Down the bot hole: Actionable insights from a one-year analysis of bot activity on Twitter
by Luca Luceri, Felipe Cardoso, and Silvia Giordano

Abstract
Social media represent persuasive tools that have been progressively weaponized to affect beliefs, spread manipulative narratives, and sow conflicts along divergent factions. Software-controlled accounts (i.e., bots) are one of the main actors associated with manipulation campaigns, especially in a political context. Uncovering the strategies behind bots’ activities is of paramount importance to detect and curb such campaigns. In this paper, we present a long term (one year) analysis of bots activity on Twitter in the run-up to the 2018 U.S. midterm elections. We identify different classes of accounts based on their nature (human vs. bot) and engagement within the online discussion and we observe that hyperactive bots played a pivotal role in the dissemination of conspiratorial narratives, while dominating the political debate in the year before the election. Our analysis, in advance of the U.S. 2020 presidential election, reveals both alarming findings of human susceptibility to bots and actionable insights that can contribute to curbing coordinated campaigns.

Contents
Introduction
Research questions
Data and methodology
Results and analysis
Discussion and conclusion

Introduction
False narratives, fake accounts, low-credibility news sources, state-sponsored operators, conspiracy theories, hate and inflammatory messages: The online ecosystem landscape appears loaded with threats and malicious entities disposed to undermine the integrity of social media discussions. Among those, bots (i.e., automated and software controlled accounts; Ferrara, Varol, Davis, et al., 2016; Cresci, 2020) and trolls (i.e., human operators often state-sponsored; Zannettou, et al., 2019a; Luceri, et al., 2020; Badawy, et al., 2019; Jamieson, 2018; Woolley and Howard, 2018; Zelenkauskaite and Niezgoda, 2017) have been recognized as the main responsible actors for manipulation and misinformation operations in diverse contexts (Ferrara, 2015), ranging from finance (Nizzoli, et al., 2020; Cresci, et al., 2018; Tardelli, et al., 2020) to public health (Ferrara, 2020; Yang, et al., 2020b), in which the rise of infodemics (i.e., the widespread diffusion of unverified information and conspiracy theories) during the Covid-19 outbreak represents the latest milestone of the misinformation age (Zarocostas, 2020; Ferrara, Cresci, et al., 2020). Moreover, the abusive behavior of these malicious actors received enormous resonance in the political domain (Luceri, et al., 2019a; Bessi and Ferrara, 2016; Metaxas and Mustafaraj, 2012; Kollanyi, et al., 2016; Alcott and Gentzkow, 2017; Badawy, et al., 2018; Gerber and Zavisca, 2016; Zannettou, et al., 2019b; Luceri, et al., 2019b; Badawy, et al., 2019), where the abuse of social platforms threatened the effective fulfillment of the democratic process, other than creating worldwide concerns for the integrity of voting events (Stella, et al., 2019; Broniatowski, et al., 2018; Stewart, et al., 2018; Ratkiewicz, et al., 2011; Ferrara, 2017; Howard, et al., 2017; Shu, et al., 2017; Stella, et al., 2018; Bovet and Makse, 2019; Grinberg, et al., 2019; Scheufele and Krause, 2019; Ruck, et al., 2019).
In such a context, recent findings showed that most of the political messages shared on Twitter are published by a small set of hyperactive accounts (Yang, et al., 2020a; Hughes and Asheer, 2019). Yang, et al., (2020a) observed that hyperactive users are more likely to publish low-credibility narratives (i.e., stories created by unreliable or hyperpartisan sources; Shao, et al., 2018; Pennycook and Rand 2019; Bovet and Makse, 2019; Grinberg, et al., 2019) with respect to ordinary users and they tend to exhibit suspicious behaviors, often linked to automated and fake accounts. In this regard, Ferrara, Chang, with respect to ordinary users and they tend to exhibit suspicious behaviors, often linked to automated and fake accounts. In this regard, Ferrara, Chang, et al. (2020) recognized how a set of conspiracy bots pushed a relevant portion of low-credibility information (e.g., QAnon conspiracy theories) within the 2020 U.S. presidential election debate on Twitter. From this perspective, it follows that, by flooding social platforms with content of questionable accuracy, hyperactive accounts can both manipulate organic users, by affecting and influencing their opinions, but also the platform mechanisms and its engagement metrics, e.g., trending topics and feed ranking algorithms (Yang, et al., 2020a). Such vulnerabilities, along with the relentless presence of inauthentic entities, highlight the need for intervention to prevent a distortion of online discussions and a manipulation of public opinion.

Therefore, developing computational tools to detect deceptive and orchestrated campaigns has been the main goal of the agenda initiated by the research community to purge online platforms. However, this represents a challenging task (Ferrara, Varol, Menczer, et al., 2016; Chen and Subramanian, 2018; Varol, Ferrara, Menczer, et al., 2017). Nevertheless, in the last few years, researchers have offered several approaches to identify bots (Varol, Ferrara, Davis, et al., 2017; Yang, et al., 2019; Cresci, et al., 2019; Mazza, et al., 2019; Sayyadiharakandeh, et al., 2020; Cresci, et al., 2020; Chavoshi, et al., 2016; Subrahmanian, et al., 2016; Chen and Subramanian, 2018), while solutions for unveiling the activity of trolls have been recently proposed (Luceri, et al., 2020; Addawood, et al., 2019; Im, et al., 2019). Other studies have focused on the detection of collective and inauthentic behaviors of malicious accounts to uncover coordinated campaigns (Sharma, et al., 2020; Nizzoli, et al., 2020; Pacheco, et al., 2020) and suspicious content diffusion (Hui, et al., 2020; Giglietto, et al., 2020; Zannettou, et al., 2020). However, social media abuse persists and the online ecosystem still presents a mix of organic and malicious users (Luceri, et al., 2019b; Im, et al., 2019), where the former class still demonstrates a moderate capability to identify the latter (Yan, et al., 2020). This also calls for a clear understanding of users’ susceptibility to content shared by malicious accounts and their interplay with them.

---

**Research questions**

In light of these considerations, more research is needed to uncover the strategies behind the activities of malicious actors on social media in order to curb the abuse of online platforms, as well as to investigate the downstream effects of users’ exposure to and interaction with malicious accounts for appraising the impact of their deceptive activity.

Along these research directions, in this paper, we inspect the activity of bot and human accounts on Twitter during the year approaching the 2018 U.S. midterm elections, which were held on 6 November 2018. To the best of our knowledge, this work represents the first attempt to study users, and in particular bots, temporal activity over an extended period of time (one year). It is worth noting that previous studies already analyzed the online discussion related to the 2018 U.S. midterm elections (Luceri, et al., 2019a; Deb, et al., 2019; Jones-Jang, et al., 2020; Trielli and Diakopoulos; 2020). However, our work advances previous efforts by offering an observational analysis of bots’ (and humans’) activity during the whole year preceding the election, which represents an unparalleled opportunity to explore the strategies developed to both inject automated accounts into the debate over time and program their sharing activities while avoiding detection. We focus on the sharing activities that Twitter users can employ to create content (i.e., original tweets), re-share others’ tweets (i.e., retweets), and respond to others’ tweets (i.e., reply). We investigate how and to what extent bots used such digital weapons (Cardoso, et al., 2020) over time to identify cues that can empower the identification of bot accounts and, accordingly, contribute to the detection of orchestrated campaigns.

More specifically, this paper aims at answering the following research questions (RQs):

- **RQ1**: How did bots perform their sharing activities over time?
  We explore the temporal dynamics of bots operations by examining the volume of published content during the year and, in particular, as the election approached.

- **RQ2**: When did bots enter into the midterm debates over the year? The rationale is to understand how bots were strategically
injected into the online conversation to perform their deceptive activity while avoiding detection.

- **RQ3**: Did hyperactive accounts play a pivotal role in broadcasting political messages? We aim to investigate the nature of hyperactive accounts, monitor their appearance in the midterm discussion, and shed light on their activity.

- **RQ4**: How did bot accounts engage and interact with human users? We explore the interactions, in terms of retweets and replies, between humans and bots, and we measure the centrality within online discussions of both classes of accounts.

In the remainder of this paper, we detail the methodology used to respond to these RQs, which includes data collection and bot detection, account classification to distinguish hyperactive and ordinary users, and the k-core decomposition to analyze user interaction and centrality within the social network. Finally, we present our analysis and corresponding findings.

---

**Data and methodology**

In this section, we present Twitter data collected to perform our analysis and methodologies followed to gather the set of data and detect bot accounts. Then, we detail the proposed approach to distinguish classes of accounts based on their engagement in the midterm debate. Finally, we describe the technique used to measure the centrality of the accounts within the Twitter conversation.

**Data collection and bot detection**

To perform our analysis, we leverage the dataset gathered by researchers of George Washington University (Wrubel, *et al.*, 2019). The authors collected Twitter messages (i.e., tweets) for one year (from December 2017 to December 2018) by using midterm elections-related keywords and hashtags. Based on Twitter’s developer policy, the authors (Wrubel, *et al.*, 2019) were allowed to publicly share only the tweet IDs of the collected tweets. Given the released tweet IDs, a process referred to as hydration was necessary to recover complete tweet information. The hydration process uses the Twitter API to obtain the entire tweet object from the corresponding tweet ID. Only 74 percent of the released 171,248,476 tweet IDs were still available online at the time we performed the hydration process (December 2019) and, as a consequence, we were able to gather only this subset of tweets.

We then considered narrowing the analysis to the tweets published by a set of accounts studied in our previous works (Luceri, *et al.*, 2019b; Deb, *et al.*, 2019) to enable comparison and validate previous findings. The rationale behind the idea of using a new dataset was to extend the analysis to a broader observation window (i.e., one year and 171M tweets) with respect to previous studies (Luceri, *et al.*, 2019b; Deb, *et al.*, 2019), which both examined the online discussion (2.6M tweets) in the month before the election. Notice that both datasets were collected by using the same mechanism (via the Twitter API) and similar trending keywords. Moreover, from the set of 997k users analyzed in (Luceri, *et al.*, 2019b; Deb, *et al.*, 2019), we found 943k accounts in the collected dataset that were responsible for the creation of the majority of the collected tweets. Specifically, more than 98M tweets (over the 126M collected) were shared by the set of 943k accounts. It should also be noted that a consistent subset of such users (62 percent), responsible for almost 80 percent of the collected tweets, was active at the time of this writing (mid-October 2020) in debates related to the U.S. 2020 presidential election (Chen, *et al.*, 2020). The interested reader can find further details in the Appendix.

Given that the set of 943k accounts was analyzed in previous efforts, we relied on account classification (bot vs. human) that we performed and validated in Luceri, *et al.* (2019b) and Deb, *et al.* (2019). The account classification was performed by means of Botometer [1], a machine learning-based tool developed by Indiana University (Davis, *et al.*, 2016; Varol, Ferrara, Davis, *et al.*, 2017). Botometer outputs a bot score: the lower the score, the higher the probability that the user is human. Besides using the bot score to classify the accounts, we further analyzed the status of such accounts, i.e., active, suspended, or deleted by leveraging the Twitter API. In Table 1, we show the statistics of
Table 1: Accounts and tweet statistics.

|                  | All            | Bots           | Humans         | Suspended | Deleted |
|------------------|----------------|----------------|----------------|-----------|---------|
| Accounts         | 943,126        | 184,545        | 758,581        | 38,164    | 30,688  |
| Account percentage |               | 19.57%         | 80.43%         | 4.05%     | 3.25%   |
| Tweets           | 98,123,612     | 42,386,269     | 55,737,343     | 9,645,522 | 2,828,963 |
| Tweet percentage |               | 43.2%          | 56.8%          | 9.83%     | 2.88%   |

Specifically, Table 1 details the number of accounts and shared tweets for each class of users. The percentage of bots (around 20 percent) and human accounts (around 80 percent) is in line with previous works (Luceri, et al., 2019b; Deb, et al., 2019). The majority of users are still active (around 93 percent) at the time of this writing (mid-October 2020), while a small fraction deleted their account (around three percent) or have been suspended by Twitter (around four percent). It is out of the scope of this paper to characterize the suspended accounts. However, an evaluation about the correlation between account suspension and the likelihood of automation is provided to the interested reader in the Appendix.

Engagement-based account classification

To better understand how both human and bot accounts acted over time, we analyze their behavior according to their engagement within the midterm debate. In particular, we consider the frequency and duration of their activity during the observation period by identifying, for each account, the following parameters:

- **first tweet**: the first time the account shared a tweet within our dataset;
- **last tweet**: the last time the account shared a tweet within our dataset;
- **tweets count**: the number of tweets shared by the account within our dataset;
- **active days**: the number of days between the last and first tweet;
- **daily tweet frequency**: the average number of tweets shared by the account during the active days time window, computed as:

  \[
  \text{daily tweets frequency} = \frac{\text{tweets count}}{\text{active days}}
  \]

Based on the **active days** and **daily tweets frequency** parameters, we propose to classify the accounts into two categories: hyperactive and ordinary accounts. User categorization has already been explored in previous studies (Brandtzæg, 2010; Graham and Wright, 2014; Zelenkauskaite and Balduccini, 2017). For example, Zelenkauskaite and Balduccini (2017) classified users according to the frequency and topic variety of their comments in a news portal. Differently from such existing efforts, we neither consider the specific typology of the user (e.g., agenda-setters or superposters as in Graham and Wright, 2014) nor the discussed topic, while we classified users based only on the frequency and lifespan of their online activity. The rationale was to discern users whose activity was continual and repeated (hyperactive) to those less involved (ordinary). To distinguish these two classes of accounts (based on the active days and daily tweets frequency parameters), we identified a threshold for each considered parameter that allows to opportunely separate hyperactive users from ordinary ones. The procedure to select these thresholds is detailed in the Appendix to provide the reader with a smooth presentation and interpretation of the results.

According to the selected filtering parameters, hyperactive accounts represent less than nine percent of the accounts in our dataset and cover about 70 percent of collected tweets. It is worth noting that bots cover about 38 percent of the hyperactive accounts, which represents a noticeable increase of automated accounts with respect to the percentage related to the full set of accounts (80 percent humans vs. 20 percent bots), other than indicating a robust presence of
bots within an online conversation. This is in line with previous findings (Yang, et al., 2020a) and is further confirmed in Figure 1, which depicts the bot score distribution for hyperactive and ordinary accounts.

![Bot score distributions comparison between hyperactive and ordinary accounts](image)

**Figure 1:** Bot score distributions comparison between hyperactive and ordinary accounts.

From Figure 1, it can be noticed that the two classes of accounts exhibit significantly different distributions, as also demonstrated by a Mann-Whitney rank test ($p$-value < 0.001). On the one hand, most of the probability mass in the distribution related to ordinary accounts is in the range [0, 0.14], suggesting that the majority of ordinary accounts are humans. On the other hand, in the distribution related to hyperactive accounts, we can notice how these users are more likely to have higher bot scores with respect to ordinary accounts, indicating a more relevant presence of automated (software-controlled) users within the set of hyperactive accounts.

**k-core decomposition**

To evaluate the centrality (or embeddedness) of the accounts within the Twitter conversation, we employ the $k$-core decomposition technique. The $k$-core decomposition aims to determine the set of nodes deeply embedded in a graph by recursively pruning nodes with degrees less than $k$. Formally, the $k$-core is a subgraph of the original graph in which every node has a degree equal to or greater than a given value $k$. The $k$-core decomposition assigns a core number to each node, which is the largest value $k$ of a $k$-core containing that node. The higher the core number $k$ is, the more embedded the node is in the network.

---

**Results and analysis**

In this section, we present our analysis and corresponding results on the temporal and behavioral dynamics of bot and
human activity over the whole year preceding the 2018 U.S. midterm elections. We focus on investigating bots sharing activities over time with the objective of understanding how such accounts have been strategically injected in the online discussion related to the midterm as the election approached. Then, we consider the distinction between hyperactive and ordinary accounts to identify peculiar and distinctive behaviors between these two classes of accounts by also considering their nature (bot vs. human). Finally, we observe the volume of interactions between bots and humans, and their embeddedness within the social network.

**RQ1: Bots activity over the midterm year**

In Figure 2a, we show the number of tweets shared by human and bot accounts per each week of the observation period. It could be noticed that the human population shared more tweets than bots over the whole year. However, this could be expected given that, as shown in Table 1, human accounts represent the largest portion (80 percent) of the accounts of the collected dataset and shared more content (about 57 percent of tweets) with respect to the bot population (about 20 percent of the accounts and 43 percent of tweets). As a consequence, bots were more active (in terms of tweets generated per account) than humans during the whole period. On average, a bot shared three times the number of tweets published by a human user (218 vs. 70 tweets/account), which is consistent with bots’ purpose of flooding online platforms with a high volume of tweets (Bessi and Ferrara, 2016).
Interestingly, bots followed a similar temporal pattern of human users over the whole year, suggesting that bots strategically attempted to mimic human sharing activity since the beginning of the online conversation related to midterm elections. As expected, the content shared by both classes of accounts increased as the election approached, and a spike of activity was noticeable during election week.

To better examine how bots performed their activity over time, we disentangle the shared messages in original tweets, retweets, and replies. In Table 2, we show the number of shared messages for each category of tweet and class of accounts over the observation period.

Figure 2: Sharing activities over time of bot and human accounts.
Table 2: Sharing activities of bot and human accounts.

|                 | Bots  | Humans          |
|-----------------|-------|-----------------|
| Accounts        | 184,545 | 758,581        |
| Original tweets | 5,677,142 (14%) | 9,635,364 (18%) |
| Retweets        | 32,746,675 (81%) | 39,427,132 (74%) |
| Replies         | 1,849,625 (5%) | 3,876,595 (8%)  |

Retweeting is the most used operation for both humans (74 percent of the time) and bots (81 percent of the time). As expected, bots heavily relied on the retweet action as it represents a simpler operation for automated accounts with respect to the creation of an original tweet or a reply, which requires the usage of more sophisticated techniques based on natural language processing and understanding (Ferrara, 2019).

Figure 2 also portrays the number of original tweets, replies, and retweets shared weekly by humans and bots over the year approaching the midterm elections. Interestingly, also in the disentangled sharing activities, bots followed a similar temporal pattern of human users over the whole year. Indeed, the number of tweets shared by humans and bots over time positively correlate (p > 0.97, p-value < 0.001) for each kind of activity (i.e., original tweets, retweets, and replies). Although original tweets and replies require to develop advanced AI techniques on software-controlled accounts, the volume of such activities performed by bots approaches the volume of human activities (see also Table 1), especially if we consider the number of messages published per each account. This suggests that more sophisticated bots operated along with less advanced spam bots (Ferrara, 2019). Also, similarly to Figure 2a, the content shared by both classes of accounts increased as the election approached, and a spike of activity was noticeable during election week. Finally, it is worth noting how bots started emulating every kind of human sharing activity since the beginning of the year, further confirming how detecting coordinated campaigns is a challenging task, even when users are monitored over an extended observation window.

RQ2: Bots injection within the midterm debate

To have a more comprehensive understanding of how both human and bot online activities evolve over time, in Figure 3a, we depict the number of human and bot accounts engaged in the debate every week of the observation period. In a given week, we denote as engaged an account that shared at least one tweet during the week under analysis.
It should be noticed that the number of engaged bots per week was almost steady and slowly increases over time, while human users engagement was less regular. This controlled engagement of bots might suggest a coordinated effort to simultaneously avoid the detection of automated accounts, while infiltrating them into the political discussion.

To shed light on this finding, we observe the number of new accounts that joined the conversation related to the midterm elections during the observation window. We define an account as new when it participates for the first time in the midterm debate. In Figure 3b, we display the number of new accounts for each week of the period under analysis. Once again, bot temporal pattern showed a stationary trend, where an almost constant number of new accounts entered into the conversation each week. As we mentioned before, this controlled, and supposedly scheduled, activity appeared a more cautious approach for injecting bots within the online discussion to allegedly avoid their detection. Moreover, we noticed that a significant number of accounts (34 percent of the accounts within our dataset) joined the midterm conversation since the end of 2017 and beginning of 2018 (first two weeks of February), as it can be observed from two peaks in the left side of Figure 3b. Interestingly, this subset of accounts was responsible for 80 percent of the collected tweets, suggesting that such users played a central role within the Twitter discussion. This holds true both for human (28 percent of the accounts) and bot (six percent of the accounts) users. It is worth, and mostly alarming, noting that 60k of such accounts were bots that have persisted online during the whole year generating around 34.2 million tweets (35 percent of the tweets) with an average rate of 570 tweets per account. It should be considered that this subset of 60k bots represented a relevant fraction (about 33 percent) of the bot accounts that participated in the midterm discussion and most of them (44k accounts) were actively involved in the 2020 U.S. presidential election debate at the time of this writing (mid-October 2020). Also, we observed that the 260k human accounts that appeared in the early stage of the year (28 percent of the accounts) shared 44.9 million tweets (45 percent of the tweets) with an average rate of 173 tweets per account, which is about 3.3 times lower than the tweet fabrication rate for bots.

A similar trend of appearance can be observed in the time around election day. Specifically, from the first week of October to the week after the election in November, we observed a peak corresponding to a new set of accounts (11.6 percent of the total users) that joined the conversation. Among such accounts, we recognized 85k human users and 24.6k bots, which were responsible for creating 210k and 158k tweets during the aforementioned period, respectively. The ratio between the sharing rate of bots (6.4 tweets/account) and humans (2.5 tweets/account) was lower than the one observed with the set of accounts that joined the conversation since the beginning of the year (3.3 vs 2.5), which allegedly indicates a more cautious strategy adopted by bots for sharing content in a relatively short time window (if
To characterize the behavior of bot accounts that participated in the midterm debate, in Table 3, we compare the sharing activities of those that joined the conversation since the beginning of the year (from now on referred to as old bots) with those that appeared during the election period (from now on referred to as election bots).

|                     | Old bots  | Election bots |
|---------------------|-----------|---------------|
| Tweets              | 34,197,174| 158,708       |
| Original tweets     | 13%       | 25%           |
| Retweets            | 83%       | 70%           |
| Replies              | 4%        | 5%            |

Both classes of bots used the retweet as their main sharing activity. However, it can be observed a relevant difference in the propensity of re-sharing and producing original tweets between old bots and election bots. Indeed, election bots significantly created more original content ($t(4,708,912)=118.2, p<0.001$), while re-sharing significantly fewer posts with respect to old bots ($t(28,213,616)=121.1, p<0.001$), which might suggest the deployment of different strategies between the two classes of bot accounts based on their lifetime within the midterm debate.

Notice that all the other bots (i.e., the ones injected between the end of February and the end of September) exhibited a hybrid behavior between the two considered classes of accounts, with a distribution of shared activities between old and election bots. This further highlights that an increasing number of bots created with a purpose of sharing original content had been strategically injected in the online conversation as the election approached.

To further investigate the nature of the analyzed bot accounts, we examined their creation date (by inspecting tweet metadata). In Figure 4, we display the date of the creation of bot and human accounts.
We observed how the number of created bots within our dataset increased as the election approached. It stands out how in 2009 (similarly to Yang, et al., 2020a), 2017, and right before the midterm elections, an elevated number of bots had been created. It was particularly concerning to note how these bots persisted online over time, allegedly participating in diverse social and political discussions. From the superimposed illustration of Figure 4, it can be appreciated how the bot accounts created from July 2018 to November 2018 (month of midterm elections) outnumbered the accounts created by humans in the same time window, which highlights that a conspicuous number of bots might have been purposely created for infiltrating into the discourse related to the midterm elections. More specifically, about 10 percent of bot accounts were created from July to November 2018, whereas 80 percent of them were created before the beginning of 2018. The remaining fraction (10 percent) was created in the first six months of 2018.

Finally, we considered the account age, computed as the time difference between the date of the first tweet of the account in our dataset and its creation date. Figure 5 illustrates the distribution of account age, measured in days, for both humans and bots. It can be appreciated how the account age for bots (average account age = 1,963.8 days) was significantly lower ($p<0.001$) with respect to human accounts (average account age = 1,313.3 days). The median account age for bots was around 994 days, whereas for humans was about 2,108 days. This result indicates that bots shared midterm-related content closer to their creation date if compared to human accounts, which further suggests that some bots might have been specifically designed to operate in the midterm debate.
**RQ3: Hyperactive accounts behavior**

Based on the engagement-based accounts classification described in the Data and methodology section, we classified users in hyperactive and ordinary accounts.

For ordinary accounts, it should be noticed that, from this set of users, we filtered out accounts that shared only one tweet during the observational window, as we want to consider these users distinctly and analyze them in contrast to hyperactive ones. Overall, we observed that 74 percent of the accounts were classified as ordinary accounts, nine percent of the accounts were classified as hyperactive accounts, and 17 percent of the accounts shared only one tweet during the whole year (see the Appendix for further details). Ordinary accounts were responsible for about 28 percent of the tweets and the proportion between bots and humans within this set of users was consistent with the full set of accounts. A similar accounts distribution was also noticed for users that shared only one tweet, publishing a negligible volume of messages.

Hyperactive users, while representing a tiny fraction of the accounts (about 83.7k), were responsible for about 72 percent of the collected tweets (about 70.7M messages). With 31.8k accounts, bots represented about 38 percent of the hyperactive accounts and published more tweets (38 percent of the collected tweets) than hyperactive humans (62 percent of the hyperactive accounts and 34 percent of collected tweets) despite the latter set of users outnumbering automated accounts, which further highlights the prevalence of bots in online discussion.

Alarmingly, we noticed that hyperactive accounts were involved in the vast majority (90 percent) of the 1.6 million
tweets related to the QAnon conspiracy theory [2] that we identified within our dataset (see the Appendix for more details). This result supports the finding of Yang, et al. (2020a), related to the low-credibility information spreading carried out by hyperactive accounts. Moreover, hyperactive bots dominated the broadcasting of such conspiratorial narratives by pushing 62 percent of the content related to the QAnon theories. It was, therefore, a cause of concern noting that the vast majority of hyperactive bots (22.5k accounts) were participating in the 2020 U.S. presidential election debate (Chen, et al., 2020) at the time of this writing.

To investigate the behavior of hyperactive accounts, we examined their activity pattern over the whole observation window. Figure 6a portrays their weekly activity. While in Figure 2 we have noticed that humans generated more tweets than bots in every week, in Figure 6a, we recognize a different trend, where hyperactive bots produced more tweets than hyperactive humans over the whole year.

In Figure 6b, we display the number of tweets shared weekly by ordinary accounts. Differently from Figure 6a, the gap between the number of shared tweets by ordinary bots and ordinary humans was remarkable, with the former class of accounts that shared about five times fewer tweets with respect to the latter class. The interested reader is referred to the Appendix for detailed trends and volumes of disentangled sharing activities (i.e., original tweets, retweets, and replies) for both hyperactive and ordinary users.

Next, we evaluate how the appearance and engagement of the accounts within both classes evolved over time. In Figure 7a, we show the number of engaged accounts within the hyperactive class for each week of the year. We recall that an account is considered engaged in a certain week if it produced at least one tweet in that week.
Down the bot hole: Actionable insights from a one-year analysis of bot activity on Twitter

Figure 7: Engaged and new accounts in the midterm debate within the hyperactive accounts class.

Figure 7: Engaged and new accounts in the midterm debate within the hyperactive accounts class.

Here, the most noticeable difference with respect to Figure 3a (where all accounts were considered) is related to the activity of hyperactive humans. In fact, as shown in Figure 7a, human activity has a progressive growth over time and a similar pattern to bot activity. This might suggest that the irregular spikes of activity in Figure 3a was mainly caused by ordinary accounts. For what pertains to bots, their weekly activity appeared similar to Figure 5a, which might indicate that bot activity was scheduled similarly for hyperactive and ordinary bot accounts.

In Figure 7b, we show the number of new hyperactive accounts that appeared in Twitter discussion for each week of the year. As discussed before, an account is considered new when it appears in the midterm-related conversation for the first time over the observation window. Also in Figure 7b, bot and human hyperactive accounts tended to have similar temporal patterns in their appearance within the political discussion. This finding, related to hyperactive accounts, was in contrast to what we noticed in Figure 3b (where all accounts were considered), which further confirms that different temporal patterns between humans and bots (in both Figure 3a and Figure 3b) were mainly due to ordinary accounts activity. Interestingly, and similarly to Figure 3b, in Figure 7b, an almost constant number of new hyperactive bot accounts joined the conversation for each week of the observation period, except for the time window around the election, when the number of new hyperactive bot accounts increased significantly.

We also notice that the majority of hyperactive accounts appeared since the end of 2017 (specifically, from December 2017 to February 2018). Such a subset of users (about five percent of the accounts) created about 60 percent of the collected tweets. This is consistent with the findings reported earlier. In particular, here, we recognize that hyperactive accounts were mainly responsible for tweets created by users that joined the midterm debate one year before the election. Indeed, of the 79M tweets created by such users, 59M were shared by hyperactive accounts, which, thus, represent the most prolific tweet fabricators. Among these accounts, bots played a relevant role by broadcasting almost one-third of the collected tweets. Indeed, 30.6M tweets were created by 19k bot accounts, which means that two percent of the accounts in our dataset were bots that shared about 31 percent of the collected tweets. This finding becomes even more concerning when considering the pivotal role of the majority of these accounts in the diffusion of narratives related to the QAnon conspiracy theory (14k accounts) and their involvement in the 2020 U.S. presidential election debate (16k accounts).

Next, we replicate the same evaluation for ordinary accounts. In Figure 8a, we display the number of engaged ordinary accounts for each week of the observation period.
Similarly to Figure 3a (where all the accounts were considered), human accounts presented an irregular pattern (with peaks and valleys) of engagement if compared to the trend of bots, which presented a more controlled and stationary pattern. The ordinary humans pattern was different from the one observed for hyperactive humans (see Figure 7a), confirming that the non-stationary pattern in Figure 3a was mainly caused by ordinary human accounts.

In Figure 8b, we depict the number of new ordinary accounts that appeared in the Twitter discussion for each week of the year. Similarly to Figure 3b and Figure 7b, bots entered into the debate at a constant and lower rate with respect to human accounts. Also, the majority of ordinary accounts joined the discussion at the beginning of 2018 and during the election month. Interestingly, by disentangling the shared tweets, we noticed different activity patterns among the bots injected from December 2017 to the first two weeks of February 2018 (old bots) and those appearing the month around the election day (election bots). In Table 4, we detail how these two classes of accounts distributed their sharing activities.

Table 4: Sharing activities of ordinary bot accounts.

|               | Old bots       | Election bots |
|---------------|----------------|---------------|
| Tweets        | 3,553,531      | 37,713        |
| Original tweets | 16%            | 37%           |
| Retweets      | 80%            | 58%           |
| Replies        | 4%             | 5%            |

The main difference was related to the creation of original content and the usage of retweets, similarly to Table 3 (where all accounts were considered). Here, the discrepancy between old and election bots is even more pronounced, with election bots that performed re-sharing activities only 58 percent of the time, while sharing original tweets 37 percent of the time. This result further confirms the different approaches used by old and election bots and highlights their diverse operational tasks.

To further examine the different behavior between hyperactive and ordinary bots, in Figure 9, we compare the activity of such classes of accounts. In particular, Figure 9a depicts the number of engaged bots for each week of the year. The
engagement of hyperactive bot accounts presents a more steady pattern, with an increasing number of accounts progressively engaged as the election approached, as compared to the trend of ordinary bots, which presents a less stationary pattern of engagement and a more pronounced growth in the month prior to the election.

This is also reflected in Figure 9b, which displays the number of new bots accounts that joined the midterm discussion over the year. From Figure 9b, we can observe how a significant number of bots were active in the conversation since the beginning of the year and new bots were weekly injected over the months preceding the election. A small and constant number of hyperactive bots joined the conversation weekly, while a larger set of ordinary bots were injected over the year and especially during the election week.

**RQ4: Interactions and embeddedness of bot and human accounts**

We now explore the interactions between human and bot accounts. In particular, we examine the interplay in terms of retweets (re-sharing of other users’ content) and replies (respond to other users’ content) exchanged between these two classes of accounts over the whole period of observation. We also investigate the embeddedness of both bots and humans within the social network with the objective of measuring their centrality in online discussion.

Figure 10 shows the interactions in terms of retweets and replies between the two classes of accounts, where each edge reports the volume of interactions (in terms of percentage) between each group. The source of the edge represents the user that re-shared (i.e., retweeted) the tweet of a target user. Node size is proportional to the percentage of accounts in each group for the considered sharing activity, while edge size is proportional to the percentage of interactions between each group.
Figure 10: Retweet and reply interactions between humans and bots.

On the one hand, we observed that human content was more re-shared (59 percent of the retweets) than bot-generated content (41 percent of the retweets). Interestingly, bots equally distributed their retweets towards human- and bot-generated content, which might indicate a strategy operated by bots to engage humans and possibly gain their endorsement (Luceri, et al., 2019a). It is important to note that more than one over three human retweets is a re-share of content produced by bots.

On the other hand, we recognize that humans performed the majority of replies and only a small percentage of those (around 12 percent) were in response to bot-generated content, which is aligned with the findings of Bessi and Ferrara (2016) related to the 2016 U.S. presidential election. Bots, in turn, focused their replies on interactions with humans, which appears in contrast with the 2016 pattern (Bessi and Ferrara, 2016), where bots interacted with other bot accounts more than with human users. We hypothesize that bots have become increasingly advanced to initiate conversations with humans (Luceri, et al., 2019b), but allegedly not sufficiently sophisticated to convincingly involve humans in their discussion. Indeed, bots received much less interaction (in terms of reply) from humans with respect to retweet interaction.

Next, we explored the embeddedness of both bots and humans within the social network by also considering (i) the extent of their sharing activity (hyperactive vs. ordinary accounts); and, (ii) the time of their appearance within the midterm debate. We recall that we identified an account as old if it shared the first midterm-related tweet by the first two weeks of February 2018. Otherwise, we here denote an account as recent. As detailed before, this choice was motivated by the evidence that a significant number of accounts joined the Twitter conversation in the early stages of
the midterm debate (see Figure 3b) and, in turn, generated a disproportionate volume of tweets as compared to recent ones.

To perform such analysis, in Figure 11, we depict the $k$-core decomposition of the retweet network for the two classes of accounts (bots and humans). A retweet network is a directed weighted graph that represents the retweet interactions between users, where nodes represent users and edges represent retweets among them. We extracted the $k$-cores from the retweet network by varying $k$ in the range between 0 and 500 (no variations can be observed after this value). Figure 11 displays the percentage of hyperactive vs. ordinary (resp. old vs. recent) users as a function of $k$, i.e., for every $k$-core we compute the proportion of accounts of each group within the accounts in the $k$-core.

**Figure 11:** $k$-core decomposition of the retweet network for bot and human accounts.

From Figure 11a, we note that the fraction of hyperactive bots increases with $k$, while the percentages related to ordinary accounts drop with $k$. It appears evident that hyperactive users, and in particular bot accounts, populate the most central position within the retweet network. To the contrary, within the ordinary accounts set, human users appear to hold a more central position within the retweet network with respect to bot accounts.

Figure 11b depicts the $k$-core decomposition of recent and old accounts. It can be noticed that old bots represent the group of accounts most embedded within the retweet network, which indicates that they were more deeply connected with respect to old humans and recent accounts. Interestingly, we observe that, as $k$ grows, the fraction of humans drops, whereas the percentage of recent bots remains almost steady. This suggests that a fraction of recent bots populate the most central area of the social network, which, in turn, shows their capability of infiltrating online discussions and rapidly gaining trust and reputation even in a relatively short time window.

To validate our findings, we repeated the evaluation related to the $k$-core decomposition (in Figure 11) for every week of the observation window (by considering the users involved in the discussion until the week under analysis) with no significant changes in results.

**Discussion and conclusion**

Bots represent one of the most recognized threats to the integrity of social media. The understanding of how bot
accounts operate in social platforms and infiltrate into crucial discussions is of pivotal importance to enable prompt and precautionary detection of coordinated campaigns. Further, it is also fundamental measuring and controlling the extent of bot interactions with organic users to assess the impact of their manipulation attempts.

Along these research directions, in this paper, we examined how bots strategically performed their online activity over time during the year approaching the 2018 U.S. Midterm elections, as well as the effects of their interplay with human users.

We observed how bots reproduced human activity patterns for every sharing activity (i.e., original tweet, retweet, and reply) since the year before the election and we recognized their propensity of flooding online platforms with retweets. This is also in line with previous findings (Luceri, et al., 2019b), further confirming that retweets have been employed as the favorite digital weapon of bots to resonate messages and create an illusion of consensus within an online community (Ferrara, Varol, Davis, et al., 2016).

However, as the election day approached, an increasing number of bots were strategically injected into the online conversation with the purpose of sharing original content.

Interestingly, we discovered that a significant fraction of bots (about one-third of the identified automated accounts) started pushing midterm–related content even one year before the election, generating a conspicuous volume of tweets (34.2 million) over the whole observation period. Alarming, this group of prolific tweet fabricators encompassed a small set of hyperactive bots (19k accounts) that acted continuously over the whole 2018 broadcasting 30.6M tweets and most of them (16k accounts) were also operating in the 2020 U.S. presidential debate, at the time of this writing. This finding becomes even more concerning when considering that such hyperactive accounts promoted a vast majority of conspiratorial narratives related to the QAnon movement, similarly to conspiracy bots identified in Ferrara, Chang, et al. (2020).

Further, we noticed that bots entered into the midterm debate at a more regular and lower rate with respect to human accounts. Specifically, a constant number of bot accounts were injected and participated in the midterm conversation every week of the year, suggesting the implementation of a scheduled and cautious strategy to infiltrate bots within online discussion, attempting to avoid detection. Besides these bots, a relevant fraction of automated accounts were created a few months prior to the election (similarly to Ferrara, 2020) and participated in the related discussion, which might indicate that such accounts had been purposely injected for operating during the midterm elections debate. Interestingly, these bots generated significantly more original content, while decreasing usage of the retweet, with respect to older bot accounts. This was particularly evident for bot accounts with a moderate engagement in the midterm debate (i.e., ordinary accounts), which suggests that different operational tasks were deployed between more recent and less recent bot accounts, also based on their degree of engagement within the conversation (i.e., hyperactive vs. ordinary accounts).

Finally, bots resulted to be more deeply connected in the social network with respect to humans, especially the hyperactive and less recent ones. A fraction of more recent bot accounts were capable of infiltrating the online discussion to the extent to rapidly gain trust, standing in a central position in the social network in a relatively short time window. In terms of interactions, bots equally distributed their retweets towards human- and bot-generated content, while their replies were more focused on the interplay with humans. For what pertains to human accounts, it is alarming noting that one over three human retweets was a re-share of content produced by bots. Such indiscriminate re-sharing, which allegedly occurred because of human ineptitude in recognizing inaccurate sources of information, represents a complex issue that might have dramatic repercussions in the spread of misinformation (Bessi and Ferrara, 2016; Shao, et al., 2018; Vosoughi, et al., 2018) and further highlights the need for intervention and regulation. However, our results also reveal that bots received less interaction from humans in terms of replies (if compared to retweets), which might represent an encouraging finding when considering user awareness of and response to the activity of malicious entities on social media.

Concluding, our study confirms how the detection of manipulation campaigns is a complex and challenging task, even when suspected malicious accounts are monitored over an extended observation window. Indeed, we have shown that a pivotal set of bots mimicked human sharing activities by emulating the volume and temporal patterns of their narratives several months before the voting event while dominating the debate and infiltrating the most influential areas of the social network. These results should be a cause of concern when considering the persuasive power of social media, especially in the political context.

However, our analysis revealed novel patterns (e.g., the increasing injection of bots aimed at creating original content...
as the election approached) and actionable insights (e.g., hyperactive bot accounts started their sharing activity one year before the election and were responsible for fueling conspiratorial narratives) that can empower the identification of malicious entities and, accordingly, contribute to the detection of coordinated campaigns. These findings, and corresponding analysis, will be validated in our future work by encompassing both other discussions on future voting events as well as social conversations on diverse contexts (e.g., public health). As we neither focused on the spread of misinformation nor on the content shared by the examined accounts, with the only exception related to the conversations about the QAnon conspiracy theories, these aspects will be investigated in our next studies.

About the authors

Luca Luceri is a PostDoctoral researcher at the University of Applied Sciences and Arts of Southern Switzerland (SUPSI) in Manno, Switzerland. To whom correspondence should be addressed: luca [dot] luceri [at] supsi [dot] ch

Felipe Cardoso is a researcher at the University of Applied Sciences and Arts of Southern Switzerland (SUPSI) in Manno, Switzerland. E-mail: felipe [dot] cardoso [at] supsi [dot] ch

Silvia Giordano is a Professor at the University of Applied Sciences and Arts of Southern Switzerland (SUPSI) in Manno, Switzerland. E-mail: silvia [dot] giordano [at] supsi [dot] ch

Acknowledgments

The authors are funded by the European Institute of Innovation & Technology via the project Virtual Machina, and by the Swiss National Science Foundation via the project Late teenagers Online Information Search (LOIS). The authors are grateful to Emilio Ferrara’s team at USC for releasing the Twitter 2020 U.S. presidential election dataset, and to Goran Muric (USC) for his support with the dataset, and to First Monday’s anonymous reviewers for their very useful comments.

Notes

1. https://botometer.iuni.iu.edu/.

2. https://www.bbc.com/news/53498434.

References

A. Addawood, A. Badawy, K. Lerman, and E. Ferrara, 2019. “Linguistic cues to deception: Identifying political trolls on social media,” Proceedings of the International AAAI Conference on Web and Social Media, volume 13, pp. 15–25, and at https://ojs.aaai.org/index.php/ICWSM/article/view/3205, accessed 11 February 2021.

H. Allcott and M. Gentzkow, 2017. “Social media and fake news in the 2016 election,” Journal of Economic Perspectives, volume 31, number 2, pp. 211–236. doi: https://doi.org/10.1257/jep.31.2.211, accessed 11 February 2021.

A. Badawy, E. Ferrara, and K. Lerman, 2018. “Analyzing the digital traces of political manipulation: The 2016 Russian interference Twitter campaign,” ASONAM ’18: Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 258–265.

A. Badawy, A. Addawood, K. Lerman, and E. Ferrara, 2019. “Characterizing the 2016 Russian IRA influence campaign,” Social Network Analysis and Mining, volume 9, article number 31. doi: https://doi.org/10.1007/s13278-019-0578-6, accessed 11 February 2021.
A. Bessi and E. Ferrara, 2016. “Social bots distort the 2016 US presidential election online discussion,” First Monday, volume 21, number 11, at https://firstmonday.org/article/view/7090/5653, accessed 11 February 2021. doi: https://doi.org/10.5210/fm.v21i11.7090, accessed 11 February 2021.

A. Bovet and H.A. Makse, 2019. “Influence of fake news in Twitter during the 2016 US presidential election,” Nature Communications, volume 10, article number 7 (2 January), at https://www.nature.com/articles/s41467-018-07761-2, accessed 11 February 2021. doi: https://doi.org/10.1038/s41467-018-07761-2, accessed 11 February 2021.

P.B. Brandtzæg, 2010. “Towards a unified media-user typology (MUT): A meta-analysis and review of the research literature on media-user typologies,” Computers in Human Behavior, volume 26, number 5, pp. 940–956. doi: https://doi.org/10.1016/j.chb.2010.02.008, accessed 11 February 2021.

D.A. Broniatowski, A.M. Jamison, S. Qi, L. AlKulaib, T. Chen, A. Benton, S.C. Quinn, and M. Dredze, 2018. “Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate,” American Journal of Public Health, volume 108, number 10, pp. 1,378–1,384. doi: https://doi.org/10.2105/AJPH.2018.304567, accessed 11 February 2021.

F. Cardoso, L. Luceri, and S. Giordano, 2020. “Digital weapons in social media manipulation campaigns,” Workshop Proceedings of the 14th International AAAI Conference on Web and Social Media, at http://workshop-proceedings.icwsm.org/pdf/2020_32.pdf, accessed 11 February 2021.

N. Chavoshi, H. Hamooni, and A. Mueen, 2016. “Debot: Twitter bot detection via warped correlation,” 2016 IEEE 16th International Conference on Data Mining (ICDM), pp. 817–822. doi: https://doi.org/10.1109/ICDM.2016.0096, accessed 11 February 2021.

E. Chen, A. Deb, and E. Ferrara, 2020. “#election2020: The first public Twitter dataset on the 2020 US presidential election,” arXiv:2010.00600 (1 October), at https://arxiv.org/abs/2010.00600, accessed 11 February 2021.

Z. Chen and D. Subramanian, 2018. “An unsupervised approach to detect spam campaigns that use botnets on Twitter,” >arXiv:1804.05232 (14 April), at https://arxiv.org/abs/1804.05232, accessed 11 February 2021.

S. Cresci, 2020. “A decade of social bot detection,” Communications of the ACM, volume 63, number 10, pp. 72–83. doi: https://doi.org/10.1145/3409116, accessed 11 February 2021.

S. Cresci, M. Petrocchi, A. Spognardi, and S. Tognazzi, 2019. “Better safe than sorry: An adversarial approach to improve social bot detection,” WebSci ’19: Proceedings of the 10th ACM Conference on Web Science, pp. 47–56. doi: https://doi.org/10.1145/3292522.3326030, accessed 11 February 2021.

S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, 2020. “Emergent properties, models, and laws of behavioral similarities within groups of Twitter users,” Computer Communications, volume 150, pp. 47–61. doi: https://doi.org/10.1016/j.comcom.2019.10.019, accessed 11 February 2021.

S. Cresci, F. Lillo, D. Regoli, S. Tardelli, and M. Tesconi, 2018. “SFAKE: Evidence of spam and bot activity in stock microblogs on Twitter,” Twelfth International AAAI conference on Web and Social Media, at https://www.aaai.org/Library/ICWSM/icwsm18contents.php, accessed 11 February 2021.

C.A. Davis, O. Varol, E. Ferrara, A. Flammini, and F. Menczer, 2016. “BotOrNot: A system to evaluate social bots,” WWW ’16 Companion: Proceedings of the 25th International Conference Companion on World Wide Web, pp. 273–274. doi: https://doi.org/10.1145/2872518.2889302, accessed 11 February 2021.

A. Deb, L. Luceri, A. Badawy, and E. Ferrara, 2019. “Perils and challenges of social media and election manipulation analysis: The 2018 US midterms,” WWW ’19: Companion Proceedings of The 2019 World Wide Web Conference, pp. 237–247. doi: https://doi.org/10.1145/3308560.3316486, accessed 11 February 2021.

E. Ferrara, 2020. “What types of COVID-19 conspiracies are populated by Twitter bots?” First Monday, volume 25, number 6, at https://firstmonday.org/article/view/10633/9548, accessed 11 February 2021. doi: http://dx.doi.org/10.5210/fm.v25i6.10633, accessed 11 February 2021.
E. Ferrara, 2019. “The history of digital spam,” *Communications of the ACM*, volume 62, number 8, pp. 82–91. doi: [https://doi.org/10.1145/3299768](https://doi.org/10.1145/3299768), accessed 11 February 2021.

E. Ferrara, 2017. “Disinformation and social bot operations in the run up to the 2017 French presidential election,” *First Monday*, volume 22, number 8, at [https://firstmonday.org/article/view/8005/6516](https://firstmonday.org/article/view/8005/6516), accessed 11 February 2021.

doi: [https://doi.org/10.5210/fm.v22i8.8005](https://doi.org/10.5210/fm.v22i8.8005), accessed 11 February 2021.

E. Ferrara, 2015. “‘Manipulation and abuse on social media’ by Emilio Ferrara with Ching-man Au Yeung as coordinator,” *ACM SIGWEB Newsletter*, article number 4. doi: [https://doi.org/10.1145/2749279.2749283](https://doi.org/10.1145/2749279.2749283), accessed 11 February 2021.

E. Ferrara, S. Cresci, and L. Luceri, 2020. “Misinformation, manipulation, and abuse on social media in the era of COVID-19,” *Journal of Computational Social Science*, volume 3, number 2, pp. 271–277. doi: [https://doi.org/10.1007/s42001-020-00094-5](https://doi.org/10.1007/s42001-020-00094-5), accessed 11 February 2021.

E. Ferrara, H. Chang, E. Chen, G. Muric, and J. Patel, 2020. “Characterizing social media manipulation in the 2020 U.S. presidential election,” *First Monday*, volume 25, number 11, at [https://firstmonday.org/article/view/11431/9993](https://firstmonday.org/article/view/11431/9993), accessed 11 February 2021.

doi: [https://doi.org/10.5210/fm.v25i11.11431](https://doi.org/10.5210/fm.v25i11.11431), accessed 11 February 2021.

T.P. Gerber and J. Zavisca, 2016. “Does Russian propaganda work?” *Washington Quarterly*, 39, 2, pp. 79–98. doi: [https://doi.org/10.1080/0163660X.2016.1204398](https://doi.org/10.1080/0163660X.2016.1204398), accessed 11 February 2021.

F. Giglietto, N. Righetti, L. Rossi, and G. Marino, 2020. “Coordinated link sharing behavior as a signal to surface sources of problematic information on Facebook,” *SMSociety’20: International Conference on Social Media and Society*, pp. 85–91. doi: [https://doi.org/10.1145/3400806.3400817](https://doi.org/10.1145/3400806.3400817), accessed 11 February 2021.

T. Graham and S. Wright, 2014. “Discursive equality and everyday talk online: The impact of ‘superparticipants’,” *Journal of Computer-Mediated Communication*, volume 19, number 3, pp. 625–642. doi: [https://doi.org/10.1111/jcc4.12016](https://doi.org/10.1111/jcc4.12016), accessed 11 February 2021.

N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, and D. Lazer, 2019. “Fake news on Twitter during the 2016 U.S. presidential election,” *Science*, volume 363, number 6425 (25 January), pp. 374–378. doi: [https://doi.org/10.1126/science.aau2706](https://doi.org/10.1126/science.aau2706), accessed 11 February 2021.

P.N. Howard, G. Bolsover, B. Kollanyi, S. Bradshaw, and L.-M. Neudert, 2017. “Junk news and bots during the U.S. election: What were Michigan voters sharing over Twitter?” *Oxford Project on Computational Propaganda, Data Memo*, 2017.1, at [https://comprop.ox.ac.uk/research/posts/junk-news-and-bots-during-the-u-s-election-what-were-michigan-voters-sharing-over-twitter/](https://comprop.ox.ac.uk/research/posts/junk-news-and-bots-during-the-u-s-election-what-were-michigan-voters-sharing-over-twitter/), accessed 11 February 2021.

A. Hughes and N. Asheer, 2019. “National politics on Twitter: Small share of U.S. adults produce majority of tweets,” Pew Research Center (23 October), at [https://www.pewresearch.org/politics/wp-content/uploads/sites/4/2019/10/PDL_10.23.19_politics.twitter_FULLREPORT.pdf](https://www.pewresearch.org/politics/wp-content/uploads/sites/4/2019/10/PDL_10.23.19_politics.twitter_FULLREPORT.pdf), accessed 11 February 2021.

P.-M. Hui, K.-C. Yang, C. Torres-Lugo, and F. Menczer, 2020. “BotSlayer: DIY real-time influence campaign detection,” *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14, pp. 980–982, and at [https://ojs.aaai.org//index.php/ICWSM/article/view/7370](https://ojs.aaai.org//index.php/ICWSM/article/view/7370), accessed 11 February 2021.

J. Im, E. Chandrasekharan, J. Sargent, P. Lighthammer, T. Denby, A. Bhargava, L. Hemphill, D. Jurgens, and E. Gilbert, 2019. “Still out there: Modeling and identifying Russian troll accounts on Twitter,” *arXiv:1901.11162* (31 January), at [https://arxiv.org/abs/1901.11162](https://arxiv.org/abs/1901.11162), accessed 11 February 2021.
Down the bot hole: Actionable insights from a one-year analysis of bot activity on Twitter

K.H. Jamieson, 2018. Cyberwar: how Russian hackers and trolls helped elect a president: what we don’t, can’t, and do know. New York: Oxford University Press.

S.M. Jones-Jang, D.H Kim, and K. Kenski, 2020. “Perceptions of mis-or disinformation exposure predict political cynicism: Evidence from a two-wave survey during the 2018 US midterm elections,” New Media & Society (20 July). doi: https://doi.org/10.1177/1461444820943878, accessed 11 February 2021.

B. Kollanyi, P.N. Howard, and S. Woolley, 2016. “Bots and automation over Twitter during the U.S. election,” Oxford Project on Computational Propaganda, Data Memo, 2016.4, at https://comprop.oii.ox.ac.uk/research/posts/bots-and-automation-over-twitter-during-the-u-s-election/, accessed 11 February 2021.

L. Luceri, S. Giordano, and E. Ferrara, 2020. “Detecting troll behavior via inverse reinforcement learning: A case study of Russian trolls in the 2016 US election,” Proceedings of the International Conference of Web and Social Media, volume 14, at https://ojs.aaai.org/index.php/ICWSM/article/view/7311, accessed 11 February 2021.

L. Luceri, A. Deb, A. Badawy, and E. Ferrara, 2019a. “Red bots do it better: Comparative analysis of social bot partisan behavior,” WWW ’19: Companion Proceedings of the 2019 World Wide Web Conference, pp. 1,007–1,012. doi: https://doi.org/10.1145/3308560.3316735, accessed 11 February 2021.

L. Luceri, A. Deb, S. Giordano, and E. Ferrara, 2019b. “Evolution of bot and human behavior during elections,” First Monday, volume 24, number 9, at https://firstmonday.org/article/view/10213/8073, accessed 11 February 2021. doi: https://doi.org/10.5210/fm.v24i9.10213, accessed 11 February 2021.

M. Mazza, S. Cresci, M. Avvenuti, W. Quattrociocchi, and M. Tesconi, 2019. “RTbust: exploiting temporal patterns for botnet detection on Twitter,” WebSci ’19: Proceedings of the 10th ACM Conference on Web Science, pp. 183–192. doi: https://doi.org/10.1145/3292522.3326015, accessed 11 February 2021.

P.T. Metaxas and E. Mustafaraj, 2012. “Social media and the elections,” Science, volume 338, number 6106 (26 October), pp. 472–473. doi: https://doi.org/10.1126/science.1230456, accessed 11 February 2021.

L. Nizzoli, S. Tardelli, M. Avvenuti, S. Cresci, M. Tesconi, and E. Ferrara, 2020. “Charting the landscape of online cryptocurrency manipulation,” arXiv:2001.10289 (28 January), at https://arxiv.org/abs/2001.10289, accessed 11 February 2021.

D. Pacheco, P.-M. Hui, C. Torres-Lugo, B.T. Truong, A. Flammini, and F. Menczer, 2020. “Uncovering coordinated networks on social media,” arXiv:2001.05658 (16 January), at https://arxiv.org/abs/2001.05658, accessed 11 February 2021.

G. Pennycook and D.G. Rand, 2019. “Fighting misinformation on social media using crowdsourced judgments of news source quality,” Proceedings of the National Academy of Sciences, volume 116, number 7 (12 February), pp. 2,521–2,526. doi: https://doi.org/10.1073/pnas.1806781116, accessed 11 February 2021.

J. Ratkiewicz, M.D. Conover, M. Meiss, B. Goncalves, A. Flammini, and F.M. Menczer, 2011. “Detecting and tracking political abuse in social media,” Proceedings of the International Conference of Web and Social Media, volume 11, pp. 297–304, at https://www.aaai.org/Library/ICWSM/icwsm11contents.php, accessed 11 February 2021.

D.J. Ruck, N.M. Rice, J. Borycz, and R.A. Bentley, 2019. “Internet Research Agency Twitter activity predicted 2016 U.S. election polls,” First Monday, volume 24, number 7, at https://firstmonday.org/article/view/10107/8049, accessed 11 February 2021. doi: https://doi.org/10.5210/fm.v24i7.10107, accessed 11 February 2021.

M. Sayyadiharikandeh, O. Varol, K.-C. Yang, A. Flammini, and F. Menczer, 2020. “Detection of novel social bots by ensembles of specialized classifiers,” arXiv:2006.06867 (14 August), at https://arxiv.org/abs/2006.06867, accessed 11 February 2021.

D.A. Scheufele and N.M. Krause, 2019. “Science audiences, misinformation, and fake news,” Proceedings of the National Academy of Sciences, volume 116, number 16 (16 April), pp. 7,662–7,669.
C. Shao, G.L. Ciampaglia, O. Varol, K.-C. Yang, A. Flammini, and F. Menczer, 2018. “The spread of low-credibility content by social bots,” *Nature Communications*, volume 9, article number 4787 (20 November), at https://www.nature.com/articles/s41467-018-06930-7, accessed 11 February 2021.

doi: https://doi.org/10.1038/s41467-018-06930-7, accessed 11 February 2021.

K. Sharma, E. Ferrara, and Y. Liu, 2020. “Identifying coordinated accounts in disinformation campaigns,” *arXiv*:2008.11308 (25 August), at https://arxiv.org/abs/2008.11308, accessed 11 February 2021.

doi: https://doi.org/10.1038/s41467-018-06930-7, accessed 11 February 2021.

K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, 2017. “Fake news detection on social media: A data mining perspective,” *ACM SIGKDD Explorations Newsletter*, volume 19, number 1, pp. 22–36.

doi: https://doi.org/10.1145/3137597.3137600, accessed 11 February 2021.

M. Stella, M. Cristoforetti, and M. De Domenico, 2019. “Influence of augmented humans in online interactions during voting events,” *PloS ONE*, volume 14, number 5, e0214210.

doi: https://doi.org/10.1371/journal.pone.0214210, accessed 11 February 2021.

M. Stella, E. Ferrara, and M. De Domenico, 2018. “Bots increase exposure to negative and inflammatory content in online social systems,” *Proceedings of the National Academy of Sciences*, volume 115, number 49 (20 November), pp. 12,435–12,440.

doi: https://doi.org/10.1073/pnas.1803470115, accessed 11 February 2021.

L.G. Stewart, A. Arif, and K. Starbird, 2018. “Examining trolls and polarization with a retweet network,” *Proceedings of WSDM Workshop on Misinformation and Misbehavior Mining on the Web*, at https://faculty.washington.edu/kstarbi/examining-trolls-polarization.pdf, accessed 11 February 2021.

V.S. Subrahmanian, A. Azaria, S. Durst, V. Kagan, A. Galstyan, K. Lerman, L. Zhu, E. Ferrara, A. Flammini, F. Menczer, A. Stevens, A. Dekhtyar, S. Gao, T. Hogg, F. Kooti, Y. Liu, O. Varol, P. Shiralkar, V. Vydiswaran, Q. Mei, and Tim Hwang, 2016. “The DARPA Twitter bot challenge,” *arXiv*:1601.05140 (21 April), at https://arxiv.org/abs/1601.05140, accessed 11 February 2021.

S. Tardelli, M. Avvenuti, M. Tesconi, and S. Cresci, 2020. “Characterizing social bots spreading financial disinformation,” In: G. Meiselwitz (editor). *Social computing and social media. Design, ethics, user behavior, and social network analysis. HCII 2020. Lecture Notes in Computer Science*, volume 12194. Cham, Switzerland: Springer, pp. 376–392.

doi: https://doi.org/10.1007/978-3-030-49570-1_26, accessed 11 February 2021.

D. Trielli and N. Diakopoulos, 2020. “Partisan search behavior and Google results in the 2018 U.S. midterm elections,” *Information, Communication & Society* (18 May).

doi: https://doi.org/10.1080/1369118X.2020.1764605, accessed 11 February 2021.

O. Varol, E. Ferrara, F. Menczer, and A. Flammini, 2017. “Early detection of promoted campaigns on social media,” *EPJ Data Science*, volume 6, article number 13, at https://epjdatascience.springeropen.com/articles/10.1140/epjs/s13688-017-0111-y, accessed 11 February 2021.

doi: https://doi.org/10.1140/epjs/s13688-017-0111-y, accessed 11 February 2021.

O. Varol, E. Ferrara, C.A. Davis, F. Menczer, and A. Flammini, 2017. “Online human-bot interactions: Detection, estimation, and characterization,” *International AAAI Conference on Web and Social Media*, pp. 280–289, and at https://www.aaai.org/Library/ICWSM/17/contents.php, accessed 11 February 2021.

S. Vosoughi, D. Roy, and S. Aral, 2018. “The spread of true and false news online,” *Science*, volume 359, number 6380 (9 March), pp. 1,146–1,151.

doi: https://doi.org/10.1126/science.aap9559, accessed 11 February 2021.

S.C. Woolley and P.N. Howard, 2018. *Computational propaganda: Political parties, politicians, and political manipulation on social media*. New York: Oxford University Press.

doi: https://doi.org/10.1093/oso/9780190931407.001.0001, accessed 11 February 2021.

L. Wrubel, J. Littman, and D. Kerchner, 2019. “2018 U.S. congressional election tweet ids,” *Harvard Dataverse*. 

---

Down the bot hole: Actionable insights from a one-year analysis of bot activity on Twitter

doi: https://doi.org/10.1073/pnas.1805871115, accessed 11 February 2021.
Appendix

Account classification

In this section, we describe the procedure followed to classify hyperactive and ordinary accounts based on active days and daily tweets frequency parameters (see section Engagement-based account classification for further details). To accomplish this task, we needed to identify a threshold for each considered parameter that permits a separation of hyperactive users from ordinary ones. More specifically, we observed how ordinary (resp. hyperactive) accounts were filtered out (resp. selected) from the whole set of accounts by varying these thresholds.

In Figure 12, we show how the percentage of retained accounts decreased as the filtering parameters (i.e., active days and tweets frequency) increased.
In particular, in the left panel of Figure 12, we show the percentage of accounts that persist to such a filtering as a function of the filtering parameters. In the right panel of Figure 12, we only focus on the two-dimensional relation between the percentage of accounts and the tweet frequency parameter. Every dot in the figure indicates a combination of the filtering parameters, while its color represents the percentage of accounts that persist to filtering. Two facts are worth noting. First, the active days parameter does not significantly affect the percentage of filtered accounts, except for the first gap (from 100 percent to 80 percent of the accounts), where the accounts active only one day were filtered out. At a fixed tweet frequency, the percentage of filtered accounts does not vary by increasing the active days parameter (dots appear to be aligned). Second, the tweet frequency parameter highly affected the percentage of filtered accounts: as this parameter increased, the percentage of accounts decreased. The most noticeable gap can be appreciated when the tweet frequency parameter value equaled 0.2 tweets/day, which resulted in a filtering of almost 70 percent of the accounts.

In Figure 13, the same evaluation was performed by considering the percentage of tweets that persisted to account filtering (based on the same set of parameters).
In Figure 13, the parameter that affected most of the percentage of filtered tweets was tweet frequency. However, in Figure 13, the percentage of tweets decreased more gradually by increasing the tweet frequency parameter if compared to the percentage of account reduction (Figure 12). Interestingly, with a tweet frequency of 0.2 tweets/day, more than 70 percent of accounts were filtered out but more than 90 percent of tweets were retained, suggesting that a small percentage of accounts was responsible for a huge amount of tweets, consistent with previous findings (Hughes and Asheer, 2019; Zelenkauskaite and Balduccini, 2017). Given the above observations, we classify as hyperactive accounts that satisfy the following conditions:

- \( \text{active days} \geq 1 \)
- \( \text{daily tweets frequency} \geq 1 \)

We consider these parameters as (i) for the active days parameter the most noticeable variation in the percentage of filtered accounts can be appreciated when this parameter was greater than 0; and, (ii) a tweet frequency of one corresponded to the knee of the curve in the account distribution (Figure 12), which can be viewed as a conservative choice to discriminate the two classes of accounts (hyperactive vs. ordinary), as no significant difference would occur in the classification by further increasing this threshold, \( i.e., \) the percentage of accounts does not significantly vary by increasing the tweet frequency parameter.

**Accounts involved in both 2018 and 2020 debates**

Here we provide detailed information about the accounts that were analyzed in this paper and were also involved, at the time of this writing (mid-October 2020), in the 2020 U.S. presidential election debate on Twitter.

We refer to this subset of accounts as \emph{users 2018 & 2020}, while we denote the whole set of accounts analyzed in this paper as \emph{users 2018}.

To attain the set of users engaged in the 2020 U.S. presidential election debate, we leveraged the Twitter dataset collected by Chen, \textit{et al.} (2020). Among the set of users involved in the more recent debate, we spotted only those that were analyzed in our paper and, accordingly, identify 582,349 \emph{users 2018 & 2020}. In \textit{Table 5}, we detail the statistics of such users and compare them with the \emph{users 2018}, whereas \textit{Table 6} shows further details related to old \emph{users 2018 & 2020}.
Table 5: *Users 2018 & 2020 and users 2018 statistics.*

|                        | All   | Bots  | Humans | Hyperactive | Ordinary | Old    | Recent |
|------------------------|-------|-------|--------|-------------|----------|--------|--------|
| **Users 2018 & 2020**  | 582k  | 93k   | 489k   | 60k         | 467k     | 251k   | 331k   |
| (Users 2018)           | (943k)| (184k)| (759k) | (84k)       | (698k)   | (322k) | (621k) |
| **Tweets by users 2018 & 2020** | 77M  | 33M   | 44M    | 56M         | 21M      | 64M    | 13M    |
| (Tweets by users 2018) | (98M)| (42M) | (56M)  | (71M)       | (27M)    | (79M)  | (19M)  |

Table 6: *Old users 2018 & 2020 statistics.*

|                              | Old hyperactive | Old ordinary |
|------------------------------|-----------------|--------------|
| (Accounts 2018 & 2020)       | 42k (51k)       | 207k (266k)  |
| (Accounts 2018)              |                 |              |
| Bots 2018 & 2020             | 16k (19k)       | 28k (40k)    |
| (Bots 2018)                  |                 |              |
| Humans 2018 & 2020           | 26k (32k)       | 179k (226k)  |
| (Humans 2018)                |                 |              |

Hyperactive accounts

Table 7 summarizes the properties related to both hyperactive bot and human accounts and the volume of their shared tweets.

Table 7: *Properties of hyperactive accounts.*

|                  | Value | % (% within hyperactive accounts) |
|------------------|-------|-----------------------------------|
| Total accounts   | 83,701| 8.87% (100%)                      |
| Bot accounts     | 31,772| 3.37% (37.96%)                    |
| Human            |       |                                   |
Down the bot hole: Actionable insights from a one-year analysis of bot activity on Twitter

| accounts | 51,929 | 5.51% (62.04%) |
|----------|--------|----------------|
| Total tweets | 70,737,005 | 72.09% (100%) |
| Bot tweets | 37,169,582 | 37.88% (52.55%) |
| Human tweets | 33,567,423 | 34.21% (47.45%) |

Figure 14 displays the sharing activities of hyperactive accounts over the observation period.

Figure 14: Weekly sharing activities for hyperactive accounts.

Ordinary accounts

Table 8 summarizes the properties related to both ordinary bot and human accounts and the volume of their shared tweets.
Table 8: Properties of ordinary accounts.

|                  | Value   | % (% within ordinary accounts) |
|------------------|---------|--------------------------------|
| Total accounts   | 698,427 | 74.05% (100.0%)                |
| Bot accounts     | 118,503 | 12.56% (16.97%)                |
| Human accounts   | 579,924 | 61.49% (83.03%)                |
| Total tweets     | 27,225,609 | 27.75% (100.0%)              |
| Bot tweets       | 5,182,417 | 5.28% (19.04%)                |
| Human tweets     | 22,043,192 | 22.46% (80.96%)              |

Figure 15 displays the sharing activities of ordinary accounts over the observation period.

Figure 15: Weekly sharing activities for ordinary accounts.
Table 9 summarizes the properties related to accounts that shared only one tweet.

|                        | Value | % (% within one-tweet accounts) |
|------------------------|-------|---------------------------------|
| **Total accounts**     | 160,998 | 17.07% (100.0%)               |
| **Bot accounts**       | 34,270 | 3.63% (21.29%)                 |
| **Human accounts**     | 126,728 | 13.43% (78.71%)               |
| **Total tweets**       | 160,998 | 0.16% (100.0%)                |
| **Bot tweets**         | 34,270 | 0.03% (21.29%)                 |
| **Human tweets**       | 126,728 | 0.13% (78.71%)                |

**Suspended accounts**

Here, we examine the correlation between account suspension and bot scores. Following the methodology proposed by Ferrara (2020), we leveraged the annotations of verified and suspended accounts and compared their bot scores. In Figure 16, we depict the bot score distributions of 38k suspended accounts and 22.7k verified accounts within our dataset.
We observe two different distributions, as statistically confirmed by the Mann-Whitney rank test ($p$-value<0.001). Indeed, bot scores of the suspended accounts were more widely distributed with respect to those of the verified accounts, which had about 80 percent of the users with a bot score lower than 0.15. To the contrary, more than half of the suspended accounts presented bots scores higher than 0.15. It should be noticed that account suspension from Twitter may occur for several reasons, not only related to the degree of automation of the account. Therefore, it was expected to also have suspended accounts with low bot scores. However, it can be noted how thousands of suspended accounts reflect a high level of automation (i.e., high bot scores), while the number of verified accounts dropped as the bot score increased. This, besides confirming a correlation between account suspension and the likelihood of automation, supports the account classification performed via Botometer (Ferrara, 2020).

**QAnon tweets**

The keywords used to gather tweets related to QAnon are listed as follows:
#qanon, #8kun, #qpatriots, #q, #qarmy, #qanon2018, #qanon2019, #qwarriors, #thegreatawakening, #wwg1wga, #hollywoodanon, #qteam, #qanons, #qclearance, #qbaby, #pedogate, #pizzagate, #darktolyight, #wakeupamerica, #thecollectiveq, #thestorm

---

**Editorial history**

Received 28 October 2020; revised 9 February 2021; accepted 9 February 2021.
To the extent possible under law, this work is dedicated to the public domain.

Down the bot hole: Actionable insights from a one-year analysis of bot activity on Twitter by Luca Luceri, Felipe Cardoso, and Silvia Giordano. *First Monday*, volume 26, number 3 (March 2021).

doi: [http://dx.doi.org/10.5210/fm.v26i3.11441](http://dx.doi.org/10.5210/fm.v26i3.11441)