Us vs. Them – Understanding the Impact of Homophily in Political Discussions on Twitter

Danula Hettiachchi, Tanay Arora, and Jorge Goncalves

School of Computing and Information Systems,
The University of Melbourne, Australia

Abstract. Analysing homophily, i.e. people's tendency to associate with others with similar social attributes, can help us unravel and better understand user behaviour in social media. In our work, we analyse the impact of homophily in discussions regarding the Citizenship Amendment Act (CAA) on Twitter. The Indian Government enacted CAA to provide relaxation in the citizenship process to religious minorities in three neighbouring countries. While it was lauded by many, it also fuelled backlash amongst some Indian citizens, resulting in the emergence of two distinctive political dispositions regarding this matter. We collected 78,004 Tweets, including 11,794 original Tweets during a period of two weeks shortly after the ruling, and examined ways of potentially reducing homophily and therefore minimise the presence of *echo chambers*. In particular, we investigated users' political dispositions and expressed sentiment, and how these two social attributes influence homophilic social ties and interactions. Further, we discuss how our findings can be used in social networks to allow people with diverse viewpoints and emotional attitudes to interact with each other in a positive and constructive manner.

Keywords: Social media · Homophily · Twitter · Political discussion.

1 Introduction

People tend to make connections and interact more with people who are similar to themselves in social characteristics such as demographics, occupation and political affiliations. This general social phenomenon, known as homophily, implies that distinction in social characteristics renders network distance, i.e. the number of connections through which any piece of information must travel to connect two individuals [59]. Homophily is also pervasive within online social networks and influences information propagation characteristics [22], which has broader implications on how people seek to form social ties [26], interact with online content [23], and develop common interests over social media channels [46].

While prior work has investigated homophily in social networks, there is limited understanding on prospects of reducing homophily in networks and promoting the positive interactions between users with different views and emotional attitudes, particularly in the Global South. Thus, we analyse homophily based on the network ties and the content generated over Twitter amidst an Indian political scenario, in which polarity among two-parties led to violent protests.

The Citizenship Amendment Act (CAA), was enacted by the Government of India on 11th December 2019 to grant citizenship to illegal migrants of Hindu, Sikh, Buddhist, Jain, Parsi, and Christian religious minorities, who escaped persecution from Pakistan, Bangladesh and Afghanistan before December 2014 or feared persecution in those countries [47]. The Act relaxed the residence requirement for naturalisation to such persecuted minorities. While it witnessed widespread support within the country, there was a backlash too. It was criticised as discriminatory based on religion since Muslims, in particular, were left out of the scope of the citizenship eligibility criterion in the Act [17, 63]. The Act gave rise to two polarising mass movements, one in the support for the Act, and one entirely protesting against it. During these events, Twitter became an active channel for disseminating information regarding the implications of the Act and fostered widespread discussions within and across communities.

Twitter is an open information network and an active platform for social and political engagement and discussions among politicians, social activists and the general public alike [2, 6, 25, 27, 66, 70]. In the past, despite facilitating many popular online movements like #MeToo [57, 80], #BlackLivesMatter [46], or against police brutalities via #NYPD [51], Twitter and other social media channels have been under a critical lens of social activists, journalist as well as academics. They have actively voiced their concerns over social media networks playing a significant role in polarising people and placing them in their own ideological bubbles, popularly coined as 'echo-chambers' [72] or 'filter-bubbles' [65]. Although there is empirical evidence in academia that suggests that homophilic attitudes give rise to the creation of polarised communities or 'echo-chambers' [32, 65, 72, 79], other studies contradict such claims [5, 12, 49].

Although there are broader studies on the effect of homophily along the party lines [52], few studies examine the impact of political dispositions on homophily in online spaces [20]. Moreover, the role of sentimental attitudes in examining homophilic behaviour in online spaces has not been studied together with political dispositions [38]. Only a few studies have drawn a comparison between the two social attributes [15]. Further, there is no consensus on whether content similarity is a driving factor for homophilic ties [1, 19, 31]. In this study, we analyse homophily based on political disposition and sentiments expressed over Twitter encircling the public discourse on Twitter over CAA and draw a comparison over the role of each attribute in giving rise to political homophily within the network. We collected Twitter data with the keyword #CAA for a period of two weeks shortly after the ruling using the Twitter Standard API. Initially, we classified all users based on political orientation and sentiment polarity. Our approach has been adapted from the works of Caetano et al. [15] who studied homophily in the context of 2016 US presidential elections. Then for each user group, we aim to uncover homophily. We also examine how the use of common hashtags and similar topical interests impact homophily.

Our findings reveal that political disposition was prevalent in how people formed social ties. A high magnitude of homophilic behaviour was observed in terms of interactions, engagement, again majorly due to the users' political disposition. Informed by our findings, we reflect on several design recommendations and future directions which can separate users from their ideological bubbles and expose them to more different views over social media channels to foster inclusiveness, a cornerstone of democracy.

2 Related Work

2.1 Political Homophily

Political homophily remains a common research area for sociologists, political psychologists, scientists and scholars in social computing alike. According to Pew Research, in the past two decades, the number of Americans with mixed ideologies has seen a significant decline and political dispositions have become distinctively liberal or conservatives [67]. For instance, in 1960, approximately 5% of Americans felt displeased with their children marrying outside the party lines, whereas by 2010 the numbers shot up to 50% for Republicans and 30% for Democrats [49]. The implications of political homophily are also quite significant in the corridors of public administration. Past research in the US [29] and South Korea [49] provide empirical evidence that similar political ideologies increase the likelihood of inter-organisational coordination and reduce the transactional cost in the collective effort of decision making. In another study concerning public life, Iyengar & Westwood [50] demonstrate the effects of political partisanship in a survey-based experiment in which individuals were asked to evaluate candidates' profiles for high-school scholarships. They found partisans were more biased towards their fellow partisans in granting the scholarships.

Nonetheless, political homophily has likely implications when it comes to our interpersonal preferences. Huber & Malhotra [42] reveal that people are more likely to engage with or show interest in profiles which are more politically compatible to them when selecting a dating partner in contemporary times. Interestingly, in the past, political similarity has taken precedence over other influential factors like ethnicity or education. In the US, it has been alleged that the Democrats and Republicans are more likely to move to neighbourhoods which they deem more politically harmonious [67], popularly coined as 'partisan sorting' [8]. However, experimental and survey-based studies [30, 61] report that although politically harmonious communities may be more desirable or highly rated by individuals [44], it does not take precedence over other factors, such as affordability.

2.2 Role of Social Media

Prior work highlight different social or cultural factors that are dominant in sorting people in their ideological cocoons. Within these decades, while other social, cultural, legal developments remain a driver for the mass polarisation across societies [67], one of the reasons which academics attribute to growing political homophily is selective exposure or confirmation bias over social media channels and news portals. Lewicka [53] describes confirmation bias as a 'survival equipment' for humans, as humans we have a tendency to automate our routines in a way to obtain preferential results free from futile deliberations in order to keep our mind space centered around more important life decisions. For example, investigating the preferences for media consumption, Iyengar & Hahn [48] reported that Republicans preferred Fox News as their preferred media source and avoided news from CNN and NPR, whereas the Democrats displayed the exact opposite behaviour.

In the context of social media, selective exposure has been studied on the basis of structurality of the network as well on the content with which people interact. Within a social network, users are treated as nodes, and edges are treated as relationships among users. Himelboim et al. [39] developed a structure-based cluster analysis method to discover patterns of selective exposure among conservatives and liberals over the U.S. President's State of the Union speech in 2012. The analysis demonstrated distinctive clusters consisting of only self-identified conservative users, while the liberal clusters illustrated a mix of self-identified liberals and mainstream media organisations. Additionally, it can be inferred from the findings that the conservative users are less likely to form links with the traditional media houses, a pattern also observed in other work [40]. In another study over Twitter, Williams et al. [79] reveal a high presence of communities with strong attitudes (activists or sceptics) with very minimal presence of moderate communities over the debate on climate change. A similar observation has been highlighted in other studies [4, 19] which have showcased strong structural connections among ideologically similar users.

On the other hand, studies have suggested that social media conversations around controversial topics tend to exhibit a high level of emotions corresponding to political views or opinions [31]. Additionally, it also has been observed that public discourse around political issues comprises highly negative emotions [3], while on the contrary some studies suggest that political discussions are dominated by positive expressions [84]. Therefore, it can be expected that people may take up similar political dispositions based on the sentiments expressed during a public debate over a political issue, given the relationship between attitudes and political behaviour have been longitudinally studied in the past [43]. Moreover, the role of content on social media in diffusing selective exposure cannot be underestimated. Although content over Twitter can take many forms, this study focuses on tweet text, hashtags and user replies. Researches in the past have shown that the aggregated hashtags facilitate engagement where conversations around particular key issues or events happen [13, 14]. Hashtags may encourage similar users to use similar hashtags while other users might be left out of the discussion. Goncalves et al. [31] were able to predict users' political leaning using natural language processing techniques and discovered latent semantic layers by aggregating their hashtag usage. Previous work also suggests that users using a mixed of ideological hashtags were able to produce more inter-ideological interactions than those who mostly use partisan hashtags [20]. On the contrary, although in the context of a cause related marketing campaign, Xu et al. [81] suggest that use of common hashtags increases the likelihood of like-minded people interacting more and at the same time alienating other users. Therefore, it becomes important to analyse whether usage of common hashtags causes people to segregate themselves into distinctive communities.

When it comes to evaluating content similarity, researchers have made use of machine learning techniques to explore semantic similarity between tweet texts. Unsupervised machine learning algorithms such as Latent Dirichlet Allocation (LDA) [9] have been useful in discovering users' topical interests across tweet texts. Wang et al. [78] present strong evidence by employing multiple community detection algorithms over follower topology of the network and demonstrate that the structure-based communities generate common interests among the community members. Kang & Lerman [52] investigated topical interests of users by using lists, lists on Twitter are like groups curated by users on Facebook. The research demonstrated that users who were more topically similar were more likely to be linked with a follower relationship than others who were less topically similar. However, a multi platform study in a wider context Bisgin et al. [7] found that the friendship ties were only 1% similar in their topical interests and over 95% of the friendship ties were less than 50% similar. Given such variations in the previous studies and the nature of the discourse, we are also interested in evaluating to what extent does topic similarity influence user connections and interactions within the network.

3 Methods

3.1 Dataset

In India, a large number of citizens engage with Twitter, and it has recorded a user base of 18.9 million active users as of 2020[1]. On Twitter, you can engage with other users by *following* a user, *retweeting* a tweet from another user, or *mentioning* the user in your own tweet. These engagements could be either unidirectional or reciprocal (*e.g.,* two users mentioning each other in their tweets). Users may also tag their tweets with a hashtag, usually with a '#' followed by the hashtag name. This functionality enables users to search tweets easily belonging to a certain category of hashtags. During the public discourse over the Citizenship Amendment Act (CAA), *#CAA* was a common trending hashtags frequently used by users on both sides of the discourse. However, there were various hashtags used in different contexts during the discourse. *#ISupportCAA*, *#IndiaSupportsCAA* emerged as trending hashtags in support of CAA and *#IndiaDoesNotSupportCAA*, *#IndiaAgainstCAA* against CAA. In conversations both for and against the protest *#CAAProtest*, *#CAANRCProtest*, *#ShaheenBagh*

[1] https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries

(site of a major protest) seemed to be trending. However, post the riots in Delhi, *#DelhiViolence*, *#DelhiRiots2020*, *#DelhiBurning*, *#DelhiGenocide* were some of the common hashtags observed in the conversations pertaining to the violence in Delhi.

Data Collection The data was collected on Twitter social network from 16th February 2020 to 1st March 2020 using the keyword #CAA through the official Twitter standard sandbox API. The mentioned duration is important due to two major geopolitical events; Delhi Riots, 2020 and President Trump's visit to India during the mentioned time period, when extensive discussions happened over Twitter regarding #CAA. We retrieved a total of 78,004 tweets with the majority being retweets (66,210) and the remaining 11,794 being original tweets posted by individuals.

As the focus of the study was on individual behaviour, we excluded all the tweets posted by media houses. However, tweets from political party handles were included, primarily because they can influence masses' opinion. Additionally, we excluded 215 tweets (1.8%) that were labelled as 'unsure' during the political analysis. The final dataset comprised a total of 9,072 original tweets posted by a total of 5,940 users. Regarding connections, the dataset contained 63,065 (reciprocal) and 145,163 (unidirectional) follow connections, 298 (reciprocal) and 1746 (unidirectional) retweet connections, and 108 (reciprocal) and 2610 (unidirectional) mention connections.

3.2 User Classi cation

All the users were classified based on their political standing (Pro-CAA, Anti-CAA, Neutral-CAA) and the sentiments derived post sentiment analysis (positive, negative) over their tweets. Consequently, 6 classes of users were obtained with the different combinations of the categorisations mentioned above.

Political Analysis In this step, we manually labelled and categorised users into three political classes such as Pro-CAA, Anti-CAA and Neutral-CAA. Previous work has highlighted how analysing message content can help understand the behaviour of advocates of political campaigns [69]. Similarly, users' engagement within the community can also be identified based on the usage of hashtags [51]. For example, we categorised a user tweeting with *#IndiaSupportsCAA* as 'Pro-CAA', and similarly a user tweeting with *#IndiaAgainstCAA* as 'Anti-CAA'. However, hashtags like *#CAAProtests*, *#ShaheenBagh* or *#DelhiRiots2020* were common across all types of users.

In order to further understand the nature of the overall discourse over CAA and Delhi riots, we referred to popular knowledge available in the media [41, 74] to formulate the classification criteria given in Table 1. Based on this criteria, we performed a semantic analysis over the tweets. Additionally, user profile attributes, such as profile description, were analysed to help classify each user's political orientation. We conducted the classification where two authors come from the Indian subcontinent, and one author is an Indian national and a proficient speaker of the two main languages of India. Furthermore, Hindi (the main language in India) and Urdu (predominantly spoken by Muslims, disadvantaged by the proposed act) are mutually intelligible languages. During this process, we excluded 215 tweets for which we could not determine the political class. This included tweets in languages such as Tamil, Telugu and Malayalam that are not directly related to the main parties involved in the CAA debate.

Political Orientation	Description
Pro CAA	In support of Government, BJP or party leaders, or criticising the opposition Criticising the protests or people who support the protests Affiliated with BJP Maligning Muslims
Anti CAA	Criticising Government, BJP or party leaders, or supporting the opposition Supporting the protests or people who support the protests Affiliation with opposition parties Maligning Hindus
Neutral CAA	Quoting News Articles or Ground Reports Quoting Speaking in the interest of peace and harmony Condemning violence from a neutral standpoint

Table 1: Political Classification Criteria

Sentiment Analysis Sentiment Analysis or opinion mining is a sub-branch of Natural Language Processing used to computationally extract sentiments, opinions, attitudes from a given text, based on the subjectivity of the text. In this study, we used VADER [45], a lexicon and rule based sentiment analysis approach especially designed to analyse social media text. We opted to use VADER because it derives sentiments from a text, based on syntactical and grammatical relationships among the words in a text. VADER considers the polarity and the intensity of the sentiments expressed by incorporating the order of the words used within the text.

Additionally, VADER's judgement is sensitive to emoticons, sentiment related acronyms and common slang words. Before applying sentiment analysis over tweet text, we removed hyperlinks, special characters such as '&amp', '/', other characters such as user mentions (@), hashtags (#), single quotes('), and extra in line spacing.

The VADER sentiment analysis outputs four types of sentiment polarity scores i.e. positive, negative, neutral, and compound scores pertaining to the subjectivity of the text. The first three types of scores above signify the proportion of the text which lies under the three sentiment categories. We used the compound score, which provides a metric by summing all the valence scores across the three categories and then normalising the score between -1 (most extreme negative) and 1 (most extreme positive) [38]. To evaluate the sentiment polarity of all 5940 users, a mean of the compound score for all posts of a user was obtained. Using the mean compound score and a suitable threshold, we then classified the user profiles as positive (score   0.05), negative (score   -0.05) or neutral (otherwise).

3.3 Homophily Analysis

In this study, the phenomenon of homophily is analysed under two categories: the structurality of the network i.e. follower, retweet and user mentions relationships; and on the basis of content generated within the network.

Structural Analysis Our dataset contained 208,228 follow, 2,044 retweet, and 2,718 mention connections. Using equation (1) [19], homophily was calculated for different user groups considering the user class (*e.g.,* Pro-CAA, positive) and type of connection (unidirectional and reciprocal).

$$H_i = S_i{=}(S_i + D_i) \tag{1}$$

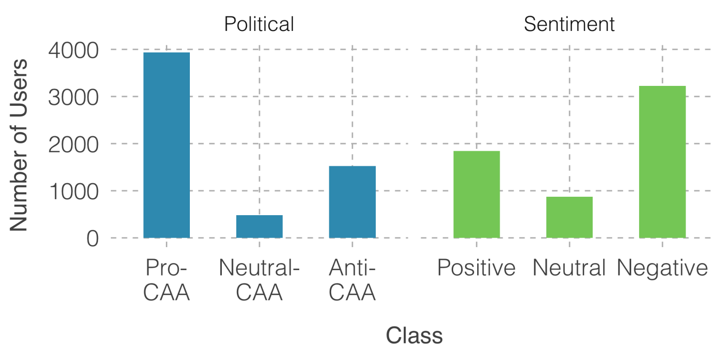

Fig. 1: Total Count vs Type of Users

H_i is the homophily index for each user group where i, S_i represents the number of homogenous connections and D_i represents the number of heterogeneous connections. As observed in Figure 1, the overall dataset is predominantly biased towards Pro-CAA users and Negative users. In order to eliminate this bias, Currarini et al. [21] recommends the inbreeding homophily index developed by Coleman [18], which we calculate using Equation 2.

$$IH_i = H_i \quad W_i{=}1 \quad W_i \tag{2}$$

Here IH_i is the normalised homophily index for each user group i and W_i is the probability of finding a user of group i within the network i.e. we obtain

W_i by dividing group size of i by the total number of users within the network. The higher the value of IH_i, greater the tendency of users of group i to form stronger connections with the similar users.

Content Analysis Next, we describe how we analysed tweets using hashtags and topic similarity.

Hashtag Analysis: Prior research shows how Twitter users leverage hashtags to disseminate ideas and opinions around social and political issues [83]. Additionally, hashtags streamline user navigation and help identify relevant conversation topics, leading to increased awareness and social media debates around the concerned issue [11, 55]. Therefore, we argue that analysing structural ties among people who use common hashtags will help uncover homophily. Specifically, we want to know if people get influenced by their structural ties and contribute to conversations similar to their connections over a public discourse. In this hashtag analysis, we have excluded #CAA as it was the initial search criteria for collecting the overall dataset.

We used three steps to analyse homophily based on common usage of hashtags. First, all the hashtags $HT(u_i)$ used by each user u_i were pooled into a user-hashtag matrix. Second, each user u_i is compared with all the other users within the matrix, and if two users were found to have used a common hashtag, they were paired together within a list. Third, using this list containing user-pairs with common hashtags, homophily was analysed within follower, retweet, mention relationships among these user-pairs using Equations 1 & 2.

Feature extraction with Topic-Modelling: We used Latent Dirichlet Allocation (LDA) [9] to extract latent features or topics within the tweets. Given a set of documents, LDA follows a probabilistic approach that postulates that the words used in each document can be subjected to a mixture of hidden or latent topics present in the overall corpus using equation (3)

$$p(\ ; z; w | \ ; \ ) = p(\ | \ ) \prod_{n=1}^{N} p(z_n | \ ) \quad p(w_n | z_n; \ ) \tag{3}$$

In equation (3), represents topic distribution for document , w represents the per-topic distribution, i.e. a list of words contained within a topic along with the probability of the words occurring within the topics, z represents per-document per-word topic word assignments. However, in this analysis only is of relevance as the goal is to identify relevant latent topics for each document.

Additionally, prior to creating the model there are several parameters to be supplied to the model i.e. the number of topics K to be formulated and also hyper-parameters and which control document-topic and word-topic distribution respectively. A higher value results in distinctive topics, whereas a higher value results in a uniform distribution across topics, hence topics being more similar to each other. Similarly, a higher value of indicates that topics are likely to made of mix of words, with less weight on dominant terms, whereas lower values of result in topics made of specific terms with more weight on dominant terms.

In order to choose the appropriate number of topics, the value of was set to 0.1 as we were interested in identifying distinctive topics which the tweets were composed of. Furthermore, model evaluation measures such as coherence score and perplexity were used to determine the values of K and respectively. Coherence score is defined as the degree to which high probability words appearing in a word are semantically similar, perplexity is defined as the log-likelihood of how well the corpus fits the model. After a series of experimentation with K = [2,8,14,20,.....,98] and = [0.01, 0.05, 0.1, 0.5, 1], the value of $K = 14$ and = 0.01 were chosen.

Although, short tweet text pose serious implications to the performance of the topics model, previous work has shown that aggregating all tweets of users into a author-tweet matrix gives better performance than non-pooled corpus [60]. Additionally, since LDA is a bag of words model, prior to performing topic modelling the tweet texts were tokenized; stop words, URLs, special characters were removed, bigrams were created and further lemmatized into their root form. Finally, the resulting corpus was used to train the LDA model using the Python gensim implementation[2].

Topic Similarity: To measure the similarity between two users, we calculated the distance between the topical distribution of the users using Jensen-Shannon distance metric [24].

$$JSD(P\|Q) = \sqrt{(0.5(D(P\|M) + D(Q\|M)))} \tag{4}$$

Equation 4 shows Jensen Shannon distance (JSD) where P and Q are probabilistic distributions and $M = 0.5(P + Q)$. JSD is a smoothed version of the Kullback-Leibler divergence (D) [24]. Topic similarity is 1 JSD where an output of 1 indicates user pairs with most similar topics.

4 Results

4.1 Structural Analysis

Figure 2 shows the tendency of forming homophilic connections among various types of users within the political and sentiment polarity classes.

Follow Connections Within the political class as observed in Figure 2 (left), pro-CAA and anti-CAA users both show high levels of inbreeding homophily, which means the presence of users with varied ideologies within the network does not affect the tendency to form follow-relationships. Moreover, the tendency to form homophilic ties is higher in reciprocal relationships than in unidirectional relationships among both pro and anti CAA users. In contrast, neutral-CAA users almost show baseline homophily, which suggests that they tend to follow pro-CAA users and anti-CAA users both by chance and not on the basis of personal preference [59].

[2] https://radimrehurek.com/gensim/

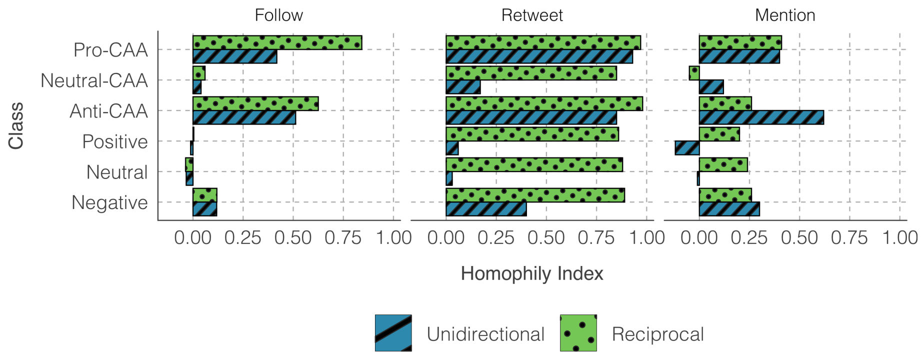

Fig. 2: Homophily-Index (IH_i) in Follow, Retweet and Mention Connections

Within the sentiment polarity class, we observe that positive and neutral sentiment users show almost baseline homophily, although negative sentiment users show marginal levels of homophily in both unidirectional and reciprocal ties.

Retweet Connections Figure 2 (middle) shows homophily analysis results among both classes of users. Unlike the follow-relationships, similar political ideology as well as similar sentiment polarity drive users to mutually retweet their similar peers.

Within the political class, it can be observed that except for neutral-CAA users in unidirectional retweet-connections, the retweeting behaviour exhibited by pro-CAA, anti-CAA, neutral-CAA users is highly preferential towards their similar peers. These findings corroborate with the findings of prior work [20, 79] suggesting that the retweeting behaviour of users is highly partisan during the discourse over a political issue.

On the other hand, within the sentiment polarity class, it is interesting to see that similarity in sentiments elicits more mutual retweet interactions than in unidirectional retweet connections. Although our results show that the users with negative sentiments do indeed show homophily in unidirectional connections, users were more likely to mutually retweet other users who resonated with their own emotional state. Therefore, sentiments in this scenario can be associated with a preference to retweet other users with similar sentiments, a behaviour also revealed in previous work by Stieglitz et al. [71] and Tsugawa & Ohsaki [75].

Mention Connections In Figure 2 (right), we observe that the mentions network appears to be less segregated than the follow and retweet networks.

Anti-CAA users display the highest level of homophily within unidirectional and lower levels of homophily within reciprocal connections. This behaviour indicates that the mutual interactions of anti-CAA users were more cross-ideological than in other unidirectional connections. Nonetheless, within reciprocal interactions, pro-CAA users appeared to be more organised among themselves than any other group, suggesting they indulged more in mutual discussions over the topic with other pro-CAA users. The neutral-CAA users display marginal levels of homophily within unidirectional and almost baseline homophily within reciprocal mention-connections. Hence, neutral-CAA users were interacting with pro and anti CAA users quite frequently.

Within the sentiment polarity class, generally all users exhibited similar modest levels of homophily in reciprocal mention-connections, indicating that users' sentiments did in fact elicit similar sentiments within the responses received by the users. However, only negative users showcase homophily in the unidirectional interactions, whereas positive users showcase heterophily and neutral users show almost baseline homophily. This behaviour suggests that posts containing negative sentiments tend to attract more attention from positive and neutral sentiment users.

4.2 Content Analysis

Hashtag Analysis We present a bi-modal analysis of usage of common hashtags i.e in terms of formed connections and the interactions between users with common usage of hashtags.

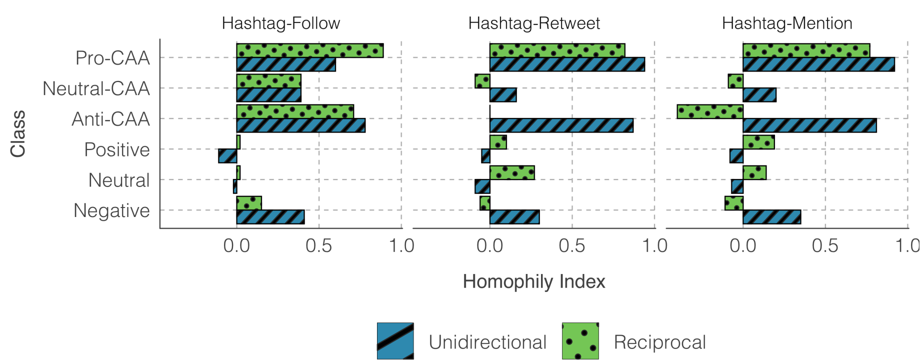

Fig. 3: Homophily-Index (IH_i) - Common Hashtag usage with Follow, Retweet and Mention Connections

Follow Connections: Figure 3 (left) shows that users who had follow-connections were more likely to use common hashtags across the political class. However, such behaviour is extreme among pro and anti CAA users than neutral-CAA users, which implies that in a political debate on Twitter, users get influenced by their followers or followee, and may contribute to a political debate with similar hashtags as their follower or followee.

In contrast, as seen in Figure 3 (left), users who discuss CAA with positive and neutral sentiments were less likely to get influenced by their follower or followee, and contributed to the discussions with a variety of hashtags. Users with negative sentiments showed a certain level of homophily with their hashtag usage, influenced by their follow-connections, and more likely to occur in unidirectional than in reciprocal connections.

Retweets: As seen in Figure 3 (middle), pro-CAA users who use common hashtags were highly likely to retweet each other and other users who used the same hashtags as them. Hence, the common usage of hashtags increased the likelihood of a pro-CAA user retweeting another pro-CAA user. On the other hand, there were no anti-CAA users who used common hashtags and retweeted each other in the dataset. However, they were highly likely to retweet another anti-CAA user if the other user used the same hashtag as them. Neutral-CAA users show low levels of homophily in unidirectional and low heterophily in reciprocal retweet connections, which means that they were more likely to get exposed to hashtags used by both pro and anti-CAA users.

Within the sentiment polarity class, users with negative and positive sentiments showed some level of homophily when they retweeted other users who used the same hashtags as them, although the likelihood of it remained low in comparison with political attitudes of users. Additionally, users with neutral sentiments in some cases did mutually retweeted other users with neutral attitudes, while their engagement in unidirectional relationships remained close to baseline homophily.

Mentions: Pro-CAA users showcase similar homophily behaviour in terms of mentions and retweets. As seen in Figure 3 (left), the use of common hashtags predominantly increased the likelihood of pro-CAA users interacting with other users. Among anti-CAA users, there are more one-sided interactions when using common hashtags. Interestingly, anti-CAA users show heterophilic behaviour within reciprocal interactions when interacting with other users with common hashtags. Moreover, the likelihood of anti-CAA users getting engaged in a cross-ideological interaction increases with the usage of common hashtags than otherwise (see Figure 2).

The use of common hashtags also decreased the likelihood of having mutual interactions with similar kinds of users among negative, neutral and positive sentiment users, who otherwise demonstrated a higher tendency to have mutual interactions with similar kinds of users.

Topic Similarity and Link Percentage In this section, we evaluate the likelihood of topically similar users connecting or interacting directly with each other. We calculated the topic similarity for each user pair based on Jensen-Shannon distance as detailed in Section 3.3. The number of links between users is binned within the intervals of 0.2 Jensen-Shannon distance units, and the percentage of links in each bin was calculated.

The results are consistent across all forms of ties i.e follow, retweet and mention connections, see Figure 4. Approximately 90% of the users across all

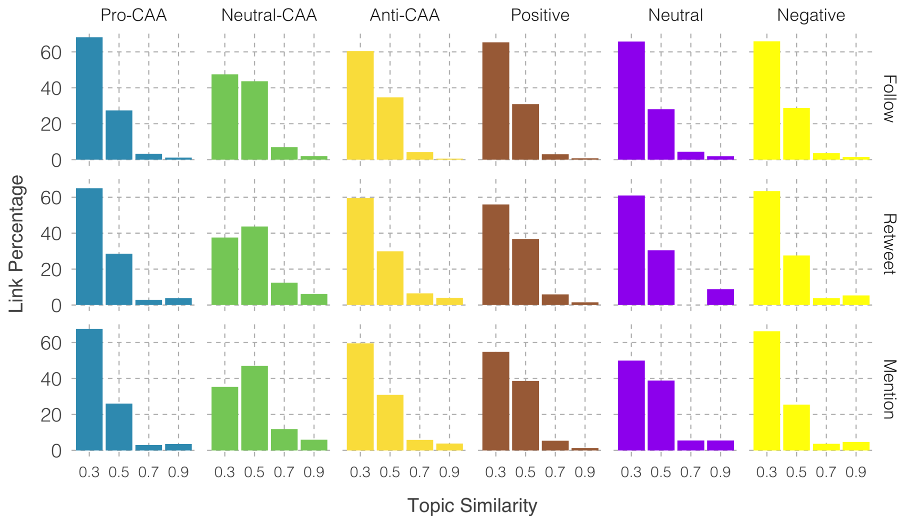

Fig. 4: Topic-Similarity vs Link-Percentage

kinds of connection were less than 60% similar in their topical interests. Based on these results, we show that similar topical interests do not influence how users seek ties or engage in discussions in a political debate over twitter.

5 Discussion

In this paper, we analyse homophily based on the structural ties and the user content on Twitter. We draw a comparison over two distinct social media user attributes, political ideology and sentiments, and further explore the significance of each attribute in the homophilic behaviour over a public discourse in India. In this section, we discuss our key findings and their practical and design implications, which may inspire future studies or implementations that aim to counter homophily in social networks.

5.1 Political Dispositions, Sentiments and Homophily

From the analysis of following and retweeting behaviour of users, more specifically within the political class, users seemed to follow and retweet their similar peers more often than dissimilar peers, a phenomenon also observed in other political debates on Twitter [19, 31, 79]. One potential cause of this homophilic behaviour in social media are personalised algorithms. Pariser [65] states that the personalisation of algorithms deployed to customise user experience exposes users to similar individuals and information repeatedly, more likely providing users easy access to information they are likely to resonate with and limiting exposure to cross-cutting content. In other words, without a feedback loop or user control on information feeds, the algorithms dictate the user experience of information consumption on social media [33, 68]. From a social behaviour standpoint, Mason [58] pointed out that partisan sorting is no longer limited to one's stance on the current issues in contemporary times. It engulfs other dimensions of one's identity, which lead us to socially sort ourselves from others who form different identities or opinions.

We also observed that the extent of homophily among users' political class was lower in mention connections. Consequently, we can infer that mention connections provide a bridge for users to have cross-ideological interactions to some extent [79]. However, we did not explore the nature of those interactions in our analysis. It is difficult to argue whether the cross-ideological interactions diluted the diversified users or made them feel more strongly about their dispositions. The social identity theory [73] suggests that the propensity to negatively engage outside the community is another way of confirming membership within a community.

Furthermore, the role of sentiments expressed by users cannot be understated when users chose to retweet. As shown in our findings, users who mainly tweeted with similar sentiments were more likely to retweet each other. This finding corroborates with prior work [16, 36] that suggests users' emotional attitudes or state influence the type of content a user may interact with. Moreover, individuals' mutual retweet behaviour in the sentiment polarity class indicates the formation of potential 'echo chambers'. We observed that users preferred to get exposed to tweets that primarily resonated with their sentimental attitudes over the discourse [40].

Overall, we conclude that political ideology was a more significant driving factor for homophilic ties over the discourse around CAA. The political ideology largely dictated with whom users prefered to form connections, retweeted or interacted with. Users' sentiments also influenced how the information was disseminated over the network. As observed in the previous research, Twitter communities grow around prominent opinion leaders who may be popular individuals, organisations, celebrities [35]. For example, in different contexts, journalists and media houses [13] and organisations [82] have emerged as influential opinion leaders during public discourses. A longitudinal study is needed to examine opinion leaders' role vs the impact of personalisation in creating homophilic ties or interactions.

Similarly, we observed that in a follow relationship, more specifically within the political class, users were more likely to use the hashtags common with their follow connections. They were also were more likely to retweet tweets of users with whom they have common hashtag usage. In other words, they were selectively exposed to hashtags that were popular in their social network, leading to selective exposure of information coming from users within the political class [4]. Additionally, this indicates that the usage of common hashtags remains a strong indicator of users segregating themselves into communities to express and also promote their ideological positions [51]. Interestingly, we noticed different behaviours across the mention connections. Pro-CAA users had more cross-ideological interactions when they did not use common hashtags as other users. However, they were highly likely to interact with other pro-CAA users whenever they used a common hashtag. In other words, their interactions became more selective with the use of common hashtags. On the contrary, when anti-CAA users used common hashtags, they were more likely to interact with users outside their ideological group. Hence, in this case, the usage of common hashtags diluted the homophily among anti-CAA users. This aried behaviour among users could also be due to the selectivity in using specific hashtags. According to Blevins et al. [10], the nature of the hashtags can be ideological (something that expresses viewpoints, positions) or conceptual (personal stories or interpretation of an event). Future studies can explore categorising hashtags further into ideological or conceptual to better understand the varied effects of homophily when using common hashtags.

Finally, based on prior research [7, 78], we analysed the likelihood of users with high topic similarity to having a follow, retweet or mention relationship. Our results indicate that topically similar users are less likely to be connected. Interestingly, on the one hand, our findings suggest that connections or engagement among users could be driven by other factors such as social reputation [56], affinity [28], social capital [54, 76], empathy [77], etc. and not topical interests. On the other hand, some topics may have attracted people with diverse ideologies and sentimental attitudes, exposing people to diverse views.

5.2 Towards Breaking the Ideological Barrier

We discuss several generalisable approaches to reduce homophily informed by our findings. In our work, political stance was the dominant factor in placing users into their ideological bubbles. Among politically engaged users, the connections and network interactions were centred around politically similar individuals. As a result, exposure to the content via hashtags also remained very selective. All these findings point towards extremely choice-based homophily [59]. Prior research has attempted addressing filter bubbles. Studying political behaviour, Nyhan & Reifler [64] indicate that direct exposure to counter opinions has mostly resulted in a phenomenon called the 'back-fire' effect [64], where people do not seem to value diversity. Hence, the recommendation systems in social media need to recommend opposing views indirectly. Some inspiration can be drawn from the work by Nagulendra & Vassileva [62] which proposed an interactive tool that visualise users' filter bubbles. The tool provides control over the algorithm by allowing users to see which topics and their connections are within their filter bubbles. Users can then choose to either stay or escape from the filter bubble. However, they only evaluated this design from a usability perspective, with less focus on the human-behaviour.

Such design suggestions seem relevant for our use case where politically engaged users were more connected with other similar users. Future research could gauge user behaviour when examining their filter bubbles containing their current connections and the users who share common topical interests but placed outside their filter bubbles. Additionally, future studies could explore whether and how people with similar topical interests get connected and continue to interact. Findings can lead to useful tools and features that can help reduce undesired homophily in social networks. We also note that in some instances, common hashtags can encourage cross-ideological interactions in politically engaged users. Hashtags can provide indirect exposure to opposing views. Tools such as the word cloud visualisation can help provide exposure to opposing views without angering users [34].

5.3 Limitations

We note several limitations in our work. First, we manually labelled the political disposition of users based on common knowledge in the media. While we assumed manual annotation is more reliable due to limited resources available around the topic, we acknowledge the potential for bias. Future research could extend our classification criteria and process. For instance, they can also utilise crowd wisdom in labelling users' political dispositions [37].

Second, our analysis was limited to tweet texts, replies and hashtags included in a tweet. Analysing the dissemination of tweets containing external URLs or links to news articles can provide further insights on selective content exposure. Mainly, understanding the trends in disseminating news articles from selected or varied sources and shared by friends within the network can provide deeper insight into users' propensity to diversify themselves or indulge in the process of 'self-brainwashing'.

Third, our final dataset was limited to 9,072 tweets across two weeks, which is not ideal for analysing and comparing topical interests. Future studies could perform topic-modelling on a larger corpus spanning over a longer duration. A broader dataset on politically sensitive issues can help discover more latent topics and the evolution of topics through temporal analysis.

6 Conclusion

This paper presents a homophily analysis using a Twitter dataset that includes a highly divisive public discourse in India. Our results indicate that the users' political dispositions predominantly dictate with whom users connect or interact with. Additionally, users also mutually retweet other users who resonate with their emotional states. However, user mentions provide scope for cross-ideological interactions, although the nature of such interactions require further investigation. With the use of common hashtags, the effect of homophily increases within follow and retweet relationships. Nevertheless, the usage of common hashtags with user mentions exhibit mixed outcomes. Pro-CAA users showed a preference to interact with other pro-CAA users. In contrast, anti-CAA users show the opposite behaviour. Hence, common usage of hashtags provides an opportunity for people with diverse points of views and emotional attitudes to interact with each other. Finally, we also show that users with congruent topical interests were less likely to connect or interact with each other, suggesting that topical similarity can bridge the users with different political dispositions or sentimental attitudes. Finally, we discuss how our findings can inform future research and implementations that aim to foster interactions among social media users with divergent viewpoints.

References

1. Aiello, L.M., Barrat, A., Schifanella, R., Cattuto, C., Markines, B., Menczer, F.: Friendship prediction and homophily in social media. In: ACM Transactions on the Web (2012). https://doi.org/10.1145/2180861.2180866
2. Ausserhofer, J., Maireder, A.: National Politics on Twitter. Information, Communication & Society **16**(3), 291–314 (4 2013). https://doi.org/10.1080/1369118X.2012.756050
3. Baek, Y.M., Wojcieszak, M., Delli Carpini, M.X.: Online versus face-to-face deliberation: Who? Why? What? With what effects? New Media and Society (2012). https://doi.org/10.1177/1461444811413191
4. Barbera, P.: Birds of the Same Feather Tweet Together. Bayesian Ideal Point Estimation Using Twitter Data. SSRN Electronic Journal (2013). https://doi.org/10.2139/ssrn.2108098
5. Barberá, P.: How Social Media Reduces Mass Political Polarization . Evidence from Germany, Spain, and the U.S. LXXIII Congress of the Midwest Political Science Association (2014)
6. Bekafigo, M.A., McBride, A.: Who Tweets About Politics? Social Science Computer Review **31**(5), 625–643 (10 2013). https://doi.org/10.1177/0894439313490405
7. Bisgin, H., Agarwal, N., Xu, X.: A study of homophily on social media. World Wide Web (2012). https://doi.org/10.1007/s11280-011-0143-3
8. Bishop, B., Cushing, R.G.: The Big Sort: Why the Clustering of Like-minded America is Tearing Us Apart. Mariner Books (2009)
9. Blei, D.M.: Probabilistic topic models. In: Communications of the ACM (2012). https://doi.org/10.1145/2133806.2133826
10. Blevins, J.L., Lee, J.J., McCabe, E.E., Edgerton, E.: Tweeting for social justice in #Ferguson: Affective discourse in Twitter hashtags. New Media and Society (2019). https://doi.org/10.1177/1461444819827030
11. Bonilla, Y., Rosa, J.: #Ferguson: Digital protest, hashtag ethnography, and the racial politics of social media in the United States. American Ethnologist (2015). https://doi.org/10.1111/amet.12112
12. Boxell, L., Gentzkow, M., Shapiro, J.: Is the Internet Causing Political Polarization? Evidence from Demographics. National Bureau of Economic Research (2017). https://doi.org/10.3386/w23258
13. Bruns, A., Burgess, J.: The use of twitter hashtags in the formation of ad hoc publics. In: European Consortium for Political Research conference, Reykjavík, 25-27 Aug. 2011 (2011)
14. Bruns, A., Moon, B., Paul, A., Münch, F.: Towards a typology of hashtag publics: a large-scale comparative study of user engagement across trending topics. Communication Research and Practice (2016). https://doi.org/10.1080/22041451.2016.1155328
15. Caetano, J.A., Lima, H.S., Santos, M.F., Marques-Neto, H.T.: Using sentiment analysis to define twitter political users' classes and their homophily during the

2016 American presidential election. Journal of Internet Services and Applications (2018). https://doi.org/10.1186/s13174-018-0089-0
16. Chen, J., Liu, Y., Zou, M.: User emotion for modeling retweeting behaviors. Neural Networks (2017). https://doi.org/10.1016/j.neunet.2017.08.006
17. CNN: Citizenship Amendment Bill explained: India's controversial bill that excludes Muslims, https://edition.cnn.com/2019/12/11/asia/india-citizenship-amendment-bill-intl-hnk/index.html
18. Coleman, J.: Relational Analysis: The Study of Social Organizations with Survey Methods. Human Organization (1958). https://doi.org/10.17730/humo.17.4.q5604m676260q8n7
19. Colleoni, E., Rozza, A., Arvidsson, A.: Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data. Journal of Communication (2014). https://doi.org/10.1111/jcom.12084
20. Conover, M., Ratkiewicz, J., Francisco, M.: Political polarization on twitter. Icwsm (2011). https://doi.org/10.1021/ja202932e
21. Currarini, S., Jackson, M.O., Pin, P.: An Economic Model of Friendship: Homophily, Minorities, and Segregation. Econometrica **77**(4), 1003–1045 (2009). https://doi.org/10.3982/ECTA7528
22. De Choudhury, M., Sundaram, H., John, A., Seligmann, D.D., Kelliher, A.: "Birds of a Feather": Does User Homophily Impact Information Diffusion in Social Media? (6 2010), http://arxiv.org/abs/1006.1702
23. De Salve, A., Guidi, B., Ricci, L., Mori, P.: Discovering Homophily in Online Social Networks. Mobile Networks and Applications **23**(6), 1715–1726 (12 2018). https://doi.org/10.1007/s11036-018-1067-2
24. Endres, D.M., Schindelin, J.E.: A new metric for probability distributions (2003). https://doi.org/10.1109/TIT.2003.813506
25. Enli, G.S., Skogerbø, E.: Personalized Campaigns in Party-centred Politics. Information, Communication & Society **16**(5), 757–774 (6 2013). https://doi.org/10.1080/1369118X.2013.782330
26. Fiore, A.T., Donath, J.S.: Homophily in online dating: When do you like someone like yourself? In: Conference on Human Factors in Computing Systems - Proceedings (2005). https://doi.org/10.1145/1056808.1056919
27. Fraisier, O., Cabanac, G., Pitarch, Y., Besançon, R., Boughanem, M.: Uncovering Like-minded Political Communities on Twitter. In: ICTIR 2017 - Proceedings of the 2017 ACM SIGIR International Conference on the Theory of Information Retrieval (2017). https://doi.org/10.1145/3121050.3121091
28. Gerbaudo, P.: Social media and populism: an elective affinity? Media, Culture & Society **40**(5), 745–753 (2018). https://doi.org/10.1177/0163443718772192
29. Gerber, E.R., Henry, A.D., Lubell, M.: Political homophily and collaboration in regional planning networks. American Journal of Political Science (2013). https://doi.org/10.1111/ajps.12011
30. Gimpel, J.G., Hui, I.S.: Seeking politically compatible neighbors? The role of neighborhood partisan composition in residential sorting. Political Geography (2015). https://doi.org/10.1016/j.polgeo.2014.11.003
31. Gonçalves, B., Perra, N., Vespignani, A.: Modeling users' activity on twitter networks: Validation of Dunbar's number. PLoS ONE (2011). https://doi.org/10.1371/journal.pone.0022656
32. Goncalves, J., Kostakos, V., Venkatanathan, J.: Narrowcasting in social media: Effects and perceptions. In: IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2013 (2013). https://doi.org/10.1145/2492517.2492570

33. Goncalves, J., Liu, Y., Xiao, B., Chaudhry, S., Hosio, S., Kostakos, V.: Increasing the reach of government social media: A case study in modeling government-citizen interaction on Facebook. Policy and Internet (2015). https://doi.org/10.1002/poi3.81
34. Graells-Garrido, E., Lalmas, M., Quercia, D.: Data portraits: Connecting people of opposing views (2013), https://arxiv.org/abs/1311.4658
35. Gruzd, A., Wellman, B., Takhteyev, Y.: Imagining twitter as an imagined community. American Behavioral Scientist (2011). https://doi.org/10.1177/0002764211409378
36. Guerra, P.C., Souza, R.C., Assunção, R.M., Meira, W.: Antagonism also flows through retweets: The impact of out-of-context quotes in opinion polarization analysis. In: Proceedings of the 11th International Conference on Web and Social Media, ICWSM 2017 (2017)
37. Hettiachchi, D., Goncalves, J.: Towards Effective Crowd-Powered Online Content Moderation. In: Proceedings of the 31st Australian Conference on Human-Computer-Interaction. pp. 342–346. ACM (12 2019). https://doi.org/10.1145/3369457.3369491
38. Himelboim, I., Cameron, K., Sweetser, K.D., Danelo, M., West, K.: Valence-based homophily on Twitter: Network Analysis of Emotions and Political Talk in the 2012 Presidential Election. New Media and Society (2016). https://doi.org/10.1177/1461444814555096
39. Himelboim, I., Mccreery, S., Smith, M.: Birds of a Feather Tweet Together: Integrating Network and Content Analyses to Examine Cross-Ideology Exposure on Twitter. Journal of Computer-Mediated Communication (2013). https://doi.org/10.1111/jcc4.12001
40. Himelboim, I., Smith, M., Shneiderman, B.: Tweeting Apart: Applying Network Analysis to Detect Selective Exposure Clusters in Twitter. Communication Methods and Measures (2013). https://doi.org/10.1080/19312458.2013.813922
41. Hindustan Times: #IndiaDoesNotSupportCAA takes Twitter by storm, https://www.hindustantimes.com/india-news/indiadoesnotsupportcaa-takes-twitter-by-storm/story-SwRmAoj4tEh2DY9OUK0mBJ.html
42. Huber, G.A., Malhotra, N.: Political homophily in social relationships: Evidence from online dating behavior. Journal of Politics (2017). https://doi.org/10.1086/687533
43. Huckfeldt, R.R., Sprague, J.: Citizens, Politics and Social Communication (1995). https://doi.org/10.1017/cbo9780511664113
44. Hui, I.: Who is Your Preferred Neighbor? Partisan Residential Preferences and Neighborhood Satisfaction. American Politics Research (2013). https://doi.org/10.1177/1532673X13482573
45. Hutto, C.J., Gilbert, E.: VADER: A parsimonious rule-based model for sentiment analysis of social media text. In: Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014 (2014)
46. Ince, J., Rojas, F., Davis, C.A.: The social media response to Black Lives Matter: how Twitter users interact with Black Lives Matter through hashtag use. Ethnic and Racial Studies (2017). https://doi.org/10.1080/01419870.2017.1334931
47. India Today: Everything you wanted to know about the CAA and NRC, https://www.indiatoday.in/india-today-insight/story/everything-you-wanted-to-know-about-the-caa-and-nrc-1630771-2019-12-23
48. Iyengar, S., Hahn, K.S.: Red media, blue media: Evidence of ideological selectivity in media use. Journal of Communication (2009). https://doi.org/10.1111/j.1460-2466.2008.01402.x

49. Iyengar, S., Sood, G., Lelkes, Y.: Affect, not ideology: A social identity perspective on polarization (2012). https://doi.org/10.1093/poq/nfs038
50. Iyengar, S., Westwood, S.J.: Fear and Loathing across Party Lines: New Evidence on Group Polarization. American Journal of Political Science (2015). https://doi.org/10.1111/ajps.12152
51. Jackson, S.J., Foucault Welles, B.: Hijacking #myNYPD: Social Media Dissent and Networked Counterpublics. Journal of Communication (2015). https://doi.org/10.1111/jcom.12185
52. Kang, J.H., Lerman, K.: Using lists to measure homophily on twitter. In: AAAI Workshop - Technical Report (2012)
53. Lewicka, M.: Confirmation Bias. In: Personal Control in Action, pp. 233–258. Springer US, Boston, MA (1998). https://doi.org/10.1007/978-1-4757-2901-6_9
54. Liu, Y., Venkatanathan, J., Goncalves, J., Karapanos, E., Kostakos, V.: Modeling what friendship patterns on facebook reveal about personality and social capital. ACM Trans. Comput.-Hum. Interact. **21**(3) (2014). https://doi.org/10.1145/2617572
55. Lovejoy, K., Waters, R.D., Saxton, G.D.: Engaging stakeholders through Twitter: How nonprofit organizations are getting more out of 140 characters or less. Public Relations Review (2012). https://doi.org/10.1016/j.pubrev.2012.01.005
56. Madden, M., Smith, A.: Reputation management and social media (2010)
57. Manikonda, L., Beigi, G., Liu, H., Kambhampati, S.: Twitter for sparking a movement, reddit for sharing the moment: #Metoo through the lens of social media (2018)
58. Mason, L.: Uncivil Agreement: How Politics Became Our Identity. University of Chicago Press (2018)
59. McPherson, M., Smith-Lovin, L., Cook, J.M.: Birds of a feather: Homophily in social networks. Annual Review of Sociology (2001). https://doi.org/10.1146/annurev.soc.27.1.415
60. Mehrotra, R., Sanner, S., Buntine, W., Xie, L.: Improving LDA topic models for microblogs via tweet pooling and automatic labeling. In: Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (2013). https://doi.org/10.1145/2484028.2484166
61. Mummolo, J., Nall, C.: Why partisans do not sort: The constraints on political segregation. Journal of Politics (2017). https://doi.org/10.1086/687569
62. Nagulendra, S., Vassileva, J.: Understanding and controlling the filter bubble through interactive visualization: A user study. In: HT 2014 - Proceedings of the 25th ACM Conference on Hypertext and Social Media (2014). https://doi.org/10.1145/2631775.2631811
63. NPR: India Passes Controversial Citizenship Bill That Would Exclude Muslims, https://www.npr.org/2019/12/11/787220640/india-passes-controversial-citizenship-bill-that-would-exclude-muslims
64. Nyhan, B., Reifler, J.: When corrections fail: The persistence of political misperceptions. Political Behavior (2010). https://doi.org/10.1007/s11109-010-9112-2
65. Pariser, E.: Filter Bubble (2012). https://doi.org/10.3139/9783446431164
66. Park, C.S.: Does Twitter motivate involvement in politics? Tweeting, opinion leadership, and political engagement. Computers in Human Behavior (2013). https://doi.org/10.1016/j.chb.2013.01.044
67. Pew Research Center: The Partisan Divide on Political Values Grows Even Wider. Tech. rep. (2017)

68. Rader, E., Gray, R.: Understanding user beliefs about algorithmic curation in the facebook news feed. In: Conference on Human Factors in Computing Systems - Proceedings (2015). https://doi.org/10.1145/2702123.2702174
69. Ranganath, S., Hu, X., Tang, J., Liu, H.: Understanding and identifying advocates for political campaigns on social media. In: WSDM 2016 - Proceedings of the 9th ACM International Conference on Web Search and Data Mining (2016). https://doi.org/10.1145/2835776.2835807
70. Small, T.A.: What the Hashtag? Information, Communication & Society **14**(6), 872–895 (9 2011). https://doi.org/10.1080/1369118X.2011.554572
71. Stieglitz, S., Dang-Xuan, L.: Emotions and information diffusion in social media - Sentiment of microblogs and sharing behavior. Journal of Management Information Systems (2013). https://doi.org/10.2753/MIS0742-1222290408
72. Sunstein, C.R.: Republic.com 2.0 (2009). https://doi.org/10.5860/choice.45-5264
73. Tajfel, H., Turner, J.: An Integrative Theory of Inter-group Conflict. In: The social psychology of intergroup relations. Oxford University Press (1979)
74. The Financial Express: 'India Supports CAA' : PM Modi launches Twitter campaign to support Citizenship Act, https://www.financialexpress.com/india-news/india-supports-caa-pm-modi-launches-twitter-campaign-to-support-citizenship-act/1807380/
75. Tsugawa, S., Ohsaki, H.: On the relation between message sentiment and its virality on social media. Social Network Analysis and Mining (2017). https://doi.org/10.1007/s13278-017-0439-0
76. Venkatanathan, J., Karapanos, E., Kostakos, V., Gonçalves, J.: Network, personality and social capital. In: ACM Web Science Conference. p. 326–329. WebSci '12, ACM (2012). https://doi.org/10.1145/2380718.2380760
77. Venkatanathan, J., Karapanos, E., Kostakos, V., Gonçalves, J.: A network science approach to modelling and predicting empathy. In: IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. p. 1395–1400. ASONAM '13, ACM (2013). https://doi.org/10.1145/2492517.2500295
78. Wang, F., Orton, K., Wagenseller, P., Xu, K.: Towards Understanding Community Interests with Topic Modeling. IEEE Access (2018). https://doi.org/10.1109/ACCESS.2018.2815904
79. Williams, H.T., McMurray, J.R., Kurz, T., Hugo Lambert, F.: Network analysis reveals open forums and echo chambers in social media discussions of climate change. Global Environmental Change (2015). https://doi.org/10.1016/j.gloenvcha.2015.03.006
80. Xiong, Y., Cho, M., Boatwright, B.: Hashtag activism and message frames among social movement organizations: Semantic network analysis and thematic analysis of Twitter during the #MeToo movement. Public Relations Review (2019). https://doi.org/10.1016/j.pubrev.2018.10.014
81. Xu, S., Zhou, A.: Hashtag homophily in twitter network: Examining a controversial cause-related marketing campaign. Computers in Human Behavior (2020). https://doi.org/10.1016/j.chb.2019.08.006
82. Xu, W.W., Sang, Y., Blasiola, S., Park, H.W.: Predicting Opinion Leaders in Twitter Activism Networks: The Case of the Wisconsin Recall Election. American Behavioral Scientist (2014). https://doi.org/10.1177/0002764214527091
83. Yang, G.: Narrative agency in hashtag activism: The case of #blacklivesmatter (2016). https://doi.org/10.17645/mac.v4i4.692
84. Yu, B., Kaufmann, S., Diermeier, D.: Exploring the characteristics of opinion expressions for political opinion classification. Proceedings of the 2008 international conference on Digital government research (2008)