Fraudulent Social Media Activities and Financial Markets

Abstract

This is the first study to examine the implications of fraudulent social media activities on financial markets, where fraudulent accounts are employed to inflate the spread of corporate-related information and manipulate public opinion. To identify and categorize fraudulent social media activities, we rely on the complete tweet histories of over 6 million randomly sampled Twitter accounts with over 2 billion tweets from 2009 to 2014, together with the tweet histories of approximately 2 million labeled fraudulent accounts. Then we investigate the potential factors that may contribute to the initiation of fraudulent social media activities and study how investors, shareholders and companies are affected from various aspects.

Introduction

Social media has become one of the most popular channels for people to disseminate information, share ideas, and express opinions. With the proliferation of social media, its relevance to financial markets has been recognized and the literature provides evidence that social media has significant implications on stock markets (e.g., Chen et al., 2014; Bartov et al., 2017).

However, along with the growing importance of social media are rising fraudulent activities on unregulated social media platforms such as Twitter and Facebook, where a large number of fraudulent accounts are inflating the spread of information and manipulating public opinion. Even though it has not been studied in the finance literature yet, anecdotal evidence suggests that fraudulent social media activities may significantly impact the financial market. In January 2013, a false tweet about the company Audience being investigated by the DOJ was suspiciously retweeted many times and spread widely on Twitter, leading to a sudden drop of 28 percent in its stock price over two days. The Securities and Exchange Commission seems to be aware of this issue and is considering monitoring and flagging posts on social media that could have an effect on financial markets.

Research Objective

We aim to examine fraudulent social media activities that are related to companies and study their consequent implications on financial markets. We use Twitter as a laboratory to study fraudulent social media activities due to the publicity of fraudulent behavior on this platform and the availability of large-scale and trackable user-level data.

First, we will identify fraudulent activities on Twitter related to companies and categorize the type of each activity. The target of fraudulent activities may be a newly-launched product of the company, an investment policy, a specific corporate event such as M&A, or simply the company's stock price. There may be many potential initiators of fraudulent activity on social media. For example, a firm may initiate a fraudulent campaign to promote itself when the firm is under substantial external pressure. Moreover, market competition may also be related as fierce competition between major market players may induce some to launch social media attacks on their competitors. In addition, other market players such as blockholders and hedge funds may take advantage of social media for the purpose of achieving their own interests. Only after having identified a large sample of fraudulent social media activities can we better understand them.

Second, we will examine the financial market implications of these fraudulent social media activities. We start by investigating the reaction of other users on social media and studying how people's attention and opinions about the involved company may be affected. Then we will examine investors' information acquisition behavior and trading behavior related to the involved company and their consequent impact on the stock market. Moreover, following fraudulent social media activities that target corporate events, we will analyze shareholders' behavior regarding these events such as voting on M&As. To summarize, we aim to systematically examine how investors, shareholders and firms will be affected by fraudulent activities in social media, which is of great significance to different aspects of financial markets.

Data

Typically, data presented by other related research on Twitter only consists of a specific type of tweets, such as tweets with "cashtags" where stock symbols are preceded by the dollar sign, or all the tweets sent by official corporate twitter accounts. In contrast, our data consists of the complete tweet histories of over 6 million randomly sampled Twitter accounts from 2009 to 2014, equating to about 1.5-2% of all English-speaking Twitter accounts. Only with the full historical tweets of an account can we identify fraudulent accounts and associated activities. In total we have collected over 2 billion tweets from these accounts, and the total value of this data according to Twitter's API offering is just over $3 million USD. The storage requirements of this dataset is approximately 5 terabytes. Beyond the randomly sampled Twitter accounts, we have also collected the tweet histories of approximately 2 million fraudulent accounts, allowing us to develop and test fraudulent account detection algorithms on a large scale.

Computing Language and Research Strategy

The computing language that we will use is Python. We will first develop and test fraudulent account detection algorithms on large-scale known sets of fraudulent accounts. Echeverria et al. (2018) showed that existing supervised fraudulent account classification approaches do not generalize to unseen account networks, suggesting that alternative approaches are required to identify the existence of fraudulent accounts. Unsupervised classification techniques such as cross-correlating account activities (Chavoshi et al. 2016) are promising for detecting fraudulent accounts but requires massive computation power. We expect that by developing smarter algorithms in combination with the use of high-performance computing it becomes possible to identify fraudulent accounts and their associated fraudulent activities. Subsequently we will deploy these algorithms on accounts that have taken part in positive and negative campaigns about companies to identify the prevalence of fraudulent activities within these campaigns.

Reference

Bartov, E., Faurel, L., & Mohanram, P. S. (2017). Can twitter help predict firm-level earnings and stock returns?. *The Accounting Review*, *93*(3), 25-57.

Chavoshi, N., Hamooni, H., & Mueen, A. (2016). Identifying correlated bots in twitter. In International Conference on Social Informatics (pp. 14-21). Springer, Cham.

Chen, H., De, P., Hu, Y. J., & Hwang, B. H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *The Review of Financial Studies*, *27*(5), 1367-1403.

De Cristofaro, E., Kourtellis, N., Leontiadis, I., Stringhini, G., & Zhou, S. (2018). LOBO: Evaluation of Generalization Deficiencies in Twitter Bot Classifiers. In Proceedings of the 34th Annual Computer Security Applications Conference (pp. 137-146). ACM.