

The Pessimistic Investor Sentiments Indicator in Social Networks

Rui Jin^{*1}, Hong-Li Zhang², Xing Wang³, Xiao-Meng Wang⁴

School of Computer Science and Technology, Harbin Institute of Technology,
Harbin, China, 150001, telp: +86-0451-86418273

Corresponding author, e-mail: jinrui@hit.edu.cn^{*1}, zhanghongli@hit.edu.cn², yeahwx@gmail.com³,
wangxiaomeng@pact518.hit.edu.cn⁴

Abstract

With the worldwide proliferation of social networks, the social networks have played an important role in the social activities. Peoples are inclined to obtain the corresponding public opinion to make decision such as shopping, education, investment and so on. Analysis of data generated by social networks has become an important field of research, however in the field of public opinion analysis of social networks the quantitative measure indexes are still lacking. In this paper, the calculation method of pessimistic investor sentiments indicator is proposed, and the index has a certain theoretical and practical value.

Keywords: social computing, information entropy, pessimistic investor sentiments indicator, public opinion analysis, social networks

Copyright © 2017 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

The formal academic study of public opinion is taken seriously in recent years, but the practical research of public opinion has a history of many years, over the years many governments try to act as the monitor of public opinion, they even established secret police forces to probe who are opposing the government [1]. With the emergence of Web 2.0, the social networks such as Facebook, Twitter and YouTube that play an important role in the social activities, the amount of freely available user-generated data has reached an unprecedented volume [2].

In China the influential social networks such as the Sina microblog, Renren and douban.com have penetrated every corner of Chinese society [3]. But how to measure people's opinions is the problem needed to be solved all this time. The corresponding researchers develop a few methods to analyze the public opinion [1], [4].

Analyzing web resources multi-dimensionally and obtaining Internet public opinion trends can take a lot advantages. In real society the public opinion in these social networks have a profound influence on the nation's politics, economy, culture, investment and so on. It will help the government in a timely manner to collect social public opinions and make quick guide decisions. Therefore the Internet public opinion information mining has become an important research topic in the study area of data mining and Internet applications [5].

When the people's opinions in social networks can be measured, the related data analysis and processing are meaningful. To develop quantitative calculation method is the precondition [1].

The research about the public opinion analysis for social networks should be included in the research field of intelligence and security informatics (ISI). The public opinion has become the critical intelligence information to the politics, economy, trade, financial transactions, and so on.

The University of Arizona has dedicated a great deal of research to "intelligence and security infor-matics (ISI)" for the sake of national security. Since 2005, the Chinese Academy of Sciences Institute of Automation has conducted research on Intelligence and Security Informatics (ISI) [6].

In 2009, David Lazer, et al., developed the concept of Computational Social Science in a paper published in the journal Science. This is a important scientific idea. They noted that a great deal of information contained in social networks such as blogs, BBS, chats, records of consumption, and E-mails are mappings of individual behavior and organizational behavior in society and argued that the data generated by these networks can be used to analyze the behavior of both individuals and groups [7]. Because the amount of the data is so massive, it is impossible for users to make sense of its whole in limited time, therefore people have tried to create systems capable of extracting information from it [2].

Although there are a large amount of multimedia data in social networks, such as images and videos, usually we only use text data and ignore other multimedia data. This is because public opinion mining requires the target data file with semantic information [6].

Nowadays the research field of online public opinion analysis is still developing, especially the lack of a reasonable evaluation indexes system. Due to the lack of the corresponding theory system and effective technical means, the efficiency of information analysis and processing is limited greatly.

In this paper, the calculation method of *pessimistic investor sentiments indicator* is put forward which can represent the changes and trends of the pessimistic public investment sentiments distinctly.

2. Related Works

2.1. The History of Public Opinion Analysis

The most common way for a modern government to learn about public opinion is through polls and elections. Polls and elections have been held at regular intervals in a lot of countries.

In the long process of human history one of the earliest expressions of public opinion is rebellion. When peasant rebellions have occurred the rulers can see that his policies and regulations are opposed, this is a clear sign that the government's policy is not be welcomed.

The formal academic study of public opinion is a relatively new research field in modern political economy, but the practical application of public opinion analysis has a long history. Governments have paid more attention to the public opinion as themselves.

For centuries rulers have even established secret police forces to find out which people oppose the government and to eliminate them. Secret police have often acted as monitors of public opinion.

Fortunately, the modern governments do not need to rely on secret police or wait for the public rebellions to obtain public opinion. Almost all modern governments have much better procedures to learn about public opinion and measure it [1].

For example, the public opinion polls are praised these years, and it has an important position in the American political system, but at the same time the merit of polls is maligned by the opponents. But regardless of the opponents' view on polls, the survey of public opinion has been developed to be one of the most important parts of American politics [4].

2.2. The Important Social Role of Public Opinion

Public opinion once occupies an important position in national politics in the history. Nowadays in the era of social networks the social role of public opinion has become more important not only to the national politics but also to the economy, investment, education, trade and so on.

The worldwide proliferation of social networks has resulted in a wave of social networking websites such as Facebook, Twitter and YouTube that played an important role in the social activities of their users. Chinese social networks such as the Sina microblog, Renren and douban.com have penetrated every corner of Chinese society [3].

In Arab spring, a series of public events through social networks led to the outbreaks of mass incidents, some countries were caught in the political situation turbulence, society chaos numerously and confusedly in every country, in slump crisis. The important social role of public opinion in the era of social networking has come to be realized gradually [8].

It also can help companies acquire the social hot spots and make business decisions quickly. Moreover, it can help enterprises to understand user experience and even trace the development directions of science, economy, technology and culture [6].

2.3. Methods Used for Measuring Public Opinion

2.3.1. Five Methods

In 1931, D. D. Droba published a paper named "Methods Used for Measuring Public Opinion". In this paper he systematically summarized five kinds of measurement methods of public opinion firstly [4].

Five methods have been proposed by investigators for measuring the public opinion.

- (1) the questionnaire method
- (2) the ranking method
- (3) the rating method
- (4) the comparison method
- (5) statement and sort

2.3.2. Social Computing and Public Opinion

In the social networks era, the researchers can obtain more social data from the internet. Social computing is one of the basic methods to analysis the public opinion.

3. Research Method

3.1. The Description of the Problem

In social networks there is a category of important social data, that is the user comments about the financial market. The relevant research into the economic crisis suggested that financial markets are more influenced by negative press rumours than fundamentals.

The viewpoints including positive or negative sentiments previously shared by consumers, clients or investors with their personal networks. These data from social networks contain abundant information, after analysis the important information related to investment intention can be obtain, and the researchers have showed great interest in it.

The information about the investor sentiments has vast implications for public companies and their stockholders [9].

3.2. The structure of Pessimistic Investor Sentiments

An analysis indicates that eight main types of sentiments concerning investments appear in social network comments, including depression, disappointment, fear, anxiety, panic, anxiety, dread, and despair.

Based on the analysis, pessimistic investor sentiments can be given a simplified structure, and it is described in Figure 1.

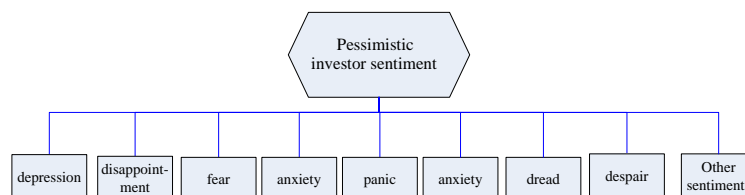


Figure 1. Pessimistic investor sentiments in social networks

3.2. The Mathematical Model for Pessimistic Investor Sentiments

Suppose that the discrete random variable X represents the "investor sentiments", the event itself is a random information system in which X_1 represents "depression", X_2 represents "disappointment", and so on through X_8 , which represents "despair". Finally, X_9 represents "other pessimistic sentiments". The discrete random variable X is equivalent to (X_1, X_2, \dots, X_9) , where (X_1, X_2, \dots, X_9) is a multidimensional random variable, or it can be represented that:

$$X \square (X_1, X_2, \dots, X_9)$$

The domain of values of X is U , the domain of values of X_i is $U_i (1 \leq i \leq 9)$.

Suppose that $q_i (1 \leq i \leq 9)$ represents the amount of X_i after obtaining the values. While X_i obtains its value in its basic set U_i of values once, $q_i = 1$.

3.3. The Method to Calculate the Entropy of Pessimistic Investor Sentiment

According to the principle of maximum entropy [10]:

$$H(X_1, X_2, \dots, X_n) = \log \left(\prod_{i=1}^n q_i \right) \quad (1)$$

The formula for pessimistic investor sentiment entropy is:

$$H(X_1, X_2, \dots, X_9) = \log(q_1 q_2 \dots q_9) \quad (2)$$

3.4. The Definition of the Pessimistic Investor Sentiments Indicator

The pessimistic investor sentiment indicator is H^- , which is the negative information entropy of all pessimistic investor sentiments in a single day.

$$H^- = -|\log(q_1 q_2 \dots q_9)| \quad (3)$$

4. Experiments and Results

We obtained 1000 stock market reviews concerning the Chinese stock market from community networks, Sina Weibo and newspaper comments. These reviews included a lot of natural emotional expression about the stock market, and it tended to be more integrated and traditional. We then counted the number of investor sentiment occurrences of each type for each day in March and April 2016.

In order to get more reasonable results the knowledge of domain experts can be considered in the computational process, the weight k_i of q_i (regardless of the value of the weight) and the weight k_i' of q_i' (considering the weighted value) are shown in Table 1.

Table 1. The weight of q_i and q_i'

X_i	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
k_i	1	1	1	1	1	1	1	1	1
k_i'	1	2	2	2	2	3	3	3	1

The corresponding statistical data are shown in Table 2 for March 1, 2016.

Table 2. The value of dimensions

X_i	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
q_i	672	753	823	708	642	832	1009	852	1
q_i'	672	1506	1643	1416	1284	2496	3027	2556	1

According to Formula (3), H^- can be calculated as follows:

$$\begin{aligned} H^- &= -|\ln(q_1 q_2 \dots q_9)| \\ &= -|\ln(672 * 753 * 823 * 708 * 642 * 832 * \\ &\quad 1009 * 852 * 1)| \\ &= -53.26 \end{aligned}$$

H' can be calculated as follows:

$$\begin{aligned} H' &= -|\ln(q_1 q_2 \dots q_9)| \\ &= -|\ln(672 * 1506 * 1643 * 1416 * 1284 * 2496 * \\ &\quad 3027 * 2556 * 1)| \\ &= -59.33 \end{aligned}$$

We can calculate the corresponding values from March 1, 2016 to March 31, 2016 as shown in Table 3, which shows the values of H^- and H' calculated using the values of q_i and q'_i .

Table 3. The values of H^- and H' in March 2016

time	Mar. 1	Mar. 2	Mar. 3	Mar. 4	Mar. 5	Mar. 6	Mar. 7	Mar. 8	Mar. 9
H^-	-53.26	-51.82	-51.92	-48.28	-46.05	-47.29	-46.49	-51.84	-52.72
H'	-59.33	-55.62	-55.72	-53.91	-51.82	-52.39	-52.01	-57.20	-59.28
time	Mar. 10	Mar. 11	Mar. 12	Mar. 13	Mar. 14	Mar. 15	Mar. 16	Mar. 17	Mar. 18
H^-	-55.73	-54.08	-53.72	-51.72	-50.24	-48.69	-46.73	-44.46	-43.15
H'	-60.73	-60.21	-58.05	-57.56	-56.29	-53.47	-51.28	-50.03	-48.29
time	Mar. 19	Mar. 20	Mar. 21	Mar. 22	Mar. 23	Mar. 24	Mar. 25	Mar. 26	Mar. 27
H^-	-42.20	-41.83	-38.33	-37.72	-38.27	-42.92	-44.35	-47.73	-49.71
H'	-48.93	-46.25	-43.82	-42.27	-41.03	-47.75	-49.06	-51.92	-54.52
time	Mar. 28	Mar. 29	Mar. 30	Mar. 31					
H^-	-51.32	-54.63	-52.06	-51.58					
H'	-56.73	-60.37	-57.81	-56.35					

As Figure 2 shows, in consideration of the weighted value the calculation results of the pessimistic investor sentiments indicator are more reasonable, and the ability to discriminate between the results are more conspicuous.

In Figure 2, the red line represents the values of H^- , and the blue line represents the values of H' .

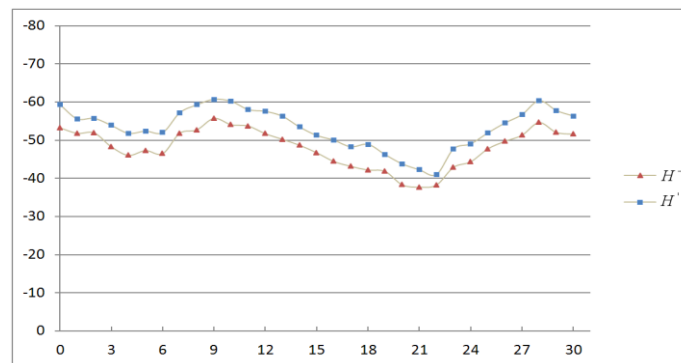


Figure 2. Comparison 1 of experimental results

In January and February 2016 the stock market of China have gone through a series of swoons, and the investor sentiments in social networks were very low. In February 2016 the economy and financial environments show some improvements, and there is corresponding some fairly optimistic about the investor sentiments, and it is described in Figure 3.

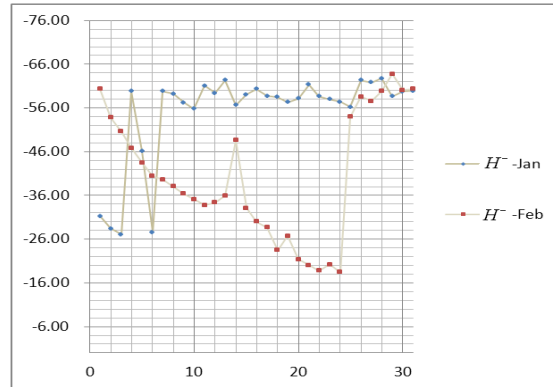


Figure 3. Comparison 2 of experimental

The results in Figure 4 show that the absolute values of the pessimistic investor sentiment indicator began to improve in April 2016, indicating that investor sentiments changed along with the stock market economic environment. This computation results can provide the reference material for the financial market analysts.

5. Results and Analysis

The calculated value of pessimistic investor sentiments indicator directly maps the pessimistic sentiments expressed by investors to the real-world stock market situation. Such mappings can help economists and investors analyze international financial markets and predict economic development trends.

In January 2016, the Chinese stock market experienced a sharp drop. As a result, a large rise in the number of pessimistic comments can be seen in social networks.

The results shown in Figure 3 and Figure 4 indicate that the absolute value of the pessimistic investor sentiment index continued to have larger average values in the first four months, 2016, which shows that investors were consistently depressed. This result is consistent with the actual investment environment.

Moreover, the results shown in Figure 4 clearly indicate a improvement in the sentiments expressed by investors in social networks between March 2016 and April 2016, as the investment environments began to recover.

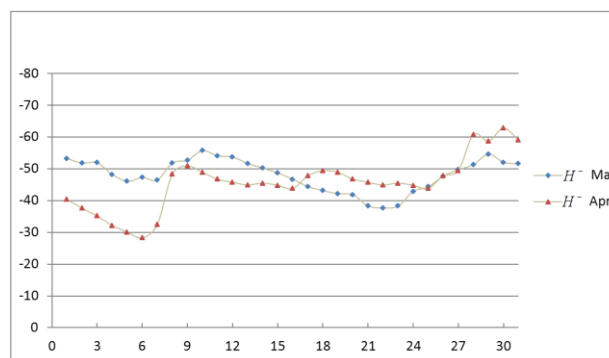


Figure 4. Comparison 3 of experimental results

6. Conclusion

This paper presents the calculation method of pessimistic investor sentiment indicator based on maximum entropy theory. The method has a solid theoretical foundation and can be used to quantitatively analyze uncertain issues using data gathered through social networks.

Acknowledgements

Supported by the Major State Basic Research Development Program of China (973 Program) under Grant 2013CB329602; the National Natural Science Foundation of China under Grant No.61202457, No.61472108, No.61402149.

References

- [1] Russell G. Brooker, Todd Schaefer. *Methods of Measuring Public Opinion*. Public Opinion in the 21st Century. 2015.
- [2] JA Balazs, JD Velásquez. Opinion mining and information fusion: a survey. *Inf. Fusion*. 2016; 27: 95-110.
- [3] Louis Y, Asur S, Huberman BA. *What Trends in Chinese Social Media*. 5th SNA-KDD Workshop'11 (SNA-KDD'11). San Diego, CA, USA. 2011.
- [4] DD Droba. *Methods Used for Measuring Public Opinion*. *American Journal of Sociology*. 1931.
- [5] K Guo, L Shi, W Ye, X Li. *A survey of internet public opinion mining*. In Proceedings of the International Conference on Progress in Informatics and Computing (PIC 2014), IEEE. Shanghai, China. 2014: 173-179.
- [6] Chen H, Wang FY, Zeng D. Intelligence and security informatics for homeland security: Information, communication, and transportation. *IEEE Trans. on Intelligent Transportation Systems*. 2004; 5(4): 329-341.
- [7] Lazer D, Pentland A, Adamic L, Aral S, Barabasi AL, Brewer D, Christakis N, Contractor N, Fowler J, Gutmann M, Jebara T, King G, Macy M, Roy D, Van Alstyne M. Computational social science. *Science*. 2009; 323(5915): 721-723.
- [8] Khondker, Habibur Haque. *Role of the New Media in the Arab Spring*. Special Forum on the Arab Revolutions. 2011: 675-679.
- [9] GNIP. *Social Media in Financial Markets: The Coming of Age*. Whitepaper. 2014.
- [10] Xian-dong L. *The Method Study about Probability Distribution Based on the Principle of Maximum Entropy*. MS Thesis. Beijing, China: North China Electric Power University; 2008