

# Twitter Volume Spikes: Analysis and Application in Stock Trading

Yuexin Mao  
University of Connecticut  
yuexin.mao@uconn.edu

Wei Wei  
FinStats.com  
weiwei@finstats.com

Bing Wang  
University of Connecticut  
bingw@engr.uconn.edu

## ABSTRACT

Stock is a popular topic in Twitter. The number of tweets concerning a stock varies over days, and sometimes exhibits a significant spike. In this paper, we investigate Twitter volume spikes related to S&P 500 stocks, and whether they are useful for stock trading. Through correlation analysis, we provide insight on when Twitter volume spikes occur and possible causes of these spikes. We further explore whether these spikes are surprises to market participants by comparing the implied volatility of a stock before and after a Twitter volume spike. Moreover, we develop a Bayesian classifier that uses Twitter volume spikes to assist stock trading, and show that it can provide substantial profit. We further develop an enhanced strategy that combines the Bayesian classifier and a stock bottom picking method, and demonstrate that it can achieve significant gain in a short amount of time. Simulation over a half year's stock market data indicates that it achieves on average 8.6% gain in 27 trading days and 15.0% gain in 55 trading days. Statistical tests show that the gain is statistically significant, and the enhanced strategy significantly outperforms the strategy that only uses the Bayesian classifier as well as a bottom picking method that uses trading volume spikes.

## Categories and Subject Descriptors

H.2.8 [Database Management]: Database Applications - Data Mining

## General Terms

Algorithm, Measurement, Performance

## Keywords

Twitter, Stock, Twitter Volume Spikes, Stock trading

## 1. INTRODUCTION

Twitter is a widely used online social media. Researchers have studied various aspects of Twitter, for instance, the general characteristics of the entire Twitter social network (e.g., [7], [8]) and

the social interactions within Twitter [6]. In addition, Twitter has been used to detect and predict real-world events including earthquakes [12], box-office revenues of movies [3], and seasonal influenza [2].

Stock is a often tweeted topic in Twitter. The number of tweets concerning a stock varies over days, and sometimes exhibits a significant spike, indicating a sudden increase of interests in the stock. In this paper, motivated by the observation of Twitter volume spikes, we aim to answer the following questions: (1) When do Twitter volume spikes occur? Are they surprises or expected? What are the potential causes of Twitter volume spikes? and (2) Are Twitter volume spikes useful for stock trading?

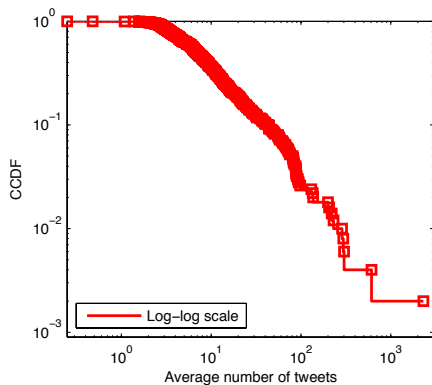
We make the following main contributions:

- We find that Twitter volume spikes often happen around earnings dates. Specifically, 46.4% of Twitter volume spikes fall into category. By comparing the implied volatility of a stock before and after a Twitter volume spike, we show that many Twitter volume spikes might be related to pre-scheduled events, and hence are expected to market participants. Furthermore, through correlation analysis, we investigate five possible causes of Twitter volume spikes including stock breakout points, large stock price fluctuation within a day and between two consecutive days, earnings days and high implied volatility. Our results show that only the last two factors show significant correlation with Twitter volume spikes.
- We develop a Bayesian classifier that uses Twitter volume spikes to assist stock trading, and show that it can provide substantial profit. We further develop an enhanced strategy that combines the Bayesian classifier and a stock bottom picking method, and demonstrate that it can achieve significant gain in a short amount of time. Simulation over a half year's stock market data indicates that it achieves on average 8.6% gain in 27 trading days and 15.0% gain in 55 trading days. Statistical tests show that the gain is statistically significant, and the enhanced strategy significantly outperforms the strategy that only uses the Bayesian classifier as well as a bottom picking method that uses trading volume spikes.

As related work, several studies use Twitter to predict stock market. A recent study [5] finds that specific public mood states in Twitter are significantly correlated with the Dow Jones Industrial Average (DJIA), and thus can be used to forecast the direction of DJIA changes. Another study [13] finds that emotional tweet percentage is correlated with DJIA, NASDAQ and S&P 500. Later on, the study [9] finds that Twitter sentiment indicator and the number of tweets that mention financial terms in the previous 1-2 days can be used to predict the daily market return. The study [10] fo-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SNAKDD'13 August 11, 2013, Chicago, United States.  
Copyright 2013 ACM 978-1-4503-2330-7 ...\$5.00.



**Figure 1: CCDF of the average number of tweets for the S&P 500 stocks.**

cuses on the daily number of tweets that mention S&P 500 stocks, and finds that the daily number of tweets is correlated with certain stock market indicators at three different levels, from the stock market, to industry sector, and then to individual company stocks. The study [11] also reports the correlation between trading volume and the daily number of tweets for individual company stocks. Our study differs from all the above in that we focus on Twitter volume spikes, and how they can be used for stock trading.

The rest of the paper is organized as follows. Section 2 describes how we collect data and identify Twitter volume spikes. Section 3 presents the analysis of Twitter volume spikes and possible causes of these spikes. Section 4 presents the trading strategies and their performance. Last, Section 5 concludes the paper and presents future work.

## 2. METHODOLOGY

### 2.1 Stock market data

We obtained daily stock market data from Yahoo! Finance [1] for the 500 stocks in the S&P 500 index. For each stock, we consider four types of data: stock daily closing price, stock daily high price, stock daily low price and stock daily trading volume.

### 2.2 Twitter data

In Twitter community, people usually mention a company's stock using the stock symbol prefixed by a dollar sign. For example, \$AAPL represents the stock of Apple Inc. and \$GOOG represents the stock of Google Inc. When collecting public tweets on S&P 500 stocks, we only search for tweets that follow the above convention (i.e., having a dollar sign before a stock symbol). This is because many stock symbols (e.g., A, CAT, GAS) are common words, and hence using search keywords without the dollar sign will result in a large number of spurious tweets. We collect Twitter data from February 21, 2012 to May 31, 2013, over 15 months.

### 2.3 Twitter volume spikes

Fig. 1 plots the CCDF (complementary cumulative distribution function) of the average number of tweets for the S&P 500 stocks. We observe that the average number of tweets for the stocks is in a wide range, varying from only a few tweets to above 2,000 tweets per day. In the rest of the paper, we only consider the stocks with average daily number of tweets larger than 10. There are 168 such stocks in S&P 500.

Consider the tweets on a stock. We say the number of tweets on a day is a spike if it is at least  $K$  times the average number of tweets in the past  $N$  days,  $K > 1$ . In this paper, we set  $N$  to 70, i.e., approximately three months of stock market trading days, and set  $K$  to 2, 3 or 4. Unless otherwise stated, the results presented in the paper use  $K = 3$ ; results when  $K = 2$  or 4 show similar trends.

## 3. TWITTER VOLUME SPIKE ANALYSIS

In this section, we investigate when Twitter volume spikes occur, whether they are surprises or not, and the potential causes of Twitter volume spikes.

### 3.1 When do Twitter volume spikes occur?

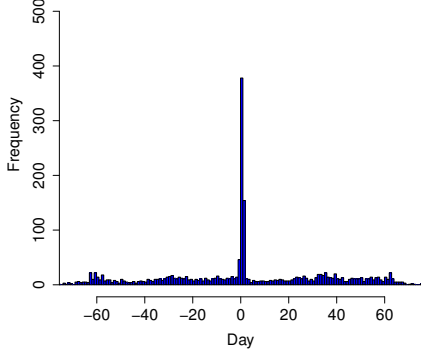
We expect that the number of tweets concerning a stock increases sharply when people show particular interests in the stock. One such occasion is company earnings dates, when a company releases earnings reports to inform public their performance during the past time period (most companies release an earnings report each quarter of a fiscal year). People may show particular interests in a company's stock when the company is going to report earnings. In the following, we investigate whether the number of tweets for a stock spikes around the earnings dates.

Suppose that a company's earnings date is day  $t$ . We investigate whether the number of tweets on the company's stock spikes around  $t$ , in particular, on days  $t - 1$ ,  $t$  and  $t + 1$ . In our data collection period, there are 509 earnings days for the stocks that we consider. We find 79.2% of them are surrounded by a Twitter volume spike, confirming our intuition that people indeed tweet more about a stock around its earnings dates. Fig. 2 plots the histogram of the time difference (in days) from an earnings day to the closest day that has a Twitter volume spike, where a negative value corresponds to the time difference to the closest Twitter volume spike in the past. We see that most of the time, the time difference is either 0 (i.e., they are on the same day), 1 (i.e., Twitter volume spike happens on the next day), or -1 (i.e., Twitter volume spike happens on the previous day).

In addition, we mark the Twitter volume spikes that coincide with earnings days, specifically, the spikes that happen within one day (earlier or later) of the earnings days, and find that 46.4% of the Twitter volume spikes fall into this category. This indicates that a significant fraction of Twitter volume spikes happen around earnings days.

### 3.2 Are Twitter volume spikes expected?

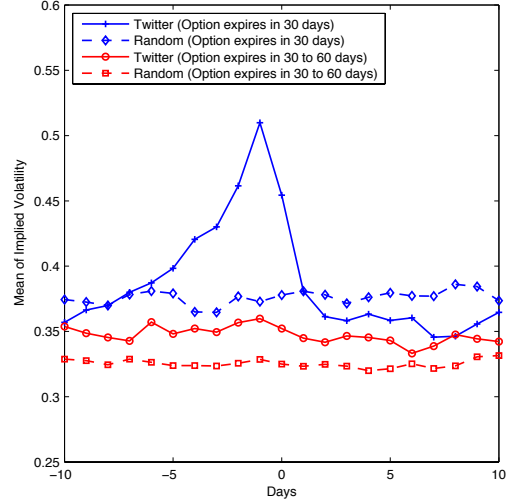
Twitter volume spikes close to earnings days are likely due to the earnings days. Since earnings days are public information that people know beforehand, these Twitter volume spikes are no surprises. Certain other scheduled events (e.g., a financial meeting) can also cause Twitter spikes. It is, however, difficult to enumerate such events one by one. On the other hand, we conjecture option implied volatility can be used as an indicator to determine whether a Twitter volume spike is expected or not, that is, whether it is related to a scheduled event. Specifically, we regard a Twitter volume spike as expected when the implied volatility is larger than usual before the spike happens, and returns back to the usual status after the spike happens. The rationale is as follows. Option implied volatility of a stock indicates how volatile the stock is expected to be based on option prices of the stock. In other words, it indicates how uncertain people feel about the stock. Consider a scheduled event. People anticipate the event, but do not know its impact, hence feel more uncertain about the stock, manifested by the higher implied volatility. Once the event happens, uncertainty reduces, and hence



**Figure 2: Time difference (in days) from an earnings day to the closest day that has a Twitter volume spike. A negative value corresponds to the time difference to the closest Twitter volume spike in the past.**

implied volatility returns back to the normal status. When a Twitter volume spike happens between an increased and back-to-normal implied volatility, it is likely related to the anticipated event, and hence is an expected spike. An earnings day described in the previous section is one special case of such expected events. In the following, we obtain the implied volatility of a stock on a day as the weighted average of the implied volatilities of all the options of the stock, where the implied volatility of an option is derived using the Black-Scholes model [4] and the weight for an option is its trading volume on that day.

To shed lights on whether Twitter volume spikes are expected or not, we compare the implied volatility of a stock before and after a Twitter volume spike occurs. Specifically, assume that for a stock, a Twitter volume spike happens on day  $t$ . Then we calculate the implied volatility of the stock from day  $t - 10$  to  $t + 10$ . We consider both short-term options, i.e., those that will expire in 30 days after  $t$ , and longer-term options, i.e., those that will expire in 30 to 60 days after  $t$ . Fig. 3 plots the average implied volatility for the ten days before and after a Twitter volume spike, where the index of the days is from -10 to 10, relative to when a Twitter volume spike happens, and the average is obtained considering all the Twitter volume spikes (there are 1245 Twitter volume spikes for all the stocks when  $K = 3$ ). For short-term options, we indeed observe that the daily average implied volatility increases before  $t$  and decreases after  $t$ . For longer-term options, the trend is not clear. This might be because option traders usually use short-term options to bet on short-term events to take advantage of higher leverages of short-term options, and hence the implied volatility considering longer-term options is not sensitive to short-term events. For comparison, we also investigate how implied volatility changes before and after a day that is chosen randomly. Specifically, suppose for a stock, a Twitter volume spike happens on day  $t$ , then we randomly choose a day  $t'$  and calculate the implied volatility of the stock from day  $t' - 10$  to  $t' + 10$ . Fig. 3 plots the average implied volatility for the ten days before and after such a randomly chosen day, where the average is obtained over all the randomly chosen days (there are 1245 such days). We see the daily average implied volatility shows no significant difference before and after a day that is chosen randomly.



**Figure 3: Daily average implied volatility in each of the ten days before and after a Twitter volume spike. Results for randomly chosen days are also plotted in the figure.**

**Table 1:  $p$ -values of the t-tests for  $\mu_{\tau}^{-} < \mu_{\tau}^{+}$ . Only consider options that will expire in 30 days after  $t$ .**

$\tau$	Twitter volume spike	Random day
5	8.14E-29	0.234
6	1.00E-24	0.510
7	9.71E-23	0.112
8	4.05E-21	0.727
9	6.54E-20	0.763
10	1.52E-17	0.692

We next use t-test to further confirm the above results. Suppose that for a stock a Twitter spike happens on day  $t$ . We only consider the options that will expire in 30 days after  $t$ . Let  $\mu_{\tau}^{-}$  denote the mean of the daily implied volatility from day  $t - \tau$  to  $t - 1$ . Let  $\mu_{\tau}^{+}$  denote the mean of the daily implied volatility from day  $t + 1$  to  $t + \tau$ . The null hypothesis is that  $\mu_{\tau}^{-} \leq \mu_{\tau}^{+}$ . Table 1 shows the  $p$ -values of the t-tests when varying  $\tau$  from 5 to 10. The very small  $p$ -values indicate that we can reject the null hypothesis, indicating that there is strong evidence that  $\mu_{\tau}^{-} > \mu_{\tau}^{+}$ , further confirming that the implied volatility before a Twitter volume spike is usually larger than that after the spike. For comparison, we also show the t-test results when choosing a random day, which exhibit large  $p$ -values, indicating no strong evidence that  $\mu_{\tau}^{-} > \mu_{\tau}^{+}$ .

The above considers the average behavior of all the Twitter volume spikes. Next we consider individual Twitter volume spikes, and identify the percentage of Twitter volume spikes that are expected. As mentioned earlier, we regard a Twitter volume spike as expected when the implied volatility is larger than usual before the spike, and returns back to the usual status after the spike. To be quantitative, we say a Twitter spike that happens on day  $t$  is expected if the implied volatility on  $t - 1$  is larger than the mean implied volatility in the past  $N$  days, i.e., from  $t - 1 - N$  to  $t - 2$ , and the implied volatility on  $t + 1$  is smaller than the mean of the past  $N$  days, i.e., from  $t + 1 - N$  to  $t$ , where  $N = 70$ , i.e., approximately three months of market trading days. We find 37.3% of the Twitter volume spikes satisfy the above condition. Note that

this percentage is a very conservative estimate due to the difficulty to quantitatively specify larger and lower than usual. Nonetheless, the result provides a lower bound, indicating that a significant percentage of Twitter volume spikes are expected.

### 3.3 Possible causes of Twitter volume spikes

We now investigate potential causes of Twitter volume spikes. Specifically, we consider the following five factors: (i) stock breakout point, (ii) intraday price change rate, (iii) interday price change rate, (iv) earnings day, and (v) stock option implied volatility. In the following, we define the first three factors (the last two factors have been defined earlier), and then calculate the correlation of each of these five factors with Twitter volume spikes.

Consider a stock. We use  $p_t^c$ ,  $p_t^h$ ,  $p_t^l$  to denote the daily closing price, daily high price, daily low price of the stock on day  $t$ , respectively. A stock breakout point is a situation where the price of the stock breaks above a resistance level and rises higher, or breaks below a support level and drops lower. In the following, we say a breakout point happens on day  $t$  if the stock closing price is larger or smaller than the closing prices in all of the past  $N$  days. We again choose  $N = 70$ , approximately three months of stock market trading days. The intraday price change rate on day  $t$  is calculated as the difference of the daily high price and daily low price, divided by the daily closing price, i.e.,  $(p_t^h - p_t^l)/p_t^c$ . The interday price change rate between day  $t - 1$  and day  $t$  is calculated as the absolute value of the relative price change between these two days, i.e.,  $(p_t^c - p_{t-1}^c)/p_{t-1}^c$ .

Intuitively, the number of tweets for a stock may increase significantly when a breakout point happens, when it is around an earnings day, or under high intraday price change rate, high interday price change rate, or high implied volatility. In the following, we calculate the correlation coefficient to quantitatively investigate the correlation of Twitter volume spikes and each of the five factors. Consider a stock. Let  $\{T_t\}$  denote the time series of Twitter volume spikes, where  $T_t = 1$  if there is a Twitter volume spike on day  $t$ , and  $T_t = 0$  otherwise. Let  $\{B_t\}$  denote the time series of stock breakout points, where  $B_t = 1$  if there is a stock breakout point on day  $t$ , and  $B_t = 0$  otherwise. Let  $\{C_t\}$  denote the time series of relative intraday price change rate, where  $C_t$  is the intraday price change rate on day  $t$  normalized by the average intraday price change rate in the past 70 days. Similarly, let  $\{D_t\}$  denote the time series of relative interday price change rate, where  $D_t$  is the interday price change rate on day  $t$  normalized by the average interday price change rate in the past 70 days. Let  $\{E_t\}$  denote the time series for earnings days, where  $E_t = 0$  by default; while if  $t$  is an earnings day, then we set  $E_t = 1$ ,  $E_{t-1} = 1$ , and  $E_{t+1} = 1$  to include one day before and after  $t$ . Last, let  $\{I_t\}$  denote the time series of relative stock option implied volatility, where  $I_t$  is the stock option implied volatility on day  $t$  normalized by the average stock option implied volatility in the past 70 days.

We now present lag 1 cross correlation between Twitter volume spikes and each of the five factors, namely, the correlation between  $T_t$  and  $B_{t-1}$ , the correlation between  $T_t$  and  $C_{t-1}$ , and so on. The reason for choosing lag 1 is that we are interested in how the value of a factor on the previous day is correlated with the Twitter volume spike on the current day. Fig. 4 plots the CDF (cumulative distribution function) of the correlations between Twitter volume spikes and each of the five factors over all the stocks. We observe that Twitter volume spike has the strongest correlation with earnings days (with median of 0.37), which confirms our earlier result that a significant fraction of Twitter volume spikes occurs around earnings days. We also can see that the correlation between Twitter volume spike and implied volatility has a median value of 0.14,

much stronger than the correlation with the rest of the three factors.

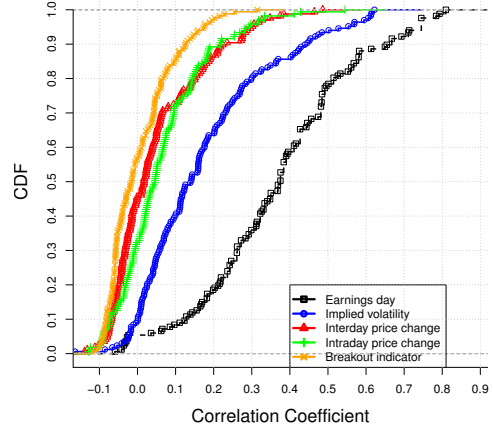


Figure 4: CDF of the lag 1 correlation coefficients between Twitter volume spikes and each of the five factors.

## 4. APPLICATION IN STOCK TRADING

After analyzing Twitter volume spikes, a natural question is whether they are useful for stock trading. We develop two trading strategies, both using Twitter volume spikes as trading signals. We next present these two strategies and their performance. For comparison, we also consider a baseline strategy that purchases a stock on a random day, and a strategy that uses trading volume spikes.

### 4.1 Strategy based on Bayesian classifier

For a stock, after observing a Twitter volume spike, a natural strategy to decide whether to buy the stock or not is as follows. We first calculate the probability that buying the stock can lead to profit after a number of days, and only buy the stock when the probability is sufficiently large. Specifically, we define two types of events, one corresponding to the events that buying the stock leads to profit, and the other corresponding to the opposite, denoted as  $\mathcal{G}$  and  $\bar{\mathcal{G}}$ , respectively. We use a set of features  $F_1, \dots, F_k$  to predict the probability that event  $\mathcal{G}$  happens, namely  $\Pr(\mathcal{G} | F_1, \dots, F_k)$ . Using Bayes rule, we have

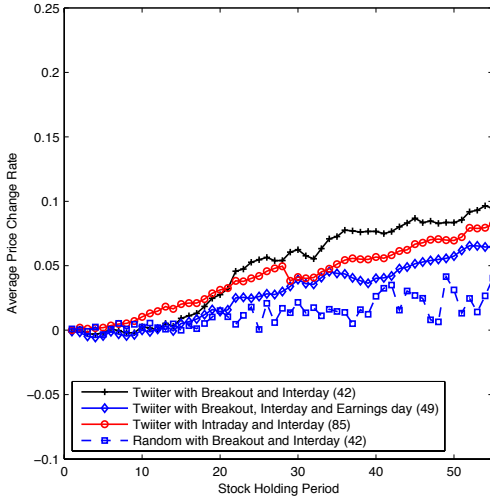
$$\Pr(\mathcal{G} | F_1, \dots, F_k) = \frac{\Pr(\mathcal{G}) \Pr(F_1, \dots, F_k | \mathcal{G})}{\Pr(F_1, \dots, F_k)} \quad (1)$$

To obtain  $\Pr(\mathcal{G} | F_1, \dots, F_k)$ , we use a training set to obtain the various probabilities on the right hand side. Obtaining  $\Pr(F_1, \dots, F_k | \mathcal{G})$  for large  $k$  (i.e., when the number of features is large) requires a large training set. For simplicity, we treat each of the features as independent, so that  $\Pr(F_1, \dots, F_k | \mathcal{G}) = \prod_{i=1}^k \Pr(F_i | \mathcal{G})$  and obtain  $\Pr(F_i | \mathcal{G})$  from the training set. Similarly, since  $\Pr(F_1, \dots, F_k) = \Pr(F_1, \dots, F_k | \mathcal{G}) \Pr(\mathcal{G}) + \Pr(F_1, \dots, F_k | \bar{\mathcal{G}}) \Pr(\bar{\mathcal{G}})$ , assuming independence, we have  $\Pr(F_1, \dots, F_k) = \prod_{i=1}^k \Pr(F_i | \mathcal{G}) \Pr(\mathcal{G}) + \prod_{i=1}^k \Pr(F_i | \bar{\mathcal{G}}) \Pr(\bar{\mathcal{G}})$ , where the various probabilities on the right hand side are obtained from training data.

To evaluate the above strategy, we use the data from February 21, 2012 to October 19, 2012 as training data, and use the data from October 20, 2012 to March 31, 2013 as test data. This results

in 573 Twitter volume spikes in the training set, and 672 Twitter volume spikes in the test set. The set of features used in the classifier is a subset of the five factors discussed in Section 3.3, excluding implied volatility which requires using option data and hence does not provide a fair comparison with other strategies. Since the stock market closes at 4pm (New York time) each day and the Twitter volume spikes are identified using the number of tweets throughout a day, when observing a Twitter volume spike on day  $t$  and we decide to buy the stock, we buy the stock on day  $t + 1$ , using the closing price on day  $t + 1$ . In training, we say buying a stock on day  $t$  makes profit if the stock closing price on day  $t + 10$  is larger than that on day  $t$ . In testing, we purchase a stock if the predicted probability,  $\Pr(\mathcal{G} | F_1, \dots, F_k)$ , is larger than 0.7.

We next report the performance of the above strategy. For a trade that buys stock on day  $t$  and sells the stock on day  $t + \tau$ , we refer to  $\tau$  as *stock holding period*, and define the *price change rate* as  $(p_{t+\tau}^c - p_t^c)/p_t^c$ , where  $p_t^c$  denotes the closing price on day  $t$ . The performance metric we use is average price change rate, defined as the average of the price change rate of all the trades. Fig. 5 plots the results for three sets of features: breakout point and interday price change rate; breakout point, interday price change rate and earnings days; intraday and interday price change rate. The stock holding period,  $\tau$ , is varied from 1 to 55 trading days. We can see that the strategy leads to substantial profit. The average price change rate is above 0 from day 10 to day 55 for all the three sets of features, which is also confirmed by t-test (detailed results of the t-test are omitted in the interest of space). In addition, we observe the average price change rate roughly increases over time. Specifically, when the features are breakout point and interday price change rate, the gain reaches 9.6% when  $\tau = 54$  trading days. Fig. 5 also plots the results of a random strategy. This random strategy differs from our strategy (when the features are breakout point and interday price change rate) in that if our strategy decides to buy stock  $s$  on day  $t$ , then it decides to buy  $s$  on a day  $t$  that is chosen randomly. We see from the figure that our strategy clearly outperforms the random strategy, which is confirmed by t-test.



**Figure 5: Performance of the strategy based on Bayesian classifier. In the legend of each setting, the number in the parentheses represents the number of trades.**

The results of the above simple strategy are encouraging, indicating that Twitter volume spikes are indeed useful in stock trad-

ing. On the other hand, the strategy does not consider the trend of a stock. For instance, it may buy a stock when the price of the stock is increasing, which may not lead to profit. In the following, we propose a further enhanced strategy that takes the trends of the stocks into account.

## 4.2 Enhanced strategy

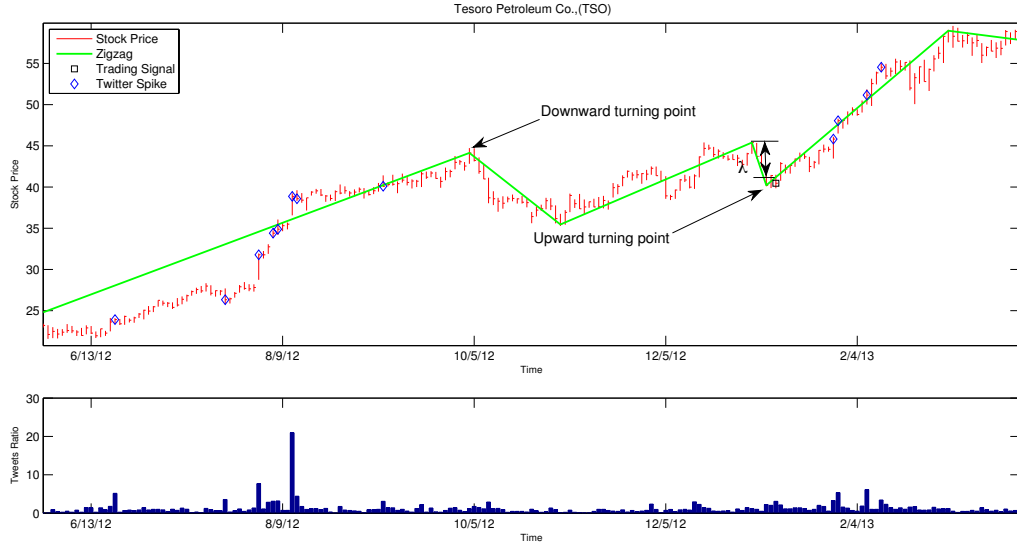
This strategy uses both Twitter volume spikes and stock turning points. We say that a stock has a turning point on day  $t$  when its trend changes on that day. Specifically, a downward turning point indicates that the stock price starts to move downward, and an upward turning point indicates that the stock price starts to move upward, as illustrated in Fig. 6. We apply a Zigzag based algorithm (based on ZigZag indicator) to identify turning points for a given movement rate,  $\lambda$ , which is defined as the minimum price difference ratio between two adjacent turning points (i.e., the relative difference between two adjacent turning points needs to be at least  $\lambda$ ). The stock price turning point identification algorithm for a given  $\lambda$  is described as follows.

### • Stock Price Turning Point Identification Algorithm

- (1) Start to search from the first point in the data set. Search forward until we find a potential turning point, i.e., one of the two conditions holds: (i) the price increases by at least  $\lambda$  from the start point, or (ii) the price decreases by at least  $\lambda$  from the start point. Continue the search.
  - (a) If condition (i) holds (i.e., the price moves upward), update the potential turning point when finding a point that is larger than the previous potential turning point. When finding a point that drops at least  $\lambda$  compared to the current potential turning point, set the current potential turning point to be a downward turning point.
  - (b) If condition (ii) holds (i.e., the price moves downward), update the potential turning point when finding a point that is smaller than the previous potential turning point. When finding a point that increases at least  $\lambda$  compared to the current potential turning point, set the current potential turning point to be an upward turning point.
- (2) Start to search from the turning point. If the turning point is an upward turning point, goes to Step (1)(a). If the turning point is a downward turning point, goes to Step (1)(b). Repeat till the end of the data set.

For each stock, we choose the movement rate,  $\lambda$ , based on a stock parameter,  $\beta$  value. The  $\beta$  value of a stock describes the correlated volatility of the stock price in relation to the volatility of the benchmark that the stock is being compared to. In our case, we use S&P 500 index as the benchmark. Specifically,  $\beta > 0$  means that the movement of the stock is in the same direction as the movement of the S&P 500 index, and  $\beta < 0$  means the opposite;  $\beta > 1$  means the movement of the stock is more than the movement of the S&P 500 index, and  $0 < \beta < 1$  means the movement of the stock is less than the movement of the S&P 500 index. For a stock, we use the historical stock closing prices from February 20, 2011 to February 21, 2012 to calculate the stock  $\beta$  value. For stocks with larger  $\beta$  values, we assign a larger movement rate rate. More specifically, we set  $\lambda$  to 10% when  $\beta > 1$ , and set  $\lambda$  to 7% otherwise. Fig. 6 illustrates both upward and downward turning points of a stock. The lines connecting two adjacent turning points form the ZigZag curve.





**Figure 6: Illustration of the turning points, ZigZag curve and bottom picking method using the price and tweets information of a stock. For the stock, the top figure shows the price over time; the bottom figure shows the tweets ratio, i.e., the number of tweets on a day over the average number of tweets in the past 70 days, over time. A day with tweets ratio above  $K$  has a Twitter volume spike.**

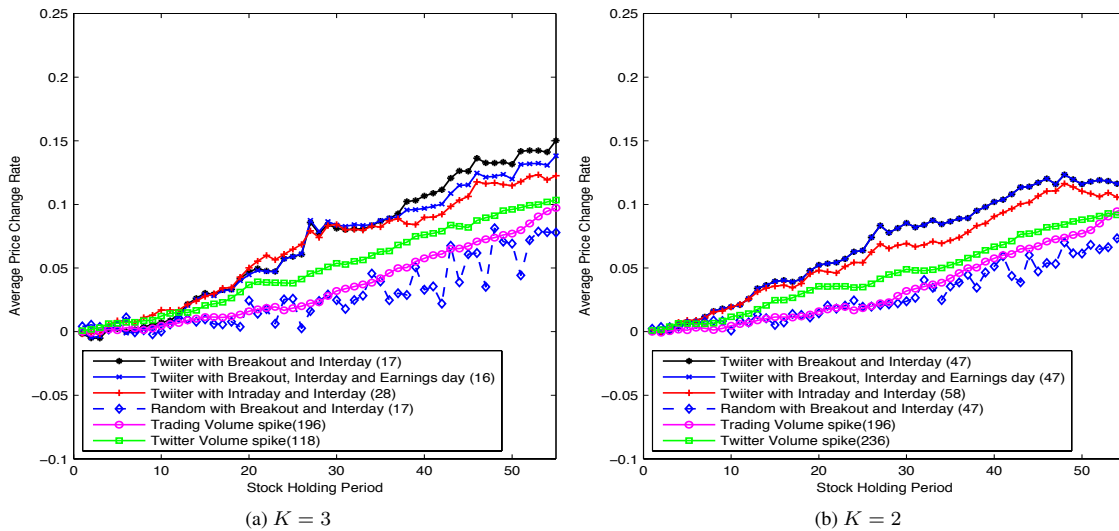
It is clear that using Twitter volume spikes that are close to the bottom of the ZigZag curve (where the stock price is a local minimum) as trading signals can make profit, as illustrated in Fig. 6. However, when a Twitter volume spike happens, we do not know whether it is close to the bottom because identifying the bottom requires future stock price information. We therefore use the following heuristic method to select Twitter volume spikes that are close to the bottom. First, we only select Twitter volume spikes that happen when the stock price is moving downward, i.e., after a downward turning point. Furthermore, suppose a Twitter volume spike happens on day  $t$ , and the previous downward turning point happens on day  $t'$ . We only choose the Twitter volume spike when the following two conditions are satisfied: (i) the price has changed by at least  $\lambda$ , i.e.,  $(p_{t'}^c - p_t^c)/p_t^c \geq \lambda$ , where  $p_t^c$  is the closing price on day  $t$ , and (ii) the closing price on day  $t$  is the minimum of the closing prices from  $t'$  to  $t$ . Since the stock price may fluctuate, we relax the above two conditions by including three earlier days,  $t-1$ ,  $t-2$  and  $t-3$ . That is, if these two conditions are satisfied on one of the four days, from day  $t-3$  to day  $t$ , then we select the Twitter volume spike as a buy signal. Fig. 6 shows one such selected Twitter volume spike. Observe that it is indeed close to the bottom. In the following, we refer to the Twitter volume spikes selected using the above bottom picking method as *valid* Twitter volume spikes.

The enhanced strategy combines the above bottom picking method with the Bayesian classifier described earlier. We again use the data from February 21, 2012 to October 19, 2012 as training data, and the data from October 20, 2012 to March 31, 2013 as test data. This results in 90 valid Twitter volume spikes in the training set, and 118 valid Twitter volume spikes in the test set. To demonstrate that Twitter volume spikes provide valuable information for stock trading, we also compare our strategy with a bottom picking method that is based on stock trading volume spikes, which only differs from our bottom picking method in that it uses stock

trading volume spikes, instead of Twitter volume spikes, as trading signals. That is, it uses the same algorithm to identify stock price turning points and the same heuristics to decide whether a day with a stock trading volume spike is close to the bottom of the ZigZag curve of the stock price. To identify stock trading volume spikes, we use the same method for identifying Twitter volume spikes (see Section 2.3), where we set  $N = 70$  and  $K = 2$ .

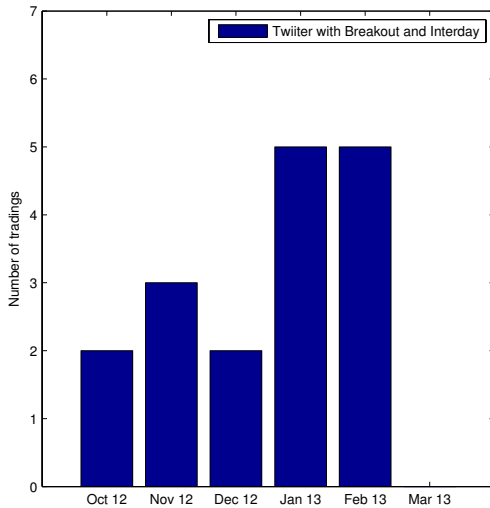
We now report the performance of the enhanced strategy. Figures 7(a) and (b) plot the average price change rate when  $K = 3$  and  $K = 2$ , respectively. The holding period,  $\tau$ , is again varied from 1 to 55 trading days. The results for three sets of features are plotted in the figure. We observe that the strategy achieves significant gain in a short amount of time. When  $K = 3$ , the average gain generally increases over time, achieving 8.6% gain when  $\tau = 27$ , and 15.0% when  $\tau = 55$ , significantly larger than the gains obtained by the strategy that only uses the Bayesian classifier. When  $K = 2$ , the gains are also significant (slightly lower than those when  $K = 3$ ), indicating that the strategy is not sensitive to the choice of  $K$ . We also observe that our strategy outperforms the random strategy, and the strategy that uses stock trading volume spikes. Last, we observe that the gain when only using valid Twitter volume spikes without using any feature (and hence it does not use the Bayesian classifier) is not as good, indicating it is important to use the bottom picking method along with the Bayesian classifier.

Table 2 presents the t-test results of the enhanced strategy when  $K = 3$ . We confirm that there is indeed strong evidence that the profit is positive, and the enhanced strategy outperforms the random strategy as well as the strategy that uses stock trading volume spikes. Specifically, when the features are breakout point and inter-day price change rate, the profit is positive from day 15 to 55 (the  $p$ -values are below 0.02), the profit is larger than that of the random strategy from day 15 to day 40 (the  $p$ -values are below 0.1), and is



**Figure 7: Performance of the enhanced strategy. In the legend of each setting, the number in the parentheses represents the number of trades.**

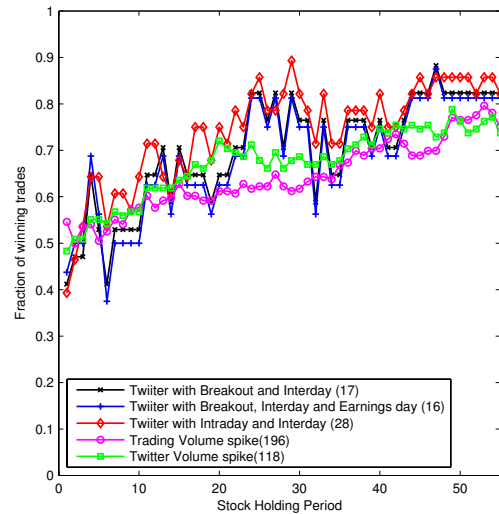
larger than that of the strategy using stock trading volume spikes from day 15 to around day 35 (the  $p$ -values are below 0.1). Fig. 8 plots the number of trades in each month when using the enhanced strategy. We can see that the trades spread in five months' testing period (no trades in March 2013), instead of in a particular month.



**Figure 8: Number of trades in each month when using the enhanced strategy (the features are breakout point and interday price change rate),  $K = 3$ .**

Fig. 9 plots the fraction of the winning trades (i.e., those that lead to profit) under the enhanced strategy, when the holding period,  $\tau$ , is varied from 1 to 55 trading days. The results for three sets of features are plotted in the figure. We observe that significant fraction of the trades lead to profit. For instance, when using intraday and interday price change rates as features, 89.3% of the trades lead to profit in 29 days. We also plot the results when only using Twitter volume spikes (i.e., without using any feature) and when

using stock trading volume spikes; both show inferior performance compared to the enhanced strategy.



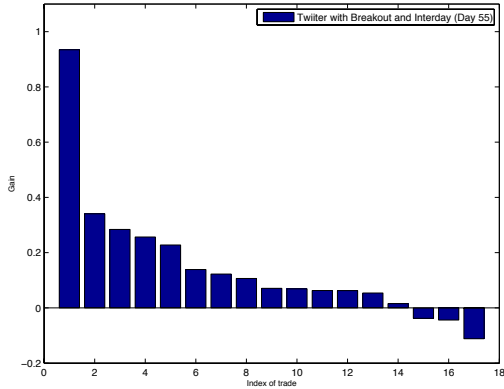
**Figure 9: Fraction of the winning trades made using the enhanced strategy,  $K = 3$ .**

Last, as an example, we present more detailed results using the enhanced strategy and the features are breakout point and interday price change rate. Fig. 10 plots the profits of the trades in decreasing order (a negative value indicates a loss in money) when the holding period,  $\tau$ , is 55 trading days. We see 14 out of the 17 trades lead to profit, and the highest gain is 93.4%. Table 3 shows the detailed information of the 17 trades, including the purchase date, purchase price, tweets ratio (i.e., the ratio of the number of tweets in the Twitter volume spike over the average number of tweets in the past 70 days), the gain when  $\tau = 55$ , the highest gain and the corresponding  $\tau$ . We see the highest gains of all the trades are

**Table 2:**  $p$ -values of the t-tests that compare the profit of the enhanced strategy (for three sets of features) with 0, with the profit using the random strategy, and with the profit using the strategy that is based on stock trading volume spikes.

$\tau$	Breakout and Interday			Breakout, Interday and Earnings day			Intraday and Interday		
	Random	0	Trading vol. spikes	Random	0	Trading vol. spikes	Random	0	Trading vol. spikes
5	0.644	0.372	0.448	0.736	0.351	0.424	0.382	0.083	0.144
10	0.267	0.250	0.410	0.428	0.308	0.466	0.201	0.043	0.122
15	0.084	0.017	0.096	0.311	0.026	0.123	0.039	0.004	0.071
20	0.097	0.012	0.065	0.093	0.020	0.091	0.072	0.000	0.012
25	0.061	0.003	0.029	0.062	0.005	0.036	0.014	0.000	0.002
30	0.037	0.006	0.057	0.055	0.007	0.059	0.012	0.000	0.009
35	0.059	0.014	0.080	0.143	0.019	0.132	0.026	0.001	0.059
40	0.079	0.018	0.158	0.181	0.032	0.219	0.053	0.003	0.158
45	0.107	0.010	0.128	0.093	0.019	0.184	0.059	0.001	0.123
50	0.148	0.017	0.178	0.281	0.030	0.242	0.086	0.002	0.164
55	0.118	0.009	0.188	0.269	0.017	0.253	0.150	0.002	0.263

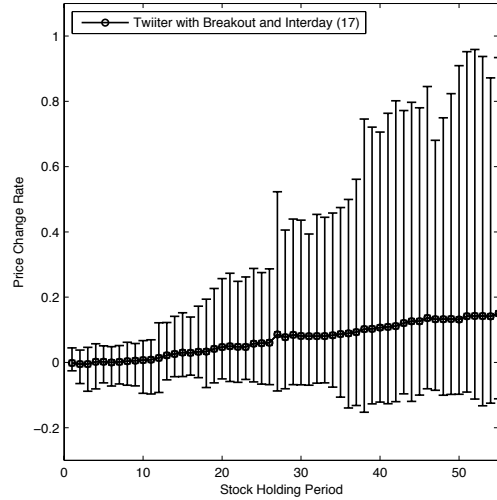
positive. The stock of MHFI is purchased twice, on 2/13/13 and 11/8/12. Fig. 11 plots the average, maximum and minimum price change rates for each value of  $\tau$  when varying  $\tau$  from 1 to 55 trading days. Of all the trades, the largest profit is 95.9%, obtained by purchasing the stock of FSLR (First Solar, Inc.) and  $\tau = 52$  trading days. The lowest profit is  $-15.3\%$  (i.e., loss of 15.3%), caused by purchasing the stock of CF (CF Industries Holdings, Inc.) and  $\tau = 38$  trading days. On the other hand, we see from Table 3 that the highest gain of the trade of CF stock is nonetheless positive.



**Figure 10:** Gains of the trades made using the enhanced strategy, where the features are breakout point and interday price change rate, the holding period  $\tau$  is 55 trading days, and  $K = 3$ .

## 5. CONCLUSION & FUTURE WORK

In this paper, we have investigated Twitter volume spikes related to S&P 500 stocks, and whether they are useful for stock trading. Through correlation analysis, we provide insight on when Twitter volume spikes occur and possible causes of these spikes. Moreover, we explore whether these spikes are surprises to market participants by comparing the implied volatility before and after these spikes. After that, we develop two trading strategies that use Twitter volume spikes, one is a basic strategy based on Bayesian classifier and the other is an enhanced strategy that combines the Bayesian classifier and a stock bottom picking method. Simulation over a half year's stock market data demonstrates that both strategies lead



**Figure 11:** Average, maximum (top bar) and minimum (bottom bar) price change rates of the trades for each value of  $\tau$ . The results are for the enhanced strategy when the features are breakout point and interday price change,  $K = 3$ .

to substantial profits, and the enhanced strategy significantly outperforms the basic strategy and a bottom picking method that uses trading volume spikes.

As future work, we are investigating in several directions to improve our trading strategies: (1) considering more sophisticated Twitter volume spike metric (e.g., based on the number of users instead of only the number of tweets), (2) adding more features to the Bayesian classifier, (3) learning  $\lambda$  from stock data instead of using a fixed  $\lambda$ . We will build an online system that implements the trading strategies. The results will be shown at <http://www.finstats.com>.

## Acknowledgements

The work was partially supported by NSF CAREER award 0746841 and NSF Grant No. 1128581. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors would also like to thank



**Table 3: Summary of the 17 trades made using the enhanced strategy when the features are breakout point and interday price change,  $K = 3$ .**

Stock ticker	Purchase date	Purchase price	Tweets ratio	Gain when $\tau = 55$	Highest Gain (corresponding $\tau$ )
FSLR	2/28/13	25.84	3.68	93.42%	95.90% (52)
CELG	11/12/12	75.66	3.10	34.07%	34.07% (55)
TSO	1/10/13	42.85	3.08	28.40%	37.62% (35)
DTV	2/19/13	49.26	3.19	25.66%	25.76% (54)
MHFI	2/13/13	44.33	3.52	22.78%	22.78% (55)
M	12/28/12	37.36	4.76	13.89%	13.89% (55)
MHFI	11/8/12	51.11	5.01	12.27%	12.27% (55)
KSS	1/4/13	42.23	4.77	10.66%	16.79% (48)
FDO	1/4/13	56.65	19.03	7.08%	8.33% (53)
DVN	11/8/12	54.02	4.09	6.90%	6.90% (55)
GME	1/4/13	24.8	4.73	6.29%	8.10% (21)
YUM	2/6/13	62.93	7.72	6.25%	14.32% (35)
EXPE	10/26/12	59.06	6.55	5.38%	9.63% (47)
SHLD	12/11/12	43.5	3.77	1.54%	11.59% (46)
T	10/25/12	34.5	3.44	-3.77%	2.58% (47)
CF	2/21/13	203.93	4.40	-4.39%	2.07% (12)
JDSU	1/31/13	14.51	6.60	-11.16%	6.69% (10)

the anonymous reviewers for their insightful comments.

## 6. REFERENCES

- [1] Yahoo! finance. <http://finance.yahoo.com/>.
- [2] H. Achrekar, A. Gandhe, R. Lazarus, S. H. Yu, and B. Liu. Twitter improves seasonal influenza prediction. In *Proc. of the 5th Annual International Conference on Health Informatics, HEALTHINF'12*, pages 61–70, Vilamoura, Algarve, Portugal, 2012.
- [3] S. Asur and B. Huberman. Predicting the future with social media. In *Proc. of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Volume 01, WI-IAT '10*, pages 492–499, Toronto, Canada, 2010.
- [4] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, pages 145–166, 1973.
- [5] J. Bollen, H. Mao, and X. Zeng. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1):1 – 8, 2011.
- [6] B. A. Huberman, D. M. Romero, and F. Wu. Social networks that matter: Twitter under the microscope. *First Monday*, pages 1 – 9, 2009.
- [7] A. Java, X. Song, T. Finin, and B. Tseng. Why we Twitter: understanding microblogging usage and communities. In *Proc. of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis, WebKDD/SNA-KDD '07*, pages 56–65, San Jose, CA, 2007.
- [8] B. Krishnamurthy, P. Gill, and M. Arlitt. A few chirps about Twitter. In *Proc. of the 1st Workshop on Online Social Networks, WOSN '08*, pages 19–24, Seattle, WA, 2008.
- [9] H. Mao, S. Counts, and J. Bollen. Predicting financial markets: Comparing survey, news, Twitter and search engine data. *arXiv:1112.1051*, 2011.
- [10] Y. Mao, B. Wang, W. Wei, and B. Liu. Correlating S&P 500 stocks with Twitter data. In *Proc. of the 1st ACM International Workshop on Hot Topics on Interdisciplinary Social Networks Research, HotSocial'12*, pages 69–72, Beijing, China, 2012.
- [11] E. J. Ruiz, V. Hristidis, C. Castillo, A. Gionis, and A. Jaimes. Correlating financial time series with micro-blogging activity. In *Proc. of the 5th ACM International Conference on Web Search and Data Mining, WSDM '12*, pages 513–522, Seattle, WA, 2012.
- [12] T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes Twitter users: real-time event detection by social sensors. In *Proc. of the 19th International Conference on World Wide Web, WWW '10*, pages 851–860, Raleigh, NC, 2010.
- [13] X. Zhang, H. Fuehres, and P. A. Gloor. Predicting Stock Market Indicators Through Twitter "I hope it is not as bad as I fear". *Procedia - Social and Behavioral Sciences*, 26(0):55 – 62, 2011.