Dynamic graph generation to simulate spreading for social platforms evaluation

Mehdi Naima, Lionel Tabourier (LIP6, SU/CNRS)
Christophe Prieur (LISIS, UGE/CNRS)

1 Context and objectives

In this project, we aim to describe **social platforms and the processes that take place there**, in particular how information spreads and how it affects exposure to opinions. For that purpose, we need to describe accurately the interactions on these social platforms at the individual level, in terms of structure and in terms of dynamics. This leads to generate synthetic structures that mimic realistically their most salient characteristics in order to predict future behaviours. In this way, we aim to contribute to the evaluation of online social platforms, in particular how their practices in terms of recommendation and moderation affect the spreading and exposure to information. It comes within the priority of **building more resilient societies** fostered by PostGenAI@Paris.

As part of the CNRS MITI project MoNReSo, we have set up **a collaboration between computer scientists from LIP6** (Sorbonne Université/CNRS) **and sociologists from LISIS** (Université Gustave Eiffel/CNRS), and currently work on establishing the main features of *ego-based* social networks, meaning networks centered on an individual. A major key to describe these dynamic networks, is to take into account the correlation between the bursty activities of the individuals which may leads to cascades of information flowing through the network. By capturing the interactions at this individual level, we aim to understand how phenomena may emerge and percolate at a larger scale.

2 Data

The adoption of the Digital Services Act (DSA) by the European Union in November 2022 **aims to create the conditions for a safer and sounder numerical environment** by allowing the public audit of the practices of the large online platforms and search engines. This regulation produces a new context for the evaluation of the platform practices by allowing an unprecedented access to data from these platforms. Since October 2025, requests can be made on the EU platform to gain access to data specifically designed for the audit by vetted researchers: `https://data-access.dsa.ec.europa.eu`.

We believe that an essential role of the academics involved in the field is to take up this issue by **proposing new methods to achieve this public audit**. It would create the conditions to ensure the regulation and ethical use of the AI technologies underpinning the functioning of the platforms, one of the core mission of PostGenAI@Paris. Having access to data through the DSA is an excellent way to investigate the complex entanglement between the structure and dynamics of information diffusion on social platforms. Our first target would be micro-blogging platforms like X/Twitter, where the spreading of misinformation has been largely documented [4, 10] and the techniques to detect it well-improved. However, the detection is made afterwards, making these systems mostly inoperative as alert warnings.

Our view is that **the simulation of spreading processes on synthetic platforms allow for a better understanding of their dynamics and allow the definition of efficient warnings**. Indeed, having access to the interaction data on these platforms (follower/followee links, re-twitting, citing, etc.) we have the elements to make realistic artificial models and simulate various spreading scenarios and actions that could efficiently limit their negative effects.

3 Challenges and scientific approach

In this PhD project, there are two main methodological aspects to investigate. The first one concerns **mining patterns in the data that allow to characterise their structure and dynamics**. Indeed, while there is a rich toolbox to identify relevant structural features on graphs, there is still a need to develop their equivalent to characterise synthetically the properties of interaction data through time. The team at LIP6 has acquired an expertise in the design and development of tools to analyse temporal data from real world interaction

networks [2, 5, 7]. Also, previous works by Christophe Prieur focused on identifying specific interaction patterns in online social networks, that are relevant to the understanding of their functioning [1, 3]. We plan to use this knowledge for the identification and measurement of the features which are specific to the data collected.

The second aspect concerns generation processes. Simple graph models, that emerged 25 years ago, helped describe some coarse-grain properties observed in real world, such as the heavy-tailed degree distribution or high clustering, but they proved limited when trying to account for more elaborate properties. Consequently, powerful techniques were proposed to generate graph models with stronger properties [6, 9, 8]. But improved access to data uncovered the important role of the temporality of interactions, especially when considering user behaviours online. There is now a **need for realistic models that also account for the dynamics of the interactions**. We think that Monte Carlo Markov Chain methods are probably the best lead to explore because they offer a possibility of unbiased generation even of complex structures. The main challenge in this area is to adapt efficiently these methods to temporal data.

4 Profile required and environment

The PhD would take place at Sorbonne Université (Jussieu Campus, Paris 5). Daily supervision would be provided by Mehdi Naima and Lionel Tabourier at LIP6 (Sorbonne-Université/CNRS), in collaboration with Christophe Prieur at LISIS (UGE/CNRS). Mehdi Naima is an Associate Professor of computer science. His field of research focuses on algorithm analysis as well as temporal graph algorithms (more information at `https://busyweaver.github.io/`). Lionel Tabourier is a Professor of computer science working in the field of of complex networks analysis. During his previous works, he contributed to realistic graph generation methods, as well as the analysis of information spreading on social platforms (more information at `https://lioneltabourier.github.io/`). Christophe Prieur is a Professor of sociology, specialized in the study of digital practices, and the transformations that they involve in society (more information at `https://pagespro.univ-gustave-eiffel.fr/christophe-prieur`).

We are looking for students whose primary interests are algorithmics and programming for data mining, but that also have a taste for interdisciplinary collaborations.

References

[1] Irène Bastard, Dominique Cardon, Raphaël Charbey, Jean-Philippe Cointet, and Christophe Prieur. What do we do on facebook?: Activity patterns and relational structures on a social network. *Sociologie*, 8(1):57–82, 2017.

[2] Alexis Baudin, Lionel Tabourier, and Clémence Magnien. Lscpm: Communities in massive real-world link streams by clique percolation method. In *30th International Symposium on Temporal Representation and Reasoning*, 2023.

[3] Raphaël Charbey and Christophe Prieur. Stars, holes, or paths across your facebook friends: A graphlet-based characterization of many networks. *Network Science*, 7(4):476–497, 2019.

[4] Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H Eugene Stanley, and Walter Quattrociocchi. The spreading of misinformation online. *Proceedings of the national academy of Sciences*, 113(3):554–559, 2016.

[5] Mehdi Naima. Temporal betweenness centrality on shortest walks variants. *Applied Network Science*, 10(1):11, 2025.

[6] Garry Robins, Pip Pattison, Yuval Kalish, and Dean Lusher. An introduction to exponential random graph (p*) models for social networks. *Social networks*, 29(2):173–191, 2007.

[7] Felix I Stamm, Mehdi Naima, and Michael T Schaub. Efficient sampling of temporal networks with preserved causality structure. *SIAM International Conference on Data Mining (SDM25)*, 2025.

[8] Lionel Tabourier and Julien Karadayi. Probabilistic k-swap method for uniform graph generation beyond the configuration model. *Journal of Complex Networks*, 12(1):cnae002, 2024.

[9] Katherine Van Koevering, Austin Benson, and Jon Kleinberg. Random graphs with prescribed k-core sequences: A new null model for network analysis. In *Proceedings of the Web Conference 2021*, pages 367–378, 2021.

[10] Soroush Vosoughi, Deb Roy, and Sinan Aral. The spread of true and false news online. *science*, 359(6380):1146–1151, 2018.