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

Quantifying the structure of controversial discussions with unsupervised methods: a look into the Twitter climate change conversation

Keywords: climate change, Twitter, retweet networks, clustering, unsupervised learning

Extended Abstract

Social media is crucial for information consumption and public opinion formation. It leverages communication channels between one-to-many and many-to-many in a decentralized way by enabling its users to choose whom to follow and interact with, thus democratizing information access and spreading [1]. However, the decentralized nature of social media promotes that users consume information mostly aligned with their beliefs and interact with like-minded individuals and communication channels, creating polarization and echo chambers within the population. [2]. Authors have studied these behaviors under the sociological frameworks of confirmation bias [3] and homophily [4]. However, the mechanisms driving these behaviors in online social media are still in dispute, and the structure of important ongoing discussions, such as the one on climate change, is relevant to characterize and understand.

Experimental design. This study aims to investigate the factors contributing to polarization and the formation of echo chambers within the Twitter discourse on climate change. To achieve this, we employ unsupervised methods that require minimal computational resources to analyze the social network dynamics of the climate change Twitter discussion. Specifically, we focus on the temporal, weighted retweet networks of the climate change conversation during the year 2019, a period marked by significant climate social movements. We gather a comprehensive dataset comprising 41.8 million climate-related tweets posted by 8.7 million users between 1st March and 1st December 2019, with retweets accounting for 73% of the total volume.

Our research examines the information sources, or *chambers*, of the audience of leading users driving the Twitter climate conversation. In essence, a leading user's *chamber* is defined as the set of users retweeted by their audience, and it functions as the many-to-many information sources associated with a leading user's audience [5]. Our goal is to analyze the ideological similarities dynamically between the leading users and to investigate how their associated chambers interact throughout the year. We are particularly interested in exploring how the network structure's characteristics relate to exogenous events and the climate ideology of the Twitter population.

Results. For the set of leading users, we take the top 50, which, collectively, explains almost 50% of the total volume of retweets in the conversation. We measure the ideological similarities between every pair of leading users by computing the Jaccard similarity between their chambers. The chamber overlap distribution is clearly bimodal, with a low-overlap peak and a significantly higher-overlap peak (see Fig 1). Thus, we classify the leading users into two groups using an unsupervised spectral clustering algorithm [6] and recognize one group as the *climate believers* and the other group as the *climate skeptics*, validating previous findings [7] (see Figure 2, where we show the block structure of the leading users' similarities).

Based on the bimodal structure of the chamber overlap distribution, we define an *echo chamber* as the union of the same group's leading users with their associated audiences and chambers. Such a bimodality implies several features that characterize echo chambers: homophilic interactions drive their formation [7], actors inside them choose to preferentially connect with the exclusion of outsiders, and attitudes and beliefs stay inside groups of like-minded people. We identify, on average per week, 15% of the total retweeting population as climate believers, 3% as climate skeptics, and only 0.3% of users classified as both believers and skeptics, suggesting that the cross-communication between echo chambers is negligible.

Furthermore, we design an *ideology score* that uncovers the ideological position of any high-impact user, even when she has not been observed previously. This score depends on the proportion of her audience that belongs to either of the echo chambers. We find that the ideology score distribution over all the high-impact users is bimodal, with modes at opposite extremes of the ideological spectrum (see Figure 3). This finding reinforces our claim that the users inside the echo chambers are mostly like-minded. We validate the ideological positions of some high-impact users discovered with the ideology score.

Using the uncovered ideologies, we *augment* our original echo chambers by combining them with the audiences of the high-impact users with the same ideological position. Thus, the augmented echo chambers constitute more than half of the weekly retweeting population (46.2% climate believers, 6.5% climate skeptics, and only 0.3% of users classified in both ideological groups). We observe that, in most cases, the number of users in each echo chamber is roughly stable throughout the year. However, we find a strong positive correlation between the dates of the main #FridaysForFuture strikes and the skeptics' echo chamber sizes but not with the believers'. We find it remarkable that we can identify the peak activity of certain parts of the population by using completely unsupervised methods. Moreover, we measure the flux of users within echo chambers as a function of time, where we find that most ($> 80\%$) users leave their echo chambers from one week to the next (see Figure 4. This result suggests that the stable properties of the echo chambers are an emergent feature of the system and not imposed by a fixed set of users.

References

[1] Andrea Ceron, Luigi Curini, and Stefano M Iacus. "First-and second-level agenda setting in the Twittersphere: An application to the Italian political debate". In: *Journal of Information Technology & Politics* 13.2 (2016), pp. 159–174.

[2] Orowa Sikder et al. "A minimalistic model of bias, polarization and misinformation in social networks". In: *Scientific reports* 10.1 (2020), pp. 1–11.

[3] Peter C Wason. "On the failure to eliminate hypotheses in a conceptual task". In: *Quarterly journal of experimental psychology* 12.3 (1960), pp. 129–140.

[4] Miller McPherson, Lynn Smith-Lovin, and James M Cook. "Birds of a feather: Homophily in social networks". In: *Annual review of sociology* 27.1 (2001), pp. 415–444.

[5] Hai Liang et al. "How did Ebola information spread on twitter: broadcasting or viral spreading?" In: *BMC public health* 19.1 (2019), pp. 1–11.

[6] Bojan Mohar. "Some applications of Laplace eigenvalues of graphs". In: *Graph symmetry*. Springer, 1997, pp. 225–275.

[7] Matteo Cinelli et al. "The echo chamber effect on social media". In: *Proceedings of the National Academy of Sciences* 118.9 (2021).

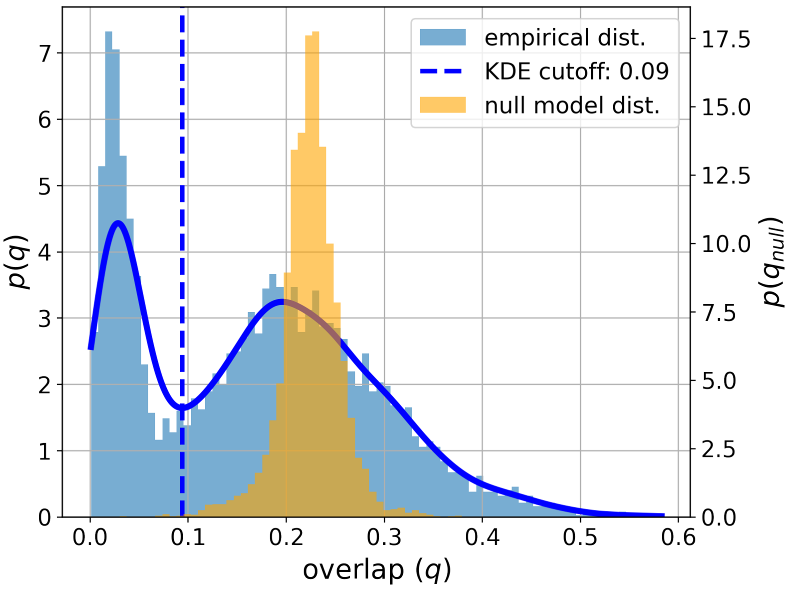
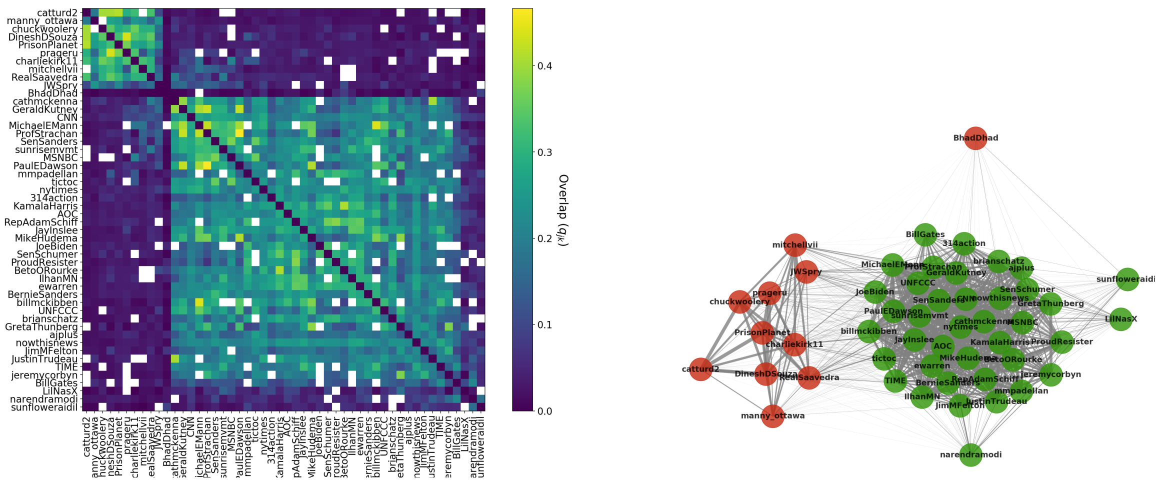


Figure 1: **Aggregate chamber overlap distributions for the empirical and null model networks.** We construct the aggregate empirical overlap distribution (blue, left y-axis) by concatenating the Jaccard overlap pairs, q^t_{ij}, for every week t on the dataset. We observe a *bimodal* structure characterized by a low-overlap peak, $q_{off} = 0.04 \pm 0.02$, and a high-overlap peak, $q_{in} = 0.23 \pm 0.08$. We compute the null-model chamber overlap distribution (orange, right y-axis) using the configuration model over the empirical degree sequences for every week. We observe that the configuration model predicts a *unimodal* overlap distribution characterized by $q_{null} = 0.22 \pm 0.03$ without any noticeable skew.



(a) Aggregate chamber overlap matrix

(b) Aggregate similarity network

Figure 2: **Aggregate chamber overlaps between the leading users**. *a)* Aggregate chamber overlap matrix, $\boldsymbol{Q}$, of every leading user pair. White pixels represent leading users that were never present simultaneously during the same week. We order the users in $\boldsymbol{Q}$ according to the rank we obtain with the *unsupervised* spectral clustering algorithm. *b)* Weighted similarity network constructed from $\boldsymbol{Q}$. We color the nodes according to the partition obtained by unsupervised spectral clustering. We identify a group of *climate believers* (green) and a group of *climate skeptics* (red) according to the users' profiles within those groups.

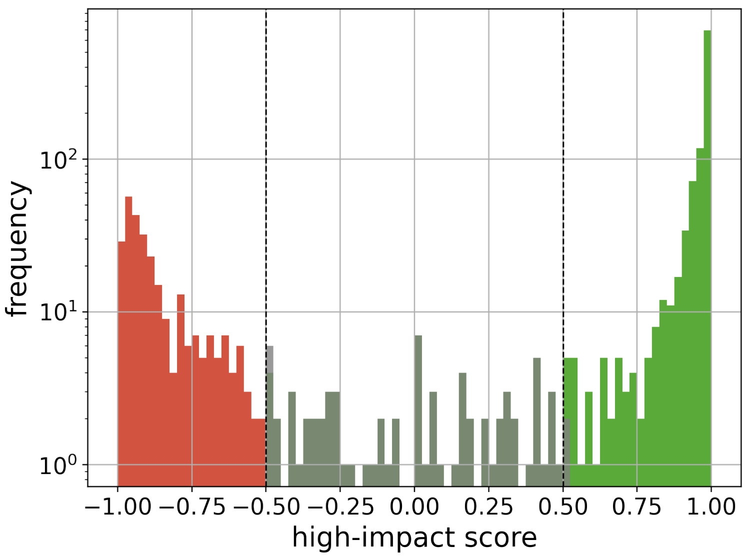


Figure 3: **Ideology scores distribution** for the high-impact users, where we find a clear bimodal structure. Users with $s_i > 0.5$ (green) are assigned as climate believers and those with $-s_i > 0.5$ (red) are assigned as climate skeptics. Users with intermediate scores are not included in the augmented echo chambers.

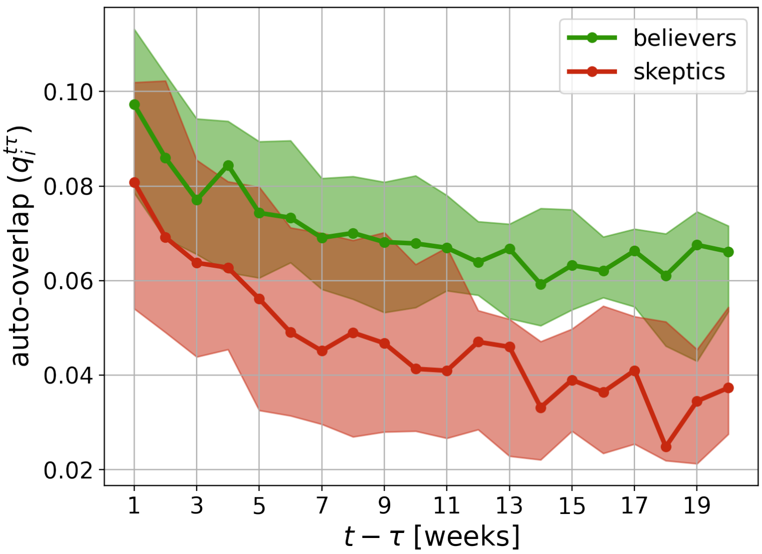


Figure 4: **Auto-overlap similarity decay for each augmented echo chamber** as a function of the time difference (in weeks) between them. The solid lines with markers indicate the median overlap similarity for each week, while the spread corresponds to the $25 - 75$ quantile spread.