

EDITORIAL

Open Access

Thematic series on Social Network Analysis and Mining



Rodrigo Pereira dos Santos^{1*} and Giseli Rabello Lopes²

Abstract

Social networks were first investigated in social, educational and business areas. Academic interest in this field though has been growing since the mid twentieth century, given the increasing interaction among people, data dissemination and exchange of information. As such, the development and evaluation of new techniques for social network analysis and mining (SNAM) is a current key research area for Internet services and applications. Key topics include contextualized analysis of social and information networks, crowdsourcing and crowdfunding, economics in networks, extraction and treatment of social data, mining techniques, modeling of user behavior and social networks, and software ecosystems. These topics have important areas of application in a wide range of fields, such as academia, politics, security, business, marketing, and science.

1 Introduction

This Thematic Series of the *Journal of Internet Services and Applications* (JISA) presents a collection of articles around the topic of Social Network Analysis and Mining (SNAM). From advances in Computer Science research and practice, the field of SNAM has become an important subject due to (i) the large amount and diversity of data that could be analyzed, (ii) the capacity of processing and solving complex analysis with efficiency, (iii) the development of new solutions for visualization of complex networks, and (iv) the application of SNAM concepts in different domains.

The study of social networks was leveraged by the social, educational and business communities. Academic interest in this field has been growing since the mid twentieth century [1], given the increasing interaction among people, data dissemination and exchange of information. In this scenario, big data sets require more accurate analyses. As such, the development and evaluation of new techniques for social network analysis and mining (SNAM) is a current key research area for Internet services and applications. These topics have important areas of application in a wide range of fields.

A social network is composed of actors who have relationships with each other. Networks can have a few

to many actors (nodes) and one or many types of relationships (arrows) between pairs of actors [2]. In our daily life, we have several practical examples of social networks: our family, friends, and colleagues from the university, gym, work, or casual meetings. Individuals and organizations – seen as nodes in social networks – can be connected due to several reasons, such as friendship and genealogy, but also values, visions, ideas, finances, disagreements, conflicts, services, computer networks, air routes etc. The structure created from such a large amount of relationships is complex. Therefore, researchers study the network as a whole from a sociocentric view (all the links referring to specific relations in a given population), or as a social structure in an egocentric view (with links selected from specific people) [3].

In addition, people join and create groups in any society [4], but the web platform fostered critical changes in the way people can interact and think about the reality. Interactions (i) become easier, (ii) allow a frequent exchange of information, and (iii) transform communications tools and social media (e.g., microblogs, blogs, wikis, Facebook) to mass communication means that are more agile and far-reaching. As such, the use of social media contributes to the sharing of different types of information, especially in real time. Some examples are personal data, location, opinions and preferences. In this context, SNAM can support the understanding of preferences and associations, the identification of

* Correspondence: rps@uniriotec.br

¹Department of Applied Informatics, Federal University of the State of Rio de Janeiro – UNIRIO, Rio de Janeiro, Brazil

Full list of author information is available at the end of the article

interactions, the recognition of influences, and the comprehension of information flow (context and concepts) among network actors.

Finally, the understanding of interactions in a specific scenario can produce concrete results. In an organization, employees should work to avoid problems regarding knowledge sharing [5]. In natural science, social networks can aid in the study of endemics and epidemics propagation [6]. In marketing, SNAM can be used as a tool for brand spread, or for the study of a market segment towards the understanding of how information propagates [7]. The last (but not least) example is the use of SNAM for the identification of criminal networks [8].

This JISA Thematic Series originates from the 6th Brazilian Workshop on Social Network Analysis and Mining (BraSNAM 2017) that was held in São Paulo, Brazil, on July 04–05, 2017. BraSNAM 2017 was affiliated with the 37th Brazilian Computer Society Congress (CSBC 2017) which is the official event of the Brazilian Computer Society (SBC). BraSNAM is focused on bringing together researchers and professionals interested in social networks and related fields. The workshop aims at providing innovative contributions to the research, development and evaluation of novel techniques for SNAM and applications. Finally, the main goal is to provide a valuable opportunity for multidisciplinary groups to meet and engage in discussions on SNAM.

Continuing in this direction, this JISA Thematic Series targets new techniques for the field of SNAM, mainly fostered by the context of Internet services and applications. We received contributions at various levels: from theoretical foundations to experiments and case studies based on real cases and applications; from modeling to mining and analysis of big data sets; and from different subjects and domains, such as entertainment, public transportation, elections, and personal social circles.

This Thematic Series presents high-quality research and technical contributions. We received six submissions as extended versions of the best papers of BraSNAM 2017. Topics included: analysis of online discussion and comments, complex networks, graph mining, government open data, power metrics, community detection, link assessment, homophily, and sentiment analysis. The five out of six submissions that were selected for publication and appear in this issue are summarized in the following section.

2 The papers

Loures et al. [9] investigate the potential that online comments have to describe television series. The authors implement and evaluated several different summarization methods. Their results reveal that a small set of comments

can help to describe the corresponding episodes and, when taken together, the series as a whole.

Caminha et al. [10] use graph mining techniques for the detection of overcrowding and waste of resources in public transport. The authors propose a new data processing methodology for the evaluation of collective transportation systems. The results show that their approach is capable of identifying global imbalances in the system based on an evaluation of the weight distributions of the edges of the supply and demand networks.

Verona et al. [11] propose metrics for power analysis on political and economic networks based on a sociology theory and network topology. The authors present a case study using a network built on data from Brazilian Elections about electoral donations explaining how the metrics can help in the analysis of power and influence of the different actors (corporations and persons) in this network.

Leão et al. [12] propose a method to handle social network data that exploits temporal features to improve the detection of communities by existing algorithms. By removing random relationships, the authors observe that social networks converge to a topology with more pure social relationships and better quality community structures.

Finally, Caetano et al. [13] propose an analysis of political homophily among Twitter users during the 2016 US presidential election. Their results showed that the homophily level increases when there are reciprocal connections, similar speeches or multiplex connections.

3 Paper selection process

The paper selection process was run during 2018 and the papers were published as soon as they were accepted and online-first versions became ready. Each submission went through two to four revisions before the final decision. We invited leading experts who are international researchers in the field of SNAM and related topics to form this Thematic Series' editorial committee. All manuscripts were reviewed by at least three members of this editorial committee. Guest editors checked the new version produced after each review cycle in order to decide whether the authors carefully addressed the reviewers' comments. Otherwise, a further review cycle was requested by the guest editors. The papers were reviewed by a total of 19 reviewers. The names of the editorial committee members are listed on the acknowledgements of this editorial.

4 Conclusion

The future of research and practice in the field of SNAM is challenging. Opportunities are many: theoretical and applied research has been published in specific conferences

and journals, but also in traditional venues since it requires a multidisciplinary arrangement. Based on the papers accepted to this Thematic Series, we can highlight some research gaps. For example, Loures et al. [9] point out the challenge of *abstractive summarization for online comments*: it is usually a much more complex task than the extractive one, since it requires a natural language generation module and a domain dependent component to process and rank the extracted knowledge. In turn, Caminha et al. [10] point out the need for a *simulator for reproducing the dynamics of human mobility* through the bus system in the case of a large metropolis. In this context, the use of data mining to estimate probability can represent the current demand for a bus system.

Regarding applications of SNAM in the context of presidential elections, Verona et al. [11] point out the challenge of *redesigning the power metric* to show relative values inside the network, instead of big absolute values. Moreover, information about company owners should be integrated in order to reveal hidden connections behind donations and politicians. In turn, Caetano et al. [13] point out the need for further investigation on the *temporal political homophily analysis* correlating it with external events that may have influenced the users' sentiments. This effort can allow user classification through data mining techniques to identify candidates' advocates, political bots, and other actors. Finally, regarding community detection, Leão et al. [12] point out the challenge of adopting *different approaches for community detection*, consider additional algorithms to explore temporal aspects or identify overlapping communities, and evaluate filtered networks. Moreover, different alternatives to measure the strength of ties should be investigated.

In the end, this Thematic Series comes out with some meaningful over-arching results:

- SNAM researchers and practitioners recognize the importance of sentiment analysis for in the identification of conflicts and agreements, as well as social trends and movements, in different domains. As such, new methods and techniques should be developed based on the large set of existing empirical studies on this topic;
- The dynamic nature of social networks makes community detection somehow a hard work. Different algorithms exist and are many. However, the treatment of randomness and noise in social relations requires further investigation. In addition, the assessment of those relations over time is also a topic of interested in SNAM;
- Another challenge in the area is the understanding of social power and the way it manifests in social networks. In this context, power is tightly related to

the notion of influence and authority. Research can vary from the development and use of SNAM algorithms and tools to the theorization based on qualitative studies (e.g., case studies, ethnography, sociotechnical approaches);

- SNAM opens opportunities to investigate different types of systems, such as (i) systems-of-systems: a set of constituent software-intensive systems that are managerially and operationally independent, and present some emergent behavior and evolutionary development (e.g., smart cities, transportation, air space, flood monitoring), and (ii) software ecosystems: a set of actors and artifacts as well as their relations over a common technological platform (e.g., iOS, Android, Eclipse, SAP);
- Finally, SNAM can support research on new trends of collaborative systems, such as crowdsourcing, free and open source software development, accountability, transparency and community engagement. A common interest lies on how to improve information visualization and recommendation based on actors' characteristics and behaviors as well as the changes in their relations over time.

Acknowledgments

We thank all the authors, reviewers, editors-in-chief, and staff for the great work, which supported this Thematic Series on a very important topic both for the research community and for the industry. In particular, we thank all the editorial committee members: Alessandro Rozza (lastminute.com Group, ITALY), Altigran Soares da Silva (Federal University of Minas Gerais, BRAZIL), Antonio Loureiro (Federal University of Minas Gerais, BRAZIL), Ari-Veikko Anttiroiko (Tampere University), Artur Ziviani (National Laboratory for Scientific Computing, BRAZIL), Bernardo Pereira Nunes (Pontifical Catholic University of Rio de Janeiro, BRAZIL), Claudio Miceli de Farias (Federal University of Rio de Janeiro, BRAZIL), Daniel Batista (University of São Paulo, BRAZIL), Flavia Bernardini (Federal Fluminense University, BRAZIL), Giacomo Livan (University College London, UK), Isabela Gasparini (Santa Catarina State University, BRAZIL), Jesús Mena-Chalco (Federal University of ABC, BRAZIL), Jonice Oliveira (Federal University of Rio de Janeiro, BRAZIL), Leandro Augusto Silva (Mackenzie Presbyterian University, BRAZIL), Luciano Antonio Digiampietri (University of São Paulo, BRAZIL), Luiz André Portes Paes Leme (Federal Fluminense University, BRAZIL), Mirella Moro (Federal University of Minas Gerais, BRAZIL), Raimundo Moura (Federal University of Piauí, BRAZIL), Yosh Halberstam (University of Toronto, CANADA). The voluntary work of these researchers was crucial for this Thematic Series.

Authors' contributions

Both authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Applied Informatics, Federal University of the State of Rio de Janeiro – UNIRIO, Rio de Janeiro, Brazil. ²Department of Computer Science, Federal University of Rio de Janeiro – UFRJ, Rio de Janeiro, Brazil.

Published online: 22 July 2019

References

1. Berkowitz SD. An introduction to structural analysis: the network approach to social research. Toronto: Butterworth; 1982.
2. Wasserman S, Faust K. Social network analysis: methods and applications. Cambridge: Cambridge University Press; 1994. p. 1994.
3. Hanneman RA, Riddle M. Introduction to social network methods. Riverside, CA: University of California, Riverside; 2005.
4. Castells, M., 2000, The Rise of the Network Society (The Information Age: Economy, Society and Culture, Volume 1, 2nd ed Wiley-Blackwell.
5. Studart RM, Oliveira J, Faria FF, Ventura LVF, Souza JM, Campos MLM. Using social networks analysis for collaboration and team formation identification. In: Proceedings of the 15th international conference on computer supported cooperative work in design, Lausanne; 2011. p. 562–9.
6. Mikolajczyk RT, Kretzschmar M. Collecting social contact data in the context of disease transmission: prospective and retrospective study designs. *Soc Networks*. 2008;30(2):127–35.
7. Kempe D, Kleinberg J, Tardos É. Maximizing the spread of influence through a social network. In: Proceedings of the 9th ACM SIGKDD international conference on knowledge discovery and data mining. Washington, D.C., USA; 2003.
8. Svenson P, Svensson P, Tullberg H. Social network analysis and information fusion for anti-terrorism. In: Proceedings of the conference on civil and military readiness. Sweden: Enköping. p. 2006.
9. Loures TC, de Melo PO, Veloso AA. Is it possible to describe television series from online comments? *J Internet Serv Appl*. 2018;9:25.
10. Caminha C, Furtado V, Pinheiro V, Ponte C. Graph mining for the detection of overcrowding and waste of resources in public transport. *J Internet Serv Appl*. 2018;9:22.
11. Verona L, Oliveira J, Hisse JVC, Campos MLM. Metrics for network power based on Castells' network theory of power: a case study on Brazilian elections. *J Internet Serv Appl*. 2018;9:23.
12. Leão, J. C., Brandão, M. A., VAZ DE Melo, P. O. S., Laender, A. H. F. "Who is really in my social circle?", *Journal of Internet Services and Applications* (2018) 9:23.
13. Caetano JA, Lima HS, Santos MF, Marques-Neto HT. Using sentiment analysis to define twitter political users' classes and their homophily during the 2016 American presidential election. *J Internet Serv Appl*. 2018;9:18.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
