Spring March 17, 2020

Rise of the Machines? Examining the Influence of Social Bots on a Political Discussion Network

Loni Hagen, *University of South Florida*
Stephen T. Neely

Article

Social Science Computer Review
1-24

DOI: 10.1177/0894439320908190
journals.sagepub.com/home/ssc
SAGE

Rise of the Machines? Examining the Influence of Social Bots on a Political Discussion Network

Loni Hagen[1], Stephen Neely[1], Thomas E. Keller[1], Ryan Scharf[2], and Fatima Espinoza Vasquez[3]

Abstract

The growing influence of *social bots* in political discussion networks has raised significant concerns, particularly given their potential to adversely impact democratic outcomes. In this study, we report the results of a case study analysis of bot activity in a recent, high-profile political discussion network. Specifically, we examine the prevalence and impact of bots in a Twitter network discussing the Special Counsel investigation into Russian interference in the 2016 U.S. elections. Using this discussion network, we conduct a "before-and-after" analysis to examine the prevalence of social bots in the discussion network as well as their influence on key network features such as (1) network structure, (2) content/messaging, (3) sentiment, and (4) influentialness. Our findings suggest that bots can affect political discussion networks in several significant ways. We found that bot-like accounts created the appearance of a virtual community around far-right political messaging, attenuated the influence of traditional actors (i.e., media personalities, subject matter experts), and influenced network sentiment by amplifying pro-Trump messaging. The results of this analysis add to a growing body of literature on the use and influence of social bots while at the same time uniquely examining their influence in a nonelectoral, political setting.

Keywords

bots, social media, Twitter, artificial intelligence, political communication, media politics

The past decade has seen a marked shift in patterns of media consumption and political engagement both here in the United States and throughout the world. Increasingly, individuals rely on web-based applications—such as digital news outlets and social media—for a variety of informational services including news and political information, product reviews, and even disaster preparedness and emergency updates (Greenwood et al., 2016; Kim et al., 2014; Mitchell et al., 2012; Stewart & Wilson, 2016). While this shift toward digital information-seeking is motivated largely by

[1] University of South Florida, Tampa, FL, USA
[2] American Integrity Insurance, Tampa, FL, USA
[3] University of Kentucky, Lexington, KY, USA

Corresponding Author:
Loni Hagen, University of South Florida, 4202 East Fowler Ave., Tampa, FL 33620, USA.
Email: lonihagen@usf.edu

convenience, some commentators have suggested that the Internet may also have a “democratizing effect” on public life, particularly insofar as it expands access to information and creates greater opportunities for civic engagement (i.e., de Zuniga et al., 2012; Papacharissi, 2004). However, as the Internet has come to occupy a more prominent place in the public sphere, others have raised concerns over the unregulated nature of web-based content and the widespread potential for emerging technologies to be used toward undemocratic ends (i.e., Persily, 2017; Sunstein, 2007).

Of particular concern in recent years has been the growing influence of *social bots* in political discussion networks, notably their potential to adversely impact democratic outcomes. Traditionally, *web bots* have been used to execute a variety of routine, online tasks (such as ranking search results and directing customer service inquiries). However, more sophisticated programs are increasingly designed to mimic human behavior—particularly on social networking sites—in an effort to misrepresent public opinion and at times even to proliferate misinformation and propaganda (i.e., Bradshaw & Howard, 2018; Broniatowski et al., 2018; Woolley, 2016). Among other instances, bots were widely employed in efforts to influence online, political discussion networks during the 2016 presidential election in the United States as well as the United Kingdom’s Brexit referendum earlier that same year (i.e., Bastos & Mercea, 2019; Guilbeault & Woolley, 2016; Leask, 2017; *The Times*, 2017).[1]

These and other high-profile cases have raised significant questions over the quality of information circulated in online discussion networks as well as our ability to accurately measure public sentiment in digital spaces (i.e., Persily, 2017; Woolley, 2016). Despite these growing concerns, relatively little is known about how these programs impact political discussion networks or the extent to which they directly influence public opinion and political behavior. Among others, Woolley (2016) has emphasized the importance of pursuing further research in this area, noting that the use of social bots in political contexts needs to “ . . . be better understood for the sake of free speech and the future of digitally mediated civic engagement” (p. 2).

In the current study, we add to a growing body of literature in this area by providing a case study analysis of bot activity in a recent, high-profile political discussion network. Specifically, we examine the prevalence, behavior, and influence of bots in a Twitter network discussing the Special Counsel investigation into Russian interference in the 2016 U.S. elections. Using this discussion network, we conduct a “before-and-after” analysis to examine the prevalence of social bots in the discussion network as well as their influence on key network features such as (1) network structure, (2) content/messaging, (3) sentiment, and (4) influentialness. Our findings suggest that bots can influence political discussion networks in several significant ways. Specifically, we found that bot-like accounts created the appearance of a virtual community around far-right political messaging, obscured the influence of traditional actors (i.e., media personalities, subject matter experts), and influenced network sentiment by amplifying pro-Trump messaging.

The results of this analysis add to a growing body of literature on the use and impact of social bots while at the same time uniquely examining their potential influence in a nonelectoral, political setting. The findings are discussed in light of their technological, informational, and political implications.

Background Information

Broadly defined, web bots are algorithm-based software programs designed to perform automated tasks. For example, bots can be used to organize search engine results, personalize advertisements based on search history, sort news stories based on interest, and even answer customer service inquiries. Generally speaking, web bots are common—and typically benign. In fact, some recent research suggests that more than 50% of all web activity may be bot generated, with much of this resulting from so-called good bots, which perform routine, web-based tasks (Zeifman, 2017).

In contrast to these traditional, task-oriented programs, a *social bot* is a web bot that "... automatically produces content and interacts with humans ... trying to emulate and possibly alter their behavior" (Ferrara et al., 2016, p. 96). These tools are increasingly being used to both generate and distribute content within social networks, often in an effort to engage in phishing activities or influence the measurement of key metrics related to phenomena such as usage, interest, and popularity. Along with uniquely operating within social networks, recent manifestations of social bots are noteworthy for their ability to mimic human behavior (i.e., patterns, language) and avoid detection (Crothers, 2019; Ferrara et al., 2016). Increasingly, these tools have been deployed toward political ends, with (Woolley & Howard, 2016) defining political bots as "... algorithms that operate over social media, written to learn from and mimic real people, so as to manipulate public opinion across a diverse range of social media and device networks" (p. 4885).

In the context of political discourse, social bots have a variety of potential applications, perhaps most notably "amplification," or the ability to create and/or widely circulate large amounts of online content in order to promote individuals and amplify messages. For example, social bots can cause topics to trend by promoting hashtags or "likes," effectively causing the impression of widespread support (or opposition) for an individual and/or idea (i.e., Broniatowski et al., 2018; Persily, 2017). Similarly, bots can be used to suggest popularity by artificially boosting the social media "follower" numbers for an individual or organization (Considine, 2012; Cresci et al., 2015), and in some instances to circulate/promote links to fake news sites, unreputable stories, and unverified rumors (i.e., Ferrara, 2016; Persily, 2017).

Ferrara (2018) notes that social bots have played a role in electioneering since at least as early as 2010, when "during the ... U.S. midterm elections, social bots were employed to support some candidates and smear others, by injecting thousands of tweets pointing to websites with fake news" (p. 2). Previously, a number of political actors have also used bots to artificially boost their social media followings in an effort to project popularity and influence (Considine, 2012; Cresci et al., 2015). However, most Americans first became familiar with this terminology following the 2016 presidential election, in which bots were found to be responsible for as much as one fifth of all election-related comments posted on Twitter and other social networking sites (Ferrara, 2018). Additionally, the extensive use of social bots was found to be one piece of a larger strategy on the part of Russian-based actors, aimed at influencing the U.S. elections through the promotion of critical stories and the spread of misinformation. For example, bots were used by Russian actors to amplify stories critical of Hillary Clinton during the 2016 election (Persily, 2017). In other instances, Russian-based actors have been accused of using similar tactics in an effort to sway and polarize political discussion networks in the United States and other Western democracies (i.e., Broniatowski et al., 2018; *The Times*, 2017).

The potential for social bots to distort the measurement of public sentiment—upon which responsive governments depend—has been flagged by many scholars and commentators as a profound threat to the effective practice of democracy, both in the United States and throughout the world (i.e., Ferrara et al., 2016; Persily, 2017; Stella et al., 2018). Along with promoting false or unreputable information, it's been suggested that the effect of these activities could be to (1) drown out legitimate grassroots/minority interest movements; (2) ascribe false legitimacy to fringe ideas by creating the impression of widespread support; and (3) possibly even to overwhelm critical information channels with noise, spam, and propaganda during crisis events and public emergencies (Ferrara et al., 2016; Woolley, 2016). In each case, the potential misuse of these technologies poses significant dangers to the public exchange of information and ideas.

As these emerging technologies are increasingly employed to influence democratic processes around the world, understanding the extent to which bots are active in political discussion networks—as well as their effect on the content, quality, and nature of political discourse—is increasingly critical. To the degree that healthy, well-functioning democracies depend on the consent and participation of an informed citizenry, there may be much at stake in this endeavor. With these considerations in mind, this article provides a case study analysis of a high-profile political discussion network (on Twitter) in order to deepen our understanding of how prominent social bots are in political discussions as well as how they influence key network attributes such as structure, sentiment, and influentialness. The section that follows briefly identifies the research questions that guided this case study analysis. It is proceeded by a summary of the case study and the data collection techniques.

Research Questions

Social bots are believed to be both common and active on social networks such as Twitter. For example, a study conducted by the Pew Research Center suggested that "an estimated two-thirds of tweeted links to popular websites are posted by automated accounts—not human beings" (Wojcik et al., 2018, p. 2). More advanced social bots are capable of imitating human behavior and executing most or all Twitter functions. For instance, they can generate new content and tweets, retweet messages from other users, proactively follow and/or unfollow accounts, send direct messages, and even "like" tweets. Twitter's application programming interface (API) makes bot programming easy and affordable to anyone with basic coding experience (Guilbeault & Woolley, 2016).

Due to the sophistication with which bots mimic human behavior, they are often difficult to identify without technical and methodological expertise. Additionally, bots can be automatic or have human curation (semiautomatic), which makes them even more difficult to distinguish for most users (Howard & Kollanyi, 2016). On top of these challenges, due to the anonymity afforded by platforms like Twitter, the bot generator's identity and location are typically undiscoverable (Ferrara, 2016; Persily, 2017). Over recent years, significant advances have been made in bot detection techniques, and several studies have used these tools to examine the prevalence of bots in high-profile instances such as the 2016 Brexit referendum election (Bastos & Mercea, 2018, 2019). However, it is still often unclear how prevalent or active bots are in political discussion networks, particularly in nonelectoral contexts.

With these concerns in mind, using emerging bot detection techniques, we begin by examining the prevalence of social bots in the Twitter discussion network surrounding the Special Counsel investigation into Russian interference in the 2016 presidential election.

Research Question 1: How prevalent are social bots in the political discussion network?

As discussed above, social bots can exhibit a variety of technical capabilities including (but not limited to) amplifying themes and messages (Lokot & Diakopoulos, 2016; Persily, 2017), creating and widely distributing information (Bolsover & Howard, 2018), and filtering as well as curating information for niche audiences (Arif et al., 2018; Geiger, 2016). In recent years, these capabilities have been exploited for a variety of purposes including legitimate efforts to spread authoritative scientific information as well as illegitimate attempts to deceive audiences through "astroturfing" (i.e., creating artificial trends; Persily, 2017). In light of these practices, significant concerns have arisen over the potential for bots to endanger democratic processes by distorting public preferences, corrupting informational channels, and even amplifying misinformation (Ferrara et al., 2016). However, the extent to which these efforts are undertaken, as well as their effectiveness at influencing public opinion, remains unclear in many instances. In an effort to better understand the influence of bot activity on a political discussion network, we next examine several ways in which bots influence the flow of information within the sampled discussion network. In doing so, we examine several important network features, including network structure, sentiment, and influentialness. This analysis is guided by three overarching research questions:

Research Question 2: In what ways—if any—do bots influence the structure and organization of the discussion network?

Research Question 3: In what ways—if any—do bots influence the "influentialness" of actors within the discussion network?

Research Question 4: In what ways—if any—do bots influence the prevailing tone or "sentiment" in the discussion network?

One particular way in which bots are believed to affect political discourse is by giving voice to nontraditional actors such as marginalized groups or those holding extreme/fringe viewpoints. Recent empirical studies have suggested that bots may be widely used as a means of amplification by nontraditional media sources and/or fringe actors. For example, during the 2016 US presidential election, around 48% of circulated news and information was found to be from "alternative" media sources such as junk news sites, WikiLeaks, or even Russian actors (Howard et al., 2017). In contrast, only 25% of content was produced by professional news organizations, while less than 4% was created by government agencies, political parties/candidates, or subject matter experts.

One particularly interesting observation from these studies has been the identification of common features that social bots frequently employ in their tweets, including an abundance of "junk news," which often relies on the use of " . . . attention grabbing techniques, lots of pictures, moving images, excessive capitalization, ad hominem attacks, emotionally charged words and pictures, unsafe generalizations and other logical fallacies" (Howard et al., 2017, p. 3). With this in mind, we also consider the extent to which bots are used to promote nontraditional information sources/ideas.

Research Question 5: To what extent—if any—do social bots appear to be amplifying content from nontraditional media sources?

Data and Methods

For this study, we conducted a mixed method case study analysis using Twitter data collected between July and August of 2017. Using data mining approaches (community detection, bot detection, network analysis, and sentiment analysis), the data were analyzed to answer the research questions posed above. In the subsections below, we will introduce the case as well as describe the data collection and methodology in detail.

Case Study and Data Collection

According to an intelligence community report published in 2017, during the United States' 2016 presidential election, the Russian government "aspired to help President-elect Trump's election chances when possible by discrediting Secretary Clinton and publicly contrasting her unfavorably to him" (Office of the Director of National Intelligence, 2017, p. ii). In this effort, Russian-based actors adopted diverse cyber activities including the use of paid social media users or "trolls." The Department of Justice appointed the Special Counsel Robert Mueller, on May 17, 2017, to investigate these activities as well as potential links or coordination between Russian-based actors and individuals working for the Trump Campaign. The investigation sparked widespread discussion across a variety of social media platforms including Twitter. We collected Twitter data pertaining to this discussion for 1 month, starting on July 25, 2017, during which time Special Counsel Mueller appointed attorneys to investigate the case.

Using the Twitter API, we collected Twitter data during the specified date range using a broad list of key words that were found to be associated with the investigation. These included:

"TrumpRussia", "TrumpRussiaCoverup", "FireproofMueller", "RobertMueller", "Mueller", "trump", "trump-russia", "Robert Mueller", "Paul Manafort", "Donald Trump Jr.", "Russia", "Lavrov", "investigation"[2]

Following conventions established by Barbera (2015), we dropped Twitter accounts with less than 25 followers and with less than 100 followings. Hagen et al. (2018) found that this filtering did not cause major biases to the subsequent analyses. We parsed English tweets only for the analysis. As a result, a total of $N = 13{,}360{,}648$ tweets and 10,204,244 retweets were present in the data. Due to the extreme volume, we sampled the first 2 weeks of data, from July 25, 2017, to August 8, 2017, for this analysis. In order to conduct the network analysis, we created edge lists using *retweet relations*. Retweet relations are arguably a better measure than *followers* or *mentions* for this type of analysis because retweets reflect user interactions that are driven by the content value without directly addressing actors as mentions or follower relations would do (Boyd et al., 2010; Cha et al., 2010).

Each node in the network is a Twitter account that retweeted at least one tweet related to any of the terms included in the query list during the study window. From these nodes and edge lists, we created a directed graph of the network. A total of 1,605,589 nodes and 7,532,332 edges were generated, from which we created the subgraph using only the top 99.9% degree in order to make visualization and network analysis feasible. This means that our network analysis focuses on the most active retweeters—those who most heavily influence network attributes and the circulation of information within the network. We tested for possible differences in the results by filtering only the top percentage and did not find notable distinctions. This filtering process yielded a total of 2,500 nodes (unique Twitter handles) and 110,455 edges for the network analysis. We used Gephi (Version 0.9.1), an open-source software for data analysis and network visualization (Bastian et al., 2009), and we used R for the rest of the analyses (R Core Team, 2015).

Mixed Method Approach

In order to conduct this case study analysis, we employed a mixed methods approach that included five distinct data analysis techniques/strategies. (1) First, we employed a community detection algorithm to define the network structure and identify unique communities based on retweet relations. (2) Second, we utilized a series of bot detection algorithms to identify the likelihood of each node in the network being a social bot. (3) Third, we calculated several commonly used centrality measures to identify influential actors in each of the detected communities. (4) Fourth, we conducted a sentiment analysis in order to better understand the tone and content of the information communicated in the network as a whole as well as in each individual community. (5) And finally, we conducted content analysis to describe the categories of user profiles in order to identify the types of actors who were most influential in the discussion network.

In each instance, we conducted a "before-and-after" analysis that included the network both prior to and following the removal of accounts identified as likely social bots. This allows us to better understand how social bots influence each of these key features and attributes of the discussion network. The details pertaining to each portion of the analysis are discussed in the following subsections.

Detecting social bots. In order to detect social bots within the discussion network, we utilized the "Botometer" program. Botometer is a bot detection algorithm that uses supervised machine learning methods to assist in determining the likelihood that a Twitter account is actually a social bot (Varol et al., 2017). Features used for the algorithm training include user-based characteristics (i.e., number of friends and followers, number of tweets, profile description, and settings), friends features (i.e., retweeting, mentioning, being retweeted, and being mentioned), network features (weighted degree, density, and clustering), temporal features (such as "average rates of tweet production over various time periods and distributions of time intervals between events"), content and language features (i.e., part-of-speech tagging, and length/entropy of tweet text), and sentiment features (Varol et al., 2017). Botometer provides "bot likelihood scores" that indicate the probability of a Twitter account being a social bot (Varol et al., 2017). Wojcik et al. (2018) suggested that a Twitter account with a Botometer score of 0.43 or higher is likely to be a social bot. Botometer is trained on thousands of manually annotated Twitter data, and the accuracy of correctly detecting bot accounts has been identified at approximately 86% (Wojcik et al., 2018).

For the purposes of comparison and confirmation, we also employed a second bot detection algorithm called "tweetbotornot" (Kearney, 2018) and observed similar results, thereby providing some level of verification for the Botometer results. In the Findings section, we report the results acquired using Botometer only.

Network structure and community detection. Virtual communities within social networks can be detected and identified based on observing patterns of interactions with the aid of *modularity* algorithms. Modularity is a measure of network structure, which groups and divides nodes into modules based on the density of interactions within the network. Communities detected by modularity have dense connections between nodes *within* the module and have sparse connections between nodes *outside* of the module. For this analysis, we initially tested with a total of six modularity algorithms including (1) clustering with eigenvectors (Newman, 2006), (2) walktrap (Pons & Latapy, 2006), (3) Louvain (Blondel et al., 2008a), (4) near-linear time algorithm (Raghavan et al., 2007), (5) maps of information flow (Rosvall & Bergstrom, 2007), and (6) fast greedy algorithm (Clauset et al., 2004). Three algorithms (fast greedy, maps of information flow, and walktrap) identified two large, distinctive communities, while the other two algorithms (Label and Louvain) identified three distinctive communities, and eigenvector clustering identified only one large community.

For the purposes of the analysis, we decided to use the Louvain algorithm as it is one of the most widely adopted methods for community detection due in part to its easy implementation and high-quality results (Blondel et al., 2008b; Ji et al., 2015). Additionally, in this instance, we found that the Louvain algorithm enabled a more granular rendering of the network's structure and polarity by detecting three distinctive communities, while most of the other algorithms detected one or two distinctive communities.

Modularity is an unsupervised clustering method, which does not include manually annotated community values to guide the learning process. To test that the three clusters were not created by random chance, we ran the modularity algorithm 10 times using random seeds and the default setting of the Louvain algorithm embedded in Gephi. We found that the topologies of the three dominant communities (responsible for over 94% of the nodes) were nearly identical during each iteration.

In order to construct an effective visual presentation of the detected communities, we initially tested three distinct spatial representation algorithms (ForceAtlas2, Fruchterman–Reingold, and Yifan Hu), with each displaying similar spatial distinctions. We report community detection results using ForceAtlas2 (Jacomy et al., 2014) as it is well suited for a spatial representation of the polarized structure of a network.

Identifying influential actors and levels of connectivity. Measures of centrality are used to understand the influentialness of nodes in a complex network. We used three distinct centrality measures (degree centrality, eigenvector centrality, and PageRank) to identify influential actors in the network communities. First, we applied degree centrality to identify nodes that were retweeted frequently. The equation for calculating degree centrality is as follows, where $d(n_i)$ is the degree (number of retweets) of ith node:

$$\text{Degree Centrality}(n_i) = d(n_i).$$

In our study, higher degree centrality means that the tweets created by a node were frequently retweeted; therefore, an entity with high degree centrality exhibits a greater tendency to influence others.

Second, while degree centrality focuses simply on the number of connections a node has, eigenvector centrality acknowledges that "not all connections are equal" (Newman, 2008). By this reasoning, the number of connections to highly influential nodes reflects greater influence for an account. Applying this concept to our data, when a tweet is frequently retweeted by important neighboring nodes, the node, which created the original tweet, is likely to have a high information spreading power (Canright & Engø-Monsen, 2006). Eigenvector centrality considers both the frequency and quality of a node's connections to detect influential nodes. The centrality x_i of a node i is proportional to the sum of the centralities of its neighbors, and the equation for this is as follows:

$$\lambda x_i = \sum_{j=1}^{n} A_{ij} x_j = (A^t x)_i,$$

where "x_i is the i component of the eigenvector of the transpose of the adjacency matrix with eigenvalue λ" (Yan & Ding, 2009, p. 4).

Third, PageRank is a variant of eigenvector centrality, formulated by Brin and Page (1998), which indicates the reliability or trustedness of a node (Caverlee et al., 2008; Giménez-Garcıa et al., 2016). In web searches, a website is considered to be highly endorsed if it has high number of *incoming links* by other important pages. For example, when two nodes have an equal number of in-links, the node with incoming links from more "important" nodes have larger PageRank. A simplified formula of PageRank is as follows (Page et al., 1999, p. 4):

$$PR(u) = c \sum_{v \in B_u} \frac{PR(v)}{N(v)}.$$

PR is PageRank of a webpage u. Bu is the set of pages pointing to (in-links) u. v is all webpages contained in Bu. $N(v)$ is the number of links from page v. c is a factor used for normalization in order to keep the total rank of all the pages to be constant (Page et al., 1999). A webpage has high rank when the sum of the ranks of its in-links is high. Similarly, in our data set, a node with a high PageRank is highly endorsed by others because its content is frequently recirculated by important nodes.

Finally, we can measure how closely actors are connected to each other in a community using a clustering coefficient. For example, the clustering coefficient of node A measures the extent to which the neighboring nodes of A form a densely clustered clique. The clustering coefficient is based on the idea that there is an increased likelihood of two users becoming "friends" when they share a common friend. Higher clustering coefficients in a network show stronger connections among actors in that community. This is closely related with a small-world phenomenon; a network concept that most nodes can be reached from each other by a relatively small number of links because neighbors of any given node are likely to be connected to each other (Watts & Strogatz, 1998). When community members are closely connected to each other (high clustering coefficient), there is a high level of redundancy in connections, which enables the rapid circulation of information within the community.

Measuring sentiment. Sentiment analysis can be used as a means of measuring the intensity of attitudes and preferences expressed in digital settings such as Twitter. While Twitter is not directly representative of the voting population, Oliveira et al. (2017) showed that sentiments expressed on the social networking platform were highly similar to those expressed in traditional opinion polls. Another study on Twitter during 2016 U.S. presidential election found that the prevailing sentiments in tweets associated with Donald Trump were consistently more positive than those associated with Hillary Clinton (Yaqub et al., 2017).

In this study, we used an automatic sentiment detection algorithm, SentiStrength, which was developed specifically for the purpose of analyzing Twitter content. SentiStrength is a dictionary-based classifier that incorporates linguistic information and rules to measure the sentiment strength in "short informal English text" such as tweets (Thelwall et al., 2012). Using the full data set, we conducted a sentiment analysis on a total of 627,721 tweets. After bot removal, we dropped any tweets created by the bots ($N = 217{,}112$) and analyzed sentiment for the remaining 410,609 tweets for comparison.

Content analysis of user profiles. Lastly, in order to better understand how social bots shape influentialness in the network, we conducted a content analysis by classifying influential actors by type, both before and after the bot removal. This portion of the analysis was conducted using publicly available user profile data as well as subsequent web searches where necessary.

Findings

Figure 1 provides a graphical depiction of the initial retweet network. Smith et al. (2014) note that network maps are valuable as descriptive tools in that they "... can provide insights into the role social media plays in our society" (p. 4), in part by helping us to visually gauge the level of polarization in a discussion network. In this case, the discussion network surrounding the Special Counsel investigation into Russian election interference fits the *polarized crowd* network structure as indicated by the presence of two distinct, densely populated clusters (aka modularity classes). In describing this network structure, Smith et al. (2014) note that "Polarized crowds on Twitter are not arguing. They are ignoring one another while pointing to different web resources and using different hashtags" (p. 2). These network structures are similar to those identified in previous analyses of contentious political discussion networks (i.e., Barberá, 2015; Conover et al., 2011; Del Vicario et al., 2017), and they appear to be consistent with the "echo chamber" hypothesis (Sunstein, 2007), suggesting that social media may contribute to the polarization and fragmentation of civic discourse by facilitating selective exposure to congenial information.

In the network depicted by Figure 1, three major communities accounted for approximately 94% of all participating nodes. Based on a manual analysis of tweets created by accounts in each of the three communities, we defined the community on the left side of the graph as *left-leaning* (politically liberal), while the communities on the right side were classified as politically conservative or *right-leaning* communities. Following this observation, we named the three major communities as L, R1, and R2 (see Figure 1). The community L includes 60% of all nodes in the network and is located on the left side of the graph. The right side of the graph contains about 34% of all nodes. R2 is topologically the greatest distance from L, emphasizing that the community appears to be ideologically further to the right on the political spectrum than both the R1 and L communities.

Prevalence of Bots in the Discussion Network (Research Question 1)

After constructing the network graph, we ran each node (i.e., Twitter account) through the Botometer bot detection algorithm. The results suggested that 23% of all active nodes (i.e., accounts) in the discussion network exhibited a high likelihood of being bots. In total, these nodes accounted for a slightly disproportionate 35% of all retweets in the network. The

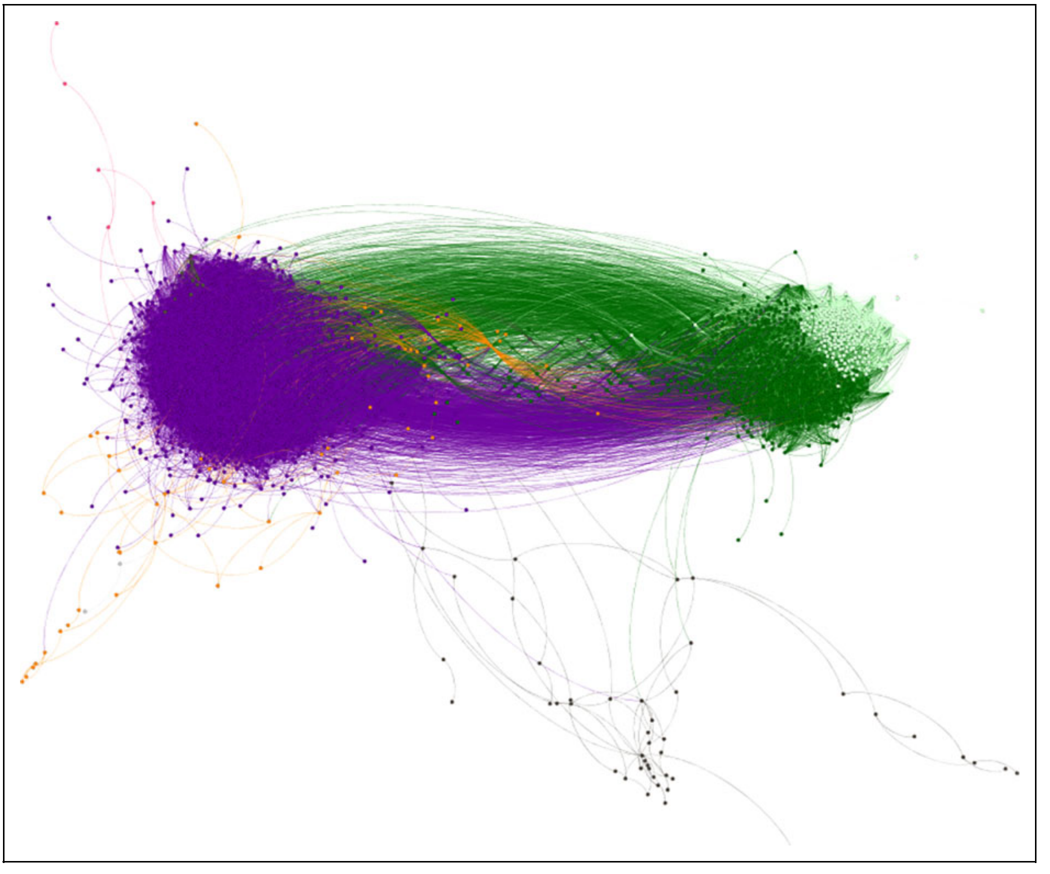

Figure 1. Initial network structure (community detection). *Note*. Colors represent different communities. Edges explain retweet relations between nodes.

concentration of bot accounts was found to be most dense in the R2 (far-right) community, where as many as 63% of the accounts exhibited a high likelihood of being bots. Figure 2 shows that the median bot probability for the far-right community (R2) was significantly higher than the other two communities with the left-leaning community having the lowest median bot probability score.

Given the higher prevalence of likely bots in the far-right community, we would anticipate seeing marked differences in Twitter behavior between the communities. Table 1 (before bot removal section) reports descriptive statistics on common Twitter behaviors for each of the three initial network communities. When viewing these data, the far-right community (R2) stands out quite noticeably as containing the youngest (most recently created) accounts, while also exhibiting an extremely high use of retweets, mentions, and hashtags compared to the other groups. The *Retweet/ Tweet Ratio* for the three groups demonstrates the extent to which the behavior of nodes in the three communities differs (Table 1, before bot removal). For example, the ratios for communities L (1.38) and R1 (1.27) were relatively similar, while nodes in R2 had an extremely high ratio of 9.32, which means that actors in this community retweeted other messages 9 times for each unique tweet that they created. As would be expected in a community with a high number of bots, these accounts seem to be more aggressively taking actions intended to amplify messages as opposed to creating original content for circulation.

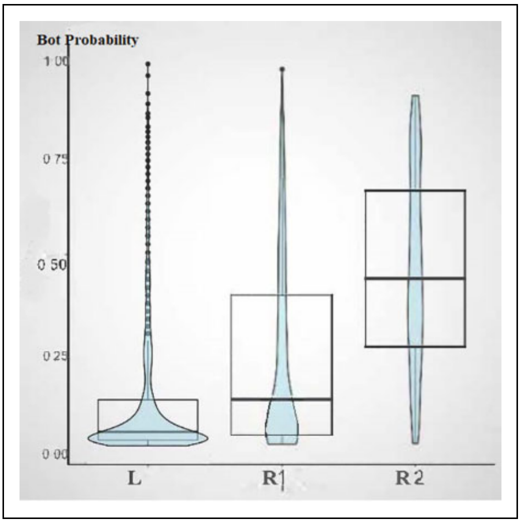

Figure 2. Bot detection results. *Note. Y*-axis depicts Botometer scores for each of the three communities.

Table 1. Descriptive Statistics by Community (Before and After Bot Removal).

Community	No. of Assigned Nodes (%)	Average Account Age (Days)	Retweet to Tweet Ratio	URLs per Account	Hashtags per User	Mentions per User	No. of Likely Bots (%)
Before bot removal							
L (left-leaning)	1,404 (60)	2,819	1.38	145	34	159	114 (12%)
R1 (right-leaning)	546 (26)	2,347	1.27	201	154	308	130 (39%)
R2 (far-right)	150 (8)	2,066	9.32	357	369	748	77 (63%)
After bot removal							
L1.wo.bot (left-leaning)	714	2,640	1.33	166	45.6	195	—
L2.wo.bot (left-leaning)	640	3,083	1.20	72	6.0	66	—
R.wo.bot (right-leaning)	684	2,527	2.18	152	120.3	258	—

Influence of Bots on the Discussion Network

Next, in order to investigate the influence of these likely bot accounts on the political discussion network (Research Question 2–Research Question 5), we conducted a before-and-after analysis of the network using both the complete data set and a subset of data created by removing all likely bot accounts. The threshold for eliminating an account from the analysis was a Botometer score of greater than 0.43 (for discussion, see Wojcik et al., 2018). We conducted an extensive network analysis to explore differences in the various communities before and after the bot removal. As set forth in Research Questions 2–5, this included community detection, an analysis of influentialness, levels of connectivity, and sentiment analysis. The results of these analyses are presented in the following subsections.

Effect of bots on network structure (Research Question 2). First, after removing likely bot accounts, we implemented the same clustering algorithm initially used to detect the three major communities. The results changed notably as shown in Figure 3B. This time, two left-leaning communities emerged, while the far-right community (R2) essentially disappeared from the network once likely bot

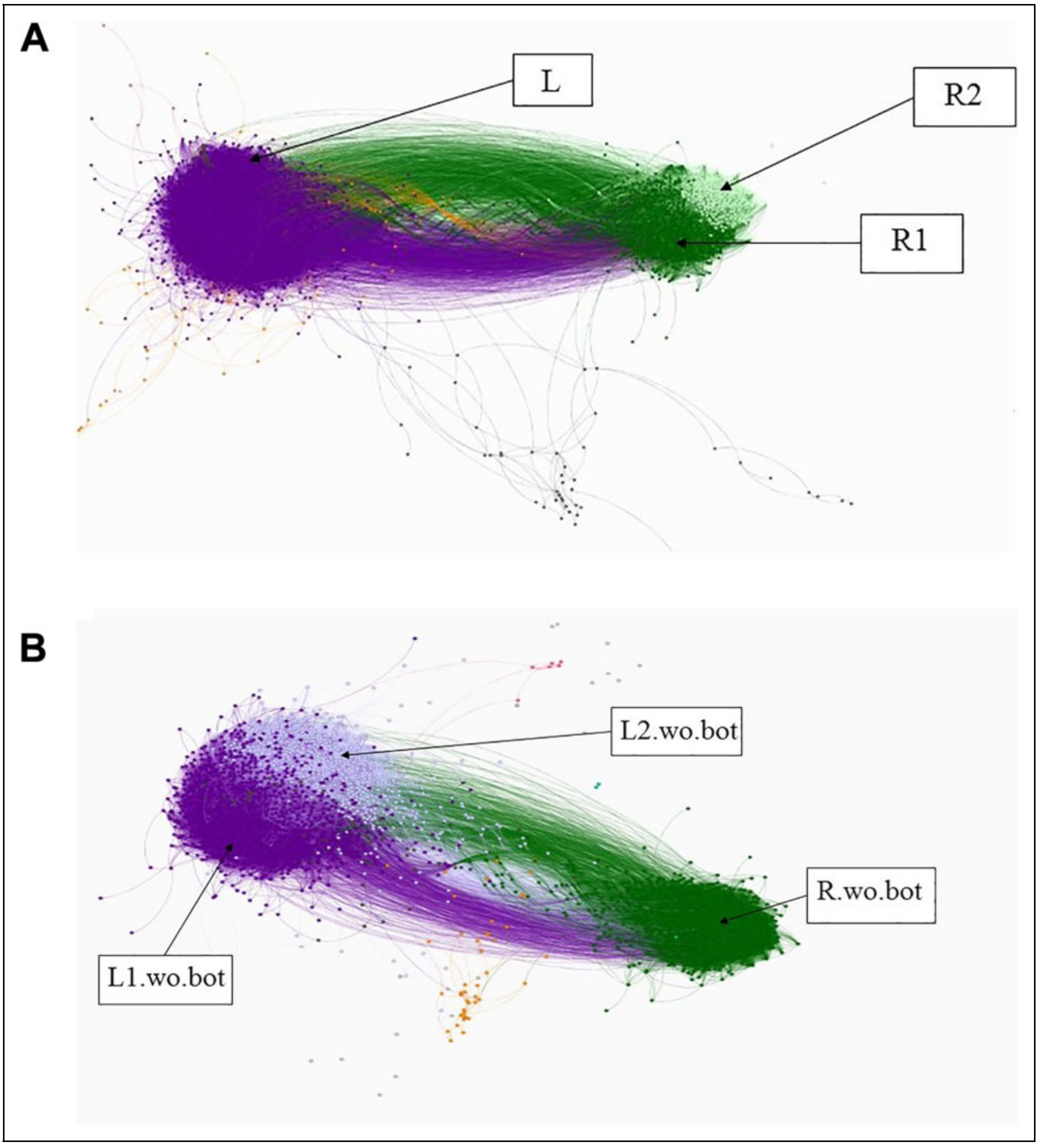

Figure 3. Community detection results (before and after bot removal). (Panel A) Community structure: Before bot removal. (Panel B) Community structure: After bot removal. *Note*. Colors represent different communities. Edges explain retweet relations between nodes.

accounts were removed from the analysis. We labeled the three communities after the bot removal as L1.wo.bot, L2.wo.bot (the two left-leaning communities), and R.wo.bot (the right-leaning community). These results suggest that bots are capable of significantly influencing the topology of a discussion network. In this instance, the effect was significant enough that the likely bot accounts appear to have created an artificial cluster, which initially might signify or portend a sense of community on particular individuals or ideas, where in fact one did not exist.

The before-and-after analysis shows that bot removal did not fundamentally alter the overall nature of the discussion network, which maintained a "polarized crowd" structure after likely bot accounts were removed (Figure 3). As was the case previously, the left-leaning communities

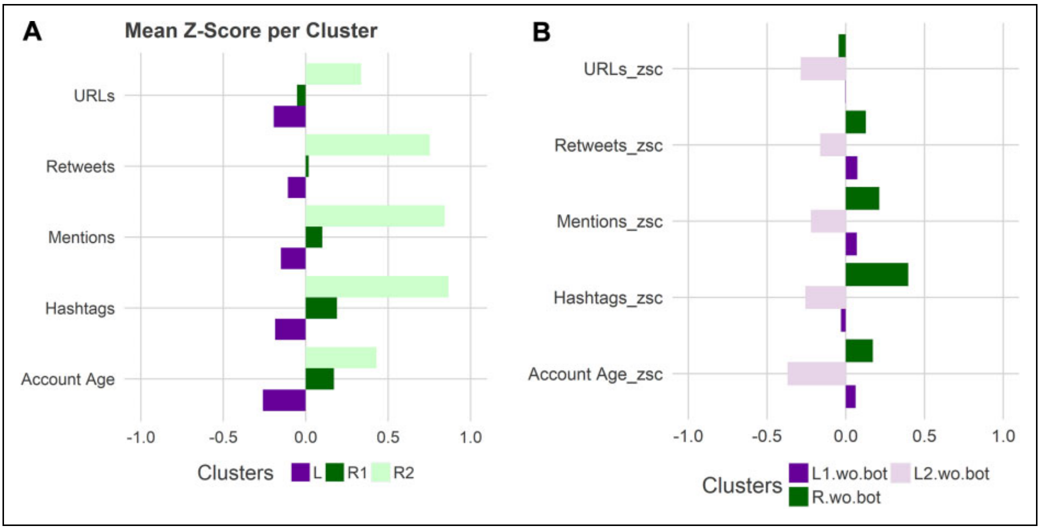

Figure 4. User behaviors across network communities (before and after bot removal). (Panel A) Before the bot removal. (Panel B) After the bot removal. *Note*. Bars represent z-scores for each cluster compared against the network mean. Higher acc_create_zsc (z-scores of account creation date) values reflect more recent dates.

continued to account for approximately 65% of all nodes in the network (a slight increase from the before analysis), while the right-leaning community accounted for approximately 30% (a slight decrease). The most notable differences were the elimination of the far-right community in the after analysis as well as the division of the left-leaning community into two slightly more unique clusters (L1.wo.bot and L2.wo.bot).

Table 1 provides a descriptive summary of changes in tweeting behavior in the network after bot removal (broken down by community). Additionally, Figure 4 presents *z*-scores to depict changes in key behaviors after likely bot accounts were removed from the network. On the whole, we found that bots contributed to creating and amplifying messages within a virtual far-right cluster. Before the bot removal, accounts in the far-right community (R2) used URLs, retweets, mentions, and hashtags much more frequently than the other two communities (Table 1 and Figure 4). Additionally, accounts assigned to the far-right community (R2) were notably younger than those in the left-leaning community (L1). After the bot removal, these distinctions are largely (though not entirely) eliminated, and behaviors in the remaining right-leaning community became more consistent with those in the left-leaning communities.

Effect of bots on influentialness (Research Question 3 and Research Question 5). The results also suggested that bots had a substantial effect on influentialness in the initial network model. For example, the PageRank and degree centrality measures (Figure 5) suggest that bots distorted the measurement of trust (PageRank) and influentialness (measured by degree centrality as well as eigenvector centrality) prior to their removal. This effect was most pronounced in the far-right community (R2). The clustering coefficients and eigenvector centrality of the right-leaning communities, before and after the bot removal, were higher than those of the left-leaning communities. This means that the right-leaning community was more densely connected to each other (clustering coefficients) with higher average levels of popularity among actors in the network (eigenvector centrality) than the liberal communities. (This also indicates that right-leaning users seemed to more broadly utilize the technological tools available through Twitter, at least as it pertains to this case study.)

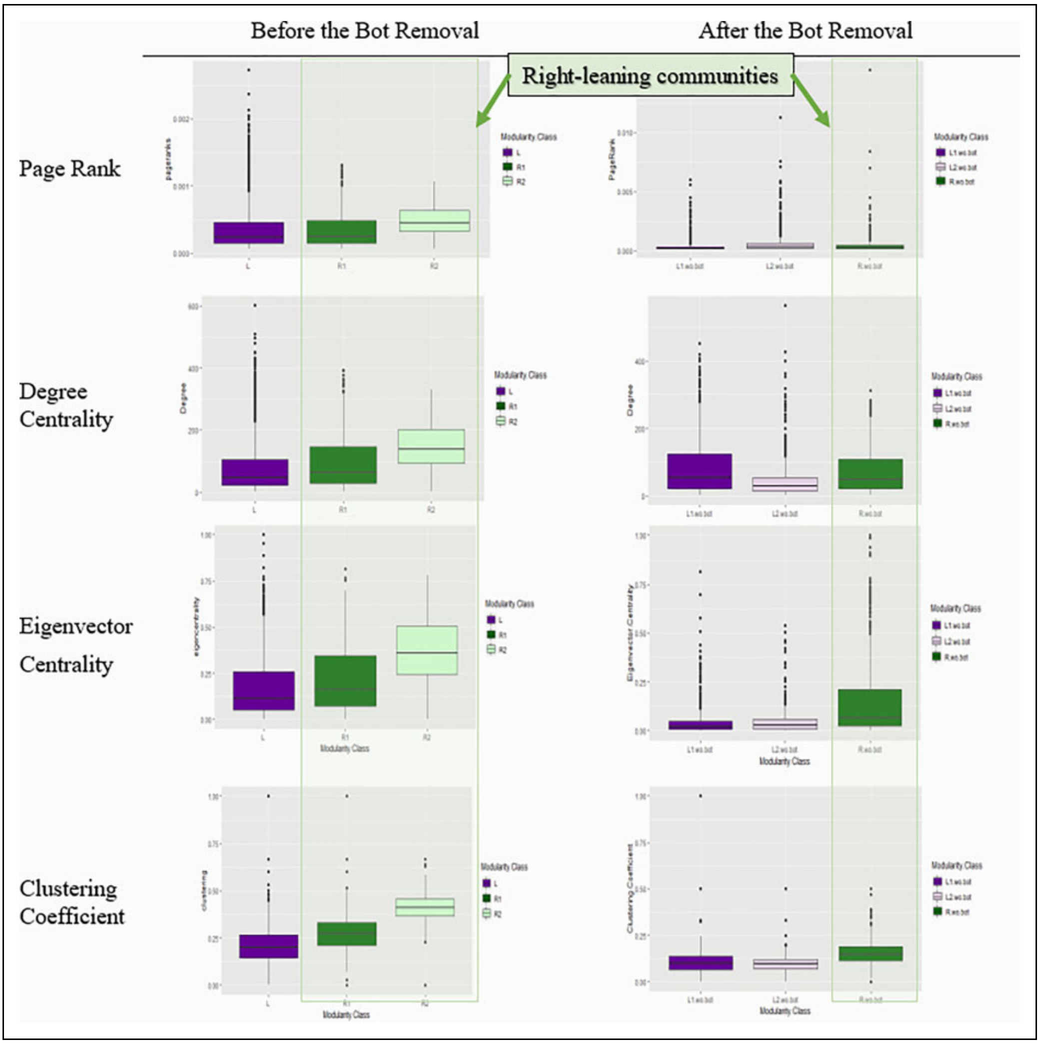

Figure 5. Centrality and clustering coefficient statistics (before and after bot removal). *Note*. R1 and R2 on the second column and R.wo.bot on the third column are right-leaning communities.

In order to better understand how bots impacted influentialness in the discussion network, we then conducted manual content analyses of user profiles for the 30 Twitter accounts with the highest PageRank (both before and after the bot removal). We used PageRank for this analysis because it reflects the level of trust that an individual node enjoys in a given community. Prior to the bot removal, the 30 accounts with the highest PageRank all belonged to the left-leaning community, and the majority of these accounts (about 85%) did not disclose their identity (Supplemental Table S1). However, after the bot removal, a majority of the most highly trusted actors changed. Among the 30 most trusted accounts postbot removal, 26 accounts were in the left-leaning community and 4 accounts (*wikileaks*, *JulianAssange*, *FoxNews*, and *lukerosiak*) were situated in the right-leaning community. And the sweeping majority of the accounts (98%) disclosed their identity (Supplemental Table S1). Below are the detailed findings from before and after the bot removal.

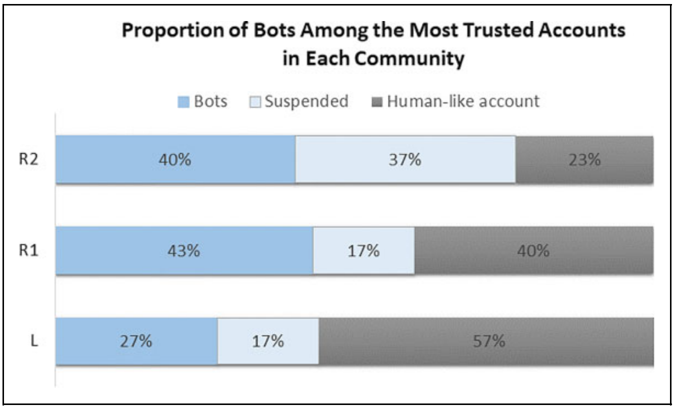

Figure 6. Proportions of bots among the 30 most trusted accounts in each community (before bot removal: $N = 30$ per community)

First, we examined the 30 accounts with the highest PageRank *from each community* before the bot removal. In doing so, we analyzed the composition of likely bot accounts, suspended accounts, and human-like accounts. In addition, we investigated occupational information for the human-like accounts to better understand the types of actors who were influential in the network. We found that the majority of the 30 accounts in the right-leaning communities (R1:77% and R2: 60%) were either likely bots or suspended accounts (Figure 6). The percentage of bots among the trusted accounts (37%) was higher than the average rate of bots (23%) in the network. This means that bots were disproportionately represented among the most influential accounts in the discussion network. Figure 6 shows that the left-leaning community (L) contained the highest proportion of human-like accounts (57%). However, even among the human-like accounts, a large majority (over 80% in each community) did not provide occupational information (Supplemental Table S1). Only two media organizations and one MSNBC producer were identified among the highly trusted actors. There are no government agencies, political parties/candidates, or subject matter experts identified among this group. This means, these Twitter users in this network exhibited high levels of trust in accounts that were often not identifiable as reputable or authoritative sources of information.

Second, we then repeated this manual analysis for the 30 actors with the highest PageRank after the bot removal. After the bot removal, only two accounts were suspended among the top 30 (PageRank) accounts from each of the three communities. Additionally, *wikileaks* became the most trusted node in the network (Figure 7). Bot removal made a noticeable change in other areas as well. The most trusted nodes of the two left-leaning communities (L1.wo.bot and L2.wo.bot) after bot removal included a larger number of major news media, subject matter experts, and political activists. Interestingly, highly trusted accounts in the right-leaning community included several alternative media sources (online and nontraditional media outlets). For example, *wikileaks*, *JulliaNAssange*, *PrisonPlanet* (a news media run by Alex Jones, a conspiracy theorist, and a member of alt-right), and *WestmonsterUK* (an online news platform to support Brexit) were found to be highly trusted alternative media sources, all of whom were included in the right-leaning community (see Figure 7). Supplemental Table S1 provides a more detailed comparison of the most influential actors in the network prior to and after the bot removal. Collectively, these data suggest that bots obscured the influence of several traditional actors and institutions in the course of amplifying unrecognizable and/or alternative information sources (Research Question 5).

Rank by PageRank	Account Name
1	wikileaks
2	cnni
3	cnnsport
4	JulianAssange
5	PeterAlexander (NBCNews national correspondence)
6	maggieNYT (White House correspondent for NYTimes)
7	FoxNews
8	CNN
9	NBCNews
10	CNNAfrica

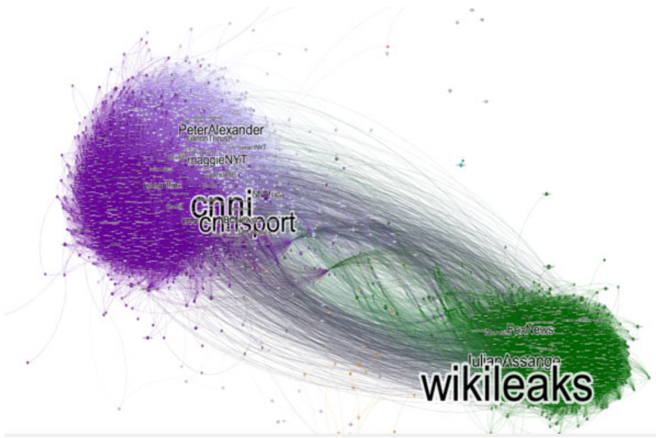

Figure 7. Network graph by PageRank (after bot removal). *Note*. The size of the account labels reflects the quantity of PageRank value such that larger usernames are correlated with higher trust.

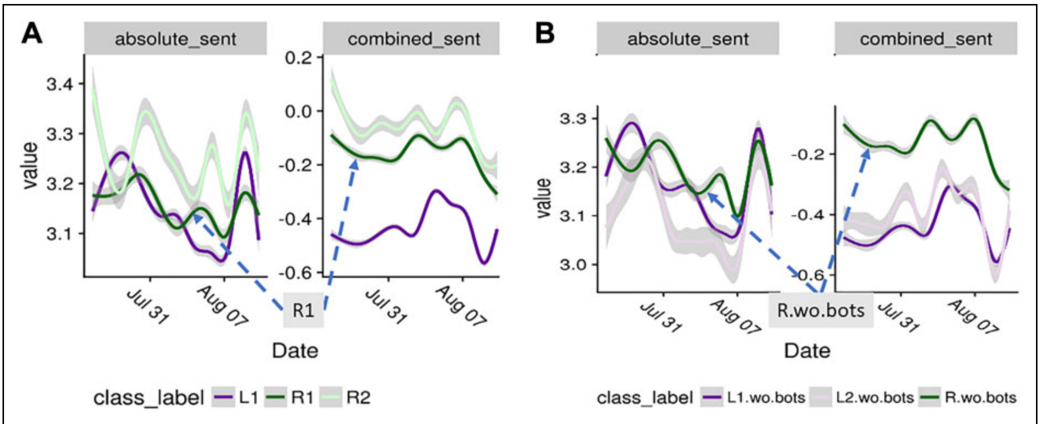

Figure 8. Sentiment analysis results (before and after bot removal). (Panel A) Before bot removal. (Panel B) After bot removal. *Note*. SentiStrength separates sentiment into positive and negative emotion. The *y*-axis depicts the average value of sentiment for the community. Scales are independent from one panel to next in order to show finer detail and variation along time. "absolute_asent" is the absolute sentiment of the accumulated sentiment from each community, and it therefore shows the intensity of sentiment. "combined_sent" is the accumulated results from positive and negative sentiment scores, which shows mean value of sentiment.

Effect of bots on sentiment in the discussion network (Research Question 4). Sentiment analysis is used to measure the "tone" or "intensity of attitudes" expressed by participants in a discussion network. *SentiStrength* is a sentiment analysis program designed specifically for the analysis of Twitter data. It distinguishes between positive and negative emotional sentiments. Figure 8 depicts differences in sentiment across the various network communities both with and without the likely bot accounts. The horizontal (x) axis depicts time, while scores are depicted on the y-axis, with more positive sentiment registering higher on the vertical axis. *Absolute sentiment* ("absolute_asent") reflects the level of absolute sentiment in each community (i.e., intensity of emotive language), and *combined sentiment* ("combined_sent") reflects the accumulated results from positive and negative sentiment scores, thereby showing the mean value of sentiment expressed in each community. Figure 8A shows that R2 exhibited the most positive (combined_sent) and highly emotional (absolute_sent) sentiment compared to the other two communities prior to the bot removal (Figure 8A). The left-leaning community (L) was highly negative in tone when discussing the Special Counsel investigation.

After dropping likely bot accounts, the intensity of sentiment expressed in the network was noticeably lower ("absolute_sent" in Figure 8B) and the overall sentiment was more negative ("combined_sent" in Figure 8B). Collectively, these data suggest that bot accounts were primarily amplifying very positively phrased, pro-Trump messaging prior to their removal. Overall, the sentiment level of left-leaning communities stayed at a relatively similar level (and remained primarily negative in tone) after dropping bots from the network. In contrast, the right-leaning community (postbot removal) reflected the more modest tone and sentiment originally associated with the R1 community, as opposed to the more positive and emotive tone of the R2 community. In sum, the before and after data suggest that bots can substantially influence the overall tone and sentiment of a political discussion network. It's important to note that while the positive/pro-Trump messaging originally circulated in the R2 community may not have directly "reached" users in the L or even R1 communities, it did substantially impact the overall measurement of key metrics in the network.

Discussion

In this study, we investigated the prevalence and influence of social bots in an online, political discussion network. Specifically, we examined a large volume of Twitter data pertaining to the Special Counsel investigation of Russian interference in the 2016 U.S. presidential election. The study was guided by several specific research questions including:

Research Question 1: How prevalent are social bots in the political discussion network?

Research Question 2: In what ways—if any—do bots influence the structure and organization of the discussion network?

Research Question 3: In what ways—if any—do bots influence the prevailing tone or "sentiment" in the discussion network?

Research Question 4: In what ways—if any—do bots influence the "influentialness" of individual actors within the discussion network?

Research Question 5: To what extent do social bots appear to be amplifying content from nontraditional media sources?

Our initial analysis showed that 23% of all participating accounts in the discussion network exhibited a high likelihood of being social bots with these accounting for 35% of all tweets in the sample. Notably, these numbers are larger than those found in prior analyses of the 2016 presidential election, where likely bots accounted for roughly 15% of the total Twitter population and were responsible for approximately 19% of the total tweet volume (Bessi & Ferrara, 2016). However, since we sampled highly active accounts with relatively high retweet frequency, the higher proportion of bots in our sample makes sense because bots tend to more actively engage with retweeting activities (Stella et al., 2018). Our analysis represents only one isolated case study, and as such, we cannot definitively say that this perceived increase is reflective of a real change in bot activity. However, these results may suggest a prevalent presence and use of social bots on Twitter during July and August 2017.

Specifically, our findings showed that the presence of bots was extremely high in the "far-right" community, where 63% of the active accounts exhibited a high likelihood of being bots. These bots artificially inflated the intensity of Twitter activity in the far-right community by amplifying the messages and viewpoints shared by members of this group. For example, the accounts in this community/cluster used a significantly higher volume of hashtags, mentions, and retweets, with the average account in this cluster retweeting more than 9 times for each new, original tweet created. Additionally, the content circulated within this community was generally very "positive" in sentiment, suggesting the widespread use of "pro-Trump" language. This is in contrast to Stella et al. (2018) that bots were used to evoke negative sentiments in the case of the Catalan referendum for independence in 2017. This suggests that bots are used to evoke either positive or negative emotions depending on the political contexts.

After removing likely bot accounts from the analysis, we found that bots had influenced key attributes of the discussion network in several important ways. For instance, while the network remained highly polarized even after the removal of likely bot accounts, the far-right community disappeared from the network in the "after" analysis, suggesting that bots contributed to the intensity of polarization in the network by amplifying "fringe" elements within one ideological cluster. One effect of this was to give the appearance of a vibrant and populated virtual community where one did not exist.

Bots also had a notable influence on the "sentiment" expressed within the discussion network. Overall sentiment scores became notably lower once likely bot accounts were removed from the analysis, suggesting that bots may have distorted the initial measurement of public sentiment by understating public displeasure over the issue of Russian interference in the 2016 election. In addition, bots had an apparent impact on influentialness in the discussion network, as the before-and-after analysis showed that bots had inflated centrality measures (and thus the measured level of trust) for many actors in the network, particularly in the far-right community. This finding is particularly noteworthy as traditional media outlets and subject matter experts only registered as "influential" network actors *after* likely bot accounts were removed from the data set. Prior to this, the majority of "highly trusted" accounts in the right-leaning communities were either likely bots or suspended accounts. This suggests that bots essentially obscured or "overpowered" the influence of traditional actors (such as mainstream media and subject matter experts) by artificially boosting the visibility and influentialness of unknown Twitter accounts.

Lastly, it has been suggested that bots help to amplify the voice of "outsiders" or fringe actors, making users believe their networks are more extensive, their ideas more popular, and their spokespeople more trusted than they actually are. Our study found that the far-right community in particular used bots extensively toward this end when compared to other groups or clusters. This finding is in line with a study by Marwick and Lewis (2017) where the authors found that "conspiracy theorists, Men's Rights advocates, trolls, anti-feminists, anti-immigration activists, and bored young people," which they labeled as "far-right" online groups, leveraged both the participatory culture and affordance of social media to widely circulate their beliefs (Marwick & Lewis, 2017, p. 3). These techniques are sometimes known as "attention hacking," wherein actors seek to "increase the visibility of their ideas through the strategic use of... bots—as well as by targeting journalists, bloggers, and influencers to help spread content" (Marwick & Lewis, 2017).

Conclusion

The results of this case study analysis underscore several key concerns that have been raised over the serious impacts of bots on political communications and democratic processes. For example, Guilbeault and Woolley (2016) have suggested that bots distort naturally occurring citizen engagement and democratic communications while solidifying polarization among citizens. Our findings appear to support this contention as discussed above.

This study contributes to the literature in the areas of technology (especially artificial intelligence) and democracy. Our findings demonstrate that during political discussions on social media, users are presented with information environments possibly manipulated by bots. More specifically, bots appear to often be utilized by actors with ideological positions that are reflective of only a small subset of the public (i.e., the far-right, nationalist community) in order to amplify their message. It should be noted that this raises concerns over stability in our social structure; bots and artificial intelligence technologies can be adopted for spreading misinformation resulting in "undermining democratic processes by fostering doubt and destabilizing the common ground that democratic societies require" (Starbird, 2019, p. 449).

Reluctantly or not, social media platforms are "the new intermediary institutions for our present politics" (Persily, 2017, p. 74) as we have witnessed through the successful use of social media during the Brexit referendum and the 2016 U.S. presidential election. Social media platforms necessarily mediate (informationally, algorithmically, and technologically) politics by selecting information to present and to make available for end users (Graham & Dutton, 2019). As was shown in our findings, smaller communities tend to implement bots to amplify their messages. However, when these niches "promote fragmentation, especially along political lines," it could bring a risk of group polarization, which could easily move toward extreme positions (Sunstein, 2018, p. 151). A similar concern regarding the same technologies is that the void created by decreasing trust in, and power of, traditional institutions may be filled by unmediated ideologies with the ability to utilize Internet technologies (Marwick & Lewis, 2017; Persily, 2017).

It should be emphasized that the case study approach used in this study is both a strength and a weakness of the analysis. We focused specifically on a subset of tweets surrounding a specific topic during a limited period of time. Therefore, the findings should be interpreted with the appropriate caveats in mind, and additional research focusing on a range of cases should be undertaken. As Papacharissi (2016) stated, "[s]ocial media presence does not convey the same impact for all issues, publics, and movements" (p. 312). Depending on ideology, goals, and context of the communication, social media presence and use varies. This is why a case study like ours can reveal specific characteristics of political communities bounded in a particular theme. We need further validations and added case analyses to come up with a better understanding of the generalizable behavior of social bots in political discussion networks as well as their possible impact on democratic processes. We also recommend that future research directly consider the influence of bots not only on the network structure and content but also on subsequent political behaviors by those individuals exposed to bot-generated content.

Data Availability

The collected Twitter data are prohibited to be distributed (Twitter API Developer Policy: https://developer.twitter.com/en/developer-terms/agreement-and-policy.html; https://gwu-libraries.github.io/sfm-ui/posts/2017-05-18-twitter-policy-change).

For replicability of our study, we will provide the list of tweet IDs for noncommercial research purposes in this link (https://github.com/TrumpRussiaInvestigationTwitterResearch/trump-russia-bot-analysis) or by responding to email requests (tekeller@usf.edu). With the provided tweet IDs, users can extract tweets and retweets including the queries specified in the article in order to create similar network created in our study.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Loni Hagen was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2017S1A3A2066084).

Software Information

The major code is available online (https://github.com/TrumpRussiaInvestigationTwitterResearch/trump-russia-bot-analysis)

- R Version 3.6.1 (https://www.r-project.org/).
  - The major R library used for the analysis: readr, rtweet, dplyr, lubridate, couplot, dplyr, botcheck, httr, xml2, RSJONIO, purr, and ggplot2.
- Java SentiStrength from sentiment analysis: Thelwall et al. (2012).
- Botometer: Varol et al. (2017).
- Gephi Version 0.9.2 (https://gephi.org/): Bastian et al. (2009).

Supplemental Material

The supplemental material is available in the online version of this article.

Notes

1. While not directly relevant to the study at hand, it's important to note that social bots have also been prominent and raised concerns in other contexts such as nonpolitical discussion networks related to finance (Cresci et al., 2019) and health (Broniatowski et al., 2018) communications.
2. A full summary of the data collection process is available from the authors upon request.

References

Arif, A., Stewart, L. G., & Starbird, K. (2018). Acting the part: Examining information operations within #BlackLivesMatter discourse. In *Proceedings of the ACM on Human-Computer Interaction* (Vol. *2*) CSCW, November 2018. ACM. https://doi.org/10.1145/3274289

Barberá, P. (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political Analysis*, *23*(1), 76–91.

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. *International Conference on Web and Social Media*, *8*, 361–362.

Bastos, M. T., & Mercea, D. (2018). The public accountability of social platforms: Lessons from a study on bots and trolls in the Brexit campaign. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2128), 20180003. https://doi.org/10.1098/rsta.2018.0003

Bastos, M. T., & Mercea, D. (2019). The Brexit botnet and user-generated hyperpartisan news. *Social Science Computer Review*, *37*(1), 38–54. https://doi.org/10.1177/0894439317734157

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. Presidential election online discussion. *First Monday*, *21*(11). https://doi.org/10.5210/fm.v21i11.7090

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008a). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008b, October 9). *The Louvain method for community detection in large networks*. https://perso.uclouvain.be/vincent.blondel/research/louvain.html

Bolsover, G., & Howard, P. (2018). Chinese computational propaganda: Automation, algorithms and the manipulation of information about Chinese politics on Twitter and Weibo. *Information, Communication & Society*, *0*(0), 1–18. https://doi.org/10.1080/1369118X.2018.1476576

Boyd, D., Golder, S., & Lotan, G. (2010). Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter. *System Sciences (HICSS), 2010 43rd Hawaii International Conference On*, 1–10. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5428313

Bradshaw, S., & Howard, P. N. (2018). *Challenging truth and trust: A global inventory of organized social media manipulation* (p. *26*). Oxford Internet Institute.

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, *30*(1–7), 107–117. https://doi.org/10.1016/S0169-7552(98)00110-X

Broniatowski, D. A., Jamison, A. M., Qi, S., AlKulaib, L., Chen, T., Benton, A., Quinn, S. C., & Dredze, M. (2018). Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. *American Journal of Public Health*, *108*(10), 1378–1384. https://doi.org/10.2105/AJPH.2018.304567

Canright, G. S., & Engø-Monsen, K. (2006). Spreading on networks: A topographic view. *Complexus*, *3*(1–3), 131–146. https://doi.org/10.1159/000094195

Caverlee, J., Liu, L., & Webb, S. (2008). Socialtrust: Tamper-resilient trust establishment in online communities. *Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries*, 104–114. http://dl.acm.org/citation.cfm?id=1378908

Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, P. K. (2010). Measuring user influence in Twitter: The million follower fallacy. *International Conference on Web and Social Media*, *10*(10–17), 30.

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, *70*(6), 066111. https://doi.org/10.1103/PhysRevE.70.066111

Conover, M., Ratkiewicz, J., Francisco, M. R., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on Twitter. *International Conference on Web and Social Media*, *133*, 89–96.

Considine, A. (2012, August 22). Buying their way to Twitter fame. *The New York Times*. https://www.nytimes.com/2012/08/23/fashion/twitter-followers-for-sale.html

Cresci, S., Fabrizio, L., Daniele, R., Serena, T., & Maurizio, T. (2019). Cashtag piggybacking: Uncovering spam and bot activity in stock microblogs on Twitter. *ACM Transactions on the Web*, *13*(2).

Cresci, S., Roberto, D. P., Marinella, P., Angelo, S., & Maurizio, T. (2015). Fame for sale: Efficient detection of fake Twitter followers. *Decision Support Systems*, *80*(1), 56–71.

Crothers, B. (2019, April 22). Internet bots are getting better at imitating humans. *New York Post*. https://nypost.com/2019/04/22/internet-bots-are-getting-better-at-imitating-humans/

de Zuniga, H. G., Jung, N., & Valenzuela, S. (2012). Social media use for news and individuals' social capital, civic engagement and political participation. *Journal of Computer-Mediated Communication*, *17*(3), 319–336.

Del Vicario, M., Zollo, F., Caldarelli, G., Scala, A., & Quattrociocchi, W. (2017). Mapping social dynamics on Facebook: The Brexit debate. *Social Networks*, *50*, 6–16. https://doi.org/10.1016/j.socnet.2017.02.002

Ferrara, E. (2016, November 8). How Twitter bots affected the US presidential campaign. *The Conversation*. http://theconversation.com/how-twitter-bots-affected-the-us-presidential-campaign-68406

Ferrara, E. (2018). Measuring social spam and the effect of bots on information diffusion in social media. In S. Lehmann & Y. Ahn (Eds.), *Complex spreading phenomena in social systems: Influence and contagion in real-world social networks* (pp. 229–255). Springer.

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM*, *59*(7), 96–104.

Geiger, R. S. (2016). Bot-based collective blocklists in Twitter: The counterpublic moderation of harassment in a networked public space. *Information, Communication & Society*, *19*(6), 787–803. https://doi.org/10.1080/1369118X.2016.1153700

Giménez-Garcıa, J. M., Thakkar, H., & Zimmermann, A. (2016). Assessing trust with PageRank in the web of data. *PROFILES 2016 3rd International Workshop on Dataset PROFIling and FEderated Search for Linked Data*. http://ceur-ws.org/Vol-1597/PROFILES2016_paper5.pdf

Graham, M., & Dutton, W. H. (Eds.). (2019). *Society and the internet: How networks of information and communication are changing our lives* (2 edition). Oxford University Press.

Greenwood, S., Perrin, A., & Duggan, M. (2016, November 11). *Social media update 2016*. Pew Research Center.

Guilbeault, D., & Woolley, S. (2016, November 1). How Twitter bots are shaping the 2016 presidential election. *The Atlantic*. https://www.theatlantic.com/technology/archive/2016/11/election-bots/506072/

Hagen, L., Keller, T., Neely, S., DePaula, N., & Robert-Cooperman, C. (2018). Crisis communications in the age of social media: A network analysis of Zika-Related Tweets. *Social Science Computer Review*, *36*(5), 523–541. https://doi.org/10.1177/0894439317721985

Howard, P. N., Bolsover, G., Kollanyi, B., Bradshow, S., & Neudert, L.-M. (2017). *Junk news and bots during the U.S. Election: What were Michigan voters sharing over Twitter?* https://comprop.oii.ox.ac.uk/research/working-papers/junk-news-and-bots-during-the-u-s-election-what-were-michigan-voters-sharing-over-twitter/

Howard, P. N., & Kollanyi, B. (2016). Bots, #StrongerIn, and #Brexit: Computational Propaganda during the UK-EU Referendum. *ArXiv:1606.06356 [Physics]*. http://arxiv.org/abs/1606.06356

Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS One*, *9*(6), e98679. https://doi.org/10.1371/journal.pone.0098679

Ji, X., Machiraju, R., Ritter, A., & Yen, P.-Y. (2015). Examining the distribution, modularity, and community structure in article networks for systematic reviews. In *AMIA Annual Symposium Proceedings* (pp. 1927–1936). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4765615/

Kearney, M. W. (2018). *tweetbotornot: R package for detecting Twitter bots via machine learning*. https://github.com/mkearney/tweetbotornot

Kim, K., Sin, S. J., & Tsai, T. (2014). Individual differences in social media use for information seeking. *The Journal of Academic Librarianship*, *40*(2), 171–178.

Leask, D. (2017). Fake Twitter accounts send 400,000 independence messages. *The Herald*, November 19, 2017.

Lokot, T., & Diakopoulos, N. (2016). News bots. *Digital Journalism*, *4*(6), 682–699. https://doi.org/10.1080/21670811.2015.1081822

Marwick, A., & Lewis, R. (2017). *Media manipulation and disinformation online* (p. 106). Data & Society.

Mitchell, A., Rosenstiel, T., & Christian, L. (2012). What Facebook and Twitter mean for news. *The Pew Research Center*. http://www.pewresearch.org/2012/03/19/state-of-the-news-media-2012/

Newman, M. E. J. (2006). Finding community structure in networks using the eigenvectors of matrices. *Physical Review E*, *74*(3), 036104. https://doi.org/10.1103/PhysRevE.74.036104

Newman, M. E. J. (2008). The mathematics of networks. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.131.8175&rep=rep1&type=pdf

Office of the Director of National Intelligence. (2017). *Background to "Assessing Russian activities and intentions in recent US elections": The analytic process and cyber incident attribution*. Office of the Director of National Intelligence.

Oliveira, D. J. S., Bermejo, P. H. de S., & Santos, P. A. dos. (2017). Can social media reveal the preferences of voters? A comparison between sentiment analysis and traditional opinion polls. *Journal of Information Technology & Politics*, *14*(1), 34–45. https://doi.org/10.1080/19331681.2016.1214094

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank citation ranking: Bringing order to the web. http://ilpubs.stanford.edu:8090/422

Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, *6*(2), 259–283.

Papacharissi, Z. (2016). Affective publics and structures of storytelling: Sentiment, events and mediality. *Information, Communication & Society*, *19*(3), 307–324. https://doi.org/10.1080/1369118X.2015.1109697

Persily, N. (2017). The 2016 U.S. Election: Can democracy survive the internet? *Journal of Democracy*, *28*(2), 63–76. https://doi.org/10.1353/jod.2017.0025

Pons, P., & Latapy, M. (2006). Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications*, *10*, 191–218.

R Core Team. (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Raghavan, U. N., Albert, R., & Kumara, S. (2007). Near linear time algorithm to detect community structures in large-scale networks. *Physical Review E*, *76*(3), 036106. https://doi.org/10.1103/PhysRevE.76.036106

Rosvall, M., & Bergstrom, C. T. (2007). Maps of information flow reveal community structure in complex networks. In *Proceedings of the National Academy of Sciences USA*, 1118–1123.

Smith, M. A., Rainie, L., Himelboim, I., & Shneiderman, B. (2014). *Mapping Twitter topic networks: From polarized crowds to community clusters*. https://www.pewresearch.org/wp-content/uploads/sites/9/2014/02/PIP_Mapping-Twitter-networks_022014.pdf

Starbird, K. (2019). Disinformation's spread: Bots, trolls and all of us. *Nature*, *571*, 449–449. https://doi.org/10.1038/d41586-019-02235-x

Stella, M., Ferrara, E., & Domenico, M. D. (2018). Bots increase exposure to negative and inflammatory content in online social systems. Proceedings of the National Academy of Sciences, *115*(49), 12435–12440. https://doi.org/10.1073/pnas.1803470115

Stewart, M. C., & Wilson, B. G. (2016). The dynamic role of social media during hurricane #sandy: An introduction to the STREMII model to weather the storm of the crisis lifecycle. *Computers in Human Behavior*, *54*, 639–646.

Sunstein, C. R. (2007). *Republic.Com 2.0*. Princeton University Press.

Sunstein, C. R. (2018). *#Republic: Divided democracy in the age of social media*. Princeton University Press.

Thelwall, M., Buckley, K., & Paltoglou, G. (2012). Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology*, *63*(1), 163–173. https://doi.org/10.1002/asi.21662

The Times. (2017, November 15). Russian dolls; Through bots and disinformation, Foreign powers have sought to affect our democracy. The Legal and Technical Loopholes which Allowed this Must be Closed. https://www.thetimes.co.uk/article/russian-dolls-krq8s05hv

Varol, O., Ferrara, E., Davis, C. A., Menczer, F., & Flammini, A. (2017). Online human-bot interactions: Detection, estimation, and characterization. In *Proceedings of the Eleventh International AAAI Conference on Web and Social Media* (ICWSM2017) (pp. 280–289).

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918

Wojcik, S., Messing, S., Smith, A., Rainie, L., & Hitlin, P. (2018, April 9). *Bots in the Twittersphere*. http://www.pewinternet.org/2018/04/09/bots-in-the-twittersphere/

Woolley, S. C. (2016). Automating power: Social bot interference in global politics. *First Monday*, *21*(4). https://doi.org/10.5210/fm.v21i4.6161

Woolley, S. C., & Howard, P. N. (2016). Automation, algorithms, and politics| Political communication, computational propaganda, and autonomous agents—Introduction. *International Journal of Communication*, *10*(0), 9.

Yan, E., & Ding, Y. (2009). Applying centrality measures to impact analysis: A coauthorship network analysis. *Journal of the American Society for Information Science and Technology*, *60*(10), 2107–2118. https://doi.org/10.1002/asi.21128

Yaqub, U., Chun, S. A., Atluri, V., & Vaidya, J. (2017). Analysis of political discourse on Twitter in the context of the 2016 US presidential elections. *Government Information Quarterly*, *34*(4), 613–626. https://doi.org/10.1016/j.giq.2017.11.001

Zeifman, I. (2017, January 24). *Bot Traffic Report 2016| Imperva*. https://www.imperva.com/blog/bot-traffic-report-2016/

Author Biographies

Loni Hagen (lonihagen@usf.edu) is an assistant professor at the School of Information at the University of South Florida. Prior to her current position, she worked for a Korean government over 10 years. Her primary research combines method development for computational analysis with the application of these methods for evidence-based decision-making. Her current domains of interest are ethics/issues with artificial intelligence, extreme online communities, privacy, cybersecurity, and health emergency communication.

Stephen Neely (srneely@usf.edu) is an associate professor in the University of South Florida's School of Public Affairs. He offers courses in public policy, media politics, and research methodology. His current research focuses on the emerging role of social media in the public sphere as well as on public affairs education.

Thomas E. Keller (tekeller@usf.edu) is a computational biologist working in the Genomics Program at the University of South Florida. His primary research themes involve applying machine learning to genomics scale problems. Additionally, he enjoys broad collaborations that involve the use of social networks and textual analysis.

Ryan Scharf (ryan.scharf@gmail.com) is a risk analyst at American Integrity Insurance Company. He received a degree in information studies from the University of South Florida and concentrated on data science and analytics. His current work relates to leveraging meteorological data to forecast claims for rate making and reinsurance purchasing.

Fatima Espinoza Vasquez (Fatima.Espinoza@uky.edu) is an assistant professor at the School of Information at the University of Kentucky. She earned her PhD in information science at Syracuse University. Her research explores how political actors negotiate the meaning, use, and configuration of information and communications technologies (ICTs). Her research provides a comprehensive account of the social, political, and material contexts in which ICTs are being used. Her multidisciplinary, multi-method, and decolonial approach provides insight into how ICTs are appropriated and repurposed by political actors under various contexts.