

# Characterizing Community Structures on Social Media Over Time: A Graph Learning Approach

María Victoria Zolezzi<sup>1</sup>, Federico Albanese<sup>2,3</sup>, and Esteban Feuerstein<sup>1,3</sup>  
 {mzolezzi, falbanese, efeuerst}@dc.uba.ar

<sup>1</sup> Departamento de Computación, Facultad de Ciencias Exactas y Naturales,  
 Universidad de Buenos Aires, Buenos Aires, Argentina.

<sup>2</sup> Instituto de Cálculo, CONICET Universidad de Buenos Aires, Buenos Aires,  
 Argentina.

<sup>3</sup> Instituto de Ciencias de la Computación, CONICET Universidad de Buenos Aires,  
 Buenos Aires, Argentina.

**Abstract.** In an age where information is more accessible than ever, it's easy to assume that people are becoming more informed and open-minded. In spite of that, people are increasingly finding themselves in echo chambers [1], surrounded by like-minded individuals and exposed mainly to content that reinforces their existing beliefs. There are, however, social media users that break with that pattern by changing the group of users they interact with over time [2]. In this study, we analyze the dynamics of interactions between users on Twitter and Reddit over extended periods, with the aim of identifying changes in community structures. We leverage the data available through these platforms' APIs to construct user interaction graphs and use several methods to classify users into communities, including SBM [3], Infomap [4] and Louvain [5], to classify users into communities. Additionally, we use NLP techniques such as Community Pooling [6], BERTopic [7] and Perspective [8], as well as graph algorithms, to characterize different user profiles in online debates. Our research analyzes how social media communities and their users evolve over time, with implications for understanding online discourse and facilitating healthy interactions on these platforms.

As a first approach, we analyzed three months of Donald Trump's tweets, finding clear signs of polarization. Regarding the user flow between communities, we found that most of the users who changed communities twice went back to their original one ( $\sim 96\%$ ).

## References

1. Quattrociocchi, W., Scala, A., & Sunstein, C. R. (2016). Echo chambers on Facebook. Available at SSRN 2795110.
2. Albanese, F., Lombardi, L., Feuerstein, E., & Balenzuela, P. (2020). Breaking the Communities: Characterizing community changing users using text mining and graph machine learning on Twitter. arXiv preprint arXiv:2008.10749.
3. Holland, P. W., Laskey, K. B., & Leinhardt, S. (1983). Stochastic blockmodels: First steps. *Social networks*, 5(2), 109-137.

4. Rosvall, M., & Bergstrom, C. T. (2007). An information-theoretic framework for resolving community structure in complex networks. *Proceedings of the national academy of sciences*, 104(18), 7327-7331.
5. Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10), P10008.
6. Albanese, F., & Feuerstein, E. (2021). Improved topic modeling in twitter through community pooling. In *String Processing and Information Retrieval: 28th International Symposium, SPIRE 2021, Lille, France, October 4-6, 2021, Proceedings 28* (pp. 209-216). Springer International Publishing.
7. Bhagat, R. K., Wilson, D. I., & Linden, P. F. (2020). Experimental evidence for surface tension origin of the circular hydraulic jump. *arXiv preprint arXiv:2010.04107*.
8. Wulczyn, E., Thain, N., & Dixon, L. (2017, April). Ex machina: Personal attacks seen at scale. In *Proceedings of the 26th international conference on world wide web* (pp. 1391-1399).

## Appendix

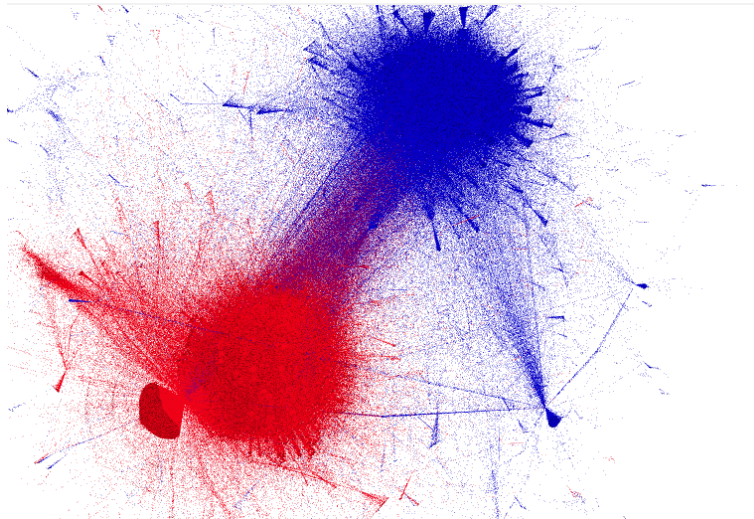


Fig. 1: The Democrat community (40.93%) is shown in blue, while the Republican community is shown in red. The Republican community comprises two sub-communities: a smaller one, consisting of users who almost exclusively retweeted Trump (2.4%), and another with a Republican majority, but also including some Democrats (56.67%).

Evolution of users between communities

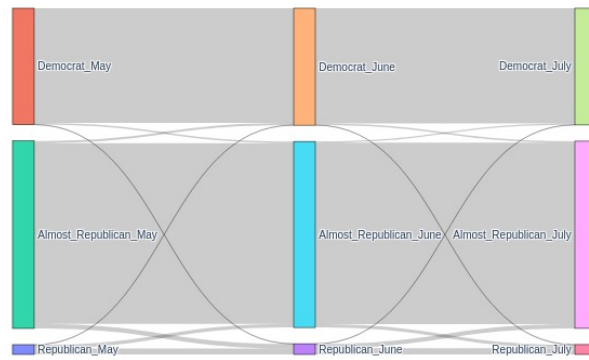


Fig. 2: Sankey diagram of the users' flow. As mentioned in the previous plot (1), there were two Republican communities: one with the people that almost exclusively retweeted Trump, and the other with a Republican majority. In this plot we call these communities *Republican* and *Almost\_Republican* respectively.