Social Media and News Sentiment Analysis for Advanced Investment Strategies

Steve Y. Yang and Sheung Yin Kevin Mo

Abstract The motivation of this chapter hinges on the growing popularity in the use of news and social media information and their increasing influence on the financial investment community. This chapter investigates the interplay between news/social sentiment and financial market movement in the form of empirical impact. The underlying belief is that news and social media influence investor sentiment, which in turn drives financial decisions and predicates the upward or downward movement of the financial markets. This book chapter contributes to the existing literature of sentiment analysis in the following three areas: (a) It provides a review of existing findings about influence of social media and news sentiment to asset prices and documents the persistent correlation between media sentiment and market movement. (b) It shows that abnormal news sentiment can be a predictive proxy for financial market returns and volatility, based on the intuition that extreme investor sentiment changes tend to have long and last effects to market movement. (c) It presents a number of approaches to formulate investment strategies based on the sentiment trend, shocks and feedback strength. The results show that the sentiment-based strategies yield superior risk-adjusted returns over other benchmark strategies. Altogether, this chapter provides a framework of existing empirical knowledge on the impact of sentiment on financial markets and further prescribes advanced investment strategies based on sentiment analytics.

Keywords Sentiment analysis · Financial investment community · Genetic algorithms · Investment strategies

S.Y. Yang (✉)
Faculty of Financial Engineering, School of Systems & Enterprises,
Stevens Institute of Technology, Hoboken, USA
e-mail: steve.yang@stevens.edu

S.Y.K. Mo
Faculty of Financial Engineering, Stevens Institute of Technology, Hoboken, USA
e-mail: smo@stevens.edu

W. Pedrycz and S.-M. Chen (eds.), *Sentiment Analysis and Ontology Engineering*,
Studies in Computational Intelligence 639, DOI 10.1007/978-3-319-30319-2_11

1 Introduction

With increasing digitization of textual information and computation capability, large-scale and sophisticated sentiment analysis has become a feasible alternative for leveraging the use of computational intelligence technologies to advance understanding of the financial markets. The field of behavioral finance, in particular, studies how psychology and cognition influence decision-making of real-world investors in an irrational manner. Past studies have relied heavily on the impact of textual form of financial documents such as news and press releases. Psychological evidence suggests that sentiment, emotion and mood play a key role in affecting investors when making financial decisions [9, 12, 17, 35]. Barberis et al. developed a theory of investor sentiment to illustrate the impact of investor overreaction and underreaction to public information on generating on post-earnings announcement drift, momentum, long-term reversals and predictive power or scaled-price ratio [7]. Daniel et al. further enriched the idea of investor sentiment with private information leading to overconfidence [14, 15]. On the empirical front, a number of studies found different measures of investor sentiment significant in explaining asset price and volatility movements. Chopra et al. showed that prior losing portfolios significantly outperform prior winning portfolio by 5–10 % annually during the next 5 years, validating the overreaction effect [11]. La Porta et al. also displayed evidence that the correction of the extreme investor sentiment tends to revert during earnings announcements when investors realize their initial beliefs were too extreme [27, 36, 43]. These studies are instrumental in demonstrating the existence of investor sentiment along with its impact on the financial markets. The motivation of this book chapter rests on the thesis that news and social media sentiment reflect the societal states that affect individual investors to react and therefore predicates upward or downward movement of the financial markets. In addition, this chapter aims to investigate the interplay between social media/news sentiment and financial market movement, and to demonstrate how the existing findings can be leveraged with computational intelligence techniques to develop advanced investment strategies. This book chapter contributes to the existing literature of sentiment analysis in the following three areas:

1. It provides a review of existing findings about influence of social media and news sentiment to asset prices and documents the persistent correlation between media sentiment and market movement. Furthermore, it shows the presence of a financial community on Twitter whose primary interests are consistently aligned with financial market-related knowledge and information. By harnessing the sentiment expressed by the influential Twitter users within the community, one can construct a better proxy for quantifying social sentiment.
2. It shows that abnormal news sentiment can be a predictive proxy for financial market returns and volatility, based on the intuition that extreme investor sentiment changes tend to have long and last effects to market movement.
3. It presents a number of approaches to formulate investment strategies based on the sentiment trend, shocks and feedback strength. The results show that the sentiment-based strategies yield superior risk-adjusted returns over other bench-

mark strategies. Altogether, this chapter provides a framework of existing empirical knowledge on the impact of sentiment on financial markets and further prescribes advanced investment strategies based on sentiment analytics.

2 Market Sentiment Analysis

This section provides a review of existing literature about common sentiment analysis techniques among financial studies. Market sentiment algorithms can be categorized into two major groups in lexicon-based approach and machine learning techniques. At the end of the literature review, two additional innovative approaches are presented for extracting market sentiment in using social media financial community and media sentiment feedback.

2.1 Lexicon Based Sentiment Approach

The lexicon-based approach refers to the use of specific sets of vocabulary to identify semantic orientation of a textual source, which has been widely used in recent research studies [30]. Moreo et al. proposed a lexicon-based news sentiment analyzer that incorporates non-standard language and generates sentiment measures based on specific topics of interest [34]. A study conducted by Schumaker et al. investigated the effectiveness of the Arizona Financial Text system, which leverages the use of OpinionFinder in identifying the tone and polarity of the underlying text [37]. In addition, Li et al. evaluated financial news articles with a lexicon-based approach using the Harvard psychological dictionary and Loughran-McDonald financial sentiment dictionary for sentiment generation [29]. Other related studies emphasis the use of emotional words such as “soar” and “fall” to enhance the sentiment measure of news articles [48].

The lexicon-based approach is popular among academic and industry studies because of its direct usage of specialized word dictionaries to generate relevant sentiment scores. A common source of reference is the SentiWordNet dictionary which is a lexical resource with words linked to sentimental scores [5]. Our studies performed in later sections also rely on the application of the lexicon-based sentiment approach. Through a four-step process, the objective of the sentiment algorithm is to convert raw text data into daily sentiment score for the empirical study. With the complex textual structure, the raw text is initially decomposed into individual words with the removal of stop words such as “a”, “but”, “how” and “to”. Lemmatization techniques are then applied to convert different inflicted forms of a word into a uniform entity. For instances, the words “rising”, “risen” and “rises” are regarded as the entity “rise”. For each word entity in the textual data, the algorithm extracts the associated score from the sentiment dictionary and finally, generates the sentiment score for each news text by averaging all individual word scores.

2.2 Machine Learning Based Sentiment Approach

Related to financial news mining, machine learning has become more popular as a feasible approach to extract text sentiment. In a recent study, Tan et al. showcased a sentiment mining analyzer based on machine learning techniques using polarity lexicon [21]. Xie et al. illustrated a novel approach in quantifying news document in semantic tree structures and then using tree kernel support vector machines to predict stock market movement [44]. In addition, the five-step procedure was illustrated related to the computerized techniques for handling news content [23] (see Fig. 1).

2.3 Social Media Financial Community Based Sentiment Approach

Apart from the mainstream sentiment analysis approaches, financial sentiment can be extracted from the social media messages of key influencers among the social media financial community. This novel approach has been developed and implemented in our recent publications [33, 46]. It performs sentiment analysis on social media

Fig. 1 Computerized news handling techniques

messages according to their relevance to key financial topics. The approach can be further described in the following components:

(a) **Financial Entity Matching**

A natural processing step of examining Twitter data is to pinpoint those messages that are related to the financial market topics. The rationale behind this critical step is to ensure that the level of noise is minimized in the study. We propose an approach to first form a list of financial entities by extracting commonly used languages from the top financial news broadcasters and traders' Twitter accounts, and then matching them against the individual words and phrases in the tweet message. Furthermore, each financial entity is quantitatively assigned a score to reflect its proximity to the financial related topics.

- ***Entity Extraction from top financial Twitter users***: Sentiment analysis does not perform well if members of the financial community tweet messages that are unrelated to financial market. We conducted an entity matching process to extract messages with financial interests. A financial keyword corpus was then created using a sample dataset from 10/05/2013 to 02/05/2014 from the top financial news broadcasters and traders' Twitter accounts. The sample dataset was categorized according to the type of message initiator. Messages from the top news agencies and traders were labeled as "key messages" while messages from less important community users were "noise messages". By running text parsing and text filtering processes for the two groups, two respective keyword lists are obtained with frequency of occurrence reported. Through text parsing, stop words were dropped and only nouns and noun phrases were extracted from tweet messages. Entities that demonstrated higher occurrence in the "key messages" group were preserved to form the entity corpus while others are discarded. This step effectively screens for financial entities that are more commonly mentioned through established financial Twitter accounts. In addition, the corpus pairs each entity with a weight based on the frequency of occurrence. Higher weight can be interpreted as closer relevance to financial market topics.
- ***Company name and ticker***: Another source of financial entity identification is the name and ticker symbol of major U.S. domestic companies. If a specific company or its ticker is mentioned, there is a high likelihood that the message contains financial-related information. As a result, the financial entity dictionary contains the names and their ticker symbols for over 6,000 companies listed on the three largest U.S. exchanges: the New York Stock Exchange (NYSE), the NASDAQ stock market (NASDAQ) and the American Stock Exchange (AMEX).
- ***Financial entity score computation***: All words and phrases from the tweet message are matched against the financial entity dictionary. Each message contains a score based on the degree of relevance to financial topics. The higher the financial entity score is, the heavier its weight counting towards its sentiment is. For instance, a tweet message containing the ticker symbol for General Electric 'GE' has a financial entity score of $+1$. If multiple financial entities are matched in one Twitter

message, the financial entity score weight is equal to the highest weight among the matched financial entities (1).

$$S^i_{entity} = \max\left(\omega\left(W^i_{message} \cap W_{fe}\right)\right) \quad i = 1, 2, \ldots, \mathrm{N} \tag{1}$$

where S^i_{entity} is the financial entity score of message i, $W^i_{message}$ is the word set split from message i, W_{fe} is the financial entity word set, $\omega(W)$ is the financial entity weights set of word set W, N is the number of message and $n_{matched}$ is the number of matched financial entity.

(b) **Message Sentiment Computation**

With the number of word occurrence and negative flag detection, the sentiment of a single tweet message can be generated. The message sentiment score ranges from -1 to 1, with -1 indicating the most negative sentiment, 0 being neutral, and $+1$ the most positive sentiment. The formula for computing sentiment score of message I can be found in (2). Among all the messages initiated by the top 2,500 critical nodes, the sentiment distribution approximately follows a normal distribution.

$$S^i_{sentiment} = \frac{\sum_j n^j_i \times s(j)}{\sum_j n^j_i} \times S^i_{entity} \times sgn(i)$$

$$sgn(i) = \begin{cases} -1 & if\ W^i_{message} \cap W_{neg} \neq \emptyset \\ 1 & others \end{cases} \tag{2}$$

where $S^i_{sentiment}$ is the sentiment score of message S^i_{entity} is the financial entity score of message i, $W^i_{message}$ is the word set split from message i, W_{neg} is the negative connotation word set, n^j_i is the number of occurrence of SentiWordNet word j in message i, $s(j)$ if the sentiment score of word j (Fig. 2).

(c) **Sentiment Daily Score Computation**

The last procedure in the sentiment analysis algorithm is to compute the sentiment score for the day of observation for a given user. Three components are factored into the computation process: the score of the financial entity, the message sentiment score and the centrality score for the message initiator (3). The algorithm then takes the average of the active user sentiment generated for each day to generate the daily sentiment score for regression studies.

$$S(t) = \frac{1}{N} \sum_{j=1}^{N} W^j_c \times \frac{\sum_{k=1}^{n_j(t)} S^k_{sentiment}(t)}{\sum_{k=1}^{n_j(t)} S^k_{entity}(t)} \tag{3}$$

Data: Twitter Messages
Result: Message Sentiment Score

while *not at the end of document* **do**
  $Message_{raw} \leftarrow$ read current message;
  Step 1: Data-Preprocessing;
  $Message_{splitted} \leftarrow$ Word-splitting(Message);
  $Message_{splitted} \leftarrow$ Remove-stop-words($Message_{splitted}$);
  $Message_{splitted} \leftarrow$ Word-stemming($Message_{splitted}$);
  Step 2: Financial Entity List Processing;
  $WordList_{FinancialEntity} \leftarrow$ Financial Entity List created from SAS;
  $MatchedWords_{FinancialEntity} \leftarrow Message_{splitted} \cap WordList_{FinancialEntity}$;
  $Score_{FinancialEntity} \leftarrow \max\{MatchedWords_{FinancialEntity}\}$;
  if $Score_{FinancialEntity} > 0$ **then**
    Step 3: Sentiment Word List Processing;
    $WordList_{Sentiment} \leftarrow SentiWordNetList$;
    $MatchedWords_{Sentiment} \leftarrow Message_{splitted} \cap WordList_{Sentiment}$;
    $Score_{SentiWordNet} \leftarrow \frac{1}{N} \sum\{MatchedWords_{Sentiment}\}$;
    Step 4: Inverse Factor;
    if *negative words/phrases (not n't) detected* **then**
      $InverseFactor \leftarrow -1$
    else
      $InverseFactor \leftarrow 1$
    end
    Step 5: Final Message Sentiment Score [-1,1];
    $Score_{Sentiment} \leftarrow Score_{FinancialEntity} \times Score_{SentiWordNet} \times InverseFactor$
  else
    Drop $Message_{raw}$; continue;
  end
end

Fig. 2 Sentiment analysis algorithm

where $S(t)$ is the daily sentiment score of day t, ω_c^j is the centrality weight of user j, $n_j(t)$ is the number of message by user j on day t, $S^k_{entity}(t)$ and $S^k_{sentiment}(t)$ are entity score and sentiment score of message k computed by Eq. (10) and Eq. (11) respectively.

2.4 Media Sentiment Feedback Effects

Feedback mechanisms have been explored in the field of finance, mainly through the examination of its effects on price and volatility. Sentiment can be quantified in the form of its feedback effects. Hirshleifer et al. presented a theoretical framework that justifies irrational investors to earn abnormal profits based on a feedback mechanism from stock prices to cash flows [18]. Crude oil prices were found to contain feedback effects along with an inverse leverage impact with its implied volatility [1]. Khanna and Sonti showed the feedback effect of stock prices on firm value through a herding equilibrium model and investigated into the incentive for traders to conduct price manipulation [25]. The volatility feedback effect is an empirical observation of feedback effect between squared volatility and stock price. Inkaya and Okur estimated the volatility feedback effect rate using Malliavin calculus and suggested its predictability of large price declines [22]. They showed that large feedback effect rate is a useful indicator for measuring market stability [22].

There is also empirical evidence that feedback trading, a self-perpetuating pattern of investor's behavior, is present in G7 stock markets, and other international markets [2, 36]. The effect of feedback trading was found to vary across business cycle [10] and the strongest influence was observed during periods of financial crisis with declining futures prices [36]. Hou and Li developed a regression model of feedback trading to analyze CSI300 stock returns and demonstrated that lagged index returns can predict market index return and conditional volatility [20]. In addition, feedback trading was found to significantly influence exchange rate movements [28]. Using a theoretical framework, Arnold and Brunner showed that positive feedback trading causes price overreaction and the impacts of feedback trading would be dampened if news is incorporated into price in time [4].

3 Market Sentiment Based Investment Strategies

This section presents the latest research findings on using market sentiment as a source to develop financial investment strategies. It covers the economic intuition behind the study, the methodology along with the data sources, the key findings and their implications. The first study centers on the use of one of the most popular social media platforms Twitter and how its messages can be used to generate sentiment for predictive signals. Using news data, the second study focuses on firm-specific sentiment and how it is correlated with the financial market returns and volatility. The last study examines the interaction effect between social media and news sentiment and showcases that the feedback effect can be quantified for developing investment strategies using a genetic programming framework. These studies revolve around the central theme of market sentiment based investment strategies and display the respective theoretical basis along with empirical evidence for their practical applications.

3.1 Twitter Financial Community and Its Predictive Relation to Stock Market

Twitter, one of the several major social media platforms, has been identified as an influential factor to financial markets by multiple academic and professional publications in recent years. The motivation of this method hinges on the growing popularity of the use of Twitter and the increasing prevalence of its influence among the financial investment community. We present an empirical evidence of the existence of a financial community on Twitter in which users' interests align with the financial market related topics. We establish a methodology to identify relevant Twitter users who form the financial community, and we also present the empirical findings of network characteristics of the financial community. We observe that this financial community behaves similarly to a small-world network, and we further identify groups of critical nodes and analyze their influence within the financial community based on several network centrality measures. Using a novel sentiment analysis algorithm, we construct a weighted sentiment measure using tweet messages from these critical nodes, and we discover that it is significantly correlated with the returns of the major financial market indices. By forming a financial community within the Twitter universe, we argue that the influential Twitter users within the financial community provide a better proxy between social sentiment and financial market movement. Hence, we conclude that the weighted sentiment constructed from these critical nodes within the financial community provides a more robust predictor of financial markets than the general social sentiment.

Our hypothesis is that Twitter sentiment reflects the market participants' beliefs and behaviors toward future outcomes and the aggregate of the societal mood can present itself as a reliable predictor of financial market movement. However, not all users are equally influential in the social media, and those influential social media users will certainly have higher impact to the societal mood or sentiment. Reported evidence demonstrates that there exists a community on Twitter whose primary concern is about financial investment. Those users who are harvesting information from these influential sources on the social media for their daily trading decisions forms the robust linkage between the social mood and financial market asset price movement. Hence this community would be more representative to market participant's beliefs, and consequently the sentiment extracted from this financial community would serve as a better predictor to the market movement.

We seek to identify the corresponding investment community and pinpoint its major influencers in the social networks context. The primary research question is whether the beliefs and behaviors of major key players in such community reveal better signals to financial market movement. From a large-scale data crawling effort, we define a financial community as a group of relevant Twitter users with interests aligned with the financial market. We first identify 50 well-recognized investment experts' accounts in Twitter and use their common keywords to create the interests of the financial investment community. By constructing the two layers of the experts' followers, we apply a multitude of rigorous filtering criteria to establish a financial community boundary based on their persistent interests in the topic of financial investment (Fig. 3 and Table 1).

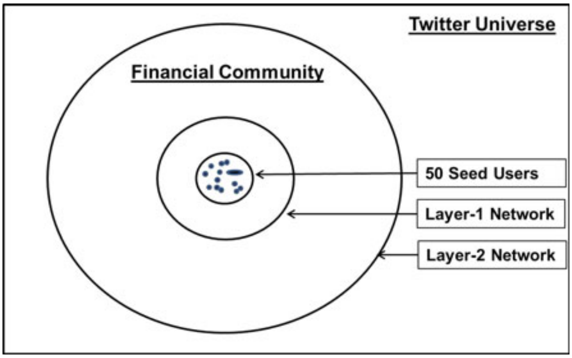

Fig. 3 Financial community from Twitter universe

Table 1 Financial community network summary statistics

Number of nodes	154,327
Number of links	4,846,805
Average out-degree centrality	35.71
Average betweenness centrality	420.2
Clustering coefficient	0.15
Network diameter	6
Average path length	2.72
Connected component	1

3.1.1 Critical Node Analysis in the Financial Community

In the financial community, there exist users who play a central role in the connectedness of the network. These users, known as critical nodes, situate at the most critical locations of the community network and therefore bear a large weight in the network dynamic properties such as connectedness and message propagation pattern. Analyzing these nodes is essential for understanding the financial community because they represent the most influential users in the community in terms of facilitating the message propagation process and stabilizing the network structure.

Through social network analysis, we identify these critical nodes by applying centrality measures: out-degree centrality, betweenness centrality and closeness centrality (4), (5) and (6). Our data captures the direction of the friend-follower relationship which contributes to the formation of a directed network. The three centrality measures incorporate direction as information propagates from the sender to their followers, and, more importantly, capture unique aspects regarding the node's relative centrality level among the community. We track the profile information and tweet messages of the top 2,500 users ranked by each of the three centrality measures. The description of each centrality measure is defined as:

1. **Out-degree centrality** measures the number of followers a node has in the network. A higher out-degree centrality value indicates that the specific node is well-connected to many nodes in the network.

$$C_D(v_i) = \sum_{j=1}^{n} a_{ij} \tag{4}$$

   where A denotes the adjacency matrix, a_{ij} is a binary term that values 1 if node i out-ties with j and values 0 otherwise, and n is the number of nodes in network [8].

2. **Betweenness centrality** captures the number of shortest paths from all vertices to other nodes in the network that passes through the specific node. With a higher number of shortest paths passing through a specific node in the network, its betweenness centrality measure will be higher.

$$C_B(v_i) = \sum_{j \neq i} \sum_{k \neq i} \frac{g_{jik}}{g_{jk}} \tag{5}$$

   where g_{jk} denotes the number of geodesic paths from node j to node k and g_{jik} denotes the number of geodesic paths from node j to node k that pass through node i [8].

3. **Closeness centrality** is calculated based on the aggregation of geodesic distances from each node to all of other nodes in the network (also known as the farness). If a specific node is located at a more central location relative to another node, the farness will be lower because of its shorter distance from all other nodes in the network.

$$C_C(v_i) = \sum_{j=1}^{n} d_{ij} \tag{6}$$

   where C is the geodesic distance matrix, d_{ij} is the geodesic distance between node i and node j, and n is the number of nodes in network [8].

Each group of these critical nodes shares certain common attributes. The profile data consists of the nodes' location, username, description, the number of messages they have posted, and the number of followers and friends. We examine the three important attributes: the number of tweeted messages, the number of followers and the number of friends. Samples of these critical nodes are provided in Table 2.

When we analyzed critical nodes with the highest out-degree centrality, we observed that they were followed by a large portion of the users in the community. A tweet message initiated by this group can be spread to a large domain in the network. Furthermore, nodes with the highest out-degree centrality can be much more influential than other nodes as their large number of followers can gain significant interest from the financial community and therefore attract more users who follow them. From the profile dataset, these top nodes with the highest out-degree centrality are broadcasters who have posted more than 10,000 tweets over the

Table 2 Critical node sample users

Critical nodes	Sample users
Out-degree centrality	@TheEconomist, @BreakingNews, @FinancialTimes, @FortuneMagazine, @CMEGroup
Betweenness centrality	@themotleyfool, @Vanguard_Group, @ReformedBroker, @TheStreet, @NYSEEuronext
Closeness centrality	@YESBANK, @currency4trades, @QNBGroup, @Bizzun, @FFinancialGroup, @sobertrader

lifetime of the account, a much higher tweeting rate than that of normal broadcasters and the community. They also appeared to have a longer account history compared to the other two groups of critical nodes.

As an illustration of the importance of centrality measures related to the community structure, we are interested in exploring how the current financial community compares against sample community structures. We select four sample communities to showcase extreme scenarios with distinct characteristics. We then calculate the normalized centrality score for every community node and compute the average among all the node scores to gauge the connectedness of the community (see Fig. 4). Figure 4a shows a symmetric structure with two nodes as the central hub. The peripheral nodes are directly attached to the center of the community and therefore the overall structure has a relatively high average closeness centrality score. Figure 4b features a similar structure with one central node with more connections. This biased phenomenon results in a lower average centrality scores across all three measures. Figure 4c is another symmetric structure featuring three nodes as the central hub and higher average centrality scores. In contrast, Fig. 4d is a fully connected community with each node linked to all other nodes. This structure is an ideal framework for message propagation as the distance between any two nodes in the community is 1. Therefore, the betweenness centrality and the closeness centrality for all nodes is 0 and 1, respectively. In summary, these sample communities feature distinct network structures in terms of key centrality measures. For the financial community, the average scores of out-degree centrality, betweenness centrality and closeness centrality are 0.720, 0.005 and 0.360, respectively. These scores are normalized from 0 to 1 to facilitate the comparison with the sample structures. The high average out-degree centrality score illustrates that the financial community contains more direct linkages among nodes, while the low average betweenness centrality score shows that the central hub among the network is widely spread among many nodes. In addition, the average closeness centrality score reflects that a substantial portion of the community users are close to the center of the network. These observations show that the financial community is more densely connected at the local level. In addition,

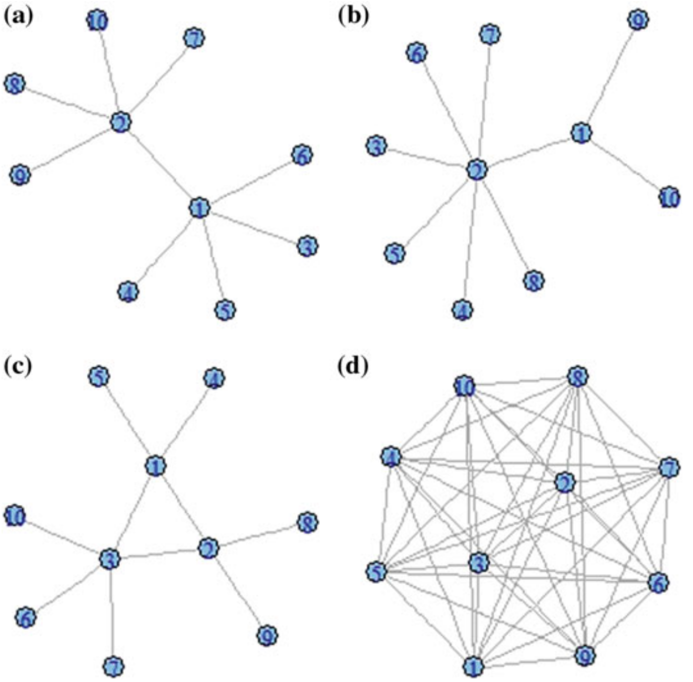

Fig. 4 Financial community network topology. **a** avg_dc = 0.36 avg_bc = 0.20 avg_cc = 0.70. **b** avg_dc = 0.26 avg_bc = 0.15 avg_cc = 0.62. **c** avg_dc = 0.40 avg_bc = 0.24 avg_cc = 0.71. **d** avg_dc = 1.00 avg_bc = 0.00 avg_cc = 1.00

the connectivity is not strong at the global level but there is evidence of clustering around the center of the network.

Based on the location data, we extract the top 2,500 users with the highest betweenness centrality from the financial community in the U.S. and map their population density (see Fig. 5). New York and California are identified as the top two states containing the largest number of critical nodes, a fact reflective of their wealth distribution and status as financial and media centers. The next level consists of the states of Massachusetts, Texas, Illinois and Florida. These are among the U.S. states with the largest and wealthiest population. It is not surprising that some of their cities, such as Boston and Chicago, serve as major hubs of the U.S. financial system. Lastly, the final level comprises mostly states on the east coast like Pennsylvania, Virginia, Georgia, New Jersey, Maryland and District of Columbia. They tend to be situated with close proximity to New York City and have significant business ties with regard to the number of financial corporation headquarters. The population map of the critical nodes has two significant interpretations: First, it reflects to a certain degree where the most influential nodes in the financial community are most likely located. Second, their tweeting activities might have direct implications on the location of

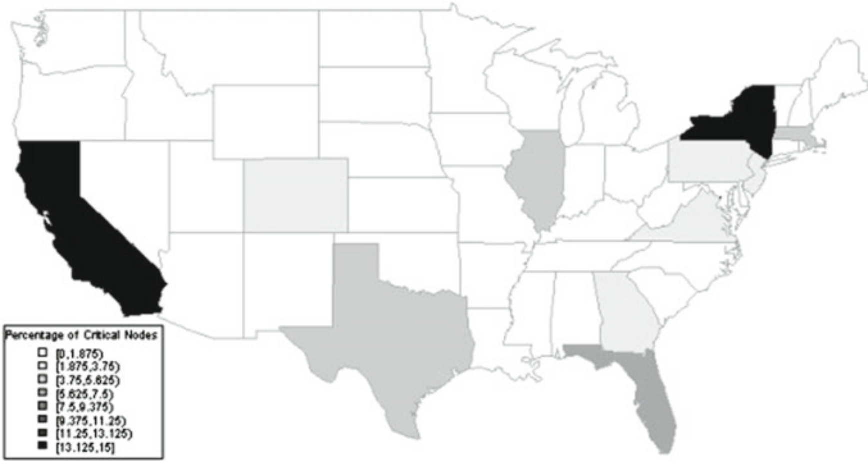

Fig. 5 Critical nodes location in financial community

the events. Knowing the location of the source provides a competitive advantage in tracking the scope of the events.

After settling on a definition, we examine how messages from key influencers in the community interact with social mood or sentiment that tend to signal an impending upward or downward swing in the market price movement. We use key network metrics such as out-degree centrality and betweenness centrality to identify the financial community influencers and we conjecture that these key influencers along with their weight of their influence in the financial community will provide better predictors of financial market movement measures.

This section demonstrates the value of extracting sentiment based on the social structure of the financial community. A key hypothesis of this approach is that Twitter sentiment extracted from the network's critical nodes serves as a reliable predictor of financial market movement. We believe that not all messages carry equivalent weight of information and therefore the message initiated from a more credible source should have larger influence on the financial community. Through regression analysis, we test whether specific critical nodes in the financial community have predictive power to key financial market measures. From the previous section, we identified three groups of 2,500 critical nodes based on key centrality measures from the financial community: betweenness centrality, out-degree centrality and closeness centrality. Through their tweet messages, we extracted their sentiments using the sentiment analysis algorithm and determined the statistical relationship with the historical daily return of major market returns and volatility indices.

3.1.2 Data Sources

We applied the regression analysis on 1,606,104 tweet messages for the period between 02/15/2014 and 06/15/2014. The daily sentiment series for the three major critical node groups were generated from their messages and then weighed with respect to the normalized centrality measures. In addition, we adopted the returns series for 6 exchange-traded funds which served as proxies for the historical daily market returns and volatility. The daily return is computed by taking the log return of market price from 02/15/2014 to 06/15/2014 (7).

$$r_t = \log(\frac{p_t}{p_{t-1}}) \tag{7}$$

where r_t denotes the return of day t and p_t is the price of day t.

3.1.3 Linear Regression Model of Sentiment

A linear regression model is applied to examine the relationship between the daily index return and the lag-1 Twitter message sentiment from the critical nodes. In particular, we investigated whether the lagged series of message sentiment have significant statistical relations with market returns and volatility. The dependent variable in the regression model is the daily return of respective market index. In this analysis, we used SPY, DIA, QQQ, and IWV as the market returns and VIX as the volatility returns. These market indices represent major distinct components of the financial market by the characteristics of the underlying stock such as market capitalization and industry sector (see Table 3). The independent variable is the lagged time series of the message sentiment from three critical node groups. In the experiment, we test the lag-1 sentiment for their respective significance to the returns series (8). If a significant relationship is observed, it suggests that returns lag behind the sentiment movement and therefore sentiment has predictive property over returns.

$$r_t = \beta_0 + \beta_1 S_{t-1} + \varepsilon \tag{8}$$

Table 3 Exchange-traded funds ticker and corresponding index

Market return	
SPY	S&P 500
DIA	Dow Jones industrial average
QQQ	NASDAQ
IWV	Russell 2000
Market volatility	
VIX	CBOE volatility index

where r_t denotes the daily return of day t and S_{t-1} denotes the daily sentiment score of day $t-1$.

3.1.4 Comparison Among Centrality Groups

It is important to examine whether there is a fundamental difference of the regression result across the three different centrality measures in relation to the financial market returns and volatility. Varying the number of critical nodes in each group, we found that the betweenness centrality (BC) group consistently outperformed the degree centrality (DC) and closeness centrality groups (CC) (see Appendix 4B). The sentiment regression model of the BC group has shown significance across all market returns at the level of 95 %. In addition, the positive coefficients of the model demonstrate the predictive capability that the more positive the message sentiment is, the higher market returns it leads to. For volatility, the betweenness centrality group is also more significant than the two other groups in terms of its significance level. The result shows that more positive sentiment leads to lower volatility level, vice versa. It is consistent with the observation that negative sentiment can cause a higher volatility spike, suggesting that bad news on Twitter increases the volatility of price return in the stock market.

3.1.5 Comparison Among Number of Top Critical Nodes

Along the same intuition that not all messages carry equivalence of information, we investigated whether there is an optimal number of a critical node for each centrality group in explaining financial market movement. For the extreme scenarios, too many critical node users may introduce unnecessary noise but too few users may omit key contributing sentiment for the regression study. Varying the number of critical nodes for each centrality group might yield results that reveal the emerging critical point and, therefore, lead to an enhanced indicator for explaining market movements. In this analysis, we investigate all three centrality measures starting from the top 100 users to 2,500 users in each group at an increment of 100 additional users. For instances, we first examine the top 100 users in the betweenness centrality group and then the top 200 users in the same group. We observe that an optimal point exists when the number of critical nodes in the group is 200 (see Fig. 6). The coefficient for the 200-user group with the highest centrality measure is the most significant and consistent among all market indices. With the incorporation of more critical nodes in the regression model, we find that the p-value remains stabilized under the 0.05 level (except VIX volatility measure) for the models against market returns. This illustrates that our regression result is robust across different number of top critical nodes (Fig. 7).

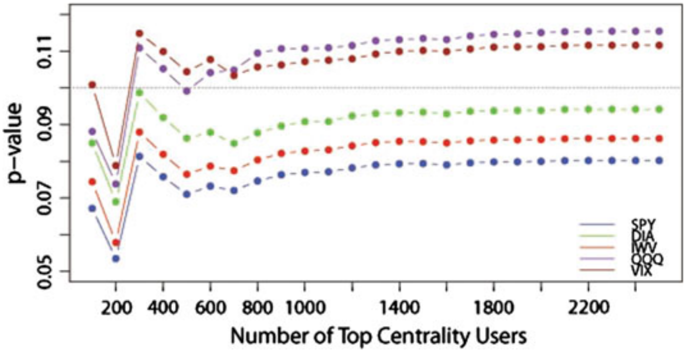

Fig. 6 Different market indices comparison (p-value)

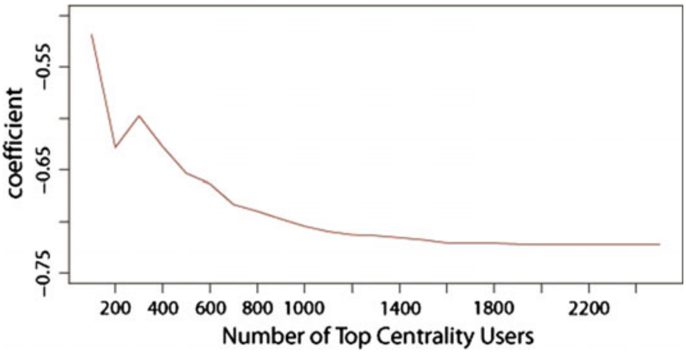

Fig. 7 Different market indices comparison (coefficient)

3.1.6 Discussion

The market sentiment regression highlights the significant influence of critical nodes towards movements in the financial market. Through performance comparison, the sentiment expressed by the betweeness centrality group is more significant and impactful than those by other community members. Critical nodes ranked by betweeness centrality degree yield the highest parameter significance close to the 99 % level. In addition, the BC group also yields the highest positive sentiment coefficients against market return at 0.15^{**}, 0.16^{**}, 0.18^{**} and 0.19^{**} for the top 200, 500, 1000 and 2000 user group respectively. In Table 4, the volatility regression exhibits similar observations of sentiment coefficients at -0.94^{*}, -1.01^{*}, -1.11^{*}, -1.15^{*} with better significance level by the critical nodes ranked by betweenness centrality. This is consistent with the definition of critical nodes because of their central contribution to network connectedness. Grouping users in financial community by centrality analysis transforms the regression result to be more precise and accurate in explaining market movements (Table 5).

Table 4 Lag-1 sentiment linear regression statistics (n = 200)

Market indices	BC score[a]		DC score[b]		CC score[c]	
	Coeff	p-value	Coeff	p-value	Coeff	p-value
SPY	0.15	0.01**	0.15	0.07*	0.11	0.35
DIA	0.12	0.03**	0.13	0.11	0.09	0.41
QQQ	0.17	0.05**	0.16	0.20	0.13	0.46
IWV	0.16	0.01**	0.17	0.07*	0.11	0.38
VIX	−0.94	0.10*	−0.81	0.29	−1.14	0.29

$^{*}p < 0.10$; $^{**}p < 0.05$
[a]Sentiment score weighted by betweenness centrality
[b]Sentiment score weighted by degree centrality
[c]Sentiment score weighted by closeness centrality

Table 5 Lag-1 sentiment linear regression statistics (n = 500)

Market indices	BC score[a]		DC score[b]		CC Score[c]	
	Coeff	p-value	Coeff	p-value	Coeff	p-value
SPY	0.16	0.01**	0.18	0.12	−0.10	0.54
DIA	0.14	0.03**	0.15	0.18	−0.06	0.69
QQQ	0.19	0.05**	0.17	0.30	−0.22	0.35
IWV	0.18	0.01**	0.19	0.11	−0.10	0.55
VIX	−1.01	0.09*	−0.95	0.36	0.91	0.53

Another significant finding is the effect of using a smaller subset of the critical node groups. The tradeoff analysis of varying the number of critical nodes reveals the optimal point in yielding significant signals to financial market returns and volatility. We found that selecting the top 200 users in the betweenness centrality group provide the most significant signal. A group with more than 200 critical node users dilutes the significance of the model but the result remains robust at a high significance level. The systematic search for the optimal number among the key centrality groups reinforces the main principle that not all messages should be treated with equal weight. Lastly, our comparison among different market indices signifies the value of extracting message sentiment based on social structure of the investment financial community. The sentiment series has potential applications on pairs trading strategy across multiple market indices.

We show that the behavior of critical nodes can be used to yield a more reliable indicator for the financial asset's price movement. It is worth noting that the current critical node analysis does not factor in the effect of tweets being read by unregistered users who may be active investors. This channel of information propagation is achieved by searching the Twitter website for specific keywords and the associated tweets would appear regardless whether the users are followers or not. This would facilitate the propagation mechanism of tweets to a wider community, including those tweets broadcasted by the critical nodes. Measuring the impact of this unobservable user group to the financial market will be a challenging problem, and we plan to address this issue in future studies. This empirical study of the financial community has three major contributions to the current literature of financial market and Twitter sentiment. First, it addresses the hypothesis that Twitter sentiment reflects the market participants' beliefs and behaviors toward future outcomes and the aggregate of the societal mood can present itself as a reliable predictor of financial market movement. Second, the concepts of leveraging critical nodes in the financial community generate a robust linkage between the social mood and financial market asset price movement. Moreover, the findings of critical nodes serve as an important guidance for regulatory authorities in paying attention to avoid manipulative or malicious actions such as the 2013 Associated Press hacking incident. Lastly, the empirical study provides insightful observations about the demographics and network structure of the financial community. By decomposing the community into unique user types, beliefs and behaviors of market participants can be better understood. With the continual growth of the Twitter universe, the dynamic characteristics of the financial community can be better understood in terms of its network structure and the message sentiment influence towards the financial market movement.

3.1.7 Conclusion

Based on the intuition that "not everyone on Twitter has the same influence on market sentiment", we document that there exists a financial community on Twitter, and the weighted sentiment of its key influencers has significant predictive power to market movement. In proving this hypothesis, we first document a methodology to identify the financial community, and then we illustrate the key properties of the financial community networks. We also demonstrate that the betweenness centrality measure of the network nodes is a better measure of influence of the nodes of the financial community networks than the other popular influential measures, i.e. the degree centrality and the closeness centrality. We show that different groups of critical nodes exert different degree of impacts on financial asset prices and volatility movement. In conclusion, we document that there is a robust correlation between the weighted Twitter financial community sentiment and financial market movement measured by lagged daily prices. This study covers the major financial market indices such as Dow Jones (DIA), S&P 500 (SPY), NASDAQ (QQQ) and Russell 3000 ETF (IWV), and the model significant levels are all less than 0.05 (p-value). A key finding of this study is that Twitter sentiment generated from critical nodes in the Twitter financial community provides a robust surrogate to predict financial market movement: the weighted sentiment of the critical nodes has significant predictive power over major market returns, and it consistently predicates market volatility (VIX) as well [45].

3.2 News Sentiment as Predictive Proxy to Financial Market

News sentiment has been empirically observed to have significant impact on financial market returns. In this study, we investigate firm-specific news from the Thomson Reuters News Analytics data from 2003 to 2014 and propose an optimal trading strategy based on a sentiment shock score and a sentiment trend score which measure extreme positive and negative sentiment levels for individual stocks. The intuition behind this approach is that the impact of events that generate extreme investor sentiment changes tends to have long and lasting effects to market movement and hence provides better prediction to market returns. We document that there exists an optimal signal region for both indicators. In addition, we demonstrate that extreme positive sentiment provides a better signal than extreme negative sentiment, which presents an asymmetric market behavior in terms of news sentiment impact. The back-test results show that extreme positive sentiment yields superior trading signals across market conditions, and its risk-adjusted returns significantly outperform the S&P 500 index over the same time period.

Many studies have demonstrated that news media can affect financial markets and often becomes drivers of market activities [3, 6, 7, 31, 32, 38, 41]. Analyzing news contents and translating them into trading signals have become an attractive research topic in both academia and industry. There have been a number of studies that further document the value of using media sentiment to make trading decisions [13, 24, 40, 49]. The motivation of this study is based on recent findings that news content affects investor sentiment and market volatility [6, 16, 39, 42]. We propose a trading strategy based on extreme news sentiment levels on individual stocks, and we further explore the effect of a long and short strategy based on extreme positive and negative sentiment on these stocks.

3.2.1 Data Source

This study utilizes the Thomson Reuters News Analytics package as the sole financial news data source. The package is used to quantify individual news events into sentiment, and its numerical form is then supplied for the trading system. The duration of the data ranges from January 2003 to December 2014. With more than 80 metadata fields in the Thomson Reuters News Analytics package, the corresponding fields are used in this study below.

- **Date/Time**: The date and time of the news article.
- **Stock RIC**: Reuters Instrument Code (RIC) of the stock for which the sentiment scores apply.
- **Sentiment Classification**: An integer number indicate the predominant sentiment value for news with respect to a stock identified by the RIC. Possible values are 1 for positive sentiment, 0 for neutral and -1 for negative sentiment.
- **Sent_POS**: Positive Sentiment Probability, the probability that the sentiment of the news article is positive for the stock. The possible value ranges from 0 to 1.

- **Sent_NEUT**: Neutral Sentiment Probability, the probability that the sentiment of the news article is neutral for the stock. The possible value ranges from 0 to 1.
- **Sent_NEG**: Negative Sentiment Probability, the probability that the sentiment of the news article is negative for the stock. The possible value ranges from 0 to 1. The sum of the three probabilities (Sent_POS, Sent_NEUT, Sent_NEG) equals 1.
- **Relevance**: A real-valued number between 0 and 1 indicating the relevance of the news item to a stock. A single news article may refer to multiple stocks, by comparing the number of occurrences within the text, the stock with the most mentions will be assigned with the highest relevance, and a stock with a lower number of mentions will have a lower relevance value.

In order to calculate a sentiment score for each stock mentioned in one news item, we first calculate the expected value of the sentiment score, and then generate the weighted expected value using its relevance value. Finally, the weighted weekly average sentiment score is calculated as follows:

$$Avg\_Sent = \frac{1}{N}\sum\nolimits_N \begin{array}{c} ((Sent\_POS \times 1 \\ +Sent\_NEUT \times 0 \\ +Sent\_NEG \times (-1)) \\ \times Relevance), \end{array} \tag{9}$$

where N is the total number of new articles for a stock within one week. The weighted weekly average sentiment is later used as the input for computing the other two sentiment scores.

The summary statistics of the news sentiment data, including the mean, standard deviation, maximum/minimum, and 5th/50th/95th percentile of each variable, is displayed below (see Table 8). We also plot the monthly aggregated average news sentiment for all 596 stocks, the total number of news articles (hereinafter "number of news") for each month, and the S&P 500 index monthly return (see Fig. 8). The data indicates that the average news sentiment is positively correlated with market return with a correlation coefficient of 0.21, while the total number of news is negatively correlated with market return with a correlation coefficient of −0.14. Through conducting a lead-lag analysis, the news sentiment is shown to lead the market return, while there is no opposite effect from market return on future news sentiment (Table 6).

Table 6 Statistics of calculated average news sentiment

	Mean	STD.	Max	Min	5 %	50 %	95 %
Average sentiment	0.09	0.24	0.83	−0.78	−0.20	0.00	0.60
Number of news items	5.18	11.69	830	0	0	2	22

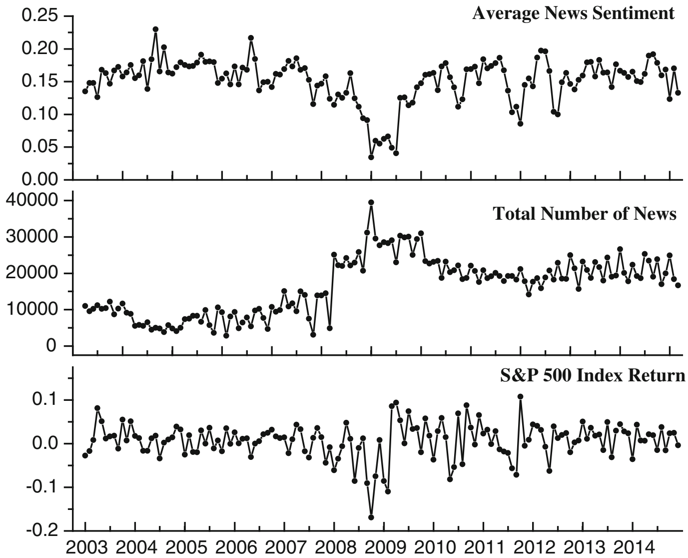

Fig. 8 Monthly aggregated news data comparison with market returns

3.2.2 Sentiment Shock and Trend Scores

With the preliminary observations on correlation, we explore the abnormal levels of the news sentiment data as a way to create more insightful and unique indicators for the trading system. We propose two sentiment scores to characterize shocks (i.e. spike up or down) and trends in the sentiment time series. The calculation is based on the average weekly news sentiment scores for each stock. In order to reduce the number of parameters in the trading strategy and avoid over-fitting, we optimize the calculation parameters for each GICS sector, so that all stocks within the same sector use the same parameter. The trading strategy is designed to monitor the calculated sentiment scores and generate buy-and-sell signals for each stock.

Sentiment shocks are the abnormal spikes observed from the time series. These sentiment shocks are often caused by the release of unexpected macroeconomic data, financial report results, and corporate actions. The sentiment shock score is calculated as below:

$$(S_{t0} - \mu)/\sigma, \tag{10}$$

where S_{t0} is sentiment value on week $t0$, $t0$ represents the current week, μμ is the mean of sentiment values from week $t0 - N$ to $t0 - 1$$S_{t0-N}$ to S_{t0-1}, and σσ is the standard deviation of sentiment values from week $t0 - N$ to $t0 - 1$. N is the total number of look-back weeks.

Sentiment trend score measures the aggregated change of sentiment over a historical period. The sentiment trend measure reveals more information than the sentiment shock approach when comparing abnormal changes of sentiment over a long period of time.

$$\sum_{i=t0-N}^{t0} \Delta S_i \tag{11}$$

where ΔS_i is the change of sentiment in week i, and represents the current week. N is the moving window size, summing the change of sentiment within it.

3.2.3 Parameters Optimization

Each of the sentiment shock or trend score has a parameter N (the look-back window) to choose. To find the best parameter, we use Spearman rank correlation as the objective value. In order to reduce the number of parameters and avoid over-fitting, N is optimized for each GISC sector, and stocks in the same sector use the same value. The method we use to optimize these parameters is to maximize the Spearman rank correlation between the sentiment scores and the next week's stock return. The Spearman rank correlation is a measure of rank dependence between two variables. For a sample of size n, the two variables X_i, Y_i are converted to ranks x_i, y_i, the correlation coefficient is computed as:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{12}$$

where $d_i = x_i - y_i$. By maximizing the rank correlation, the calculated sentiment scores are most informative for future stock return.

3.2.4 Trading Strategy Construction

Using the optimized look-back windows for each sector, the time series of sentiment shock and sentiment trend scores are constructed for each stock from 2003 to 2014. The design of our trading strategy is based on the hypothesis that extreme sentiment has a persistent effect on subsequent stock returns. Therefore, the trading strategy involves establishing long positions on stocks with unusually high positive sentiment scores and vice versa.

In terms of the trading strategy parameters, the cutoff percentile between extreme sentiment score versus normal score is defined according to the empirical probability distribution of sentiment scores during the training period (see Fig. 9). A threshold of bottom 5 % means the sentiment score in the 5 % bottom percentile in the training data is the break point of extreme negative sentiment, and scores lower than that threshold are considered as short signals. As suggested in Fig. 9, the majority of the sentiment scores are centered on zero. With the selection of larger thresholds, the sentiment value diminishes quickly and only the extreme values are of significance in providing signals. This effect is particularly pronounced in the sentiment trend

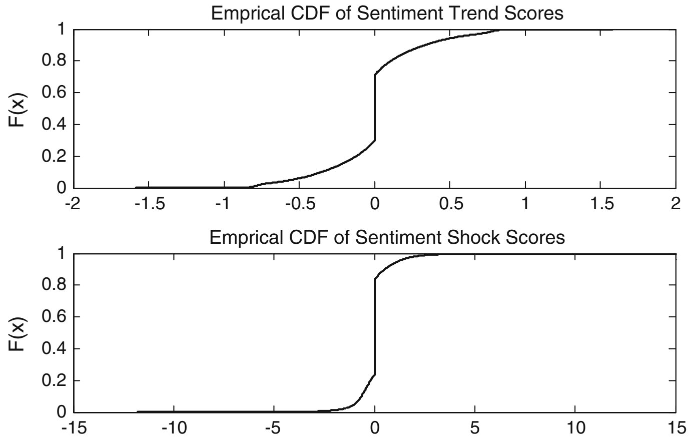

Fig. 9 Empirical CDFs of sentiment trend and sentiment shock scores

score, as it has a broader non-zero region in the cumulative distribution function than that in the sentiment shock score. We will further discuss the optimal selection criterion for this threshold next (Table 7).

3.2.5 Trading Strategy Implementation

We design a dynamic trading framework with an evaluation period of 4 weeks (see Fig. 10). We determine the threshold of extreme sentiment with 90 % as positive threshold and 10 % as negative threshold. Each firm-specific sentiment score is then evaluated and compared with the threshold to make trading decision. If the firm's sentiment exceeds the positive threshold, we establish a long position or vice versa. The portfolio is rebalanced every 4 weeks, with the risk control of a 10 % stop loss limit order in place. In the trading system, the strategy return is recorded weekly. During the training process, we use 4 years of data from 2003 to 2006 as the training period so that sufficient out-of-sample data is available including the 2008 financial crisis. Table 8 summarizes the optimized look-back windows of sentiment indicators for each sector (Table 9).

3.2.6 Strategy Performance and Discussion

The proposed trading strategy using sentiment shock and trend scores was back-tested from 2007 to 2014. For the long strategy, both shock indicator and trend indicator yield higher Sharpe ratios than the S&P 500 index (see Table 10). Interestingly, higher cutoff percentile led to the phenomenon that Sharpe ratio rises to a peak and then gradually flattens out. This can be explained from the two perspectives. (1) When the

Table 7 Optimized number of weeks for sentiment scores by sector

Sector name	Sentiment shock	Sentiment trend
Consumer discretionary	15	14
Information technology	11	30
Consumer staples	18	19
Materials	15	16
Industrials	21	18
Utilities	16	28
Health care	10	15
Energy	25	20
Financials	11	25
Telecommunication services	19	24

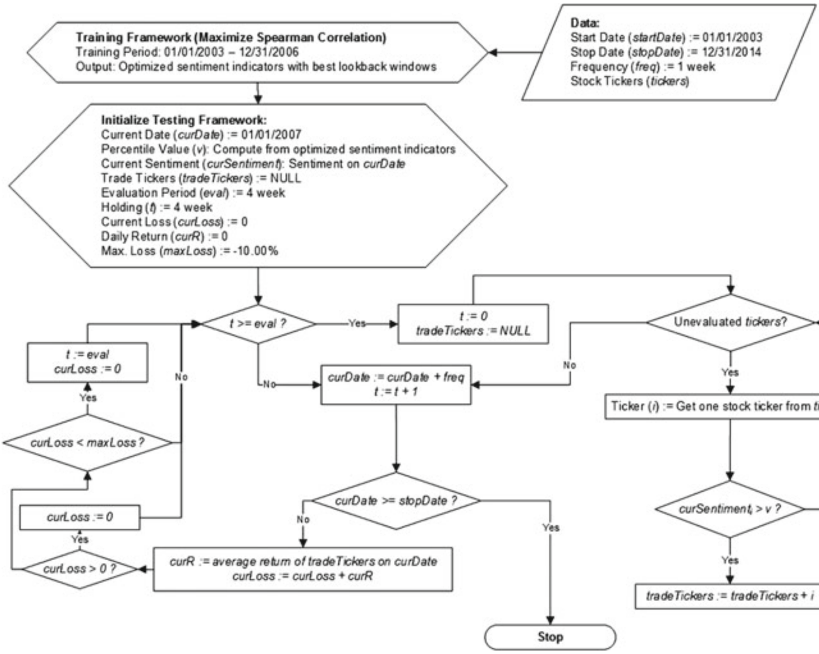

Fig. 10 Trading strategy diagram (long strategy)

cutoff percentile rises, more stocks are added into the trading portfolio. The Sharpe ratio increases in the initial stage because more companies with superior returns are included for better diversification. (2) The subsequent decline of Sharpe ratio is due to the diminishing effect of news influence. The other key result is that the trend indicator strategy performs better than the shock indicator strategy in terms of higher Sharpe ratio consistently across all cutoff percentiles. The distinctively

Table 8 Backtest statistics for long strategies

Strategy	Max. drawdown (%)	No. of trades	No. of wining trades	No. of losing trades	Avg. holding period (weeks)
Sentiment trend	49.09	673	449	224	9.51
Sentiment shock	56.37	896	576	320	6.96

different results from the long and short strategies demonstrate the asymmetric market response to extreme positive and negative sentiment. In order to test the robustness of the trading strategies, we recorded the trading activities for each strategy (Table 8). For both sentiment indicators, the number of winning trades almost doubles the number of losing trades, which suggests the consistency of the strategy in generating positive return (Fig. 11).

The top chart of Fig. 12 shows the cumulative returns of our trading strategies, with benchmark of the buy-and-hold strategy return of the S&P 500. The bottom chart of Fig. 12 illustrates the market volatility at corresponding period. The long strategies outperform the buy-and-hold strategy for the entire test period, confirming that the sentiment trend trading strategy has superior performance over the sentiment shock strategy. To validate the performance of the trading strategies in different market conditions, we split the back-test period into high and low volatility regime using 6-month realized market volatility. The high volatility regime during 2003 to

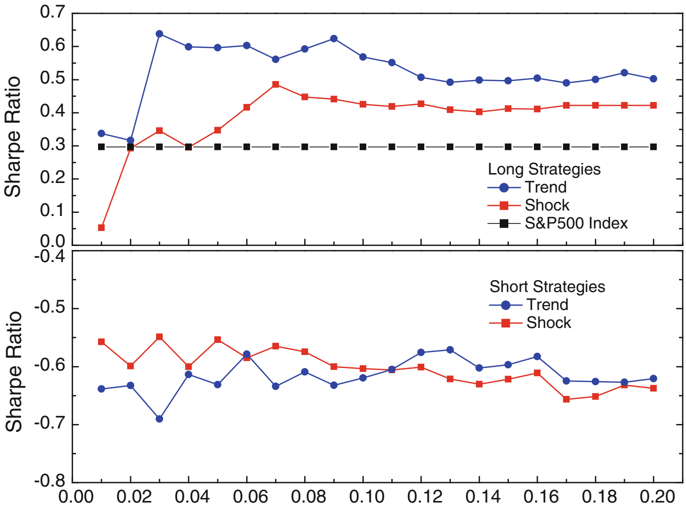

Fig. 11 Strategy sharpe ratios by changing extreme sentiment selection percentile. *Top chart* shows long strategy with threshold from *top* 1 to 20 % and *bottom chart* shows short strategy with threshold from *bottom* 1 to 20 %

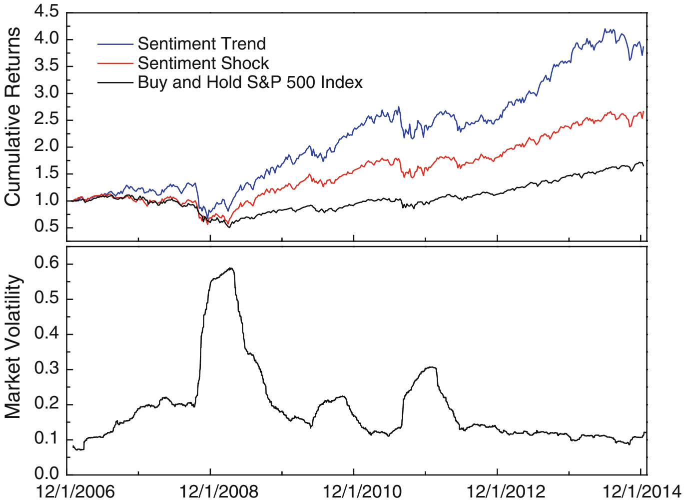

Fig. 12 Cumulative returns of sentiment trend and shock strategy benckmarked with Buy and Hold S&P 500 Index. *Top chart* shows the long strategy, *bottom chart* shows the market volatilty for the same time peroid

2014 was from 10/2008 to 05/2009. Both sentiment indicator strategies show higher profitability than the benchmark strategy in high volatility regime. In the low volatility regime that was bull market period, the trend indicator outperforms the benchmark in terms of higher return and Sharpe ratio. The shock indicator exhibits the same level of performance compared with the buy-and-hold strategy with a slightly lower Sharpe ratio (see Table 9). This result demonstrates that both sentiment indicators have good performance in predicting subsequent market returns in the long run, and the sentiment trend indicator provides more robust trading signals than the sentiment shock indicator. As shown in Table 9, the long strategies using the extreme positive sentiment outperform the S&P 500 index in both high and low market volatility regimes.

3.2.7 Conclusion

We demonstrate the value of using news sentiment data in formulating potential trading signals. Specifically, we use firm-specific news data from Thomson Reuters News Analytics, and we propose a sentiment shock score and a sentiment trend score for individual stocks to identify extreme sentiment levels and used them as trading signals. A previous study has shown that abnormal news sentiment, like sentiment shocks and trends, are predictive for future market return and volatility [47]. For individual stock level, the same intuition still applies that a big jump of the sentiment or a trend of sentiment change in the same direction will trigger persistent

Table 9 Backtest results in different market conditions

Strategy	Annualized performance measures		
	Mean return (%)	Volatility (%)	Sharpe ratio
Total backtest period			
Sentiment trend	16.90	26.50	0.64
Sentiment shock	12.24	25.22	0.49
Buy and Hold S&P 500	6.30	21.18	0.30
High volatility regime			
Sentiment trend	56.40	52.60	1.07
Sentiment shock	49.21	54.28	0.91
Buy and Hold S&P 500	−19.94	47.92	−0.42
Low volatility regime			
Sentiment trend	13.61	23.11	0.59
Sentiment shock	9.15	21.18	0.43
Buy and Hold S&P 500	8.49	17.29	0.49

impact on stock price movement. The back-test results of the trading strategy support the intuition that extreme positive sentiment has an impact on market returns and volatility. Our results show that the extreme positive sentiment for individual stocks generates more reliable trading signals than the extreme negative sentiment, which suggests the asymmetric response of the market to positive and negative sentiment.

3.3 Sentiment Genetic Programming Based Trading Strategies

This approach is motivated by the empirical findings that news and social media Twitter messages (tweets) exhibit persistent and predictive power on financial market movement. Based on the evidence that tweets is faster than news in revealing new market information, whereas news is regarded broadly a more reliable source of information than tweets, we propose a superior trading strategy based on the sentiment feedback strength between the news and tweets using generic programming optimization method. The key intuition behind the feedback strength based approach is that the joint momentum of the two sentiment series leads to significant market signals, which can be exploited to generate above-average trading profits. With the trade-off between information speed and its authenticity, we aim to develop a trading strategy with the objective to maximize the Sterling ratio. We find that the sentiment feedback based strategy yields superior market returns with small standard deviation over the two years from 2012 to 2014. When compared across the Sterling ratio and other risk measures, the proposed sentiment feedback based strategy generates better results over both the technical indicator-based strategy and the basic buy-and-hold strategy.

Using both news and tweets sentiments, we present a novel framework for developing an optimal trading strategy using genetic programming. It leverages existing empirical findings on the relationships observed among new sentiment, tweets sentiment and market returns. The key intuition behind the sentiment indicator is that the joint momentum of the two sentiment series leads to significant market anomalies which can be exploited in the form of above-average trading profits. For instance, if both news and tweets sentiments show strong momentum in trending in one direction, the market return is likely to follow in the same direction. An investor can therefore establish a long position when the sentiment indicator generates such signal and exits when the reversal appears. In addition, the two information sources also display key distinguishable characteristics that the trading rules can be constructed by choosing the optimal tradeoff between the speed of information release and the authenticity of the publishers. Our previous studies demonstrate that news sentiment has a more delayed impact than tweets sentiment on financial market returns, and we extend the findings by formulating an optimization problem to maximize risk-adjusted returns with the sentiment indicator [33, 45]. We prefer the use of genetic programming because of its flexibility in handling character strings and capability to search in a large population. In the study, we find that the sentiment-based genetic programming approach yields an above-average trading profit over the two years from 2012 to 2014. The out-performance suggests that the sentiment-based indicator can be regarded as a valuable source of information along with technical indicators.

3.3.1 Genetic Programming as an Optimization Approach

Genetic programming is a special class of genetic algorithm, which was first developed by John Holland in 1992. Genetic algorithm was built on the premise of the natural selection process that individual action with condition is evaluated with a pre-specified fitness function until the optimal combination is reached. Holland illustrated that "a population of fixed length character strings can be genetically bred using the Darwinian operation of fitness proportionate reproduction and the genetic operation of recombination" [26]. The central goal of using genetic algorithm is to exploit a vast region in the search space and at the same time to manipulate variations of strings [19]. The difference between genetic programming and genetic algorithm lies on the representation of the varying string length in the search space. Genetic programming allows solutions to be represented by a flexible string length with the Boolean connectors connecting the sentiment indicator with other technical indicators. For example, we can construct solutions with different combinations of indicators and parameters in contrast to the fixed set of indicators that we have to use for each search. Moreover, GP requires input solutions to be represented in a tree structure to accommodate the flexibility. Three major genetic operators are applied to a given problem during the optimization process: mutation, crossover and encoding. These

1: Randomly create an initial population of programs from the available primitives;
repeat
  2: Execute each program and ascertain its fitness;
  3: Select one or two program(s) from the population with a probability based on gitness to participate in genetic operations;
  4: Create new individual program(s) by applying genetic operations with specified probabilities;
until *an acceptable solution is found or some other stopping condition is met*;
return the best-so-far individual;

Fig. 13 Genetic programming algorithm

operators are crucial in the effectiveness of the genetic programming framework to converge towards the optimal solution within the search space.

We apply the standard framework of genetic programming to locate the optimal trading strategy with the proposed sentiment indicator (see Fig. 13). The framework allows the comparison of the large number of combinations among variations of indicators and parameters. The goal of the algorithm is to locate the optimal trading strategy with the highest Sterling ratio, which is set to be the fitness function. In the genetic programming (GP) framework, we first initialize the population of programs constructed from the sentiment feedback strength indicator and the two technical indicators. Through the search process, we incorporate the Boolean operators, "AND" and "OR", for allowing different combinations of the indicators. For each strategy, the algorithm generates trading signals in the form of "TRUE/FALSE" signal at each time period. Since there is no short position allowed, the "FALSE" signal is effective only when an existing position is open. For a "TRUE" signal, the system records the cumulative returns over the holding period until a reversal of the trading signal appears. For example, if a position is established on day 1 and closed on day 10, the trading return is calculated as the cumulative returns over the 10-day period. With the trading record of the strategy, the algorithm ascertains its fitness with the Sterling ratio and then performs genetic operations in crossover and mutation with specified probabilities. We select the crossover rate and mutation rate as 0.50 and 0.90 respectively to extend the search population and increase the likelihood of solution achieving global convergence. The GP algorithm is implemented through multiple iterations to generate the optimal combinations of the indicators connected by Boolean operators and numeric parameters in each indicator.

3.3.2 Trading Strategy Performance Comparison

This section presents the performance comparison of the two sentiment feedback strength based trading strategies against two benchmark strategies. The first benchmark strategy utilizes the genetic programming framework to generate trading signals based on entirely technical indicators only. The rationale behind this strategy is that GP can generate useful technical trading rules with optimal set of parameters. The

Table 10 GP Algorithm parameter set

Parameter	Value
Population size	100
Number of iterations	5,000
Selection method	Roulette wheel
Fitness function	Stirling ratio
Boolean operators	"AND", "OR"
Numeric parameters	U(0,1)
Crossover rate	0.5
Mutation rate	0.9
Max initial tree depth	5
Max following tree depth	5

second benchmark is the traditional buy-and-hold strategy that is commonly utilized by small investors and mutual funds.

The results show that the combined trading strategy with both sentiment feedback strength and technical indicators provides a clear edge over the other three trading strategies in terms of higher Sterling ratio and total profit/loss. The optimal strategy generates over 21.3 % Sterling ratio compared to 15.5 %, 12.1 % and 8.6 % from the feedback strength-only strategy, technical indicators-only strategy, and the buy-and-hold strategy respectively (see Table 11). For the comparison of the total cumulative returns over the evaluation period, the combination of sentiment feedback strength and technical indicators yields the best performance at 21.7 % (see Fig. 14). On the other hand, the result related to the trading strategy based entirely on sentiment feedback strength only suggests that sentiment provides support in controlling loss indicated by the significantly lower maximum dropdown at −0.6 % in contrast to −2.1 % for the three strategies. We find that the percentage of winning trades is higher at 73.61 % despite the lower number of trades at 72. From a standpoint of evaluating the strategy risk, the standard deviation of the daily returns is also lower at 0.30 % compared to 0.47 % and 0.58 % respectively.

Through the genetic programming optimization, we find that the lag-1 news sentiment and lag-2 tweets sentiment are the most dominant factors in the formulation of the optimal trading strategy. In other words, the trading signals based on the feedback strength indicator rely significantly on business news articles published one day ago and tweet messages generated by the Twitter financial community two days ago. The results suggest that the combination of the two factors yields the optimal performance in terms of Sterling ratio and the percentage of winning trades. Furthermore, the lag-2 tweets sentiment exhibits a stronger effect on triggering trading signals over the lag-1 news sentiment, demonstrated by the higher optimal weight determined by the algorithm. On the contrary, the lag-1 news sentiment displays a greater sensitivity in affecting market returns, reflected by the lower summation threshold in the sentiment feedback strength indicator.

Table 11 GP optimization trading strategy performance

	Sentiment + Technical indicators	Sentiment indicator	Technical indicators	Buy-and-hold
Number of observations	132	132	132	132
Number of trades	102	72	92	132
Trading %	77.3	54.5	69.7	100
Number of winning trades	74	53	61	81
Percentage of winning trades (%)	72.5	73.6	66.3	61.4
Total profit/loss (%)	21.7	15.6	12.3	8.8
Average profit/loss per trade (%)	0.21	0.22	0.13	0.07
Standard deviation (%)	0.44	0.30	0.47	0.58
Maximum drawdown (%)	−2.1	−0.6	−2.1	−2.1
Stirling ratio (%)	21.3	15.5	12.1	8.6

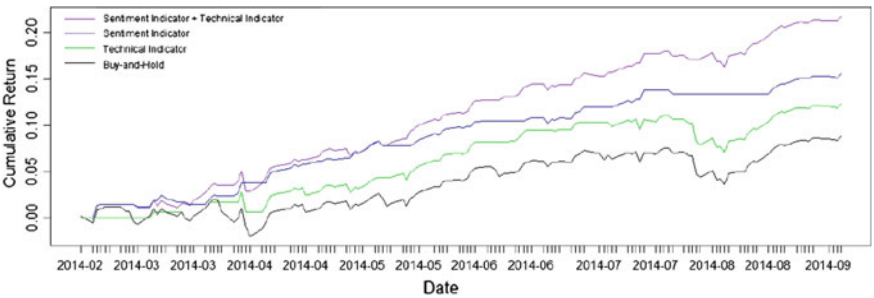

Fig. 14 Cumulative returns of sentiment-based trading strategy

3.3.3 Conclusion

We introduce a genetic programming approach to develop an optimal trading strategy with news and tweet sentiments. The proposed feedback strength indicator, a measure of the joint momentum between the news and tweet sentiments, was found to provide a significant improvement in trading performance over the S&P 500 financial market index ETF. By quantifying the joint momentum of the sentiment series, we can detect significant market anomalies that can be exploited in the form of modest trading returns. We find that the sentiment-based genetic programming approach yields positive market returns with small standard deviation over the two years from 2012 to 2014. When comparing the Sterling ratio and other risk measures, the proposed strategy is superior to the technical indicators and the traditional buy-and-hold strategy. The out-performance suggests that news and tweet sentiments can be regarded as valuable sources of information in constructing meaningful trading system along with technical indicators [45].

4 Summary and Conclusion

This chapter focuses on showcasing the capabilities of financial market sentiment analysis. It explores the impact of news and social media on investor sentiment and financial markets. Through their empirical relations, it further introduces how existing findings can be combined with artificial intelligence techniques to develop advanced investment strategies. The main contribution of the book chapter rests on surveying existing literature findings related to sentiment analysis and financial market and addressing research questions within the field of behavioral finance whether investor sentiment can be quantified through news and tweet sentiment. Research has shown that financial market sentiment can be leveraged as a source to develop practical financial solutions in the form of a modestly profitable trading strategy. It reinforces the empirical evidence in the literature of behavioral finance that there exists opportunities in the area of investor sentiment for generating above-average returns. Using advanced techniques in processing and analyzing textual sources of information, the specific contributions of this book chapter are summarized as follows:

- The concept of community construction using social network analysis is a novel approach in the field of analyzing social media impact on financial market. By harnessing the sentiment expressed by the influential Twitter users within the community, the empirical study provides a better proxy over existing studies between social sentiment and financial market movements
- Using news sentiment data, a firm-specific trading strategy is developed based on the detection of abnormal sentiment levels. The study shows that news sentiment can be a proxy for future market returns and volatility. The above-normal positive sentiment for individual stocks generates more reliable trading signals than the extreme negative sentiment. This is also an indication of the asymmetric response of the market to positive and negative sentiment.
- Based on the finding of the time scale difference between news and social sentiment, we present a novel approach to formulate an optimization problem for identifying a trading strategy based on the sentiment feedback strength using genetic programming. The proposed sentiment feedback based strategy shows better performance over the technical indicator-based strategy and the basic buy-and-hold strategy and further validates the value of both news and tweets sentiment in exploiting trading opportunities.

Acknowledgments The authors would like to thank the Financial Engineering Division at Stevens Institute of Technology for providing a state-of-the-art research environment with data access and hardware support. They would also like to acknowledge the support from the Civil Group of Northrop Grumman Corporation.

References

1. Aboura, S., Chevallier, J.: Leverage vs. feedback: which effect drives the oil market? Finan. Res. Lett. **10**(3), 131–141 (2013)
2. Antoniou, A., Koutmos, G., Pericli, A.: Index futures and positive feedback trading: evidence from major stock exchanges. J. Empirical Finan. **12**(2), 219–238 (2005)
3. Antweiler, W., Frank, M.Z.: Is all that talk just noise? The information content of internet stock message boards. J. Finan. **59**(3), 1259–1294 (2004). http://doi.org/10.1111/j.1540-6261.2004.00662.x
4. Arnold, L.G., Brunner, S.: The economics of rational speculation in the presence of positive feedback trading. Q. Rev. Econ. Finan. (2014)
5. Baccianella, S., Esuli, A., Sebastiani, F.: SentiWordNet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: Calzolari, N., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., ... Tapias, D. (eds.) Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), pp. 2200–2204. European Language Resources Association (ELRA) (2010)
6. Barber, B.M., Odean, T.: All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. Rev. Finan. Stud. **21**(2), 785–818 (2007)
7. Barberis, N., Shleifer, A., Vishny, R.: A model of investor sentiment. J. Finan. Econ. **49**, 307–343 (1998)
8. Borgatti, S.P., Everett, M.G.: A Graph-theoretic perspective on centrality. Soc. Netw. **28**(4), 466–484 (2006)
9. Brown, G.W., Cliff, M.T.: Investor sentiment and the near-term stock market. J. Empirical Finan. **11**(1), 1–27 (2004)
10. Chau, F., Deesomsak, R.: Business cycle variation in positive feedback trading: evidence from the G-7 economies. J. Int. Finan. Mark. Inst. Money (2014)
11. Chopra, N., Lakonishok, J., Ritter, J.R.: Measuring abnormal performance. J. Finan. Econ. **31**, 235–268 (1992)
12. Cohen, G., Kudryavtsev, A.: Investor rationality and financial decisions. J. Behav. Finan. **13**(1), 11–16 (2012)
13. Crone, S.F., Koeppel, C.: Predicting exchange rates with sentiment indicators: an empirical evaluation using text mining and multilayer perceptrons. In: 2014 IEEE Conference on Computational Intelligence for Financial Engineering and Economics (CIFEr), pp. 114–121 (2014)
14. Daniel, K.D., Hirshleifer, D., Subrahmanyam, A.: Overconfidence, arbitrage, and equilibrium asset pricing. J. Finan. **LVI**, 921–965 (2001)
15. Daniel, K., Hirshleifer, D., Subrahmanyam, A.: Investor psychology and security market under- and overreactions. J. Finan. **53**, 1839–1885 (1998)
16. DiBartolomeo, D., Warrick, S.: Linear Factor Models in Finance. Elsevier (2005)
17. Hilton, D.J.: The psychology of financial decision-making: applications to trading, dealing, and investment analysis. J. Psychol. Finan. Mark. **2**(1), 37–53 (2001)
18. Hirshleifer, D., Subrahmanyam, A., Titman, S.: Feedback and the success of irrational investors. J. Finan. Econ. **81**(2), 311–338 (2006)
19. Holland, J.: Genetic algorithms. Sci. Am. 66–72 (1992)
20. Hou, Y., Li, S.: The impact of the CSI 300 stock index futures: positive feedback trading and autocorrelation of stock returns. Int. Rev. Econ. Finan. **33**, 319–337 (2014)

21. Im, T.L., San, P.W., On, C.K., Alfred, R., Anthony, P.: Impact of financial news headline and content to market sentiment. Int. J. Mach. Learn. Comput. **4**(3), 237–242 (2014)
22. İnkaya, A., Yolcu Okur, Y.: Analysis of volatility feedback and leverage effects on the ISE30 index using high frequency data. J. Comput. Appl. Math. **259**, 377–384 (2014)
23. Johnson, B.: Algorithmic Trading & DMA: An Introduction to Direct Access Trading Strategies, p. 574. 4Myeloma Press (2010)
24. Kaya, M.: Stock price prediction using financial news articles. In: 2nd IEEE International Conference on Information and Financial Engineering ICIFE 2010, pp. 478–482 (2010)
25. Khanna, N., Sonti, R.: Value creating stock manipulation: feedback effect of stock prices on firm value. J. Finan. Mark. **7**(3), 237–270 (2004)
26. Koza, J.R.: Evolution of a Subsumption Architecture that Performs a Wall Following Task for an Autonomous Mobile Robot via Genetic Programming, pp. 1–39 (1992)
27. La Porta, R., Lakonishok, J., Shleifer, A., Vishny, R.W.: Good news for value stocks: further evidence on market efficiency. J. Finan. **52**, 859–874 (1997)
28. Laopodis, N.T.: Feedback trading and autocorrelation interactions in the foreign exchange market: further evidence. Econ. Modell. **22**(5), 811–827 (2005)
29. Li, X., Xie, H., Chen, L., Wang, J., Deng, X.: News impact on stock price return via sentiment analysis. Knowl-Based Syst. **69**, 14–23 (2014)
30. Medhat, W., Hassan, A., Korashy, H.: Sentiment analysis algorithms and applications: a survey. Ain Shams Eng. J. **5**(4), 1093–1113 (2014)
31. Mitchell, M., Mulherin, H.: The impact of public information on the stock market. J. Finan. **49**(3), 923–950 (1994)
32. Mitra, L., Mitra, G.: The Handbook of News Analytics in Finance. Wiley, Chichester, West Sussex, UK (2011)
33. Mo, S.Y.K., Liu, A., Yang, S.: News sentiment to market impact and its feedback. SSRN Working Paper (Submitted) (2015)
34. Moreo, A., Romero, M., Castro, J.L., Zurita, J.M.: Lexicon-based comments-oriented news sentiment analyzer system. Expert Syst. Appl. **39**(10), 9166–9180 (2012)
35. Nofsinger, J.R.: Social mood and financial economics. J. Behav. Finan. **6**(3), 144–160 (2005)
36. Salm, C.A., Schuppli, M.: Positive feedback trading in stock index futures: international evidence. Int. Rev. Finan. Anal. **19**(5), 313–322 (2010)
37. Schumaker, R.P., Zhang, Y., Huang, C.-N., Chen, H.: Evaluating sentiment in financial news articles. Decis. Support Syst. **53**(3), 458–464 (2012)
38. Scott, J., Stumpp, M., Xu, P.: News, not trading volume, builds momentum. Finan. Anal. J. **59**(2), 45–54 (2003)
39. Smales, L.A.: Asymmetric volatility response to news sentiment in gold futures. J. Int. Finan. Mark. Inst. Money **34**, 161–172 (2015)
40. Song, Q., Liu, A., Yang, S.Y., Datta, K., Deane, A.: An extreme firm-specific news sentiment asymmetry based trading strategy. In: 2015 IEEE Conference on Computational Intelligence for Financial Engineering and Economics (CIFEr 2015) (2015)
41. Tetlock:. Giving content to investor sentiment: the role of media in the stock market. J. Finan. **62**(3), 1139–1168 (2007)
42. Tetlock, P.C., Saar-Tsechansky, M., MacSkassy, S.: More than words: quantifying language to measure firms' fundamentals. J. Finan. **63**, 1437–1467 (2008). http://doi.org/10.1111/j.1540-6261.2008.01362.x
43. Thaler, R.H.: Advances in Behavioral Finance, vol. 2. Princeton University Press, Princeton, NJ (2005)
44. Xie, B., Wang, D., Passonneau, R.J.: Semantic feature representation to capture news impact. In: Twenty-Seventh International Florida Artificial Intelligence Research Society Conference (2014)
45. Yang, S., Mo, S.Y.K., Liu, A., Kirilenko, A.: Genetic programming optimization for a sentiment feedback strength based trading strategy. SSRN Working Paper (2015)
46. Yang, S.Y., Mo, S.Y.K., Liu, A.: Twitter financial community sentiment and its predictive relationship to stock market movement. Quant. Finan. **15**(10), 1637–1656 (2015)

47. Yang, S.Y., Song, Q., Yin, S., Mo, K., Datta, K., Deane, A.: The impact of abnormal news sentiment on financial markets. J. Bus. Econ. **6**(10), 1682–1694 (2015)
48. Yu, L.-C., Wu, J.-L., Chang, P.-C., Chu, H.-S.: Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news. Knowl.-Based Syst. **41**, 89–97 (2013)
49. Zhang, W., Skiena, S.: Trading Strategies To Exploit Blog and News Sentiment (2010). Chan (2003)