Characterizing Social Bots Spreading Financial Disinformation

Serena Tardelli[1], Marco Avvenuti[2], Maurizio Tesconi[3], and Stefano Cresci[3](✉)

[1] IIT-CNR and Department of Information Engineering,
University of Pisa, Pisa, Italy
serena.tardelli@iit.cnr.it
[2] Department of Information Engineering, University of Pisa, Pisa, Italy
marco.avvenuti@unipi.it
[3] IIT-CNR, Pisa, Italy
{maurizio.tesconi,stefano.cresci}@iit.cnr.it

Abstract. Despite the existence of several studies on the characteristics and role of social bots in spreading disinformation related to politics, health, science and education, *financial social bots* remain a largely unexplored topic. We aim to shed light on this issue by investigating the activities of large social botnets in Twitter, involved in discussions about stocks traded in the main US financial markets. We show that the largest discussion spikes are in fact caused by mass-retweeting bots. Then, we focus on characterizing the activity of these financial bots, finding that they are involved in speculative campaigns aimed at promoting low-value stocks by exploiting the popularity of high-value ones. We conclude by highlighting the peculiar features of these accounts, comprising similar account creation dates, similar screen names, biographies, and profile pictures. These accounts appear as untrustworthy and quite simplistic bots, likely aiming to fool automatic trading algorithms rather than human investors. Our findings pave the way for the development of accurate detection and filtering techniques for financial spam. In order to foster research and experimentation on this novel topic, we make our dataset publicly available for research purposes.

Keywords: Social bots · Disinformation · Deception · Financial spam · Stock markets · Twitter

1 Introduction

Nowadays, social bots play a pivotal role in shaping the content of online social media [25]. Their involvement in the spread of disinformation ranges from the promotion of low-credibility content, astroturfing, and fake endorsements, to the propagation of hate speech propaganda, in attempts to manipulate public opinion and to increase societal polarization [26]. Indeed, recent studies have observed the presence of artificial tampering in a wide variety of online topic

G. Meiselwitz (Ed.): HCII 2020, LNCS 12194, pp. 376–392, 2020.
https://doi.org/10.1007/978-3-030-49570-1_26

debates, including political discussions, terrorist propaganda, and health controversies [8].

A growing field under scrutiny is the online financial ecosystem, in which social bots now pervade [14,24]. Indeed, such an ecosystem has proven to be of great interest as a valuable ground to entice investors. Although the leverage of social media content for predicting trends in the stock market has promising potential [6], the presence of social bots in such scenarios poses serious concerns over the reliability of financial information. Examples of repercussions of financial spam on unaware investors and automated trading systems include the real-world event known as the Flash Crash – the one-day collapse of the Dow Jones Industrial Average in 2010 induced by an error in the estimation of online information by automated trading systems [19]. Another most notable example is the hacking of the US International Press Officer's official Twitter account in 2013, when a bot reported the injury of President Obama following a terrorist attack, causing a major stock market drop in a short time[1]. Finally, we witnessed to the abrupt rise of *Cynk Technology* in 2014 from an unknown unprosperous small company to a billions-worth company, due to a social bot orchestration that lured automatic trading algorithms into investing in the company's shares based on a fake social discussion, which ultimately resulted in severe losses[2]. Therefore, investigating such manipulations and characterizing them is of the utmost importance in order to protect our markets from manipulation and to safeguard our investments.

Contributions. In an effort to shed light on the little-studied activity of social bots tampering with online financial discussions, we analyze a rich dataset of 9M tweets discussing stocks of the five main US financial markets. Our dataset is complemented with financial information collected from Google Finance, for each of the stocks mentioned in our tweets. By comparing social and financial information, we report on the activity of large botnets perpetrating speculative campaigns aimed at promoting low-value stocks by exploiting the popularity of high-value ones. We highlight the main characteristics of these financial bots, which appear as untrustworthy, simplistic accounts. Based on these findings, we conclude that their activity is likely aimed at fooling automatic trading algorithms rather than human investors.

Our main contributions are analytically summarized as follows:

- We outline the activities and role of social bots in the spread of financial disinformation on Twitter.
- We uncover the existence of several large botnets, actively involved in artificially promoting specific stocks.
- We characterize social bots tampering with financial discussions from various perspectives, including their content, temporal, and social facets.

[1] https://www.bbc.com/news/world-us-canada-21508660.

[2] https://www.cnbc.com/2014/07/25/mysterious-stock-cynk-plummets-after-reopening.html.

Our findings provide an important contribution towards understanding the role and impact of social bots in the financial domain, and pave the way for the development of accurate detection and filtering techniques for financial spam.

2 Related Work

As anticipated, in the present study we are interested in analyzing the activity, the behavior, and the characteristics of financial social bots. For this reason, in this section we do not survey previous works related to the *detection* of social bots – which is a different topic covered by many through studies [8] – but we rather focus on those works related to the *characterization* of malicious accounts.

Given the many issues caused by malicious accounts to our online social ecosystems, a large body of work analyzed the behavior of bots and trolls in disinformation campaigns aimed at influencing a variety of debates. To understand how trolls tampered with the 2016 US Presidential elections, previous work characterized the content they disseminated, and their influence on the information ecosystem [30]. Among other findings, authors discovered that trolls were created a few weeks before important world events, and that they are more likely to retweet political content from normal Twitter users, rather than other news sources. In [31], authors evaluated the behavior and strategy changes over time of Russian and Iranian state-sponsored trolls. By exposing the way Iranian trolls changed behavior over time and started retweeting each other, they highlighted how strategies employed by trolls adapt and evolve to new campaigns. Authors in [27] detected and characterized Twitter bots and user interactions by analyzing their retweet and mention strategies, and observed a high correlation between the number of friends and followers of accounts and their bot-likelihood. In [1], authors characterized Arabic social bots spreading religious hatred on Twitter, and discovered they have a longer life, a higher number of followers, and an activity more geared towards creating original content than retweets, compared to English bots [5]. The previous works remarked the importance of understanding the inherent characteristics of bots and trolls. In fact, despite showing signs of bot-likelihood, bots do not often get caught in time, thus potentially affecting the polarization and outcome of essential debates. Moreover, previous works also highlighted how bots and trolls evolve and adapt to new contexts. Despite such consistency in previous results, the characteristics of social bots disseminating financial information are yet to be explored. The few previous works that tackled automation and disinformation in online financial discussions, went as far as providing evidence of the presence of financial spam in stock microblogs and raised concerns over the reliability of such information [13,14]. However, the detection and impact estimation of such bots in social media financial discussions still represent largely unexplored fields of study. Conversely, the leverage of social bots in other sectors has been extensively examined, with previous works focusing on the interference of bots in health issues [2], terrorist propaganda [3], and political election campaigns in the US [5], France [16], Italy [10], and Germany [7], to name but a few.

In this work, we aim at filling in the missing piece of the puzzle – that is, the characterization of social bots in online financial conversations, with a focus on how they are organized and how they operate.

Table 1. Statistics about the financial and social composition of our dataset.

Markets	Financial data			Twitter data		
	Companies	Median cap. ($)	Total cap. ($B)	Users	Tweets	Retweets (%)
NASDAQ	3,013	365,780,000	10,521	252,587	4,017,158	1,017,138 (25%)
NYSE	2,997	1,810,000,000	28,692	265,618	4,410,201	923,123 (21%)
NYSEARCA	726	245,375,000	2,227	56,101	298,445	157,101 (53%)
NYSEMKT	340	78,705,000	256	22,614	196,545	63,944 (33%)
OTCMKTS	22,956	31,480,000	45,457	64,628	584,169	446,293 (76%)
Total	30,032	–	87,152	467,241	7,855,518	1,802,705 (23%)

3 Dataset

By leveraging Twitter's Streaming API [15], we collected all tweets mentioning at least one of the 6,689 stocks listed on the official NASDAQ Web site[3]. Companies quoted in the stock market are easily identified on Twitter by means of *cashtags* – strings composed of a dollar sign followed by the ticker symbol of the company (e.g., `$AAPL` is the cashtag of `Apple, Inc.`). Just like the hashtags, cashtags serve as beacons to find, filter, and collect relevant content [18].

Our data collection covered a period of five months, from May to September 2017, and resulted in the retrieval of more than 9M tweets. We also extended the dataset by gathering additional financial information (e.g., capitalization and industrial classification) about the companies mentioned in our tweets, by leveraging the Google Finance Web site[4]. Table 1 shows summary statistics about our dataset, which is publicly available online for research purposes[5].

4 Uncovering Financial Disinformation

In this section we describe the various analyses that allowed us to uncover widespread speculative campaigns perpetrated by several botnets. For additional details on the subsequent analyses, we point interested readers to [14].

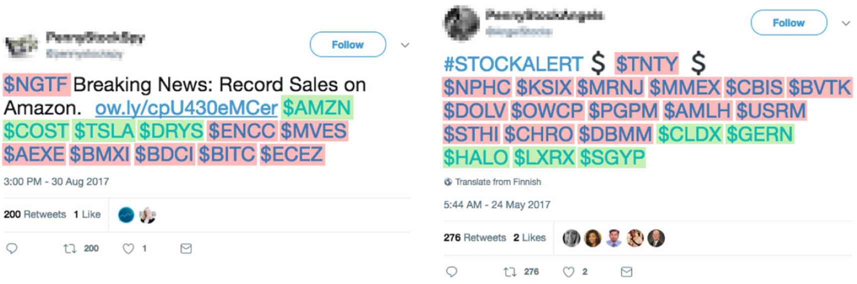
[3] http://www.nasdaq.com/screening/company-list.aspx.

[4] https://www.google.com/finance.

[5] https://doi.org/10.5281/zenodo.2686862.

4.1 Dataset Overview

Each tweet in our dataset mentions at least one of the 6,689 stocks of the NASDAQ list. Companies from this list typically feature a large market capitalization and are traded in the 4 main US financial markets – namely, NASDAQ, NYSE, NYSEARCA, and NYSEMKT. However, among tweets mentioning our 6,689 stocks, we also found many mentions of other, less known, stocks. In particular, as shown in Table 1, overall tweets of our dataset also mention 22,956 stocks traded in the OTCMKTS market. Contrarily to the main stock exchanges, OTCMKTS has less stringent constraints and mainly hosts stocks with a small capitalization. Unsurprisingly, if we analyze our whole dataset, no company from OTCMKTS appears among those that are discussed the most. In fact, the most tweeted companies in our dataset are in line with those found in previous works [18], and include well-known and popular stocks such as $AAPL, $TSLA, and $FB. Nonetheless, a few concerns rise if we consider the rate of retweets for OTCMKTS stocks, which happens to be as high as 76% and in sharp contrast with the much lower rates measured for all other markets. Since automated mass-retweets have been frequently exploited by bots and trolls to artificially boost content popularity [23], this result might hint at the possibility of a manipulation related to OTCMKTS stocks.


[redacted]

$NGTF Breaking News: Record Sales on Amazon. ow.ly/cpU430eMCer $AMZN $COST $TSLA $DRYS $ENCC $MVES $AEXE $BMXI $BDCI $BITC $ECEZ

3:00 PM - 30 Aug 2017

200 Retweets 1 Like

[redacted]

#STOCKALERT $ $TNTY $ $NPHC $KSIX $MRNJ $MMEX $CBIS $BVTK $DOLV $OWCP $PGPM $AMLH $USRM $STHI $CHRO $DBMM $CLDX $GERN $HALO $LXRX $SGYP

5:44 AM - 24 May 2017

276 Retweets 2 Likes

Fig. 1. Examples of tweets in which a few high-capitalization companies (green-colored) co-occur with many low-capitalization ones (red-colored). (Color figure online)

4.2 Investigating Financial Discussion Spikes

In order to deepen our analysis of financial conversations, we now focus our attention on discussion spikes about the 6,689 stocks of our starting list. For identifying discussion spikes, we compute for each stock the hourly time series of the volume of tweets mentioning that stock, to which we apply a simple anomaly detection technique. In detail, we label as anomalies those peaks of discussion that exceed the mean hourly volume of tweets by more than 10 standard deviations, finding in total 1,926 financial discussion spikes.

Within the discussion spikes, we found more retweets than in the rest of the dataset – namely, 60% retweets for spikes *vs* 23% for the whole dataset, on average. This finding alone does not necessarily imply a coordinated inauthentic activity, since also organic surges of interest in social media typically result in many retweets. What is unusual however, is that tweets posted during the identified discussion spikes contain, on average, many more cashtags (i.e., mentioned stocks) than the ones in the rest of the dataset. Moreover, such co-occurring stocks seem largely unrelated, and the authors of those tweets do not provide any information to explain the co-occurrences, as shown in the examples of Fig. 1.

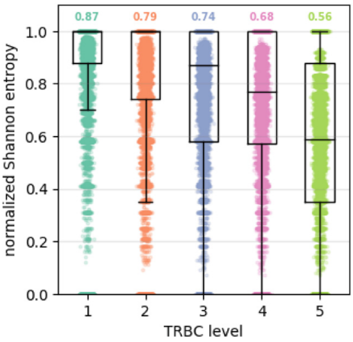


Fig. 2. Entropy of the industrial classes of co-occurring stocks in discussion spikes. As shown, the high measured entropy implies that co-occurring companies are largely unrelated.

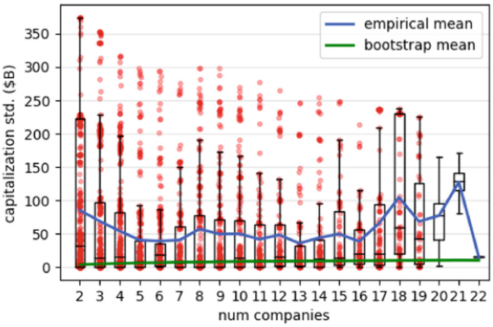


Fig. 3. Standard deviation of the capitalization of co-occurring companies in discussion spikes, and comparison with a bootstrap. The large measured standard deviation implies that high-cap companies co-occur with low-cap ones.

4.3 Co-occurring Stocks

To investigate the reasons behind this large number of co-occurring stocks, we follow two different hypotheses: (i) stocks might co-occur because of a similar industrial sector (i.e., companies involved in the same business are more likely to be mentioned together) or (ii) they might co-occur because of a similar market value (i.e., high capitalization companies are more likely to be compared to others with similar capitalization).

To assess whether our co-occurring stocks have a similar industrial sector, we leverage Thomson Reuters Business Classification (TRBC)[6]. In particular, we compute the normalized Shannon entropy between the TRBC classes of co-occurring stocks for each tweet that contributes to a discussion spike. This analysis is repeated for all 5 TRBC levels. Each entropy value measured for a given

[6] TRBC is a 5-level hierarchical sector and industry classification, widely used in the financial domain for computing sector-specific indices: https://en.wikipedia.org/wiki/Thomson_Reuters_Business_Classification.

TRBC level for discussion spikes is then compared with the corresponding one computed out of the whole dataset. Results of this analysis are shown in Fig. 2 and depict a situation characterized by large entropy values (i.e., $\simeq$1, which is the maximum possible value of normalized entropy). In turn, this implies that co-occurring companies in discussion spikes are almost completely unrelated with regards to their industrial classification. Moreover, entropy values measured for discussion spikes are always higher than those measured for the whole dataset.

Regarding financial value, we assess the extent to which co-occurring companies have a similar value by measuring the standard deviation of their market capitalizations. To understand whether the measured standard deviation is due to the intrinsic characteristics of our dataset (i.e., the underlying statistical distribution of capitalization) or to other external factors, we compared mean values of our empiric measurements with a bootstrap. Results are shown in Fig. 3 and highlight a large empiric standard deviation between the capitalization of co-occurring companies, such that a random bootstrap baseline – accounting for the intrinsic characteristics of our dataset – can not explain it. These results mean that not only high-capitalized companies indeed mostly co-occur with small-capitalized ones, as shown in Fig. 1, but also that this phenomenon is rather the consequence of some external action.

In summary, we demonstrated that tweets responsible for generating financial discussion spikes mention a large number of unrelated stocks, some of which are high-cap stocks while the others are low-cap ones.

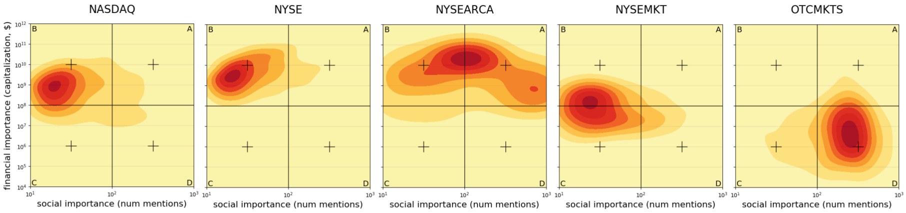


Fig. 4. Kernel density estimation investigating the relation between social and financial importance, for stocks of the 5 considered markets. `OTCMKTS` stocks have a suspiciously high social importance despite their low financial importance.

4.4 Financial vs Social Importance

Several existing systems for forecasting stock prices leverage the positive correlation between discussion volumes on social media around a given stock, and its market value [22]. In other words, it is generally believed that stocks with a high capitalization (i.e., high *financial* importance) are discussed more in social media (i.e., high *social* importance) than those with a low capitalization. In this section we verify whether this expected positive relation exists also for the stocks in our dataset.

In fact over our whole dataset, we measure a moderate positive Spearman's rank correlation coefficient of $\rho = 0.4871$ between social and financial importance, thus confirming previous findings. However, when focusing on discussion spikes only, we measure a suspicious behavior related to OTCMKTS stocks, which feature a negative $\rho = -0.2658$, meaning that low-value OTCMKTS stocks are more likely to appear in discussion spikes than high-value ones. To thoroughly understand the relation between social and financial importance, in Fig. 4 we report the results of a bi-dimensional kernel density estimation of social and financial importance for stocks of the five considered markets. Confirming previous concerns, OTCMKTS stocks feature a suspiciously high social importance, despite their low financial importance, in contrast with stocks of all other markets.

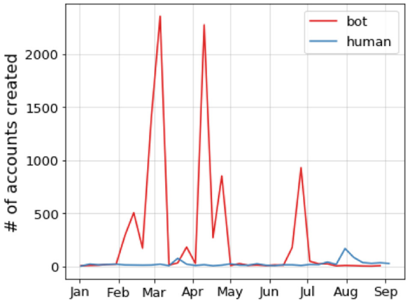


Fig. 5. Number of accounts created per week in 2017. Bot accounts display coordinated creation activities, while humans are more evenly spread across the year.

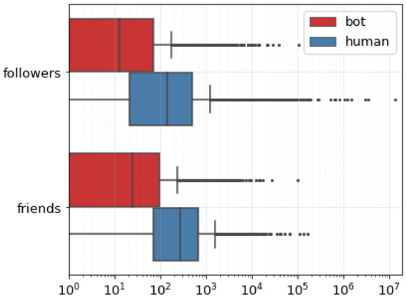


Fig. 6. Distribution of the number of followers and friends. Bot accounts show a lower number of followers and friends with respect to human accounts.

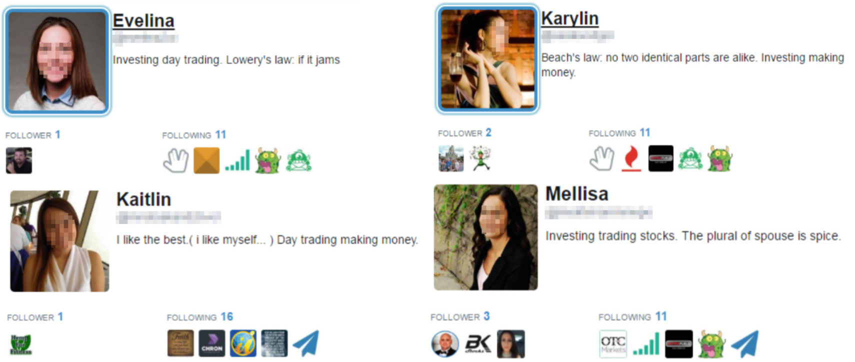
5 Bot Detection and Characterization

In the previous section, we described several suspicious phenomena related to stock microblogs. In detail, discussion spikes about high-value stocks are filled with mentions of low-value (mainly OTCMKTS) ones. Such mentions are not explained by real-world stock relatedness. Moreover, the discussion spikes are largely caused by mass retweets.

5.1 Bot Detection

In order to understand whether the previously described disorders in financial microblogs are caused by organic (i.e., human-driven) or rather by synthetic activity, here we discuss results of the application of a bot detection technique to all users that contributed to at least one of the top-100 largest discussion spikes. In this way, we analyzed roughly 50% of all our dataset, both in terms of tweets and users, in search for social bots.

To perform bot detection, we employ the state-of-the-art technique described in [10], which is based on the analysis of the sequences of actions performed by the investigated accounts. Strikingly, the technique classified as much as 71% of all analyzed users as bots. Moreover, 48% of the users classified as bots were also later suspended by Twitter, corroborating our results. Given these important findings, we conclude that social bots were responsible for perpetrating the financial disinformation campaigns that promoted `OTCMKTS` low-value stocks by exploiting the popularity of high-value ones. In the remainder of this section we report on the general characteristics of the 18,509 users classified as financial social bots and we compare them to other bots and trolls previously studied in literature as well as to the 7,448 accounts classified as humans.



Evelina
Investing day trading. Lowery's law: if it jams
FOLLOWER 1
FOLLOWING 11

Karylin
Beach's law: no two identical parts are alike. Investing making money.
FOLLOWER 2
FOLLOWING 11

Kaitlin
I like the best.(i like myself...) Day trading making money.
FOLLOWER 1
FOLLOWING 16

Mellisa
Investing trading stocks. The plural of spouse is spice.
FOLLOWER 3
FOLLOWING 11

Fig. 7. Examples of a subset of users classified as bots. The accounts show similarities in their names, screen names, numbers of followers and followings, and in their description. Such similarities support the hypothesis that these accounts are part of large, organized botnets.

5.2 Profile Characteristics of Financial Bots

The *creation date* is an unforgeable characteristic of a social media account that has been frequently used to spot groups of coordinated malicious accounts (e.g., bots and trolls) [29]. Its usefulness lies in the impossibility to counterfeit or to masquerade it, combined with the fact that "masters" typically create their bot and troll armies in short time spans[7]. As a consequence, large numbers of accounts featuring almost identical creation dates might represent botnets or troll armies. Given this picture, the first characteristic of financial social bots that we analyze is the distribution of their account creation dates. The creation dates of the accounts in our dataset are distributed between 2007 and 2017. However, the majority of bots (53%) were created in 2017, as opposed to humans (12%). Figure 5 shows the distribution of creation dates of bots and humans in

[7] https://www.nytimes.com/interactive/2018/01/27/technology/social-media-bots.html.

2017, at a weekly granularity. Interestingly, bots display coordinated creation activities, while the creation of human accounts is more evenly distributed across the year. In detail, 45% of bots were created between February and April, with a particularly significant spike of 1,346 bots created on March 2. These findings further confirm the manufactured nature of the accounts classified as bots, and their pervasive presence in stock microblogs.

Table 2. Top-5 words and 3-grams used in account descriptions by bots and humans. Descriptions for humans are more heterogeneous and repetitions are less frequent.

Social bots				Humans			
Word	Freq.	3-gram	Freq.	Word	Freq.	3-gram	Freq.
Trading	8,138	Day trading making	848	Love	314	Follow follow back	7
Day	4,195	Trading making money	848	Life	209	Never say never	7
Money	4,173	Investing day trading	838	Follow	168	Always strive prosper	6
Stocks	4,056	Trading stocks investing	821	Music	164	Live life fullest	5
Trading	4,047	Investing trading stocks	814	Like	112	Stock market investor	4

Colluding groups of bots and trolls have also been associated to peculiar patterns in their *screen names* [21]. This is because they represent fictitious identities whose names and usernames are typically generated algorithmically. Looking for artificial patterns in the screen name, we first analyze the distribution of the screen name length. Interestingly, 50% of bots have a screen name length between 14 and 15 characters, while only 26% of humans share such characteristic. By examining the structure of suspiciously long bot screen names, we observe two main patterns. The first denotes the presence of screen names composed of a given name, followed by a family name. Such users also use the given name, which in almost all the cases is a female English name, as their display name. The second pattern exposes bots with a screen name composed of exactly 15 random alpha-numeric characters, accompanied by a given name as a display name. Such phenomenon has been observed before for numerous bot accounts involved in two different political-related events [4], and it's a strong confirmation of the malicious nature of our accounts labelled as bots. Figure 7 provides some examples of such bots. Moreover, by cross-checking information related to the creation dates, we observe that 11% of such bots are created on the same day.

Next, we inspect account *descriptions* (also known as *biographies*). We find a total of 575 users (3%) sharing the exact same description with at least 3 other users. In other words, there are 174 small groups of at least 3 users having in common the same description. Such repeated descriptions follow a specific pattern – in particular, they are composed of a famous quote or law, and of a set of financial keywords that are totally unrelated with the rest of the description. Interestingly, the use of famous quotes by bots to attract genuine users has already been documented before, for bots acting in the political domain [11].

We find 373 occurrences of such pattern, and none amongst the users with this pattern is classified as human. Some bot accounts exhibiting this characteristic are shown in Fig. 7. Table 2 summarizes the words and 3-grams mostly used in account descriptions by bots and humans. As shown, striking differences emerge. In summary, all previous findings support the hypothesis that users classified as bots did not act individually but that are rather part of large, organized and coordinated botnets.

Finally, we measure differences between bots and humans with respect to their *social relationships*. In particular, Fig. 6 shows differences in the distributions of followers and followings. Bots are characterized by a significantly lower number of both followers and followings, indicating accounts with few social relationships. It has been demonstrated that accounts with many social relationships in online social platforms are perceived as more trustworthy and credible [9]. Thus, to this regard, our financial bots appear as rather untrustworthy and simplistic accounts. Having few social connections also implies a difficulty in amplifying and propagating messages. In other words, only few users can read – and possibly re-share – what these bots post.

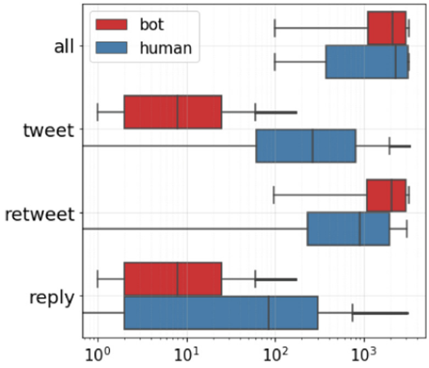


Fig. 8. Distribution of the number of tweets per user. Although bots and humans feature similar volumes of shared tweets, financial bots tend to retweet rather than to create original content.

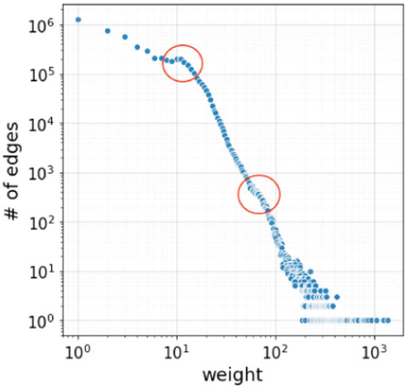


Fig. 9. Edge weights distribution in the user similarity network. The distribution approximates a power law with 2 notable exceptions, marked with red circles. (Color figure online)

5.3 Tweeting Characteristics of Financial Bots

Studying general profile characteristics, as we have done in the previous subsection, allows to assess the credibility and trustworthiness of financial bots (or lack thereof). Instead, in the remainder of this section we focus on their tweeting

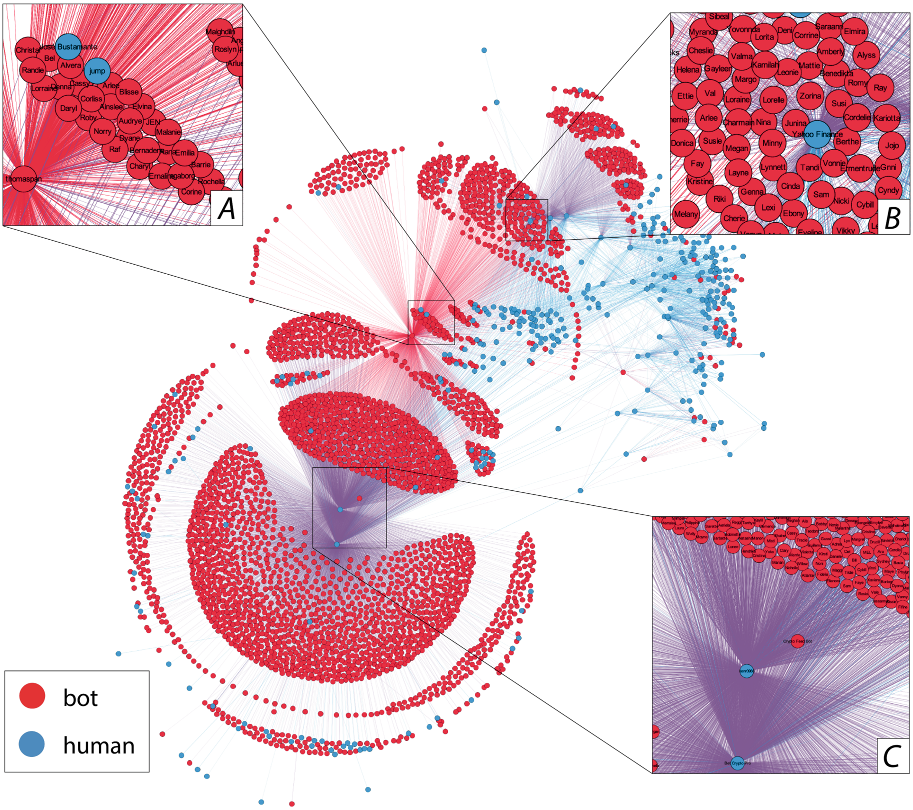


Fig. 10. A portion of the user similarity network. Nodes are colored according to their classification as bots or humans. Several different botnets are clearly visible as dense clusters. Bots are typically connected to a single human-labeled user, with which they share the majority of their mentioned companies.

activity. Our aim for the following analyses is to understand the likely target of financial bots as well as their inner organization.

We first analyze the *distribution of the number of tweets* posted by bots and humans, for each possible type of tweet (that is, original tweets, retweets and replies). As displayed in Fig. 8, financial bots and humans share a comparable total number of tweets. In other words, financial bots do not seem to post excessively (i.e., to spam), as other simplistic types of bot do [8], but instead they have an overall content production that is similar to that of humans. However, bots exhibit a strong preference for retweeting rather than for creating original content or for replying. Therefore, retweets are the primary mechanism used by financial bots to propagate content. It is worth noting however that the focus of these financial bots is likely posed on the retweet itself, rather than on retweets as an efficient mean to rapidly reach broader audiences [17,23]. This is because financial bots are characterized by few social relationships, as discussed in the previous section. As such, few users would be exposed to their retweets. This strategy, applied to the financial context, may nonetheless deceive trading algorithms listening to social conversations in search for hot stocks to invest in. As a consequence, synchronized mass-retweets of stock microblogs may contribute to artificially overstate the interest associated with specific stocks.

We conclude our analyses by studying the use of *cashtags* by bots and humans. Here, we are particularly interested in identifying groups of users that systematically tweet about the same stocks, because this might reveal the inner structure of financial disinformation botnets. Interesting questions are related as to whether we are witnessing to a single huge botnet or whether there are multiple botnets individually promoting different sets of stocks.

To answer these questions, we first build the bipartite network of users (comprising both bots and humans) and companies. In detail: Twitter users are one set of nodes, companies represent the other set of nodes, and a link connects a user to a company if that user mentioned that company in one of its tweets. This bipartite network is directed and weighted based on the number of times a user mentions given companies. In order to study similarities between groups of users, we then project our bipartite network onto the set of users. This process results in two users being linked to one another if they both mentioned at least one common company. The projected network, henceforth called *user similarity network*, is undirected and weighted. The weight of a link connecting two users measures the number of companies mentioned by both users.

For the sake of clarity, in the following we report results of the analysis of a subset of the user similarity network. In particular, Fig. 9 shows the distribution of edge weights in the considered portion of the network. As shown, the edge weights distribution approximates a power law, with 2 notable exceptions marked in figure with red circles. Peculiar patterns that deviate from the general law for specific portions of a network distribution have been previously associated with malicious activities [20]. For this reason, we focus subsequent analyses on the network nodes and edges that are responsible for the deviations highlighted in Fig. 9. In particular, Fig. 10 shows the resulting user similarity network, visualized via a force-directed layout, where nodes are colored according to their classification as bots or humans. Interestingly, the vast majority of nodes in this network were previously labeled as bots, during our bot detection step. This explains the deviations observed in the edge weights distribution plot. In addition, the vast majority of bots is organized in a few large distinct clusters. Each cluster of bots is typically connected to a single human-labeled user, with which bots share the majority of their mentioned companies. In other words, the visualization of Fig. 10 clearly allows to identify several distinct botnets, as well as the accounts that they are promoting. The few human-labeled users of the network show more diverse patterns of network connections. They are not organized in dense clusters and, in general, feature more heterogeneous connectivity patterns with respect to the bots, confirming previous literature results [12]. A few interesting portions of the network are magnified in the ***A***, ***B*** and ***C*** insets of Fig. 10, and allow to identify the users to which the botnets are connected (including the *@YahooFinance* account visible in inset ***B***), as well as the similar names (e.g., all English and female, as shown in insets ***A*** and ***B***) of the accounts that constitute the botnets.

6 Discussion

Results of our investigations highlighted the widespread existence of financial disinformation in Twitter. In particular, we documented a speculative campaign where many financially unimportant (low-cap) stocks are massively mentioned in tweets together with a few financially important (high-cap) ones. In previous work, this fraud was dubbed as *cashtag piggybacking*, since the low-value stocks are piggybacked "on top of the shoulders" of the high-value ones [14]. Considering the already demonstrated relation between social and financial importance [22], a possible outcome expected by perpetrators of this advertising practice is the increase in financial importance of the low-value stocks, by exploiting the popularity of high-value ones. To this regard, promising directions of future research involve assessing whether these kinds of malicious activity are correlated to, or can influence, stock prices fluctuations, the stock market's performance, or even the macroeconomic stability [14].

Analyses of suspicious users involved in financial discussion spikes, revealed that the speculative campaigns are perpetrated by large groups of coordinated social bots, organized in several distinct botnets. We showed that the financial bots involved in these manipulative activities present very simple accounts, with few details and social connections. Among the available details, many signs of fictitious information emerge, such as the suspicious profile descriptions where some financial keywords are mixed with other unrelated content. The simplistic characteristics of these bots, their relatively recent and bursty creation dates, and their limited number of social connections give the overall impression of untrustworthy accounts. The financial social bots discovered in our study have different characteristics with respect to the much more sophisticated social bots recently emerged in worldwide political discussions [8,11]. Financial social bots thus appear as a rather easy target for automatic detection and removal, as also confirmed by the large number of such bots that has already been banned by Twitter.

Based on these findings, we conclude that these bots should not pose a serious threat to human investors (e.g., noise traders) looking for fresh information on Twitter. However, the aim of financial bots could be that of fooling automatic trading algorithms. In fact, to the best of our knowledge, the majority of existing systems that feed on social information for predicting stock prices, do not perform filtering with regards to possibly fictitious content. As such, these systems could potentially be vulnerable to coordinated malicious practices such as that of *cashtag piggybacking*. The fact that no study nor existing system actually hunted financial bots before our present works, could also possibly explain the simplistic characteristics of these bots. In fact, it is largely demonstrated that recent social bots became so evolved and sophisticated as an evasion mechanism for the plethora of existing bot detection techniques [8]. In other words, financial bots could be this simple, just because nobody ever hunted them. If this proves to be the case however, we should expect financial bots to become much more sophisticated in the near future. A scenario that would pose a heavier burden on our side with regards to their detection and removal.

The user-centric classification approach that we adopted in this study demands the availability and the analysis of large amounts of data, and requires intensive and time-consuming computations. This is because, in order to assess the veracity of a discussion spike, all users involved in that discussion are to be analyzed. This could easily imply the analysis of tens of thousands of accounts for evaluating a single spike of discussion. On the contrary, another – more favorable – scenario could involve the classification of the discussion spikes themselves. In other words, future financial spam detection systems could analyze high-level characteristics of discussion spikes (e.g., their burstiness, the number of distinct accounts that participate, market information of the discussed stocks, etc.), with the goal of promptly detecting promoted, fictitious, or made up discussions. This approach, previously applied to other scenarios [28], is however still unexplored in the online financial domain. As such, it represents another promising avenue of future research and experimentation.

7 Conclusions

Our work investigated the presence and the characteristics of financial disinformation in Twitter. We documented a speculative practice aimed at promoting low-value stocks, mainly from the `OTCMKTS` financial market, by exploiting the popularity of high-value (e.g., `NASDAQ`) ones. An in-depth analysis of the accounts involved in this practice revealed that 71% of them are bots. Moreover, 48% of the accounts classified as bots have been subsequently banned by Twitter. Finally, bots involved in financial disinformation turned out to be rather simplistic and untrustworthy, in contrast with recent political bots that are much more sophisticated.

Our findings about the characteristics of fake financial discussion spikes as well as those related to the characteristics of financial bots, could be leveraged in the future as features for designing novel financial spam filtering systems. Hence, this work lays the foundations for the development of specific – yet still unavailable – methods to detect online financial disinformation, before it harms the pockets of unaware investors.

Acknowledgements. This work is partially supported by the European Community's H2020 Program under the scheme `INFRAIA-1-2014-2015: Research Infrastructures`, grant agreement #654024 *SoBigData: Social Mining and Big Data Ecosystem* and the scheme `INFRAIA-01-2018-2019: Research and Innovation action`, grant agreement #871042 *SoBigData++: European Integrated Infrastructure* *for Social Mining and Big Data Analytics*, and by the Italian Ministry of Education and Research in the framework of the *CrossLab Project* (Cloud Computing, Big Data & Cybersecurity), Departments of Excellence.

References

1. Albadi, N., Kurdi, M., Mishra, S.: Hateful people or hateful bots? Detection and characterization of bots spreading religious hatred in Arabic social media. In: Proceedings of the ACM on Human-Computer Interaction (HCI), vol. 3, no. CSCW, pp. 1–25 (2019)
2. Allem, J.P., Ferrara, E.: Could social bots pose a threat to public health? Am. J. Public Health (AJPH) **108**(8), 1005 (2018)
3. Berger, J.M., Morgan, J.: The ISIS Twitter census: defining and describing the population of ISIS supporters on Twitter. The Brookings Project on US Relations with the Islamic World, vol. 3, no. 20, pp. 1–4 (2015)
4. Beskow, D.M., Carley, K.M.: Its all in a name: detecting and labeling bots by their name. Comput. Math. Organ. Theory (CMOT) **25**(1), 24–35 (2019). https://doi.org/10.1007/s10588-018-09290-1
5. Bessi, A., Ferrara, E.: Social bots distort the 2016 U.S. Presidential election online discussion. First Monday **21** (2016). https://doi.org/10.5210/fm.v21i11.7090
6. Bollen, J., Mao, H., Zeng, X.: Twitter mood predicts the stock market. J. Comput. Sci. (JCS) **2**(1), 1–8 (2011)
7. Brachten, F., Stieglitz, S., Hofeditz, L., Kloppenborg, K., Reimann, A.: Strategies and influence of social bots in a 2017 German state election - a case study on Twitter. In: The 28th Australasian Conference on Information Systems (ACIS 2017) (2017)
8. Cresci, S.: Detecting malicious social bots: story of a never-ending clash. In: Grimme, C., Preuss, M., Takes, F.W., Waldherr, A. (eds.) MISDOOM 2019. LNCS, vol. 12021, pp. 77–88. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-39627-5_7
9. Cresci, S., Di Pietro, R., Petrocchi, M., Spognardi, A., Tesconi, M.: Fame for sale: efficient detection of fake Twitter followers. Decis. Support Syst. (DSS) **80**, 56–71 (2015)
10. Cresci, S., Di Pietro, R., Petrocchi, M., Spognardi, A., Tesconi, M.: Social fingerprinting: detection of spambot groups through DNA-inspired behavioral modeling. IEEE Trans. Dependable Secure Comput. (TDSC) **15**(4), 561–576 (2017)
11. Cresci, S., Di Pietro, R., Petrocchi, M., Spognardi, A., Tesconi, M.: The paradigm-shift of social spambots: evidence, theories, and tools for the arms race. In: The 26th International Conference on World Wide Web Companion (WWW 2017 Companion), pp. 963–972 (2017). IW3C2
12. Cresci, S., Di Pietro, R., Petrocchi, M., Spognardi, A., Tesconi, M.: Emergent properties, models and laws of behavioral similarities within groups of Twitter users. Comput. Commun. **150**, 47–61 (2020)
13. Cresci, S., Lillo, F., Regoli, D., Tardelli, S., Tesconi, M.: $FAKE: evidence of spam and bot activity in stock microblogs on Twitter. In: The 12th International AAAI Conference on Web and Social Media (ICWSM 2018). AAAI (2018)
14. Cresci, S., Lillo, F., Regoli, D., Tardelli, S., Tesconi, M.: Cashtag piggybacking: uncovering spam and bot activity in stock microblogs on Twitter. ACM Trans. Web (TWEB) **13**(2), 11:1–11:27 (2019)

15. Cresci, S., Minutoli, S., Nizzoli, L., Tardelli, S., Tesconi, M.: Enriching digital libraries with crowdsensed data. In: Manghi, P., Candela, L., Silvello, G. (eds.) IRCDL 2019. CCIS, vol. 988, pp. 144–158. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11226-4_12
16. Ferrara, E.: Disinformation and social bot operations in the run up to the 2017 French Presidential election. First Monday **22**(8) (2017). https://doi.org/10.5210/fm.v22i8.8005
17. Giatsoglou, M., Chatzakou, D., Shah, N., Faloutsos, C., Vakali, A.: Retweeting activity on Twitter: signs of deception. In: Cao, T., Lim, E.-P., Zhou, Z.-H., Ho, T.-B., Cheung, D., Motoda, H. (eds.) PAKDD 2015. LNCS (LNAI), vol. 9077, pp. 122–134. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-18038-0_10
18. Hentschel, M., Alonso, O.: Follow the money: a study of cashtags on Twitter. First Monday **19**(8) (2014). https://doi.org/10.5210/fm.v19i8.5385
19. Hwang, T., Pearce, I., Nanis, M.: Socialbots: voices from the fronts. Interactions **19**(2), 38–45 (2012)
20. Jiang, M., Cui, P., Beutel, A., Faloutsos, C., Yang, S.: CatchSync: catching synchronized behavior in large directed graphs. In: The 20th SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD 2014), pp. 941–950. ACM (2014)
21. Lee, S., Kim, J.: Early filtering of ephemeral malicious accounts on Twitter. Comput. Commun. **54**, 48–57 (2014)
22. Mao, Y., Wei, W., Wang, B., Liu, B.: Correlating S&P 500 stocks with Twitter data. In: The 1st International Workshop on Hot Topics on Interdisciplinary Social Networks Research (SIGKDD 2012 Workshops), pp. 69–72. ACM (2012)
23. Mazza, M., Cresci, S., Avvenuti, M., Quattrociocchi, W., Tesconi, M.: RTbust: exploiting temporal patterns for botnet detection on Twitter. In: The 11th International ACM Web Science Conference (WebSci 2019), pp. 183–192. ACM (2019)
24. Nizzoli, L., Tardelli, S., Avvenuti, M., Cresci, S., Tesconi, M., Ferrara, E.: Charting the landscape of online cryptocurrency manipulation. arXiv preprint arXiv:2001.10289 (2020)
25. Shao, C., Ciampaglia, G.L., Varol, O., Yang, K.C., Flammini, A., Menczer, F.: The spread of low-credibility content by social bots. Nat. Commun. **9**(1), 4787 (2018)
26. Stella, M., Ferrara, E., De Domenico, M.: Bots increase exposure to negative and inflammatory content in online social systems. Proc. Natl. Acad. Sci. **115**(49), 12435–12440 (2018)
27. Varol, O., Ferrara, E., Davis, C.A., Menczer, F., Flammini, A.: Online human-bot interactions: detection, estimation, and characterization. In: The 11th International AAAI Conference on Web and Social Media (ICWSM 2017). AAAI (2017)
28. Varol, O., Ferrara, E., Menczer, F., Flammini, A.: Early detection of promoted campaigns on social media. EPJ Data Sci. **6**(1), 1–19 (2017). https://doi.org/10.1140/epjds/s13688-017-0111-y
29. Viswanath, B., et al.: Strength in numbers: robust tamper detection in crowd computations. In: The 2015 ACM Conference on Online Social Networks (COSN 2015), pp. 113–124. ACM (2015)
30. Zannettou, S., Caulfield, T., De Cristofaro, E., Sirivianos, M., Stringhini, G., Blackburn, J.: Disinformation warfare: understanding state-sponsored trolls on Twitter and their influence on the Web. In: The 2019 World Wide Web Conference Companion (WWW 2019 Companion), pp. 218–226 (2019). IW3C2
31. Zannettou, S., Caulfield, T., Setzer, W., Sirivianos, M., Stringhini, G., Blackburn, J.: Who let the trolls out? Towards understanding state-sponsored trolls. In: The 11th International ACM Web Science Conference (WebSci 2019), pp. 353–362. ACM (2019)