Final version of the accepted paper. Cite as: *Muhammad Abulaish and Mohd Fazil, Socialbots: Impacts, Threat-Dimensions, and Defense Challenges, IEEE Technology and Society Magazine, 39(3), IEEE, pp. 52-61, 2020.*

Socialbots: Impacts, Threat-Dimensions, and Defense Challenges

Muhammad Abulaish, *SMIEEE* and Mohd Fazil

***Abstract*— Online Social Networks (OSNs) are the modern communication media that are under threat by socialbots. The rise of socialbots and their role has posed diverse challenges in the contexts of political astroturfing, fake news, and spear phishing. This article presents a concise and multifaceted study of socialbots. It commences with a comparative analysis of social botnets with conventional web botnets and further presents an experimental analysis of socialbots' impact on network infiltration. We also categorize the threat landscape of socialbots into four dimensions and present a detailed discussion of each threat-dimension and its impact on different OSN stakeholders. The paper also presents different categories of defense challenges against the socialbots to understand the complications of the problem, which will help in devising future mitigation strategies. Finally, we present a brief overview of the current trends in the direction of socialbot research and their role in the context of COVID-19 pandemic.**

***Index Terms*— Socialbot infiltration, Socialbot detection, Socialbot defense**

I. Introduction

ONLINE Social Networks (OSNs) are the modern communication media at the top of the Internet technology that facilitate their users to get connected with friends and celebrities, and to disseminate breaking news and updates on the real-life incidents. Connections and interactions among OSN users generate a massive amount of data containing a rich set of knowledge that could be useful to various real-world data modeling and predictive analytics problems like *open-source intelligence*, *business intelligence*, *event prediction*, and *product recommendations*. On the Web, adversaries explore and target products and services for misuse and vulnerabilities, and OSNs are no exception. The large user-base, easy to use functionality, and open-nature further attract anti-social elements to these OSNs. In OSNs, cybercrimes and illicit activities are generally committed using various forms of fake profiles, such as *clone profiles*[1], *sybil*[2], and *bots*. In OSNs, fake profiles creation is easy, but their manual handling is neither economically feasible nor scalable. Therefore, adversaries automate fake profiles, which are known as socialbots and given different nomenclature like political bots, spambots, click bots, and cyborgs depending on their end misuse. Socialbots are sophisticated and advanced threat entities from adversaries. They are program-controlled OSN profiles, which imitate human behavior to camouflage the identity and project themselves as human beings. Socialbots are programmed to perform the required OSN functionalities. In the initial stage of injection on an OSN, socialbots behave like normal users, and this is called the reputation building process. Once the reputation is built, socialbots abuse it to perform illicit activities of their masters. Socialbots do not have any direct repercussions like service disruption, but indirect consequences, such as online protest orchestration [1], electoral campaign distortion [2], terrorism funding and recruitment [3], and content pollution [4] that are damaging. Therefore, characterization and detection of socialbots and other threat entities are essential for the healthy growth of OSN platforms. To design an effective socialbots prevention and detection mechanism, analysis of their ecosystem, operating characteristics, temporal evolution, and infiltration impacts are of significant importance. Recently, researchers have started analyzing the socialbots problem from different perspectives, publishing a number of articles. Figure 1 presents the trend of year-wise research publications during 2011-2019. It includes all articles containing the word socialbot or its variants[3] either in title or in abstract. The growth rate of papers over the year shows severity of the problem. Existing literature reports massive misuse of socialbots in political astroturfing, rumor diffusion, and spamming [2, 4]. A comprehensive analysis and study of socialbots is important from both service providers and users' perspectives. In this direction, Ferrara et al. [5] described both benign and malicious socialbots, and discussed anecdotes of malicious socialbots in political astroturfing, stock market manipulation, personal information theft, and misinformation diffusion. They also proposed a grouping of the existing socialbot detection approaches into four categories.

In line to Ferrara et al. [5], we present a different taxonomy of the existing socialbot detection approaches and group them into five categories. In addition to categorization, we also present a characterization of socialbots from different perspectives. Socialbots operate like Web bots, an old phenomenon since the Turing days of computers. Starting with

M. Abulaish (corresponding author) is currently working as an Associate Professor at the Department of Computer Science, South Asian University, New Delhi-110021, India (e-mail: abulaish@ieee.org).

M. Fazil is currently with the Department of Computer Science, Jamia Millia Islamia (A Central University), New Delhi-110025, India (e-mail: mohdfazil.jmi@gmail.com).

[1] *A fake profile created using the information of a real user to deceive other users of the network.*

[2] *A group of fake accounts controlled by a bot or human for malicious purpose.*

[3] *social bot, socialbots, social bots, socialbotnet, socialbotnets, social botnet, and social botnets.*

a technological and operational differences between the two variants of bots, we present some analytical observations from a socialbots injection experiment performed on Twitter to perceive the potential of socialbots infiltration. Further, we categorize the areas which are under threat from socialbots, along with a brief description of each threat category. We also discuss the enabling factors and challenges in defending the malicious socialbots and group them into three categories – platform-, user-, and detection-related challenges. Finally, we discuss the trends of current socialbot research along with a brief description of very recent studies analyzing the role of socialbots on OSN sphere in the context of COVID-19 pandemic.

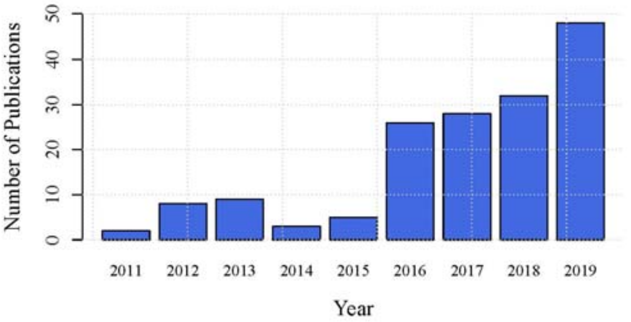


Fig. 1. Year-wise trend of research publications containing the term socialbot and its variants (source: Google Scholar).

II. Web Bot vs Socialbot

Bots are not new, rather they exist since the development of the Internet technology. Web-based bots are compromised entities, which generally operate in groups. On the other hand, due to huge user-base and open-nature OSNs provide a fertile venue to adversaries, who create new and sophisticated threats like socialbots. Socialbots are the OSN-specific Web bot that are sophisticated and dangerous due to their deceiving nature. The network of socialbots is generally controlled by a botmaster, which is a human-controlled automation software to command and control the activities of the socialbots in OSNs. The command to be executed and any update in the execution behavior of socialbots is controlled through the botmaster. In other words, botmaster acts as a command and control hub to check the overall behavior of the socialbot network or Web-based botnet. A network of bots is called botnet. Web-based botnet exploits a combination of social engineering and software and hardware vulnerabilities of the computer systems to infect and make them zombies. The conventional or Web-based botnet, instructed and controlled by a botmaster, keeps on operating until observed and resolved by the users of the infected hosts. Figure 2 presents the architectural difference between the operational behavior of Web-based botnet and social botnet. Though the creation and operation of socialbots are easy, their characterization and detection are difficult. In the absence of online-offline identity binding, easy profile creation process, and lack of effective control mechanisms, OSNs are at risk of higher-order sophisticated attacks from socialbots. A study estimates that around 9 to 15% of Twitter profiles are bots [6]. Table I presents a comparison between the Web- and OSN-based botnets based on various parameters.

TABLE I
A COMPARISON BETWEEN WEB AND SOCIAL BOTNETS

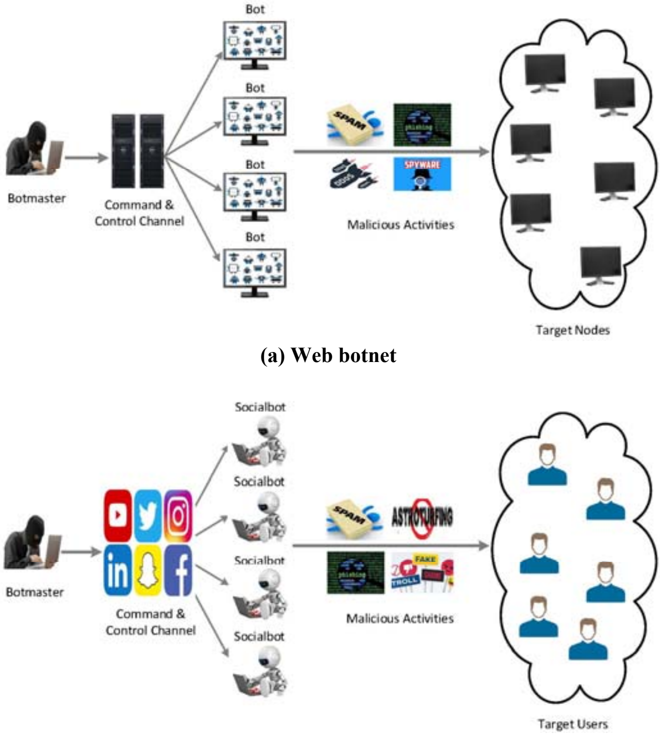
Parameter	Web botnet	Social botnet
Bot	Malware-infected computer	An automated OSN profile emulating human behavior
Botnet	Group of web bots controlled by a botmaster	Group of socialbots controlled by a botmaster
Botmaster	A human-controlled software to control a group of malware-infected computers	A human-controlled software to control a group of socialbot profiles
Command & control channel	Communication protocols, such as TELNET, IRC, and P2P to control and instruct web bots	Uses OSNs as command and control channel
Threat dimensions	Network & host hijacking, distributed denial-of-service attack, spamming, and information stealing.	Astroturfing and propaganda diffusion, fake news and rumor spreading, spamming, and trolling.
Threat space	Public and private computer networks	Online social networks
Attack strategy	Software vulnerabilities and social engineering	Social engineering

III. Impact Analysis Through a Socialbots Injection Experiment

To do an empirical analysis of socialbots impact on infiltration, we performed a socialbots injection experiment on Twitter. A detailed description of the experiment and underlying observations is presented in one of our previous works [7]. In the experiment, 98 socialbots associated with top-six Twitter using countries were manually created within two months. The number of socialbots assigned to each country was proportional to its user-base. We operated the socialbots network for four weeks and logged all the activities for analysis. Socialbots were programmed to perform connection formation, tweet posting, retweeting, and tweet liking. However, socialbots were not programmed to generate content due to fear of detection and reporting from normal users [8]. We performed a comprehensive analysis of the logged data and perceived very interesting observations, which are reported in the following paragraphs.

- During the four weeks of operation, socialbots were successful in infiltrating a significant number of users with an infiltration rate of 29 users per socialbot.
- In a significant observation, we found that verified users, including prominent actors and songwriters, are also at

risk of being trapped[4] by the socialbots. Among the trapped verified users, one has approximately 1.72 million followers, which is a very significant number on Twitter.



(a) Web botnet

(b) Social botnet

Fig. 2. An architectural difference between web botnet and social botnet.

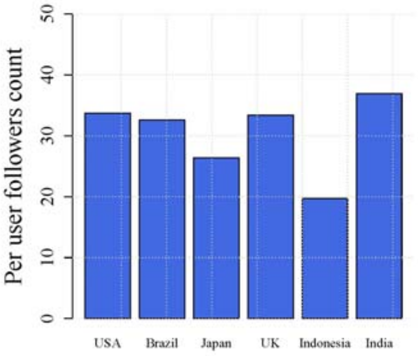


Fig. 3. Country-specific socialbots' average followers.

- We also found that the infiltration performance of socialbots depends on their regional association, as shown in figure 3 because users of certain geographies are more unconscious while accepting friend requests from unknown users.
- Profile gender showed a conditional impact on infiltration because socialbots associated to certain geographies and having exposing profiles were more infiltrative, as shown in figure 4.
- Among various socialbot activities, we found "following" as the most affluent activity in terms of alluring and duping users to follow socialbots.
- To observe the malicious and botnet behavior among the trapped users, we performed an individual- and group-level analysis of the trapped users. On analysis, we found a significant number of suspicious users that were later suspended by Twitter. We also tracked a set of trapped profiles, operating in a coordinated manner and following a single socialbot. All these profiles were posting similar tweets; although their content were not malicious. Moreover, these profiles were created at nearly the same point in time and later suspended by Twitter.

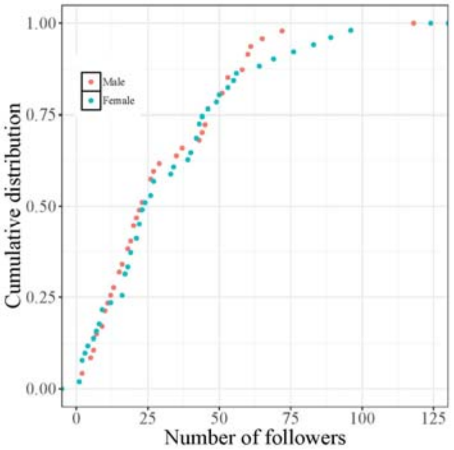


Fig. 4. Gender-wise cumulative distribution of the socialbots' followers count.

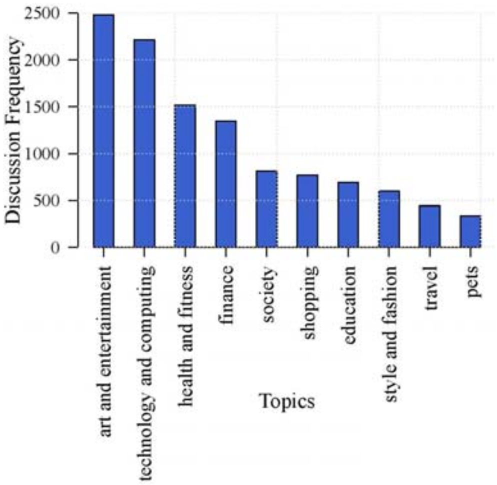


Fig. 5. Top-10 topics discussed by the socialbot followers.

- We also performed topical analysis of trapped users content to have a glimpse of their topical inclination.

[4] *A trapped user is one who followed our injected socialbots*

Figure 5 shows top-10 topics and depicts that the most relevant and discussed topics, such as finance, computer accessories, and shopping are advertising- related topics. They also discussed important topics like education and society, but relatively less frequently.

IV. Threat Dimensions

In the early stage of research, socialbots were developed and deployed in OSNs to observe their infiltration potential and impact analyses. As an open-source initiative, Coburn and Marra [9] conducted a socialbots injection experiment on Twitter under the *Realboy* project. In the project, socialbots imitated the *sense-think-act*[5] model of physical robots to observe the information propagation behavior of social robots (socialbots) and underlying impact on the social graph. This was an early study, and since then researchers are repeatedly reporting misuse cases in different contexts of the OSNs [1, 2, 3]. Mitter et al. [10] presented a generic framework to categorize socialbot attacks based on five parameters – *target*, *account type*, *vulnerability*, *attack method*, and *result*. However, authors have not categorized socialbots based on their intention and functionality. In this study, we have identified different threat dimensions where socialbots abuse OSN platforms for deceptive activities.

A. Socialbots as Political Tools

OSNs are modern discussion platforms and open in nature where people debate government policies, daily life events, social issues, and so on. In this era, users purchasing behavior, political inclination, etc. are influenced by social media events and discussion [11, 12]. The business organizations and political parties are using social media platforms for targeted advertisements, to log and resolve customer grievances, to observe users' opinion regarding policies and products, and so on. The huge user-base, open nature, and easy to use functionalities of OSNs have also attracted adversaries. Socialbots handler exploit the open nature of OSNs at the top of social engineering to carry out smear campaigns against the political rivals [13, 14]. The first misuse of socialbots as a political tool was reported in the 2010 U.S. mid-term election [15]. Since OSNs are generally used to express views regarding current incidents, political scenarios, and daily life events, maligning the discourse will hamper the ultimate objectives and growth of the underlying OSN platforms. Political parties around the world are using socialbots to distort public opinion and create artificially manufactured campaigns [14]. Dickerson et al. [13] also reported the use of socialbots in the 2014 Indian general election and presented a sentiment and linguistic features-based technique for socialbots detection. The existing studies report the use of these computational agents in most of the larger democracies. The problem is not just the use of socialbots but their impact on results. Moreover, government authorities inject socialbots to bias the public opinion for pro-regime discussions [1]. Socialbot as a political tool is used for astroturfing, fame fostering, opinion manipulation, etc., as given in Figure 6.

B. Socialbots as Fake News Distributors

The fake news problem is not new and dates back to the printing press era [16]. In traditional news media, scope and repercussions of fake news are limited, and its after-effect can be controlled and sometimes revoked. Traditional news media does not have the magnitude of user-base and real-time proliferation power like OSNs which has provided an unprecedented opportunity to anti-social elements to misuse these platforms for rumor and fake news propagation. The malicious users exploit socialbots for fake news and misinformation diffusion to pollute the social media discourse. Social media are modern news and information sharing platform, therefore, ensuring the credibility and authenticity of its content is vital to ensure the confidence of all stakeholders. Recently, the world economic forum[6] has expressed concern over the use of automatic agents in fake news and misinformation diffusion. A study reports that during the 2016 U.S. presidential election, socialbots distorted the political discourse and massively diffused fake reports and false information [2]. Studies also report the involvement of socialbots in fake news propagation in the form of misinformation, disinformation, and hoaxes [4]. The socialbots network along with humans can be misused to spread unsubstantiated information and make it influential. Apart from socialbots, humans are also responsible for spreading fake news [17].

C. Socialbots as Spammers

Spams even existed during the ARPANET period. However, as technologies evolved, spammers also evolved to abuse them. On the Web, spammers exploited bot technology to automate and carry out illegitimate activities, such as phishing, hacking, and identity theft. Similarly, spammers are also targeting OSNs using various forms of fake profiles. They also automated the spamming process using automated handling of these profiles. Spamming campaigns on OSNs are now generally operated using bots due to the easy accessibility of the OSN application program interface (API), real-time proliferation, and large user-base [18]. Initially, spambots were conducting spam campaigns on the Web for content pollution, advertising, and so on. However, with the inception of OSN, a new breed of spammers, called social spambot, came into existence with more deceptive nature and consequences [19]. Social spambots, a variant of socialbot, are created to perform stealthy spamming activities like spear phishing, identity theft, information harvesting, and account hacking [4, 20]. These are defining new dimensions of spamming in terms of deception and sophistication level. In a seminal work, Cresci et al. [19] analyzed the behavioral difference among the normal users, conventional spambots, and social spambots. On analysis, they found that human annotators can easily differentiate between conventional spambots and real

[5] *A robotic paradigm which describes the three primitives of a robot. "Sense" represents data gathering (e.g., motion detection) process of robots; "think" represents information processing power; and finally "act" represents the decision making ability.*

[6] *https://www.weforum.org/agenda/2017/02/how-can-we-defeat-fake-news-automate-the-right-to-reply/*

users, but failed to differentiate social spambots and real users. It is because social spambots imitate human users to build a reputation in the network and then exploit it to carry out stealthy spamming activities. As a result, it is difficult to distinguish social spambots from real users. Researchers are devising novel techniques for social spambot detection [21, 22]. However, as approaches become mature and advance, spammers use more sophisticated mechanisms like triad closure property, human-imitating profiles, and modern natural language generation methods to evade detection, and thereby resulting in a "cat and mouse game". At the individual level, social spambots imitate the behavior of normal users and perform illicit activities, but not as frequently as conventional spambots. Hence, node-level socialbot detection is a challenging problem, and we need approaches to detect social spambots that operate in a coordinated manner [21, 23].

D. Socialbots as Infiltrators and Influence Manipulators

In OSN, users' influence is determined based on the size of their social network and the impact of their content in terms of reach and diffusion on the network. Influence clinching in OSN is not trivial because users generally follow or send friend requests to known users, celebrities, and politicians. In other words, users' influence depends on their infiltration ability because it expands their social space. Section 3 illustrates through an injection experiment on Twitter that socialbots can easily infiltrate an OSN to expand their social network. Existing literature has various studies showing the influence manipulation ability of socialbots. Aiello et al. [24] injected a socialbot on aNobii, a popular OSN of book lovers in Spain, and observed that the injected socialbot, without trust and reputation, reached among one of the most influential users of the network. On analysis, authors found that simply browsing user profiles by socialbot did the job as aNobii sends a notification about the browsing activity on user profiles. In response, notified users generally examine the profile of the browsing users, raising the chance of connection creation. On analysis, the Klout score[7] of the injected socialbot was higher than many celebrities. The Klout score of a user represents online social influence based on his/her activities in nine different OSNs – Bing, Facebook, Foursquare, Google+, Instagram, LinkedIn, Twitter, YouTube, and Wikipedia. Elyashar et al. [25] performed a socialbots injection experiment on Facebook and found that even technologically aware users are not conscious enough while accepting friend requests from unknown users. Socialbots exploited the triad closure[8] to inflate the connection formation probability with the target users. In a seminal work, Boshmaf et al. [26] observed that Facebook network can be infiltrated with a success rate of 80%. They also reported that based on privacy settings, users' personal information can easily be harvested after infiltration. Further, they reported various OSN vulnerabilities and enabling factors that are abused by socialbots at various stages of the injection and operation process. In an incident, LinkedIn filed a case against unknown handles of a social botnet[9], who were scrapping users' personal information and violating the terms and use conditions of the service.

Apart from the four major misuse cases discussed above, socialbots are also misused by adversaries for other malicious activities, such as trolling and bullying. Socialbots are used to dupe OSN users to steal sensitive information, such as username, password, credit card detail, and so on. Shafahi et al. [20] performed a socialbots injection experiment to observe the social engineering power of socialbots for phishing. In the experiment, socialbots were successful in duping a significant number of users from different geographies without being detected until the end of the experiment.

V. Defense Challenges against socialbots and their evolution

Socialbots are sophisticated, modern, and advance threat entities in OSN. Researchers from different disciplines are working for the development of various approaches to tackle this cyber-menace. In the initial phase of injection, socialbots build trust[10] in the network and then exploit it to carry out illicit activities. This process makes their detection difficult. Existing approaches for socialbots detection can be grouped into five categories, namely machine learning, graph-based, behavior-based, crowd-sourcing-based, and hybrid approaches. Moreover, existing detection techniques are not adaptive, and deceptive socialbots can easily evade them. The prevention and detection of socialbots are also challenging due to several enabling and operational factors, such as inherent vulnerabilities of OSNs, development of automation technologies, existence of bogus, fake, and unconscious users, and dynamic nature and behavioral resemblance of socialbots with benign users [26].

A. Platform and Technology-Based Challenges

The first level of challenge while enforcing defense against malicious socialbots is due to the inherent vulnerabilities of the OSN platforms and the development of advanced artificial intelligence technologies. Boshmaf et al. [27] broadly recognized three categories of inherent vulnerabilities – web automation, identity binding, and usable security. Socialbots exploit technological developments and abuse exposed functions by OSN platforms to automate all the activities starting from account creation to executing various OSN functions. The exploitation of web automation is an arms race between the service provider and socialbots which does not seem to end soon. A strict measure at this level will hamper the legitimate use of exposed functions such as the use of APIs by legitimate third-party applications. Identity binding has the

[7] *https://en.wikipedia.org/wiki/Klout*

[8] *A social network concept to model and understand network evolution and community growth.*

[9] *https://yro.slashdot.org/story/16/08/16/134239/linkedin-sues-100-individuals-for-scraping-user-data-from-the-site*

[10] *We define trust of a user in an OSN, based on other users' response to the connection requests of the user. For Twitter, it can be defined as the ratio of the 'followers count' to the 'followings count'.*

potential to prevent the admission of socialbots exposing their identity at the time of account creation. It requires government-authorized policy and regulation to ensure online identity like social security number of every individual, which must have acceptability and trust among common users. However, such mechanisms generally have certain privacy concerns. Moreover, OSN platforms should provide control to users to decide their level of security and privacy. The users should also be informed regarding the repercussions of every provided flexibility. In addition to the vulnerabilities listed by Boshmaf et al. [27], the fourth and most important challenge is the fierce competition among OSNs for user-base because the monetary value of an OSN is determined based on the size of its active user-base. As a result, due to fear of loss of user-base, most of the platforms do not have stringent account registration mechanism or service usage policy. It makes the admission and abuse of OSN services easy for adversaries through bots and fake profiles.

B. User-Based Challenges

In OSNs, the trust and influence of users are based on their followers/friends count in the network [28]. Once a socialbot gets injected into a social network, it starts building reputation in the network through infiltration and connection formation with other users. In the process, they are facilitated by users who accept their connection requests either knowingly or unknowingly. The challenge at this level is to prevent socialbots from reputation building in the network. In this direction, three challenges are – (i) accounts trading, (ii) users' awareness, and (iii) interest exploitation avoidance. In OSN, socialbots gain friends either purchasing them from third-party vendors or establishing connections with random users, or creating fake profiles and befriending with them. Among the three means of connection creation, purchasing friends is the most effective and feasible solution due to the availability of economical black market vendors. In contrast, connecting with random users is not trivial, whereas creating fake profiles and connecting with them is not feasible in terms of reputation and scale. Therefore, tracking and controlling the black market of followers trading is vital to fight against socialbots. Although, large-scale trading of followers/friends is illegal in OSN and researchers are devising methods to fight against it, this is still happening. Socialbots also exploit human inclination and hunger to increase followers/friends, and choosing users having higher friendship acceptance rate as the target. In OSN, some users accept most of the friend requests and do not show any consciousness while accepting friend requests from unknown users. In addition, there are users who are more interested in number of followers/friends rather than their authenticity. Such unconscious and follower-hungry users are potential facilitator and followers of the malicious socialbots. Therefore, creating awareness among OSN users and the development of friendship assistance mechanisms, showing the trust and reputation levels of the senders, to assist users in friend selection and request acceptance process is important. Though socialbots can create connections, they can hardly deceive genuine users who are the real asset of the network. In case of benign and conscious users, socialbots use social engineering tactics like exploitation of common profile attributes (e.g. interest, place, education) and triad closure property to deceive them. In the infiltration process, socialbots target attribute or structurally coherent users raising the probability of acceptance of the friendship requests.

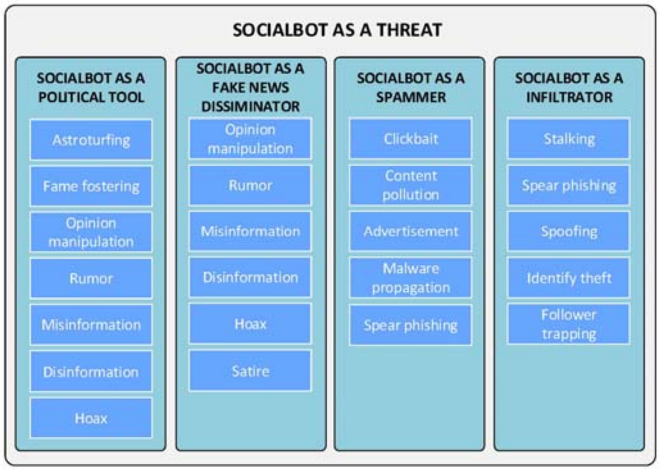


Fig. 6. Different types of socialbots misuse and related threats.

C. Detection Challenges

To tackle the socialbot problem, researchers from academia and industry have proposed various approaches and still working on it. We group the existing approaches into five categories, namely machine learning, graph-based, behavior-based, crowd-sourcing-based, and hybrid approaches. In machine learning approaches, classification models using a set of predefined features are trained to predict the label of a new user [22]. As new features are devised and classification models are updated, socialbots change their behavior to bypass the detection systems, making the detection a challenging problem. Machine learning approaches use hand- crafted features which is a manual and time-consuming task. As a result, it incorporates human biases and deficiencies. Moreover, the efficacy of a machine learning-based classification system depends on the set of defined features. Therefore, if they are of low quality, the performance of the respective trained classification system will be automatically low. Graph-based approaches partition a social graph into multiple sub-graphs representing sybil and non-sybil regions [29]. Socialbots deceive this line of techniques by creating attack edges with benign users. Further, graph-based approaches do not guarantee the detection of individual socialbot. Researchers have also modeled the operational and behavioral patterns of users to discriminate between malicious and benign users [21, 30]. These approaches model synchronization among user activities to detect suspicious behavior. Presently, these approaches appear most effective in comparison to other categories of approaches. In a crowd-sourcing approach, expert annotators are hired to label an account as a socialbot or benign based on its profile and activity information. However, it is neither economically nor technically feasible for OSNs having millions of users [5]. Researchers have also proposed hybrid approaches for socialbots detection.

VI. Current Direction of Socialbot Research and its Role in COVID-19

The focal point of earlier socialbot research was the impact and infiltration analysis [7, 24, 25], and development of detection systems [21, 22, 30]. Recently, researchers have started exploring the role of socialbots in the context of various threat dimensions, such as political tools [31, 32], fake news and rumor spreaders [4, 33], trolls [12, 34], as depicted in Figure 6. Recently, bots are also misused in diffusing public health infodemic and amplifying vaccination debates [35, 36, 37]. In the global coronavirus (COVID-19) crisis, most of the countries around the world are under lockdown to contain the spread of the virus. As a result, people are staying at home and spending more time on social media platforms. Researchers are conducting studies to observe insights from user-generated OSN content in the context of COVID-19. In such a study, Ferrara [38] analyzed the role of socialbots and found their use in both good and malicious purposes. The author observed the use of socialbots to foster democratic discussions, which may otherwise be censored. On the other hand, the author also found the use of socialbots in promoting divisive political campaigns and conspiracy theories. However, a comprehensive analysis regarding the role of bots in diffusing different types of weaponized information like artificial political campaigns, health infodemic, xenophobic behavior, rumor, and fake news in the contexts of COVID-19 is not studied yet. These are good directions for research to investigate the contextual role of the socialbots.

VII. Conclusion

Based on the discussions and analysis presented in this article, we can conclude that the socialbot problem is still evolving and targeting newer areas. Existing literature also lacks in terms of efficient techniques for prevention and detection of the sophisticated version of socialbots like political bots, influence bots, and spambots. In this paper, we have presented a study of socialbots from four different perspectives – their differentiating characteristics with the web bots, infiltration ability on Twitter, different threat-dimensions, and various categories of defense challenges. We have also presented a comprehensive discussion regarding the role of socialbots in the contexts of various OSN threats. Finally, we have presented a brief overview of the current trends in socialbots research and their role in the context of COVID-19 pandemic. This article could be very enlightening and informative for a comprehensive understanding of socialbots and related research challenges.

References

[1] N. Abokhodair, D. Yoo, and D. W. McDonald, “Dissecting a social botnet: Growth, content and influence in Twitter,” in Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing, (Vancouver, BC, Canada), pp. 839–851, 2015.

[2] A. Bessi and E. Ferrara, “Social bots distort the 2016 U.S. presidential election online discussion,” First Monday, vol. 21, no. 11, 2016.

[3] J. M. Berger and J. Morga, “The ISIS Twitter census defining and describing the population of ISIS supporters on Twitter,” vol. 20, 2015.

[4] C. Shao, G. L. Ciampaglia, O. Varol, K. C. Yang, A. Flammini, and F. Menczer, “The spread of low-credibility content by social bots,” Nature Communications, vol. 9, no. 11, 2018.

[5] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, “The rise of socialbots,” Communications of the ACM, vol. 59, no. 7, pp. 96-104, 2016.

[6] O. Varol, E. Ferrara, C. A. Davis, F. Menczer, and A. Flammini, “Online human-bot interactions: detection, estimation, and characterization,” in Proceedings of the 11th International Conference on Web and Social Media, (Montreal, Canada), pp. 280–289, 2017.

[7] M. Fazil and M. Abulaish, “Why a socialbot is effective in Twitter? a statistical insight,” in Proceedings of the 9th International Conference on Communication Systems and Networks, Social Networking Workshop, (Bengaluru, India), pp. 562–567, 2017.

[8] C. de Freitas, F. Benevenuto, S. Ghosh, and A. Veloso, “Reverse engineering socialbot infiltration strategies in Twitter,” in Proceedings of the International Conference on Advances in Social Networks Analysis and Mining, (Paris, France), pp. 25-32, 2015.

[9] Z. Coburn and G. Marra, “Realboy: Believable Twitter bots, http://ca.olin.edu/2008/realboy/index.html,” 2008.

[10] S. Mitter, C. Wagner, and M. Strohmaier, “A categorization scheme for socialbot attacks in online social networks,” in Proceedings of International Conference on Web Science, (Bellevue, USA), pp. 269- 278, 2013.

[11] A. K. Fotiadis and N. Stylos, “The effects of online social networking on retail consumer dynamics in the attractions industry: The case of ‘e-da’ theme park, Taiwan,” Technological Forecasting and Social Change, vol. 124, no. 11, 2017.

[12] C. A. Bail, B. Guay, E. Maloney, A. Combs, D. S. Hillygus, F. Merhout, D. Freelon, and A. Volfovsky, “Assessing the russian internet research agency’s impact on the political attitudes and behaviors of American Twitter users in late 2017,” Science, 2020.

[13] J. P. Dickerson, V. Kagan, and V. S. Subrahmanian, “Using sentiment to detect bots on Twitter: Are humans more opinionated than bots?,” in Proceedings of the International Conference on Advances in Social Networks Analysis and Mining, (Beijing, China), pp. 620-627, 2014.

[14] S. C. Woolley, “Automating power: Social bot interference in global politics,” First monday, vol. 21, no. 4, 2016.

[15] J. Ratkiewicz, M. Conover, M. Meiss, B. Gonc¸alves, S. Patil,A. Flammini, and F. Menczer, “Truthy: Mapping the spread of astroturf in microblog streams,” in Proceedings of International Conference on World Wide Web, (Hyderabad, India), pp. 249–252, 2011.

[16] J. M. Burkhardt, “History of fake news,” Tech. Rep., Library Technology Reports, 2017.

[17] S. Vosoughi, D. Roy, and S. Aral, “The spread of true and false news online,” Science, vol. 359, no. 6380, 2018.

[18] H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B. Y. Zhao, “Detecting and characterizing social spam campaigns,” in Proceedings of the 10th SIGCOMM Conference on Internet Measurement, (Melbourne, Australia), pp. 35-47, 2001.

[19] S. Cresci, R. D. Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, “The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race,” in Proceedings of the 26th International Conference on World Wide Web, (Perth, Australia), pp. 963–972, 2017.

[20] M. Shafahi, L. Kempers, and H. Afsarmanesh, “Phishing through social bots on Twitter,” in Proceedings of the International Conference on Big Data, (Washington DC, USA), pp. 3703-3712, 2016.

[21] S. Cresci, R. D. Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, “DNA-inspired online behavioral modeling and its application to spambot detection,” IEEE Intelligent System, vol. 31, no. 5, pp. 58- 64, 2016.

[22] C. A. Davis, O. Varol, E. Ferrara, A. Flammini, and F. Menczer, “Botornot: A system to evaluate social bots,” in Proceedings of International Conference on World Wide Web, (Montreal, Canada), pp. 273–274, 2016.

[23] M. Fazil and M. Abulaish, “A hybrid approach for detecting automated spammers in Twitter,” IEEE Transactions on Information Forensics and Security, vol. 13, no. 11, pp. 2707-2719, 2018.

[24] L. M. Aiello, M. Deplano, R. Schifanella, and G. Ruffo, “People are strange when you’re a stranger: impact and influence of bots on social networks,” in Proceedings of the 6th International Conference on Weblogs and Social Media, (Dublin, Ireland), pp. 10-17, 2012.

[25] A. Elyashar, M. Fire, D. Kagan, and Y. Elovici, “Homing socialbots: intrusion on a specific organization’s employee using socialbots,” in Proceedings of the International Conference on Advances in Social Networks Analysis and Mining, (Niagara Falls, Canada), pp. 1358-1365, 2013.

[26] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu, “Design and analysis of social botnet,” Computer Networks, vol. 57, no. 2, pp. 556-578, 2013.

[27] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu, “Key challenges in defending against malicious socialbots,” in Proceedings of the 5th USENIX Workshop on Large-Scale Exploits and Emergent Threats, (San Jones, USA), pp. 1-4, 2012.

[28] M. Cha, H. Haddadi, F. Benevenuto, and K. P. Gummadi, “Measuring user influence in twitter: The million follower fallacy,” in Proceedings of International Conference on Weblogs and Social Media, (Washington DC, USA), pp. 10-17, 2010.

[29] N. Z. Gong, M. Frank, and P. Mittal, “Sybilbelief: A semi-supervised learning approach for structure-based sybil detection,” IEEE Transactions on Information Forensics and Security, vol. 9, no. 6, pp. 976-987, 2014.

[30] N. Chavoshi, H. Hamooni, and A. Mueen, “Debot: Twitter bot detection via warped correlation,” in Proceedings of 16th International Conference on Data Mining, (Barcelona, Spain), pp. 817-822, 2016.

[31] L. Luceri, A. Deb, S. Giordano, and E. Ferrara, “Evolution of bot and human behavior during elections,” First Monday, vol. 24, no. 9, 2019.

[32] A. Badawy, E. Ferrara, and K. Lerman, “Analyzing the digital traces of political manipulation: The 2016 Russian interference Twitter campaign,” in Proceedings of International Conference on Advances in Social Networks Analysis and Mining, (Barcelona, Spain), pp. 258- 265, 2018.

[33] N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, and D. Lazer, “Fake news on Twitter during the 2016 U.S. presidential election,” Science, vol. 363, no. 6425, 2019.

[34] L. Luceri, S. Giordano, and E. Ferrara, “Detecting troll behavior via inverse reinforcement learning: A case study of Russian trolls in the 2016 U.S. election,” in Proceedings of 34th International Conference on Artificial Intelligence, (New York, USA), pp. 1-11, 2020.

[35] D. A. Broniatowski, A. M. Jamison, S. Qi, L. AlKulaib, T. Chen, A. Benton, S. C. Quinn, and M. Dredze, “Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate,” American Journal of Public Health, vol. 108, no. 10, 2018.

[36] J. Sutton, “Health communication trolls and bots versus public health agencies’ trusted voices,” American Journal of Public Health, vol. 108, no. 10, 2018.

[37] J.-P. Allem, E. Ferrara, S. P. Uppu, T. B. Cruz, and J. B. Unger, “E-cigarette surveillance with social media data: Social bots, emerging topics, and trends,” JMIR Public Health and Surveillance, vol. 3, no. 4, 2017.

[38] E. Ferrara, “#covid-19 on Twitter: Bots, conspiracies, and social media activism,” Computing Research Repository, vol. 2020arXiv200409531F, 2020.