Proceedings of the Twelfth International AAAI Conference on Web and Social Media (ICWSM 2018)

$FAKE: Evidence of Spam and Bot Activity in Stock Microblogs on Twitter

Stefano Cresci,[1] Fabrizio Lillo,[2] Daniele Regoli,[3] Serena Tardelli,[1] Maurizio Tesconi[1]
[1]Institute of Informatics and Telematics, IIT-CNR, Italy [name.surname]@iit.cnr.it
[2]Department of Mathematics, University of Bologna, Italy fabrizio.lillo@unibo.it
[3]Scuola Normale Superiore, Italy daniele.regoli@sns.it

Abstract

Microblogs are increasingly exploited for predicting prices and traded volumes of stocks in financial markets. However, it has been demonstrated that much of the content shared in microblogging platforms is created and publicized by bots and spammers. Yet, the presence (or lack thereof) and the impact of fake stock microblogs has never systematically been investigated before. Here, we study 9M tweets related to stocks of the 5 main financial markets in the US. By comparing tweets with financial data from Google Finance, we highlight important characteristics of Twitter stock microblogs. More importantly, we uncover a malicious practice perpetrated by coordinated groups of bots and likely aimed at promoting low-value stocks by exploiting the popularity of high-value ones. Our results call for the adoption of spam and bot detection techniques in all studies and applications that exploit user-generated content for predicting the stock market.

Introduction

The exploitation of user-generated content in microblogs for the prediction of real-world phenomena, has recently gained huge momentum. An important application domain for such approach is that of finance, and in particular, stock market prediction. Indeed, a number of works developed algorithms and tools for extracting valuable information from microblogs and proved capable of predicting prices and traded volumes of stocks in financial markets (Bollen, Mao, and Zeng 2011). Meanwhile, evidence of spam and automated (bot) activities in social platforms is being reported at a growing rate (Ferrara et al. 2016b). The existence of fictitious, synthetic content appears to be pervasive since it has been witnessed both in online discussions about important societal topics (e.g., politics, terrorism, immigration), as well as in discussions about seemingly less relevant topics, such as products on sale on e-commerce platforms, and mobile applications (Cresci et al. 2017a).

Thus, on the one hand, user-generated content in microblogs is being exploited for predicting trends in the stock market. On the other hand, without a thorough investigation, we run the risk that much of the content we rely on, is actually fake and possibly purposely created to mislead algorithms and users alike. Should this risk materialize, real-world consequences would be severe. This study moves in the direction of investigating the presence of spam and bot activity in stock microblogs, thus paving the way for the development of intelligent financial-spam filtering techniques. Specifically, we first collect a rich dataset comprising 9M tweets posted between May and September 2017, discussing stocks of the 5 main financial markets in the US. We enrich our dataset by collecting financial information from Google Finance about the 30,032 companies mentioned in our tweets. Cross-checking discussion patterns on Twitter against official data from Google Finance uncovers anomalies in tweets related to some low-value companies. Further investigation of this issue reveals a large-scale speculative campaign perpetrated by coordinated groups of bots and aimed at promoting low-value stocks by exploiting the popularity of high-value ones.

Related work

Finance and social media Starting from the general assumption that user-generated messages about a company's future prospects provide a rich and diverse source of information, much effort has been devoted towards the detection of correlations between metrics extracted from social media posts and stock market prices. In particular, *sentiment* metrics have been widely used as a predictor for stock prices and other economic indicators (Gilbert and Karahalios 2010; Bollen, Mao, and Pepe 2011; Sprenger 2011). Others have instead proposed to exploit the overall volume of tweets about a company (Mao et al. 2012) and the topology of stock networks (Ruiz et al. 2012) as predictors of financial performance. However, subsequent work (Zheludev, Smith, and Aste 2014) evaluated the informativeness of sentiment- and volume-derived predictors, showing that the sentiment of tweets contains significantly more information for predicting stock prices than just their volume. The role of *influencers* in social media has also been identified as a strong contributing factor to the formation of market trends (Cazzoli et al. 2016). Others have instead used weblogs for studying the relationships between different companies (Kharratzadeh and Coates 2012). In detail, co-occurrences of stock mentions in weblogs have been exploited to create a graph of companies, which was subsequently clustered. Authors have verified that companies belonging to the same clusters feature strong correlations in their stock prices. This

markets	financial data		twitter data	
	companies	*median capitalization*	*users*	*tweets*
NASDAQ	3,013	365,780,000	252,587	4,017,158
NYSE	2,997	1,810,000,000	265,618	4,410,201
NYSEARCA	726	245,375,000	56,101	298,445
NYSEMKT	340	78,705,000	22,614	196,545
OTCMKTS	22,956	31,480,000	64,628	584,169

Table 1: Overall statistics about our dataset.

methodology can be employed for market prediction and as a portfolio-selection method, which has been shown to outperform traditional strategies based on company sectors or historical stock prices. Nowadays, results of studies such as those briefly surveyed in this section are leveraged for the development of automatic trading systems that are largely fed with social media-derived information (Feldman 2013). As a consequence, such automatic systems can potentially suffer severe problems caused by large quantities of fictitious posts.

Spam and bots in social media Many developers of spammer accounts make use of bots in order to simultaneously and continuously post a great deal of spam content. This is one of the reasons why, despite bots being in rather small numbers when compared to legitimate users, they nonetheless have a profound impact on content popularity and activity in social media (Gilani, Farahbakhsh, and Crowcroft 2017). In addition, bots are driven so as to act in a coordinated and synchronized way, thus amplifying their effects (Ratkiewicz et al. 2011). Another problem with bots is that they *evolve* over time, in order to evade established detection techniques (Cresci et al. 2017a). Recently, a general-purpose overview of the landscape of automated accounts was presented in (Ferrara et al. 2016a). This work testifies the emergence of a new wave of social bots, capable of mimicking human behavior and interaction patterns in social media better than ever before. A subsequent study (Cresci et al. 2017a) compared "traditional" and "evolved" bots in Twitter, and demonstrated that the latter are almost completely undetected by platform administrators, by users themselves, and even by the majority of state-of-the-art bot detection techniques. The combination of automatic systems feeding on social media data and the pervasive presence of spam and bots, motivates our investigation on the presence of spam and bots in stock microblogs.

Dataset

Our dataset for this study is composed of: (i) stock microblogs collected from Twitter, and (ii) financial information collected from Google Finance.

Twitter data collection Twitter users follow the convention of tagging stock microblogs with so-called *cashtags*. The cashtag of a company is composed of a dollar sign followed by its ticker symbol (e.g., `$AAPL` is the cashtag of *Apple, Inc.*). Similarly to hashtags, cashtags can be used as an efficient mean to filter content on Twitter and to collect data about given companies (Hentschel and Alonso 2014). For this reason, we based our Twitter data collection on an official list of cashtags. Specifically, we first downloaded a list of 6,689 stocks traded on the most important US markets (e.g., NASDAQ, NYSE) from the official NASDAQ Web site[1]. Then, we collected all tweets shared between May and September 2017, containing at least one cashtag from the list. Data collection from Twitter has been carried out by exploiting Twitter's Streaming APIs. After our 5 months data collection, we ended up with ∼9M tweets (of which 22% are retweets), posted by ∼2.5M distinct users, as shown in Table 1. As a consequence of our data collection strategy, every tweet in our dataset contains at least one cashtag from the starting list. However, many collected tweets contain more than one cahstag, many of which are related to companies not included in our starting list. Overall we collected data about 30,032 companies traded across 5 different markets.

Financial data collection We enriched our Twitter dataset by collecting financial information about each of the 30,032 companies found in our tweets. Financial information have been collected from public company data hosted on the Google Finance Web site[2]. Among collected financial information, is the *market capitalization* (market cap) of a company and its *industrial classification*. The capitalization is the total dollar market value of a company. For a given company i, it is computed as the share price $P(s_i)$ times the number of outstanding shares $|s_i|$: $C_i = P(s_i) \times |s_i|$. In our study, we take the market cap of a company into account, since it allows us to compare the financial value of that company with its social media popularity and engagement. In Table 1 we report the median capitalization of the companies for each considered market. As shown, important markets such as NYSE and NASDAQ trade, on average, stocks with higher capitalization than those traded in minor markets. Industrial classification is expressed via the Thomson Reuters Business Classification[3] (TRBC). TRBC is a 5-level hierarchical sector and industry classification, widely used in the financial domain for computing sector-specific indices. In our study, we compare companies belonging to the same category, across all 5 levels of TRBC.

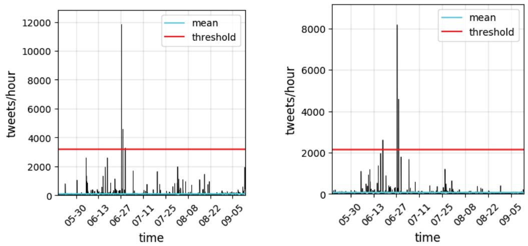
Analysis of stock microblogs

Dataset overview Surprisingly, the vast majority (76%) of companies mentioned in our dataset do not belong to the NASDAQ list and are traded in OTCMKTS, as shown in Table 1. Having so many OTCMKTS companies in our dataset is already an interesting finding, considering that our data collection grounded on a list of high-capitalization (high-cap) companies. OTCMKTS is a US financial market for over-the-counter transactions, thus with far less stringent requirements than those needed from NASDAQ, NYSE, NYSEARCA, and NYSEMKT. For this reason, many small companies opt to be traded in OTCMKTS instead of the more requiring markets. Thus, from a company viewpoint, our dataset is dominated by OTCMKTS. However, OTCMKTS companies play

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

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

[3]https://financial.thomsonreuters.com/en/products/data-analytics/market-data/indices/trbc-indices.html



(a) `$AAPL` (*Apple, Inc.*). (b) `$FB` (*Facebook, Inc.*).

Figure 1: Examples of stock time series.

a marginal role from both a financial and social viewpoint, having low capitalization and small numbers of tweets, the vast majority of which are retweets. In contrast, companies from `NASDAQ` and `NYSE` have high capitalization and are mentioned in many tweets, with low percentage of retweets.

Stock time series analysis In order to uncover possible malicious behaviors related to stock microblogs, we build and analyze the hourly time series of each of the 6,689 stocks downloaded from the `NASDAQ` Web site. Given a stock i, its time series is defined as $\mathbf{s}_i = (s_{i,1}, s_{i,2}, \ldots, s_{i,N})$, with $s_{i,j}$ being the number of tweets that mentioned the stock i during the hour j. Figure 1 shows some examples of our stock time series. As shown in figure, stock time series are characterized by long time spans over which tweet discussion volumes remain rather low, occasionally interspersed by large discussion spikes. To give a better characterization of this phenomenon we ran a simple anomaly detection technique on all the 6,689 time series. Our anomaly detection technique is designed so as to detect a peak $p_{i,j}$ in a time series $\mathbf{s}_i$ *iff* the tweet volume for the hour j deviates from the mean tweet volume $\bar{\mathbf{s}}_i$ by a number $K = 10$ of standard deviations. Next, we are interested in analyzing the tweets that generated the peaks (henceforth, *peak tweets*: $t \in \mathbf{t}$). We find out that, on average, 60% of tweets $t \in \mathbf{t}$ are retweets. In other words, the peaks identified by our anomaly detection technique are largely composed of retweets. In addition, considering that our time series have hourly granularity, those retweets also occurred within a rather limited time span, in a *bursty* fashion. This finding is particularly interesting also considering that in all our dataset, we had only 22% retweets, versus 60% measured for peak tweets.

Co-occurence analysis We also analyzed tweets $t \in \mathbf{t}$ by considering the co-occurrences of stocks. From this analysis we see that tweets $t \in \mathbf{t}$ typically contain many more cashtags than tweets $t \notin \mathbf{t}$. Indeed, the mean number of cashtags per tweet is 6 for $t \in \mathbf{t}$, versus 2 for the whole dataset. The cashtags that co-occur in peak tweets seem unrelated, and the authors of those tweets don't provide further information to explain such co-occurrences. As an example, Figure 2 shows 4 of such suspicious tweets. In figure, in every tweet, a few cashtags of high-capitalization (high-cap) stocks co-occur with many cashtags of low-cap stocks. Furthermore, in order to assess the real-world relatedness of

Figure 2: Examples of suspicious peak tweets.

stocks, we evaluated the extent to which co-occurring stocks belong to the same (or to different) TRBC class(es), for all the 5 hierarchical levels of TRBC. As a measurement for the difference in TRBC classes across stocks in a tweet, we leveraged the notion of *entropy*. Results show that the normalized median entropy ~ 1 for all 5 TRBC levels, meaning that co-occurring stocks in peak tweets are almost unrelated. The results of this experiment seem to suggest that, overall, co-occurrences of stocks in peak tweets are not motivated by the fact that stocks belong to the same industrial or economic sectors.

Since real-world relatedness (as expressed by industrial classification) is not a plausible explanation for co-occurring stocks, we now turn our attention to market capitalization. For instance, legitimate peak tweets could mention multiple stocks with similar capitalization. Conversely, malicious users could try to exploit the popularity of high-cap stocks by mentioning them together with low-cap ones. One way to evaluate the similarity (or dissimilarity) in market capitalization of co-occurring stocks is by computing statistical measures of spread, such as the *standard deviation* (std.). Thus, for each peak tweet $t \in \mathbf{t}$ we computed the std. of the capitalization of all companies mentioned in t. Results highlight a large empiric std. between the capitalization of co-occurring companies. This means that in our peak tweets, high-cap companies co-occur with low-cap ones. Moreover, the measured std. is larger than that obtained with a random baseline (i.e., a bootstrap). In turn, this means that the large difference in capitalization can not be explained by the intrinsic characteristics of our dataset, but it is rather the consequence of an external action.

Analysis of suspicious users

In previous sections we identified a wide array of suspicious phenomena related to stock microblogs. In this section we employ a state-of-the-art bot and spam detection system – specifically developed for spotting malicious group activities – to classify suspicious users (Cresci et al. 2016; 2017b). The goal of this experiment is to assess whether users that shared/retweeted the suspicious peak tweets we previously identified, are classified as bots. In turn, this would bring substantial evidence of bot activities in the stock microblogs that we analyzed.

Because of the computationally intensive analyses performed by (Cresci et al. 2016; 2017b), we constrained this experiment to the 100 largest peaks (i.e., those generated by the greatest number of tweets) of our dataset. Starting from those top-100 peaks, we then analyzed the $25,988$ distinct users that shared or retweeted at least one peak tweet. Data needed by the detection system to perform user classification have been collected by crawling the Twitter timelines of such $25,988$ users. Notably, the bot detection system classified as much as 71% $(18,509)$ of the analyzed users as bots. A manual analysis of a subset of bots allowed to identify characteristics shared between all the users (e.g., similar name, join date, profile picture, etc.), supporting the hypothesis that they are part of a larger botnet. Users classified as bots also feature very high retweet rates (ratio of retweets over all posted tweets), thus explaining the large number of retweets in our peaks and among `OTCMKTS` stock microblogs. We obtained these results by analyzing only the 100 largest detected peaks, therefore analyses of minor peaks might yield different results. Nonetheless, the overwhelming ratio of bots that we discovered among large peaks discussing popular stocks, raises serious concerns over the reliability of stock microblogs.

Conclusions

Motivated by the widespread presence of social bots, we carried out the first large-scale, systematic analysis on the presence and impact of spam and bot activity in stock microblogs. By cross-checking 9M stock microblogs from Twitter with financial information from Google Finance, we uncovered a malicious practice aimed at promoting low-value stocks by exploiting the popularity of high-value ones. In detail, many stocks with low market capitalization, mainly traded in `OTCMKTS`, are mentioned in microblogs together with a few high capitalization stocks traded in `NASDAQ` and `NYSE`. We showed that such co-occurring stocks are not related by economic and industrial sector. Moreover, the large discussion spikes about low-value stocks are due to mass, synchronized retweets. Finally, an analysis of retweeting users classified 71% of them as bots. Given the severe consequences that this new form of financial spam could have on unaware investors as well as on automatic trading systems, our results call for the prompt adoption of spam and bot detection techniques in all applications and systems that exploit stock microblogs.

Acknowledgments

This research is supported in part by the EU H2020 Program under the schemes INFRAIA-1-2014-2015: Research Infrastructures grant agreement #654024 *SoBigData: Social Mining & Big Data Ecosystem*.

References

Bollen, J.; Mao, H.; and Pepe, A. 2011. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In *ICWSM'11*. AAAI.

Bollen, J.; Mao, H.; and Zeng, X. 2011. Twitter mood predicts the stock market. *Journal of computational science*.

Cazzoli, L.; Sharma, R.; Treccani, M.; and Lillo, F. 2016. A large scale study to understand the relation between Twitter and financial market. In *ENIC'16*. IEEE.

Cresci, S.; Di Pietro, R.; Petrocchi, M.; Spognardi, A.; and Tesconi, M. 2016. DNA-inspired online behavioral modeling and its application to spambot detection. *IEEE Intelligent Systems*.

Cresci, S.; Di Pietro, R.; Petrocchi, M.; Spognardi, A.; and Tesconi, M. 2017a. The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race. In *WWW'17 Companion*. ACM.

Cresci, S.; Di Pietro, R.; Petrocchi, M.; Spognardi, A.; and Tesconi, M. 2017b. Social fingerprinting: detection of spambot groups through DNA-inspired behavioral modeling. *IEEE Transactions on Dependable and Secure Computing*.

Feldman, R. 2013. Techniques and applications for sentiment analysis. *Communications of the ACM*.

Ferrara, E.; Varol, O.; Davis, C.; Menczer, F.; and Flammini, A. 2016a. The rise of social bots. *Communications of the ACM*.

Ferrara, E.; Varol, O.; Menczer, F.; and Flammini, A. 2016b. Detection of promoted social media campaigns. In *ICWSM'16*. AAAI.

Gilani, Z.; Farahbakhsh, R.; and Crowcroft, J. 2017. Do bots impact Twitter activity? In *WWW'17 Companion*. ACM.

Gilbert, E., and Karahalios, K. 2010. Widespread worry and the stock market. In *ICWS'10*. AAAI.

Hentschel, M., and Alonso, O. 2014. Follow the money: A study of cashtags on Twitter. *First Monday*.

Kharratzadeh, M., and Coates, M. 2012. Weblog analysis for predicting correlations in stock price evolutions. In *ICWSM'12*. AAAI.

Mao, Y.; Wei, W.; Wang, B.; and Liu, B. 2012. Correlating S&P 500 stocks with Twitter data. In *KDD'12 Workshops*. ACM.

Ratkiewicz, J.; Conover, M.; Meiss, M. R.; Gonçalves, B.; Flammini, A.; and Menczer, F. 2011. Detecting and tracking political abuse in social media. In *ICWSM'11*. AAAI.

Ruiz, E. J.; Hristidis, V.; Castillo, C.; Gionis, A.; and Jaimes, A. 2012. Correlating financial time series with micro-blogging activity. In *WSDM'12*. ACM.

Sprenger, T. O. 2011. TweetTrader.net: Leveraging crowd wisdom in a stock microblogging forum. In *ICWSM'11*. AAAI.

Zheludev, I.; Smith, R.; and Aste, T. 2014. When can social media lead financial markets? *Scientific reports*.