

Analyzing the Hidden Causal Interactions in Large-Scale Social Networks: A Case Study on GameStop

Gang Zhou

*Institute of Automation, Chinese Academy of Sciences
School of Artificial Intelligence,
Chinese Academy of Sciences
Beijing, China
zhougang2020@ia.ac.cn*

Rongrong Jing

*Institute of Automation, Chinese Academy of Sciences
School of Artificial Intelligence,
Chinese Academy of Sciences
Beijing, China
jingrongrong2019@ia.ac.cn*

Xiaolong Zheng*

*Institute of Automation, Chinese Academy of Sciences
School of Artificial Intelligence,
Chinese Academy of Sciences
Beijing, China
xiaolong.zheng@ia.ac.cn*

Xingwei Zhang

*Institute of Automation, Chinese Academy of Sciences
School of Artificial Intelligence,
Chinese Academy of Sciences
Beijing, China
zhangxingwei2019@ia.ac.cn*

Hu Tian

*Institute of Automation, Chinese Academy of Sciences
School of Artificial Intelligence,
Chinese Academy of Sciences
Beijing, China
tianhu2018@ia.ac.cn*

Daniel Zeng

*Institute of Automation, Chinese Academy of Sciences
School of Artificial Intelligence,
Chinese Academy of Sciences
Beijing, China
dajun.zeng@ia.ac.cn*

Abstract—A war between individual investors and Wall Street institutions that occurred in January 2021 has attracted attention worldwide. There were many participants in this event, and the causality behind that was also complex. Due to the COVID-19, people work from home, so social media has become the main battlefield of GameStop events, especially Reddit's subreddit: r/WallStreetBets ("WSB" for short). Analyzing the causal interaction in social media helps people gain a more comprehensive and profound understanding of this online Wall Street movement. We use the WSB data set to construct causal networks and demonstrate the evolution of causality between GameStop stock price and WSB Redditors sentiment. We further analyzed the causality between GameStop stock price, WSB, and cryptocurrency. The discovery convinced us that Redditor in WSB did dominate and promote this movement and extended the battle to cryptocurrency market.

Keywords—gamestop, causality, cryptocurrency

I. INTRODUCTION

In the past decades, social media platforms have had a profound impact on global financial markets. The latest GameStop episode is an excellent case of this information revolution [1]. Some studies have shown that users can lead their friends behave similarly through their social influence, either explicitly or implicitly [2]. Moreover, the impact of social media on the stock market through emotional trading makes it incredibly effective in predicting stock market trends [3]. Therefore, it is confusing whether Reddit users are manipulating the stock market in the GameStop event [4]. On the other hand, Massimo et al. believe that Reddit communities replicate GameStop's success on the cryptocurrency markets [5].

Although some methods have discussed the relationship between WSB and GME prices [6], they never used causality detection methods. In order to determine whether the WallStreetBets's users promoted this event and whether the same operation was performed on the cryptocurrency, we adopted a causal analysis method based on the post text data expressed by the Redditors. Overall, the main contributions of this article are as follows: (1) Analyze the causal interaction between WSB's users sentiment and GME's price (2) Analyze

the causality between WSB's users and cryptocurrency by constructing some causal networks.

The rest of this article is organized as follows. Section II provides detailed information about the data set. Then, in Section III, we introduced how to perform an emotional, causal analysis on the text content of WSB. Section IV shows the experimental results of the cryptocurrency causal network analysis. Finally, section V summarized some instructive conclusions.

II. DATA SET

A. Reddit Data Set

Reddit is a social news website. It ranks fourth in size in the US, behind Google, YouTube and Facebook. There is a subreddit on Reddit dedicated to discussing stocks: r/WallStreetBets(WSB). In this study, we collected historical data of WSB, covering the period from December 11, 2020, to February 11, 2021. Fig. 1 shows Google Trends with WallStreetBets as the keyword during this period. On January 28, WallStreetBets' Google Trends value soared from 22 to the highest value of 100. The "epic war" between retail investors and institutional investors in the U.S. stock market became intensified on the same day. Fig. 2 shows the stock price trend of GameStop during this period. Obviously, the trend of Fig. 1 and Fig. 2 is very similar.

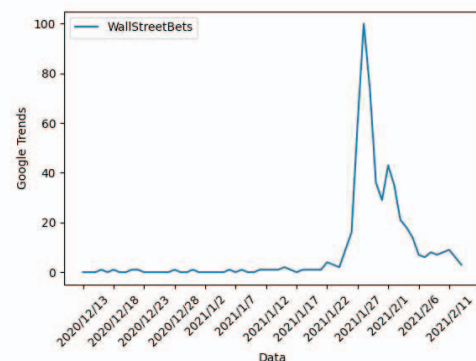


Fig. 1. Google Trends of WallStreetBets.

* Corresponding author

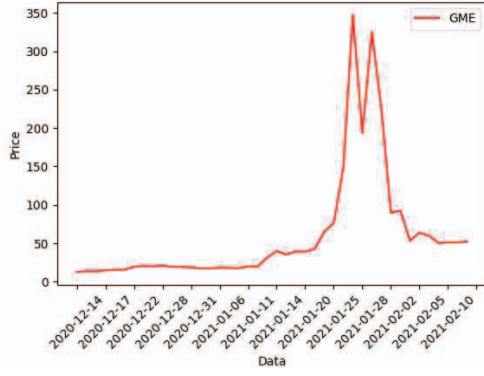


Fig. 2. Stock price trend of GameStop.

B. Yahoo Finance Data Set

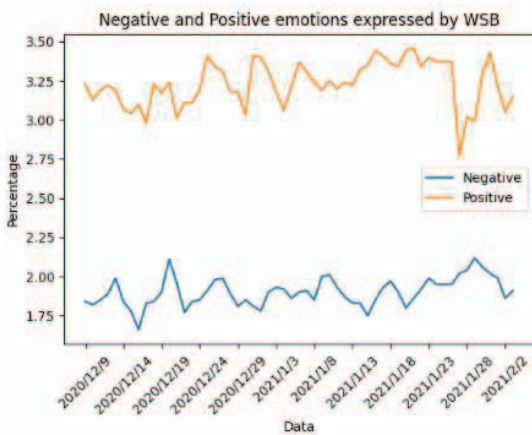
Yahoo Finance is a financial channel website owned by Yahoo. It has always been the most famous financial news and financial research website in the United States. Since January 2008, the website has had more than 37.5 million unique visitors every month. In this article, we obtain timely public stock prices and cryptocurrency prices from Yahoo Finance.

III. EMOTIONAL CAUSALITY ANALYSIS

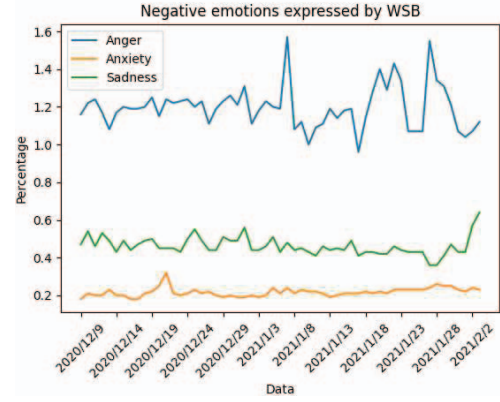
In this part, we first perform an analysis of posts and comments in WSB to reveal Redditors' sentiments. Then we further show the empirical results of their personality traits and reveal the causality between emotion, personality traits, and GME price.

A. Sentiment Analysis

LIWC [7] is the software that can quantitatively analyze word categories (primarily psychological words) of text content and calculate the different types of words in the imported text file, such as the percentage of causal words, inspirational words, and cognitive words in the entire text. The LIWC2015 dictionary defines more than 70 different categories (negative emotions, sex, work, sleep, etc.), mainly containing dozens or hundreds of words. The score of each category is calculated by dividing the number of all words' occurrence in the category by the total number of words in the text. We use the LIWC to count the number of sentiment words and describe the trend of sentiment words by date. Fig. 3 shows the changes over time of the positive and negative emotions of posts and comments in WSB.



(a)



(b)

Fig. 3. Emotions expressed by WSB (a) positive and negative emotions. (b) negative emotions.

Fig. 3(a) shows the WSB Redditors always have more positive emotions than negative emotions, which reveals that they are optimistic about success in this incident. However, in terms of negative emotions in Fig. 3(b), anger is the most, followed by anxiety and sadness, indicates that some things make them angry.

B. Personality Analysis

In order to determine the personality traits of WSB Redditors, we follow the work of Yarkoni [8] to count the number of emotional words using LIWC. For the big-five personality traits (agreeableness, conscientiousness, extraversion, neuroticism, and openness) from the Redditors of WSB, we calculate it as the dot multiply results of the 46 attributes [9] generate by LIWC 2015 (we neglected the remaining 20 attributes reported by LIWC 2015) and the corresponding coefficients that are significant in Yarkoni's work. Fig. 4 shows five personality traits of WSB Redditors.

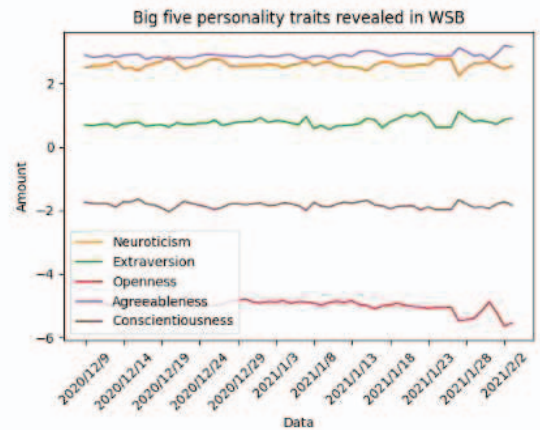


Fig. 4. Big-five personality traits revealed in WSB.

From Fig. 4, we can observe that the two prominent personality traits of WSB Redditors are agreeableness and neuroticism, which are the same as our intuition. However, WSB Redditors lack openness and conscientiousness personality traits.

C. The Causality analysis

We want to know whether there is a causality between WSB Redditor and GME's prices and what sentiments have affected the stock prices. A new causality analysis method

based on manifold reconstruction in geometry—the Convergent Cross Mapping (CCM) method has been widely used in recent years. The CCM algorithm can analyze causality in variable, non-stationary time series. At the same time, it can identify the causality in weakly coupled dynamic systems and short-term time series. Furthermore, CCM can distinguish bidirectional, unidirectional causality between variables and detect the strength of causality. According to existing work, this method is very effective in social media causal inference [10]. In addition, this article uses the FNN [11] algorithm to determine the embedding dimension and embedding delay required by the CCM algorithm.

We used the CCM algorithm to calculate the strength of the causality between GME price and WSB Redditors sentiment and thus established an emotional, causal network centered on GME price. Fig. 5 shows the causal network, from which we can observe that positive emotion, negative emotion, and GME price all have a bidirectional causality. However, the causality between positive emotion and GME price is significantly higher than the causality between negative emotion and GME price. It reveals that WSB Redditors have indeed affected GME's stock price trend and that the positive emotion of WSB Redditors has a more significant impact than the negative emotion.



Fig. 5. GME-centric emotion causal intensity network.

Ye et al. [12] proposed Extend-CCM in 2015, which can detect the causal influence delay between variables. We further use Extend-CCM to establish a causal delay network, as shown in Fig. 6



Fig. 6. GME-centric emotion causal time-delay network.

According to the theory of Extend-CCM, it can infer from Fig. 6 that the causal delay of positive sentiment on stock price and the causal delay of GME stock price on negative sentiment is very short, almost instantaneous. It reveals that the impact of positive emotion on GME's price is direct causality and; GME price on negative emotion is also a direct causality.

IV. CAUSAL ANALYSIS OF THE CRYPTOCURRENCY

During the surge in GME's price, cryptocurrencies such as Bitcoin and Dogecoin increased in price. On February 1, 2021, the subscribers to subreddit: r/Bitcoin reached 2,184,941, and this amount was 1 million 18 months ago. Therefore, we believe that cryptocurrency may be the next GME.

In this part, we set the date of December 11, 2020, to January 11, 2021, as the first stage, and January 11, 2021, to February 11, 2021, as the second stage. The reason for choosing January 11, 2021, is Chewy founder Cohen announced that he would join the GameStop board of directors on this day, making GameStop shares rise by 57%, 27%, respectively, during the period from January 13 to 14. Therefore, this day can be regarded as an essential beginning of the GameStop event. We constructed a total of 8 different causal networks based on two different periods and different network centers. The result is shown in Fig. 7, and their basic properties are shown in Table I and Table II.

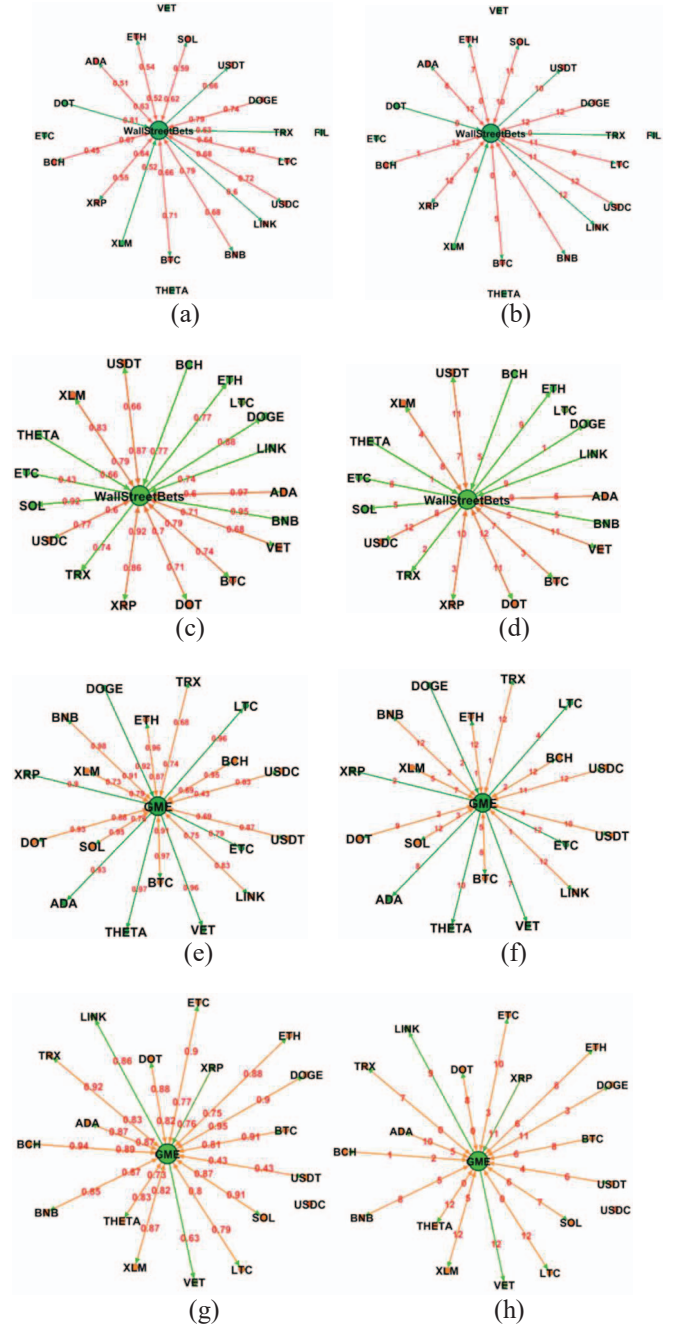


Fig. 7. (a)First stage of WallStreetBets-centric causal strength network.(b) First stage of WallStreetBets-centric causal time-delay network. (c) Second stage of WallStreetBets-centric causal strength network. (d)Second stage of WallStreetBets-centric causal time-delay network. (e)First stage of GME-centric causal strength network.(f) First stage of GME -centric causal time-delay network. (g) Second stage of GME-centric causal strength network. (h)Second stage of GME-centric causal time-delay network.

TABLE I. WALLSTREETBETS-CENTRIC CAUSAL NETWORK PROPERTIES

Phases	Average causal strength	Average causal delay	Degree
1	0.632	6.9	50
2	0.7824	6.1	50

TABLE II. GME-CENTRIC CAUSAL NETWORK PROPERTIES

Phases	Average causal strength	Average causal delay	Degree
1	0.85	6.89	50
2	0.84	6.55	62

Table I shows several attributes of the causal network centered on WallStreetBets. First, we found that the strength of the causal relationship in the second stage is significantly higher than that in the first stage, which indicates that after GameStop, WSB users are more active in the cryptocurrency market. At the same time, the causal delay of the second stage is also significantly smaller than that of the first stage, which shows that WallStreetBets has a more direct causal effect on cryptocurrencies. Table II shows several attributes of a causal network centered on GME; we find that the average causal strength and average causal delay are unchanged, indicating that GME stock price and cryptocurrency have influence each other and have not changed because of this incident. However, the degree of the network has become larger, indicating that GME's stock price has interacted with more cryptocurrencies.

V. CONCLUSION

This article uses Reddit and Yahoo Finance data to conduct a large-scale empirical study on the causality of the GameStop episode from two aspects: sentiment analysis and cryptocurrency. We draw the following remarkable conclusions:

(1) The Redditors in WSB did affect the price of GME. These Redditors have two significant personality traits, agreeableness and neuroticism, always full of positive sentiment, prompted them to affect GME's stock price directly.

(2) There is a causal interaction between WSB and GME prices with cryptocurrencies in two stages, which shows that Redditor in WSB have been paying attention to cryptocurrencies. Moreover, the average causal strength of

the WSB-centric network has increased significantly in the second stage, indicating that WSB Redditor has indeed expanded the battlefield to the cryptocurrency market.

ACKNOWLEDGMENT

This work is supported the Ministry of Science and Technology of China under Grant Nos:2020AAA0108401, 2019QY(Y)0101, the Ministry of Health of China under Grant No. 2017ZX10303401-002 and 2017YFC1200302, and the Natural Science Foundation of China under Grant No. 71602184 and 71621002.

REFERENCES

- [1] Z. Umar, M. Gubareva, I. Yousaf, and S. Ali, "A tale of company fundamentals vs sentiment driven pricing: The case of GameStop," *Journal of Behavioral and Experimental Finance*, vol. 30, 2021.
- [2] X. Zheng, Y. Zhong, D. Zeng, and F.-Y. J. F. o. C. S. Wang, "Social influence and spread dynamics in social networks," vol. 6, no. 5, pp. 611-620, 2012.
- [3] R. Burnette, "What Were the Factors that led to the GameStop Short Squeeze?," 2021.
- [4] S. Win, "Unstoppable GameStop: The Legislature's Gap in Regulating Open Market Manipulations," 2021.
- [5] M. La Morgia, A. Mei, F. Sassi, and J. J. a. p. a. Stefa, "The Doge of Wall Street: Analysis and Detection of Pump and Dump Cryptocurrency Manipulations," 2021.
- [6] S. Van Kerkhoven, and S. J. E. T. I. R. J. O'Dubhghaill, "Game Stop: How online'Degenerates' Took on Gamestop," 2021.
- [7] Y. R. Tausczik, J. W. J. J. o. l. Pennebaker, and s. psychology, "The psychological meaning of words: LIWC and computerized text analysis methods," vol. 29, no. 1, pp. 24-54, 2010.
- [8] T. J. J. o. r. i. p. Yarkoni, "Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers," vol. 44, no. 3, pp. 363-373, 2010.
- [9] X. Zheng, X. Wang, Z. Li, R. Jing, S. Xu, T. Wang, L. Li, Z. Zhang, Q. Zhang, H. Jiang, Z. Guo, X. Zhang, and F.-Y. Wang, "Donald J. Trump's Presidency in Cyberspace: A Case Study of Social Perception and Social Influence in Digital Oligarchy Era," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 2, pp. 279-293, 2021.
- [10] C. Luo, X. Zheng, and D. Zeng, "Causal inference in social media using convergent cross mapping," pp. 260-263.
- [11] M. B. Kennel, R. Brown, and H. D. J. P. r. A. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," vol. 45, no. 6, pp. 3403, 1992.
- [12] H. Ye, E. R. Deyle, L. J. Gilarranz, and G. Sugihara, "Distinguishing time-delayed causal interactions using convergent cross mapping," *Sci Rep*, vol. 5, pp. 14750, Oct 5, 2015.