Bot-ivistm: Assessing Information Manipulation in Social Media Using Network Analytics

Matthew C. Benigni, Kenneth Joseph, and Kathleen M. Carley

Abstract Social influence bot networks are used to effect discussions in social media. While traditional social network methods have been used in assessing social media data, they are insufficient to identify and characterize social influence bots, the networks in which they reside and their behavior. However, these bots can be identified, their prevalence assessed, and their impact on groups assessed using high dimensional network analytics. This is illustrated using data from three different activist communities on Twitter—the "alt-right," ISIS sympathizers in the Syrian revolution, and activists of the Euromaidan movement. We observe a new kind of behavior that social influence bots engage in—repetitive @mentions of each other. This behavior is used to manipulate complex network metrics, artificially inflating the influence of particular users and specific agendas. We show that this bot behavior can affect network measures by as much as 60% for accounts that are promoted by these bots. This requires a new method to differentiate "promoted accounts" from actual influencers. We present this method. We also present a method to identify social influence bot "sub-communities." We show how an array of sub-communities across our datasets are used to promote different agendas, from more traditional foci (e.g., influence marketing) to more nefarious goals (e.g., promoting particular political ideologies).

M. C. Benigni · K. M. Carley (✉)
Institute for Software Research, Carnegie Mellon University, Pittsburgh, PA, USA
e-mail: kathleen.carley@cs.cmu.edu

K. Joseph
Computer Science and Engineering, SUNY Buffalo, Buffalo, NY, USA
e-mail: kjoseph@buffalo.edu

N. Agarwal et al. (eds.), *Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining*, Lecture Notes in Social Networks, https://doi.org/10.1007/978-3-319-94105-9_2

1 Introduction

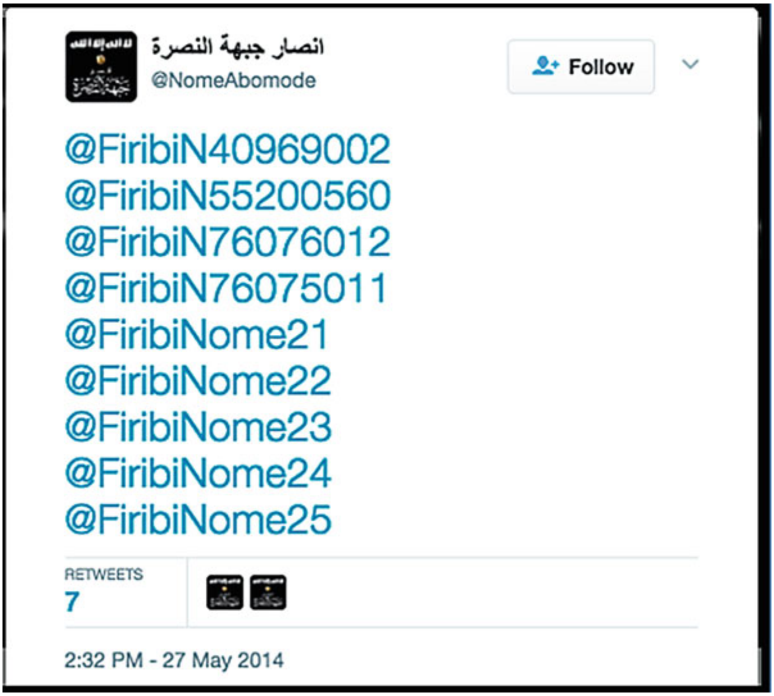
How can individuals and groups be manipulated in social media? What messaging strategies can be used to shape behavior and alter opinions? In general, opinions and behaviors are a function of social influence [1]. That is, whom you know impacts not just what you know, but also your opinions and behavior. Consequently, group members, particularly those in tightly knit groups can come to share the same opinions and behaviors. Herein, we examine this process in social media. We identify a new type of bot that operates through social influence, the social influence bot network (SIBN). We then demonstrate the information strategies employed by such SIBNs to manipulate individuals and groups in social media.

Initially used to spread spam [2] and malware [3], a substantial literature now documents the use of bots on Twitter to influence global politics [4–7]. Many of the original bots discovered were individual bots acting in isolation to effect social goals. These bots or "social bots" [7, 8] have been used to shape discussions during political revolutions [9, 10] and in recruiting and propaganda efforts for terrorist groups [11, 12]. Social bots, and individual bots, were also pervasive and highly active in online conversations during the 2016 US presidential election [13].

More recently, we find evidence of concerted coordinated effort using networks of bots. Thus, we deviate from the existing literature on social bot networks in that we study the creation and use of a new form of social bot—the social influence bot network (SIBN). Most prior work has assessed the network structure of bots on the Twitter follower network [14] or on the directed @mention network [8]. In the latter case, the focus is generally on how bots mention real users in order to gain attention. SIBNs use @mentions in this way and in doing so change the effective influence of those users whom they @mention. Importantly, SIBNs also use @mentions to manipulate the Twitter social network by *@mentioning each other*. In other words, they operate by altering the social network structure, and so impact who has social influence on whom.

An example of the way social influence bots in our data use @mentions is shown in Fig. 1. The tweet in the figure was sent by a bot in our Syrian revolution dataset and contains only a string of mentions to nine other similarly named accounts. Shortly after, the bot sending this tweet was itself mentioned in similarly structured tweets by the other bots mentioned in Fig. 1. The sole purpose of these tweets is to artificially manipulate the reciprocal @mention, or co-mention graph, creating networks of bots with strong ties in the @mention network. More specifically, this kind of behavior produces a mention core of bots—a sub-community of social influence bots displaying "core-like" behavior [15] that have anomalously strong connections in the co-mention network. To distinguish this particular form of social bot network, we will refer to those social bot networks that have a mention core as **social influence bot networks (SIBNs)**.

Social influence bots can impact the Twitter ecosystem on at least three levels. First, they can be used at the "content-level" to rapidly spread specific tweets and/or particular URLs and make them appear artificially popular [8]. Second, they can be



انصار جبهة النصرة
@NomeAbomode

@FiribiN40969002
@FiribiN55200560
@FiribiN76076012
@FiribiN76075011
@FiribiNome21
@FiribiNome22
@FiribiNome23
@FiribiNome24
@FiribiNome25

RETWEETS
7

2:32 PM - 27 May 2014

Fig. 1 Depicts core-bot behavior in the Firibinome social influence bot network

used at the "user-level" to artificially inflate the importance of themselves and/or target users as measured by standard metrics of influence, like a user's number of followers [13]. Finally, social influence bots can be used at the "community-level," forging fake communities or embedding themselves into communities of real users to promote particular ideologies [9]. In other words, they can and do manipulate what is being said, who is communicating to whom, and the relative influence of particular actors. These means of manipulation are mutually reinforcing, and thus, are often used simultaneously as part of a social influence information campaign.

Social influence [1] is the process through which the opinions and behaviors of actors are influenced by the opinions and behaviors of those with whom they have a structural relation. As the network of ties among actors change, so too does who has influence, the rate at which opinions and behaviors change, the rate at which the group reaches consensus, etc. Ties among actors, however, are continually constructed as actors learn new information, share opinions, and so forth [16, 17]. The result is the mutual development of who is tied to or has relations with whom and who shares what beliefs or opinions with whom. In general the more social relations or ties within a group, the higher the density, the faster the social influence process. At its extreme, if there are only a few connections, but it is still a connected network, if all the actors have the same information, same beliefs, the network will quickly become fully connected. The social network and the knowledge network connecting actors are co-evolutionary.

These high dimensional networks where actors are connected by social ties and knowledge/opinion ties are topic groups. In Twitter to find topic groups we use as the social network the mentions network (the network formed by all instances of who mentions or retweets or replies to whom) and as the knowledge network the shared hashtag network (the network formed by all instances of who shared the same hashtags and note that it could be same concepts in the tweet content). Using the IVCC process [18], which has now been realized in ORA-PRO [19], these two networks are used to find a set of topic groups. As the connections in both the social network and the shared knowledge network increase the topic group begins to grow into an echo-chamber. In other words, echo-chambers are a set of actors who are densely connected in the social network and in the shared knowledge network; hence, in Twitter this would be a set of users who are highly connected in the mentions network and in the shared hashtag network.

In the extreme, a topic group formed of a set of actors who all connect to each other and all share the same knowledge is a pure echo-chamber. That is, a pure echo-chamber is a completely connected subgraph in both the social network and the shared knowledge network. As topic groups become more echo-chamber like in nature emotions can escalate, language variation can decrease, and ideas will flow faster. The more echo-chamber like a topic group is the more prone it is to groupthink. Pure echo-chambers can be very influential as all members spread the same message, and it is difficult for those outside the group to get their message into the group.

Social influence bot networks (SIBNs) influence topic groups and have as their core an echo-chamber. An SIBN is composed of a network of core and non-core bots, such that the core bots form an echo-chamber. SIBNs connect into topic groups and alter the makeup of these groups. Due to this constructural process, one of the most powerful mechanisms for increasing the manipulative power of social influence bots is the construction of network ties within topic groups. SIBNs, through the use of @mentions, are altering the social networks of groups on Twitter and so impacting who is influencing whom. While these *social influence bot networks* (social botnets or SIBNs [7, 9])[1] are in some ways easier to detect than isolated accounts, they are also much more able to achieve desired manipulations of the Twitter ecosystem. In addition to allowing faster spread of content [8], creating a network of social ties among bots can inflate the importance of bots in network metrics like degree and betweenness centrality [22].

The present work presents an exploration of the impact of SIBNs in three very different political discussions—“alt-right” themed discussions of the 2016 US presidential elections, the Syrian revolution, and the Euromaidan movement in the Ukraine. For each topic, we collect a different dataset via snowball sampling of the follower network. We seed each snowball sample with a set of core members

[1] Social bot networks are also referred to as Sybils in the computer security literature [20, 21].

of these discussion communities, allowing us the ability to explore the social structure of the users and SIBNs associated with those networks and influencing user behavior. Within each of these discussion communities there are topic groups. Associated with these topic groups are one or more SIBNs. Despite the distinct contexts in which these three communities are set, we observe high levels of similarity in how SIBNs within them impact user-level metrics and community structure. Specifically, we observe that:

- SIBNs operate as echo-chambers. Each SIBN is a set of bots that @mention other bots forming a fully connected social network, and which tweets about the same topics in similar ways. By mentioning each other, the SIBN forms a dense core thus giving weight to the importance of the individual bots in the SIBN in terms of message prioritization.
- SIBNs manipulate the apparent influentialness of key actors. SIBNs are used to greatly inflate complex influence metrics on the social network induced by reciprocal mentioning behavior on Twitter. We find that some accounts drop in influence by as much as 60% in important network measures like coreness [15] after removing bot-like behavior from our data.
- SIBNs manipulate what information is flowing to achieve various goals. Each SIBN appears to have a distinct goal. These included SIBNs with more traditional foci (e.g., explicit influence marketing) as well as those aimed towards more nefarious goals (e.g., promoting particular political ideologies).
- SIBNs build network ties among users so as to better manipulate those users. Bot creators within our communities used directed social engineering to accomplish particular goals. This included, for example, a dataset sharing lewd images of women to attract young men interspersed with calls to violence in our Euromaidan data.

The obvious questions that remain, of course, are why bot creators manipulate the co-mention graph in this way and how SIBNs are used in our datasets.

With respect to the "why" of SIBNs, the creation of a mention core in the co-mention graph allows the SIBN to confound radial-volume centrality measures like PageRank [23] and coreness [15] that are supposedly more robust to spam-like behavior. Because of this, owners of SIBNs can inflate the centrality of **promoted accounts**, which are fully or semi-automated accounts that are mentioned by the core bots in the SIBN and that attempt to influence a specific online community of interest. These promoted users thus have influenced level metrics that distort our understanding of true **community influencers**. While we find that true community influencers, individuals with real influence on the non-automated discussion community of interest, tend to be mentioned frequently by SIBNs as well, they can be differentiated from promoted users by the fact that their ranking on radial-volume does not depend significantly on bot activity. Assume that Twitter users are more likely to have recommended to them to follow more central actors, to follow topics mentioned by more central actors, and are more likely to have messages from such actors prioritized in their lists. If this is indeed the case, and it appears to be so, then promoting accounts, and promoting certain messages, will alter who is connected to whom in the social network and who shares what information in the knowledge network.

With respect to how SIBNs are used, a graph-level perspective is required. In other words, the utility of SIBNs is best understood by first uncovering the different SIBNs that exist in our dataset, and then considering how each may be used in different ways to manipulate influence across particular sets of users and themes using various social engineering strategies. To uncover the different SIBNs within our dataset, we develop a methodology based on dense subgraph detection [24]. We use this method to survey multiple different SIBNs in our datasets, each of which have unique goals.

2 Related Work

2.1 Overview

For an expansive overview of social bots, we point the reader to [7]. Here, we restrict ourselves to a discussion only of recent work related to the study of social bot networks and automated manipulation of influence metrics.

Boshmaf et al. define a social bot network as a set of social bots with three components: the botherder who controls the social bots, the social bots that carry out tasks assigned by the botherder, and a Command and Control channel used to facilitate task assignment [25, 26]. In order to construct social bot networks, as opposed to just a series of isolated bots, botherders must engage in some form of link creation between bots and *link farming* to external accounts. Ghosh et al. [14] define link farming as the process by which “users, especially spammers, try to acquire large numbers of follower links in the social network.” They find, somewhat surprisingly, that bots do not necessarily create mutual following ties only with other bots, but also with other “real” Twitter users with low thresholds for reciprocity. Here, we discuss how bots attempt to use @mentions in similar ways. However, our focus is also on how links between the bots themselves corrupt influence metrics.

In turn, several works have shown the ease with which a social bot network can artificially inflate such metrics, from obvious ones like number of followers [26] to more opaque metrics like Klout scores [3, 8, 27]. This research shows that these more opaque metrics can be influenced by both following and @-mentioning behavior. Herein, we show one way in which these opaque metrics might be impacted by @mentioning behavior among bots themselves, which can significantly impact important social network metrics in non-obvious ways.

2.2 Defending Against Social Bot Networks

The most common approach to defending against social bot networks is to identify individual bot accounts using machine-learning methods [28–30]. As noted by Ferrara et al. [7], the best performing of such methods tend to take into account a variety of behavioral features of accounts. Such methods are quite fragile. In contrast, we detect SIBNs using only social graph features. This effort aligns more closely to a host of Sybil-detection algorithms in the literature (for a nice review, see [20]). Our method differs from most prior work; however, in that we focus on reciprocal mentioning behavior, rather than following [14] or friend (on Facebook) [26] ties. While "real" users have been observed to reciprocate friendships or following ties with social influence bots (leading to difficulties in Sybil detection [7]), the reciprocal @mention network has generally been found to signal strong social relationships [31, 32]. Perhaps more importantly, such directed communication is unlikely to occur in reciprocal formats between non-automated and automated accounts.

Moving beyond bot detection at the account level is important, however, because an increasing amount of automated, malicious activity happens within "cyborg" accounts [7]. These cyborg accounts, which blend human behavior with automated behavior produced by Twitter applications the user has subscribed to, are extremely difficult to detect using existing bot detection methods. Cresci et al. [33] find that existing methods perform at near-chance levels on classifying these accounts that blend automated and (at least seemingly) non-automated behavior.

These findings are worrying for two reasons. First, as it becomes increasingly difficult to determine whether or not an account is controlled by a human, analysts and scholars will have an increasingly difficult time studying socio-theoretic models of human behavior [34, 35]. Second, these results suggest it is increasingly difficult for "real" Twitter users to tell the different between bots and non-bots. This leads to an increasingly level of social interaction between "real" users and at least semi-automated accounts. The present work therefore focuses on automated behavior at the level of the *edge* (interaction), rather than the *node* (the individual account). We identify anomalous edge behavior and then assess how this behavior impacts user influence metrics and leverage it to identify SIBNs.

The prevalence of cyborgs also implies that differentiating "bots" from "humans" is a potentially misleading problem [7]. Scholars have thus also began to focus on developing methods to address potential impacts of automated behaviors on social media ecosystems without having to determine which accounts are bots [14, 36, 37]. These *tolerance-based* [20] defense schemes rely on tactics like down-weighting influence scores for accounts that are "bot-like" and assessing the percentage of "bot-like" activity around a certain topic. Our approach to differentiating community influencers from promoted accounts is an example of a tolerance-based approach.

2.3 The Impact of Social Bot Networks

A growing literature details the way in which social bot networks are used for promotion of political agendas of varying forms [9, 10, 38–40]. Abokhodair et al. [9] describe the 35-week lifespan of a social bot network on Twitter designed to express opinion, testimony of ongoing events, and engage in preliminary conversation associated with the Syrian revolution. Similar techniques have been observed in the case of ISIS' online video dissemination as well [12]. The present work complements these efforts by showing how SIBNs can conflate not only topical discussions but also social network metrics, i.e., they impact both the social and the knowledge network. This is demonstrated using the evidence of highly related behavior in three different datasets.

Scholars have also observed that social bot networks were responsible for a significant amount of content produced around conspiracy theories central to the 2016 US presidential election [40]. Nied et al. observed in this context that in a cluster of bot-like accounts, "[u]nexpectedly, 62% of all mentions were between users of which neither followed the other." Our work sheds significant insight into this behavior.

This recent work exploring the impact of social bot networks falls, however, on a backdrop of a large literature analyzing Twitter data without considering the impact of bots or only considering bots superficially using heuristics. The present work shows that in addition to follower networks [26], directed mention networks [41], and topical themes [10] being susceptible to social bot network behavior, so too is the network of reciprocal mentioning behavior.

3 Data

The present work studies three distinct datasets. Each has been collected using snowball sampling [42]. In a snowball sample, a set of individuals are chosen as "seed agents," and individuals with social ties to one or more seed agents are added to the sample. This technique can be iterated in a series of steps or "hops." On each "hop," the set of accounts sampled on the prior hop are used as seed for collection in the next hop. Thus, a "1-hop" snowball sample collects the social ties of the seed agents; a "2-hop" sample collects the social ties of the seed agents' social ties ("friends of friends"), and so on. For each of our datasets, we seed searches with a set of users that are known members of a community of interest. We then collect additional users based on a snowball sample using relationships defined by a dataset-specific social graph. For each user added to the snowball sample, we collect up to the last 3200 tweets sent by the user, as allowed by the Twitter API.

This method of collecting data is different from most prior work on the influence of bots on geopolitical discussions [9, 40], which focuses collection around tweets mentioning one or more terms of interest. It is potentially this form of sampling, rather than a focus on collecting tweets, that allows us to more easily observe SIBN behavior.

The first dataset we collect is the **alt-right community (ALT16)** dataset. In October 2016, we seeded a one-hop snowball sample of 2482 users who each followed five influential Twitter users associated with the Alt-Right political movement: Richard Spencer, Jared Taylor, American Renaissance, Milo Yiannopoulos, and Pax Dickinson. The search resulted in 106K users and 268 million tweets.

The second dataset is the **Euromaidan community (EUR17)** dataset. The Euromaidan revolution occurred as a wave of demonstrations starting in Ukraine in November 2013 and resulted in the removal of Ukrainian President Viktor Yanukovych from power. In an attempt to study messaging themes used within the Euromaidan movement we conducted a one-hop snowball sample of starting from 1209 and following their ties in the directed @mention network (e.g., everyone any of the seed users have mentioned) from March 2014 to September 2017. The 1209 seed users were collected via a combination of automated search and manual annotation from subject matter experts [43]. The snowball search resulted in 92,706 Twitter users and 212 million tweets.

The final dataset is the **Syrian revolution (SYR15)** dataset. With research objectives similar to our Euromaidan Study, we conducted a one-hop snowball sample of the friend graph for 13,949 accounts that had been identified as ISIS-supporting in prior work [18]. We then removed all nodes with degree less than 2 in the following graph. The search resulted in 87,046 Twitter users and 179 million tweets.

4 Observing Anomalous @Mentioning Behavior

Much like [9], our datasets were originally collected for the study of online political activism but were found to be deficient for these purposes until bot activities within it were better understood. Here, we briefly detail how the anomalous @mention behavior we observe was discovered and how it motivated the proposed methodology. Because we assume readers will be most familiar with the context and users central to the ALT16 dataset, we use it as a running example throughout this section.

As noted above, prior work has suggested that *co-mentioning* behavior—when two users reciprocate @mentions of each other—is a strong signal of social relationships between two Twitter users [31, 32]. Correspondingly, we began our analysis of these two datasets by constructing the *directed mention graph*. Mathematically, the weighted mention graph $M(V, E)$ is a weighted directed graph with vertices V: $\{v_1, \ldots, v_n\}$ consisting of the users returned from our search and edges E_m: $\{e_{M,1}, \ldots, e_{M,m}\}$ defined as the number of unique tweets where $user_i$ mentions $user_j$. We then constructed the primary graph of interest, the reciprocal or *co-mention graph*, $R(V, E)$. R has the same vertices as M but had undirected

Table 1 Top users in the ALT16 dataset co-mention graph using weighted degree centrality

Rank	Top ALT16 users
1	nayami_rescue
2	Socialfave
3	sundoghigh
4	TheMisterFavor
5	SarCatStyX
6	saravastiares
7	KoichicCheryl
8	realDonaldTrump
9	Easy_Branches
10	lupash7

edges with weights set to the minimum number of times $user_i$ mentions $user_j$, or vice versa.

As we were interested in understanding who the influential users in the co-mention graph, we began by considering weighted degree centrality of the nodes in this network. Weighted degree centrality, which simply measures the total weight of each user's edges in the co-mention graph, is one of the easiest and most common network metrics to compute. Table 1 presents the top 10 users in the ALT16 dataset according to weighted degree centrality. These users were not necessarily those one would expect to be central in a snowball sample seeded with highly politically motivated accounts. Upon closer inspection, we found some of these users to have sent several tweets containing only a string of @mentions to other accounts, e.g., "@user1 @user2 @user3 @user4 Hey!". Looking at the profiles of @user1, @user2, and @user3, we further observed that some of these accounts were cyborgs leveraging the same kind of automated behavior. Others were clearly bot accounts, sending only these @mention tweets, with no attempt to appear human or perform human-like behaviors.

As it is well known that bots can influence simple network metrics like degree centrality, results in Table 1 were discouraging but not surprising. We then turned to more complex network metrics that are better able to withstand spam and bot-like behavior. Specifically, we consider PageRank [23] and coreness [15], two common, more complex metrics for measuring influence in complex networks.

Both metrics are drawn from the family of *radial-volume centralities* [44], which attempt to quantify an account (or more generally, a network vertex) x's influence based on the influence of the other vertices to which x is linked. These measures can be viewed on a continuum based on the length of walk considered in the neighborhood of a given vertex. Coreness [45] is calculated based on the concept of *K-shells* within a graph. A K-shell is defined as the maximal subgraph of a given graph G, where all vertices are of degree greater than or equal to k. A vertex's coreness, K_s, indicates the greatest value k for with the node remains in the corresponding K-shell. In addition to coreness, we also look at PageRank, or eigenvector centrality, which is roughly defined as "accounts who are popular with other accounts who are popular" [46].

Table 2 Top users in the ALT16 dataset co-mention graph using PageRank (left) and coreness (right)

Rank	Top ALT16 users (PageRank)	Top ALT16 users (Coreness)
1	realDonaldTrump	2020sahara
2	HillaryClinton	JuliaZek
3	YouTube	Jerz_Gal
4	POTUS	DilrubaLees
5	FoxNews	7artistai
6	CNN	nayami_rescue
7	nytimes	wanderingstarz 1
8	timkaine	JulezPooh
9	NASA	Dollhouse
10	wikileaks	Edward733

Table 2 presents the top 10 nodes on PageRank and coreness for the ALT16 dataset. We see that results from the PageRank algorithm pass tests of face validity, while this is still not true of coreness. Upon further investigation of these coreness metrics, we again found the anomalous @mention behavior we observed to confound measures of degree also impacted measures of more complex measures like coreness. Below, we will show that even for PageRank, designed specifically for spam-like behavior, rank ordering of nodes below the top 10 can be significantly impacted by this behavior.

SIBNs thus inhibited our ability to analyze influential users in the ALT16 dataset, as well as the other two datasets we do not discuss here. We were left with two questions for which methods needed to be introduced. First, we wanted to be able to "tolerate" [20] the influence of automated @mentioning behavior on our analysis of influential users in our datasets. In the terminology used in the present work, we wanted to be able to differentiate between *promoted accounts* and actual *community influencers*. Second, we were interested in better understanding the structure of SIBNs in our dataset. In particular, we wanted to understand if many SIBNs existed within each dataset, and if so, what their utility was. To do so, we needed to find a means of detecting mention cores, the foundation of SIBNs, in our data. The following section introduces the straightforward, scalable methods we adopt to address these issues.

4.1 Methodology

4.1.1 Differentiating Promoted Accounts from Community Influencers

We use a straightforward procedure to differentiate between promoted users and community influencers. Our strategy is based on the fact that if we are only concerned with identifying any core bot, any bot that is likely to be part of a mention core, rather than distinct SIBNs, we can simply use a measure of @mention activity. Correspondingly, we first define γ_p as the mention per tweet ratio for each user that is greater than a percentage p of users within a given dataset. For example, $\gamma_{0.995}$ would represent the mentions per tweet ratio that is greater than the mentions per tweet ratio for 99.5% of the users in a given dataset.

By setting p to a large number, we can isolate core bots of any SIBNs in our dataset, as they are the only accounts to produce only tweets that had only @mentions. Upon removing these accounts and recomputing network metrics like coreness and PageRank, we can then differentiate promoted users from community influencers by how the network metrics of these accounts change. Metrics of promoted users should suffer significantly when core bots are removed. In contrast, true community influencers should not be impacted by the removal of core-bot behavior.

4.1.2 Detecting SIBNs

Because mention cores, the base of an SIBN, create highly dense communities, *dense subgraph detection* offers a logical means to detect them. Dense subgraph detection allows an analyst to extract only subgraphs of a network in which all nodes within that subgraph have an anomalously high density of (weighted) interactions among them. Dense subgraph detection can therefore be preferable when a complete clustering of the data is not desired, but only dense substructure is of interest [24, 47]. This is precisely the case in our setting we would like to extract out only the dense mention cores of SIBNs.

Our method for detecting dense subgraphs follows the work of [24], with one difference. Specifically, Chen and Saad [24] define the density of a subgraph as an unweighted link count. Because of the repetitive use of @mentioning by mention cores, we instead choose to define subgraph density by summing link weights instead of link counts.

Algorithm 1 Find core botnet members of an SIBN

Input: Given a large sparse, weighted, reciprocal mention graph R, and density threshold $d_{\min}$

Output: Set $D{:}d_1(v,e),\ldots,d_n(v,e)$ such that each subgraph contains a subset of users exhibiting behavior that meets the definition of a SIBN core bot member.

1. Compute Matrix C_R as defined in (2)
2. Sort the largest t non-zero entries of C_R in ascending order, where $t = nz(A)$. Denote Q the sorted array.
3. Construct the hierarchy T according to the sorted vertex pairs designated by Q.
4. Extract Subgraphs of where $d_G \geq d_{\min}$
5. Manually inspect subgraphs for SIBN like behavior

While we leave a detailed discussion of the model to [24], we provide a summary of our approach here and in Algorithm 1. Just as presented in [24], we search for dense subgraphs (mention cores) by constructing A_R, a weighted adjacency matrix of R. Let us define C_R, the cosine matrix, as:

$$C_R\left(i, j\right) = \frac{\left\langle A_G\left(:, i\right), A_R\left(:, j\right)\right\rangle}{\left\|A\left(:, i\right)\right\| \left\|A\left(:, j\right)\right\|} \tag{1}$$

We then set $t = 2 \times |E_R|$ and sort the largest t non-zero entries of C_R in ascending order. We denote this sorted array as Q and construct a hierarchy T based on sorted vertex pairs based on the measure C_R. We then extract the largest distinct subgraphs from T that meet a minimum size, s_{min} and density threshold, d_{min}. We have set s_{min} to 30 and d_{min} to 0.75 in this work based on manual inspection.

In sum, the algorithm we use, defined here and in Algorithm 1, takes as input a weighted co-mention graph for a particular dataset and produces a set of SIBNs. Each SIBN has at least s_{min} accounts (here, 30) and a weighted density of d_{min} (here, 0.75). Note that because the method we propose is unsupervised, it is possible that it captures human communities as well as SIBNs. In practice, we find that the method returns almost exclusively subgraphs of bots, and that where this is not the case we can easily differentiate SIBNs from “real” sub-communities.

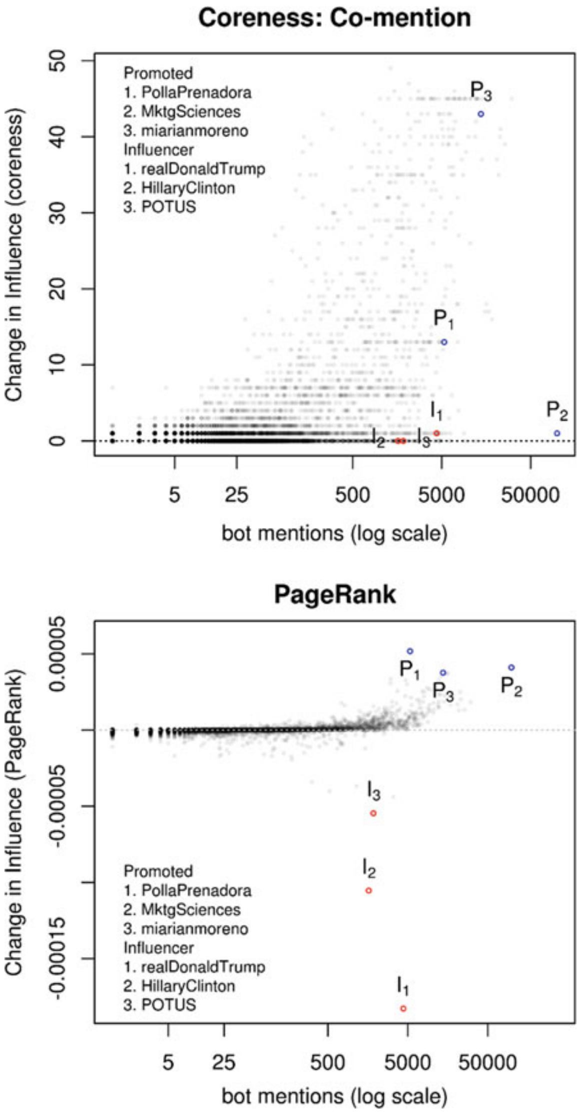
5 Results

5.1 *Differentiating Promoted Accounts from Community Influencers*

Figure 2 depicts the effect of core-bot activity on non-core-bot accounts in the ALT16 dataset with respect to coreness (top panel) and PageRank (bottom panel). The figure shows that the removal of core bots significantly worsens (increases) the ranking on these metrics for a large set of promoted accounts. For PageRank, this removal significantly improves (decreases) the rankings for a smaller set of community influencers. For coreness, we simply see no change in these rankings. Examples of highly promoted accounts (blue labeled points) and community influencers (red labeled points) are shown in Fig. 2.

To perform the analysis for Fig. 2, we use $\gamma_{0.995} = 5.94$ to select 507 core bots. For each panel, in Fig. 2, the x-axis depicts the (log-scaled) number of times a given user is mentioned by these 507 accounts. The y-axis depicts change in centrality with respect to these two metrics for all non-core-bot accounts when core-bot edges are removed. Larger positive values with respect to the y-axis therefore indicate inflated metrics based on core-bot activity. These are promoted accounts. Large negative values on the bottom panel imply that PageRank improves once core bots are removed. These accounts, although highly mentioned by core bots, actually increase in centrality when core bots are removed. These users are community influencers.

Fig. 2 Depicts the effect of core-bot activity on non-core-bot accounts with respect to coreness (top panel) and PageRank (bottom panel) in ALT16 dataset. In each panel the *x*-axis depicts the number of times a given user is mentioned by core bots in log-scale. The *y*-axis depicts change in coreness and PageRank in the top and bottom panels, respectively, when core-bot edges are removed



To further characterize promoted users and community influencers across our datasets, Tables 3 and 4 depict top community influencers and promoted accounts in the ALT16 and EUR17 datasets, respectively. To determine this set of users, we first sample the set of accounts that represent the top 0.99 percentile of users mentioned by core bots. Promoted users in Table 3 are then the users in this collection who are most negatively impacted by the removal of core bots. In contrast, the community influencers are those that most benefit on these metrics from the removal of bots.

Table 3 Top 0.99 percentile of users mentioned by core bots in the ALT16 dataset

Rank	Promoted accounts	Community influencers
1	PollaPrenadora	realDonaldTrump
2	MktgSciences	HillaryClinton
3	miarianmoreno	POTUS
4	monicasloves	YouTube
5	saravastiares	FoxNews
6	Alicelovelb	CNN
7	webcamfamosas	nytimes
8	NudeArt6969	mitchellvii
9	jimkoz69.jim	seanhannity
10	verovvp	Cernovich

The left column depicts promoted users whose PageRank decreases the most when core-bot activity is removed. The right column depicts the top 10 users whose PageRank increases most when core-bot activity is removed

Table 4 Top 0.99 percentile of users mentioned by core bots in the EUR17 dataset

Rank	Promoted accounts	Community influencers
1	dilruba_lees	YouTube
2	PollaPrenadora	rianru
3	goodenough03	Pravitelstvo_RF
4	NudeArt6969	MID_RF
5	saravastiares	KremlinRussia
6	Isobg69	history_RF
7	monicasloves	mod_russia
8	patdefranchis	zvezdanews
9	V_Samokhova	kpru
10	lenlekk	wordpressdotcom

The left column depicts promoted users whose PageRank decreases the most when core-bot activity is removed. The right column depicts the top 10 users whose PageRank increases most when core-bot activity is removed

The top community influencers in Tables 3 and 4 consist of accounts one would expect to be influential within each datasets centered around far right political and the Euromaidan movement, respectively. In contrast, manual inspection of promoted accounts highlights a variety of promotional but not necessarily relevant interests. One obvious promotional interest is pornography, highlighted by promoted users 1, 5, 7, 8, and 10 in Table 3, and nearly all promoted users in Table 4. Interestingly, although datasets were seeded from an entirely different set of users, we see that promoted accounts 1, 4, 5, and 8 appear in both datasets. Promoted accounts 2, 6, and 9 in Table 3 all provide links to third party community management applications that facilitate core-bot activity on behalf of "cyborg" subscribers. Finally, promoted accounts 3 and 4 appear to be bloggers who subscribe to some type of community management application as their timelines appear to contain posts containing human language, while others contain strings of mentions.

Although similar behavior is observed in the SYR15 dataset, it is not at the same scale as what we observe in the EUR17 and ALT16 datasets. We expect this is due in part to the conservative choice of p used here. Consequently, the proposed method may be improved by leveraging a more automated mechanism of selecting the p parameter of γ.

Regardless, however, results show that the methods we develop do a good job of differentiating between community influencers and promoted accounts. After doing so, we found that promoted accounts served a variety of different purposes, most frequently pornography or subscriptions to Twitter applications that promote cyborg behavior. In the following section, we delve deeper into the different SIBNs we found in each dataset and how they help to understand these various purposes for promoted accounts.

5.2 Detected SIBNs

We ran the method and found 11, 4, and 1 distinct SIBNs in the ALT16, EUR17, and SYR15 datasets, respectively. We will now discuss how these networks function and how they can be used for a variety of promotional ends.

To illustrate how SIBNs function, we use an SIBN from the SYR15 dataset. We call this SIBN "Firibinome" because of its easily identifiable naming convention, as is illustrated in one of its core-bot tweets depicted in Fig. 1. Firibinome is another politically motivated SIBN—each of the core-bot accounts shares the same profile image, a flag associated with Jabhat al-Nusra the predominant al-Qaeda affiliate in Syria.

Figure 3 depicts the directed mention graph of all accounts mentioned by the Firibinome SIBN. Nodes represent users within the neighborhood of the SIBN and are colored based on coreness within the co-mention graph; they are sized based on in-degree coreness which emphasizes users who are mentioned by people who are highly mentioned. Mention core edges are depicted in red. The promoted user is highlighted in blue and represents a charitable foundation for Syrian children alleged to be a revenue generating scam for Jabhat al-Nusra.

The hashtags in posts retweeted by the Firibinome SIBN are summarized in Fig. 4 and are clearly consistent with Jabhat al-Nusra's interests in the region. Without analyzing SIBNs at the graph level, one would fail to identify how the mention cores network structure influences more sophisticated measures of influence.

While SIBNs all seem to serve the purpose of inflating the importance of particular users[2] it is important to understand that within each dataset we find

[2] And, although we have not discussed it here, likely content as well.

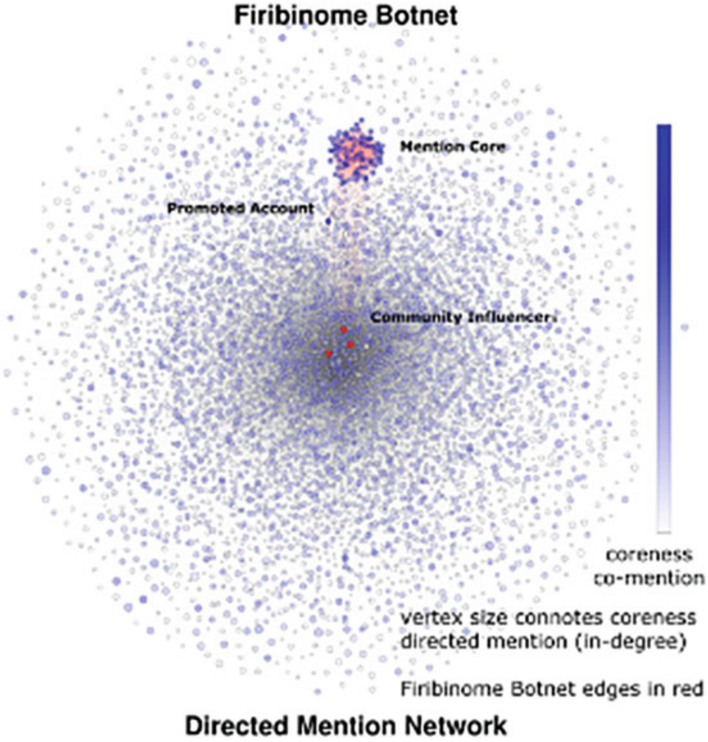


Fig. 3 Depicts the Firibinome SIBN. Nodes represent users within the neighborhood of the SIBN and are colored based on coreness within the co-mention graph and sized based on in-degree coreness in the mention graph. Mention core edges are depicted in red. The promoted user is highlighted in blue and example community influencers in red. Examples of community influencers are in red and represent pro-Nusra propaganda as well as an Islamic Scholar Hani al Sibai a well-known al-Qaeda supporter

multiple distinct SIBNs that have different goals with respect to which users and themes they are interested in promoting.

For example, the 507 ALT16 core bots form 11 distinct SIBNs when we run dense subgraph detection. As depicted in Fig. 7, many of these SIBNs have identifiable and differentiable promotional agendas. Figure 7 summarizes the most dense, distinct subgraphs consisting of between 30 and 200 users within the Alt-Right Twitter Search and found with the dense subgraph detection method described above. The *x-axis* connotes weighted subgraph density, while the *y-axis* connotes size in terms of users. Black circles in Fig. 7 display collections of users that we have manually identified as clearly SIBNs. We identify each group's promotional objective through manual inspection by summarizing the type of content retweeted. The two subgraphs showing highest density refer to third party applications designed to increase the social influence of subscribers and profile descriptions of core bots state these objectives explicitly. Religious and politically

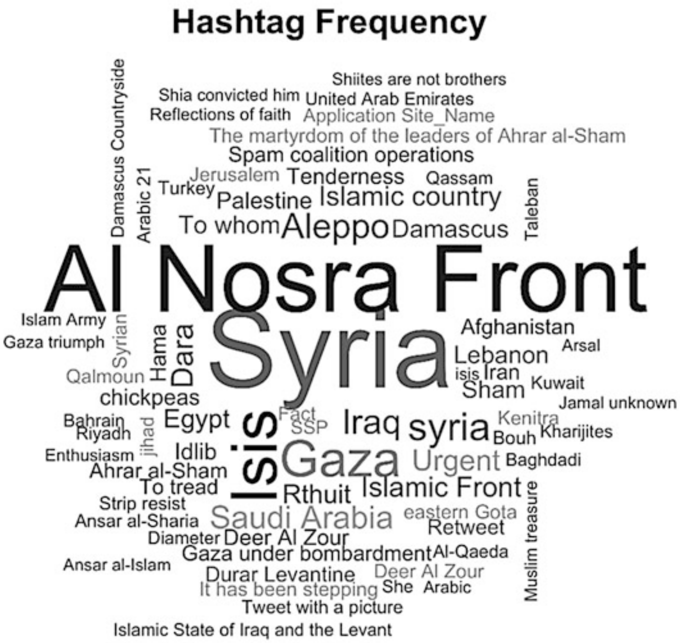


Fig. 4 Depicts translated hashtags retweeted by core bots within the Firibinome SIBN

focused cores like the Evangelical, #opisis, and #BoycottIsrael can clearly be identified by summarizing retweeted hashtags in a manner similar to those depicted in Fig. 4. In each of these cases, the communities could be formed by real users via third party applications, which will be discussed later in this section.

SIBNs can also be observed to leverage particular content to achieve very targeted objectives. For example, an SIBN in the Euromaidan dataset was found that consists of 14 Twitter accounts designed to bridge two distinct online communities. In the Spring of 2015, each tweet from the core of this SIBN shared nude images of women and strings of mentions that pointed to accounts that either shared similar images or inflammatory news updates covering Ukrainian government corruption and Russian occupation of Crimea. The most frequent hashtags shared by the botnet are translated to English and depicted in Fig. 5. Black terms highlight those hashtags associated with news and propaganda.

The community structure of this SIBN is depicted in Fig. 6 and is parameterized identically to Fig. 3. The density of red edges bridging the two communities is highlighted in the figure, and when the botnet's combined betweenness centrality makes it the most powerful bridge in the entire EUR17 dataset. Although the objective of this SIBN is unknown, such behavior would be consistent with campaigns designed to expose young men to recruiting propaganda.

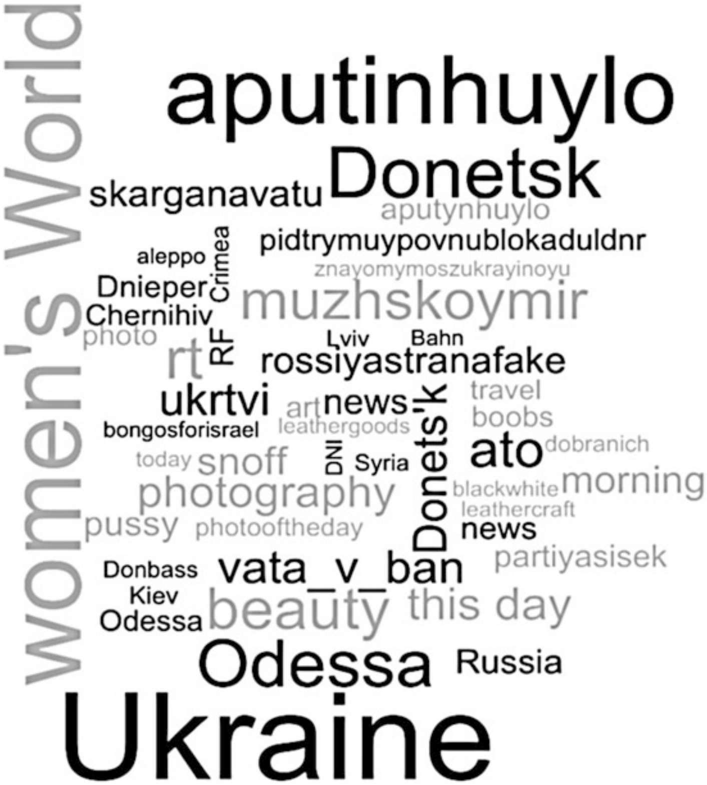


Fig. 5 Depicts translated hashtags retweeted by core bots within the Euromaidan Image Sharing SIBN. Terms in black highlight content associated with Euromaidan propaganda and Russian occupation of Crimea. Terms in grey are predominantly associated with the sharing of pornographic pictures

In some cases, SIBNs are the product of third party applications advertised as “influence enhancers” or “community managers.” Four examples, #Influence-Marketers, Evangelical, #opisis, and #followback, are depicted in Fig. 7. In each example, these communities consist of accounts that appear to be real users who retweet and post content that is core-bot-like with respect to mentions. In each case the post source annotated in the tweet json indicates that they were posted from third party applications like Follow Friday Assistant or commune.it which can generate core-bot behavior. For example, Follow Friday Assistant generates and post lists of accounts a user has interacted with in the past week, day, or month, and will split long lists into multiple tweets. In these cases, core-bot accounts are “cyborgs” or bot-assisted-humans [48]. The content associated with each SIBN summarized in Fig. 7 suggests these third party applications are leveraged for politically motivated reasons. These applications allow cyborg users to post with high volume and generate large online activist communities. In fact, one could argue

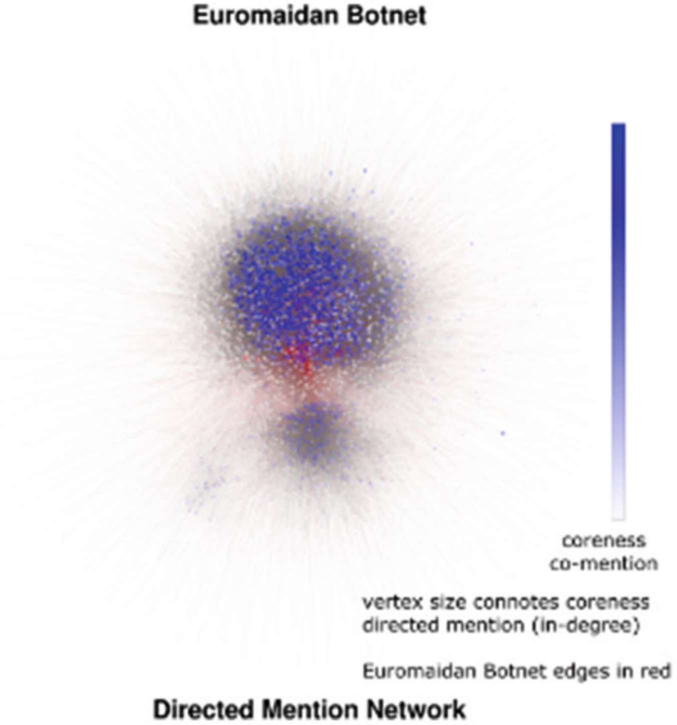

Euromaidan Botnet
coreness
co-mention
vertex size connotes coreness
directed mention (in-degree)
Euromaidan Botnet edges in red
Directed Mention Network

Fig. 6 Depicts the Euromaidan Image Sharing SIBN. Nodes represent users within the neighborhood of the SIBN and are colored based on coreness within the co-mention graph and sized based on in-degree coreness in the mention graph. Mention core edges are depicted in red. This SIBN appears to be designed to first build a community through image sharing and then to expose young men in that community to recruiting propaganda

that these third party applications slower the technical threshold for users to apply directed social engineering to accomplish their marketing objectives. This activity makes understanding online influence increasingly complex.

6 Conclusion

In this paper, we have studied the problem of automated social engineering in Twitter. We have introduced a specific class of social bot networks, social influence bot networks (SIBNs). SIBNs manipulate topic groups by creating an artificial core-like group in Twitter mention and co-mention graphs, promote particular users and cites, build connections among actors, and alter messages. We employed novel network techniques using meta-networks with multiple types of nodes to identify these SIBNs and to understand their usage. Specifically with respect to SIBNs we show:

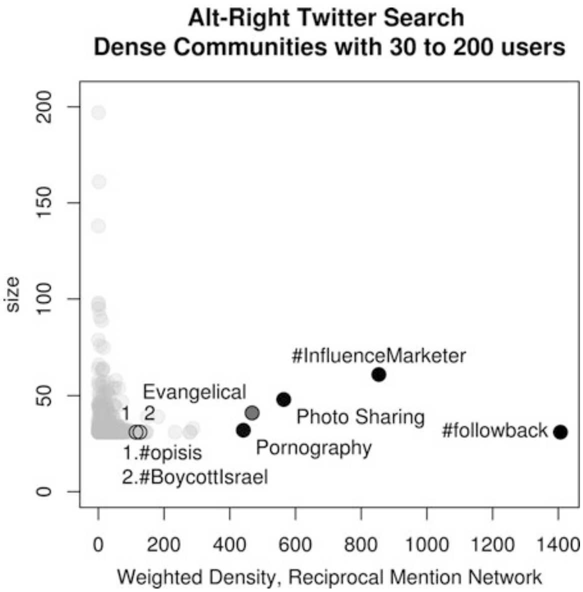


Fig. 7 Summarizes the most dense, distinct subgraphs consisting of between 30 and 200 users within the Alt-Right Twitter Search. The *x*-axis connotes weighted subgraph density, while the *y*-axis connotes size in terms of users. Black circles are SIBNs, but many of the subgraphs showing high graph density exhibit core-bot-like behavior and identifiable promotional agendas

- SIBNs are pervasive in a large number of domains and are influencing many different topic-groups.
- SIBNs are used to greatly inflate complex influence metrics on the social network induced by reciprocal mentioning behavior on Twitter.
- SIBNs formed multiple sub-communities of bots or cyborgs within each dataset, each with distinct intentions. These included SIBNs with more traditional foci (e.g., explicit influence marketing) as well as those aimed towards more nefarious goals (e.g., promoting particular political ideologies).
- Bot creators and cyborg users within our communities used directed social engineering to accomplish particular goals.

This work is limited in that we have not attempted to study these SIBNs influence with respect to following ties or content diffusion. For example, similar methods could be used to generate trending hashtags or disseminate URLs. Furthermore, this work implies the need for a detailed study of third party applications and their use for social engineering.

SIBNs have at their core an automated echo-chamber. This social structure raises interesting ethical questions. In effect, the methods described herein can be used to promote a perception of popularity or validity that does not necessarily reflect real user opinion making it increasingly difficult to determine trustworthiness. Moreover, an industry designed to market within this online ecosystem has emerged leveraging these behaviors and others with great effect. As Twitter has become one of the world's largest publication platforms the implications of providing an API that facilitates such activity merits future research.

Bots, cyborgs, and the special form of these as SIBNs are pervasive. They appear to operate in part by exploiting the social media architecture's recommendation systems and the way in which we as humans learn, recall, and make sense of information. The SIBNs are particularly powerful at manipulating groups in social media because they engage in strategic information maneuvers that create change at two levels—the social network and the shared knowledge network. No doubt, the specific form these bots will take will evolve, as will the form of these bots on other social media platforms. However, we anticipate that those bots and cyborgs that are most effective will be those that like SIBNs manipulate human activity in terms of both who is talking to whom and who shares what information or opinions.

References

1. Friedkin, N.E.: A Structural Theory of Social Influence, vol. 13. Cambridge University Press, Cambridge (2006)
2. Yardi, S., Romero, D., Schoenebeck, G., Boyd, D.: Detecting spam in a twitter network. First Monday 15, 1 (2009). http://firstmonday.org/ojs/index.php/fm/article/view/2793/2431?utm_source=twitterfeed&utm_medium=twitter
3. Zhang, J., Zhang, R., Zhang, Y., Yan, G.: On the impact of social botnets for spam distribution and digital-influence manipulation. In: 2013 IEEE Conference on Communications and Network Security (CNS), pp. 46–54. IEEE (2013). http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6682691
4. Ratkiewicz, J., Conover, M., Meiss, M., Gonçalves, B., Flammini, A., Menczer, F.: Detecting and tracking political abuse in social media. ICWSM. **11**, 297–304 (2011). http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/viewFile/2850/3274/
5. Ratkiewicz, J., Conover, M., Meiss, M., Bruno, G., Patil, S., Flammini, A., Menczer, F.: Truthy: mapping the spread of astroturf in microblog streams. In: Proceedings of the 20th International Conference Companion on World Wide Web, pp. 249–252. ACM, New York (2011). http://dl.acm.org/citation.cfm?id=1963301
6. Woolley, S.C.: Automating power: social bot interference in global politics. First Monday 21, 4 (2016). http://journals.uic.edu/ojs/index.php/fm/article/view/6161
7. Ferrara, E., Varol, O., Davis, C., Menczer, F., Flammini, A.: The rise of social bots. Commun. ACM. **59**(7), 96–104 (2016). http://dl.acm.org/citation.cfm?id=2818717
8. Zhang, J., Zhang, R., Zhang, Y., Yan, G.: The rise of social botnets: attacks and countermeasures. IEEE Trans. Dependable Secure Comput. **99**, 1 (2016). https://doi.org/10.1109/TDSC.2016.2641441
9. Abokhodair, N., Yoo, D., McDonald, D.W.: Dissecting a social botnet: growth, content and influence in Twitter. In: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, pp. 839–851. ACM, New York (2015). https://doi.org/10.1145/2675133.2675208
10. Wei, W., Joseph, K., Liu, H., Carley, K.M.: The fragility of Twitter social networks against suspended users. In: Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 9–16. ACM, New York (2015)
11. Berger, J.M., Morgan, J.: The ISIS Twitter Census: defining and describing the population of ISIS supporters on Twitter. The Brookings Project on US Relations with the Islamic World 3, 20 (2015)
12. Al-khateeb, S., Agarwal, N.: Examining botnet behaviors for propaganda dissemination: a case study of ISIL's beheading videos-based propaganda. In: 2015 IEEE International Conference on Data Mining Workshop (ICDMW) (2015-11), pp. 51–57 (2015). https://doi.org/10.1109/ICDMW.2015.413

13. Bessi, A., Ferrara, E.: Social bots distort the 2016 U.S. Presidential election online discussion. First Monday 21, 11 (2016). http://firstmonday.org/ojs/index.php/fm/article/view/7090
14. Ghosh, S., Viswanath, B., Kooti, F., Sharma, N.K., Korlam, G., Benevenuto, F., Ganguly, N., Gummadi, K.P.: Understanding and combating link farming in the twitter social network. In: Proceedings of the 21st International Conference on World Wide Web, pp. 61–70. ACM, New York (2012). http://dl.acm.org/citation.cfm?id=2187846
15. Liu, Y., Tang, M., Zhou, T., Do, Y.: Core-like groups result in invalidation of identifying superspreader by k-shell decomposition. arXiv preprint arXiv:1409.5187 (2014)
16. Carley, K.M.: Group stability: a socio-cognitive approach. Adv. Group Process. **7**(1), 44 (1990)
17. Carley, K.M., Martin, M.K., Hirshman, B.R.: The etiology of social change. Top. Cogn. Sci. **1**(4), 621–650 (2009)
18. Benigni, M., Joseph, K., Carley, K.M.: Online extremism and the communities that sustain it: detecting the ISIS supporting community on Twitter. PLoS One. **12**(12), e0181405 (2017)
19. Carley, K.M.: ORA: a toolkit for dynamic network analysis and visualization. In: Alhajj, R., Rokne, J. (eds.) Encyclopedia of Social Network Analysis and Mining. Springer, New York (2017). https://doi.org/10.1007/978-1-4614-7163-9_309-1
20. Viswanath, B., Mondal, M., Clement, A., Druschel, P., Gummadi, K.P., Mislove, A., Post, A.: Exploring the design space of social network-based Sybil defenses. In: 2012 Fourth International Conference on Communication Systems and Networks (COMSNETS 2012), pp. 1–8 (2012). https://doi.org/10.1109/COMSNETS.2012.6151333
21. Yang, Z., Wilson, C., Wang, X., Gao, T., Zhao, B.Y., Dai, Y.: Uncovering social network sybils in the wild. ACM Trans. Knowl. Discov. Data (TKDD). **8**(1), 2 (2014). https://doi.org/10.1145/2556609
22. Messias, J., Schmidt, L., Oliveira, R., Benevenuto, F.: You followed my bot! Transforming robots into influential users in Twitter. First Monday 18, 7 (2013)
23. Page, L., Brin, S., Motwani, R., Winograd, T.: The PageRank Citation Ranking: Bringing Order to the Web. Technical Report. Stanford InfoLab. 2 (1999)
24. Chen, J., Saad, Y.: Dense subgraph extraction with application to community detection. IEEE Trans. Knowl. Data Eng. **24**(7), 1216–1230 (2012). https://doi.org/10.1109/TKDE.2010.271
25. Boshmaf, Y., Muslukhov, I., Beznosov, K., Ripeanu, M.: The social bot network: when bots socialize for fame and money. In: Proceedings of the 27th Annual Computer Security Applications Conference (2011) (ACSAC ’11), pp. 93–102. ACM, New York (2011). https://doi.org/10.1145/2076732.2076746
26. Boshmaf, Y., Muslukhov, I., Beznosov, K., Ripeanu, M.: Design and analysis of a social botnet. Comput. Netw. **57**(2), 556–578 (2013)
27. Freitas, C., Benevenuto, F., Ghosh, S., Veloso, A.: Reverse engineering social bot infiltration strategies in Twitter. In: Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 25–32. ACM, New York (2015). http://dl.acm.org/citation.cfm?id=2809292
28. Zhang, C.M., Paxson, V.: Detecting and analyzing automated activity on Twitter. In: Passive and Active Measurement, pp. 102–111. Springer, Berlin, Heidelberg (2011). https://doi.org/10.1007/978-3-642-19260-9_11
29. Wu, L., Hu, X., Morstatter, F., Liu, H.: Adaptive Spammer Detection with Sparse Group Modeling. ICWSM (2017)
30. Ferrara, E., Varol, O., Menczer, F., Flammini, A.: Detection of promoted social media campaigns. In: Tenth International AAAI Conference on Web and Social Media (2016)
31. Romero, D.M., Tan, C., Kleinberg, J.: On the interplay between social and topical structure. In: Proceedings of the 7th International AAAI Conference on Weblogs and Social Media (ICWSM) (2013)
32. Joseph, K., Carley, K.M.: Culture, Networks, Twitter and Foursquare: Testing a Model of Cultural Conversion with Social Media Data (2015)
33. Cresci, S., Di Pietro, R., Petrocchi, M., Spognardi, A., Tesconi, M.: The paradigm-shift of social spambots: evidence, theories, and tools for the arms race. arXiv preprint arXiv:1701.03017 (2017). https://arxiv.org/abs/1701.03017

34. Ruths, D., Pfeffer, J.: Social media for large studies of behavior. Science. **346**(6213), 1063–1064 (2014). https://doi.org/10.1126/science.346.6213.1063
35. Tufekci, Z.: Big questions for social media big data: representativeness, validity and other methodological pitfalls. In: ICWSM '14: Proceedings of the 8th International AAAI Conference on Weblogs and Social Media (2014). http://arxiv.org/abs/1403.7400
36. Viswanath, B., Bashir, M.A., Zafar, M.B., Bouget, S., Guha, S., Gummadi, K.P., Kate, A., Mislove, A.: Strength in numbers: robust tamper detection in crowd computations. In: Proceedings of the 2015 ACM on Conference on Online Social Networks (COSN '15), pp. 113–124. ACM, New York (2015). https://doi.org/10.1145/2817946.2817964
37. Kakhki, A.M., Kliman-Silver, C., Mislove, A.: Iolaus: securing online content rating systems. In: Proceedings of the 22nd International Conference on World Wide Web (WWW '13), pp. 919–930. ACM, New York (2013). https://doi.org/10.1145/2488388.2488468
38. Gupta, A., Lamba, H., Kumaraguru, P., Joshi, A.: Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy. In: Proceedings of the 22nd International Conference on World Wide Web, pp. 729–736. ACM, New York (2013). http://dl.acm.org/citation.cfm?id=2488033
39. Conover, M., Ratkiewicz, J., Francisco, M.R., Gonçalves, B., Menczer, F., Flammini, A.: Political polarization on twitter. ICWSM. **133**, 89–96 (2011). http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/viewFile/2847/3275.pdf
40. Conrad Nied, A., Stewart, L., Spiro, E., Starbird, K.: Alternative narratives of crisis events: communities and social botnets engaged on social media. In: Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, pp. 263–266. ACM, New York (2017). http://dl.acm.org/citation.cfm?id=3026307
41. Zhang, X., Li, Z., Zhu, S., Liang, W.: Detecting spam and promoting campaigns in Twitter. ACM Trans. Web. **10**(1), 4:1–4:28 (2016). https://doi.org/10.1145/2846102
42. Goodman, L.A.: Snowball sampling. Ann. Math. Stat. **32**(1), 148–170 (1961). https://doi.org/10.1214/aoms/1177705148
43. Benigni, M.: Detection and analysis of online extremist communities. Ph.D. Thesis, School of Computer Science, Carnegie Mellon University (2016)
44. Bonacich, P.: Power and centrality: a family of measures. Am. J. Sociol. **92**(5), 1170–1182 (1987)
45. Kitsak, M., Gallos, L.K., Havlin, S., Liljeros, F., Muchnik, L., Eugene Stanley, H., Makse, H.A.: Identification of influential spreaders in complex networks. Nat. Phys. **6**(11), 888–893 (2010)
46. Kwak, H., Lee, C., Park, H., Moon, S.: What is Twitter, a social network or a news media. In: Proceedings of the 19th International Conference on World Wide Web, pp. 591–600. ACM, New York (2010)
47. Kumar, R., Raghavan, P., Rajagopalan, S., Tomkins, A.: Trawling the Web for emerging cyber-communities. Comput. Netw. **31**(11), 1481–1493 (1999)
48. Chu, Z., Gianvecchio, S., Wang, H., Jajodia, S.: Who is tweeting on Twitter: human, bot, or cyborg? In: Proceedings of the 26th Annual Computer Security Applications Conference, pp. 21–30. ACM, New York (2010). http://dl.acm.org/citation.cfm?id=1920265