AI Communications 29 (2016) 55–56
DOI 10.3233/AIC-150688
IOS Press

Introduction

Social Network Analysis in Applications

Katarzyna Musial [a,*], Piotr Brodka [b] and Matteo Magnani [c]

[a] *Faculty of Science and Technology, Data Science Institute, Bournemouth University, Bournemouth, UK*
E-mail: kmusialgabrys@bournemouth.ac.uk
[b] *Department of Computational Intelligence, Wroclaw University of Technology, Wroclaw, Poland*
E-mail: piotr.brodka@pwr.edu.pl
[c] *Department of Information Technology, Uppsala University, Uppsala, Sweden*
E-mail: matteo.magnani@it.uu.se

Social networks have been investigated for many years, but until recently the scope of the analyses was limited due to the small size of the available data samples, usually collected through questionnaires and interviews. As a consequence there were no or limited efficiency requirements for the analysis methods. Nowadays, vast amounts of data about people and their interactions can be gathered using computer systems, coming from technological-based services such as Online Social Networking Sites, telecommunication services and e-mails. The variety of information gathered about people ranges from shopping habits through social contact to medical records. It gives researchers and practitioners the opportunity to dive into this Big Data. Nowadays, we are able to extract and analyse social networks consisting of millions of nodes and connections. Due to scale, complexity and dynamicity, these networks are extremely difficult to study using traditional social network analysis methods not optimised in terms of performance. At the same time, data about human communication, common activities and collaboration provide new opportunities for innovative applications.

This special issue contains extended papers presented during the Third Workshop on Social Network Analysis in Applications. Both the special issue and the workshop are devoted to the analysis of social structures and more specifically to the identification of the areas where social network analysis can be applied and provide knowledge not accessible through other types of analysis. The articles in this issue cover such topics as (i) information that people share vs data coming from physical sensors, (ii) prediction of human mobility patterns, (iii) malicious socialbots that act as social network "friends", (iv) identification of campaigns of malicious profiles on social networking sites and (v) selection of the best initial set of nodes for the collective classification process.

The first paper reasons why "human sensors", namely citizens who share information about their surroundings via social media, can supplement, complement, or even replace the information measured by physical sensors. It presents a methodology using probabilistic language models to extract and visualise the perceptions of initiatives from social media updates. Over six million geo-tagged tweets collected from regions with varying geographical, social, cultural and political characteristics illustrate the method's capability.

The massive amounts of geo-located data collected from mobile phone records have also sparked an ongoing effort to understand and predict the mobility patterns of human beings.

The second paper studies how social phenomena are reflected in mobile phone data, and shows that information about social events can be used to improve the prediction of users' location. It further proposes a method for the automatic detection of such events and discusses their relation to the social fabric as derived from mobile phone communications.

The third paper demonstrates the relative ease of creating malicious socialbots that act as social network

[*] Corresponding author: Katarzyna Musial, Faculty of Science and Technology, Data Science Institute, Bournemouth University, Bournemouth, UK. E-mail: kmusialgabrys@bournemouth.ac.uk.

"friends", resulting in online social network users unknowingly exposing potentially harmful information about themselves and their places of employment. The paper introduces an algorithm for infiltrating specific users who were employees of targeted organizations, using the topologies of organizational social networks and utilizing socialbots to gain access to these networks.

The fourth paper presents a full stack methodology for the identification of campaigns of malicious profiles on social networking sites, composed of maliciousness classification, campaign discovery and attack profiling. The methodology, named REPLOT (REtrieving Profile Links On Twitter) has been applied by authors to a real world dataset, with a view to understanding the links between malicious profiles, their attack methods and their connections.

The fifth paper investigates which nodes' characteristics should be considered while selecting nodes from which the process of active learning and inference should start in order to ensure the highest collective classification accuracy. This is especially important in the situations when the resources are limited and thus we are not able to manually classify all nodes in the network. For example if we consider a marketing campaign, the question is which customers should be initially targeted in order to optimise the return on investment (ROI) of the entire campaign.

This special issue is an outcome of a hard work of a number of people and could not happen without the support of our collaborators. We would like to thank the editors in chief of the *AI Communications* journal, Daniele Magazzeni, Maria Fox and Eva Onaindia, for their kind support and help during the entire process of publication. This was possible thanks to the work of the researchers who provided their anonymous reviews. Finally, we are most grateful to the authors for their valuable contributions and for their willingness and efforts to improve their papers in accordance with the reviewers' suggestions and comments.

We hope that we can meet and socialize during future editions of the SNAA workshops http://snaa.pwr.edu.pl/.