Polarization in large socio-informational networks: models and measures

Doctoral research project for the 2022 EDITE Campaign

Abstract

The prevalence of algorithmic recommendations in social media platforms has raised concerns about undesired societal effects. A central threat is the risk of polarization. Despite the numerous recent studies that investigate polarization, no clear answer has emerged as to what polarization is and how it should be measured, let alone what is the role of Recommender Systems in the phenomenon and what algorithmic design principles could be adopted to manage it. As many emerging phenomena in complex systems, polarization is hard to conceptualize, measure, and operationalize in Recommender Systems research. This doctoral project proposes to build on graph models for networks representing digital traces on large web platforms, and the development of measures from information theory, to understand and redefine polarization as a formal property of large complex networks, and the effect that recent Recommender Systems have on it. This highly interdisciplinary project will draw from recent advances but also from applied mathematics, graph formalisms, and empirical geometrical opinion space embeddings for social networks developed in the emergent field of computational social science.

Keywords: social network analysis, recommender systems, complex networks, information theory

1 Introduction

Social polarization has been an object of research and debate for several decades. Recently, observers have questioned the responsibility of recommendation algorithms in web platforms, suggesting they might be fostering filter bubbles and even radicalization [Flaxman et al., 2016]. However, we still lack empirical data and formal metrics to audit the role of recommendations in real-world settings. These concerns have inspired a wealth of studies, but mostly limited to two types: 1) purely topological approaches that investigate how recommenders connect or isolate groups of nodes in a graph (*e.g.*, do algorithms recommend conservative content to liberals, and vice-versa? [Bakshy et al., 2015]), and 2) spatial opinion approaches that study how recommenders change the distribution of users on some opinion scale (*e.g.*, on liberal-conservative opinion scales [Santos et al., 2021]). The former prove inadequate in settings where users cannot classified into categorical types (e.g., in two-party systems with binary social divides, which excludes French, European, and most settings). The latter rely on synthetic data and simulations due to the unobservability of opinions, offering little empirical or actionable results.

Improving the state of the research around polarization in complex networks and recommendations is difficult for several reasons. It requires a deep understanding of the mechanisms at the core of recommendations, of complex networks dynamics, and mathematical theories to formalize polarization as a property. All of the above must also be tackled in the background of existing models of social dynamics. This project builds on recent results in information-theoretical measurements of network properties, on recommender systems audit, and on inferring opinion space from online social networks. The project aims at combining these results to move forward in the analysis of the societal effects of algorithmic recommendation.

2 Related Works

Graph models of complex networks. Recently, a variety of graph formalisms has been applied to model complex social and informational systems, such as online social networks, online media, and news platforms. The Heterogeneous Information Network (HIN) [Shi et al., 2016] graph formalism allows to model edges of different nature (indicating, *e.g.*, interactions such becoming friends, retweeting, posting, commenting, sharing, clicking or viewing) between nodes of different types (*e.g.*, users, comments, messages and news articles or URLs in general). The HIN formalism is also at the center of a very active research community in Recommender Systems [Shi et al., 2018]. The typological flexibility of the HIN is complemented by graph formalisms capable of modeling the temporal dynamics of networks [Latapy et al., 2018], which can also be leveraged in recommendations [Tabourier et al., 2019].

Formalizing emergent properties in complex systems. Besides modeling complex systems and computing algorithmic recommendations, a different but related line of research using graphs takes interest in the measurement of emerging properties. Related to polarization, measuring *diversity* in social and information platforms has seen important recent developments [Bakshy et al., 2015]. Diversity is a property accounting for the variety or balance of how users interact among themselves and with items (such as content pieces or comments) online. Diversity has been measured in the context of algorithmic recommendations [Ramaciotti Morales et al., 2020] and online navigation traces [Ramaciotti Morales et al., 2019], using information theoretical concepts in graph settings [Ramaciotti Morales et al., 2021b]. Similar approaches are lacking in relation to polarization.

Polarization as a concept calls for geometrical models. While several authors have taken interest in purely topological metrics for polarization (*e.g.*, degrees of [dis]connection as measured by graph modularity [Guerra et al., 2013]), numerous social systems do not lend themselves to categorical classification on which these types of metrics can be used. A striking illustration is the case of political polarization: while in the US many users can be classified as Democrat- or Republican-leaning, in Europe, users may display political attitudes that do not equate to party sympathies, and that are best captured by continuous opinion scales (ranging from most opposed to most favorable towards several issues: *e.g.*, immigration, environment protection, etc.). In this context, polarization would be better described by the geometrical clustering of individuals around poles in opinion spaces defined by these scales. Recent empirical works have shown that graph embedding methods applied to social networks can yield positions of users in such spaces [Ramaciotti Morales et al., 2021a]. In these multidimensional opinion spaces, users can be positioned along several dimensions that act as indicators of positive or negative opinion towards several issues [Ramaciotti Morales et al., 2022]. Much like diversity, a few notions capturing polarization can be derived within axiomatic –information– theories [Duclos et al., 2004]. These families of information-theoretical measures have also been used for auditing recommendation algorithms in opinion spaces [Ramaciotti Morales and Cointet, 2021]. Still, a general theory of polarization which would encompass a broader set of notions underlying polarization is largely missing. Such a theory would allow for the definition of families of measures: functions which would account for both geometrical and topological descriptions of online social networks.

3 The Doctoral Research Project

Consequently, this project seeks to substantially advance the understanding and the modeling of polarization in online platforms. These advances in modeling will be conceived to investigate the impact of algorithmic recommendations. This project seeks to do so by developing work in several interacting areas of research:

Axiomatic theories of polarization in embedded networks. Similar to how recent research has axiomatized the measurement of diversity in complex systems ([Ramaciotti Morales et al., 2021b] at LIP6), an important lane of work is to develop formal measures for polarization. Such measures would have to include geometrical aspects besides graph and information theory. There is a need for geometrical frameworks to account for continuous and multidimensional models of opinions of online users. A starting point for this is to build on existing axiomatic theories [Duclos et al., 2004] to adapt them to the context of social networks embedded in opinion spaces. Connected to polarization in

networks, a special attention is given to heterogeneous and dynamic networks of users interacting between them and with online content. To produce results on large real-world data, it demands to control the algorithmic complexity of the measures defined.

Impact of recommender systems in polarization. The project is expected to investigate the effect of recommender systems in social and informational online systems. For this, the doctoral candidate will gain a deep understanding of recommenders from the state of the art, and conduct simulations based on both, synthetic and real-world data. An important research axis is understanding how representations learned by recommenders (*e.g.*, in weight spaces in GNN-based recommenders, or latent spaces in matrix factorization-based recommenders) impact polarization.

Real-world data. This project will benefit from unique large-scale, real-world data through co-supervision with médialab at Sciences Po (see the Supervision and Host Laboratory section). This should enable significant practical results besides the theoretical ambition of this project. In particular, the doctoral student will gain access to social graphs datasets and online activity traces of large number of users, with meta-data including positioning in opinion spaces.

[Inter]disciplinarity. This computer science project is mainly concerned with methods and algorithms adapted to graph representations. A special attention is given to graph models for heterogeneous and dynamic systems. This project is also concerned with information-theoretical measures in probability theory for the formalization of measures of polarization. The field of algorithmic recommendation being at the core of this project, having a deep knowledge of recommender systems is of central importance for this project. Finally, the issue of polarization is an emerging source of preoccupation in several areas related to social sciences. This project is thus connected with the disciplines that deal with them, leveraging large volumes of digital traces of online behaviors.

4 Supervision and Host Laboratory

This doctoral project will be hosted at *Laboratoire d'Informatique de Paris 6* (LIP6), by the Complex Networks Group, and co-supervised by

- **Lionel Tabourier**, *Maître de Conférences* (HDR), head of the Complex Networks Group, LIP6, Sorbonne Université/CNRS;
- **Pedro Ramaciotti Morales**, *Research Scientist*, médialab, Sciences Po & *Chair of AI in Computational Social Sciences*, CRI/LPI Université Paris Cité.

The Complex Networks Group at LIP6 (www.complexnetworks.fr) is specialized in the study of empirical graphs, graph modeling and algorithms, data mining on graphs.

The médialab at Sciences Po (medialab.sciencespo.fr) is an interdisciplinary laboratory gathering mathematicians, computer scientists, designers, sociologists, and political scientists. The médialab leverages resources from different disciplines to study digital informational and social ecosystems such as social networks and media internet platforms. Additional interactions are also expected with the Centre de Recherches Interdisciplinaires (CRI/LPI, www.cri-paris.org). This research unit of the Université Paris Cité fosters innovative interdisciplinary approaches to societal challenges, which should benefit to the doctoral project.

5 Required Profile of the Doctoral Student

The candidate for this project must have solid bases in applied mathematics, graph algorithms, and data mining. Having knowledge or experience with recommender systems and complex network modeling is also an advantage, as it is experience with data from internet platforms.

The candidate is not required to have previous knowledge or experience in computational social sciences, but must display scientific openness and interest to interact with researchers in this emerging field and understand the context and broader motivation of this project.

References

[Bakshy et al., 2015] Bakshy, E., Messing, S., and Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132.

[Duclos et al., 2004] Duclos, J.-Y., Esteban, J., and Ray, D. (2004). Polarization: concepts, measurement, estimation. *Econometrica*, 72(6):1737–1772.

[Flaxman et al., 2016] Flaxman, S., Goel, S., and Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly*, 80(S1):298–320.

[Guerra et al., 2013] Guerra, P., Meira Jr, W., Cardie, C., and Kleinberg, R. (2013). A measure of polarization on social media networks based on community boundaries. In *Proceedings of the international AAAI conference on web and social media*, volume 7, pages 215–224.

[Latapy et al., 2018] Latapy, M., Viard, T., and Magnien, C. (2018). Stream graphs and link streams for the modeling of interactions over time. *Social Network Analysis and Mining*, 8(1):1–29.

[Ramaciotti Morales and Cointet, 2021] Ramaciotti Morales, P. and Cointet, J.-P. (2021). Auditing the effect of social network recommendations on polarization in geometrical ideological spaces. In *RecSys' 21: 15th ACM Conference on Recommender Systems*.

[Ramaciotti Morales et al., 2021a] Ramaciotti Morales, P., Cointet, J.-P., and Muñoz Zolotoochin, G. (2021a). Unfolding the dimensionality structure of social networks in ideological embeddings. In *Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 333–338.

[Ramaciotti Morales et al., 2022] Ramaciotti Morales, P., Cointet, J.-P., Zolotoochin, G. M., Peralta, A. F., Iñiguez, G., and Pournaki, A. (2022). Inferring attitudinal spaces in social networks. *In revision at Social Network Analysis and Mining*.

[Ramaciotti Morales et al., 2021b] Ramaciotti Morales, P., Lamarche-Perrin, R., Fournier-S'Niehotta, R., Poulain, R., Tabourier, L., and Tarissan, F. (2021b). Measuring diversity in heterogeneous information networks. *Theoretical Computer Science*, 859:80–115.

[Ramaciotti Morales et al., 2020] Ramaciotti Morales, P., Tabourier, L., and Fournier-S'Niehotta, R. (2020). Testing the impact of semantics and structure on recommendation accuracy and diversity. In *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 250–257. IEEE.

[Ramaciotti Morales et al., 2019] Ramaciotti Morales, P., Tabourier, L., Ung, S., and Prieur, C. (2019). Role of the website structure in the diversity of browsing behaviors. In *Proceedings of the 30th ACM Conference on Hypertext and Social Media*, pages 133–142.

[Santos et al., 2021] Santos, F. P., Lelkes, Y., and Levin, S. A. (2021). Link recommendation algorithms and dynamics of polarization in online social networks. *Proceedings of the National Academy of Sciences*, 118(50).

[Shi et al., 2018] Shi, C., Hu, B., Zhao, W. X., and Philip, S. Y. (2018). Heterogeneous information network embedding for recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 31(2):357–370.

[Shi et al., 2016] Shi, C., Li, Y., Zhang, J., Sun, Y., and Philip, S. Y. (2016). A survey of heterogeneous information network analysis. *IEEE Transactions on Knowledge and Data Engineering*, 29(1):17–37.

[Tabourier et al., 2019] Tabourier, L., Bernardes, D. F., Libert, A.-S., and Lambiotte, R. (2019). Rankmerging: a supervised learning-to-rank framework to predict links in large social networks. *Machine Learning*, 108(10):1729–1756.