9^{th} International Conference on Computational Social Science IC^2S^2
July 17-20, 2023, Copenhagen, Denmark

Separating Polarization from Noise: Comparison and Normalization of Structural Polarization Measures

network science, polarization, twitter, community detection, statistical significance

Extended Abstract

Political polarization is a complex and significant phenomenon that affects social and political systems [1, 3]. Measuring the level of polarization is crucial to understanding and studying this phenomenon [2]. In recent years, various methods have been developed to measure polarization in social networks by examining their structural characteristics. These methods are typically based on a pipeline where the network is first clustered and the result of this clustering process is fed into a polarization metric calculation, which makes them susceptible to overfitting: one might not even reliably distinguishing polarized networks from random networks that have similar characteristics.

To address this issue, we analyzed eight commonly used methods for measuring structural polarization in social networks [4]. We found that all of these methods produced high polarization scores even for random networks with similar density and degree distributions to typical real-world networks. Moreover, some of the methods were sensitive to the degree distributions and relative sizes of the polarized groups, indicating that they may not be suitable for all types of networks.

Fig. 1A gives a detailed view of the poplarization scores (three out of eight scores are shown in this extended abstract, see [4] for the similar results for all the remaining scores). It can be used to read scores for each of the original networks and the corresponding random networks. Modularity and Adaptive EI Index are some examples of methods for which the average degree (d = 0 model) already explains most of the observed scores, and for the Random Walk Controversy score the degree sequence (d = 1) is usually a very good predictor. It is concerning to note that some randomized networks exhibit higher polarization scores than the corresponding original network, as demonstrated in the same figure.

To improve the reliability of these methods, we proposed a normalization technique for the existing scores and a minimal set of tests that a score should pass to be suitable for separating polarized networks from random noise. To remove the effect of network size and the degree distribution, we computed the averaged polarization score for multiple instances of the network shuffled with the configuration model, and subtracted it from the observed polarization score value. The normalization increased the performance of the scores by 38%-220% in a classification task of 203 manually curated networks, demonstrating its effectiveness (Fig. 1B). We also found that the choice of method is not as important as normalization, after which most of the methods have better performance than the best-performing method before normalization.

Overall, this work provides a critical assessment and comparison of various methods for measuring structural polarization in social networks. The proposed normalization technique and tests could be used to improve the reliability of these methods and facilitate their application in various contexts.

References

[1] CHEN, T. H. Y., SALLOUM, A., GRONOW, A., YLÄ-ANTTILA, T., AND KIVELÄ, M. Polarization of climate politics results from partisan sorting: Evidence from finnish twittersphere. *Global Environmental Change 71* (Nov. 2021). Publisher Copyright: © 2021 The Author(s).

[2] GARIMELLA, K., MORALES, G. D. F., GIONIS, A., AND MATHIOUDAKIS, M. Quantifying controversy on social media. *ACM Transactions on Social Computing 1*, 1 (2018), 1–27.

[3] MASON, L. "i disrespectfully agree": The differential effects of partisan sorting on social and issue polarization. *American journal of political science 59*, 1 (2015), 128–145.

[4] SALLOUM, A., CHEN, T. H. Y., AND KIVELÄ, M. Separating polarization from noise: Comparison and normalization of structural polarization measures. *Proc. ACM Hum.-Comput. Interact. 6*, CSCW1 (apr 2022).

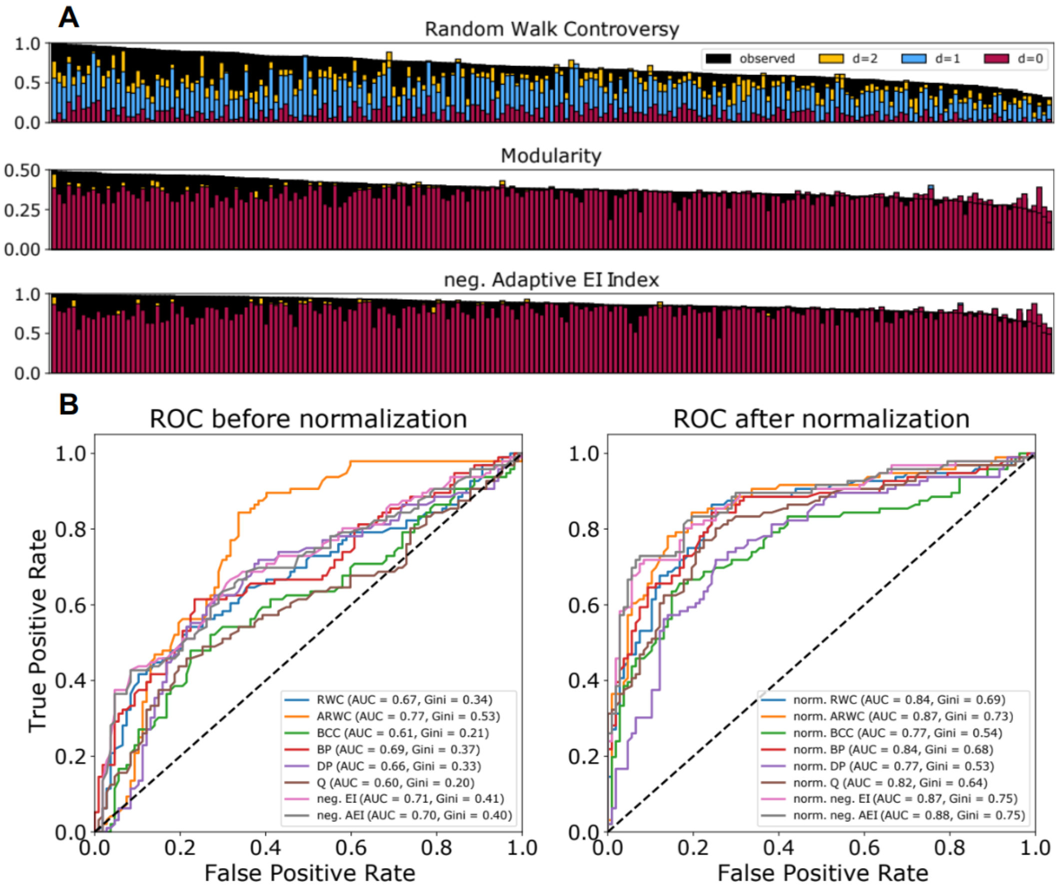

Figure 1: **A** Polarization scores for the 203 observed networks and their shuffled versions. Each bar corresponds a score, and scores for a network and its randomized versions are on top of each other, ordered from bottom to top in the following order: observed network (black) and randomized networks where degree-degree sequence (d = 2, yellow), degree sequence (d = 1, blue), or average degree (d = 0, red) is preserved. An interpretation for the figure is that, the amount of color that is shown tells how much of the total bar height (the score value) is explained by the corresponding network feature. Note that in some cases, the randomized networks produce higher scores than the original network and in this case the black bar is fully covered by the colored bar(s). In this case we draw a black horizontal line on top of the colored bars indicating the height of the black bar. **B** ROC curves, Gini coefficient values, and AUC values for the task of predicting manually curated labeling of polarized and non-polarized networks. The results shown (left) for the score values before the normalization and (right) after the normalization with denoised scores. Figures are based on all the 203 empirical networks.