




Exploring the disparity of influence between users in the discussion of Brexit on Twitter

Twitter influence disparity in Brexit If so, write it here

Amirarsalan Rajabi¹ · Alexander V. Mantzaris² · Kuldip Singh Atwal³ · Ivan Garibay^{1,3} 

Received: 22 June 2020 / Accepted: 10 March 2021 / Published online: 30 March 2021
© The Author(s) 2021

Abstract

The topic of political polarization has received increased attention for valid reasons. Given that an increased amount of the social exchange for opinions happens online, social media platforms provide a good source of information to investigate various aspects of the phenomena. In this work, data collected from Twitter are used to examine polarization surrounding the topic of the Brexit referendum on the membership of the European Union. The analysis specifically focuses on the question of how different tiers of users in terms of influence can project their opinions and if the polarized conditions affect the relative balance in the broadcast capabilities of the tiers. The results show that during polarization periods, users of the higher tier have increased capabilities to broadcast their information in relation to the lower tiers thereby further dominating the discussion. This validates previous modeling investigations and the hypothesis that polarization provides an opportunity for influencers to increase their relative social capital.

Keywords Polarization · Social media · Brexit · Influence · Analytics

Mathematics Subject Classification 62-07 · 91F99 · 91D99

✉ Ivan Garibay
igaribay@ucf.edu

¹ Department of Computer Science, University of Central Florida (UCF), Orlando, FL, USA

² Department of Statistics and Data Science, University of Central Florida (UCF), Orlando, FL, USA

³ Department of Industrial Engineering and Management Systems, University of Central Florida (UCF), Orlando, FL, USA

Introduction

Societal dialogue has been a corner stone for democracies and its use in building bridges between groups in education is long standing [36]. It is a tool for exchanging ideas and propagating important information at an individual's level and that of a societal conscious [4] for making important decisions. Voting systems rely upon a populous which is informed and in a state of mind for making "good" and rational decisions [23]. Recently, there has been an increased amount of attention in the academic community investigating the issue of polarization since it can affect the results of the democratic process and jeopardize its long-term stability [8]. This is not a new concern as it has been raised in the works of [2, 13, 18, 21, 38, 39] which spans a period from the 80s to this current year. A key case study for phenomenon of political polarization is the *Brexit* referendum on leaving the European Union (E.U.) [9]. The insightful work of [21] provides a more detailed understanding of the placements of the identities formed out of the stance for those voting on this matter. It reinforces, by reminder, the concept of how a democracy requires mutual respect and a "willingness to talk across political divides". The surveys conducted show that social identities were formed during the EU referendum and that they provided a personally important component which connects to stereotyping and biases beyond preexisting political partisanship. The work of [16] provides a model for how this effect can be produced in a stochastic message exchange model and this work here uses data from Twitter to support the hypothesis that polarization provides an opportunity for *social media influencers* to gain social capital in the process. The results provide evidence for the increased role that influencers have during polarized discussions, and subsequently how the lower tiers of users have diminished broadcasting roles.

With the advent of social media platforms and their growing influence on public discourse, polarization in social media is a new field of interest. The work of [43] reviews the literature on more traditional sociological approaches into investigating the connections between social media and political polarization. With the increasing amount of social exchange happening on the platforms of Twitter, Reddit and Facebook, for instance, online social networks have modified and transformed much of our today's communication [28, 31]. With ascending number of users, there is an increasing attention paid to studying the development of links in such networks [27]. These networks provide an environment in which individuals might isolate themselves from the viewpoints that contradict theirs, and aid the formation of *echo chambers*. The adoption of social media platforms has also led to the rapid increase of social bots [29]. The work of [20] proposes an automated method to detect social media bots. In such an environment, the role of users with high influence and elevated capability to broadcast their message, especially in the context of a polarized discussion, is imperative. The work of [10] has investigated the effect of polarization on retweet and user-to-user mention networks in Twitter, and has concluded that while the former is highly polarized during a polarized discussion, the latter is not, and is also observed in the results shown in Sect. 4.

On June 23rd 2016, a UK-wide referendum called Brexit was held in which 52% of participants voted for the UK to ‘leave’ European Union, in contrast with 48% of participants who voted for ‘remain’. The Brexit referendum is an important case of how online social networks can have a pivotal effect in determining the outcome of a major polarized political discussion [19]. Among online social network platforms, given its multi-directional interactivity and its popularity among politicians and journalists, Twitter is considered one of the most important ones [19, 30]. The conversations of Brexit on Twitter provide an apt case study to investigate polarization in the context of online social networks, given its dichotomous nature of discussion and outcome. Polarization in Twitter has been studied by a wide range of data-driven approaches and on multiple topics; polarization in Venezuela expressed on Twitter [26], religious polarization in Egypt [44], political polarization in Japan [42], and polarization in Twitter regarding Canada’s perspective [17].

In an effort to understand how polarization forms to explore approaches of slowing or reversing the effects, computational modeling can provide ways to experiment with preventative measures after the effects can be reproduced in silico. [7] provides a macroscopic modeling approach to the topic of polarization with interesting results of how it began from a top economic tier to the lower strata over time. The work of [24] and [35] explore polarization and consensus within the context of opinion dynamics models. [32] and [12] provide opinion dynamics models under which competition and polarization results in stable configurations of coexistence. The work of [33] show how intervention to counter the spread of disinformation, might result in an undesirable effect which is further polarizing the social network.

The work of [16] provides a stochastic network model of dynamic communicators. Users have an influence value, which determines their ability to get their message broadcast by other users. The model undergoes non-polarized and polarized phases which provide an identity per user affecting the way cross group messages can then be propagated (retweet effect). The users are then divided into two equally sized tiers in terms of influence: T_{top} and T_{bottom} . These users in T_{top} have higher influence values while users in T_{bottom} have lower influence values. The authors show that introduction of polarization into the model, enables users of high influence to accumulate even higher relative influence and consequently getting more message broadcast. The implications are that the dialogue then becomes dominated by a smaller group of accounts rather than allowing more equally weighted exchanges. The modeling approach allows users to accumulate social capital over time as type of currency which can mirror the effect of the memory of an account by other users over time. This way influencers can find an opportunity to use such a period of polarized discourse to capitalize upon it providing a direct motive for their activity.

Social media influencers are users/accounts with the ability to broadcast their message to a relatively larger set of users. The role and effect of influencers on topics ranging from opinion leaders in political discussions to forming customer attitudes has been investigated [3, 15, 25, 41]. Several methods have been proposed to measure a user’s influence on Twitter, and at the intersection of most of these methods is the centrality of a user [22, 34, 40], which is captured well with the number of followers. Users with higher number of followers are able to broadcast their

message to a bigger audience, which in turn makes their message more prone to getting retweeted.

This study follows a similar approach to [16] in that users are divided into two tiers of influence T_{top} and T_{bottom} . The objective for classifying users and their tweets into two tiers is to investigate the relative capability of members of the two groups to broadcast their message and acquire attention, comparing to the same quantity before the hyper-polarizing event. This seeks to answer this question as to whether, and how, a polarizing event causes disparity among users participating in the discussion of Brexit, within the Twitter network. Reinforcing the modeling approach of the results of [16] with a data-driven investigation, the outcome is that a polarizing event such as an election, results in an increase in relative influence of users of the top tier compared to users of bottom tier. This hypothesis is tested by investigating quantities such as Gini coefficient, retweet ratio and favorite ratio of top and bottom tiers. Also, the work of [1] puts forward the evidence that exposure to opposing views can induce polarization and in this paper we show that it is the event which causes people to see that changes affecting their views are possible which induces the polarization. It can then be speculated that influencers could do this on smaller scale disagreements, e.g. manufactured scandals or to actively participate in arguments with followers of the account to witness.

The next section discusses the data used for the investigation, the subsequent section presents the key metrics used upon the data for the analysis of the topic, then the results and the description of the findings are displayed, and the final section provides a discussion with a brief summary of the outcomes.

Data

The data used in this study arise from Twitter where the subselection is based upon the discussion surrounding Brexit. The reference to the dataset [5] (Harvard DataVerse) provides a collection of tweet ids and user ids on the topic of Brexit which was collected between January 2016 and October 2019. The dataset contains more than 50 million tweet ids, which can be used to retrieve the original tweet objects through the Twitter API which is valuable for inferring specifics regarding the Brexit context. This dataset also provides a list of users whose tweets are present in the dataset along with the number of tweets each user has produced. Each user also contains an attribute field with a *BotScore* that is produced using a bot detector approach named BotOrNot ([11]). This assigns a score in the range (0,1) as the probability that a user is a bot or not. The values closer to 1 are more supportive of the probability that the user is a bot. Each user is also attributed with a stance: “Leave”, “Remain”, or “Neutral” according to the tweets present in the dataset by each user. The stance classification task is performed using stance-indicative (SI) hashtags, and machine learning approaches for the users that lack stance-indicative hashtags in their tweets [6].

From the complete dataset, 802 users were chosen, the stance was cross examined by 2 human curators independently on the topic of Brexit. The 802 users were chosen based on the following criteria: (1) they have been active during the

2016 Brexit referendum and the 2017 UK snap election period. (2) They have a low bot score, and (3) they are diverse in terms of the number of followers, so that there exists users with a high number of followers, as well as users with average and low numbers of followers. The corresponding tweets were retrieved from the Twitter API. More than 90% of accounts had a BotScore of less than 0.3. Of the chosen accounts, 731 users were created before 2016 and the rest were created before 2017. Of the labelled users, 457 were labelled as “remain”, 31 were labelled as “neutral”, and 314 were labelled as “leave”. In the end, all the tweets in the dataset which were created by these labelled users were then retrieved by Twitter API.

The corresponding dataset created includes 485,383 tweet objects over the span of 45 months. The dataset is then analyzed with the metrics discussed in Sect. 3 where users are first sorted based on their followers count number and then divided into two tiers T_{top} and T_{bot} , with each tier containing equal number of unique users. Each day contains a number of active unique users. Figure 1 shows daily number of active users during Brexit referendum 2016 (subfigure a) and UK snap election 2017 periods (subfigure b). There are dashed lines for the exact day of the election. Both periods observe an increase in the number of unique users participating in the discussion right after the events.

As discussed in Sects. 3 and 4, a ‘cross group interaction’ investigation is performed. Since between-group events (members of opposite stance labels interacting) are found to be sparse, this particular investigation shown in Fig. 6 utilizes a larger dataset where more events are present. The dataset used for this investigation differs from the previous one in which users stance labels were manually validated after the automated procedure. This new dataset arises from the same source [5], but the human validation is not included to verify the correctness of the automated approach. It contains 1,902,173 tweets, which are created by 906,608 “remain” users and 352,632 “leave” users. The stance attributed to each user was performed automatically and is present in the dataset, as explained in [6].

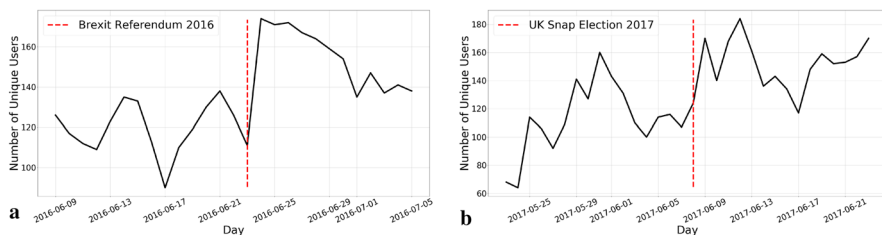


Fig. 1 Daily number of active unique users in each day. **a** The daily number of active unique users during Brexit referendum 2016 period and **b** daily number of active unique users during snap election of 2017 period

Methodology

Several metrics are utilized for analyzing the Twitter dataset described in Sect. 2 with the purpose of uncovering the user influence characterizations also described in [34]. One of the unique features of social media platforms is that they allow users to control a sharing option that can replicate their own content to their friends/followers or that of other users' content, therefore, acting like a proxy for the amplification and spread of influence. This is discussed in more detail in Sect. 1. The measures described here are those used in Sect. 4 to produce the results shown.

The users in the dataset are first sorted based on their followers count number, and are subsequently divided into two tiers, with each containing an equal number of unique users. The top tier T_{top} consists of users with higher numbers of followers, and the bottom tier T_{bottom} is formed by users with lower number of followers. Furthermore, each tweet belongs to one of the classes of $\{T_{top}, T_{bottom}\}$; tweets whose original author belongs to T_{top} and the ones whose original author belongs to T_{bottom} .

Gini coefficient

The Gini coefficient as a measure of statistical dispersion has been applied in economics to capture and measure income inequality [14]. In this work, we consider the number of times a tweet has been retweeted, as an analogous unit for the 'wealth' of that tweet produced by the user. This coefficient is calculated daily for all tweets, with respect to the number of times each tweet has collected retweets. Figure 3 compares the Gini coefficient calculated for tweets considering their retweet number (wealth). Each point G_t in the traces are calculated as in [37]:

$$G_t = \frac{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|}{2n^2\bar{x}}. \quad (1)$$

Here t is the time point for which the coefficient is calculated, e.g. day in the timeline. x_i is the number of retweets that tweet i has received and equivalently x_j is the retweets for another tweet so that all the pairs are explored. \bar{x} is the average number of retweets (per tweet) and n is the number of tweets, in that specific day. A lower number is indicative of a more uniform distribution in comparison to a larger number representative of some disproportionate number of retweets.

Retweet ratio

Here is presented the measure used for measuring the disparity between the number of retweets received by the users in different tiers in the dataset. In Fig. 4, its use is presented with both a daily retweet ratio (subfigure a and b) and a weekly one (subfigure c and d). At each day the total number of retweets that the tweets of users of T_{top} have acquired is recorded. Then the total number of retweets that the tweets

of users of T_{bottom} have acquired is also recorded. These two quantities are divided by each other, and the retweet ratio corresponding to that specific day is calculated. This process is performed for all days and recorded in S_t .

$$S_t = \frac{\sum_{i \in T_{top}} R_i}{\sum_{i \in T_{bottom}} R_i} \quad (2)$$

and t is the time point for which the retweet ratio is calculated e.g. day in the timeline. R_i is the number of retweets that user i has received. In the end, a 3-day centered moving average is then taken. Note that in subfigures c and d of Fig. 4, the unit of time considered is a week, so each point shows the ratio of number of retweets that users of T_{top} received, divided by that of users of T_{bottom} , for each specific week.

Favorite ratio

This measure captures the disparity between the number of favorites that users of T_{top} receive and the number of favorites that users of T_{bottom} receive. For each day, the total number of favorites that the tweets of users of T_{top} have received is recorded. Then the total number of retweets that the tweets of users of T_{bottom} have received is also recorded. The favorite ratio is then obtained by dividing the first quantity by the second quantity:

$$S_t = \frac{\sum_{i \in T_{top}} F_i}{\sum_{i \in T_{bottom}} F_i} \quad (3)$$

Here t is the time point (day) in the timeline for which the favorite ratio is calculated, and F_i is the number of retweets that user i has received. In the end, a 3-day centered moving average is then taken.

Cross-group interaction

This part explains the measure used for creating Fig. 6. An interaction is defined as an event in which a user does one of the following: (1) quote, (2) mention, or (3) reply. In the Twitter network, these events always take place between at least two users. Also, as explained in Sect. 2, each user is considered to belong to one of the following sides in Brexit discussion: (1) *remain*, (2) *leave*, (3) *neutral* (where the labels in the dataset are provided). The purpose of this measure is to monitor over time the within and between group activities as ‘interactions’.

An interaction which has taken place between users that share the same stance on Brexit, are all *remain* users or all are *leave* users in those actions, is classified as an inter-group interaction I_{inter} . The second type of interaction is called cross-group interaction and is one in which the participating users belong to both sides of Brexit discussion, some belong to the *remain* side and some belong to *leave* side.

Figure 6 shows what percentage of all interactions between users of *remain* and *leave* in each day, is of the type cross-group interaction I_{cross} . For each day, the total number of cross-group interactions I_{cross} is divided by total interactions (I_{inter} and I_{cross}).

$$S_t = \frac{\sum I_{cross}}{\sum I_{inter} + \sum I_{cross}}. \quad (4)$$

Here t is the time point (day) in the timeline for which the ratio is calculated.

Results

This section presents the results of applying the metrics and measures described in Sect. 3 to the datasets described in Sect. 2. The overarching goal is to explore whether a key political event can induce a change in the distribution of social media exchanges between active users. Particularly to investigate if users of higher rank can benefit from times of polarized discussion in terms of being able to spread content with greater relative value than their peers and having their content favorited with larger relative values. The data used are taken from Twitter for exchanges relevant to the Brexit vote events; the UK referendum of 2016 and the 2017 UK snap election. It was hypothesized in previous work that the events can offer an opportunity for influencers to regain relative broadcast rankings which are diluted during non-polarized exchanges. Here the findings show increased influence during and surrounding these events. Key findings are that more users participate in the discussion during key events (as hypothesized from modeling considerations), and as well during these events that there is a larger disparity between the upper tiers of influencers and lower tiers during critical political events. The increase in activity does not maintain the same distribution of as before, but that the influencers gain not just in the quantity but also the relative impact. This alludes to a possible anticipation and capitalization of the process through dialogue choices. The results show that the capitalization does not involve influencers producing more content but receiving activity on their content by other users and those in a lower tier (increasing the relative social capitol).

Activity ratio between tiers of users in the Twitter Brexit discussions

Figure 2 looks at the amount of relative activity of the 2 tiers of the users participating in the Twitter Brexit discussion in the timelines of the 2016 Brexit referendum and the 2017 UK snap election (subfigure a and b), respectively, with the dashed lines for the voting day events). The timelines show that there is a decrease in the ratios around the event dates. This means that there is an increase of activity from the bottom tier in relation to the group influencers and can allude to the activity being uniform, but as shown in Fig. 4 it can be seen how the top tiers receive more retweets during these events. As a result this shows that the lower tier activity is

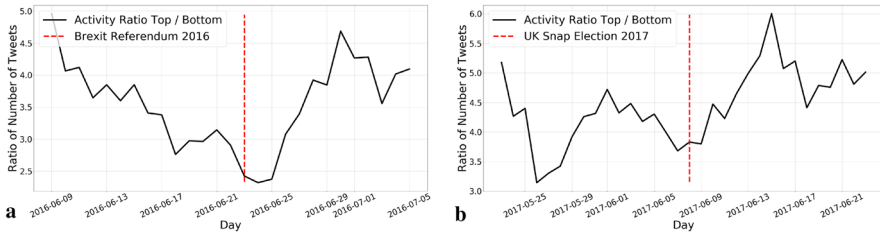


Fig. 2 Looking at the relative activity of users of the top and bottom tiers. Here the ratio of the number of tweets that users of the top tier create divided by that of the users of bottom tier is shown, for each day, measured independent of the previous days. **a** The dashed line is the Brexit referendum date June 2016, and **b** the UK snap election of June 2017. It is showing that the relative activity of the users of two tiers decreases during the events, meaning that the subsequent plots and results of this work are not due to higher activity of users of top tier

there to support and propagate the influencer content with more activity than without the polarizing discussion as discussed in [16].

Examining the retweet disparity between users over time

The Gini coefficient is measured along the dataset timeline, Sect. 3.1, to investigate the disparity between tiers of users with respect to the number of retweets in each group. Retweets are an indication of content popularity of a user and the Gini coefficient provides a measure to see the extent for which the upper level of users, in respect to retweet counts, is increasing or decreasing over time. The timeline is taken surrounding the dates of two important political decisions for the UK; the Brexit referendum vote of 2016 and the UK snap election of 2017.

Figure 3 shows the results of the Gini coefficient exploration for the number of retweets that each tweet gains, taken each day, and independent of previous days. A 3-day centered moving average is taken. **a**) The timeline surrounding the Brexit referendum and **b**) the timeline around the snap election, where a dashed line indicated the day in which the event takes place. It can be seen how in each of the plots there

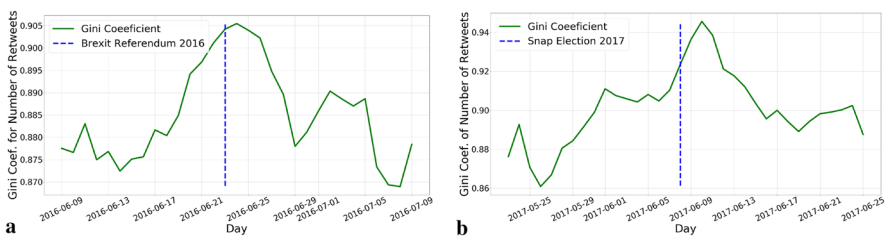


Fig. 3 Daily Gini coefficient values for retweet number per tweet. Here we consider a timeline for the Brexit discussion on Twitter and calculate the Gini coefficient for the population of tweets, binned each day and taken independently of previous days. The dashed line in **a** is the day of the Brexit referendum election and in **b** the day of the UK 2017 snap election. Increased values show that there was more of a bias for a smaller set of tweets to be retweeted

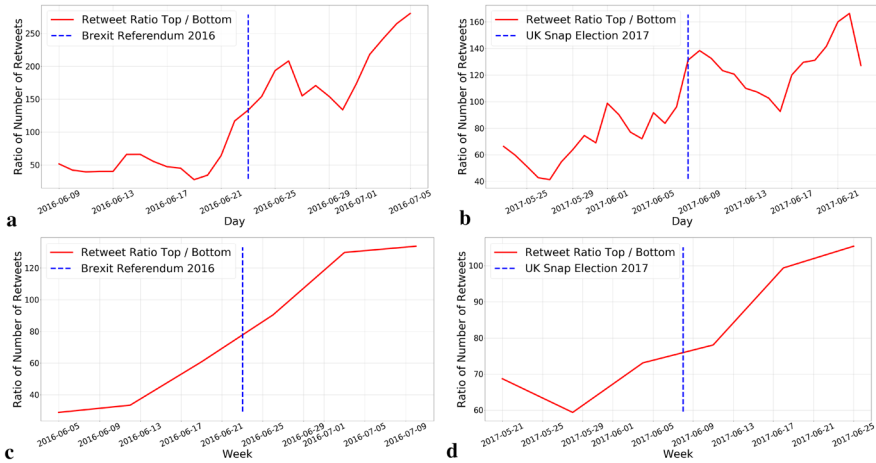


Fig. 4 Looking at the disparity between the top and bottom tiers for the number of retweets. The daily trend is plotted for the users' retweet count ratio in the 2 tiers where the tiers are set to have the same number of unique users. **a** The dashed line is the Brexit referendum date June 2016, and **b** the UK snap election of June 2017. These events start an increasing trend in the ratio of the number of retweets that users of the top tier receive divided by the bottom tier

is a general increase of the Gini coefficient value around the dashed lines. Here the population of tweets on the daily level are used in calculating the Gini coefficient and that the lowest values correspond to a distribution that is closer to a uniform distribution and the larger values to when the distribution is skewed. It can be seen that the distribution becomes less uniform during the public discussion around the key event time points. As a result, this means that a smaller subset of the active participants have effectively utilized the sharing platform's features to express themselves and that the social influence becomes more dominated by a smaller group.

Exploring the trace of the disparity in retweet ratio

The disparity in the retweet ratio used here is described in Sect. 3.2. Users active in the timeline are placed into two tiers, T_{Top} and T_{Bottom} with equal number of unique users placed in these categories. Examining this ratio over time produces further understanding for the disparity between the upper level of the influencer rank and the lower tier. To see whether there is an even set of opportunities for expression about the times when key political decisions are being made, this retweet ratio helps investigate whether this is being deviated from which would ideally have a uniform distribution representing a democratic exchange.

Figure 4 displays the results of examining this ratio upon 2 timelines; that of the Brexit 2016 referendum day and the 2017 snap election with those days indicated by a dashed line. The ratio of the number of retweets that users of T_{Top} receive divided by the number of retweets that users in T_{Bottom} receive, for each day. (a) and (b) correspond to using a 3-day centered moving average for the Brexit and snap election

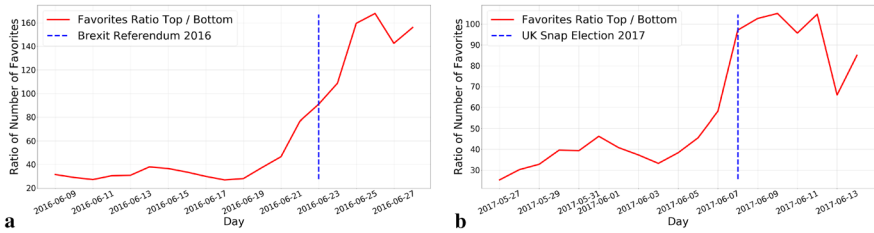


Fig. 5 The ratio of the favorites obtained by users separated into 2 tiers per day surrounding the Brexit discussion. Each day the tweets are designated as belonging to 2 tiers of users with equal numbers based upon the number of favorites obtained from other users (favorites, aka ‘likes’). **a** Time around the Brexit referendum with that date on a dashed line and **b** the snap election in 2017 on the dashed line

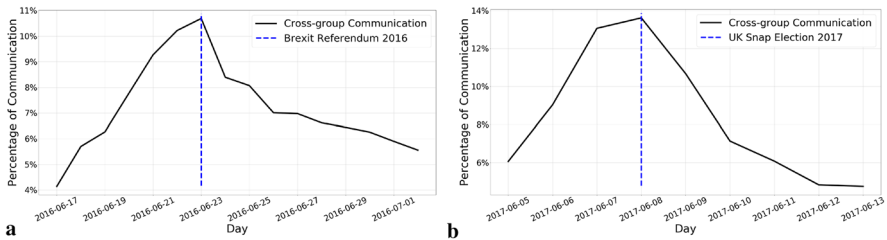


Fig. 6 Examining the trends of pro and anti Brexit group interactions surrounding key election time points. Each user is labelled as being pro or anti Brexit and then cross communication events are labelled when there is a ‘mention’, ‘reply’ or ‘quote’ is produced between members of different labels. The percentage of those cross community interactions over the total number of interactions per day is produced. **a** A timeline surrounding the Brexit referendum day with a dashed line for the event and **b** for the 2017 UK snap election with the dashed line designating that day

vote respectively. (c) and (d) are analogous investigations taken at centered, 3-week rolling averages instead of it being daily. What can be seen in these timelines is that there is a general increase in the disproportionate retweet ratios between tiers and that these trends are related to Brexit events. The ability for a subset of users to have increased message broadcast capabilities during these discussions not only means that it is possible that fewer users are able to influence the public discussion and possibly the outcome, it also means that these users irrespective of the outcome can increase their social relative influence. This was a key assumption in the work [16] where the model definition supported that a division in social cohesion would provide an opportunity for high end influencers to further their relative ranking amongst others.

Disparity in favorite ratio between tiers

In the Twitter platform, the ‘likes’ a user can attribute to tweets and content shared is recorded as a ‘favorite’ in the data retrieved from Twitter. This feature is also important to see the distribution of the attention and focus of the community surrounding the discussion. In comparison to the results shown in Sect. 4.3 which are based upon the

‘retweets’, the favorites do not indicate an ability to propagate or amplify the content users produce. Users with larger retweet counts can influence more people and accumulate more social capital than those which do not manage to have other users support them through sharing. The favorites, although do not provide an increase in reach or amplification of content through proxy accounts, it does provide at least an indication of interest in the community.

Figure 5, shows the results of the disparity ratio between top tier and bottom tier for the favorites as described in Sect. 3.3 with a 3-day centered moving average. The 2 subfigures a and b show the timelines for the Brexit referendum of 2016 and the UK snap election in 2017, respectively (dashed lines show the specific dates). What can be seen is that there are relatively stable equal accumulations of ‘favorites’ between users in the 2 tiers until the events take place. The trends are relatively stable until the timeline approaches the events. It can be seen here more clearly than in Fig. 4, that the events instigate a change in the ratio as the changes here are less gradual. This supports the claims and the previous results shown that the events induce an overall skewed distribution of the attention and focus onto a subset of the network even if the users are not restricted from applying their focus more evenly.

Cross-group interaction

Section 3.4 of the methodology defines what is meant here by a ‘cross-group interaction’. The ‘interaction’ is when a user i quotes/mentions/replies with another user set j in the dataset. As described in the Sect. 2, the users are labeled as belonging to one of three sides in the discussion, pro-Brexit/anti-Brexit/neutral (leave/remainer/neutral in the Brexit parlance). When users between, pro and anti, produce an ‘interaction’ with users of another side (category label) it is classified as a ‘cross-group interaction’. Over the timeline the percentage of the interaction events which classify as cross-group interaction is plotted (compared to all interaction events). The stance attributed to each user in this dataset is performed by [6] where these labels are already present.

Figure 6 displays the results for the question of what percentage of total interactions among users takes place between users of opposite sides of the Brexit discussion, *remainers* and *leavers* referred to here as cross-group interactions (a 3-day centered moving average is taken). Subfigure a shows the timeline for the 2016 Brexit referendum and subfigure b for the 2017 snap election (dashed lines show the actual events days). The figures support the previous results that the elections change the trends, and also provides evidence that more users engage with content from different sides of the discourse. This also supports the concept speculated and introduced as a modeling paradigm in [16], that polarizing events will produce interactions that will result in increased support for influencers’ content rather than produce their own.

Discussion

The work presented here investigated the changes in the ability for different tiers of users in Twitter to express themselves during the exchanges regarding the Brexit topic. There is previous work which discusses the modeling dynamics for how polarization can facilitate an increase in social capital for users in the upper tier that act as ‘influencers’ and the data-driven analysis here supports those theoretical findings. A key aspect is that during periods of polarization influencers’ content receive a disproportionate increase in attention from the community. The modeling assumptions of the previous work and the data traces from the Tweets do show that although users of the lower tier increase their participation their content was retweeted in a lower proportion outside of a voting event.

This is important for the reason that a democratic process would ideally have this proportion become more uniform for key events rather than influencers providing anchors for the discussion. The other factor is that influencers through experience, or from insight, can foresee these dynamics and capitalize on them if not actually participate in invigorating. If there is an avenue to increase social capital on the platform and subsequently financial capital, this provides a motive for them to further capitalize from the process. This makes the task of providing depolarizing incentives more challenging.

Future work could look into the prevalence of this phenomenon on other platforms as well. As well, the investigations could explore the question as to what degree do the followers of the most influentials become polarized due to their discussion leaders.

Acknowledgements This research was supported by DARPA program HR001117S0018. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., et al. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, *115*(37), 9216–9221.
2. Baldassarri, D., & Gelman, A. (2008). Partisans without constraint: Political polarization and trends in American public opinion. *American Journal of Sociology*, *114*(2), 408–446.
3. Booth, N., & Matic, J. A. (2011). Mapping and leveraging influencers in social media to shape corporate brand perceptions. *Corporate Communications: An International Journal*, *16*(3), 184–191.
4. Burton, M., & Kagan, C. (1994). The verbal community and the societal construction of consciousness. *Behavior and Social Issues*, *4*(1–2), 87–96.
5. Calisir, E., & Brambilla, M. (2020). Twitter dataset about brexit.
6. Calisir, E., & Brambilla, M. (2020). Wide-spectrum characterization of long-running political phenomena on social media: the brexit case. In *Proceedings of the 35th annual ACM symposium on applied computing* (pp. 1869–1876).
7. Callander, S., & Carbajal, J. C. (2020). Cause and effect in political polarization: A dynamic analysis. Peruvian Economic Association.
8. Carothers, T., & O’Donohue, A. (2019). *Democracies divided: The global challenge of political polarization*. Washington: Brookings Institution Press.
9. Clarke, H. D., Goodwin, M. J., Goodwin, M., & Whiteley, P. (2017). *Brexit*. Cambridge: Cambridge University Press.
10. Conover, M. D., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on twitter. In *Fifth international AAAI conference on weblogs and social media*.
11. Davis, C. A., Varol, O., Ferrara, E., Flammini, A., & Menczer, F. (2016). Botornot: A system to evaluate social bots. In *Proceedings of the 25th international conference companion on world wide web* (pp. 273–274).
12. Del Vicario, M., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrocioni, W. (2017). Modeling confirmation bias and polarization. *Scientific Reports*, *7*, 40391.
13. DiMaggio, P., Evans, J., & Bryson, B. (1996). Have American’s social attitudes become more polarized? *American Journal of Sociology*, *102*(3), 690–755.
14. Dorfman, R. (1979). A formula for the Gini coefficient. *The review of economics and statistics* (pp. 146–149).
15. Freberg, K., Graham, K., McGaughey, K., & Freberg, L. A. (2011). Who are the social media influencers? A study of public perceptions of personality. *Public Relations Review*, *37*(1), 90–92.
16. Garibay, I., Mantzaris, A. V., Rajabi, A., & Taylor, C. E. (2019). Polarization in social media assists influencers to become more influential: Analysis and two inoculation strategies. *Scientific Reports*, *9*(1), 1–9.
17. Gruz, A., & Roy, J. (2014). Investigating political polarization on twitter: A Canadian perspective. *Policy & Internet*, *6*(1), 28–45.
18. Hall, R. G., Varca, P. E., & Fisher, T. D. (1986). The effect of reference groups, opinion polls, and attitude polarization on attitude formation and change. *Political Psychology*, *7*(2), 309–321.
19. Hänska, M., & Bauchowitz, S. (2017). Tweeting for Brexit: How social media influenced the referendum. In J. Mair, T. Clark, N. Fowler, R. Snoddy, & R. Tait (Eds.), *Brexit, Trump and the Media*. Bury St Edmunds, UK: Abramis academic publishing, pp. 31–35.
20. Heidari, M., & Jones, J. H. (2020). Using bert to extract topic-independent sentiment features for social media bot detection. In *2020 11th IEEE annual ubiquitous computing, electronics & mobile communication conference (UEMCON)*. IEEE (pp. 0542–0547).
21. Hobolt, S., Leeper, T., & Tilley, J. (2020). Divided by the vote: Affective polarization in the wake of the Brexit referendum. *British Journal of Political Science*. <https://doi.org/10.1017/S0007123420000125>.
22. Kundu, S., Murthy, C. A., Pal, S. K. (2011). A new centrality measure for influence maximization in social networks. In *International conference on pattern recognition and machine intelligence* (pp. 242–247). Springer.
23. Lau, R. R., & Redlawsk, D. P. (2006). *How voters decide: Information processing in election campaigns*. Cambridge: Cambridge University Press.

24. Li, L., Scaglione, A., Swami, A., & Zhao, Q. (2013). Consensus, polarization and clustering of opinions in social networks. *IEEE Journal on Selected Areas in Communications*, 31(6), 1072–1083.
25. Lim, X. J., Radzol, A. R. M., Cheah, J. H., & Wong, M. W. (2017). The impact of social media influencers on purchase intention and the mediation effect of customer attitude. *Asian Journal of Business Research*, 7(2), 19–36.
26. Morales, A. J., Borondo, J., Losada, J. C., & Benito, R. M. (2015). Measuring political polarization: Twitter shows the two sides of venezuela. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(3), 033114.
27. Mutlu, E. C., Oghaz, T., Rajabi, A., & Garibay, I. (2020). Review on learning and extracting graph features for link prediction. *Machine Learning and Knowledge Extraction*, 2(4), 672–704.
28. Rasmus, K. N., & CS, Kim. (2014). The relative importance of social media for accessing, finding, and engaging with news: An eight-country cross-media comparison. *Digital Journalism*, 2(4), 472–489.
29. Oentaryo, R. J., Murdopo, A., Prasetyo, P. K., & Lim, E.-P. (2016). On profiling bots in social media. In *International conference on social informatics* (pp. 92–109). Springer.
30. Oghaz, T. A., Mutlu, E., Jasser, J., Yousefi, N., & Garibay, I. (2020). Probabilistic model of narratives over topical trends in social media: A discrete time model. In *Proceedings of the 31st ACM conference on hypertext and social media, HT '20, New York, NY, USA, 2020* (pp. 281–290). Association for Computing Machinery.
31. Patino, A., Pitta, D. A., & Quinones, R. (2012). Social media's emerging importance in market research. *Journal of Consumer Marketing*, 29(3), 233–237. <https://doi.org/10.1108/07363761211221800>.
32. Quattrocioni, W., Caldarelli, G., & Scala, A. (2014). Opinion dynamics on interacting networks: Media competition and social influence. *Scientific Reports*, 4, 4938.
33. Rajabi, A., Gunaratne, C., Mantzaris, A. V., & Garibay, I. (2020). On countering disinformation with caution: Effective inoculation strategies and others that backfire into community hyper-polarization. In *International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation* (pp. 130–139). Springer.
34. Riquelme, F., & González-Cantergiani, P. (2016). Measuring user influence on twitter: A survey. *Information Processing & Management*, 52(5), 949–975.
35. Salzarulo, L. (2006). A continuous opinion dynamics model based on the principle of meta-contrast. *Journal of Artificial Societies and Social Simulation*, 9(1), 13.
36. David, L. S., & Sylvia, H. (2001). *Intergroup dialogue: Deliberative democracy in school, college, community, and workplace*. Ann Arbor: University of Michigan Press.
37. Sen, A., Sen, Master Amartya, Foster, J. E., et al. (1997). *On economic inequality*. Oxford: Oxford University Press.
38. Shlapentokh, V. (1990). Public opinion in gorbachev's USSR: Consensus and polarization. *Media, Culture & Society*, 12(2), 153–174.
39. Sides, J., & Hopkins, D. J. (2015). *Political polarization in American politics*. New York: Bloomsbury Publishing.
40. Simmie, D., Vigliotti, M. G., & Hankin, C. (2014). Ranking twitter influence by combining network centrality and influence observables in an evolutionary model. *Journal of Complex Networks*, 2(4), 495–517.
41. Soares, F. B., Recuero, R., & Zago, G. (2018). Influencers in polarized political networks on twitter. In *Proceedings of the 9th international conference on social media and society* (pp. 168–177).
42. Takikawa, H., & Nagayoshi, K. (2017). Political polarization in social media: Analysis of the “twitter political field” in Japan. In *2017 IEEE international conference on big data (big data)* (pp. 3143–3150). IEEE.
43. Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., & Nyhan, B. (2018). Social media, political polarization, and political disinformation: A review of the scientific literature. *Political polarization, and political disinformation: a review of the scientific literature (March 19, 2018)*.
44. Weber, I., Kiran, G., Venkata, R., & Batayneh, A. (2013). Secular vs. islamist polarization in egypt on twitter. In *Proceedings of the 2013 IEEE/ACM international conference on advances in social networks analysis and mining* (pp. 290–297).