Exploiting investors social network for stock prediction in China’s market

Xi Zhang\(^{a,}\)\(^{,*}\), Jiawei Shi\(^{a}\), Di Wang\(^{a}\), Binxing Fang\(^{a,b}\)

\(^{a}\) Key Laboratory of Trustworthy Distributed Computing and Service, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China

\(^{b}\) Institute of Electronic and Information Engineering of UESTC in Guangdong, Dongguan, Guangdong 523808, China

**Abstract**

Recent works have shown that social media platforms are able to influence the trends of stock price movements. However, existing works have major focused on the U.S. stock market and lacked attention to certain emerging countries such as China, where retail investors dominate the market. In this regard, as retail investors are prone to be influenced by news or other social media, psychological and behavioral features extracted from social media platforms are thought to well predict stock price movements in the China’s market. Recent advances in the investor social network in China enables the extraction of such features from web-scale data. In this paper, on the basis of tweets from Xueqiu, a popular Chinese Twitter-like social platform specialized for investors, we analyze features with regard to collective sentiment and perception on stock relatedness and predict stock price movements by employing nonlinear models. The features of interest prove to be effective in our experiments.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Social networks such as Twitter, Weibo, Facebook, and LinkedIn have attracted millions of users to post and acquire information, which have been well studied by various works [1–4]. In addition to these general social networks, there is another breed of smaller, more focused sites that cater to niche audiences. Here we look at a social site designed for traders and investors, that is, Xueqiu. Xueqiu is a specialized social network for Chinese investors of the stock market, and due to the increasing number of retail investors, Xueqiu has attracted millions of users. Xueqiu enables investors to share their opinions on a twitter-like platform, or post their portfolios, demonstrating their trading operations and returns. Different from those general social networks, almost all the information on Xueqiu is related to stocks, making it a natural data source to collect investors’ perceptions, which may be useful for stock market prediction in China.

The literature on stock market prediction was early based on the Efficient Market Hypothesis (EMH) and random walk theory [5]. However, investors’ reactions may not support a random walk model in reality. Behavioral economics has provided plenty of proofs that financial decisions are significantly driven by sentiment. The collective level of optimism or pessimism in society can affect investor decisions [6,7]. Besides, investor perceptions on the relatedness of stocks can also be a potential predictor. Firms may be economically related with one another [8,9]. Therefore, there is a probability that one stock’s price movement can influence its peer’s due to the investment reactions driven by investors’ perceptions on such relatedness.

Sentiment and perception are psychological constructs and thus difficult to measure in archive analyses. News articles have been used as a major source for textual content analysis. For example, news articles are employed to analyze public mood [10], by which stock price movements can be predicted. However, this type of content has an obvious drawback that news articles directly reflect their authors’ sentiment rather than the investors’. Online social platforms have provided us with more direct data and enable opportunities for exploring users’ sentiment and perception. In recent studies, it is found in [11] that collective mood derived from Twitter feeds improved the prediction accuracy of DowJones Industrial Average (DJIA). Facebook’s Gross National Happiness (GNH) index is shown to have the ability to predict changes in both daily returns and trading volume in the U.S. stock market [12]. The predictability of StockTwits (Twitter-like platform specialized on exchanging trading-related opinions) data with respect to stock price behavior is reported in [13].

---

\(^{*}\) Corresponding author.

E-mail addresses: zhangx@bupt.edu.cn (X. Zhang), 201121278@bupt.edu.cn (J. Shi), jgxzw@bupt.edu.cn (D. Wang), langbx@bupt.edu.cn (B. Fang).

https://doi.org/10.1016/j.jocs.2017.10.013

1877-7503/© 2017 Elsevier B.V. All rights reserved.

Fig. 1. Tweets number of Xueqiu vs. trading volume and turnover.

Most of the existing studies have focused on the U.S. stock market and lacked attention to certain emerging countries such as China, where the stock market is inefficient exhibiting a considerable non-random walk pattern [14]. The China’s stock market (also denoted as the A-share market) differs remarkably from other major markets in the structure of investors. Specifically, unlike other major stock markets, which are dominated by institutional investors, retail investors account for a greater percentage in China’s market. Importantly, retail investors are more likely to buy rather than sell stocks that catch their attention and thus tend to be influenced by news or other social media [15]. Therefore, in this paper, we study the China’s stock market based on a unique dataset from a popular Chinese Twitter-like social platform specialized for investors, namely Xueqiu (which means ‘snowball’ in Chinese), aiming to fill this gap in the literature.

To demonstrate how closely Xueqiu is related to the China’s stock market, Fig. 1(a) shows the daily published tweets volume of all stocks on Xueqiu and the daily trading volume of the A-share market from November 2014 to May 2015. It can be observed that the fluctuation trends of these two curves show great synchronicity, especially when high trading volume volatility occurs. When we look at the individual stocks, the synchronicity between the movement of daily tweets volume and the movement of daily turnover rate still holds, as displayed in Fig. 1(b), where one of the most popular stocks in Xueqiu, that is, the CITIC Securities, is taken as an example. On the basis of the tweets from Xueqiu, we analyze features with regard to collective sentiment and perception. The sentiment and perceived stock relatedness are proposed to be formed on the basis of two types of networks extracted from Xueqiu. One is the user network, and the other is the stock network perceived by users. Combined with the network characteristics, the features can exhibit better predictive performance. In contrast to previous works that only study a small subset of the stocks, we evaluate our proposal on all the active stocks (more than 2000) in the A-share market, indicating it’s a feasible approach.

In the remainder of the article, we first briefly introduce related research in Section 2. The online social platform Xueqiu and the crawled dataset are described in Section 3. Then, we describe the methodology in Section 4 and present the experiment of predicting stock price movements in Section 5. Finally, the article is concluded in Section 6.

2. Sentiment, perception, and stock market

2.1. Stock prediction with historical price data

Most of previous studies utilize historical stock prices to make predictions with various models [16–19]. A Support Vector Machine-based model is proposed by using the selected subset of financial indexes as the weight inputs [20]. A multi-layer perceptron method is proposed for short-term stock prediction in [21]. Multiple techniques of Artificial Neural Network (ANN) in stock market prediction are evaluated in [22]. However, these works only use the historical price data and ignores the impacts of social media.

2.2. Sentiment and stock price movement

A variety of studies have found that financial news can have significant effects on stock price movements [23–27]. Recent studies try to extract events from the news with natural language techniques for event-driven prediction [28, 29]. News sentiments are measured and the combined effect of Web news and social media on stock markets are studied in [30]. The investors’ sentiments can also be extracted from social networks, media platforms, and blogs. It is reported that social networks such as Twitter [11, 31] and Facebook [12] are important sources of sentiment data. Specialized social networks, such as StockTwits, has also shown its predictive power [13]. A method to measure the collective hope and fear on each day and analyze the correlation between these indices and the stock market indicators is proposed in [32]. A topic-based sentiment time series approach is proposed to predict the market [33]. This work is extended to further exploit the social relations between stocks from the social media context. A stock network is built with Twitter by co-occurring relationships, and a labeled topic model is employed to jointly model the tweets and the network structure to assign each node and each edge a topic respectively. Then, a lexicon-based sentiment analysis method is used to compute the sentiment score for each node and edge topic. Last, the sentiment time series and price time series are used for prediction [34]. Financial trend prediction can be boosted with Twitter moods based on deep network models [35]. Sentiments and events are integrated with a tensor for stock prediction in [36].
دریافت فوری متن کامل مقاله

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات