Bots in Nets: Empirical Comparative Analysis of Bot Evidence in Social Networks

Ross Schuchard[1(✉)], Andrew Crooks[1,2], Anthony Stefanidis[2,3], and Arie Croitoru[2]

[1] Computational Social Science Program, Department of Computational and Data Sciences, George Mason University, Fairfax, VA 22030, USA
rschucha@gmu.edu
[2] Department of Geography and Geoinformation Science, George Mason University, Fairfax, VA 22030, USA
[3] Criminal Investigations and Network Analysis Center, George Mason University, Fairfax, VA 22030, USA

Abstract. The emergence of social bots within online social networks (OSNs) to diffuse information at scale has given rise to many efforts to detect them. While methodologies employed to detect the evolving sophistication of bots continue to improve, much work can be done to characterize the impact of bots on communication networks. In this study, we present a framework to describe the pervasiveness and relative importance of participants recognized as bots in various OSN conversations. Specifically, we harvested over 30 million tweets from three major global events in 2016 (the U.S. Presidential Election, the Ukrainian Conflict and Turkish Political Censorship) and compared the conversational patterns of bots and humans within each event. We further examined the social network structure of each conversation to determine if bots exhibited any particular network influence, while also determining bot participation in key emergent network communities. The results showed that although participants recognized as social bots comprised only 0.28% of all OSN users in this study, they accounted for a significantly large portion of prominent centrality rankings across the three conversations. This includes the identification of individual bots as top-10 influencer nodes out of a total corpus consisting of more than 2.8 million nodes.

Keywords: Bots · Online social networks · Social network analysis

1 Introduction

The increased dependency on online social networks (OSNs) for information and the unprecedented ability to instantaneously message global populations provides an opportunity to control or exploit the narrative of online conversations. Attempting to control or exploit the narrative of a certain topic becomes much easier in OSNs as ‘digital gatekeepers’ can employ social bots—computer algorithms designed to mimic human behavior and interact with humans in an automated fashion—to amplify a specific position or drown out its opposition at scale. This includes increasing the

L. M. Aiello et al. (Eds.): COMPLEX NETWORKS 2018, SCI 813, pp. 424–436, 2019.
https://doi.org/10.1007/978-3-030-05414-4_34

spread of fake news by orders of magnitude through a directed bot campaign [1]. The evolvement of social bot sophistication is a primary concern, as it has become very hard for humans to discern whether they are engaging in dialogue with a human or a bot [2]. Given that recent studies estimate that social bots account for 9–15% of all Twitter accounts [3, 4], it is essential to understand the implications associated with human and machine dialogue, either intentional or not.

Recent social bot research continues to build initial essential knowledge on the classification and detection of social bots [4–8]. However, the establishment of social bot norms is difficult and predictively elusive given the evolving nature of bot sophistication. For this reason, studies continue to discover bot activity that does not align with previously published conceptions [9]. Beyond the necessary continued work associated with improved bot detection methods to move closer to ground truth discovery, there is also a growing need to present novel evaluation methodologies to better understand the effects of currently detected bots within social media conversations. Promising recent studies applying multidisciplinary approaches to social bot analysis include classifying bot emotion [10], determining the political agenda of bots [11] and distorting political discourse with bots [12–14].

In this paper, we present a unique methodological framework to comparatively analyze evidence of social bots found within OSN Twitter conversations about three major global events in 2016: (1) the United States Presidential Election, (2) the Ukraine Conflict and (3) Turkish Online Political Censorship. First, we conducted a comparative descriptive statistical analysis of these Twitter conversations to determine the characteristics of human and social bot tweeting patterns. We then sought to determine the relative influence of social bots by applying social network analysis techniques to each of the associated conversation's constructed retweet networks. In total, we evaluated more than 30.4 million tweets generated by 5.2 million distinct Twitter users, of which, we recognized 14,661 users as bots responsible for 2.1 million tweets.

The results of this study showed that social bot communication patterns were fairly consistent across the various observed online conversations. We found bots to have a higher engagement rate than humans for both in-group and cross-communication. Most interestingly, although online conversation participants recognized as social bots comprised only 0.28% of all OSN users in this study, they accounted for a significantly large portion of prominent centrality rankings across the three online conversations. In total, this work provides a new contribution to the growing study of social bots by applying social network analysis techniques across multiple online conversations to help determine the relative pervasiveness and importance of detected bots.

2 Related Work

The term bot has broad meaning in the context of technology and Internet applications, since all automated services or applications could be construed as bots. For the purpose of this paper, we restrict our definition of bots, or social bots, to automated software or computer algorithms designed to mimic human behavior and/or engage with human actors within online social networks. Many recent works have contributed to the growing corpus of knowledge capturing social bot features that differentiate social-bot generated activity from human-generated activity in OSNs [6, 7, 15].

Some researchers have not only published their research on bot detection methodologies and findings but have also transitioned their work to open-source bot detection platforms for other researchers to use via a web application or an application programming interface (API). Davis et al. [7] provide access to *Botometer* (formerly known as *BotOrNot*), which assesses the likelihood of a Twitter account being a bot by using a supervised Random Forest applied to extracted account features. Chavoshi et al. [5] published *DeBot*, which employs an unsupervised warped correlation model to detect Twitter bots rather than feature extraction.

Published research analyzing detected bots in specific OSNs has increased as the prevalence of bots has risen. Such studies include examining bot evidence in the following use-cases: the 2016 U.S. presidential election [4, 16], Venezuelan political public opinion [13], the Syrian civil war [9], the Brexit Referendum [14], the Ukrainian conflict [11, 17] and Russian politics [8]. Most methodologies are limited to initial descriptive statistical and temporal analyses of the human versus bot tweet volumes. Although highly relevant contributions, these efforts focus on single events. As Kušen and Strembeck [10] point out in their recent analysis of bot emotion across multiple events, bot studies focused on sole events make it difficult to generalize findings across this growing topic of interest.

3 Methodology

In order to understand the patterns of bots across multiple global events and determine the relative bot impact within associated online conversations, this study employed a combination of comparative descriptive statistical analysis and social network analysis applications. This multi-faceted approach expands the literature of social bot analysis by comparatively analyzing multiple OSN use-cases and contributes new techniques to the field of bot research by adapting social network analysis methods to measure and define the impact or influence of social bots. The remainder of this section will present in detail the methodology steps used in this study as depicted in Fig. 1.

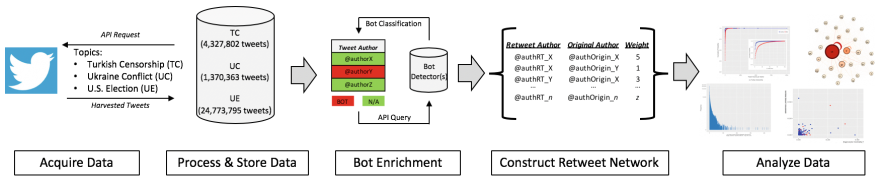

Fig. 1. Overall methodology to analyze bot evidence across multiple Twitter OSN conversations.

3.1 Data

This study focused on three major global online conversations harvested solely from Twitter in 2016. Summarized descriptions of each event conversation are as follows: *(1) U.S. Presidential Election (Feb. 1–29, 2016):* a one-month period which captured the narrative surrounding the Republican and Democratic party primary races prior to the U.S. general election when it became apparent that then-candidate Donald Trump could win his party's nomination, *(2) Ukraine Conflict (Aug. 1–31, 2016):* a one-month period which captured the narrative surrounding the ongoing conflict in Ukraine as military activity and political rhetoric intensified between Russia and Ukraine around the 25th anniversary of Ukrainian independence from Russia, *(3) Turkish Political Censorship (Dec. 1–31, 2016):* a one-month period which captured Turkish political conversations before, during and after two distinct periods of censorship when the Turkish government banned Turkish citizens from using Twitter.

We crafted and submitted relevant key words for each of these events to extract associated tweets from the Twitter Standard Search API. The volumes of tweets returned were as follows: 24.8 million (U.S. Presidential Election), 1.4 million (Ukraine Conflict), 4.3 million (Turkish Censorship). Given the resulting large tweet volumes, all initial data storage and pre-processing for normalization took place in an Amazon Web Services EC2 t2.2xlarge instance (8 vCPUs/32GiB). This allowed for rapid processing and the creation of individual graph objects for more rapid data analysis use at the local compute level.

3.2 Bot Enrichment

To determine the presence of bots within the acquired Twitter conversations, we leveraged the DeBot open-source bot detection platform [5]. Our decision to use DeBot was two-fold. First, our corpus of tweets came from 2016, so we required access to historical bot evidence, which only DeBot currently provides. Second, the performance of DeBot's unsupervised warped correlation process has outperformed other bot detection platforms to date [18]. To determine bot presence, we extracted tweet author names from our harvested tweet corpus and submitted them for classification via the DeBot API. We then merged the returned results with our existing database and labeled each tweet user as a bot (or not) and annotated the source of bot classification. We purposely created automated scripts to execute this enrichment phase with the hope of accounting for other bot detection services in the future.

In total, this enrichment process classified 14,661 Twitter users as bots, which accounted for just 0.28% of total tweet corpus users. This relatively small population of users classified as bots was responsible for publishing 2.1 million tweets, or 6.8% of all tweets in this study. Table 1 provides detailed values for each event conversation.

Table 1. Harvested Twitter Corpus Overview

Corpus	Tweets	Retweets	Users
United States Election	24,773,795	14,321,387	3,472,114
Bot Source *(% of total)*	1,882,809 *(7.60%)*	1,452,155 *(10.14%)*	6,875 *(0.20%)*
Ukraine Conflict	1,370,363	681,806	383,237
Bot Source *(% of total)*	55,718 *(4.07%)*	34,938 *(5.12%)*	2,486 *(0.65%)*
Turkey Censorship	4,327,802	2,837,059	1,390,362
Bot Source *(% of total)*	126,352 *(2.92%)*	83,582 *(2.95%)*	5,300 *(0.38%)*

3.3 Construct Retweet Network

Retweets accounted for 57.8% of all tweets in this study, with the Turkey Censorship conversation exhibiting the highest retweet density at 65.6%, followed by 57.8% for the U.S. Election conversation and 49.8% for the Ukraine Conflict conversation. The parsed retweets from the originally harvested tweets served as the basis for the construction of retweet networks for each conversation. These resulting retweet networks serve as the primary artifacts required to examine the conversation via social network applications that include centrality analysis and community detection.

To reveal the network structure from the harvested Twitter conversations, we constructed retweet networks for each of the events in this study. The act of a Twitter user 'retweeting' a message of an originally authored tweet establishes the basis for an edge between two nodes, or users, in the retweet network. Specifically, when a Twitter user (X) retweets an original tweet message from a given user (Y), then we assign a directed edge weight value of 1 for initial retweets or add to the cumulative weight for existing edges. The resulting directed networks for each of the conversations were as follows: 2,557,805 nodes / 8,985,736 edges (U.S. Election), 250,541 nodes / 537,459 edges (Ukraine Conflict), 1,075,833 nodes / 2,224,939 edges (Turkish Censorship).

3.4 Analyze Data

The final phase of this study's methodology was the application of a multi-faceted data analysis approach to the processed data from the three online conversations. Recall that the main purpose of this work was to identify potential common characteristics of social bots across multiple online conversations and ascertain any in-group (bot-to-bot) or cross-group (bot-to-human/human-to-bot) tendencies. Additionally, we sought to classify the overall relative importance of bots within the conversations by examining bot positions within the social structure of the retweet networks and associated bot membership within any emergent communities of said networks. Section 4 follows with detailed subsections discussing the specific methods used to achieve the purpose described above.

4 Results and Discussion

4.1 Bot and Human Participation Rates

To directly compare the conversation participation rates between bot and human authors, we constructed a cumulative distribution frequency (CDF) plot depicting tweet volume per author for each of the online conversations. The resulting CDFs serve as comparative artifacts between the author types and the various conversations. In addition, we conducted a two-sample Kolmogorov–Smirnov (KS) test to return a D statistic metric that captures the absolute max distance between the bot and human distributions for each of the conversations.

The *CDF*s, depicted in Fig. 2, show similar general participation rate trends for both bots and humans across all conversations. The resulting distributions all exhibit a 'many-some-few' fat-tail distribution, with most of the authors having extremely low tweet volume (*i.e. fewer than 10 tweets*), some authors with higher tweet volumes (*i.e.* $10 < x < 1000$) and very few authors with high tweet volumes (*i.e.* $x = 1000+$). Additionally, we observed that human authors account for the largest tweet volumes per author across all conversations and have a higher concentration of low volume authors accounting for all tweet volumes.

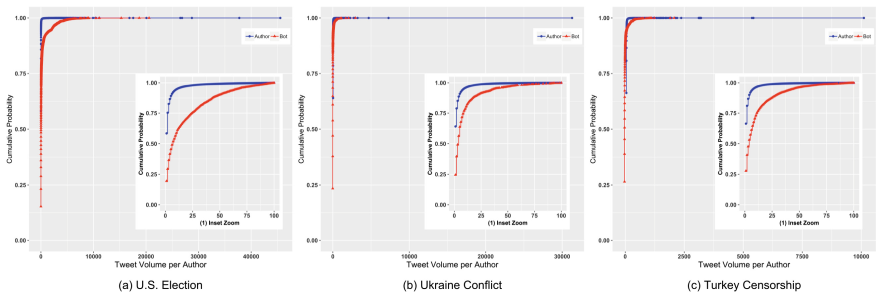

Fig. 2. Cumulative distribution (CDF) plots of tweet volume per human (blue) and bot (red) for each online conversation: (a) U.S. Election, (b) Ukraine Conflict and (c) Turkish Censorship. Inset zooms provide granularity to capture the high density of authors with low tweet volumes.

The *KS* test results between bot and human authors highlight the major difference in low tweet volume authors accounting for much larger portions of the entire tweet conversation by humans. The conversations returned *D* statistic values of 0.529, 0.408, and 0.419 for the U.S. Election, Ukraine Conflict and Turkish Censorship conversations, respectively. These maximum values were all observed where the tweet volume per author was a single tweet as shown in each plot's associated inset zoom.

4.2 In-Group and Cross-Group Communications

Figure 3 presents a consolidation of all in-group and cross-group communication frequencies observed in this study. We define in-group communication as retweet edges between like types of authors (i.e. bots retweeting bots or humans retweeting humans), while cross-group communication refers to retweets between different types of authors (i.e. bots retweeting humans or humans retweeting bots). While low retweet volumes appear to dominate for in-group and cross-group conversations across all of

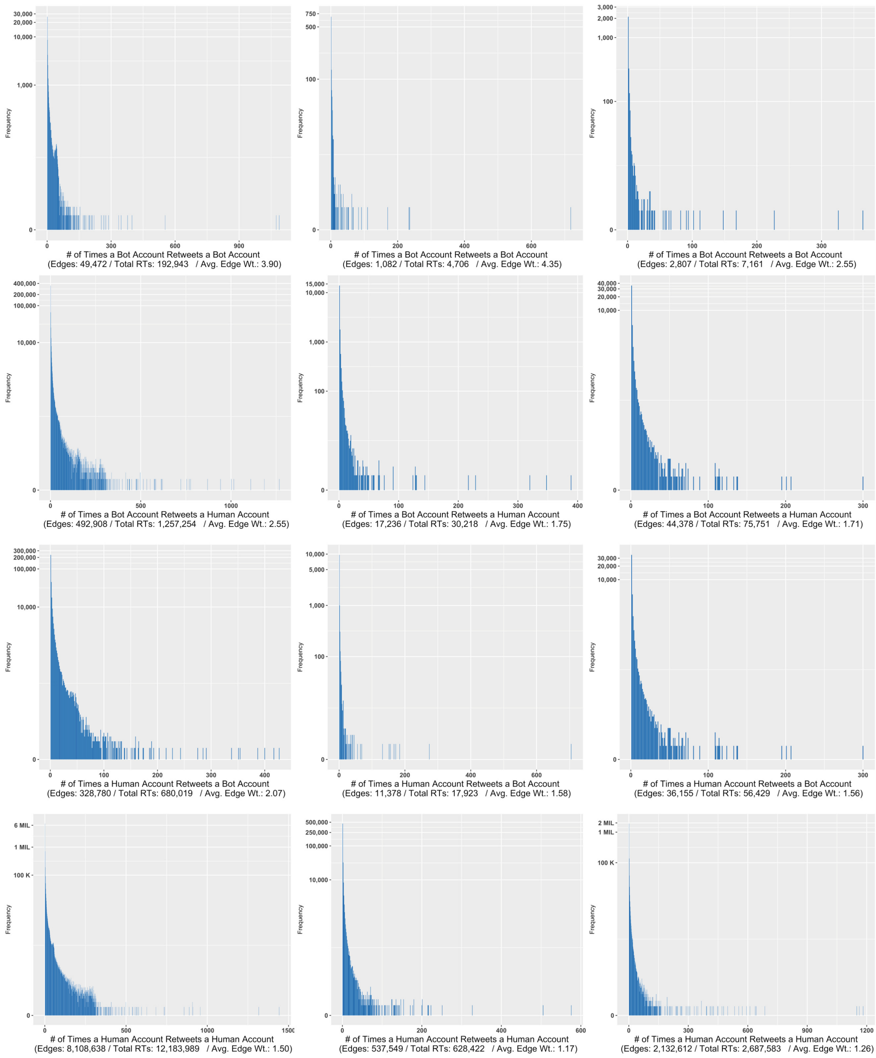

Fig. 3. Frequency distribution plots for (a) U.S. Election, (b) Ukraine Conflict and (c) Turkish Censorship retweets of in-group bot conversations (row 1), cross-group bot and human conversations (rows 2 and 3) and in-group human conversations (row 4).

the online conversations, we see increased retweet rates for all conversations initiated by a bot author, as opposed to a human author. For all three online conversations, each bot-to-bot in-group and bot-to-human cross-group conversation has a relatively higher average edge weight. The bot-to-bot author average edge weight is 160%, 272% and 102% higher than the human-to-human author average edge weight for the U.S. Election, the Ukrainian Conflict and Turkish Censorship, respectively. This suggests that either bots seek persistent contact more so than humans, or the high rate of single retweet volumes between so many different human edges dilutes any persistent human-to-human connections that exist.

4.3 Centrality Analysis

In social network analysis, centrality measurements allow for us to distinguish nodes in a network as more prominent, or important, than other nodes based on their relative position in the structure of the network [19]. In terms of our study, we sought to classify the overall relative importance of bots within our online conversations of interest by using centrality measures. To do so, we calculated three relatively common centrality measures (degree, eigenvector, and betweenness) for each online conversation. Degree centrality is the most straightforward centrality, as it is calculated from the total number of direct connections a node shares with other nodes throughout the network. One could view degree centrality as a level of popularity in a network. Eigenvector centrality is a weighted sum of both direct and indirect connections for a given node that is based on the individual degree centrality score of each node with which it shares an edge [20]. Thus, we can infer eigenvector centrality as a level of entire network influence. Betweenness centrality is the degree to which a node falls on the shortest path between other nodes in the network [21]. Therefore, we can characterize betweenness as a potential measure of information flow in a network.

The consolidated results for the three centrality measure calculations across all three conversations are presented in Fig. 4. We binned the results to capture the density of bots falling within the Top-N centrality valuations (*where, N= 1000, 100, 50 or 10*). Of note, we provide the raw number of bots and the total percentage of bots comprising the given population of Top-N centrality values. The results clearly show that authors identified as bots, though they comprise just 0.28% of total conversation authors in this study, account for a significantly large portion of prominent centrality rankings for each of the centrality measures across all conversations. Showing penetration into conversations as an influencer, the eigenvector valuations show that bots account for 43% of the top-100 nodes in the U.S. Election conversation, to include four of the top-10 centrality value positions. In the Ukraine Conflict dialogue, bots show a gaining dominance of top eigenvector values, as the bot population accounts for 21%, 30% and 50% at the top-100, top-50 and top-10 bins respectively.

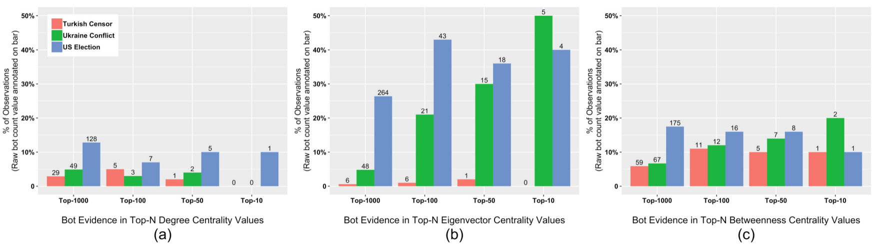

Fig. 4. Bot evidence in Top-N (N = *1000/100/50/10*) [(a) degree (b) eigenvector (c) betweenness] centrality values for: U.S. Election (*blue*), Ukraine Conflict (*green*) and Turkish Censorship (*red*).

Many studies point to the positive correlation of computed centrality values given the conceptual overlap that exists between the inputs required of the calculations [22]. Given an expected correlation of centrality values, lack of correlation evidence provides an opportunity to further investigate a node for interesting behavior. We conducted such an analysis by plotting correlation plots against each other as depicted in Fig. 5.

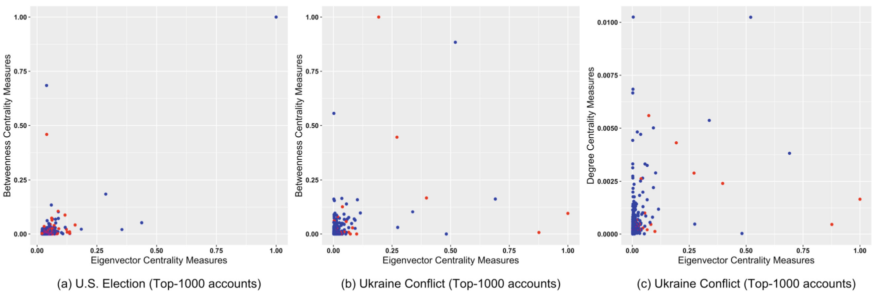

Fig. 5. Correlation of centrality measures for select centrality comparisons: (a) U.S. Election eigenvector versus betweenness analysis, (b) Ukraine Conflict eigenvector versus betweenness analysis and (c) Ukraine Conflict eigenvector versus degree analysis.

The depicted centrality correlation plots in Fig. 5 provide compelling insights into some of the observed conversations. First, in the U.S. Election conversation plot (Fig. 5a), we see very few correlation outliers on the plot. Interestingly, the top eigenvector and betweenness centrality node is the same human author, in this case, then-candidate Donald Trump (*@realDonaldTrump*). Conversely, we see far more correlation outliers in the Ukraine conflict conversations. Specifically, the most divergent nodes are bots, which could be cause for greater investigation as to their specific tweeting behavior. In the eigenvector versus degree Ukraine plot (Fig. 5c), the two most 'influential' nodes according to eigenvector centrality, which are bots, are actually not that popular given low degree centralities. This suggests these bots were able to infiltrate the conversation network by acquiring connections with popular nodes, while avoiding popularity, or detection, themselves.

4.4 Community Detection

Community detection is another common application in social network analysis that allows researchers to uncover localized sub-graphs, or communities, of highly connected nodes that are otherwise less connected to the remainder of the network [23]. The Louvain [24] method is one such community detection algorithm that is highly applicable for the identification of emergent community structure in large-scale network analyses. It seeks an undefined number of emergent communities by executing a two-stage greedy heuristic that iteratively optimizes modularity locally and culminates when global network modularity reaches a maximum value. For our purposes, we sought to observe the density of bots within any defined community structure of the online conversations. Specifically, we wanted to determine if bots clustered among themselves or if they dispersed among the larger human author communities, which would provide further explanation for our in-group and cross-group communication findings in Sect. 4.2.

Table 2 outlines the evidence of bot density within the most populated emergent communities detected for each online conversation. In total, we discovered 71.2% of all bots within the top-5 most populated communities for the U.S. Election conversation, with 75.9% and 53.1% for the Ukraine Conflict and Turkish censorship conversations, respectively. Although we see a dispersal of bot populations throughout all of the top communities, there are multiple instances in which the bot density is much greater than the community population percentage in relation to the total network population. This is representative of the higher in-group communication rates found between bots in Sect. 4.2, while the general dispersal of bots supports the findings of cross-group communication evidence.

Table 2. Bot density of largest emergent communities.

Comm.	U.S. Election		Ukraine Conflict		Turkish Censorship	
	Bot Count (% of comm.)	Comm. Size (network %)	Bot Count (% of comm.)	Comm. Size (network %)	Bot Count (% of comm.)	Comm. Size (network %)
1	901 (15.69%)	1,009,872 (39.48%)	454 (21.25%)	58,397 (23.30%)	787 (16.39%)	268,311 (24.94%)
2	1305 (22.73%)	900,076 (35.19%)	166 (7.78%)	45,330 (18.09%)	277 (5.77%)	146,350 (13.60%)
3	1345 (23.43%)	308,040 (12.04%)	267 (12.50%)	29,310 (11.69%)	1172 (24.40%)	107,224 (9.97%)
4	337 (5.87%)	84,733 (3.31%)	12 (0.06%)	15,536 (6.20%)	287 (5.98%)	86,550 (8.04%)
5	242 (4.20%)	59,441 (2.32%)	616 (28.84%)	15,439 (6.16%)	27 (0.56%)	48,813 (4.54%)

5 Conclusion and Future Work

In summary, we presented a framework to characterize the pervasiveness and relative importance of bots in various OSN conversations of three significant global events in 2016. In total, we harvested more than 30 million tweets from the U.S. Presidential Election, the Ukrainian Conflict and Turkish Political Censorship and compared the conversational patterns of bots and humans within each event. We further examined the social network structure of each online conversation to determine if bots exhibited particular influence in a network, while also determining bot participation in key emergent network community subgraphs. The results showed that although Twitter participants identified as social bots comprised only 0.28% of all OSN users in this study, they accounted for a significantly large portion of prominent centrality rankings across the three conversations. This includes the identification of individual bots as top-10 influencer nodes out of a total corpus consisting of more than 2.8 million nodes. Additionally, we observed that the most influential social bots had relatively low popularity, or degree centrality, suggesting influence can be obtained without popularity. In the case of social bots, popularity could be seen as a negative characteristic if trying to avoid detection. This finding is supported by previous findings in social media studies showing influence in a network is not necessarily driven by popularity [25].

While this study contributes to the nascent literature of social bot analysis by introducing a comparative analysis framework based on social network analysis techniques, there are limitations to take into consideration. As Tufekci [26] asserts, social media analyses must state their limitations in terms of validity and representativeness when attempting to account for issues such as the over-emphasis of single platforms and sampling biases. These issues are not unique to this study. However, we did limit our research to just one platform (i.e. Twitter) that includes a sampling bias. Though the methodology presented is not bound to a particular social media platform type, we were limited to currently available bot detection sources, which focus solely on Twitter. As the literature expands in the near future, we hope to account not only for additional bot detection services using Twitter, but additional social media platform sources as well. Specifically, we will seek to determine if the findings produced here hold with other bot detection algorithms. Further extensions of this initial work will closely examine any observable characteristics differentiating the emergent communities of interests. This will include narrative analysis through natural language processing to determine any attempts by bots to polarize particular populations within the conversations. The results from such an analysis could increase the relevancy of this study by potentially extending the observable influence of social bots beyond online social networks and into other social activities.

Acknowledgments. Special thanks to Nikan Chavoshi from New Mexico State University for support and access to DeBot. This research was partially supported through funding from the Seth Bonder Foundation.

References

1. Lazer, D., et al.: The science of fake news. Science **359**, 1094–1096 (2018)
2. Ferrara, E., Varol, O., Davis, C., Menczer, F., Flammini, A.: The rise of social bots. Commun. ACM **59**(7), 96–104 (2016)
3. Subrahmanian, V., et al.: The DARPA Twitter bot challenge. Computer **49**(6), 38–46 (2016)
4. Varol, O., Ferrara, E., Davis, C., Menczer, F., Flammini, A.: Online human-bot interactions: detection, estimation, and characterization. In: AAAI Web and Social Media, pp. 280–289 (2017)
5. Chavoshi, N., Hamooni, H., Mueen, A.: DeBot: Twitter bot detection via warped correlation. In: 2016 IEEE 16th International Conference on Data Mining (ICDM), pp. 817–822 (2016)
6. Chu, Z., Gianvecchio, S., Wang, H., Jajodia, S.: Detecting automation of Twitter accounts: are you a human, bot, or cyborg? IEEE Secur. Comput. **9**(6), 811–824 (2012)
7. Davis, C.A., Varol, O., Ferrara, E., Flammini, A., Menczer, F.: BotOrNot: a system to evaluate social bots, pp. 273–274 (2016)
8. Stukal, D., Sanovich, S., Bonneau, R., Tucker, J.A.: Detecting bots on Russian political Twitter. Big Data **5**, 310–324 (2017)
9. Abokhodair, N., Yoo, D., McDonald, D.W.: Dissecting a social botnet: growth, content and influence in Twitter. CSCW **2016**, 839–851 (2015)
10. Kušen, E., Strembeck, M.: Why so emotional? An analysis of emotional bot-generated content on Twitter. COMPLEXIS **2018**, 13–22 (2018)
11. Hegelich, S., Janetzko, D.: Are social bots on Twitter political actors? Empirical evidence from a Ukrainian social botnet. In: ICWSM, pp. 579–582 (2016)
12. Bessi, A., Ferrara, E.: Social bots distort the 2016 U.S. Presidential election online discussion. First Monday 21(11) (2016)
13. Forelle, M., Howard, P.N., Monroy-Hernández, A., Savage, S.: Political bots and the manipulation of public opinion in Venezuela. In: SSRN, pp. 1–8 (2015)
14. Howard, P.N., Kollanyi, B.: Bots, #StrongerIn, and #Brexit: Computational Propaganda during the UK-EU Referendum. SSRN (2016). URL: https://ssrn.com/abstract=2798311
15. Boshmaf, Y., Muslukhov, I., Beznosov, K., Ripeanu, M.: Design and analysis of a social botnet. Comput. Netw. **57**, 556–578 (2013)
16. Howard, P.N., Woolley, S., Calo, R.: Algorithms, bots, and political communication in the US 2016 election. J. Inf. Technol. Polit. **15**, 81–93 (2018)
17. Zhdanova, M., Orlova, D.: Computational Propaganda in Ukraine: Caught Between External Threats and Internal Challenges. COMPROP, Oxford, UK. Working Paper (2017)
18. Chavoshi, N., Hamooni, H., Mueen, A.: Temporal patterns in bot activities. In: 26th International Conference on World Wide Web, pp. 1601–1606 (2017)
19. Wasserman, S., Faust, K.: Social Network Analysis: Methods and Applications. Cambridge University Press, Cambridge (1994)
20. Bonacich, P.: Some unique properties of eigenvector centrality. Soc. Netw. **29**(4), 555–564 (2007)
21. Freeman, L.C.: A set of measures of centrality based on betweenness. Sociometry **40**(1), 35–41 (1977)
22. Valente, T.W., Coronges, K., Lakon, C., Costenbader, E.: How correlated are network centrality measures? Connections **28**(1), 16–26 (2008)
23. Girvan, M., Newman, M.E.: Community structure in social and biological networks. Proc. Natl. Acad. Sci. **99**(12), 7821–7826 (2002)

24. Blondel, V.D., Guillaume, J-L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. J. Stat. Mech. **2008**(10) (2008)
25. Cha, M., Haddadi, H., Benevenuto, F., Gummadi, K.P.: Measuring user influence in Twitter: the million follower fallacy. In: Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media (ICWSM 2010), pp. 10–17 (2010)
26. Tufekci, Z.: Big questions for social media big data: representativeness, validity and other methodological pitfalls. In 8th International Conference on Weblogs and Social Media (ICWSM 2014), pp. 505–514 (2014)