Modelling how social network algorithms can influence opinion polarization

Henrique Ferraz de Arruda [1], Felipe Maciel Cardoso[2], Guilherme Ferraz de Arruda [1], Alexis R. Hernández[3], Luciano da F. Costa[4], and Yamir Moreno[1,2,5]

[1]CENTAI Institute, Turin 10138, Italy
[2]Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Spain
[3]Institute of Physics, Rio de Janeiro Federal University, Rio de Janeiro, RJ, Brazil
[4]São Carlos Institute of Physics, University of São Paulo, São Carlos, SP, Brazil
[5]Department of Theoretical Physics, University of Zaragoza, 50018 Zaragoza, Spain
*h.f.arruda@gmail.com

Due to the considerable online social networks and their impact on society, scholars have been studying the dynamics behind these processes (1; 2; 3). Here, we proposed a novel approach to simulate online discussions on social networks, in which users and their friendship as the nodes and edges of a network. Our opinion model is based on the idea of adding information from external sources and how users and social network algorithms handle the information. The first step of this process represents a piece of news obtained outside the online social network, simulated as a random number. Next, two separate steps are associated with this news piece: *post transmission* and the *post distribution*. The *post transmission* models the willingness of a user to share the information, creating a post in the social network. To do so, we used different probability functions based on the difference between the post and the user's opinion. Among the possibilities of this function, we considered a function based on the cosine-squared, leading the users to post information they strongly agree with or disagree with. The latter represents the scenarios in which users react to the news. The social network algorithm must define the users who receive the information if the post is transmitted in a real social network. In order to simulate this step, we modeled the *post distribution*. More specifically, with the basis of the difference between the opinions of the posting user and their neighbors, another probability function is defined. Again, several different *post distributions* were tested. Finally, if the users receive posts that disagree with their opinions, there is a chance of breaking the friendship, which is modeled as a third probability function. If the friendship is broken, the respective edge is rewired. Several variations of *post transmission* and the *post distribution* were tested, as well as different network structures. Fig. 1 illustrates one step of the opinion dynamic and four possible results. Our dynamic converged into many different scenarios, including opinion consensus, polarization, and the formation of echo chambers. For various dynamic configurations, friendship rewiring can help promote echo chamber formation. However, even without allowing friendship rewiring, this effect can also occur for specific networks with well-defined community structures. Finally, we compare the results with real social networks. We show that the outcomes of our model are similar to the real scenario in terms of polarization and echo chamber formation. Our outcomes suggest that the social network algorithm can be important to mitigating or promoting opinion polarization. The work described in this abstract has been published in (4), and the source codes can be found at `https://github.com/hfarruda/OpinionPolarization`.

References

[1] C. Castellano, S. Fortunato, and V. Loreto, Reviews of modern physics **81**, 591 (2009).
[2] J. Lorenz, International Journal of Modern Physics C **18**, 1819 (2007).
[3] K. Sznajd-Weron, J. Sznajd, and T. Weron, Physica A: Statistical Mechanics and its Applications **565**, 125537 (2021).
[4] H. F. de Arruda, F. M. Cardoso, G. F. de Arruda, A. R. Hernández, L. da Fontoura Costa, and Y. Moreno, Information Sciences **588**, 265 (2022).

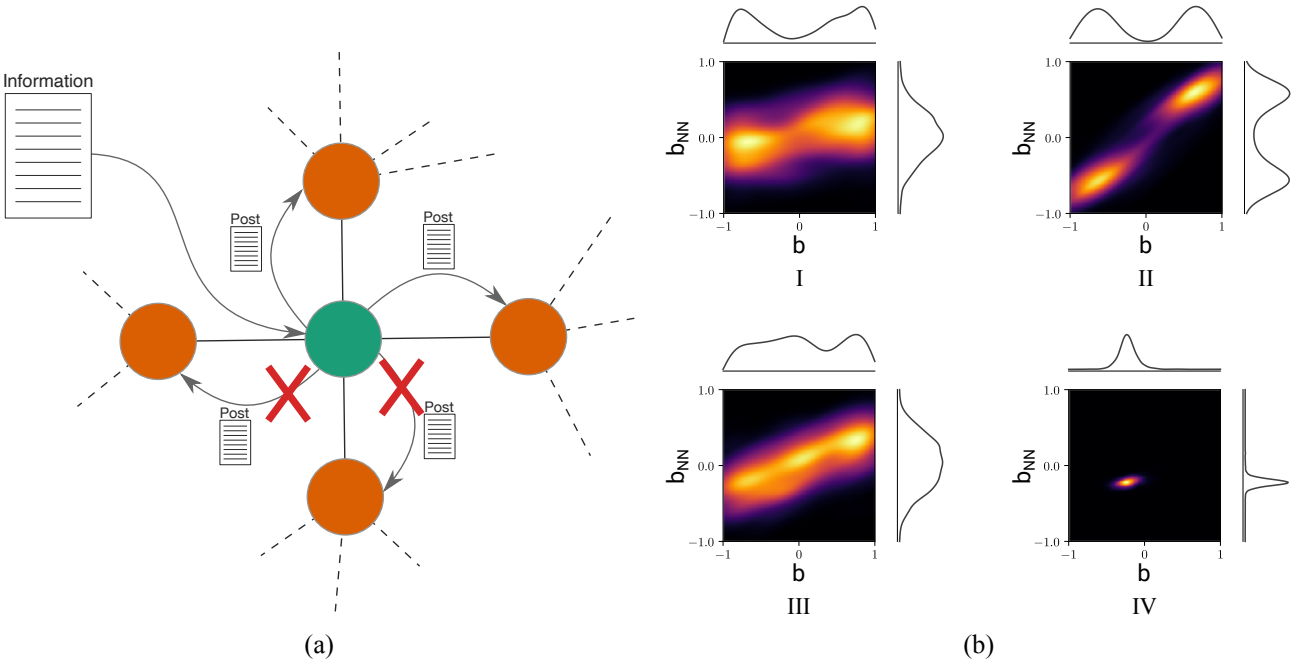

Figure 1: The scheme of the proposed model is shown in panel (a), in which a randomly selected user (green node) posts a piece of information, and two orange nodes receive the post. Panel (b) shows outcomes obtained from the execution of our dynamics, which illustrates some possibilities of the resultant opinion distributions. These plots represent the opinion assigned to nodes, b, against the average opinions of their neighbors (b_{NN}). Heatmap I represents the cases in which the opinions are polarized, but there is no echo chamber formation. This is evidenced by the fact that values of b_{NN} tend to be close to the average. In heatmap II the opinions are also polarized, but the values of b_{NN} tend to be closer to 1 or -1, representing the echo-chamber formation. In heatmap III, there is no polarization nor consensus, but extreme opinions exist. Another possible scenario is when the consensus is reached, illustrated in heatmap IV.