rees of Springer Verlag, Germany for the editorial assistance and excellent cooperative collaboration to produce this important scientific work. We hope that the reader will share our excitement to present this volume on social networks and will find it useful.

Machine Intelligence Lab (MIR Lab) *Ajith Abraham*
Cairo University, Egypt *Aboul Ella Hassanien*
VSB Technical University, Czech Republic *Václav Snášel*

Contents

Part I Social Network Mining Tools

1 An Overview of Methods for Virtual Social Networks Analysis 3
Alessia D'Andrea, Fernando Ferri, and Patrizia Grifoni

2 Discovering Sets of Key Players in Social Networks 27
Daniel Ortiz-Arroyo

3 Toward Self-Organizing Search Systems 49
Stanislav Barton, Vlastislav Dohnal, Jan Sedmidubsky, and Pavel Zezula

4 DISSECT: Data-Intensive Socially Similar Evolving Community Tracker 81
Alvin Chin and Mark Chignell

5 Clustering of Blog Sites Using Collective Wisdom 107
Nitin Agarwal, Magdiel Galan, Huan Liu, and Shankar Subramanya

6 Exploratory Analysis of the Social Network of Researchers in Inductive Logic Programming 135
Nada Lavrač, Miha Grčar, Blaž Fortuna, and Paola Velardi

7 Information Flow in Systems of Interacting Agents as a Function of Local and Global Topological Features 155
Andre S. Ribeiro

Part II Social Network Evolution

8 Network Evolution: Theory and Mechanisms 191
Saeed Omidi and Ali Masoudi-Nejad

9 Vmap-Layout, a Layout Algorithm for Drawing Scientograms 241
Arnaud Quirin and Oscar Cordón

10 Nature-Inspired Dissemination of Information in P2P Networks 267
Christophe Guéret

11 Analysis and Visualization of Relations in eLearning 291
Pavla Dráždilová, Gamila Obadi, Kateřina Slaninová, Jan Martinovič, and Václav Snášel

12 Interdisciplinary Matchmaking: Choosing Collaborators by Skill, Acquaintance and Trust 319
Albert Hupa, Krzysztof Rzadca, Adam Wierzbicki, and Anwitaman Datta

13 Web Communities Defined by Web Page Content 349
Miloš Kudělka, Václav Snášel, Zdeněk Horák, Aboul Ella Hassanien, and Ajith Abraham

14 Extended Generalized Blockmodeling for Compound Communities and External Actors 371
Radoslaw Brendel and Henryk Krawczyk

15 Analyzing Collaborations Through Content-Based Social Networks 387
Alessandro Cucchiarelli, Fulvio D'Antonio, and Paola Velardi

Part III Social Network Applications

16 IA-Regional-Radio – Social Network for Radio Recommendation 413
Grzegorz Dziczkowski, Lamine Bougueroua, and Katarzyna Wegrzyn-Wolska

17 On the Use of Social Networks in Web Services: Application to the Discovery Stage 437
Zakaria Maamar, Leandro Krug Wives, and Khouloud Boukadi

18 Friends with Faces: How Social Networks Can Enhance Face Recognition and Vice Versa 453
Nikolaos Mavridis, Wajahat Kazmi, and Panos Toulis

Index 483