Social Network Analysis and Mining (2020) 10:68
https://doi.org/10.1007/s13278-020-00676-1

ORIGINAL ARTICLE

Analyzing the relationship between relevance and extremist discourse in an alt-right network on Twitter

Javier Torregrosa[1] · Ángel Panizo-Lledot[1] · Gema Bello-Orgaz[1] · David Camacho[1]

Received: 7 March 2020 / Revised: 18 June 2020 / Accepted: 15 July 2020 / Published online: 13 August 2020

Abstract

Nowadays, Twitter is used by several political extremist groups to establish close communities on which the opinions are amplified following an echo-chamber effect. However, few literature analyses the effect of the use of an extremist discourse in relation to the relevance of these users on their online network. With the aim of analyzing this effect, this work studies the relationship between the use of indicators of extremist discourse from users belonging to an alt-right network on Twitter and their relevance on it. The network of alt-right users is created using the retweets of 96 accounts where the user relevance is measured by five different types of centrality metrics, including *in-degree*, *eigenvector*, *k-shells*, *betweenness*, and *closeness*. Both the linguistic indicators and the tone were analyzed using LIWC and VADER software. The network analysis outcomes show that user relevance on the network is indeed related to the use of an extremist discourse. Finally, this relationship is also tested on different corpus of texts and about different topics, being found that this relationship is more clear on retweets made by the users and when discussing about hate speech topics.

Keywords Social network analysis · Centrality metrics · Linguistic · Political extremism · Psychology · Far-right

1 Introduction

Among the different Social Media platforms, Twitter has become especially popular thanks to its public interaction nature (Camacho et al. 2020; Yardi and Boyd 2010). This open platform allows to quickly spread messages and media content to the user's network, and potentially to the network of the target users. Features such as the mention or the retweet are used with the aim of sharing content, which in turn creates "*an ideal platform for users to spread information*" (Stieglitz and Dang-Xuan 2012).

Political information remains as one of the most popular content shared by users (Kreis 2017), who constantly share their opinion both supporting their ideologies and criticizing the contraries. These opinions, nonetheless, are usually in consonance with the environment of the user and thus can be influenced by relevant people of their social network. In fact, relevant users of specific networks can guide the direction of the public opinion of that network and lead information propagation (Cha et al. 2010; Romero et al. 2011). Their relevance on the network and the interactions they made with the rest of the people can modulate different flows, including behavior, opinion or emotions. These users, who are often called "opinion leaders," can be analyzed in terms of their role on the network (Riquelme and González-Cantergiani 2016).

As other communities, political groups also have "opinion leaders" in their social networks. Most of the users involved establish themselves in homogeneous online communities (Conover et al. 2011), creating an echo-chamber on which the dominant messages and opinions are reinforced. As they interact more with same-minded people than with ones with different political views, the echo-chamber effect is increased. Among these political groups, extremist movements have attracted more attention on recent years, due to the relevance they have reached on the current worldwide political context. Understanding

✉ David Camacho
david.camacho@upm.es

Javier Torregrosa
franciscojavier.torregrosa@upm.es

Ángel Panizo-Lledot
angel.panizo@upm.es

Gema Bello-Orgaz
gema.borgaz@upm.es

[1] Departamento de Sistemas Informáticos, Universidad Politécnica de Madrid, Madrid, Spain

how their networks are built, and how they manifest their ideology, can facilitate the understanding of how extremist users behave, feel and act inside their communities (Zannettou et al. 2018).

The relevance of a user in a network is an ill-defined term and is measured in different ways in social network analysis (SNA) , such as *centrality*, but also as *popularity* or *prestige*. Using Twitter as an example, one can consider an account relevant if its messages are retweeted by a lot of users. On the contrary, one can consider an account relevant if its messages are retweeted by a few but very important users. What's more, one can consider a user relevant regardless of the retweets received, focusing instead on its ability to act as a bridge, or hub, between different group of users. As a matter of fact, this diversity of criteria has made assessing the key users in a network an active research field, on which several metrics have been proposed over the years (Das et al. 2018; Bello-Orgaz et al. 2016). Nevertheless, in recent years centrality measures have been the most commonly applied method in SNA domain to measure the relevance of users within a network.

Considering all of the above, this work presents an exploratory observational analysis of the relation between the extremist discourse of a user and their relevance inside the network, understanding relevance as scoring high in certain centrality. Because there are several types of centrality measures, each one with its own criteria, and little insight into how they relate to radical discourse in Social Media, it is impossible to take an educated guess of what metrics should be used. Therefore, this works aims to start building that insight by studying how the most common centrality measures found in SNA are related to the written discourse of the users within an extremist network, especially of users from an alt-right community on Twitter (Graham 2016; van der Vegt et al. 2019), a subtype of far-right group with high online activity. The extremist discourse is understood as the particular terminology used by extremist groups on their messages, which reflects their ideas, cognitive processes, and emotional state, and which is measured using different linguistic indicators (Alizadeh et al. 2019; Figea et al. 2016; Forscher and Kteily 2017; Greven 2016; Grover and Mark 2019; Kaati et al. 2016; Lyons 2017; Panizo-LLedot et al. 2019; Torregrosa et al. 2019; Zhang et al. 2018). These indicators have been chosen after checking the literature about the discourse of different extremist groups. The objective of this study is to relate the use of these indicators with the relevance that users reach on the network. To do that, the next research questions are formulated:

- *RQ1: How are different relevance criteria related to discourse?* To do so, different centrality measures are compared with the linguistic indicators used by each user through all the content they share.
- *RQ2: Does the discussed topic affect the relationship between relevance and discourse?* To do so, different centrality measures are compared with the linguistic indicators used in different topics, determined by the hashtags used by the users.
- *RQ3: Does the corpus of analyzed text affect the relationship between relevance and discourse?* To do so, different centrality measures are compared with the linguistic indicators from the different corpus of text.

We address these questions through SNA and Natural Language Processing (NLP) techniques. Specifically, we use five different centrality metrics (in-degree, closeness, betweenness, eigenvector and k-shells) as a way to measure the relevance of an alt-right user. These five measures are calculated over a *directed weighted graph* that captures the retweeting activity among the Alt-Right accounts. Besides, we use two lexicon-based NLP tools for analysis the discourse of user accounts, LIWC2015 (Pennebaker et al. 2015) and VADER (Hutto and Gilbert 2014). The former is a well-known tool used to extract linguistic patterns from corpus of text. The later is a sentiment analysis tool specifically attuned for social media.

The main contributions of this work can be summarized as follows: (1) Studying the relationship between the use of words and user's relevance in an alt-right community on Twitter, in order to understand if the use of an extremist language increases the relevance of the actor on the network. (2) Analyzing which linguistic indicators of the extremist discourse are the most used by the most relevant users of the network by topic. This detailed analysis will show the differences that exist in the extremist discourse using the topics as different contexts. (3) Identifying how different relevance criteria (centrality measures) relate to the discourse. For this reason, the study has been carried out using various measures of centrality that consider different topological aspects of the network for its computation. (4) Increasing the literature about the extremist discourse on far-right groups, as the chance of studying an online-native group allows a better understanding about their interactions and terminology.

The rest of the paper has been structured as follows: Section 2 provides a brief introduction to the alt-right speech and reviews previous works that also study the relationship between speech and user relevance. Section 3 describes the process followed to extract insights from the outcomes. Section 4 presents the results, whereas Section 5 provides a detailed and exhaustive discussion of these results. Finally, Section 6 draws some conclusions and future research lines of work.

2 State of the art

2.1 SNA techniques on the study of the relationship between user relevance and user discourse

In social networks, there is a type of users who can guide the direction of public opinion and have the ability to propagate information and spread trends in a short time (Cha et al. 2010). These users are capable of expressing their ideas with a greater impact than other people in the network. Thus, being able to analyse and detect these relevant or highly influential on social networks becomes a crucial issue in areas such as politics and security, where extremist movements try to propagate their political ideologies using social networks in media platforms (Conover et al. 2011).

From a political perspective, social networks in media platforms have played a fundamental role in the current changing political landscape, playing a critical role in certain events such as national elections in different countries, including France, Germany or the USA. For this reason, several works have appeared in recent years analyzing this issue using SNA techniques (Davidson et al. 2020; Tumasjan et al. 2010; Nguyen and Jung 2019). For example, Davidson et al. (2020) explored the behavioral differences for the Twitter usage in the elections in France and USA. They modeled the Twitter data as a complex graph where the edges represent the retweets between the users and where the vertexes are labeled with behavioral data (vector of the hashtags used by the users) to check the hashtag usage by different communities. The outcomes show that Twitter was used mainly for debate in the French election, whereas it was intensively used to energize others in the same community during the US election.

Other work following this same research line was presented by Tumasjan et al. (2010). This work analyzed whether the content of Twitter reflected the political sentiment during the German federal election. Using LIWC software for text analysis, the authors conducted a content analysis of over 100,000 tweets containing a reference to either a political party or a politician published in the weeks leading up to the federal election of the national parliament in Germany (which took place on September 27th, 2009). They found evidence of a lively political debate on Twitter, but this discussion was dominated by a small number of users: only 4% of all users accounted for more than 40% of the messages. This study also found that the sentiment profiles of politicians and parties reflect many nuances of the political campaign, e.g. the similar profiles of Angela Merkel and Frank Walter Steinmeier, mirroring the consensus-oriented style of their grand coalition.

Both works analyze how Twitter content may have affected some specific election campaigns, but do not take into account the relevance of users who published such content on the social network. In fact, there are studies showing that relevant users can guide the direction of the public opinion of the social network and lead information propagation (Cha et al. 2010; Romero et al. 2011). On this research, we propose a methodology that not only analyzes the content that users publish on the social network, but also relates it to how relevant they are within this network.

To address the issue of measuring user relevance in a social network, analysis of the graph-based representation of user interactions in it is the most commonly used method in SNA area. The structural information from graphs is very useful to identify their most important nodes (users in this case). These relevant users are usually key members of the network. So, centrality measures defined in graph theory can be used to calculate the degree of relevance of users into the network (Das et al. 2018): degree centrality, betweenness centrality, closeness centrality, eigenvector centrality, etc. For this reason, this study will use centrality measures to estimate of the relevance of the users.

There are research papers that focus on other important aspects, besides the message content, such as the topics approached. For example, Romero et al. (2011) shown that the relationship between hashtag exposure and hashtag adoption depends on the topic. This study shows that, in the political domain, repeated exposure to the same hashtag increases the chances of adopting it. In contrast, repeated exposure to hashtags in the music domain does not increase probabilities of adoption. Therefore, in order to analyze if the relevance of a user within a social network is related to his or her discourse, the topics that the user uses shall be taken into account.

In the literature, there are studies that consider both aforementioned aspects (graph-based and linguistic content analysis) for detecting and studying the behavior of relevant users in social networks. For example, Deturck et al. (2017) present a hybrid system combining both aspects where relevance detection is performed by analyzing a graph-based representation of user interactions in a social network and by measuring the impact of the linguistic content of user messages in online discussions. The authors first extract structural information to estimate the user influence among interaction networks (retweet and flow graphs between Twitter users) applying different centrality measures. Then, they identify the linguistic features of the influencing behaviors based on a morphosyntactic analysis performed by the Eloquant Semantic Solutions parser (Tan et al. 2011).

With a similar approach than the previous article, the research of Chen et al. (2020) studied the sentiment correlation for different groups of users on Twitter to examine the influential factors for retweeting a message with a certain sentiment. In this study, authors propose a method for measuring the sentiment of tweets to categorize the Twitter users into different groups by using several characteristics, which are: the follower count; the betweenness connectivity; a combination of follower count and betweenness centrality; and the number of tweets. This study found that users with higher betweenness centrality and higher tweets amount tend to exhibit a higher sentiment correlation. The users with medium-level followers_count show the highest sentiment correlation compared to the low-level and high-level followers_count.

Based on the aforementioned studies, which analyze the relationship between relevance and both morphosyntactic or sentiment analysis, the methodology proposed in this article combines both approaches to conduct a more detailed study. In this case, instead a common political group, this article will focus on an online native extremist community, the alt-right, which can offer an opportunity to analyze the relationship between discourse and user relevance in a far-right community.

2.2 The alt-right's discourse on social networks

The alt-right, a highly decentralized extremist movement originated on Internet blogs, is currently one of the most active far-right groups on Twitter. Taking advantage of Social Media platforms, they use them as speakers to encourage discussions about politics with extremist approaches, such as racism, supremacism or antisemitism (Graham 2016). This group, which was born in the context of different webs and blogs related (such as 4chan), has nowadays a great presence on Twitter. Firmly defenders of different far-right policies, they have gained relevance since the "Unite the Right" rally on 2017 (van der Vegt et al. 2019). However, their discourse and symbolism makes them easier to trace than other groups, and therefore to study their interactions on Twitter. For this reason, we have selected this extremist community in this paper to study the relationship between user relevance and the use of an extremist discourse.

The alt-right's discourse became specially relevant during the 2016 presidential elections on USA, when they acted sharing and amplifying Donald Trump's discourse, specially concerning immigration (Zhang et al. 2018). They took a critic position about different topics related to immigration, nationalism, traditional media or political correctness, among other issues. As many other political extremist groups, they used Twitter as a speaker to send their messages about those topics, following the campaign of the Republican Candidate (Graham 2016).

The discourse of the alt-right sympathizers is determined by their ideology, which is deeply rooted on a racist and anti-immigrant position (Mirrlees 2018). Even though it shares certain viewpoints with different ultra-conservative groups (especially those concerning nationalism), there are some ideological insights which makes this group different from others, like their position about religion or traditional family. In fact, as stated on Panizo-LLedot et al. (2019), who analyzed an alt-right Twitter network during the 2018 mid-term elections on USA, most of the topics discussed by the group remained stable during time periods, and include a few isolated communities talking about non-politically related topics. Religion and family were not present among the topics found. Previous research also shows similar patterns (Forscher and Kteily 2017; Greven 2016; Lyons 2017). On the basis of these studies, the specific topics to be analyzed in this paper have been selected.

Beyond the topics analyzed, the terms used on the discourse are also an important characteristic of any extremist group. To the knowledge of the authors, few research has been conducted on this topic, with most of this work focused on general extremist and far-right groups. Alizadeh et al. (2019) found that, regardless of their ideology, political extremists on Twitter use less positive emotion and more negative emotion words compared to non-extremist users in USA. Also, they found that far-right extremists used more positive emotion words and less negative emotion words than the far-left users. A possible explanation of the authors was that having a Republican president made far-right users more satisfied with the policies, and then they use more positive emotions. To address the issue of emotion in discourse in this paper, the methodology proposed includes a part of sentiment analysis as part of the process to be applied.

Torregrosa et al. (2019) found that, compared with a random sample from Twitter, Islamic radical supporters had more negative discourse, with more third person plural and less first person singular pronouns, and more words related to religion, anger, certainty and death. As these outcomes were in line with previous research on Jihadi extremism, they theorized that these features could be generalized to other extremists discourses. These dictionaries have also been used to analyse far-right forums and extract general tendencies about worries, racism and aggression with promising outcomes (Figea et al. 2016; Kaati et al. 2016).

Concerning the topics commonly addressed by far-right groups, Grover and Mark (2019) proposed dictionaries which included terms commonly used by alt-right supporters. These dictionaries included words related to racism, white race references, different ethnicities, immigration and racial slang. These dictionaries, which were useful for the objective of detecting warning behaviors online, are open on the aforementioned article.

Finally, following the dictionaries found on the literature for the extremist discourse shown by the above-mentioned studies, this paper will compare the use of the terms they contain with different relevance measures from the users of an alt-right community.

3 Methodology

This section describes the methodology followed to analyze the relationship between a user's relevance and its discourse. All of the conducted processes are described in detail below, from the data collection to the study of the user's centrality and their discourse. This methodology is designed as a 6-step process with the next phases: (1) *Data gathering*: a collection of tweets is built, (2) *Retweet network construction*: a directed weighted network where nodes are accounts and edges join two nodes if any of them have retweeted each other is built, (3) *Centrality analysis*: five different centrality measures are calculated for each node, (4) *Linguistic analysis*: for each account, several linguistic analyses with different types of messages (tweet/retweet) are done, (5) *Hashtag analysis and text categorization*: for each account, several linguistic analyses with messages of different topics are done and (6) *Correlation analysis*: correlation between the centrality measures and the different linguistic analysis is done. Please notice that only one network is generated, and therefore, each account or user will have only five centrality values. Contrarily, each user will have several linguistic analyses each one conducted over a different set of tweets published by that account on a specific topic. This means that an account does not necessarily participate in all the correlations performed. For example, if a user has not published any tweets talking about politics that account will not be considered in the correlation between the topic Politics and the different centrality measures.

3.1 Data gathering

We built a dataset, periodically checking Twitter's API, which contains 116, 387 tweets posted by 96 alt-right Twitter accounts between the 1st of July and the 1st of November of 2019 (4 months). The 96 alt-right accounts were picked from the dataset created in Thorburn et al. (2018) work. That work used twelve twitter accounts advertised in posters of 2017 Unite the Right Rally in Charlottesville (Virginia) as seeds. Next, 546 accounts were randomly selected from all the accounts that followed six or more of these twelve aforementioned accounts and have published at least 20 tweets. Finally, those 546 accounts were manually labeled, reading their published tweets, to ensure that they were a valid representation of the alt-right ending with a dataset of 422 alt-right accounts. It is worth mentioning that at the time of this writing all the aforementioned accounts have been shut down and the selected period corresponds to their last four months of activity.

Nevertheless, the accounts were fairly active before being ceased. The median number of tweets published by an account is 173 with an interquartile range of 750. Furthermore, approximately 80% of the accounts have posted at least 5 days and 50% of them at least 30. Figure 1a–c shows several graphs related to user activity.

From the 116, 387 tweets, $\approx 43\%$ of them are original tweets, and the remaining $\approx 57\%$ are retweets. This trend repeats itself for each month; the number of retweets published is slightly higher than the number of tweets. Apart from this little difference in activity intensity, both (tweets and retweets) follow a similar activity pattern as shown in Fig. 1d.

Regarding the hashtag activity, 8, 389 messages ($\approx 7\%$ of the dataset) contain one or more hashtags. From those 8, 389 messages, 2, 982 ($\approx 3\%$) where original tweets and the remaining 5, 407 messages ($\approx 4\%$) are retweets. As expected from the message posting activity of the dataset, we observe a higher number of retweets with hashtags than tweets.

3.2 Retweet network construction

In this research, retweets are used to build the network, as they have been found to represent interest and trust in another one's content across the literature (Metaxas et al. 2015). We build the retweet network using a *directed weighted graph* where nodes are Twitter accounts, and an edge joins two nodes (A, B) if A has retweeted B. The weight of an edge, between node A and node B, is the number of times that A has retweeted B. Therefore, node A is connected to node B with a weight of 50, and to node C with a weight of 1, which means that a tweet from node A can be more easily retweeted by node B than by node C. (So, we can use this weight as a diffusion probability.)

To build this retweet network, we use only the retweets where the original account posting the message is not the same as the account retweeting it (is not an auto retweet), and where the messages are not duplicated (a message is duplicated if an account has published the same exact text twice). Also, both accounts (the one who published the original tweet and the one who retweets it) have to be members of the confirmed 96 alt-right accounts. Then, only the biggest connected component of the aforementioned graph is used. After applying these criteria, we get 4208 retweets and a graph with 91 nodes and 567 edges with an average in/out-degree ≈ 6. Figure 2 shows the aforementioned graph.

3.3 Centrality analysis

Five different centrality measures, *in-degree, closeness, eigenvector, betweenness* and *k-shells*, are calculated for each of the 91 nodes in the retweet network. Each centrality measure considers a node relevant in the network for different reasons (Das et al. 2018). Therefore, scoring high in any

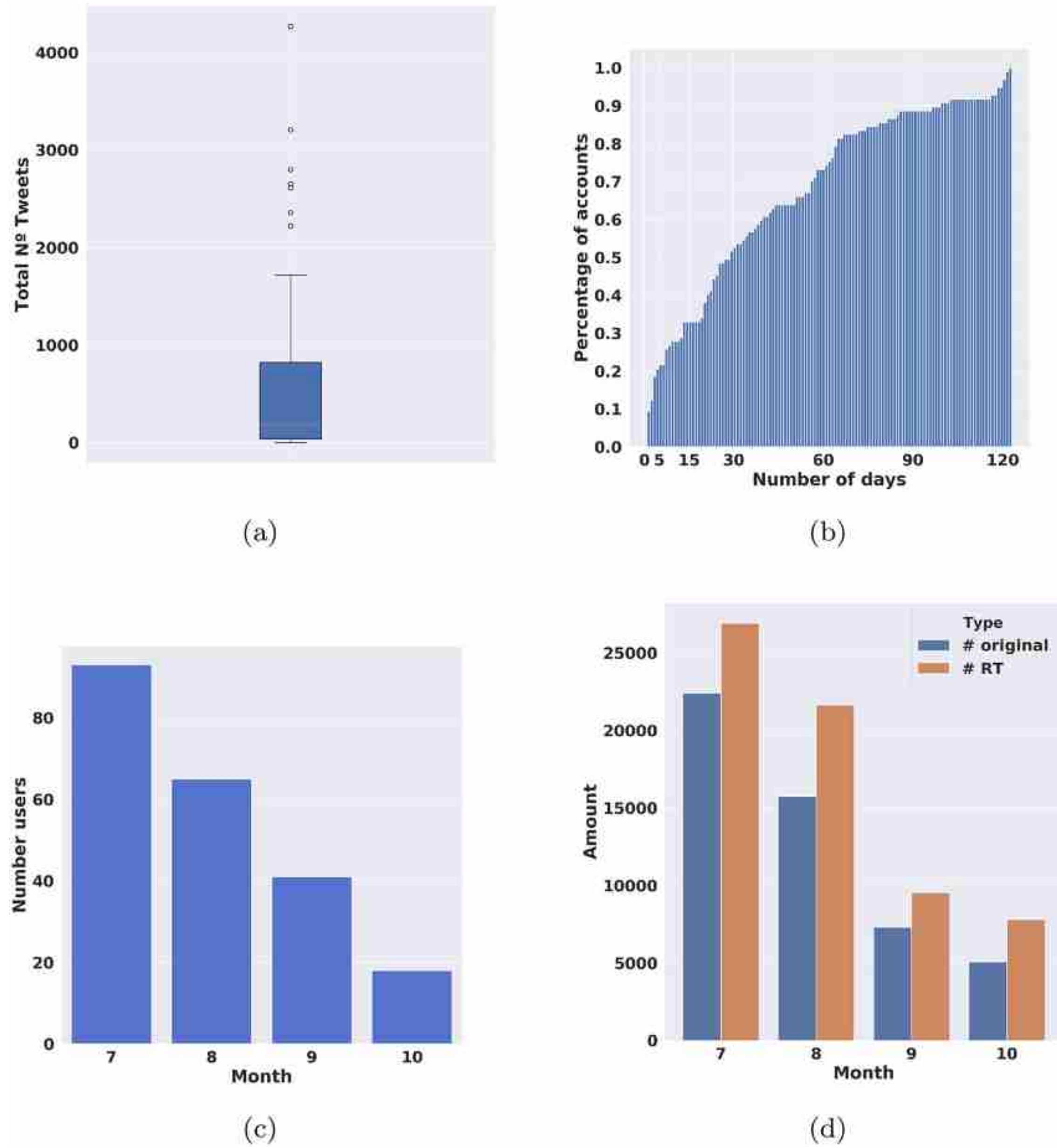

Fig. 1 **a** Shows the distribution of the total number of tweets published by each account during the four months, **b** shows the cumulative histogram of the number of days that each account has posted, **c** shows the number of users that have published each month, and **d** shows the number of original tweets (blue) and retweets (orange) published each month.

of them means different things inside the retweet network. Below we list each of the aforementioned centrality measures and its meaning inside the retweet network:

1. *In-degree* Freeman (1978): counts the number of connections pointing to a node. The users with the maximum in-degree in our network are the ones that have received the highest number of retweets, not taking into account who has retweeted them and they will be considered *influencers*.
2. *Closeness* Freeman (1978): measures how close a node is to every other node in the network. The idea of this measure is that the more central the nodes are, the easier it will be to reach other nodes. It is calculated as the sum of the length of the shortest paths between the node and all other nodes in the network. Therefore, the smaller the average shortest path length is, the centrality value of the node will be higher, and more central a node is. The users with the highest closeness in our retweet network are the ones whose messages can reach all accounts (not

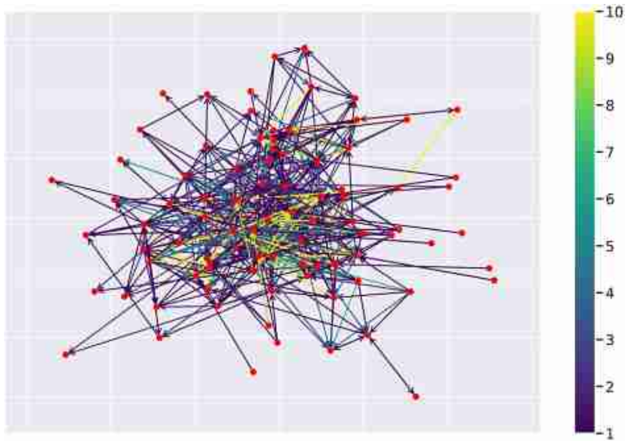

Fig. 2 Retweet network, the color of an edge indicates how strong is that connection. On the right, the color bar indicates the mapping between color and strength

necessarily the most important ones) more easily and they will be considered *disseminators*.

3. *Betweenness* Freeman (1978): measures the number of shortest paths that pass through a specific node. This measure shows how central is a node connecting any other pair of nodes into the network. It is an indicator of the control over the flow of information between nodes. The node with the higher betweenness in our network is the one that has more control over the message flow. In other words, if this node disappears, all the other nodes messages have more difficulties to reach the rest of the accounts in the retweet network. Nodes scoring high in betweenness will be considered *bridges*.
4. *Eigenvector* Bonacich (1987): considers that a node is important if it is connected to other important nodes. Measure how influential a node is inside the retweet network. This measure tries to generalize the degree centrality by incorporating also the importance of the neighbors. It assigns relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. The node with the highest eigenvector centrality is a node whose messages reach easily other influential nodes and will be considered *authorities*.
5. *K-shells* Kitsak et al. (2010): measurement that came from studies of diffusion models. It is based on the concept that ideas can be spread likewise infectious diseases propagate within a society. In these diffusion models, there are cases where the best spreaders do not correspond to the most highly connected or the most central nodes of the network. This method is better than the degree centrality in many real networks with a core-like structure. In this measure, the network is decomposed into many shells to identify the most influential spreaders based on the assumption that nodes in the same shell have a similar influence, and nodes in higher shells are likely to infect more nodes. The nodes with the highest k-shells centrality will be considered *super-spreaders*.

3.4 Linguistic analysis

Before doing the linguistic analysis of the text, we clean all the tweets in the dataset. First, all the mentions(@someone), hashtags, and URLs were removed from the posted text. Then, all the empty messages, the ones containing only punctuation signs and the duplicated ones, were removed too. Next, we group all messages published by each account into four different text corpora. Each text corpus is composed of different types of messages (tweet/retweet) to study whether there are differences between the type of discourse a person uses when publishing original opinions or retweet. The text corpora considered are:

1. *all_messages* corpus: contains all the messages published by an account whether they are tweets or retweets.
2. *original_tweets* corpus: contains only the original messages (tweets).
3. *RT_done* corpus: contains the retweets that the account has done.
4. *RT_received* corpus: contains the retweets that the account has received.

Finally, the linguistic analysis is done over all the messages posted by an account in each text corpus independently. Hence, we will calculate four different vectors of linguistic variables (one for each text corpus) for every node (user account) in the retweet network.

For the linguistic analysis, two lexicon-based NLP tools are used: LIWC-2015 (Pennebaker et al. 2015) and VADER (Hutto and Gilbert 2014). On the one hand, LIWC is a well-known text analysis tool that extracts psychological insights from the way people use words when expressing themselves. It is based on word-counting and comes with 108 dictionaries representing different concepts like Sadness, Power, Verbs, or Past-Tense. Moreover, LIWC is customizable and allows enhancing its base dictionaries with task-specific ones. On the other hand, VADER (Hutto and Gilbert 2014) is a sentiment analysis tool specially designed to be used on social media. Its provides four different scores on the sentiment of a text (*positive*, *negative*, *neutral* and *compound*)

Not all the scores given by both tools are used in the linguistic analysis. From VADER, only the *compound* score is used in this work. This score encodes the polarity of a text into one single real value between -1 (most negative) and +1 (most positive). Regarding LIWC-2015, dictionaries previously found related to an extremist discourse are used.

These include *power, death, relig, anger, negemo, posemo, sixltr, i, we, they, you* and *shehe*. In addition to those scores, we include four new task-specific dictionaries proposed by T.Grover and G.Mark related to far-right speech in social media (Grover and Mark 2019) (*whiteEthno, JewOrBlack, OtherRacial* and *RacialSlang*). Table 1 contains a short description of the meaning of each dictionary.

3.5 Hashtag analysis and text categorization

The linguistic analysis described above considers all the messages posted by an account regardless of the topics they talk about. In this step of our methodology, we want to consider the topics of the messages to analyze if there are differences in the type, and tone, of the speech according to the topics. To perform this task, we first need to classify the tweets by topic using the hashtags included in them. Depending on the nature of those hashtags, a tweet will be classified into one or more different topics. To decide the topics in which the hashtags will be categorized, we use our previous research as a starting point (Panizo-LLedot et al. 2019). In that paper, the topics discussed during the 2018 US midterm elections, by the same twitter accounts using in this work, were analyzed and six general topics were detected. Those topics were: *Racial Discourse, Politics, Immigration, Media, Elections* and *Policy Debate*.

After manually checking all the hashtags available in our dataset and filtering those that have not been used at least 5 times, we decide to use five categories: *Conspiracy, Hate Speech, Media, Neutral* and *Politics*. Table 2 contains the hashtags included in each category. Although the selected five topics are very similar to the six found in the past investigation, some differences exist. First, we do not find enough hashtags related to *Immigration* to justify having its own category. Therefore, we decide to merge the *Racial discourse* and *Immigration* categories into the *Hate Speech* one. In addition, we observed an increase in hashtags related to conspiracy theories. In the midterm elections research, the *Conspiracy* topic was part of the *Media* one. However, this time, we decide to take *Conspiracy* out of *Media* and into its own category. Finally, two topics were not present anymore: *Elections* and *Policy Debate*. The disappearance of these topics was expected as no ballot occurs during the period of study.

3.6 Correlation analysis

To analyze the relationship between centrality measures (relevance) and the four different vectors of linguistic variables (one for each text corpus) calculated for each user account in the retweet network, a correlation analysis is conducted. Due to the nonparametric nature of the distributions on several dictionaries, it has been decided to use Spearman's Rho as a correlation test. Spearman's Rho works as a more robust correlation parameter than Pearson's R when analyzing data without normal distribution. Among the three different

Table 1 Description of the selected linguistics indicators

Dictionary	Description
LIWC	
power	words related to domination like superior, bully, administrator, ambition...etc
death	words related to death like kill, dead, alive, autopsy...etc
relig	words mostly focused on Christians but can include terms related to other religions like agnostic, god, heaven, altar...etc
anger	words related to anger like rage, abuse, yelling, threatens...etc
negemo	words related to negative emotions like :(, adverse, angry, burden...etc
posemo	words related to positive emotions like :), agree, win, triumph, sunny...etc
sixltr	words with more than six letters (long words)
i	first person singular pronouns, words like I, me, mine, methinks...etc
you	second person singular pronouns, words like thy, u, ye, thee...etc
shehe	third person singular pronouns, words like thy, u, ye, thee...etc
we	first person plural pronouns, words as we, us, our, lets...etc
they	third person plural pronouns, words like he, she, his, her...etc
WhiteEthno	words related to white identity like white, aryan, ethnostate...etc
JewOrBlack	words related to jewish or black identity like jews, black, ashkenazi, zionist...etc
OtherRacial	words related to other identities like asian, mexican or arab like muslim, mestizo, ethnic, asian...etc
RacialSlang	slang terms related to the alt-right community like goy, kebab, alien, wuz...etc
VADER	
Emotional Tone	Global metric that uses all the words in the text and ranges from −1.0(negative text) to +1.0(positive text).

Table 2 Hashtags included in each of the five selected topics

Topic	Hashtags
Politics	MAGA, AmericaFirst, Editorial, MAGA2020, Trump2020, KAG2020, Trump, KAG, Brexit, DemDebate, impeach, TrumpForever, Antifa, TrumpRally, Thread, KeepAmericaGreat, 4MoreYears, MuellerHearing, MuellerHearings, California, SaluteToAmerica, GunControlNow, DemocraticDebates, LosAngeles, 2A, EU, poll, ElPasoShooting, FourMoreYears, ElPaso, HiteAmerican, DemocraticDebate, RedFlagLaws, China, Democrats, RINOCARE, BorisJohnson, GiletsJaunes, AmericaLast, USA, TradeWar, BlueToRed, WomenForTrump, Analysis, Portland, DemocratsAreDestroyingAmerica, BREXIT, Greenville, Ukraine, NoDealBreixt, StraightPrideParade, London, DemDebate2, CNNDebate
Media	CNNCredibilityCrisis, Killstream, FakeNews, CNNsotu, Poll, G7Summit, EnemyOfThePeople
Hate Speech	WhiteGenocide, PAxJudaica, DeportThemAll, BuildTheWall, DIVERSITY, WalkAway, Antifa, HateHoax, WhitePeople, ICEraids, NoOn1044, IllegalAliens, NeverForget, MeToo, HR1044, UndocumentedImmigrants, AntifaTerrorists, Pantheon, Kabbalah, redpilled, LOL, IllegalImmigration, WWII, WhitePeopleAgainstRacism, antiwhite, REDPILLED, ANTIFA, Universality, TeamWhite, WakeUp, WhiteLivesMatter, Diversity, Tucker, ChasingDownWhites, Whites, deported, BlackMagic, REMINDER, LGBTQ, Paris, antifa, Immigration, WHITEGENOCIDE, Reminder, USWomensSoccerTeam, StandUpForEurope, WSHIS
Neutral	AmericansHelpingAmericans, Bitcoin, Blockchain, AI, FridayThoughts, BREAKING, TuesdayThoughts, BASED, IndependenceDay, Futurist19, FridayFeeling, WashingtonElite, SundayThoughts, WednesdayWisdom, BAltimoreCleanup, ThursdayThoughts, July4th, MondayMotivation, based, SaturdayMorning, SMH, ItalianArt, YellowVests, GlobalBoost, GroyperSeason, FF, Vienna, WednesdayThoughts, DIACWFWG, crypto, FFXIV, FrenchArt, StillWithHer, Israel, IceBae, MiningDisrupt, pubg, SmartCity, HongKong, RIP, FourthOfJuly, 4thofJuly, Toronto, DogRight, SundayMorning, TuxedoPepeTwitter, OWNED
Conspiracy	Epstein, Freemasonry, KalergiPlan, EpsteinSuicide, ClintonBodyCount, JeffreyEpstein, Freemasons, epstein, NewWorldOrder, Freemason, JeffreyEpsteinArrest, israel, EpsteinMurder

perspectives stated by Akoglu (2018) to interpret the outcomes of the correlation, who includes the psychological, the political and the medicine approaches, it was decided to use the former, due to the nature of the data analyzed. In this perspective, the Spearman's Rho correlations are interpreted according to the next index (the sign of the value determines the direction of the correlation):

- $|1|$: perfect relationship.
- $|0.7 - 0.9|$: strong relationship.
- $|0.4 - 0.6|$: moderate relationship.
- $|0.1 - 0.3|$: weak relationship.
- $|0|$: no relationship.

4 Network analysis outcomes

To perform a detailed analysis of each centrality measure, this section is divided into five subsections (one per centrality). Each subsection includes a heatmap that shows the correlations between a specific centrality and the linguistic variables obtained in the different corpus of text. The first four rows of the heatmap show the correlations obtained for each of the text corpus defined in Section 3.4 (linguistic analysis): RT_done, RT_received, all_messages, and original_tweets. Furthermore, the last five rows contain the correlations obtained for the five discriminated topics described in the Methodology section (*Conspiracy*, *Hate Speech*, *Media*, *Neutral* and *Politics*). This division of the corpus will allow us to carry out a study of the differences between the type of discourse that a user employs when publishing original opinions or retweets, and when talking about different topics.

Regarding the columns of the heatmaps, each column represents a linguistic variable extracted from the analysis performed using the two lexicon-based NLP tools (VADER and LIWC2015). Finally, each significant correlation appearing in the heatmap is marked using one to three asterisks whose meaning is: * ($p < 0.05$), ** ($p < 0.01$), or *** ($p < 0.001$).

4.1 In-degree

Nodes with higher in-degree have more connections with adjacent other nodes (more edges). In a retweet network, this represents the accounts that receive more direct retweets.

Analyzing correlations by the text corpus shown in Fig. 3, it can be seen that the language does not seem to follow a specific pattern. Original tweets of the users with high in-degree scores tend to have fewer words of negative emotion, more third-person singular words, and words related to minorities (JewOrBlack column). In the text that high in-degree users retweet from others (RT_done row), there are more words related to death, racial slang, and more first-person plural pronouns. The retweets that other users make from them (RT_received row) include significantly more first-person plural pronouns and more references to other racial minorities (for example, black people or jews). Finally, when analyzing together all the text generated on the timeline of the users, it can be seen that only anger and negative emotion words are negatively correlated with this centrality.

Focusing now on the topics, it can be seen that high in-degree centrality is not related to any of the dictionaries on

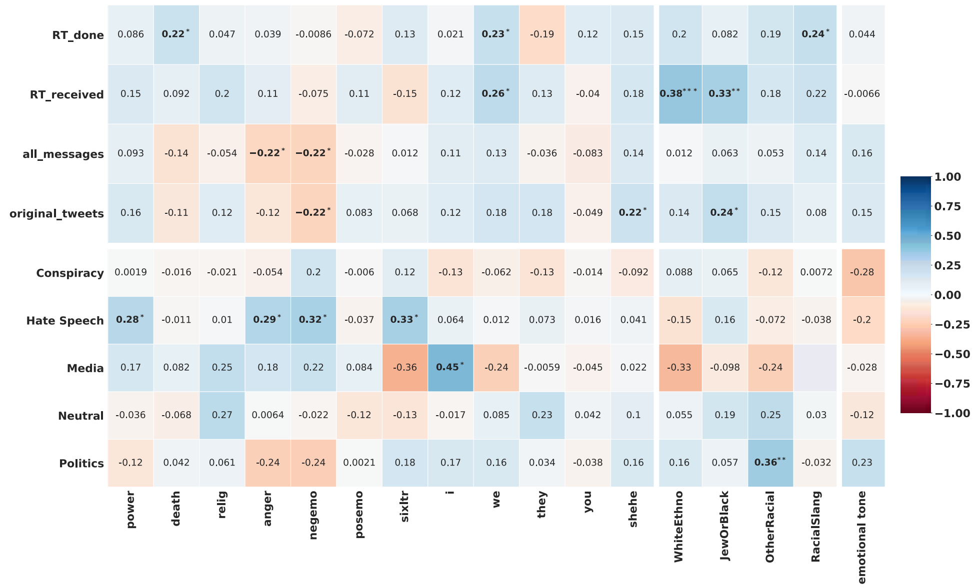

Fig. 3 Correlation heatmap for in-degree centrality. Positive correlations are shown in blue. Contrarily, negative correlations are shown in red. Finally, the color intensity of a cell indicates the strength of a correlation and the significance of it is shown using one to three asterisks whose meaning is: *($p < 0.05$), **($p < 0.01$), or ***($p < 0.001$)

Neutral or Conspiracy themes. When users are talking about Politics, some racial terms are directly related to the centrality score, and the same happens with first-person singular pronoun when talking about Media. However, when the discussion is about Hate Speech topics (racism, anti-feminism, immigration, etc...), the users with more in-degree centrality use more words related to power, anger, negative emotions, together with longer words (six letters or more).

4.2 Eigenvector

A higher score on the eigenvector centrality means that a user has strong ties to other users of the network that also score high on this centrality. In our case, this represents authorities inside the network.

Figure 4 shows the correlation between eigenvector centrality and the linguistic variables. One of the most outstanding outcomes found analyzing this figure is that power words are significantly correlated with high centrality accounts in all the text corpus (original tweets, all messages, retweets made by the user, and retweets that a user received). A similar thing happens with the third-person singular pronouns, except in the retweets done by the account. Finally, first-person plural pronouns seem to be directly related to the centrality for the retweets a user receives and its original tweets.

When analyzing the correlations between the eigenvector centrality and the language on the different topics, it can

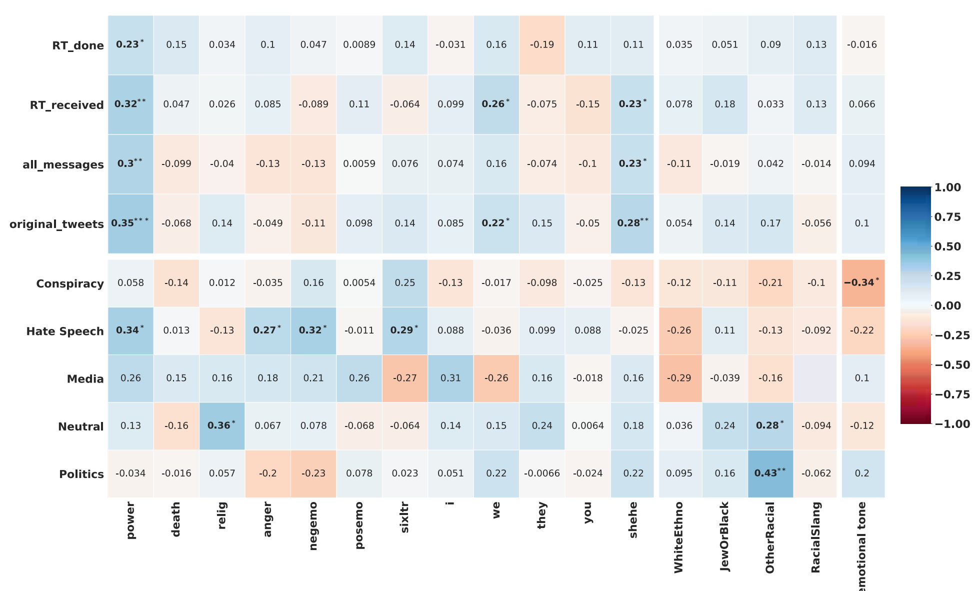

Fig. 4 Correlation heatmap for eigenvector centrality. Positive correlations are shown in blue. Contrarily, negative correlations are shown in red. Finally, the color intensity of a cell indicates the strength of a correlation and the significance of it is shown using one to three asterisks whose meaning is: *($p < 0.05$), **($p < 0.01$), or ***($p < 0.001$)

be seen that several indicators are more significantly used by the highest centrality users: for example, power, anger, negative emotions, and long words. Also, the tone of the high central users tends to be more negative when they are talking about Conspiracy topics. Moreover, religion appears to be significantly related to Neutral topics, along with racial expressions, which are also directly related to this centrality when talking about Politics.

4.3 Closeness

Closeness centrality measures the distance from one node to the rest of the nodes of the network. In the methodology section, the high scorers on this centrality were presented as *disseminators*, as they can reach all the accounts on the network more easily.

Analyzing the correlation results shown in Fig. 5, the use of power words seems to be positively related to closeness centrality, both when looking at the original tweets and the retweets that a user receives. Concerning the original tweets of a user, third-person singular pronouns and references to minorities (JewOrBlack) were also positively correlated with closeness. In the case of the retweets that a user receives, references to the white race and minorities are also positively related to this centrality measure. Regarding the tweets done by the user, the only dictionary related to closeness seems to be third-person plural pronouns, with a negative relationship. No significant relationship was found between the whole text from an account's timeline and this centrality.

Dividing the analysis by topic, there is no relationship among closeness centrality and the dictionaries when talking Neutral or Conspiracy content. Politics only shows a positive relationship between other racial terms and closeness, while Media shows a positive relationship with first person plural. Again, Hate Speech shows the biggest amount of significant relationships, with high-scoring users using more power, anger, negative emotions and longer words (sixltr).

4.4 Betweenness

The users with higher betweenness are critical for the information flow on the network. If they disappear, the difficulty for the messages to reach other nodes increases.

Figure 6 shows the relationship between betweenness centrality and discourse. Analyzing the retweets made by the high-scoring users, the most prominent pattern found is that they use more words related to death, first-person plural and third-person singular pronouns, and words related to the far-right discourse (white race, minorities references, racial slang, and other racial terms). Besides, religion and minorities references seem to be directly correlated with betweenness on the retweets that a user receives. On the original tweets, religion, references to the white race, and racial slang also seem to be directly correlated with the centrality. Finally, taking together all the text on the timeline of an account, negative emotion words are negatively related to this centrality, while racial slang remains positively related.

When approaching the relationship between language and centrality depending on the topic, it can be seen that few relationships are found now, even in the Hate Speech corpus. In that section, only references to minorities are significantly related to the centrality. Moreover, in the Neutral topic, it can be found a direct relationship between this centrality, religion, and other racial terms. The religion dictionary also shows a significant correlation on the Politics topic.

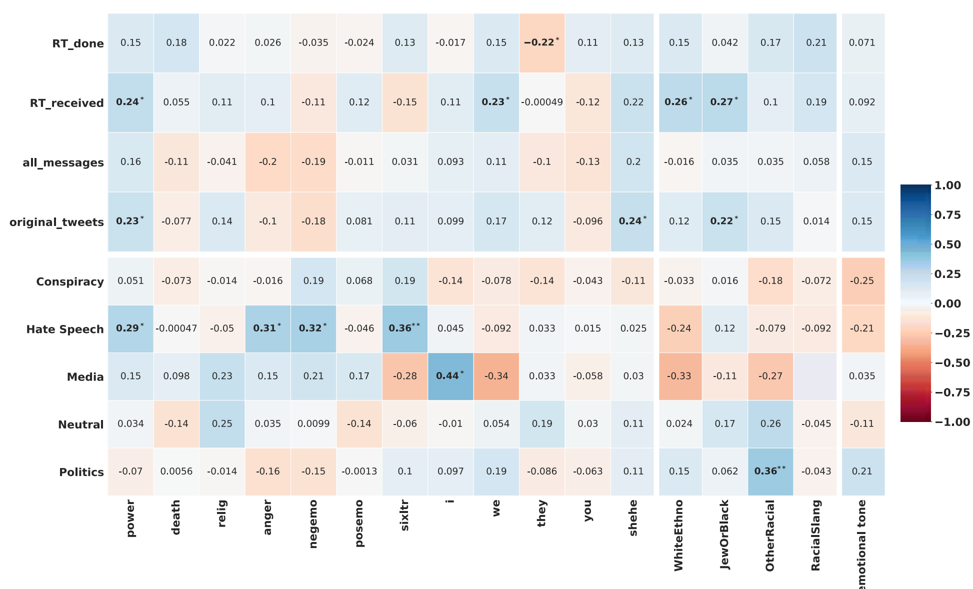

Fig. 5 Correlation heatmap for closeness centrality. Positive correlations are shown in blue. Contrarily, negative correlations are shown in red. Finally, the color intensity of a cell indicates the strength of a correlation and the significance of it is shown using one to three asterisks whose meaning is: *($p < 0.05$), **($p < 0.01$), or ***($p < 0.001$)

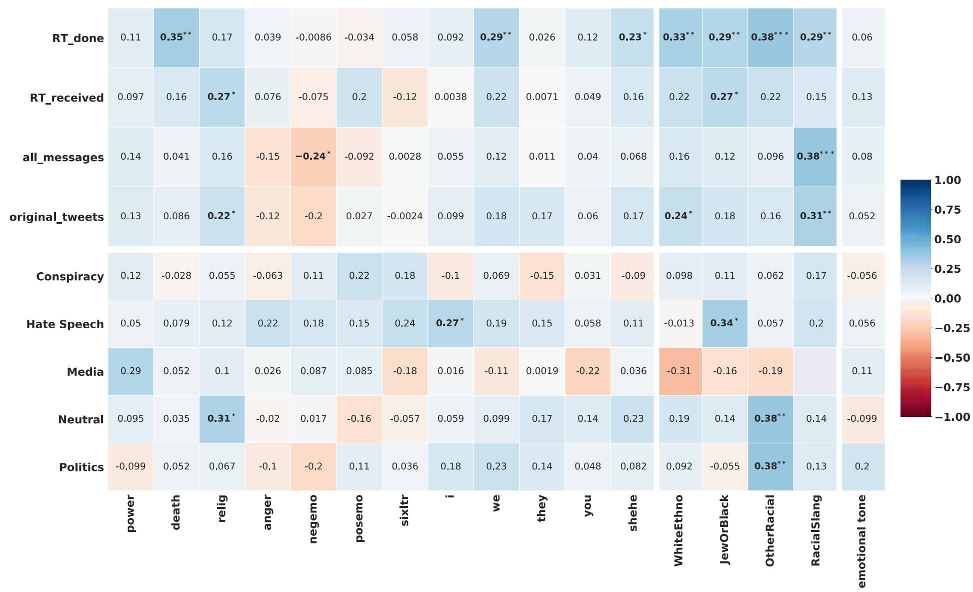

Fig. 6 Correlation heatmap for betweenness centrality. Positive correlations are shown in blue. Contrarily, negative correlations are shown in red. Finally, the color intensity of a cell indicates the strength of a correlation and the significance of it is shown using one to three asterisks whose meaning is: *($p < 0.05$), **($p < 0.01$), or ***($p < 0.001$)

However, no relationship was found on the Conspiracy and Media corpora.

4.5 K-shell

Users high-scoring on the K-shell centrality are more capable of spreading information on the network. This measure came from studies of diffusion models, and it is better than the degree centrality in networks with core-like structure. Figure 7 includes all the linguistic variables on the different corpora and their relationship with k-shell centrality.

For most of the text corpus, white race references, minorities mentions, racial terms, and racial slangs are correlated with k-shell centrality. On the retweets done by the user, death, religion, first-person plural, third-person singular, and second-person pronouns are also directly related to the centrality. Besides, first-person plural pronouns are also correlated with this centrality on the original tweets from a user. Finally, concerning the retweeted that a user receives, religion words are positively related to k-shell centrality.

Concerning the different topics, k-shell centrality on Hate Speech content seems to be directly related to anger, negative emotion words, and minorities related terms. When addressing Political content, only a few racial terms are significantly related to this centrality. First-person plural pronouns are negatively related to this centrality when

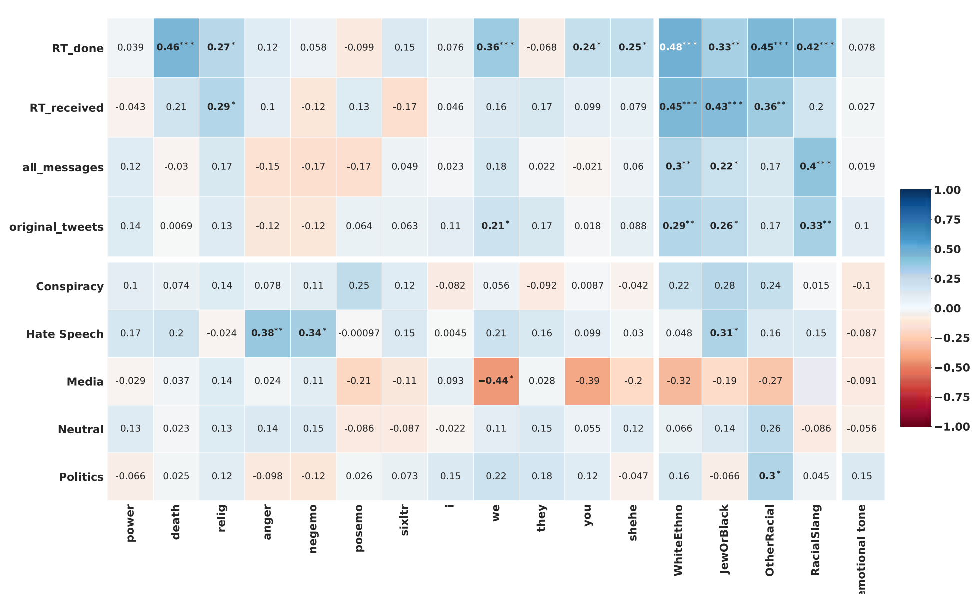

Fig. 7 Correlation heatmap for k-shell centrality. Positive correlations are shown in blue. Contrarily, negative correlations are shown in red. Finally, the color intensity of a cell indicates the strength of a correlation and the significance of it is shown using one to three asterisks whose meaning is: *($p < 0.05$), **($p <$ 0.01), or ***($p < 0.001$)

talking about Media. Finally, no relationship was found between this centrality and Conspiracy or Neutral topics.

4.6 Centrality measures and hashtag use

After analyzing the outcomes from the different centralities and noticing the similarities between k-shell centrality and betweenness, it was decided to make a final analysis concerning centrality and hashtag use. To do so, firstly the use of hashtags from the top-5 scoring users on each centrality was calculated and compared, to determine if there was a radical difference in the number of hashtags used in each top. Table 3 shows that the use of hashtags was relatively proportional among all the centrality measures. After this, a final heatmap correlating centrality and hashtag use was calculated. Figure 8 shows that only k-shell and betweenness centrality seem to be correlated with the use of hashtags from different topics, with a small exception on the in-degree centrality and the Neutral hashtags. This means that relevant super-spreaders and bridges tend to use significantly more topics than other kinds of relevant users.

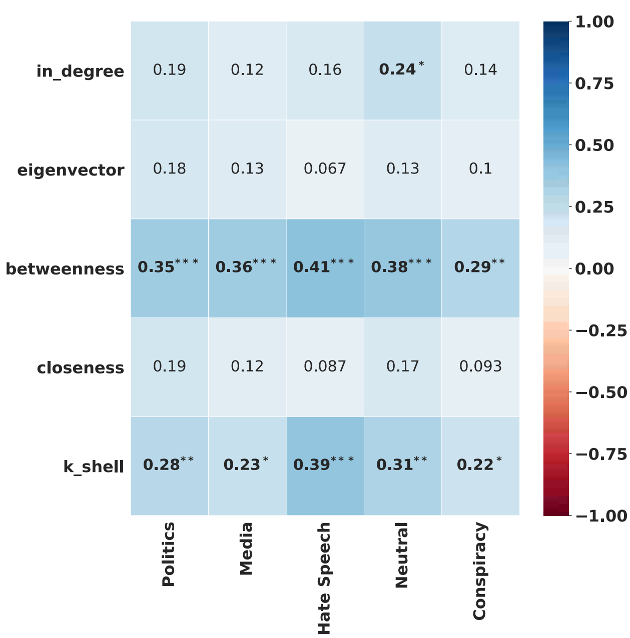

Fig. 8 Correlation heatmap between the different centralities and hashtag use. Positive correlations are shown in blue. Contrarily, negative correlations are shown in red. Finally, the color intensity of a cell indicates the strength of a correlation and the significance of it is shown using one to three asterisks whose meaning is: *($p < 0.05$), **($p < 0.01$), or ***($p < 0.001$)

Table 3 Condensed number of topics used for the Top 5 scoring users in each centrality

Topic	In_degree	Eigenvector	Betweenness	Closeness	K_shell
Politics	283	565	385	287	141
Media	33	16	45	34	48
Hate Speech	92	100	107	92	29
Neutral	33	26	38	33	15
Conspiracy	7	7	11	7	11
Total	448	714	586	453	244

5 Discussion

The present article aimed to check the relationship between a user's relevance (centrality) and its extremist discourse in an alt-right social network on Twitter. In the Introduction section, three different questions were raised concerning this objective. After comparing the language used on the different text corpus and topics with the five centrality measures applied, these questions can now be answered.

RQ1 How is different relevance criteria related to discourse?

To answer this question, it shall be determined if there is a correlation between linguistic variables and the different centrality measures (relevance). To do so, we analyzed together all the text related to the user, including original tweets and retweets. Overall, though the relationships between both variables are generally weak (between 0.2 and 0.3) and occasionally moderate (with a score of ≈ 0.4), *it can be confirmed that user relevance on the network is related to the linguistic patterns of the extremist discourse.*

Table 4 Summary of the correlations found between the centralities and all the linguistic variables of all the messages published by an account

	All messages
In-degree	∇*anger** ∇*negemo**
Eigenvector	Δ***power***** Δ*shehe**
Closeness	
Betweenness	∇*negemo** Δ***RacialSlang******
K-shells	Δ***WhiteEthno***** Δ*JewOrBlack** Δ***RacialSlang******

Rows show the centralities and cells the linguistic variables that have shown a significant correlation. Also, a ∇ indicates a negative correlation and a Δ a positive one. Besides, the significance of it is shown using asterisks: *($p < 0.05$), **($p < 0.01$), or ***($p < 0.001$). Finally, moderate correlations are shown in bold while the weak ones are not

This relationship is, however, dependent on how each centrality index is calculated. As we can see in Table 4, there are no common patterns among all the centrality measures. Users with high scores on *k-shells* show moderate correlations with far-right speech, especially on the dictionaries related to racism, nationalism, and minorities (variables WhiteEthno, JewOrBlack, and Racial Slang). This means that users who are, together with their neighbors, on the core on the network composition (high k-shells), and that are more capable of "spreading" efficiently their messages on different layers of the network, use significantly more extremist terms.

Users with high scores on *betweenness* also have a higher use of pejorative terms against other races, but do not use significantly more words related to white supremacism or directly mention other races. Therefore, users that facilitate the connection on the network use more pejorative terms directed to races, but do not appear as racist as the users with high k-shells. Also, high betweenness users do share a pattern with high *in-degree* users: they tend to use less words related to negative emotions. The latter also use less anger words, which means that users with more direct links on the network are more positive on their messages.

Users in close contact with other relevant nodes on the network (high *eigenvector*) show more use of third person singular pronouns, and more words related to power. The first finding can be related to the different hashtags used by the users, such as #HateHoax and #WomenForTrump, that mentioned women. The former is related to a campaign against a woman on Twitter, and the latter is related to a pro-Trump campaign. Therefore, it is difficult to clearly identify the reasons behind this pattern. However, as power words have been theorized to be related to a more aggressive discourse Cohen et al. (2016), it can be hypothesized that users related to other relevant nodes of the network use a more gender-oriented and aggressive discourse.

Finally, high *closeness* actors do not seem any significant pattern when compared to low scorers, which means that the proximity of the actor regarding the whole network do not seem to be related to the extremist indicators of their discourse.

RQ2 Does the discussed topic affect the relationship between relevance and discourse?

Concerning the dictionaries related to the alt-right discourse (Table 5), it is worth highlighting that all centrality measures correlate positively with the dictionary "otherRacial" when talking about *Politics*. This dictionary includes references to different cultures (Asian and Muslim/Arabic) as well as racism and ethnics. This pattern can be understood through two different insights. First, as Donald Trump's policies mention both Muslims and Asians due to the conflict in the Arabic countries and the economic conflict with China. Second, alt-right sympathizers share some behavioral patterns with white supremacists groups, who use different hashtags and comments related to race, inverse racism, and the victimization of white race (Thorburn et al. 2018).

Regarding *Hate Speech*, it can be seen that super-spreaders and bridges use more references to jews and blacks. This trend could also be understood using previous insight about behavioral patterns in white supremacists groups. However, it is interesting to see that all these centralities do not correlate with far-right specific dictionaries on *Hate Speech* discussions. This means that, while they use a more aggressive vocabulary, they do not necessarily use linguistic extremist terms. High in-degree, eigenvector, and closeness users share a common pattern: when discussing *Hate Speech* topics, there exists a direct relationship between these centralities and the use of long words (six letters or more), words related to power, words related to anger, and words related to negative emotions. This makes sense since long words (six letters or more) are used to create psychological distance with other groups (Pennebaker 2011), and words related to power, anger, and negative emotions imply an aggressive and dominant discourse (which is also present on the high k-shells users). This trend matches the *Hate Speech* topic because it includes hashtags referring to immigration, racial topics, and white nationalism, which also implies an aggressive, distant, and dominant discourse. It is worth mention that these patterns have already been found in other extremist groups (Vergani and Bliuc 2015). The only pattern which appears contradictory with previous literature is the higher use of first person singular pronouns by the users who facilitate the connectivity of the network, as the first-person plural pronouns are more commonly used to create an identity on extremist groups (Grover and Mark 2019).

Both *Neutral* and *Media* topics show little patterns, and they are not common among the different centralities. When referring to *Neutral* topics, only one patterns can be identified: both "bridges" and "authorities" present a positive correlation with religion. As stated before, it could be related to the term "God bless America," which is used as part of campaign slogans. Finally, while influencers and disseminators use more first-person singular pronouns when talking about media, super-spreaders have the opposite pattern: they use less first-person plural pronouns .

Finally, no pattern but one was found on the *Conspiracy* topic: a significantly more negative tone was used by users close to other relevant users on their messages. Concerning the emotional tone, an increase in the use of negative words has been reported on the literature when analyzing extremist political groups, including far-right communities (Alizadeh et al. 2019). However, this patterns does not seem to appear on this network.

RQ3 Does the corpus of analyzed text affect the relationship between relevance and discourse?

Table 5 Summary of the correlations found between the centralities and all the linguistic variables of the messages published by an account about the different topics

	Conspiracy	Hate Speech	Media	Neutral	Politics
In-degree		Δ*power** Δ*anger** **Δ*negemo**** **Δ*sixltr****	**Δ*i****		**Δ*OtrRacial*****
Eigenvector	**∇*emo.tone****	**Δ*power**** Δ*anger** **Δ*negemo**** Δ*sixltr**		**Δ*relig****	**Δ*OtrRacial*****
Closeness		Δ*power** **Δ*anger**** **Δ*negemo**** **Δ*sixltr*****	**Δ*i****		**Δ*OtrRacial*****
Betweenness		Δ*i** **Δ*JewOrBlack****		**Δ*relig****	**Δ*OtrRacial*****
K-shells		**Δ*anger***** **Δ*negemo**** **Δ*JewOrBlack****	**∇*we****		**Δ*OtrRacial****

Columns show the topics, rows the centralities and cells the linguistic variables that have shown a significant correlation. Also, a ∇ indicates a negative correlation and a Δ a positive one. Besides, the significance of it is shown using asterisks: *($p < 0.05$), **($p < 0.01$), or ***($p < 0.001$). Finally, moderate correlations are shown in bold, while the weak ones are not

Table 6 Summary of the correlations found between the centralities and all the linguistic variables of the different type of messages published by an account

	RT_done	RT_received	Original_tweets
In-degree	Δ*death** Δ*we** Δ*RacialSlang**	Δ*we** **Δ*WhiteEthno****** **Δ*JewOrBlack*****	∇*negemo**
Eigenvector	Δ*power**	**Δ*power***** Δ*we** Δ*shehe**	**Δ*power**** Δ*we** Δ*shehe***
Closeness	∇*they**	Δ*power** Δ*we** Δ*WhiteEthno** Δ*OtrRacial**	Δ*power** Δ*shehe** Δ*JewOrBlack**
Betweenness	**Δ*death***** Δ*we*** Δ*shehe** **Δ*WhiteEthno***** Δ*JewOrBlack** **Δ*OtrRacial****** Δ*RacialSlang***	Δ*relig** Δ *JewOrBlack**	Δ*relig** Δ*WhiteEthno** **Δ*RacialSlang*****
K-shells	**Δ*death****** Δ*relig** **Δ*we****** Δ*you** Δ*shehe** **Δ*WhiteEthno****** **Δ*JewOrBlack***** **Δ*OtrRacial****** **Δ*RacialSlang******	Δ*relig** **Δ*WhiteEthno****** **Δ*JewOrBlack***** **Δ*OtrRacial*****	Δ*we** Δ*WhiteEthno*** Δ*JewOrBlack** **Δ*RacialSlang*****

Columns show the types of messages, rows the centralities and the cells the linguistic variables that have shown a significant correlation. Also, a ∇ indicates a negative correlation and a Δa positive one. Besides, the significance of it is shown using asterisks: *($p < 0.05$), **($p < 0.01$), or ***($p < 0.001$). Finally, moderate correlations are shown in bold, while the weak ones are not

When dividing all the messages from the users on different corpus of text, it can be seen that several underlying patterns appear (Table 6).

Starting with the *retweets done by the users*, it can be seen that the more outstanding pattern is that super-spreaders speech (high score on k-shell centrality) presents significantly more words related to all the far-right dictionaries, but also with death, religion and different pronouns, specially the first person plural. A similar but weaker pattern can be found on high betweenness users. The rest of the high-scoring users on in-degree, eigenvector and closeness do not present patterns as aggressive as the former's in nearly all the text corpus. Therefore, it can be seen that super-spreaders and users who are more important to facilitate the information flow on the network tend to retweet more aggressive and extremist content than the users who have more direct retweets, are in contact with other relevant users, or take less time to reach the rest of the users.

Regarding the *retweets received by the users*, the patterns found for betweenness and k-shells seem to moderate, but the contrary happens with in-degree, eigenvector and closeness. Users with more retweets tend to be retweeted when their messages include references to white nationalism and racism, along with more first person plural words. A similar pattern is found on users who are closer to the rest of the network (high closeness), but also including more power words, which shall be remembered that are related to a more aggressive language). Users closer to other relevant users also present more first person plural and power words, but the content other users retweet from them do not seem to include significantly more references to nationalism, racism or pejorative terms against other minorities. In the case of retweets made from high betweenness users, religion and racist terms are significantly more common. Retweets made from high k-shells users still present references to white nationalism and other minorities, such as Asian people.

Finally, when focusing on the *original tweets written by the users*, it can be seen that the pattern is relatively similar to the ones found on the retweets received. High betweenness and k-shell users tend to publish original content including racist and supremacist references, with the latter including more first person plural pronouns as an indicative of an intent to create a group identity. High closeness users also present references on their original content to jews and blacks, but also references to third person singular pronouns and power terms. Interestingly, the first person singular pronouns and the powers words are also significantly common on the high eigenvector centrality users. The only distinctive pattern can be found on high in-degree users, which present less negative emotion words on their original content.

As a summary of the previous findings, it can be seen that a deeper analysis of the different corpus shows that the retweets done by super-spreaders and users acting like "bridges" for the network tend to be more much more aggressive than the message they originally create. Also, the presence of racist, aggressive, religious and group-conforming terms tend to be common on the retweets received by users with a high relevance on the network. High eigenvector and closeness users tend to retweet less aggressive tweets that the one they publish. But the retweets received and done by the most directly retweeted users (high in-degree) also include more racist content than their original tweets.

6 Conclusions and future work

This research aimed to expand the literature about the extremist discourse on the far right. In this case, focusing on a Twitter community and its organization, this article represents one of the few approaches to study the organization among these groups and their links with the psychological use of words. The outcomes of this research also allowed to confirm that the different relevance criteria (or centrality measures) have weak to moderate relationships with the extremist discourse of the users of an alt-right network on Twitter. These relationships varied depending on the topic discussed and the corpus of text screened. But the main discoveries were that high k-shells users tend to have a more racist discourse in all the conditions, hate speech topic is the one where the most aggressive language is used by all the high relevant users (no matter the criteria), and that retweets done by high betweenness and k-shells users tend to include significantly more aggressive, racist and supremacist content. Finally, the retweets received by high relevant users (no matter the criteria) also include more aggressive, racist, supremacist and group-directed language.

Future work should continue expanding this topic, and focusing on the different characteristics of the extremist and non-extremist use of language, along with its impact on the interactions among users on a social network, in both Twitter and other social contexts. Besides, the methodology presented should be improved to address some of its limitations. A first extension could be enhancing the text categorization phase, notice that only $\approx 8\%$ of the available tweets could be analyzed using the proposed approach. The use of hashtags for categorizing the tweets into topics, despite being accurate, misses most of the data available in the dataset. Hence, it would be convenient to use more sophisticated techniques that are capable of exploiting more parts of the dataset. A second improvement could be expanding the relationship analysis between the different centralities and the speech. While the present article studies each linguistic variable independently, interactions between them may exist. Therefore, expanding the analysis to consider them would be highly interesting.

Acknowledgements This work has been co-funded by the following grants: Comunidad Autonoma de Madrid under grant S2018/TCS-4566 (CYNAMON: Cybersecurity, Network Analysis and Monitoring for the Next Generation Internet); Spanish Ministry of Science and Education and Competitivity (MINECO) and European Regional Development Fund (FEDER) under grant TIN2017-85727-C4-3-P (DeepBio).

References

Akoglu H (2018) User's guide to correlation coefficients. Turk J Emerg Med 18(3):91–93
Alizadeh M, Weber I, Cioffi-Revilla C, Fortunato S, Macy M (2019) Psychology and morality of political extremists: evidence from twitter language analysis of alt-right and antifa. EPJ Data Sci 8(1):17
Bello-Orgaz G, Jung JJ, Camacho D (2016) Social big data: recent achievements and new challenges. Inf Fusion 28:45–59
Bonacich P (1987) Power and centrality: a family of measures. Am J Sociol 92(5):1170–1182
Camacho D, Panizo-LLedot A, Bello-Orgaz G, Gonzalez-Pardo A, Cambria E (2020) The four dimensions of social network analysis: an overview of research methods, applications, and software tools. Inf Fusion 28:45
Cha M, Haddadi H, Benevenuto F, Gummadi KP (2010) Measuring user influence in twitter: the million follower fallacy. In: fourth international AAAI conference on weblogs and social media
Chen J, Hossain MS, Zhang H (2020) Analyzing the sentiment correlation between regular tweets and retweets. Social Netw Anal Min 10(1):13
Cohen K, Kaati L, Shrestha A (2016) Linguistic analysis of lone offender manifestos. In: International conference on CyberCrime and computer forensics (ICCCF)
Conover MD, 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

Das K, Samanta S, Pal M (2018) Study on centrality measures in social networks: a survey. Soc Netw Anal Min 8(1):13
Davidson I, Gourru A, Velcin J, Wu Y (2020) Behavioral differences: insights, explanations and comparisons of french and us twitter usage during elections. Soc Netw Analy Min 10(1):6
Deturck K, Patel N, Avouac PA, Lopez C, Nouvel D, Partalas I, Segond F (2017) Detecting influencial users in social networks: analysing graph-based and linguistic perspectives. In: IFIP international workshop on artificial intelligence for knowledge management, pp. 113–131. Springer
Figea L, Kaati L, Scrivens R (2016) Measuring online affects in a white supremacy forum. In: 2016 IEEE conference on intelligence and security informatics (ISI), pp. 85–90. IEEE
Forscher PS, Kteily N (2017) A psychological profile of the alt-right
Freeman LC (1978) Centrality in social networks conceptual clarification. Soc Netw 1(3):215–239
Graham R (2016) Inter-ideological mingling: White extremist ideology entering the mainstream on twitter. Sociol Spect 36(1):24–36
Greven T (2016) The rise of right-wing populism in Europe and the united states. A Comparative Perspective. Friedrich Ebert Foundation, Washington DC Office
Grover T, Mark G (2019) Detecting potential warning behaviors of ideological radicalization in an alt-right subreddit. Proc Int AAAI Conf Web Soc Media 13:193–204
Hutto CJ, Gilbert E (2014) Vader: A parsimonious rule-based model for sentiment analysis of social media text. In: Eighth international AAAI conference on weblogs and social media
Kaati L, Shrestha A, Cohen K, Lindquist S (2016) Automatic detection of xenophobic narratives: a case study on Swedish alternative media. In: 2016 IEEE conference on intelligence and security informatics (ISI), pp. 121–126. IEEE
Kitsak M, Gallos LK, Havlin S, Liljeros F, Muchnik L, Stanley HE, Makse HA (2010) Identification of influential spreaders in complex networks. Nat Phys 6(11):888–893
Kreis R (2017) # refugeesnotwelcome: anti-refugee discourse on twitter. Discour Commun 11(5):498–514
Lyons MN (2017) Ctrl-alt-delete: The Origins and Ideology of the Alternative Right. Political Research Associates, Somerville, MA
Metaxas P, Mustafaraj E, Wong K, Zeng L, O'Keefe M, Finn S (2015) What do retweets indicate? results from user survey and meta-review of research. In: Ninth international AAAI conference on web and social media
Mirrlees T (2018) The alt-right's discourse on "cultural marxism": a political instrument of intersectional hate. Atlant Criti Stud Gender Cult Soc Justice 39(1):49–69
Nguyen HL, Jung JJ (2019) Social event decomposition for constructing knowledge graph. Future Gener Comput Syst 100:10–18
Panizo-LLedot A, Torregrosa J, Bello-Orgaz G, Thorburn J, Camacho D (2019) Describing alt-right communities and their discourse on twitter during the 2018 us mid-term elections. In: International conference on complex networks and their applications, pp. 427–439. Springer
Pennebaker JW (2011) The secret life of pronouns. New Sci 211(2828):42–45
Pennebaker JW, Boyd RL, Jordan K, Blackburn K (2015) The development and psychometric properties of liwc2015. Technical report
Riquelme F, González-Cantergiani P (2016) Measuring user influence on twitter: a survey. Inf Process Manag 52(5):949–975
Romero DM, Meeder B, Kleinberg J (2011) Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In: Proceedings of the 20th international conference on World wide web, pp. 695–704
Stieglitz S, Dang-Xuan L (2012) Political communication and influence through microblogging–an empirical analysis of sentiment in twitter messages and retweet behavior. In: 2012 45th Hawaii international conference on system sciences, pp. 3500–3509. IEEE
Tan LKW, Na JC, Theng YL (2011) Influence detection between blog posts through blog features, content analysis, and community identity. Online Inf Rev 2011:1
Thorburn J, Torregrosa J, Panizo Á (2018) Measuring extremism: Validating an alt-right twitter accounts dataset. In: International conference on intelligent data engineering and automated learning, pp. 9–14. Springer
Torregrosa J, Thorburn J, Lara-Cabrera R, Camacho D, Trujillo HM (2019) Linguistic analysis of pro-isis users on twitter. Behav Sci Terrorism Politi Aggress 2019:1–15
Tumasjan A, Sprenger TO, Sandner PG, Welpe IM (2010) Predicting elections with twitter: What 140 characters reveal about political sentiment. In: Fourth international AAAI conference on weblogs and social media
van der Vegt I, Mozes M, Gill P, Kleinberg B (2019) Online influence, offline violence: Linguistic responses to the'unite the right'rally. arXiv preprint arXiv:1908.11599
Vergani M, Bliuc AM (2015) The evolution of the isis'language: a quantitative analysis of the language of the first year of Dabiq magazine. Sicur Terror soc 2:7–20
Yardi S, Boyd D (2010) Dynamic debates: an analysis of group polarization over time on twitter. Bull Sci Technol Soc 30(5):316–327
Zannettou S, Bradlyn B, De Cristofaro E, Kwak H, Sirivianos M, Stringini G, Blackburn J (2018) What is gab: a bastion of free speech or an alt-right echo chamber. Compan Proc Web Conf 2018:1007–1014
Zhang Y, Wells C, Wang S, Rohe K (2018) Attention and amplification in the hybrid media system: the composition and activity of Donald Trump's twitter following during the 2016 presidential election. New Media Soc 20(9):3161–3182