2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining

How the Tables Have Turned: Studying the New Wave of Social Bots on Twitter Using Complex Network Analysis Techniques

Pujan Paudel, Trung T. Nguyen, Amartya Hatua and Andrew H. Sung
School of Computing Sciences and Computer Engineering, The University of Southern Mississippi
Hattiesburg, Mississippi 39406, U.S.A.
Email: {pujan.paudel, trung.nguyen, amartya.hatua, andrew.sung}@usm.edu

***Abstract*—Twitter bots have evolved from easily-detectable, simple content spammers with bogus identities to sophisticated players embedded in deep levels of social networks, silently promoting affiliate campaigns, marketing various products and services, and orchestrating or coordinating political activities. Much research has been reported on building accurate machine learning classifiers to identifying bots in social networks; recent works on social bots have started the new line of research on the existence, placement, and functions of the bots in a collective manner. In this paper, we study two families of Twitter bots which have been studied previously with respect to spamming activities through advertisement and political campaigns, and perform an evolutionary comparison with the new waves of bots currently found in Twitter. We uncover various evolved tendencies of the new social bots under social, communication, and behavioral patterns. Our findings show that these bots demonstrate evolved core-periphery structure; are deeply embedded in their networks of communication; exhibit complex information diffusion and heterogeneous content authoring patterns; perform mobilization of leaders across communication roles; and reside in niche topic communities. These characteristics make them highly deceptive as well as more effective in achieving operational goals than their traditional counterparts. We conclude by discussing some possible applications of the discovered behavioral and social traits of the evolved bots, and ways to build effective bot detection systems.**

***Index Terms*—social bots, twitter, network analysis**

I. Introduction

In the recent years, online social networking platform, such as Twitter has seen a massive increase in the presence of automated accounts, or bots. The bots have evolved their range of activities from polluting the content of Twitter feed [1], affiliate marketing [2], and link farming [3], to manipulate elections [4] and public opinions [6] by spreading malicious content on Twitter.

There has been recent progress on detection of bots on Twitter, such as the work by Botometer [7], which leverages more than one thousand features (based on users, friends, network, temporal, linguistic and sentiment) of tweets and users to classify a Twitter account as a bot. For the classification of malicious and benign accounts on Twitter different machine learning methods are applied using different features from user profiles [8], graph-based features [9] and temporal features [10]. However, as the detection algorithms and strict policy to monitor those accounts have improved, the bots on Twitter have also evolved. The spambots have evolved into social bots [11], adopting complex content posting and social interaction patterns. Some of the emerging trends in Twitter bot research have deviated from crafting machine learning features to propose new social dimensions to fight and overcome the new wave of social bots. Some of those new dimensions include study of lockstep behaviors between user tweets [12], detection of latent group anomalies in graph [13], and similarity between online user behaviors modeled by digital DNA sequences [14].

The work of Cresci et al. [11] highlights the paradigm shift of social spambots by introducing a novel dataset of Twitter bots active on three different cases. They extend the idea of analyzing the collective behaviors of social bots, rather than separating individual accounts for malicious behaviors. The new social bots identified has a very high survival rate with 96.5% of them still being active on Twitter, which demonstrates the highly deceptive capability of the bots. Surprisingly, the new wave of social bots presented by [11] were able to fool human annotators on crowdsourcing campaigns as the annotators obtained an accuracy of less than 24% in identifying the social spambots with a heavy proportion of False Negatives. Even the state-of-the-art public bot detection service, Botometer demonstrated very low recall on detecting such bots. The social bots also remain largely undetected through other techniques of spambot classification like supervised classification, unsupervised classification, and graph clustering-based approaches, which demonstrates real threat of the new wave of social bots. Despite a relatively dormant spamming behavior of the social bots, those bots were able to generate replies and retweet interactions from genuine human accounts.

The contribution of our work is a detailed comparative analysis of the evolution of new wave of social bots from

ASONAM '19, August 27-30, 2019, Vancouver, Canada

ACM ISBN 978-1-4503-6868-1/978-1-4503-6868-1/19/08
http://dx.doi.org/10.1145/3341161.3342898

three different dimensions of study. We conclude our paper by discussing possible exploratory bot detection directions for further discovery of social bots on Twitter. Our work is largely inspired by the direction of recent paradigm shift of Twitter bots discussed in the works of [11], [12], [13]. The major objectives of our paper are:

- To contribute to the literature of new waves of social bots by studying in detail two different types of such bots on Twitter: 1) Political Bots, 2) Advertisement Bots and comparing them with their traditional counterparts for each of the types.
- To analyze the evolution of social bots in terms of network structure by answering three important questions: ***RQ1)*** How do the new wave of social bots differ from traditional bots in terms of social network statistics and their organization of Core-Periphery structure? ***RQ2)*** How embedded are the social bots in their social as well as communication networks? ***RQ3)*** How do the networks of the social bots perform under robustness attack?
- To study the information diffusion and communication patterns of the social bots by answering the following questions: ***RQ4)*** How does the information cascade profiles of the social bots look like? ***RQ5)*** Do the bots have different communication leaders across different forms of communication networks?
- To perform detailed content analysis of the tweets produced by those bots by answering the following questions: ***RQ6)*** Do the social bots have any specific patterns of topic distribution over time? ***RQ7)*** Do the bots have some community specific content spreading behavior? Are there any niche topic communities?
- To discuss possible exploratory network analysis directions and advanced features for machine learning classifiers to detect ever-evolving plethora of novel social bots.

II. Datasets Used

We used four different datasets of Twitter bots coming from three different works for our study and comparative analysis of the Twitter bots. Throughout the rest of the paper, we refer to the previously studied bots as traditional bots and the new wave of bots as social bots. In this study advertisement and political bots are used. These bots are publicly available and previously studied in the literature. The results of the experiments are often effected by the different criteria of data collection process. Thus, we are applying a sound and verified criteria of separation between the two generation of bots. Our line of separation for the traditional bots is based on the content polluting behavior discussed on previous works on literature and the high proportion of deleted Twitter accounts of bots previously showing involvement in political campaigns. High survival rates as well as deceptive spamming patterns was considered to identify the bots as social bots.

A. Traditional Advertisement Bots

For the category of traditional advertisement bots, we used the dataset originally collected by [1]. Their dataset lacks any type of separation or grouping by advertisement campaigns. Therefore, we used the idea that content polluters mostly make use of URLs to spread spam about the campaign they are advertising, to cluster the domains of the URLs posted by the spammers. One of the top domains of URLs the content polluters tweeted about, called “Aweber”, an email marketing service provider was identified and spammers who tweeted the hashtags and URLs of it were filtered as traditional advertisement bots.

B. Traditional Political Bots

For the traditional political spammers, we used the labeled dataset used in the work of [15]. The bots in this dataset tweeted about Arab Spring activity in Libya, from February 3rd, 2011 to February 21st, 2013, and were labeled as bots based upon the deletion and suspension of accounts by Twitter services. We were only able to partially reconstruct the users and tweets from the dataset provided by the authors as most of the users had already been suspended by the Twitter platform.

C. Social Advertisement Bots

For our category of social advertisement bots, we used social spambots #2 dataset from the work of [11], which consists of manually labeled bots that spent several months promoting a mobile application, called Talnts, using the #TALNTS hashtag.

D. Social Political Bots

Finally, for our category of social political bots, we used the social spambots #1 dataset from [11] which consists of novel group of social bots being active on the 2014 Mayoral election of Rome employed by one of the runner-ups of the election to publicize his policies.

After expanding the metadata of all the valid tweets and the users, the statistics of the final dataset for the respective categories of Twitter bots under study is shown in Table I. It can be observed that the size of the bots and their tweets in the different classes is not comparable. This is mostly due to the largely missing data of the traditional bots, which have been taken down by Twitter since they were initially reported on their respective works. However, the methodology we design to answer our research questions are minimally impacted by the difference in data size. The hashtag based and topical analysis are not directly dependent on the number of tweets being studied. The network based studies we perform, are comparisons within the communication and social networks of the same dataset. We only analyze the behavior of the changed network structures observed between the datasets to avoid direct comparisons. The other remaining methodology of corelation analysis, uses proportional and relative ranking based measures within the same dataset.

Table I: Statistics about the datasets of categorical bots

Category	Tweets	Bots
Traditional Advertisers	37922	165
Social Advertisers	1418558	465
Traditional Political Spammers	1967	152
Social Political Spammers	1610016	992

III. Methodology: Creation of Social Interaction and Communication Networks

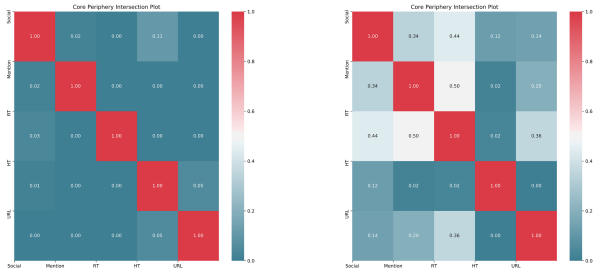
We create the following social and communication networks for each of the bot types under study. A) **Social Network:** is a directed network where there is an edge from a Bot A to Bot B if Bot A follows Bot B. B) **RT Network:** is an undirected network where there is an edge from Bot A to Bot B if both of them have retweeted *thres_rt* or a higher number of similar Twitter users. C) **Mention Network:** is an undirected network where there is an edge from Bot A to Bot B if both of them have mentioned *thres_mention* or a greater number of similar Twitter users. D) **URL Network:** is an undirected network where there is an edge from Bot A to Bot B if both of them have tweeted *thres_url* or a higher number of similar unique URLs. E) **HT Network:** is an undirected network where there is an edge from Bot A to Bot B if both of them have tweeted *thres_ht* or a higher number of similar unique hashtags. We applied the concept of slicing weighted networks to learn the value of threshold weights for the different communication networks. The respective thresholds were determined based upon the extraction of giant connected component from the dense networks. This also made the community detection algorithms computationally feasible. For the relatively sparse networks of traditional advertisement bots, social advertisement bots and traditional political bots, we used the threshold of 5 for *thres_rt*, *thres_mention*, *thres_url* and *thres_ht*. For a relatively denser network of social political bots, we used thresholds of 20, 30, 20 and 50 for *thres_rt*, *thres_mention*, *thres_url* and *thres_ht* respectively.

Throughout the rest of the paper, we refer to RT Network, Mention Network, URL Network and HT Network collectively as communication network of the bots.

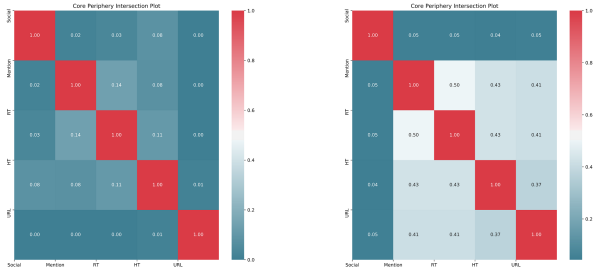
IV. Results

A. RQ1: Network Statistics and Core Periphery Structure

To gain a basic understanding of the social and communication networks we are analyzing, we initially studied the network statistics of the bots. The social network of the traditional advertisement and political bots were found to be more densely connected, having higher clustering and transitivity than that of social advertisement bots. On the other hand, the communication networks of the same traditional bots were very sparse, with lesser clustering and transitivity values. In the case of both types of social bots, the communication networks were denser, with higher clustering and transitivity than their traditional counterparts.



(a) i & ii



(b) iii & iv

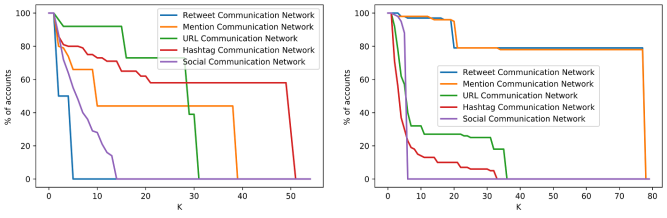
Figure 1: Network Core Members Intersection Plot of i) Traditional Advertisement Bots ii) Social Advertisement Bots iii) Traditional Political Bot iv) Social Political Bots

The social and communication networks of both types of traditional bots have a smaller core and a larger periphery which is connected very weakly to the core. The social bots in both cases have a larger, strongly connected network core and a relatively smaller size of peripheral nodes which are strongly connected to the core. The results demonstrate a close-knit and more focused network structure in the social bots, closer to the findings of human communication networks discussed in the work of [18].

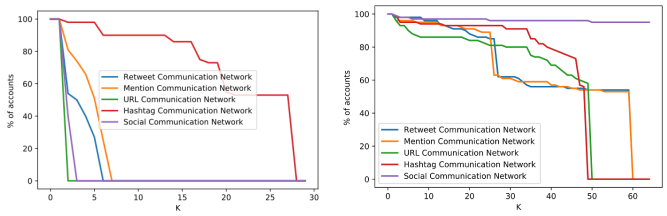
Next, we wanted to study if the core nodes operating on the network structure remain stable across the communication networks by plotting the intersection of the members of core on each of the network. As seen in the core members intersection plot in Figure 1, we found that the core nodes in the social bots remain more intact across the social and communication networks, whereas there is very little intersection between the core nodes of traditional advertisement and political bots. This demonstrates the evolution of social bots towards stable sets of principal actors in the central core structure, across the different communication channels.

B. RQ2: K-Core Decomposition

We apply K-Core decomposition to study the structure and embeddedness of bots in the graph. In Figure 2, we plot the



(a) i & ii



(b) iii & iv

Figure 2: K-Core Decomposition Analysis of i) Traditional Advertisement Bots ii) Social Advertisement Bots iii) Traditional Political Bot iv) Social Political Bots

percentage of accounts retained (y-axis) as we increase the value of K in the K-core (x-axis).

When we compare the social networks of traditional advertisement bots and social advertisement bots, we notice that even after considerable amount of increased core size of the traditional bots, much of the network core is retained. This behavior signifies the deeply embedded bots in their social networks, whereas, the social graphs of social bots appear to disrupt easily with the slightest increase of K. When we observe the decomposition graphs of communication channels of those same bots, social bots outnumber the traditional bots as the value of K increases. The results are similar with the case of political bots as with the increase of K, the accounts become more active in their communication channels than their traditional counterparts. This behavior is an evolved tendency of social bots to populate areas of the communication networks being more central and better connected, whereas being shallowly embedded in their social graphs. From the point of view of capital required to design and operate the bots, it is easy for bots to get deeply embedded in their social networks, but at the same time equally risky in terms of large scale moderation and detection of collectively acting bots. On the other hand, disrupting the bots through multiple dimensions of communication channels (Retweet, Hashtag, Mention) is more costly and requires more effort from a platform moderator or a bot detection service, which is exactly the strategy used by the new wave of social bots.

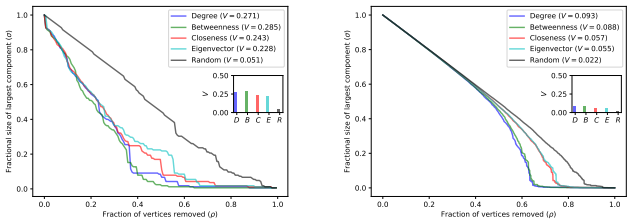
C. RQ3: Network Robustness

For Robustness Analysis of the social and communication networks, we used the open source implementation by [17] and adopted the sequential targeted attack approach. In this approach, centrality measures (degree, betweenness, closeness and eigenvector) is calculated for all the vertices in the initial network, and the vertex with highest centrality measure is removed. The centrality measures are recalculated for all vertices in the new network and the highest ranked vertex is removed, with the process repeating until desired fraction of vertices has been removed.

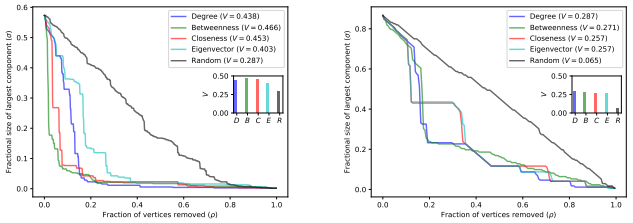
We subject the five different networks for all of the bots, traditional as well as social, used in our study to robustness analysis. In Figure 3 and Figure 4, we plot the fractional size of the largest component of the network against the proportion of removed network nodes in decreasing order of various centrality measures for different social and communication networks. The results show that social advertisement bots are highly resilient to network attacks, with significant proportion of size of largest component dropping only after more than 60% of its vertices being removed. We can also see that there is no specific vulnerability of centrality in the social graph of them as removing nodes based on different types of centrality leaders result in similar effect on the decrease on size of largest component. In contrast, size of the largest connected component in the social graph of traditional bots decreases substantially, before even 40% of the vertices are removed. The social graph of traditional advertisement bots is extremely prone to network failure, especially when the centrality leaders on Closeness and Eigenvector are attacked first. The other communication networks are equally vulnerable to robustness attacks through identical centrality leaders. Similarly, the outcomes of robustness attacks are identical, for traditional political bots and social political bots, with all forms of networks of social bots showing more resiliency to robustness attack than their traditional counterparts. An observation we noted is that amongst the communication networks of social political bots, the Mention network are relatively less robust, through the point of attack of betweenness centrality, while the Hashtags network of social advertisement bots are relatively more vulnerable to network failure.

D. RQ4: Information Diffusion

We collect the top 10 retweeted tweets from each of the dataset of bots in our study, which has at least a single occurrence of the campaign related hashtags, or URLs and which were retweeted more than 10 times by user accounts outside of the bot datasets. We then expanded the information diffusion timeline of the tweets by collecting information of the metadata of their retweets. We made sure our study analyzes the tweets that are primarily authored by the bots under study, by inspecting the “retweet_status_id” property of the tweet and expand the diffusion network. We study the temporal visualizations of retweeted tweets of different bot datasets by plotting the cumulative frequency of retweets originated by a tweet against the time (in hours) since the origin of the tweet in the x axis. The steeper the line, the faster the information was spread, and the height of the line (y axis) indicates the number of times it was retweeted. The timeline for the respective datasets of bots are in Figure 5.

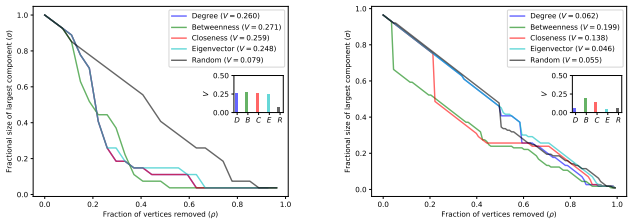


(a) i & ii

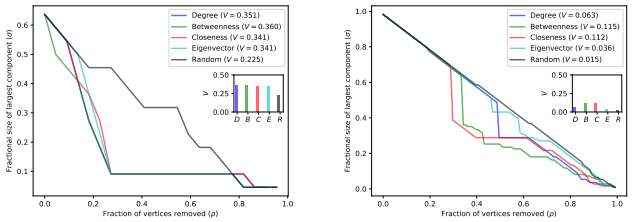


(b) iii & iv

Figure 3: Robustness Attack Test i) Friendship Network of Traditional Advertisement Bots ii) Friendship Network of Social Advertisement Bots iii) Hashtag Network of Traditional Advertisement Bots iv) HashTag Network of Social Advertisement Bots



(a) i & ii



(b) iii & iv

Figure 4: Robustness Attack Test i) Mention Network of Traditional Political Bots ii) Mention Network of Social Political Bots iii) RT Network of Traditional Political Bots iv) RT Network of Social Political Bots

The lines on the diffusion graph of traditional advertisement bots show fast and abrupt diffusion of information shared by very few actors. The tweets gathered a certain number of retweets within the first hour of their origin abruptly but were unable to gather further retweet along with time. In contrast, the social advertisement bots were more effective in diffusing information, showing both the patterns of slow diffusion occurring over a long period of time, shared by a few actors as well as fast diffusion of information shared by many actors. The results were identical in the case of social political bots, as the political bots were able to generate even larger number of multi-user, fast and sustained retweet chains of information diffusion of the social political bots.

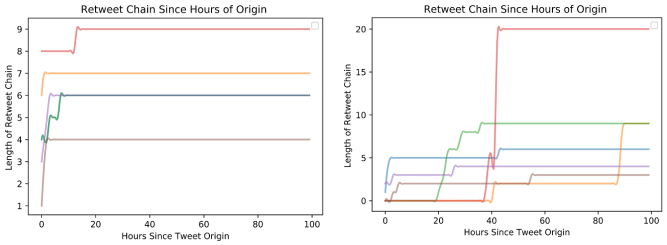
E. RQ5: Leaders Across Communication Networks

We examine if the leaders of centrality in the communication and social networks are consistent across the different communication networks, or new communication leaders arise within the communication networks. We calculated the betweenness centrality of the communication networks for each type of bots and calculated Kendalls correlation coefficient between the centralities of bots in the networks. Kendalls correlation plot allows us to visualize the ranked correlation between the leaders of the communication networks. Betweenness centrality was chosen out of the multiple available centrality measures as this measure has been used in prior work [19] to identify influencers in legitimate or criminal organizations, whose removal could seriously destabilize the organization.

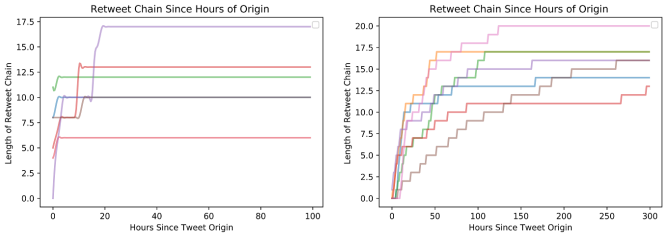
Our results show that in the case of traditional advertisement bots, they have positive values for correlation between social communication network leaders and leaders of other forms of communication. The other forms of communication networks also show increased positive correlation between each other than the social advertisement bots. The results are similar in the case of traditional political bots and social political bots, as shown in Figure 7. This lack of strong positive correlation between communication centralities leaders shows that the new wave of social Twitter bots deploy different role leaders across different communication channels to meet their campaign objectives, while the same leaders of social networks reoccur in the different communication networks in case of traditional bots.

F. RQ6: Topic Usage Over Time

In this study, we first build the topic model over the pre-processed tweets and divide all the tweets from the respective datasets into equal monthly time buckets and infer the tweets of those buckets against the learnt topic models. We apply basic pre-processing strategies to clean the text. We used Python's Gensim library to build Latent Dirichlet Allocation (LDA) topic models. For the LDA model, we set the number of topics to 10 and the number of iterations to 100. The normalized topic distribution weights, "norm_topic_weight" representing the (y-axis) with the increasing time (in months), from the initial tweet of the respective bot dataset is displayed in Figure 6. The differently colored lines represent the 10 different topics inferred from the topic model.

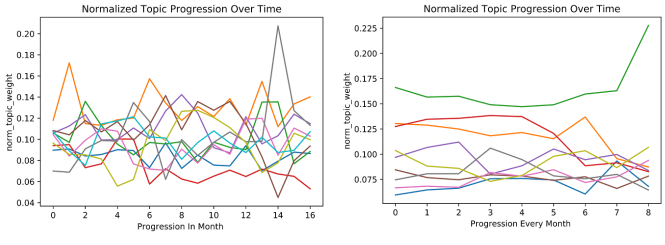


(a) i & ii

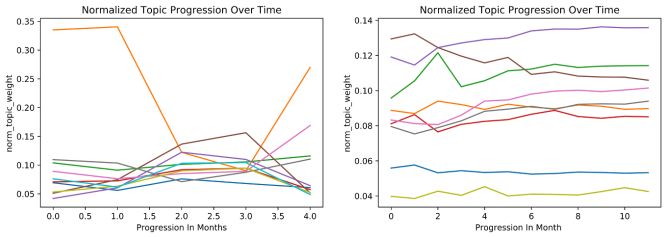


(b) iii & iv

Figure 5: Information Diffusion Timeline of i) Traditional Advertisement Bots ii) Social Advertisement Bots iii) Traditional Political Bot iv) Social Political Bots



(a) i & ii



(b) iii & iv

Figure 6: Normalized Topic Over Time Distribution of i) Traditional Advertisement Bots ii) Social Advertisement Bots iii) Traditional Political Bot iv) Social Political Bots

We compared the social bots with their traditional counterparts on the evolution of topics over time. As shown in Figure 6, the normalized topic distribution weight for almost all topics of the traditional advertisement bots increases and decreases abruptly over time. The transition between time buckets of the topic distribution not being smooth, with lots of edgy crests and falls, reflects on the bots instability on their topic's distribution throughout the time. In contrast, the topic distributions for the advertisement networks are distributed evenly across time. The results on this study are similar in the comparison of social political bots and their traditional counterparts, demonstrating the topical stability of the social bots over time.

G. RQ7: Niche Topic Community

Finally, we study if the bots spreadout across communities who intermingle their campaign tweets with tweets related to niche topics. We applied Louvain [16] community detection algorithm on the networks of the bots under comparison and separated them into communities. We then extracted the hashtags used by the users of the top communities of the networks and investigated the top hashtags from each community.

As seen on Table II, top hashtags used by the top communities of the traditional advertisement bots in each of the communities are almost similar (newsletter, Email Marketing, Marketing Tips). We also found some communities, tweeting numerals (#1, #2, #3) as their top tweets. Upon investigating those tweets, we found that those tweets were again related to Email Marketing tips (Tip #1, Tip #2). The top hashtags across the communities are fairly suspicious with regards to the intent of the deployed bots.

Table II: Top Hashtags Across Communities of Traditional Advertisement Bots

Comm 1	Comm 2	Comm 3	Comm 4
newsletter	emailmarketing	EmailMarketing	emailmarketing
photography	integrations	Autoresponder	integrations
LearnPhoto...	EmailMarketing	FreeEbook	marketing
eBook	marketingtips	SocialMedia	email
#1	cmworld	FreeTrial	network-marketing

Whereas, the top hashtags adopted by the social advertisement bots are different across different communities, displayed in Table III. We can observe a community dedicated on tweeting about a particular music artist (Community-1), another community which tweets about the public event happening at that time (Community-4), and also a community tweeting about web series (Community-2). One interesting observation we found was a specific community (Community-3), which interacted with a very dedicated community of Indie Music artists. They tweeted about promoting Indie Rap Music, and EDM sound, with the effort to recruit premium members inside their application, TALNTS. We also noted that their primary hashtag for promotion of their app, #Talnts, is ranked lower and dominated by some other hashtags in two different communities.

Similarly, the top hashtags across the community of traditional political bots, listed in Table IV are similar across the different communities, and they are solely related to the campaign they are deployed for, including hashtags which are potentially sensitive to the community. Finally, in the

Table III: Top Hashtags Across Communities of Social Advertisement Bots

Comm 1	Comm 2	Comm 3	Comm 4
TALNTS	ReekSpeaks	TALNTS	VoteUK-Arianators
WeLove-Justin	TALNTS	EDM	BaNvAfG
music	gossipgirl	EDM-SoundofLA	WorldCuP
HBD-JustinBieber	Halloween	iPromoteYou	MARCH
myxmusic-awards	TheOriginals	hiphopmusic	DubaiWorldCup
nowplaying	ISM2014	hotnewhiphop	LincolnTrialsAt-Wolves

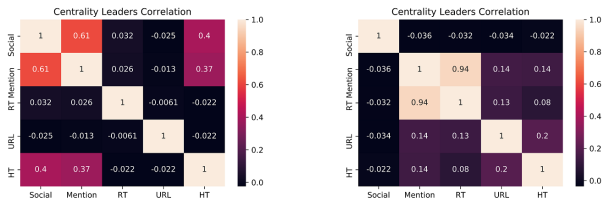
Table IV: Top Hashtags Across Communities of Traditional Political Bots

Comm 1	Comm 2	Comm 3	Comm 4
Benghazi	Benghazi	Benghazi	libya
Obama	CorruptMedia	Israel	muslim
BenghaziGate	StandDown	pakistani	arab
Obama	Libya	CNN	impeach-obama
Libya	Obama	News	Benghazi
GOP	7HoursOfHell	Obama	Obama
MSNBC	RedNationRisin	BenghaziGate	CNN
CorruptMedia	MediaBias	AddressTheNation	islamofascist

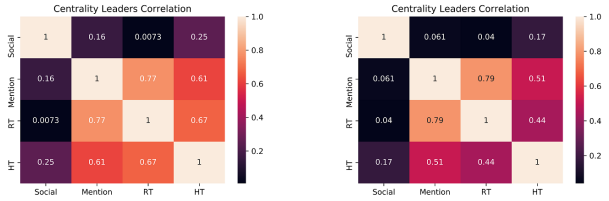
case of social political bots, the top hashtags across the communities, as well as the ranking of the top campaign related hashtags vary. We can see in Table V that all of the communities tweet about varieties of topics mixed with campaign related tweets with some communities even having a few external events as their top hashtags. Alongside mixing of campaign related hashtags with external hashtags, upon cross referencing the top hashtags with the event context, we found that the political bots have some communities dedicated for special sub campaigns within the larger political campaign.

Table V: Top Hashtags Across Communities of Social Political Bots

Comm 1	Comm 2	Comm 3	Comm 4
Renzi	Belloe-Democratico	Gomorra-LaSerie	dosomet-hingedgy
Berlusconi	Proverbi	SlotMachine	sVapevatelo
Napolitano	Governo	Palestina	Manovra
sapevatelo	Senato	Israele	curvy
jeep	WorldCup	roma	PeroniFWC
ibiza14	missitalia2013	Gaza	stopinvasione



(a) i & ii



(b) iii & iv

Figure 7: Centrality Leaders Correlation Plot of i) Traditional Advertisement Bots ii) Social Advertisement Bots iii) Traditional Political Bot iv) Social Political Bots

For example, Community-1 tweeted mostly about leaders and political figures associated with the event, Community-2 tweeted about senate related activity, Community-3 tweeted about an external event, related to other political turmoil in Palestine, mixed with their tweets, whereas, Community-4 was focused more on tweeting slogans and political ideologies.

Our observation of community-based hashtag study demonstrates how the social advertisement as well as political bots are effectively distributed across communities tweeting across diverse topics. Alongside tweeting about their primary campaign hashtags, in an attempt to remain undetected as well as gain a level of mutual trust in the human community they are embedded in, they tweet about specific topics and interact with specific communities. We were also able to discover sub-campaign related communities within the bots, which intermingle their sub-campaign objectives alongside varied, genuine looking tweets.

V. Towards new dimensions for Exploratory Social Bots Detection

The comparative social, behavioral, and content-based differences of the new wave of social bots with their traditional counterparts helps us to gain better understanding on the strategy that they adopt to remain unnoticed for a long time, the communication channels they utilize to gain interaction from genuine human users. Moreover, the understanding we gained by the evolutionary traits of these bots can be used on future to develop more robust tools to detect similar waves of social bots, which might have remained unnoticed till now. From the viewpoint of an exploratory detection of social bots, we demonstrated that graph-based study like K-Core decomposition of various signal-based networks can uncover the bots through communication mediums utilized by them.

Similarly, the depth of Information Diffusion trees, augmented by the temporal mining of retweet patterns could also be an equally effective further avenue for studying coordinated Twitter bots.

The results of network robustness attacks on various communication networks of social bots suggests that certain networks, like Mention and Hashtag (HT) network could be attacked from vulnerability points of centrality leaders, to disintegrate the network, and possibly explore the channels of communication that could be blocked to minimize the effects of the social bots. Based upon the economy of social bot design and objectives, we can also argue that the intersection of core-periphery structure across the communication networks can be another pattern to study for the presence of bots. From the network centrality point of view, relative ranked positioning of the centrality leaders across the social network as well as the communication networks could help us in revealing the information diffusion setup, even for the more sophisticated of the bots. Expanding the graph-based analysis, we also identified behavioral traits of the social bots, from a content analysis point of view. We studied how the new wave of social bots demonstrate human like content patterns on their tweets, by tweeting with similar intensity about a similar distribution of topics for a long amount of time. The utilization of niche topic communities for the promotion of their campaigns should be studied in depth for the bots who are continuously evolving and fighting against the adversaries of bot detection systems.

VI. Conclusion

Bots in social networks are becoming increasingly sophisticated and present great opportunities for novel applications in various areas, as well as tremendous challenges for their detection. In this work, we studied in detail the evolution of traditional spam bots into social bots by evaluating them with respect to three different dimensions of social activity.

Our study contributes to the literature of the growing study about the new wave of social bots in Twitter, which act in highly deceptive, yet effective coordinated fashions, and have shown human like interaction, and behavioral traits compared to the bots studied previously in the literature. The application of these behavioral traits we have discovered to explore different variety of coordinated Twitter bots in the wild would be a very interesting avenue for future work. We would also like to test how generalized our findings are, across different varieties and multiple case studies of modern social bots, after exploring them. Converting the experimental findings of our study to quantitative and statistical measures, which could possibly be extended to a real-time expert detection system of social bots, is a major remaining challenge as we look forward to join forces on bringing down these new waves of bots on Twitter.

References

[1] K. Lee, B. D. Eoff, and J. Caverlee, "Seven months with the devils: A long-term study of content polluters on twitter," In Fifth International AAAI Conference on Weblogs and Social Media, July 2011.

[2] A. A. Amleshwaram, N. Reddy, S. Yadav, G. Gu, and C. Yang, "Cats: Characterizing automation of twitter spammers," In 2013 Fifth International Conference on Communication Systems and Networks (COMSNETS), pp. 1–10, IEEE, January 2013.

[3] S. Ghosh, B. Viswanath, F. Kooti, N. K. Sharma, G. Korlam, F. Benevenuto, and K. P. Gummadi, "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, April 2012.

[4] A. Bessi, E. Ferrara, "Social bots distort the 2016 US Presidential election online discussion," First Monday, 21(11), November 2016, Available: https://firstmonday.org/article/view/7090/5653 (visited on July 12, 2019).

[5] A. Badawy, E. Ferrara, and K. Lerman, "Analyzing the digital traces of political manipulation: The 2016 russian interference twitter campaign," IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 258–265, IEEE, August 2018.

[6] M. Forelle, P. Howard, A. Monroy-Hernndez and S. Savage, "Political bots and the manipulation of public opinion in Venezuela," arXiv preprint arXiv:1507.07109, July 2015.

[7] C. A. Davis, O. Varol, E. Ferrara, A. Flammini, and F. Menczer, "Botornot: A system to evaluate social bots," 25th International Conference Companion on World Wide Web, pp. 273–274, International World Wide Web Conferences Steering Committee, April 2016.

[8] P. C. Lin, and P. M. Huang, "A study of effective features for detecting long-surviving Twitter spam accounts," 15th International Conference on Advanced Communications Technology (ICACT), pp. 841–846, IEEE, January 2013.

[9] J. Wang, and I. C. Paschalidis, "Botnet detection based on anomaly and community detection," IEEE Transactions on Control of Network Systems, 4(2), pp. 392–404, June 2017.

[10] A. Duh, M. Slak Rupnik, and D. Koroak, "Collective behavior of social bots is encoded in their temporal twitter activity," Big Data, 6(2), pp. 113–123, June 2018.

[11] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, "The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race," 26th International Conference on World Wide Web Companion, pp. 963–972, International World Wide Web Conferences Steering Committee, April 2017.

[12] M. Jiang, P. Cui, A. Beutel, C. Faloutsos, and S. Yang, "Inferring lockstep behavior from connectivity pattern in large graphs," Knowledge and Information Systems, 48(2), pp. 399–428, August 2016.

[13] R. Yu, X. He, and Y. Liu, "Glad: group anomaly detection in social media analysis," ACM Transactions on Knowledge Discovery from Data (TKDD), 10(2), Article no. 18, October 2015.

[14] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, "DNA-inspired online behavioral modeling and its application to spambot detection," IEEE Intelligent Systems, 31(5), pp. 58–64, September 2016.

[15] F. Morstatter, L. Wu, T. H. Nazer, K. M. Carley, and H. Liu, "A new approach to bot detection: striking the balance between precision and recall," 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 533–540, IEEE, August 2016.

[16] P. De Meo, E. Ferrara, G. Fiumara, and A. Provetti, "Generalized louvain method for community detection in large networks," In 2011 11th International Conference on Intelligent Systems Design and Applications, pp. 88–93, IEEE, November 2011.

[17] S. Iyer, T. Killingback, B. Sundaram, and Z. Wang, "Attack robustness and centrality of complex networks," PloS one, 8(4), e59613, April 2013.

[18] T. Khaund, K. K. Bandeli, M. N. Hussain, A. Obadimu, S. Al-Khateeb, and N. Agarwal, "Analyzing Social and Communication Network Structures of Social Bots and Humans," 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 794–797, IEEE, August 2018.

[19] C. Morselli, and J. Roy, "Brokerage qualifications in ringing operations," Criminology, 46(1), pp. 71–98, February 2008.