Ideological, semantic and geographic proximity drive following behaviour on Twitter

Armin Pournaki, Pedro Ramaciotti Morales, Jean-Philippe Cointet

8^{th} International Conference on Computational Social Science IC^{2}S^2
July 19-22, 2022, The University of Chicago Booth School of Business, Chicago, IL, USA

Ideological, semantic and geographic proximity drive following behaviour on Twitter

Armin Pournaki, Pedro Ramaciotti Morales, Jean-Philippe Cointet

médialab, Sciences Po, Paris

Keywords: social networks, graph embedding, homophily, Twitter

The set of followed accounts substantially determines the content Twitter algorithmically delivers to a given timeline. Previous studies suggest that users can get trapped in echo chambers in which they are only exposed to content aligned to their own ideological stance (Kwak et al. 2010). However, one may wonder whether ideological alignment is the sole and unique determinant of social ties in social media. This work aims at identifying alternative mechanisms that bind users together and to measure how determinant they are. We show that geographical proximity as well as shared interests also play a significant role in the formation of links, jeopardizing the classical model of social media as a place where people only exchange in-between politically aligned individuals.

To examine the relation of political alignment and following, we focus on a set of accounts for which ideological positions can be inferred. We consider 230,911 accounts that follow French parliamentarians (MPs) active on Twitter. To this bipartite graph of MPs and their followers, the so-called ideological embedding procedure from Ramaciotti Morales, Cointet, and Muñoz Zolotoochin (2021) is applied as a starting point for the inferred opinion space. In essence, the method consists of embedding the bipartite graph of users to MPs in a latent lower-dimensional space through Correspondence Analysis (CA), which produces a low-dimensional *homophily embedding* in which users that are close to each other have a higher probability of following the same parliamentarians . To interpret the meaning of the emerging latent dimensions, the positions of some reference points along each dimension are correlated with positions in opinion scales available in political opinion surveys. This produces a score for each account along three ideological scales: one related to Left-Right cleavages (”left-right”), a second related to attitudes towards trade protectionism and European integration (”local-global”), and a third related to attitudes towards immigration and multiculturalism (”immigration”). We interpret these scores as positions in a three-dimensional latent ideological space, in which we can compute pairwise distances between the accounts. Collecting the follower relations between the users allows us to relate distance in the ideological space to link creation in the social network. Furthermore, we go on to infer two other latent spaces from the metadata: (1) the latent *social* space, inferred from the actual adjacency matrix of the *full* follower network (not the bipartite) using CA and (2) the latent *semantic* space, inferred from the Twitter account descriptions using state-of-the-art sentence embeddings [1]. Finally, we compute the geographic space inferred from each account's location entry, leaving out locations that contain only ”France”, or satirical locations such as ”somewhere in space”.

For each metadata-based distance metric at our disposal, we divide the overall set of edges in the graph in into 10 equally-populated bins, each bin corresponding to the interval of distance for which 10% of the accounts are connected to one another. We then proceed to measure the propensity of link formation which compares i) the number of actual links between a given set

[1] We used ”distiluse-base-multilingual-cased” in the Python library `sentence-transformers`.

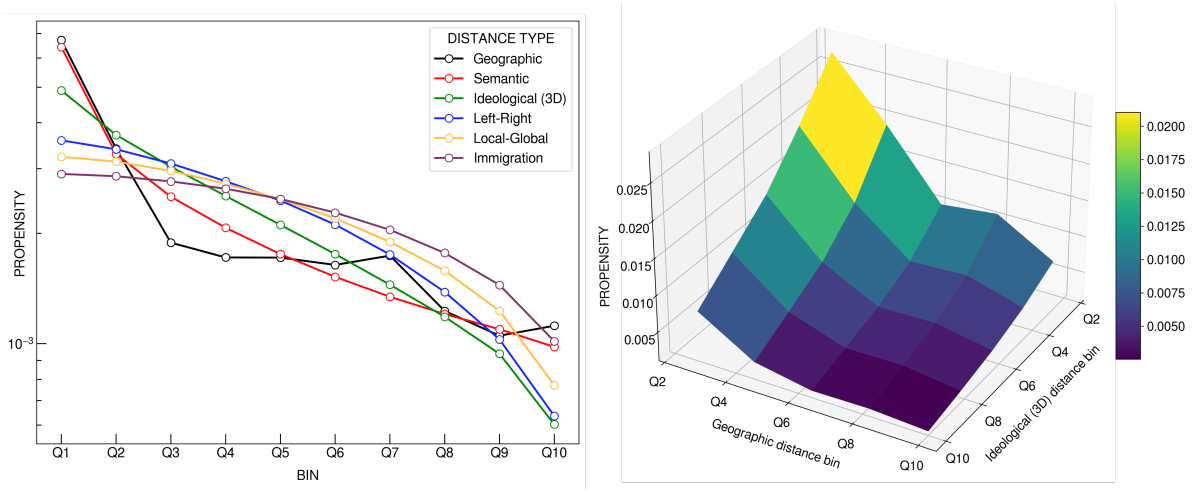

Figure 1: Link propensity scores measuring the effect of different metadata on following behaviour. The propensity decreases monotonically for all distance types except for the geographic distance between Q3 (300km) and Q7 (700km), a distance interval that is neither local nor too distant. The two-dimensional propensity with respect to semantic and geographic distance (right) shows that even though accounts may be far geographically, their chance of following each other is non-negligible if they are close ideologically.

of node pairs to ii) the maximum number of links that could exist between them. The propensity for different distance types on varying distance intervals is shown in Fig. 1 (left). High values of propensity indicate a strong likelihood for link formation, which occur at low ideological, geographic and semantic distance. This is in line with previous results that show the importance of geographic proximity as a driver for tie formation on Twitter (Takhteyev, Gruzd, and Wellman 2012). Our finding further substantiates the observation that social media draw a digital map of existing geographic structures (Barnidge 2022). Furthermore, the visible effect of semantic proximity, computed through a sentence embedding of Twitter descriptions, reveals other factors of preferential attachment related to shared features such as profession, interests or socio-economic status. Examining the interplay of ideological and geographic distance in Fig 1 (right) shows that there is non-negligible propensity value when users are geographically distant but ideologically aligned. By comparing the effects of ideological, semantic and geographic proximity on link creation, this contribution sheds light on the different mechanisms that shape the homophilous structure of social networks.

References

Barnidge, Matthew (2022). "How Geographic Mobility Contributes to Exposure to Political Difference on Social Media Platforms". In: *Telematics and Informatics*, p. 101781.

Kwak, Haewoon et al. (2010). "What is Twitter, a social network or a news media?" In: *Proceedings of the 19th international conference on World wide web*, pp. 591–600.

Ramaciotti Morales, Pedro, Jean-Philippe Cointet, and Gabriel Muñoz Zolotoochin (2021). "Unfolding the dimensionality structure of social networks in ideological embeddings". In: *Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 333–338.

Takhteyev, Yuri, Anatoliy Gruzd, and Barry Wellman (2012). "Geography of Twitter networks". In: *Social networks* 34.1, pp. 73–81.