Detection of antagonism and polarization on social media through community boundaries

Alexis Guyot, Annabelle Gillet, Eric Leclercq

Noname manuscript No.
(will be inserted by the editor)

Detection of antagonism and polarization on social media through community boundaries

An application to the study of conspiracy theories within tweets related to COVID-19 vaccines

Alexis Guyot · **Annabelle Gillet** · **Éric Leclercq**

Received: date / Accepted: date

Abstract Keywords Social networks · Graph mining · Communities · Polarization · Communities boundaries

1 Extended Abstract

Nowadays, Online Social Networks (OSN) play an important role in data analytics thanks to their ability to capture a massive amount of interactions between individuals. They can be represented by graphs with people as vertices and their interactions as weighted and directed links. They usually exhibit scale-free properties [3] and allow the detection and the study of communities [5, 9, 11] to better understand the roles and behaviours of users.

However, for a detailed interpretation of some phenomenons, it is also necessary to adopt a broader view by studying communities themselves and their interactions. An example of phenomena is polarization, defined in social sciences by Isenberg [8] as the process whereby a social group can be divided into two opposed sub-groups having conflicting and contrasting positions on a given subject, with few individuals remaining neutral or in an intermediate position.

In our contribution [7], we propose new tools to detect polarization traces and to study community boundaries [4, 6] on graphs extracted from OSN (more particularly from Twitter). Our proposal extends Guerra's work [6], formalized for undirected and unweighted graphs. Because human interactions are not necessarily symmetrical and a lot of contextual information is contained in link direction (source/destination), we decide to exploit this feature in addition to

Laboratoire d'Informatique de Bourgogne - EA 7534
Université de Bourgogne Franche-Comté
Dijon, France
E-mail: alexis.guyot@depinfo.u-bourgogne.fr
E-mail: annabelle.gillet@depinfo.u-bourgogne.fr
E-mail: eric.leclercq@u-bourgogne.fr

weights, which allow a better granularity by quantifying the strength of the connection. Using the same original idea, we want to obtain a more precise tool, better suited to real data extracted from OSN.

Just like Guerra's, our method exclusively uses a structural analysis. We do not exploit any Natural Language Processing (NLP) or Sentiment Analysis method in order to overcome the language barrier and spelling approximations, unlike Alamsyah and Adityawarman's proposal [2]. Contrary to Morales et al. [10] and Al Amin et al.'s [1] methods, our analysis also does not need any additional domain knowledge.

Our method takes a graph decomposed into communities as an input and produces a more detailed structure of each community with two different areas: internals and boundaries. They are built and populated for each pair of communities with different users according to formalized structural properties, intuitively described as: 1) internals, are sets of nodes without any neighbour in the other community; and 2) boundaries, are sets of nodes with at least one neighbour located in the other community and another one in the internal area of their own community.

Interactions between community areas can be used to assess antagonism, on the assumption that more an individual is involved in his community, more likely he is committed to the group opinion, and so more likely he will be vehicle for antagonism. Thus, we estimate this value by calculating the internal interactions ratio for each node located in a boundary area. The method produces a matrix containing the average antagonism score in boundaries for each pair of communities.

Based on our formally defined boundaries, we propose a new measure: the porosity of boundaries, which is a rate estimated from the number of nodes having at least as much interactions with the other community as with the internal area of their own community. This measure can be used to comment on cohesion inside a given group.

We also supply an algorithm and an open source implementation in R[1] to compute the different areas (internals and boundaries), the antagonism matrix and the porosity of boundaries from a weighted and directed graph. The algorithm uses an intermediate data structure called the *Structural Matrix* which allows to reduce the algorithmic complexity by precomputing each vertex role within its community compared to each other.

Finally, we apply our method and run through a way to interpret its results on a case study. We look for polarization in a corpus containing more than 9 millions tweets related to COVID-19 vaccines. Among the detected communities with the Louvain method, we highlight the absence of a well-defined anti-vaccine group whereas pro-vaccine users stand out in their own community. We suggest this could be an indicator that anti-vaccine users either don't share too much their opinion on Twitter or split themselves on different communities with other concerns, for example anti-government ones.

[1] `https://github.com/AlexisGuyot/CommunityBoundaries`

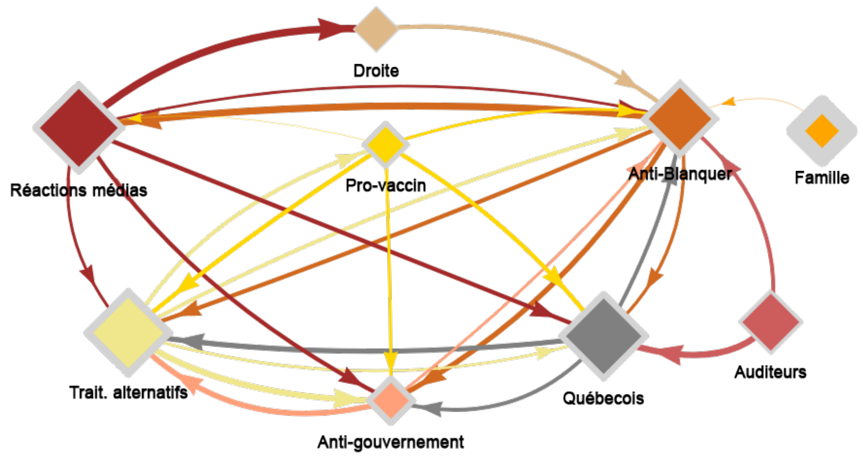

Fig. 1 Overview of communities rivalries in our case study.

We also point out the omnipresence of conspiracy hashtags inside all the communities main topics, such as *#GreatReset*, *#BlanquerMent*, *#Plandemie*, *#ComplotVaccinObligatoire* or *#JeNeSaisPasJeDemande*. [2] This observation is pretty surprising considering the fact that no keyword used to build the corpus mentions or refers to any conspiracy theory and underlines a likely mistrust in the COVID-19 vaccination campaign in France.

Even if there is no anti-vaccine community to be polarized with the pro-vaccine one, the antagonism values allow to detect polarization with the community supporting alternative medications to treat the disease. Finally, we introduce the hypothesis that boundary porosity could be used to measure how much a community as seen by a graph mining algorithm is close to a social community.

Overviews like the one presented in figure 1 can be used to highlight results of case studies. Each diamond represents a community. Shapes have proportional size according to: 1) community size for diamonds; 2) porosity value for light grey borders; and 3) antagonism score between communities pair for arrows. This kind of view allows analysts to easily interpret values returned by our method, for example to quickly identify communities relationships and rivalries or to compare porosity values between groups with different sizes.

In a near future, we plan to run more experiments on real data to confront our method with other subjects and problems, leaning on the interdisciplinary project Cocktail[3]. We also are working on extensions to handle overlapping communities, to better develop and use the porosity concept and to integrate our algorithm in a complete analysis workflow of communities in a graph. More

[2] *#GreatReset*, *#BlanquerLies*, *#Plandemic*, *#MandatoryVaccineConspiracy* or *#JustAsking*

[3] `https://projet-cocktail.fr/`

experiments will also be done to precisely estimate the algorithmic complexity of our method.

References

1. Al Amin, M.T., Aggarwal, C., Yao, S., Abdelzaher, T., Kaplan, L.: Unveiling polarization in social networks: A matrix factorization approach. In: IEEE INFOCOM 2017-IEEE Conference on Computer Communications, pp. 1–9. IEEE (2017)
2. Alamsyah, A., Adityawarman, F.: Hybrid sentiment and network analysis of social opinion polarization. In: 2017 5th International Conference on Information and Communication Technology (ICoIC7), pp. 1–6. IEEE (2017)
3. Barabási, A.L., Bonabeau, E.: Scale-free networks. Scientific american **288**(5), 60–69 (2003)
4. Clauset, A.: Finding local community structure in networks. Physical review E **72**(2), 026,132 (2005)
5. Fortunato, S.: Community detection in graphs. Physics reports **486**(3-5), 75–174 (2010)
6. Guerra, P., Meira Jr, W., Cardie, C., Kleinberg, R.: A measure of polarization on social media networks based on community boundaries. In: Proceedings of the International AAAI Conference on Web and Social Media, vol. 7 (2013)
7. Guyot, A., Gillet, A., Leclercq, É.: Frontières des communautés polarisées : application à l'étude des théories complotistes autour des vaccins (to be published). In: INFormatique des ORganisations et des Systèmes d'Information et de Décision (INFORSID) (2021)
8. Isenberg, D.J.: Group polarization: A critical review and meta-analysis. Journal of personality and social psychology **50**(6), 1141 (1986)
9. Lancichinetti, A., Fortunato, S.: Community detection algorithms: a comparative analysis. Physical review E **80**(5), 056,117 (2009)
10. Morales, A.J., Borondo, J., Losada, J.C., Benito, R.M.: Measuring political polarization: Twitter shows the two sides of venezuela. Chaos: An Interdisciplinary Journal of Nonlinear Science **25**(3), 033,114 (2015)
11. Reichardt, J., Bornholdt, S.: Statistical mechanics of community detection. Physical review E **74**(1), 016,110 (2006)