Lecture Notes in Social Networks

More information about this series at http://www.springer.com/series/8768

Jalal Kawash • Nitin Agarwal • Tansel Özyer
Editors

Prediction and Inference from Social Networks and Social Media

Springer

Editors
Jalal Kawash
Department of Computer Science
University of Calgary
Calgary, AB, Canada

Nitin Agarwal
Information Science Department
University of Arkansas at Little Rock
Little Rock, AR, USA

Tansel Özyer
Department of Computer Engineering
TOBB University
Ankara, Turkey

ISSN 2190-5428 ISSN 2190-5436 (electronic)
Lecture Notes in Social Networks
ISBN 978-3-319-51048-4 ISBN 978-3-319-51049-1 (eBook)
DOI 10.1007/978-3-319-51049-1

Library of Congress Control Number: 2017932668

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Social networks (SN) have brought an unprecedented revolution in how people interact and socialize. SN are used not only as a lifestyle but also in various other domains, including medicine, business, education, politics, and activism. SN have as well grown in sizes to include billions of users. As of this writing, Twitter claims to have 313 million monthly active users. Facebook grew by the end of 2016 to 1.71 billion users with 1.13 billion daily active users. Facebook now has a "population" that surpassed the population of India! Online social media (OSM), media produced by SN users, have offered a real and viable alternative to conventional mainstream media. OSM are likely to provide "raw", unedited information, and the details can be overwhelming with the potential of misinformation and disinformation. Yet, OSM are leading to the democratization of knowledge and information. OSM are allowing almost any citizen to become a journalist reporting on specific events of interest. This is resulting in unimaginable amounts of information being shared among a huge number of OSM participants. For example, Facebook users are generating several billion "likes" and several hundred million posted pictures in a single day. Twitter users are producing 6000 tweets per second. This immense amount of OSM poses increasing challenges to mine, analyse, utilize, and exploit such content. One grand challenge in OSM is mining its content to make useful inferences or predict future behaviour of SN users. This book includes nine contributions that examine new approaches that relate to predication and inference in OSM content. What follows is a quick summary of these chapters included in this book.

Mood prediction in SN is of great utility, especially in the medical field. For instance, predicting a patient's mood can be crucial to identify signs for depression. In this book, Roshanaei et al.'s approach is to design accurate personalized classifiers that can predict a person's emotions based on features extracted from OSM postings. By developing techniques to mine features such as social engagement, gender, language and linguistic styles, and various psychological features in a patient's tweets, they are able to infer the patient's mode as positive, negative, or neutral. In a different chapter, Kaya investigates the prediction of future

symptoms of patients from current patients' records. Kaya's approach consists of the construction of a weighted symptom network and, then, through unsupervised link prediction, building the evolving structure of symptom network with respect to patients' ages. Medical SN are also the subject of a chapter by Ayadi et al. Their objective is the automatic inference of indexing medical images. Their approach automatically extracts and analyses information from specialist's analysis and recommendations. The approach is multilingual, applying to different languages.

Shahriari et al. study signed SN in another chapter. Their focus is the significance of overlapping members in signed networks, and the intension is to discover overlapping communities in these networks. Different features are used to investigate the significance of overlapping members.

Alhajj looks at link prediction as a class of recommendation systems, predicting recommendations (links) between users and items. Efficiently finding hidden links and extracting missing information in a network aid in identifying a set of new interactions. In this chapter, Alhajj approaches the problem by exploiting the benefits of social network analysis tools and algorithms. For better scalability and efficiency, Alhajj utilizes a graph database model, as opposed to a traditional relational database.

The prediction of the quality of Wikipedia articles is studied in a chapter by De La Robertie et al. Wikipedia is not immune to problems relating to article quality, such as reputability of third-party sources and vandalism. The huge number of articles and the intensive edit rate make it infeasible to manually evaluate the content quality. De La Robertie et al. propose a quality model that integrates both temporal and structural features captured from the implicit peer review process enabled by Wikipedia. Two mutually reinforced factors are taken into account: article quality and author authority.

Charitonidis et al. realize that microblogs, such as Twitter, can be used for a good or a bad cause. In this chapter, their concern is the prediction of collective behaviour before it happens. Their approach is to analyse social media content to detect what they call "weak" signals; these are indicators that initially appear isolated but can be early indicators of large-scale, real-world phenomena. The 2011 London riots and tweets pertaining to them are their test-bed for their study.

Discussion forums of Massive Open Online Courses (MOOCs) are the subject of a chapter by Hecking et al. Their objective is to infer the structure of knowledge exchange in MOOC forums. The first step is the extraction of dynamic communication networks from forum data. Next is characterizing users according to information-seeking and information-giving behaviours and analysing the development of individual actors. Finally comes the reduction of a dynamic network to reflect knowledge exchange between clusters of actors and its evaluation.

Bouanan argues that the definition of a unique social network is too restrictive since in reality people are interlinked by several relationships, rather than by only one relationship. Hence, multidimensional networks are a better representation of relations among humans. They also define distinctive rules for the simulation of

message diffusion. Hence, Bouanan's model includes agents interacting through multiple channels or with different relationships, and information disseminates differently on different link categories. The modelling and simulation of multidimensional networks are the subject of this chapter.

Calgary, AB, Canada — Jalal Kawash
Little Rock, AR, USA — Nitin Agarwal
Ankara, Turkley — Tansel Özyer

Contents