Characterizing the Twitter Network of Prominent Politicians and SPLC-defined Hate Groups in the 2016 US Presidential Election

Raazesh Sainudiin · Kumar Yogeeswaran · Kyle Nash · Rania Sahioun

Preprint accepted for publication at Social Network Analysis and Mining, 2019.
The Final Publication is available at Springer via `https://link.springer.com/journal/13278`

Abstract We characterize the Twitter networks of the major presidential candidates, Donald J. Trump and Hillary R. Clinton, with various American hate groups defined by the US Southern Poverty Law Center (SPLC). We further examined the Twitter networks for Bernie Sanders, Ted Cruz, and Paul Ryan, for 9 weeks around the 2016 election (4 weeks prior to the election and 4 weeks post-election). We carefully account for the observed heterogeneity in the Twitter activity levels across individuals through the null hypothesis of apathetic retweeting that is formalized as a random network model based on the directed, multi-edged, self-looped, configuration model. Our data revealed via a generalized Fisher's exact test that there were significantly many Twitter accounts linked to SPLC-defined hate groups belonging to 7 ideologies (Anti-Government, Anti-Immigrant, Anti-LGBT, Anti-Muslim, Alt-Right, White-Nationalist and Neo-Nazi) and also to @realDonaldTrump relative to the accounts of the other four politicians. The exact hypothesis test uses Apache Spark's distributed sort and join algorithms to produce independent samples in a fully scalable way from the null model. Additionally, by exploring the empirical Twitter network we found that significantly more individuals had the fewest retweet degrees of separation simultaneously from Trump and each one of these seven hateful ideologies relative to the other four politicians. We conduct this exploration via a geometric model of the observed retweet network, distributed vertex programs in Spark's GraphX library and a visual summary through neighbor-joined population retweet ideological trees. Remarkably, less than 5% of individuals had three or fewer retweet degrees of separation simultaneously from Trump and one of several hateful ideologies relative to the other four politicians. Taken together, these findings suggest that Trump may have indeed possessed unique appeal to individuals drawn to hateful ideologies, however such individuals constituted a small fraction of the sampled population.

Keywords Donald Trump · Twitter · 2016 US Presidential election · US hate groups · configuration model · scalable generalized Fisher's exact test · Apache Spark · directed degrees of separation · empirical geometric retweet model · population retweet ideological trees

R.Sai. was partly supported by the chaire Modélisation Mathématique et Biodiversité of Veolia Environnement–École Polytechnique–Museum National d'Histoire Naturelle–Fondation X. Distributed cloud computing of Twitter data by R.Sai. was supported by Databricks Academic Partners Program.

R. Sainudiin
Department of Mathematics,
Combient Competence Centre for Data Engineering Sciences,
Uppsala University, Uppsala Sweden
E-mail: raazesh.sainudiin@math.uu.se

K. Yogeeswaran
Department of Psychology,
University of Canterbury, Christchurch, New Zealand
E-mail: kumar.yogeeswaran@canterbury.ac.nz

K. Nash
Department of Psychology,
University of Alberta, Edmonton, Canada
E-mail: knash@ualberta.ca

R. Sahioun
Department of Psychology,
University of Canterbury, Christchurch, New Zealand
E-mail: rouna79@yahoo.com

1 Introduction

The 2016 US presidential election will be remembered as one of the most divisive in recent history with two of the least liked candidates competing for the White House [26,1]. During the election, the mainstream media appeared baffled by the rise of Donald J. Trump, a businessman with no prior experience in government. Two narratives emerged to explain Trump's rise to fame. On the one hand, it was argued that Trump's rise was driven by uncertainty and angst in the American public [60,63,39] as well as frustration and distrust of a partisan political system [59,65,8]. This first narrative broadly reflects a number of psychological theories that commonly demonstrate that uncertainty and angst can catalyze the formation of rigid, nationalistic, and extreme political beliefs; these are supported by research on right-wing authoritarianism [4], uncertainty-identity theory [40], need for closure [85], and motivated social cognition [46]. By contrast, a second narrative argued that Donald Trump's rise was driven by a growing sense of threat among White Americans regarding their changing status in America as a result of growing cultural diversity and the improving status of minority groups made salient by the election of Barack Obama. Such changes were argued to promote minority scapegoating and a campaign that pitted 'us' versus 'them' [85,46–48,10,43,5]. Along with uncertainty-identity theory [40], this second narrative is reflected in psychological research on relative deprivation theory [66,83], intergroup threat theory [76], and recent evidence that ethnic diversity can threaten the status, identity, and distinctiveness of majority group members [51,18,88].

Though both sets of factors may have collectively played a role in Trump's rise to power, there has been considerable debate within the media and around dinner tables about the latter narrative that Trump was able to capitalize on White Americans' fears about their changing status and social identity in an unstable national landscape. Critics of Donald Trump point toward Trump's own rhetoric on the campaign trail, which often featured insults leveled against Mexicans, Muslims, immigrants, and women [37,74]. Such rhetoric, it was argued, took advantage of the increased appetite for nationalism and outgroup derogation due to perceived threats to status and identity for the majority group. By doing so, Trump was argued to have forged a coalition where normative bigotry was thought to be acceptable [19]. As a consequence, Trump's candidacy was openly supported by the Ku Klux Klan (KKK) and Neo-Nazi groups [64,62]. Though Trump's campaign was careful to distance itself from the endorsement of such extreme hate groups [21], critics point to the rise of hate crimes since the election [38,61,16,56] and the Trump administration's decision to remove government focus on right-wing extremism as evidence of Trump's ties to hate groups [58,62]. Even months after the election, people continued to debate the role of bigotry in Trump's candidacy [86,80,52].

Twitter has become a major platform for communication between politicians and their followers. Consequently, Twitter activity has been used to gauge political sentiments, predict elections [79,82] and reportedly influence the 2016 election [14,34,87] [1]. President Trump uses the platform to directly communicate with the American public [73,25]. Twitter has also become a major avenue through which hate groups spread their ideologies with larger sections of the population [36]. For all these reasons, Twitter was chosen as the ideal medium to observe naturalistic relationships between various individuals and groups.

In the present work, we examined the Twitter networks of several key politicians and hate groups or their leadership during the 2016 US Presidential election to shed light on two major narratives regarding Trump's rise to fame. As most academic studies in the Twitterverse we were limited by the freely available Twitter data. So our study of a specifically sampled population, despite our attempts at optimizing the sampling design, may not be representative of the US Twitter population. This caveat should be borne in mind when generalizing this study without the full data that can be purchased from Twitter.

The political leaders included in our study were the two major presidential candidates (i.e., Hillary Clinton with mentionable screen-name as @HillaryClinton and Donald Trump as @realDonaldTrump). Additionally, we chose to include Ted Cruz (as @tedcruz) and Bernie Sanders (as @BernieSanders) who were the second most popular Republican and Democratic presidential candidates in the 2016 election. And finally, we chose to include Paul Ryan (as @SpeakerRyan) as he was the Speaker of the House of Representatives at the time of the election. These individuals all possess active Twitter accounts with more than 2 million followers and engage in frequent Twitter activity. We collected publicly streaming data from these five politicians' official Twitter accounts.

To examine the links between these five US politicians and hateful ideologies, we used the Southern Poverty Law Center's (SPLC's) definition and ideological classification of hate groups and hate group lead-

[1] This article is the subject of separate legal complaints on behalf of Cambridge Analytica LLC and SCL Elections Limited, and Sophie Schmidt.

ers; hereafter referred to as hate groups. Although the SPLC has been criticized for including groups and individuals with politically diverging views from their own [29, 78, 45, 33], it is the most comprehensive database for hate groups in the USA that is available to the public with records over the last 40 years [31, 17, 68, 30]. Moreover, our mathematical models, statistical tests, and the data science pipelines are generic and can therefore be used to examine Twitter networks between political leaders and any database of hate groups as long as these have enough activity on Twitter. More generally, our approach formally only requires a population of users, and two sub-populations of interest.

The SPLC does not consider all groups or individuals on its list as violent or engaged in criminal activities, but rather identifies any group or individual "whose beliefs or practices attack or malign an entire class of people, typically for their immutable characteristics." [15]. The database does not include foreign hate groups or extremist groups such as ISIS, Al Qaeda, or Boko Haram, as the focus of the database is on American hate groups. In the present work, we focused only on hate groups which the SPLC has specifically designated as promoting a particular hateful ideology in order to examine the appeal of major politicians for subscribers to that ideology. We collected publicly streaming data from 52 Twitter accounts belonging to eight hateful ideologies including: Neo-Nazi or NN (e.g., National Socialist Movement), Black-Separatist or BlS (e.g., Nation of Islam), Anti-LGBT or AL (e.g., Westboro Baptist Church), Anti-Immigrant or AI (e.g., American Border Patrol), Anti-Muslim or AM (e.g., ACT for America), Alt-Right or AR (e.g., American Renaissance), Anti-Government or AG (e.g., Oath Keepers), and White-Nationalist or WN (e.g., Aryan Brotherhood). As argued by others [32], studying blatant bigotry and hate is better suited using online communication as opposed to more traditional means of conducting prejudice research via self-report measures due to social norms against such blatant expressions. In line with this, the present work examines linkages between hate groups and politicians by focusing on social media communication on Twitter that offers such hate groups a shroud of anonymity.

We use the *retweet network*, which represents each retweet as a directed edge from the the tweeter, i.e., the user who posted the original tweet to the retweeter, i.e., the user who made the retweet, to model the communication links between users in our dataset. A retweet not only represents the strongest index of interest in the message, but also endorsement and trust in the communicator [44, 53]. Retweet analysis via random network models (as used in this study) also circumvents the ambiguity and further uncertainty associated with statistical algorithms in natural language understanding (NLU) that will be required when working with (a) quoted tweets, where one is allowed to add a comment to the retweet, (b) reply tweets, or (c) reply of quoted tweets, where one can reply in possible disagreement, etc. Thus, retweets are a simple and natural signal of directional concurrence, i.e., concurrence with the user who posted the (original) tweet. This is especially the case when one retweets another multiple times. Examining retweet networks would therefore be ideal as the research questions of interest focus on observing linkages between people endorsing blatantly hateful ideologies and endorsing specific political candidates during the 2016 Presidential election.

Using the observed retweet network we examine three specific questions: (Q1) Is Trump preferentially retweeted by hate groups when compared to other politicians against the null random network model of *apathetic retweeting*? (Q2) What frequency of unique users retweeted both a politician and a hate group more than one would expect under the null model? And (Q3) What is the joint distribution of the *degrees of separation*, measured through the lengths of the most retweeted directed paths in the observed Twitter network, to each user from each of the five politicians and the eight most prolific hateful ideologies on Twitter? The last question is of a data-exploratory nature, as such a joint distribution can give insights into certain projections of ideological profiles of politically active Twitter users involving *population ideological trees* that naturally extend typical projections involving a binary spectrum, where users are distinguished only up to points along an interval from one extreme to another.

Our analyses observed Twitter networks without consciously filtering out bot accounts. As bots may have played a role in influencing the 2016 election [87, 42], we simply observed the public Twitter streams associated with these politicians and hate groups. However, we do retrospectively study the effect of potential bot accounts with high entropy scores and the accounts identified by Twitter as being Russian trolls [81] as discussed later on.

Using exact tests, we address questions (Q1) and (Q2) by attempting to reject the null hypotheses of *apathetic retweeting* where each user is apathetic (i.e., indifferent with a lack of interest or concern[2]) about who they retweet while preserving the number of times they retweet others and the number of times others retweet them. We formalize this null model as the *apathetic retweet network* from the directed versions of the con-

[2] "Apathy." Merriam-Webster.com. Accessed July 10, 2018. https://www.merriam-webster.com/dictionary/apathy.

figuration network model [57,13] (see [3] and references therein for an overview of models for stochastic social dynamics). Thus, we can preserve the observed innate differences among users' tweet and retweet rates in this null model when we attempt to reject it in favor of its complementary alternative hypothesis of non-apathetic retweeting of one set of users by another set of users. By leveraging Apache Spark [89], a unified engine for big data processing, our scalable and fault-tolerant version of CutPermuteAndRewire, the original Monte Carlo sampling algorithm due to [57], can produce independent samples from this null model. These samples can in turn be used to obtain the null distribution with consistent estimates of the p-value of any test statistic of the apathetic retweet network, for instance. Thus we have an exact nonparametric test that can computationally scale for arbitrarily large directed networks. Finally, to answer (Q3) we use distributed vertex programs over the empirically estimated *geometric retweet network* model to obtain the degrees of separation from a politician and a hate group and visualize the pairwise distance in the degrees of separation between users through a neighbor-joined retweet ideological tree of the sampled population.

There is a large body of literature on online social media and hate. In general, such studies are either descriptive (see for e.g. [11] and the references therein and [55] for a recent example) or predictive (for e.g. hate speech detection [72,28]). Thus, the quantitative literature around social media and hate is more focused on mining and modeling patterns in the data for predictive and descriptive purposes, as opposed to testing specific hypotheses of interest to social scientists using custom-designed experiments. This work is a significant extension of [70] that primarily focused on testing specific hypotheses. Here we not only focus on classical statistical hypothesis testing with interpretable probabilistic models of the null hypothesis, but also on a systematic exploration of the data through scalable data science pipelines for conducting (designing, testing and exploring) one's own statistical experiments in Twitter by leveraging the latest advances in elementary distributed algorithms for sorting, joining and vertex programming.

2 Data, Design, Models and Methods

2.1 Data and Experimental Design

2.1.1 Communications in Twitter

Twitter, a popular micro-blogging service, provides an observable social network with millions of users [44]. Twitter allows users to communicate or update their status in many ways. One can post a message called a *tweet* that is no more than 140 characters in length (in 2016), *follow* another Twitter user, and receive the status updates of every user they follow. A tweet that is shared publicly with one's own followers is known as a *retweet*.

2.1.2 Dataset

Our dataset comprises of 21,749,868 communication events in Twitter over a total of 9 weeks centered around the US 2016 election. The tweet IDs are available from `https://tinyurl.com/y8lamxzx`. The dataset was collected in our designed experiment using Twitter's streaming and REST APIs [22] that were extended for Apache Spark [89]. Each of these communication events were parsed and classified into seven types of events using SparkSQL modules.
Streaming data collector: Over 17 million events were collected from Twitter's public streams by directly tracking communications related to the Twitter accounts of the five political candidates and 52 hate group accounts. Because only 78% of the SPLC identified hate groups had a valid Twitter account that we could track, our study is limited to a further subset of 52 user accounts who were active in the Twitter public streams. Thus, our approach is not exhaustive, in terms of being able to track every account of each hateful ideology, but is nonetheless representative of the public Twitter activity[3].
Retrospective data augmentation: The remaining 5 million events were obtained retrospectively using Twitter's REST API as follows. Due to most retweets being an immediate reaction to a tweet that one finds interesting or concurs with [49], nearly 7 million of the 10.5 million retweets in the 17 million events collected from the public streams happened within the same day of the original tweet and over 98% of the retweets happened within a week of the original tweet. Furthermore, since our tweet collector is only recording events that are directly related to one of our tracked users, we do not know of any other Twitter interactions by those who retweeted one of our tracked users. Thus, to better understand the recent retweet behavior of at least some

[3] Although the SPLC also identifies ideologies of the Ku Klux Klan (KKK), Neo-confederates, Racist Skinheads, and Holocaust Deniers, within their database, there was not enough activity on the Twitter accounts of these groups that we were able to track given just a few thousand followers. It may be that these groups have a smaller online presence, or they simply use a different social media platform (including closed discussion forums; see [52]) making it harder to follow as we focused our analyses only on hate groups with public streams.

of the politically active Twitter users with other non-tracked users on Twitter, we focus on October 19 2016, the day of the 3rd US Presidential debate, and obtain a seed set of users who retweeted either @HillaryClinton or @realDonaldTrump on this day. The communication intensity reached over 120 events per second around the debate in our streaming data collector. Our seed set is made up of a random sample of about a third of all users (including all verified and geo-enabled accounts) who retweeted either Clinton or Trump on October 19 2016 and thus constitutes an evenly represented sample of politically active Twitter users from the two parties. For each user in the seed set, at the end of the 9 week period, we added all the retweets from their 200 most recent status updates that occurred in the 9 week period. This strategy involves a breadth-first expansion about the seed set of users in the much larger retweet network on Twitter as it allows us to expand our 9-week-long retweet network by focusing on the recent retweet timelines of those who retweeted either of the two final presidential candidates during the day of the last debate. Crucially, this augmented data added another 0.3 million users to our network, increased the number of retweet events from 10.5 million to 13.7 million and made the retweet network into a single connected component.

2.2 Modeling Retweets

2.2.1 Retweet Network

Retweets are a simple and natural signal of *directional concurrence*, i.e. concurrence *with* the user who posted the (original) tweet *by* the user who retweeted it (i.e., between the tweeter and the retweeter), as they express interest in the message and also endorsement and trust in the communicator [44,53], especially when retweeted multiple times. Retweet analysis via random network models as done in this study circumvents the ambiguity and uncertainty associated with statistical algorithms [9] in natural language understanding (NLU) that will be required when working with (a) *quoted tweets* where one is allowed to add a comment to the retweet, (b) *reply tweets* or (c) *reply of quoted tweets* where one can reply in possible disagreement.

The number of retweets per day captured from the public stream during the 9 week period reached over 1 million during the day of the third US Presidential debate and the days leading up to the election. The proportion of retweets is known to vary greatly depending on the features of the Twitter subnetwork under study [49]. There were over 13.7 million retweets (63.1%) of an original tweet and over 2.7 million (12.6%) original tweets in our dataset. We ignore the remaining 25% of the events that require further NLU and focus our analysis instead on the retweet network obtained from 75% of all events in our dataset as explained in the next section.

We allowed the set of users who tweet and retweet, i.e., tweeters and retweeters, to form the nodes of the retweet network. Each retweet was allowed to represent a directed edge or arc from its tweeter to its retweeter, i.e. the tweeter-retweeter pair, as a signal of directional concurrence in the retweet network. Because one can retweet more than one tweet posted by any user including oneself, we allowed for parallel or multiple edges between the same pair of users including self-loops, i.e. edges from and to the same user. There were more than 4.4 million unique tweeter-retweeter pairs out of 2.5 million unique users representing over 16.4 million tweets and retweets in our dataset with over 21.7 million communication events during the 9-week period around the US presidential election. Mathematically, the retweets in our dataset are represented by a *directed multi-edged self-looped network*. Thus the *out-degree* (number of outgoing edges) and *in-degree* (number of incoming edges) for a user in the retweet network gives the number of times that the user is retweeted by others and the number of times the user retweets others (including oneself), respectively. Similarly, the number of distinct users who are retweeted by the user is given by their *in-nbhd* and the number of distinct retweeters of the user is given by their *out-nbhd*.

The observed retweet network is highly heterogeneous and largely dominated by the two presidential candidates as depicted in Table 1. Due to our Twitter collector's design with retrospective data-augmentation, our network forms a single connected component when viewed with undirected edges.

2.2.2 Apathetic Retweet Network Model

The *configuration model* for directed networks [57] is a random network model that produces samples uniformly from the set of networks that preserve the in-degree and out-degree of each node in given observed network.

By representing each retweet as a directed edge between the author of the tweet and its retweeter, we define the null apathetic model for a random retweet network as the *directed, multi-edged, self-looped configuration model*, whereby each user apathetically or non-preferentially retweets another user, by choosing uniformly at random from the set of all users, while preserving their observed in-degree and out-degree. However, the model frees up information on the users in-

Table 1 Retweet Network statistics of the five political accounts

Politician	in-degree	in-nbhd	out-degree	out-nbhd
Donald Trump	40	12	5,952,257	958,262
Hillary Clinton	225	121	2,774,111	943,995
Bernie Sanders	107	62	762,209	356,718
Paul Ryan	769	158	68,973	28,902
Ted Cruz	322	189	49,479	27,663

volved in retweeting, i.e. who retweets whom, provided the in-degree and out-degree of each user is identical to that of the observed network. This null model thus preserves the innate tweet and retweet rates of each user while modeling complete apathy or indifference for who one chooses to retweet.

2.2.3 Cut, Permute and Rewire

CutPermuteAndRewire is a Monte Carlo algorithm to generate independently sampled networks from the directed multi-edged self-looped random configuration model, i.e. our null random network model for apathetic retweeting. CutPermuteAndRewire is a distributed, scalable, and fault-tolerant version of the standard construction involving random pairings of out-bound and in-bound half-edges [57] through the following three steps: (i) *cutting* the directed edges representing the retweets in our observed retweet network into out-bound and in-bound half edges, (ii) *permuting* the in-bound half edges by sorting them according to pseudo-random numbers that are generated and associated with them and (iii) *rewiring* the original out-bound half edges with the permuted in-bound half edges using a distributed join. Note that the in-degree and out-degree of each node in the observed retweet network is preserved after the three steps by construction.

By taking advantage of the fastest available distributed sorting and optimized distributed joining algorithms [89], CutPermuteAndRewire can produce independent sample networks with tens of millions of retweets or edges in a small Apache Spark cluster over six commodity compute nodes. The null distribution of any test statistic of the retweet network can be directly obtained from applying it to each independent Monte Carlo sample from our scalable fault-tolerant randomized algorithm with probability given by:

$$1/(\text{number of retweets})!$$

under the null model. By comparing the observed test statistic to Monte Carlo samples from the null distribution, one can directly obtain consistent estimates of the p-value in order to attempt to reject the null hypothesis of apathetic retweeting in favor of the alternative hypothesis of non-apathetic retweeting in the framework of generalized Fisher's exact test.

2.2.4 Degrees of Separation in Geometric Retweet Network Model

To gain deeper insights into the empirical retweet network beyond the immediate retweet neighborhood of each user, i.e. those retweeted by a user and those who retweet the user, and explore the alternative hypothesis space if the null hypothesis is rejected, we adapt Milgram's concept of *six degrees of separation* [54] that all people in the world are six or fewer steps away from each other so that a sequence of "a friend of a friend" relationships can be made to connect any two people in a maximum of six steps. We adapt the concept in three major ways.

First, we focus on Twitter users and replace the mutual or undirected relationship of being a friend by the directed relationship of being a retweeter. This adaptation accounts for the main difference between various social and technological networks [84] as well as other communication networks [50] that are characterized by mutually reciprocal relationships, and the directed relationships in Twitter where a path from a user to another may follow several distinct sequences while not existing in the reverse direction [49].

Second, instead of the degrees of separation between every pair of users in an undirected sense, we are interested in the *retweet* degrees of separation in a directional sense, measured by the length of the most retweeted path that originates from a few influential users, such as politicians or hate groups, and terminates at each user.

Third, we account for the strength of the retweet relationship when defining the most retweeted path by incorporating $r_{a,b}$, the observed number of retweets between the user who is the source of the original tweets, i.e. the tweeter a, and the user who retweets them, i.e. the retweeter b, through the directed edge-weight given by $\hat{p}_{a,b} = 1/(1 + r_{a,b})$ that is used to specify the probability of an independent geometric random variable giving the number of retweets of user a by user b.

Table 2 Relative frequency of retweets by any one of the hate groups for any original tweet made by one of the politicians

Politician, observed test statistic: marginal interval for the region of acceptance at 0.001 significance level				
Donald Trump	Hillary Clinton	Bernie Sanders	Paul Ryan	Ted Cruz
0.987 : (0.6008,0.6013)	0 : (0.2708,0.2709)	0 : (0.0677,0.0682)	0: (0.00411,0.00413)	0.0131 : (0.0024,0.0028)

Let G be such a weighted directed retweet network with nodes as users and directed edges with weights:

$$\{\hat{p}_{a,b} = (1 + r_{a,b})^{-1} : \forall \text{ edge } (a, b) \text{ in } G \text{ with } r_{a,b} > 0\}.$$

The collection of independent but non-identical geometric random variables with probability parameters given by these weights is our empirically estimated *geometric retweet network* model for the joint distribution of the number of retweets of each user by another in out dataset.

We derive our estimated geometric retweet network model by recalling the well-known relationship between Poisson, exponential, and geometric random variables. If the random variable $R_{a,b}$, giving the total number of retweets of a by b is Poisson distributed with a random mean parameter $\xi_{a,b}$ that is drawn from the exponentially distributed random variable with rate parameter $1/\lambda_{a,b}$, then $R_{a,b}$ is geometrically distributed with probability parameter $p_{a,b} = 1/(1 + \lambda_{a,b})$ and expectation $\lambda_{a,b}$. We can estimate the parameters from the observed number of retweets $r_{a,b}$ via the moment estimate $\hat{\lambda}_{a,b} = r_{a,b}$ and model the number of retweets during the 9 week period according to the geometric random variable with probability parameter $\hat{p}_{a,b} = (1 + r_{a,b})^{-1}$ for each directed edge between a tweeter and a retweeter in the retweet network.

A small weight $\hat{p}_{a,b} = (1 + r_{a,b})^{-1}$ corresponds in an inversely proportionate manner to a large number of retweets and the shortest path from user a to user b through this directed weighted network corresponds to the path with a large number of retweets from user a to user b. Thus, the estimated geometric model interpretation of the weighted retweet network G allows us to use a straightforward distributed vertex program via Pregel API in Apache Spark's GraphX library to obtain the shortest path that is composed of a sequence of d tweeter-retweeter pairs of edges, say,

$$E = \{(a, v_1), (v_1, v_2), \ldots, (v_{d-1}, b)\},$$

with the lowest sum of weights given by $\sum_{(u,v)\in E} \hat{p}_{u,v}$ among all possible paths between an influential user of interest a and every other user b in G. Crucially, this $\hat{p}$-weighted shortest path is called *the most retweeted path* as it is composed of the same sequence of edges with the correspondingly high sum of retweet counts given by $\sum_{(u,v)\in E} r_{u,v}$. The length d of the most retweeted path from a to b, known as the *(retweet) degrees of separation* from a to b, has a clear interpretation as the length of the sequence of "retweeter of a retweeter" statements along the most retweeted path that is needed to link user a to user b by considering the retweet activities of every user in the network. When we have a set of influential users $A = \{a_1, a_2, \ldots, a_k\}$ of interest, say k accounts of a hateful ideology, we define the shortest path from A to any user b as the minimum of the k shortest paths from each user in A to b.

3 Results & Discussion

Given the observed heterogeneity in the activity levels of Twitter users, including the most influential user @realDonaldTrump, who was more active than the other four politicians with over twice as many retweets by about the same number of retweeters as the next most influential user @HillaryClinton (Table 1), any hypothesis test needs to control for this heterogeneity. Our approach uses retweet networks sampled independently from the null apathetic retweet network model in order to obtain the null distribution of any test statistic while preserving the observed differences in the activity levels, in terms of the observed in-degree and out-degree of each user in the retweet network. Note that we do not have multiple testing issues when attempting to reject the null hypothesis of apathetic retweeting in favor of its complementary alternative hypothesis of non-apathetic retweeting since we obtain independent Monte Carlo samples from the joint distribution of the statistics under the null model to estimate the acceptance and rejection regions at the significance level of 0.001.

In the following two sections we conducted two hypothesis tests involving frequencies of direct retweets and that of retweeters of politicians and hate groups to address our questions (Q1) and (Q2). Then we see if the tests are robust to any effects by bot and/or troll accounts and in the last section we explore the empirical retweet network.

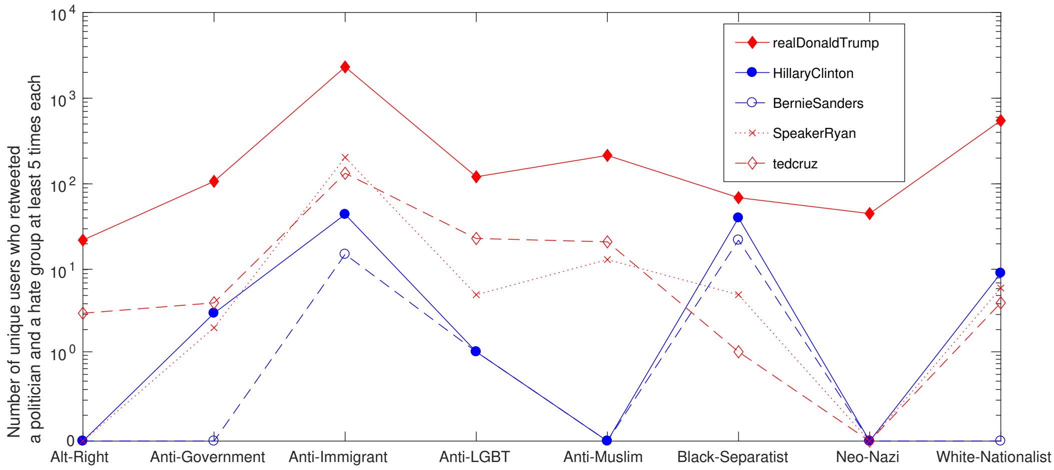


Fig. 1 Number of unique users who retweeted a politician and a hate group at least five times each (Note: The y-axis is in log-scale in powers of 10).

3.1 Frequencies of direct retweets of politicians by hate groups

None of the 194,098 retweets of an original tweet by one of the hate group accounts were made by any one of our five political accounts. However, out of the 7,233 retweets made by one of the hate group accounts, 151 were retweets of Trump, 2 were retweets of Cruz, and none were of the remaining three politicians. These 151 retweets of Trump were split among NN (87), WN (55), AM (6) and AG (3) groups, while the 2 retweets of Cruz were from NN and WN hate groups.

One of the simplest statistics is the relative frequency of retweets by any one of the hate groups of any original tweet made by each of the five politicians. This statistic of relative frequencies is given by five proportions that sum to 1. Given that Trump has more than twice as many retweets as Clinton despite having about the same number of retweeters (see out-degree and out-nbhd in Table 1), one may expect Trump to be retweeted more by hate groups even if they were retweeting the five politicians apathetically without showing any preference for Trump. To control for this effect, we used the CutPermuteAndRewire algorithm to obtain samples from the joint distribution of the relative frequencies under the null hypothesis of apathetic retweeting. As shown in Table 2, the observed statistic lies outside the acceptance region obtained from 1000 Monte Carlo samples from the null model and we thus reject the null hypothesis of apathetic retweeting with a p-value less than 0.001 in favor of non-apathetic retweeting with over 98% of retweets by hate groups preferring Trump well above the upper bound of 60.13% under the null model, while simultaneously avoiding any retweeting of Clinton or Sanders with a relative frequency of 0% that is well below their respective lower bounds of 27.08% and 6.77% under the null model, respectively. The observed relative frequencies for Ryan and Cruz, on the other hand, deviate relatively little from the corresponding marginal intervals of their acceptance region.

3.2 Frequencies of retweeters of politicians and hate groups

For a robust test statistic we looked one degree deeper into the retweet network and obtained the number of Twitter users who retweeted both a politician and a specific hate group at least five times. This provided a more conservative estimate of the number of users retweeting hate groups or politicians as it would only include those actively retweeting both. The observed frequency of such users is shown in Fig. 1 and Table 3 along with their 0.1% marginal intervals of the acceptance region for the null hypothesis of apathetic retweeting. Since the observed frequencies jointly lie well outside the acceptance region obtained from 1000 Monte Carlo samples drawn from the null model using the CutPermuteAndRewire algorithm, we reject the null hypothesis of apathetic retweeting in favor of non-apathetic retweeting with a p-value less than 0.001.

When the observed counts are in the rejection region by falling below or above the respective marginal intervals of the acceptance region, we can obtain statistically meaningful insights of the observations under the favored alternative hypothesis of non-apathetic retweet-

Table 3 Observed frequency of distinct users who retweeted a politician and an account within a hate group at least 5 times each

Ideology	Politician				
	Donald Trump	Hillary Clinton	Bernie Sanders	Paul Ryan	Ted Cruz
	observed test statistic: marginal interval for the region of acceptance at 0.001 significance level				
Anti-Government (AG)	*107 : (0, 1)	3 : (0, 3)	0 : (0, 1)	*2 : (0,1)	*4 : (0, 1)
Anti-Immigrant (AI)	*2314 : (375, 498)	°44 : (373, 492)	°15 : (369, 485)	*204 : (47, 95)	*133 : (18, 54)
Anti-LGBT (AL)	*121 : (0, 4)	1 : (0, 4)	1 : (0, 4)	*5 : (0, 3)	*23 : (0, 3)
Anti-Muslim (AM)	*215 : (0, 3)	0 : (0, 3)	0 : (0, 3)	*13 : (0, 3)	*21 : (0, 3)
Neo-Nazi (NN)	*45 : (0, 1)	0 : (0, 1)	0 : (0, 1)	0: (0, 1)	0: (0, 1)
White-Nationalist (WN)	*548 : (0, 12)	9 : (0, 10)	0 : (0, 10)	6: (0, 8)	4: (0, 7)
Black-Separatist (BlS)	°69 : (653, 811)	°40 : (649, 808)	°22 : (645, 801)	°5 : (72, 128)	°1 : (28, 66)
Alt-Right (AR)	*22 : (0, 0)	0 : (0, 0)	0 : (0, 0)	0 : (0, 0)	*3 : (0, 0)

ing. However, such insights need further models and tests for a refined and rigorous understanding within the space of alternative hypotheses.

In Table 3 the observed frequencies that lie in the rejection region above and below the marginal intervals for the acceptance region are indicated by * and °, respectively. Thus, for example, a statistically significant frequency of users have retweeted Trump, Ryan or Cruz and Anti-Immigrant groups at least five times each in a non-apathetic and preferentially approving manner since these frequencies are above their null intervals, i.e. the marginal intervals of the acceptance region for the null model of apathetic retweeting at a significance level of 0.001. These same frequencies are below the null intervals for Clinton and Sanders in favor of the non-apathetic retweeting alternative hypothesis, but in a preferentially disapproving manner. Only Trump has a significant frequency of retweeters who also frequently retweet AG, AI, AL, AM, NN, WN and AR hate groups in a non-apathetic and preferentially approving manner. Note that the Black-Separatist hate group forms a "natural control" as their retweeters non-apathetically retweet all five politicians in a preferentially disapproving manner.

Our results are robust to changes in the threshold number of retweets being at least 5, albeit at the expense of fewer observations for larger thresholds. This is because the pattern of user frequencies across hateful ideologies and the five politicians shown in Fig. 1 is preserved across other thresholds. The results are also robust when the test statistic is an expectation of the frequencies taken over all threshold levels greater than 1. We chose the naturally interpretable threshold of at least 5 retweets over the 9-week period of observation in Fig. 1 and Table 3 to select users with more than one retweet per fortnight, on average.

3.3 Effect of Bot and Troll Accounts

While we did not actively filter out highly sophisticated bot accounts given the difficulty in their detection [27], we studied the effect of removing retweets related to accounts with very high scores for two types of information entropy based on: (i) the distributions of time lags between successive posts and (ii) the content of words within the posts. This is because accounts that send messages at uniform time intervals or post messages with unusually static or identical text content might be bots or cyborgs with unusual entropy scores. Accounts with more than 10 posts that crossed 2.5 standard deviations from the mean for either entropy score were flagged as potential bot accounts using a scalable variant of a practical approach [77]. Additionally, we retrospectively examined our retweet network against the 2,752 now-deactivated Twitter accounts that the company identified as being allegedly tied to Russia's *Internet Research Agency* troll farm [81]. Out of the 12,984,331 retweets in our dataset, less than 0.1% were related to a troll account (293 were retweeted by and 12,347 were originally tweeted by a troll account) and out of 2,451,081 distinct users in our retweet network, only 172 were related to a troll account. Interestingly, removal of these troll-related retweets from the retweet network did not alter the statistical tests in Tables 2 and 3. The observed test statistics remained the same when we removed the troll related retweets and the acceptance region did not vary enough to alter any of the reported results. Similarly, removal of nearly 1% of all retweets that were affiliated with one of 2045 potential bot accounts or one of the troll accounts did not affect the test results. The robustness of the tests to presence or absence of the potential bot and identified troll accounts suggests that our results are independent of the direct influence of these bot and troll accounts, at least with regard to our Twitter study focused around the activities of the five politicians in relation to the 52 hate groups.

3.4 Degrees of Separation from Politicians and Hate Groups

For each user node in the estimated geometric retweet network model, we obtained the length of the most retweeted path, i.e. the degrees of separation, from one of the five politicians (DT = @realDonaldTrump, HC = @HillaryCinton, BS = @BernieSanders, PR = @SpeakerRyan, TC = @tedcruz) and eight hateful ideologies (AI = Anti-Immigrant, AM = Anti-Muslim, WN = White-Nationalist, AL = Anti-LGBT, AG = Anti-Govt, NN= Neo-Nazi, BlS=Black-Separatist, AR=Alt-Right). Thus each user had a *retweet ideological profile* or simply a profile given by their thirteen degrees of separation (from the five politicians and eight hateful ideologies). In order to focus on users who were politically active, we excluded (i) unreachable users who had infinite degrees of separation from all five politicians, i.e. those users who were unreachable from any one of the five politicians by a sequence of retweets, and (ii) users whose most retweeted path did not have an average number of retweets greater than 4 per edge.

We thus obtained profiles for 2,137,712 politically active users over the 9 week period (our monitored population) and report in Table 4 the frequencies of users with the same retweet ideological profile given by thirteen numbers along with the corresponding percentage of the overall sample.

Using the Manhattan distance between profiles, i.e. sum of the absolute value of the differences between the profiles (i.e. the 13 degrees of separation), we clustered the top 15 groups of users with the most frequent profiles constituting 95% of the monitored population as shown by the classification tree [71,12] in the first column of Table 4. Since this tree is built from the most retweeted path-lengths from influential users promoting various ideologies, we call this the *retweet ideological tree* of the sampled population.

Broadly there are three clades in the tree (only the top 15 most populous leaf nodes are shown). First, the politically ambivalent clade is made up of three profiles (with leaf id's 05, 07 and 12) that constitute nearly 3% of the monitored population and is characterized by the smallest possible degrees of separation not only from @realDonaldTrump (DT), but also from either @HillaryClintom (HC) or @BernieSanders (BS). Second, the Democratic-leaning cluster that is characterized by having fewer degrees of separation from at least one of the two Democratic politicians: HC and BS. It is shown by the clade with leaf id's 02, 04, 03, 13, and the more diverged id 11 that is farthest from all hate groups and from the Politicians but relatively closer to HC. The three most populous leaf nodes in the Democratic-leaning clade with id's 02, 03 and 04, constitute about 37%, 11% and 3.5% of the monitored population. Third, the Republican-leaning cluster that is characterized by having fewer degrees of separation from at least one of the three Republican politicians: @realDonaldTrump (DT), @SpeakerRyan (PR) and @tedcruz (TC). It is shown by the larger clade with the following seven leaf id's: 01, 15, 14, 08, 09, 06 and 10 in Table 4. The most populous leaf nodes in the Republican-leaning clade are made up of two of the 15 groups of users with id's 01 and 15, and constitute nearly 38% of the monitored population with distinct profiles that are primarily characterized by the smallest possible degrees of separation of 1 from DT. Similarly, the other two less populous Republican-leaning id's (08, 09, 06 and 10) are primarily characterized by different degrees of separation from DT, PR and TC, whereas id 14, with its own branch, is notably hateful by having the smallest possible degrees of separation from both DT and Anti-Immigrant groups.

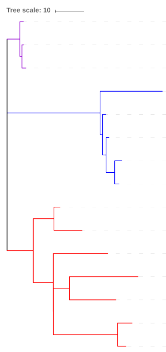
Note that more than 91% and 95% of users have their degrees of separation from all eight hateful ideologies no smaller than 4 and 3, respectively. This is reflective of little direct influence by hate groups for over 95% of the monitored population. In order to get insights on users whose most retweeted paths from at least one of the hateful ideologies is within 3 or fewer retweet degrees of separation, we zoomed into the less frequent profiles of the remaining 5% of the monitored population of politically active users. These results are depicted as cumulative percentages of the monitored population for the following nine pairs of retweet degrees of separation from one of the five politicians and a given hateful ideology:

$$\{(1,1),(1,2),(2,1),(1,3),(2,2)(3,1),(2,3),(3,2),(3,3)\}$$

in Fig. 2. For example, over 0.6% and 1% of the population is at 1 degree of separation from @realDonaldTrump and within 1 and 2 degrees of separation from Anti-Immigrant ideologies, as specified by the y-axis values corresponding to the x-axis values of $(1,1)$ and $(1,2)$, respectively, in the top left sub-plot of Fig. 2. Similarly, nearly 2.4% of the population is within 3 degrees of separation from Trump and Anti-Immigrant ideologies as specified by the y-axis value corresponding to the x-axis value of $(3,3)$ in the same sub-plot. Thus, the cumulative percentage of the politically active monitored population that is within a given ordered pair of degrees of separation (from a politician and a hateful ideology) is depicted similarly by the sub-plots of Fig. 2.

Although the number of distinct retweeters of @realDonaldTrump and @HillaryClinton are roughly the

Table 4 The top 15 groups of users according to their profiles of most retweeted path-lengths from the five politicians⋆ and eight hateful ideologies° given by their id, frequency, percentage of population and their classification given by the retweet ideological tree with leaf nodes as the ids.

ideological tree	id	frequency	percentage of population	Politician					Hate Group							
				DT	HC	BS	PR	TC	AI	AM	WN	AL	AG	NN	BlS	AR
	05	42853	02.005	1	1	2	4	4	5	5	7	6	4	7	7	7
	07	11481	00.537	1	2	1	4	4	5	5	7	6	4	7	7	7
	12	5868	00.274	1	1	1	4	4	5	5	7	6	4	7	7	7
	11	5972	00.279	4	2	3	5	7	8	8	9	9	7	10	10	10
	02	791286	37.016	3	1	2	4	6	7	7	8	8	6	9	9	9
	04	74126	03.468	3	1	1	4	6	7	7	8	8	6	9	9	9
	03	232093	10.857	3	2	1	6	6	7	7	9	8	6	9	9	9
	13	5173	00.242	3	1	1	6	6	7	7	8	8	6	9	9	9
	01	811586	37.965	1	4	7	4	4	5	5	7	6	4	7	7	7
	15	3892	00.182	1	4	7	1	4	5	5	7	3	4	7	7	7
	14	4011	00.188	1	4	7	4	4	1	5	3	5	4	5	7	7
	08	10460	00.489	3	5	9	1	3	3	3	5	3	6	7	9	9
	09	8069	00.377	3	3	3	3	1	4	3	3	3	6	5	6	9
	06	29997	01.403	2	3	3	3	3	5	3	3	5	5	5	3	3
	10	6257	00.293	1	3	3	4	4	5	3	3	5	4	5	3	3

⋆DT = @realDonaldTrump, HC = @HillaryCinton, BS = @BernieSanders, PR = @SpeakerRyan, TC = @tedcruz
°AI = Anti-Immigrant, AM = Anti-Muslim, WN = White-Nationalist, AL = Anti-LGBT, AG = Anti-Govt, NN= Neo-Nazi, BlS=Black-Separatist, AR=Alt-Right

same at around 950,000 (see out-nbhd column in Table 1), the proportion of the population that is within 1 or 2 degrees of separation from @realDonaldTrump and from any one of the six hateful ideologies (AI, AM, AL, WN, AG, and NN), whose test statistic fell above the acceptance region and thus suggestive or indicative of non-apathetic preferential retweeting, is consistently higher than that from any other politician (unlike those for the Black-Separatists, our "natural control" in the exact test, and Alt-Right ideologies partly due to minimal Twitter activity induced by SPLC's reclassification during the course of our data collection). As mentioned earlier, the null model of apathetic retweeting was strongly rejected using the frequency of users who retweeted both a politician and a hateful ideology more than four times each in favor of the alternative hypothesis of non-apathetic retweeting in which these same six ideologies were found to be retweeted in a preferentially approving manner by those who also retweeted Trump (Table 3). Thus, zooming into the less frequent profiles of the remaining 5% of the monitored population that is within 3 retweet degrees of separation from a politician and a hate group is a conservative exploration of the empirical geometric retweet network (see Fig. 2).

Recall that the degrees of separation has a clear interpretation as the length of the sequence of "retweeter of a retweeter" statements along the most retweeted path that links a set of influential users to each user by considering the global retweet activities of every user in the network. Thus, this exploratory analysis provides global insights into the most retweeted pathways from influential sets of users (politicians and leaders of hateful ideologies), through the joint distribution of the lengths of these paths, to each of the $2,137,712$ politically active users. It further provides an ideological tree to classify users based on the Manhattan distance of these path-lengths that allows one to ask where one lies on a well-defined ideological tree, a mathematically natural generalization [24,23] of the line giving the usual bi-polar spectrum of left versus right or Democrat versus Republican, a common restriction in social media research [9, for e.g.], and thus allowing us to consider the effects of transmissions of "memes" [20, p. 192][4] on social media networks [69] through ideological trees using more sophisticated mathematical notions[6,7].

4 Conclusion

Using 9-weeks of Twitter data collected around the 2016 US presidential election involving nearly 22 million communication events, the present research examined the Twitter linkages between five major American political leaders (Donald Trump, Hillary Clinton, Ted Cruz, Bernie Sanders, and Paul Ryan) with American hate groups. Using two different approaches to the data, we found converging evidence that Donald Trump possessed unique appeal to a variety of American hate groups.

First, utilizing direct retweets, we found through a generalized Fisher's exact test that Twitter users who

[4] See [75].

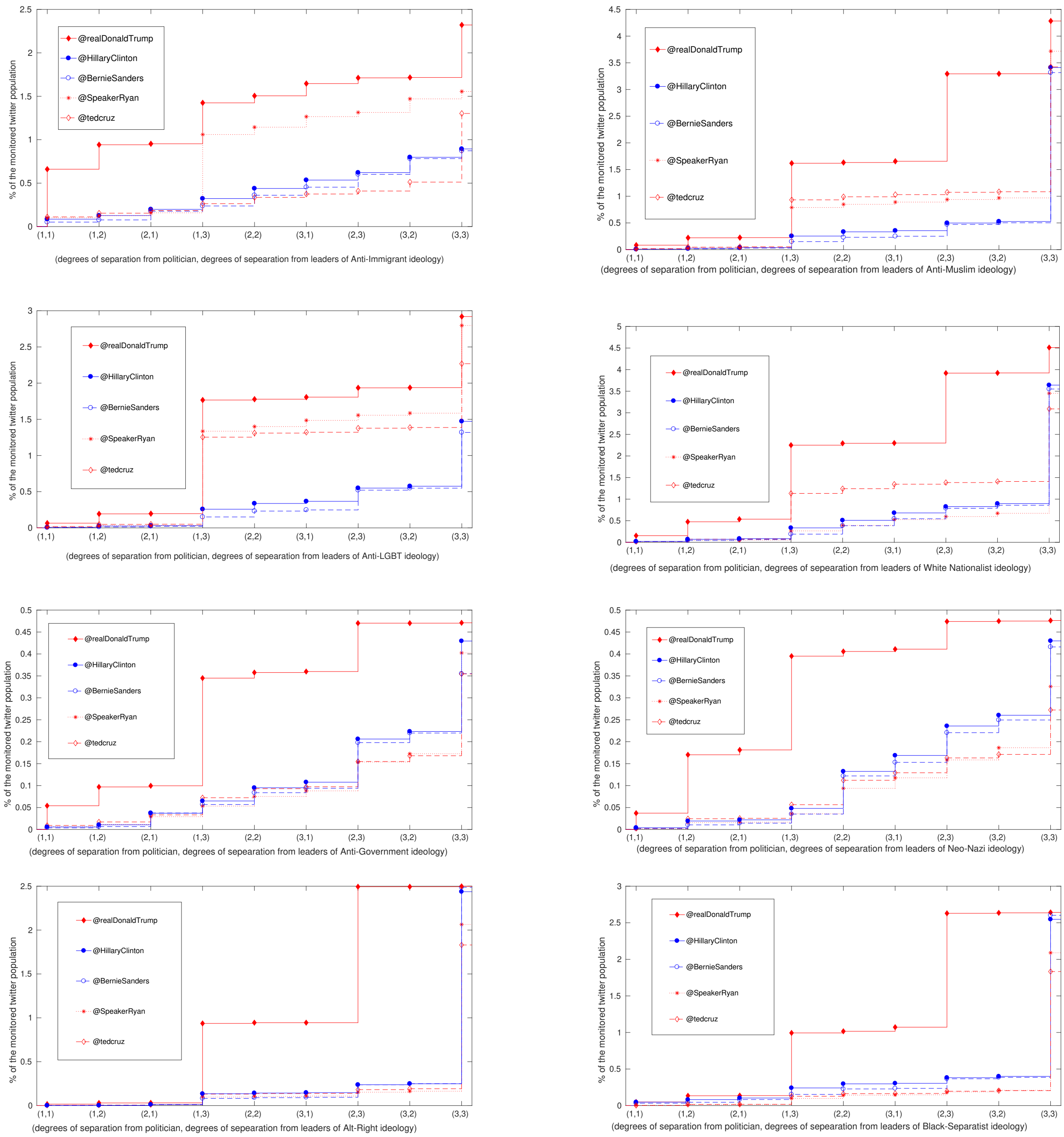


Fig. 2 Cumulative percentage of the monitored Twitter population who are within a given in-degree of separation from a politician and a hateful ideology.

frequently retweeted hate groups (i.e., at least 5 times in the span of 9 weeks) were significantly more likely to retweet Trump over any other politician. A significant number of retweeters of nearly all hateful ideologies, except for Black-Separatists (BlS), were found in Trump's Twitter network including Anti-Immigrant (AI), Anti-Muslim (AM), Anti-Government (AG), Neo-Nazi (NN), Alt-Right (AR), White-Nationalists (WN) and Anti-LGBT (AL). By comparison, retweeters of only AI, AM, and AL were significantly linked to Paul Ryan or Ted Cruz, but even these linkages were quite small relative to those observed in Trump's network.

Second, using most retweeted path lengths, we found that a significantly larger proportion of people were within one, two and three degrees of separation from Donald Trump and from AG, AL, AM, AI, WN or NN ideologies, relative to the corresponding degrees of separation from Hillary Clinton, Bernie Sanders, Ted Cruz, or Paul Ryan and the same hate groups. None of the political candidates appeared to hold particular appeal to users retweeting BlS hate groups. While there is some debate about the relationship between hate groups and hate crimes [2,35,67, for e.g.], it is still important to examine the social media networks of prominent politicians and hate groups as political leadership influences social norms about what is acceptable and what is not in wider society [41].

In the present work, our retweet network provides for the second narrative we mentioned in the introduction that Trump held unique appeal to those espousing hateful ideologies more so than other Republican or Democratic candidates. Nevertheless, it is important to note that only a small fraction of those retweeting Trump were within a few degrees of separation from hateful ideologies suggesting that most of Trump's support appear removed from people with hateful views.

References

1. ABC News/Washington Post: Clinton hits a new high in unpopularity; on par with Trump among registered voters. Post poll, Langer Research Associates. Available from `http://www.langerresearch.com/wp-content/uploads/1144-59ClintonTrumpFavorability.pdf`. Accessed on May 28, 2017 (2016)
2. Adamczyk, A., Gruenwald, J., Chermak, S.M., Freilich, J.D.: The relationship between hate groups and far-right ideological violence. J Contemp Criminal Justice **30**(3), 310–332 (2014)
3. Aldous, D.: Interacting particle systems as stochastic social dynamics. Bernoulli **19**(4), 1122–1149 (2013)
4. Altemeyer, B.: The Authoritarian Specter. Harvard University Press (1996)
5. Anderson, C.: Donald Trump is the result of white rage, not economic anxiety. Time (2016). Nov 16 2016. Available from `http://time.com/4573307/donald-trump-white-rage/`. Accessed on May 28, 2017
6. Athreya, S., Eckhoff, M., Winter, A.: Brownian motion on $\mathbb{R}$-trees. Transactions of the American Mathematical Society **365**(6), 3115–3150 (2013)
7. Athreya, S., Löhr, W., Winter, A., et al.: Invariance principle for variable speed random walks on trees. The Annals of Probability **45**(2), 625–667 (2017)
8. Ball, M.: Donald Trump and the politics of fear. The Atlantic (2016). September 2 2016. Available from `https://www.theatlantic.com/politics/archive/2016/09/donald-trump-and-the-politics-of-fear/498116/`. Accessed on May 28, 2017
9. Barberá, P., Jost, J.T., Nagler, J., Tucker, J.A., Bonneau, R.: Tweeting from left to right. Psychological Science **26**(10), 1531–1542 (2015)
10. Beauchamp, Z.: White riot. Vox (2017). THE BLOG 10/21/2016 07:15 am ET — Updated Oct 21, 2016. Available from `https://www.vox.com/2016/9/19/12933072/far-right-white-riot-trump-brexit`. Accessed on May 28, 2017
11. Bliuc, A.M., Faulkner, N., Jakubowicz, A., McGarty, C.: Online networks of racial hate: A systematic review of 10 years of research on cyber-racism. Computers in Human Behavior **87**, 75 – 86 (2018). DOI https://doi.org/10.1016/j.chb.2018.05.026. URL `http://www.sciencedirect.com/science/article/pii/S0747563218302528`
12. Boc, A., Diallo, A.B., Makarenkov, V.: T-rex: a web server for inferring, validating and visualizing phylogenetic trees and networks. Nucleic Acids Research **40**(W1), W573 (2012)
13. Bollobás, B.: Random Graphs, 2 edn. Cambridge Studies in Advanced Mathematics. Cambridge University Press (2001). DOI 10.1017/CBO9780511814068
14. Cadwalladr, C.: The great British Brexit robbery: how our democracy was hijacked. The Guardian (2017). The Observer. 2017, May 7, Available from `https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy`. Accessed on May 28, 2017
15. Southern Poverty Law Center: Hate map. SPLC (2016). October 11, 2013 4:00 AM, Available from `https://www.splcenter.org/hate-map`. Accessed on May 28, 2017
16. Southern Poverty Law Center: The year in hate and extremism. SPLC (2017). Available from `https://www.splcenter.org/fighting-hate/intelligence-report/2017/year-hate-andextremism`. Accessed on May 21, 2019
17. Chermak, S., Freilich, J., Suttmoeller, M.: The organizational dynamics of far-right hate groups in the United States: Comparing violent to non-violent organizations. Studies in Conflict & Terrorism **36**, 193–218 (2013)
18. Craig, M.A., Richeson, J.A.: On the precipice of a 'majority-minority' America: Perceived status threat from racial demographic shift affects White Americans' political ideology. Psych Sci **25**(6), 1189–1197 (2014)
19. Crandall, C.S., White, M.H.I.: Trump and the social psychology of prejudice. Undark (2016). VARIABLES/Essays & Opinion, 11.17.2016, Available from `https://undark.org/article/trump-social-psychology-prejudice-unleashed/`. Accessed on May 28, 2017
20. Dawkins, R.: The Selfish Gene. Oxford University Press, Oxford, UK (1976)
21. Detrow, S.: KKK paper endorses Trump; campaign calls outlet repulsive. NPR (2016). November 2, 2016 10:05 AM ET, Available from `http://www.npr.org/2016/11/02/500352353/kkk-paper-endorses-trump-campaign-calls-outlet-repulsive`. Accessed on May 28, 2017
22. Developers, T.: Twitter developer documentation. `https://dev.twitter.com/docs` (2018). Accessed on May 17 2017
23. Dress, A., Terhalle, W.: The real tree. Advances in Mathematics **120**(2), 283 – 301 (1996)
24. Dress, A.W.: Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: A note on combinatorial properties of metric spaces. Advances in Mathematics **53**(3), 321 – 402 (1984)
25. Editors: The trump tweet tracker. The Atlantic (2017). Live BLOG May 18, 2017, Available from `https://www.theatlantic.com/liveblogs/2017/05/donald-trump-twitter/511619/`. Accessed on May 28, 2017
26. Enten, H.: Americans' distaste for both Trump and Clinton is record-breaking. FiveThirtyEight.com (2016). May 5 2016 at 8:29 AM, Available from `https://fivethirtyeight.com/features/americans-distaste-for-both-trump-and-clinton-is-record-breaking/`. Accessed on May 28, 2017

27. Ferrara, E., Varol, O., Davis, C., Menczer, F., Flammini, A.: The rise of social bots. Commun. ACM **59**(7), 96–104 (2016)
28. Fortuna, P., Nunes, S.: A survey on automatic detection of hate speech in text. ACM Comput. Surv. **51**(4), 85:1–85:30 (2018). DOI 10.1145/3232676. URL http://doi.acm.org/10.1145/3232676
29. Freilich, J.D., A., P.W.: Mismeasuring militias: Limitations of advocacy group data and of state-level studies of paramilitary groups. Justice Quarterly **23**(1), 147–162 (2006)
30. Gemignani, M., Hernandez-Albujar, Y.: Hate groups targeting unauthorized immigrants in the us: Discourses, narratives, and subjectivation practices on their websites. Eth Rac Studies **38**(15), 2754–2770 (2015)
31. Gilliard-Matthews, S.: The impact of the economic downturn, immigrants, and political representation of White supremacist organization in the United States. Sociol Focus **44**(3), 255–279 (2011)
32. Glaser, J., Dixit, J., Green, D.P.: Studying hate crime with the internet: What makes racists advocate racial violence? J Soc Issues **58**(1), 177–193 (2002)
33. Graham, D.: How did Maajid Nawaz emd up on a list of Anti Muslim extremists? The Atlantic (2016). October 29, 2016, Available from https://www.theatlantic.com/international/archive/2016/10/maajid-nawaz-splc-anti-muslim-extremist/505685/. Accessed on June 28, 2017
34. Grassegger, H., Krogerus, M.: The data that turned the world upside down. Stanford Public Policy Program (2017). VARIABLES/Essays & Opinion, 01.28.2017, Available from https://publicpolicy.stanford.edu/news/data-turned-world-upside-down. Accessed on June 15, 2017
35. Green, D.P., Rich, A.: White supremacist activity and cross burning in north carolina. J Quant Criminology **14**, 263–282 (1998)
36. Hale, W.C.: Extremism on the world wide web: A research review. Criminal Justice Studies **25**(4), 343–356 (2012)
37. Hananoki, E.: The complete history of Donald Trump's relationship with the White nationalist movement. Media Matters for America (2016). Blog, August 18, 2016 2:40 PM EDT. Available from https://www.mediamatters.org/blog/2016/08/18/complete-history-donald-trumps-relationship-white-nationalist-movement/212502. Accessed on May 28, 2017
38. Hayden, M.E.: US civil rights groups worry anti-Muslim sentiment is fueling right-wing extremism. ABC News (2017). Jun 12, 2017, 6:57 AM ET, Available from http://abcnews.go.com/US/us-civil-rights-groups-worry-anti-muslim-sentiment/story?id=47840271. Accessed on June 15, 2017
39. Hochschild, A.R.: Strangers in Their Own Land: Anger and Mourning on the American Right. New Press (2016)
40. Hogg, M.A.: Uncertainty–identity theory. Adv Exp Soc Psychol **39**, 69–126 (2007)
41. Hogg, M.A., Reid, S.A.: Social identity, self-categorization and communication of group norms. Comm Theory **16**(1), 7–30 (2006)
42. Howard, P.N., Kollanyi, B., Woolley, S.: Bots and automation over twitter during the us election. Zugriff am **14** (2016). URL http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2016/11/Data-Memo-US-Election.pdf
43. Ingraham, C.: Two new studies find racial anxiety is the biggest driver of support for Trump. Washington Post (2016). Wonkblog, June 6 2016. Available from https://www.washingtonpost.com/news/wonk/wp/2016/06/06/racial-anxiety-is-a-huge-driver-of-support-for-donald-trump-two-new-studies-find/?utm_term=.d5b2d1ca159b. Accessed on May 28, 2017
44. Jansen, B.J., Zhang, M., Sobel, K., Chowdury, A.: Twitter power: Tweets as electronic word of mouth. J Amer Soc Info Science Tech **60**(11), 2169–2188 (2009)
45. Jonsson, P.: Annual report cites rise in hate groups, but some ask: what is hate? Christian Science Monitor (2013). February 23, 2011, Available from http://www.csmonitor.com/USA/Society/2011/0223/Annual-report-cites-rise-in-hate-groups-but-some-ask-What-is-hate. Accessed on May 28, 2017
46. Jost, J.T., Glaser, J., Kruglanski, A.W., Sulloway, F.J.: Political conservatism as motivated social cognition. Psychol Bull **129**(3), 339–375 (2003)
47. Kharakh, B., Primack, D.: Donald Trump's social media ties to White supremacists. Fortune (2016). March 22, 2016, 10:03 AM EDT. Available from http://fortune.com/donald-trump-white-supremacist-genocide/. Accessed on May 28, 2017
48. Knowles, E.D., Tropp, L.R.: Donald Trump at the rise of White identity politics. Huffington Post (2016). THE BLOG 10/21/2016 07:15 am ET — Updated Oct 21, 2016. Available from http://www.huffingtonpost.com/the-conversation-us/donald-trump-and-the-rise_b_12584616.html. Accessed on May 28, 2017
49. Kwak, H., Lee, C., Park, H., Moon, S.: What is twitter, a social network or a news media? In: Proceedings of the 19th International Conference on World Wide Web, WWW 10, pp. 591–600. ACM, New York, NY, USA (2010)
50. Leskovec, J., Horvitz, E.: Planetary-scale views on a large instant-messaging network. In: J. Huai, R. Chen, H.W. Hon, Y. Liu, W.Y. Ma, A. Tomkins, X. Zhang (eds.) WWW, pp. 915–924. ACM (2008)
51. Major, B., Blodorn, A., Major Blascovich, G.: The threat of increasing diversity: Why many White Americans support Trump in the 2016 presidential election. Group Proc Intergr Rel (2017). In press
52. Martin, T.: Dissecting Trump's most rabid online following. FiveThirtyEight (2017). Available from https://fivethirtyeight.com/features/dissecting-trumps-most-rabid-online-following/. Accessed on June 14, 2017
53. Metaxas, P., Mustafaraj, E., Wong, K., Zeng, L., O'Keefe, M., Finn, S.: What do retweets indicate? results from user survey and meta-review of research. In: International AAAI Conference on Web and Social Media. ACM (2015)
54. Milgram, S.: The small-world problem. Psychology Today **1**(1) (1967)
55. Mondal, M., Silva, L.A., Benevenuto, F.: A measurement study of hate speech in social media. In: Proceedings of the 28th ACM Conference on Hypertext and Social Media, HT '17, pp. 85–94. ACM, New York, NY, USA (2017). DOI 10.1145/3078714.3078723. URL http://doi.acm.org/10.1145/3078714.3078723
56. Müller, K., Schwarz, C.: Making america hate again? twitter and hate crime under trump. Tech. rep., Available from http://dx.doi.org/10.2139/ssrn.3149103. Accessed on May 20, 2019 (2018)

57. Newman, M.E.J., Strogatz, S.H., Watts, D.J.: Random graphs with arbitrary degree distributions and their applications. Phys. Rev. E **64**, 026118 (2001)
58. O'Connor, T.: Trump, alt-right movement: Nazis, white supremacists praise new security measures against radical Islam. International Business Times (2017). Available from https://tinyurl.com/y9j8oq72. Accessed on July 21, 2017
59. Packer, G.: Head of the class. The New Yorker (2016). COMMENT May 16 2016 issue. Available from http://www.newyorker.com/magazine/2016/05/16/how-donald-trump-appeals-to-the-white-working-class. Accessed on May 28, 2017
60. Pew Research Center: Beyond distrust: How Americans view their government. Research report, Pew Research Center. Available from http://www.people-press.org/2015/11/23/beyond-distrust-how-americans-view-their-government/. Accessed on May 28, 2017 (2015)
61. Piggott, S.: Neo-Nazi Billy Roper Organizing ACT for America's "March Against Sharia" in Arkansas. Southern Poverty Law Center (2017). Hate Watch, June 6, 2017, Available from https://www.splcenter.org/hatewatch/2017/06/06/neo-nazi-billy-roper-organizing-act-americas-march-against-sharia-arkansas. Accessed on June 15, 2017
62. Raymond, L., Pyke, A.: 'Trump is setting us free': White supremacists celebrate reports that Trump will dial down scrutiny. Think Progress (2017). Available from https://tinyurl.com/znmy9xe. Accessed on May 28, 2017
63. Reicher, S.D., Haslam, S.A.: Trump's appeal: What psychology tells us. Scientific American Mind (2017)
64. Reuters: Ku Klux Klan newspaper declares support for Trump. Reuters (2016). POLITICS — Wed Nov 2, 2016 — 4:22pm EDT, Reporting by Mohammad Zargham; Editing by Jonathan Oatis, Available from http://www.reuters.com/article/us-usa-election-trump-kkk-idUSKBN12X2IG. Accessed on May 28, 2017
65. Roussos, G.: Donald Trump's psychological manipulation of the American people. Research blog, Yale Institution for Social and Policy Studies. Available from https://tinyurl.com/y9yacnyr. Accessed May 28, 2017 (2016)
66. Runciman, W.: Relative deprivation and social justice: a study of attitudes to social inequality in twentieth-century England. Reports of the Institute of Community Studies. University of California Press (1966)
67. Ryan, M.E., Leeson, P.T.: Hate groups and hate crime. Intl Rev Law Econ **31**(4), 256–262 (2011)
68. Ryan, M.E., Leeson, P.T.: Hate groups and hate crimes. Intl Rev Law and Econ **31**, 256–262 (2011)
69. Sainudiin, R., Welch, D.: The transmission process: A combinatorial stochastic process for the evolution of transmission trees over networks. Journal of Theoretical Biology **410**, 137–170 (2016)
70. Sainudiin, R., Yogeeswaran, K., Nash, K., Sahioun, R.: Rejecting the null hypothesis of apathetic retweeting of US politicians and splc-defined hate groups in the 2016 US presidential election. In: U. Brandes, C. Reddy, A. Tagarelli (eds.) IEEE/ACM 2018 International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2018, Barcelona, Spain, August 28-31, 2018, pp. 250–253. IEEE Computer Society (2018). DOI 10.1109/ASONAM.2018.8508555. URL https://doi.org/10.1109/ASONAM.2018.8508555
71. Saitou, N., Nei, M.: The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution **4**(4), 406 (1987)
72. Schmidt, A., Wiegand, M.: A survey on hate speech detection using natural language processing. In: Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media, pp. 1–10. Association for Computational Linguistics, Valencia, Spain (2017). DOI 10.18653/v1/W17-1101. URL https://www.aclweb.org/anthology/W17-1101
73. Selk, A.: Twitter co-founder: I'm sorry if we made trump's presidency possible. Washington Post (2017). May 21, 2017, Available from https://tinyurl.com/yckypjku. Accessed on May 28, 2017
74. Shalby, C.: A brief history of the Trump campaign's controversies with women. Los Angeles Times (2016). October 9, 2016, 7:00 AM. Available from http://www.latimes.com/politics/la-na-pol-trump-campaign-insults-women-20161009-snap-htmlstory.html. Accessed on May 28, 2017
75. Solon, O.: Richard Dawkins on the internet's hijacking of the word 'meme'. WIRED Magazine (2013). 20 June 13, Available from http://www.wired.co.uk/news/archive/2013-06/20/richard-dawkins-memes. Accessed on June 22, 2017
76. Stephan, W.G., Stephan, C.W.: An integrated threat theory of prejudice. In: S. Oskamp (ed.) Reducing Prejudice and Discrimination, Claremont Symposium on Applied Social Psychology Series, chap. 2, pp. 23–45. Taylor & Francis (2000)
77. Steve Kramer: Identifying viral bots and cyborgs in social media. Tech. rep., Available from https://www.oreilly.com/ideas/identifying-viral-bots-and-cyborgs-in-social-media. Accessed on Feb 18, 2018 (2017)
78. Torres, A.: To the southern poverty law center, believing homosexual activity is sinful is 'hate' and 'bigotry'. National Review (2013). October 11, 2013, Available from http://www.nationalreview.com/article/360968/splc-and-hate-groups-alec-torres. Accessed on May 28, 2017
79. Tumasjan, A., Sprenger, T., Sandner, P., Welpe, I.: Predicting elections with twitter: What 140 characters reveal about political sentiment. In: Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, pp. 178–185 (2010)
80. Tuttle, I.: Racism (still) didn't elect Trump. National Review (2017). Available from http://www.nationalreview.com/corner/446808/washington-post-racism-motivated-trump-voters. Accessed on June 14, 2017
81. Twitter Public Policy: Update on Twitter's review of the 2016 US election. Tech. rep., Available from https://blog.twitter.com/official/en_us/topics/company/2018/2016-election-update.html. Accessed on Feb 18, 2018 (2018)
82. Vaccari, C., Valeriani, A., Barbera, P., Bonneau, R., Jost, J.T., Nagler, J., Tucker, J.A.: Political expression and action on social media: Exploring the relationship between lower and higher-threshold political activities among twitter users in italy. J Comp Med Comm **20**(2), 221–239 (2015)
83. Walker, I., Pettigrew, T.F.: Relative deprivation theory: An overview and conceptual critique. Brit J Soc Psychol **23**, 301–310 (1984)
84. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature **393**, 440–442 (1998)

85. Webster, D.M., Kruglanski, A.W.: Individual differences in need for cognitive closure. J Per Soc Psychol **67**(6), 1049–1062 (1994)
86. Wood, T.: Racism motivated Trump voters more than authoritarianism. Washington Post (2017). Available from `https://tinyurl.com/yavzdm5s`. Accessed on June 14, 2017
87. Woolley, S., Guibeault, D.: Computational propaganda in the United States of America: Manufacturing consensus online. Samuel Woolley and Philip Howard, Eds. Working Paper No. 2017.5 Project on Computational Propaganda, Oxford, UK pp. 1–29 (2017). URL `http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Comprop-USA.pdf`
88. Yogeeswaran, K., Dasgupta, N.: The devil is in the details: Abstract versus concrete construals of multiculturalism differentially impact intergroup relations. J Pers Soc Psychol **106**(5), 772–789 (2014)
89. Zaharia, M., Xin, R.S., Wendell, P., Das, T., Armbrust, M., Dave, A., Meng, X., Rosen, J., Venkataraman, S., Franklin, M.J., Ghodsi, A., Gonzalez, J., Shenker, S., Stoica, I.: Apache spark: A unified engine for big data processing. Commun. ACM **59**(11), 56–65 (2016)