Research on the Abnormal Recognition of Corporate Overseas Public Opinion Based on News Sentiment Analysis

LiXiongLuluTangZhaoranXu
Corresponding Author:Lulu Tang.
School of Management, Shanghai University

ABSTRACT:*For international companies, quickly collecting large amounts of relevant news information and conducting corporate sentiment analysis in a short period of time is a vital capability to help companies identify crises and make the right decisions. The paper tries to calculate the sentiment score of the news through the semantic-based sentiment analysis algorithm, combined with the S-H-ESD algorithm to identify the abnormal point of the public opinion of the enterprise, and finally use the data visualization to analyze the cause of the abnormal occurrence of the public opinion. Taking Huawei companies as the research object, the results show that the emergence of Huawei's positive public opinion in the international market is related to the release of Huawei's new technology and new product, the emergence of Huawei's negative public opinion is related to international political and economic factors such as "national security" and "subsidies".*
KEY WORD*: Enterprise public opinion monitoring, sentiment analysis,outlier detection*

I.   INTRODUCTION AND LITERATURE REVIEW

In today's world, increasing trade friction and geopolitical conflicts have affected the rapid transformation of multinational companies. Because of their suddenness and destructiveness, the overseas crisis of multinational companies has become the focus of international public opinion. For example, Alibaba was accused of concealing unfavorable information after listing in 2015, ZTE was accused of violating U.S. Iran sanctions law in 2017, and Huawei's CFO was arrested in Canada in 2018, etc. It can be seen that for multinational companies whose home countries are China, such as Alibaba and Huawei, due to the impact of factors such as international politics, economy, culture, geographical location and other factors, the public opinion environment and risk challenges are also more complicated. Therefore, how do Chinese companies conduct overseas public opinionanalyze, grasp the complex changes in the overseas economic and political situation, and promptly detect crises has become a topic worth studying.

From the enterprise's perspective, monitoring and management of public opinion helps improve corporate performance and increase corporate profits. Corporate public opinion refers to the sum of the public's use of the Internet to publish and disseminate different perceptions, opinions, and attitudes about corporate events (Liyan, 2013). Compared with personal public opinion and national public opinion, corporate public opinion has the characteristics of higher timeliness, large-scale outbreak and spread in a short time, and the goal of promoting sales (Xun et al., 2017). Prospect theory proposes that consumers have a loss-avoidance mentality(Tversky,1979). Compared with positive public opinion, negative public opinion can affect consumers' behavior more and cause serious losses to enterprises. Therefore, multinational companies need to pay close attention to foreign media reports, especially the negative public opinion of the company, and quickly make countermeasures to reduce losses.

Detecting emotions in social media can help companies understand the views of related products and events(Cheng et al., 2017). The core of enterprise public opinion monitoring is to obtain massive network public opinion data through web crawlers and monitor public opinion according to certain indicators. The key technologies include data collection technology, natural language processing technology, sequence pattern mining technology, sentiment analysis technology, etc.(Jun-Ling, 2016). Among them, sentiment analysis can be regarded as a classification problem(Korayem et al., 2016), which aims to classify text or video as positive, negative, or neutral(Chaturvedi et al. 2017).

Currently, sentiment analysis is mainly divided into semantic-based analysis and machine learning-based analysis. The former uses sentiment analysis tools such as WordNet or HowNet to identify sentiment tendencies by constructing sentiment dictionaries(Giatsoglou et al. 2017). The latter constructs a binary classifier by training a large number of sentiment polarities, and allows the classifier to determine the emotional tendency. Commonly used classifiers include naive Bayes classifier, decision tree, k-nearest neighbor classifier, support vector machine, artificial neural network, etc. For example, Rahab (2017) used support vector machines

---

to classify the sentiments of Algerian newspaper websites, Sabatovych(2019) trained a simple Bayesian sentiment classifier to analyze tweet sentiment to predict the possibility of Brexit. At present, neural networks are also quite popular in sentiment analysis applications, and can be used to identify the sentiment of a single word in a review or even the sentiment of text and video data(Chaturvedi et al., 2019).

Anomaly detection refers to identifying data objects that deviate from standard or normal values. The data objects that deviate significantly from expected values are also called outliers, such as short-term peaks. Anomaly detection methods can be categorized into two types. The first method is mainly to classify data objects so that each data object is marked as anomalous or non-anomalous. Commonly used algorithms include classification regression trees, neural networks, etc. Another method is mainly to predict the future trend of data based on statistical models, set a confidence interval, and measure the degree of deviation between the actual value and the predicted value. If the degree of deviation is too large, it is regarded as abnormal. Commonly used algorithms include exponential smoothing, ESD algorithm, ARIMA model, etc. At present, there are few studies on anomaly detection based on time series data in the field of corporate public opinion, most of which are concentrated on decision trees, neural networks, and other means.This paper uses the Seasonal Hybrid ESD (SH-ESD) algorithm(Hochenbaum et al., 2017) to detect anomalies in time series data. The algorithm is a statistical-based anomaly detection method, which can detect local anomalies and global anomalies with low complexity, fast calculation speed, and strong generalization ability.

Based on the important role of enterprise public opinion early warning in enterprise management, it is necessary to conduct special research on intelligent judgment and abnormal identification of positive and negative public opinion of the enterprise. To this end, this paper raises research questions: (1) How to assess the emotional tendencies toward Chinese multinational companies implied by overseas news media reports? (2) How to identify sudden anomalies of corporate public opinion so that companies can respond quickly?

This article takes Huawei, a Chinese multinational company, as a case analysis object, uses sentiment analysis methods to calculate the sentiment value of the company's overseas news public opinion, and uses anomaly detection algorithms to identify the abnormal points of the company's public opinion, so as to further analyze the events and reasons leading to the abnormal appearance of enterprise public opinion.

II. RESEARCH DESIGN

2.1 Research Framework

The research framework for anomaly analysis of enterprise public opinion is shown in Figure 1, which is mainly divided into three steps: sentiment score calculation of news, outlier detection and enterprise public opinion analysis. First of all, this article takes Huawei's corporate public opinion as the research object, and crawls articles published by foreign news media as data sources. The main work at this stage includes news data collection, news text data pre-processing, construction of MPQA sentiment dictionary, and calculation of sentiment scores of overseas news through semantic-based sentiment analysis technology. Secondly, after calculating the comprehensive sentiment score of news with time series data characteristics, the S-H-ESD algorithm is used to detect anomalies in the time series. Finally, not only determine whether the emotions reflected by the outliers are negative or positive, but also track the news events of the outliers, and then visually analyze the news media's concerns through the word cloud map, so that the company can be informed of the abnormality in time and make correct decisions.

Figure 1:Research framework for analysis of abnormal corporatepublic opinion
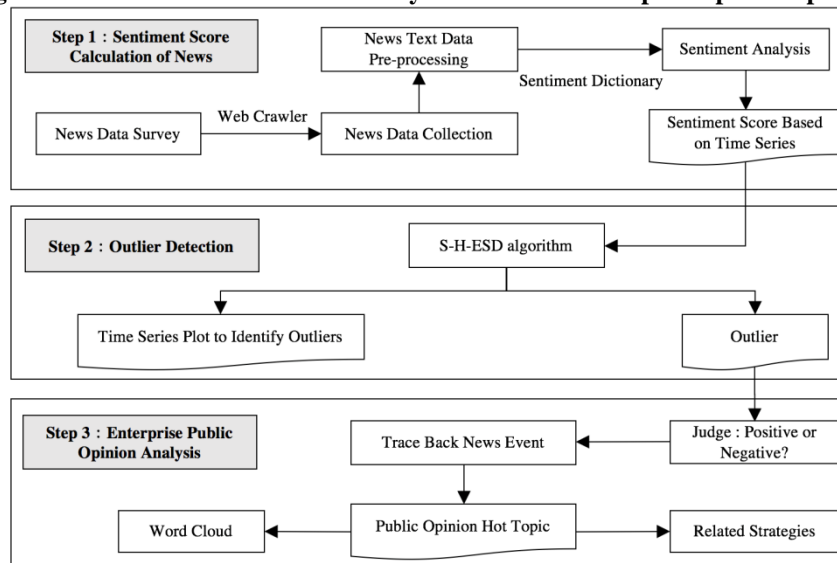
2.2 Data Collection and Pre-processing

In terms of news data collection, this paper collected the time, title, source and content of reports published by overseas news media. A total of six authoritative news sites were selected, including The Washington Post, USA Today, New York Post, Chicago Tribune, The New York Times, Los Angeles Times. The search keyword was "Huawei", and the data collection date was from January 1, 2012 to December 20, 2019. After deleting the news reports with video and other unstructured data as the main content, 1595 overseas news reports about Huawei were kept as the research object. The distribution of news data for each year is shown in Figure 2. It is worth noting that in 2018 and 2019, due to the US-China trade war and Huawei's 5G technology development and other reasons, overseas news media have paid great attention to Huawei companies, and the number of overseas news releases has also increased sharply.

Figure 2:Annual distribution of overseas news reports with Huawei as the keyword
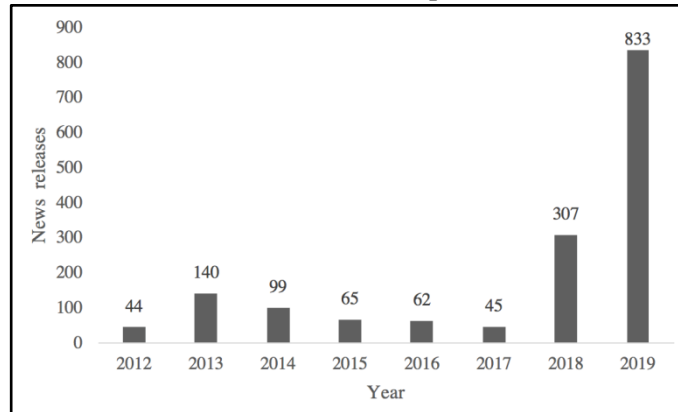
2.3 News Sentiment Analysis

Sentiment analysis in different fields needs to build a sentiment dictionary suitable for that field. Multiple-Perspective Question Answering Corpus contains 535 manual-annotated news articles from different news sources. Therefore, this paper uses the sentiment dictionary in MPQA Corpus to calculate the positive sentiment score Pos_{ij} and negative sentiment score Neg_{ij}. The results are shown in Table 1. Among them, **i** represents the corresponding date, **j** represents different media, and polarity represents the emotional polarity of the news. If the positive sentiment score is greater than the negative sentiment score, the emotional polarity of the news is determined to be **Positive**, otherwise, it is **Negative**. If the sentiment scores are equal, it is determined to be **Neutral**.

Since the media does not report the same company every day, on the day when the media does not publish news reports about the company, the missing sentiment score is automatically defaulted to 0, and the news's sentiment polarity is defaulted to Neutral. The accuracy rate of the sentiment analysis algorithm reaches 89%, indicating that the algorithm in this paper can correctly divide the sentiment polarity of news.

Table 1:News sentiment score and sentiment polarity

News Source	Date	News Title	Pos_{ij}	Neg_{ij}	Polarity
None	2019-01-01	None	0	0	Neutral
None	2019-01-02	None	0	0	Neutral
The Washington Post	2019-01-03	Three things that the trade war with China won't change	31.7	12.4	Positive
The New York Times	2019-01-03	Chinese Consumers' Confidence Sags, Casting a Pall Over the Global Economy	29.8	21.9	Positive
The New York Times	2019-01-03	U.S. Renews China Travel Warning, Citing Risk of Arbitrary Detention	8.1	12.8	Negative
The Washington Post	2019-01-03	How China's Spies Became Key Players in the Trade War	27.0	7.9	Positive
The New York Times	2019-01-04	In Price and Value, Chinese Phone Makers Outpace Apple in Much of the World	29.7	15.0	Positive
Los Angeles Times	2019-01-04	Retired general warns against letting China dominate 5G networks	28.2	12.2	Positive
The Washington Post	2019-01-04	Canada's leaders must reevaluate ties to China	30.4	16.2	Positive
The New York Times	2019-01-05	Why My Chinese Dad Switched From an iPhone to a Huawei	28.3	8.1	Positive
……	……	……	……	……	……
The New York Times	2019-12-20	At the Edge of the World, a New Battleground for the	16.8	18.7	Negative

The Washington Post	2019-12-20	U.S. and China As Trump shuns US multilateralism, China ups diplomatic ante	48.0	18.4	Positive

Considering that different media may have published reports on the company on the same day, for example, in Table 2, on January 3, 2019, four news media have made news reports on Huawei companies. Therefore, this article establishes the average positive sentiment score $AvgPos_i$, **n** indicating the number of news of a certain company on that day:

$$AvgPos_i = \frac{\sum_1^n Pos_{ij}}{n}$$

Similarly, the average negative sentiment score $AvgNeg_i$ is:

$$AvgNeg_i = \frac{\sum_1^n Neg_{ij}}{n}$$

Based on these two indicators, a comprehensive sentiment score $NetPos_i$ is constructed. The formula is as follows:

$$NetPos_i = AvgPos_i - AvgNeg_i$$

$NetPos_i$ reflects the overall sentiment evaluation of all media to enterprise on the day. If $NetPos_i$ is positive, it means that the overall evaluation of the enterprise is optimistic, on the contrary, if $NetPos_i$ is negative, it indicates that the enterprise is currently facing or is about to face a threat. The study simplifies the time series data by constructing a comprehensive sentiment score, which will be beneficial to the subsequent detection of corporate public opinion anomalies. Taking "day" as the unit, the comprehensive sentiment score based on time series is shown in Table 2 (taking the time window data from January 1, 2019 to December 20, 2019 as an example).

Table 2:Comprehensive sentiment score based on time series

Date	$NetPos_i$
2019-01-01	0
2019-01-02	0
2019-01-03	10.40
2019-01-04	14.83
2019-01-05	20.20
……	……
2019-12-20	13.85

2.4 Anomaly recognition and analysis of corporate public opinion

This paper imports the comprehensive sentiment score data based on time series from 2012 to 2019, and runs the S-H-ESD algorithm. The results are shown in Figure 3 and Table 3. Taking 2019 as an example, Figure 3 shows that there were 16 negative outliers and 5 positive outliers.
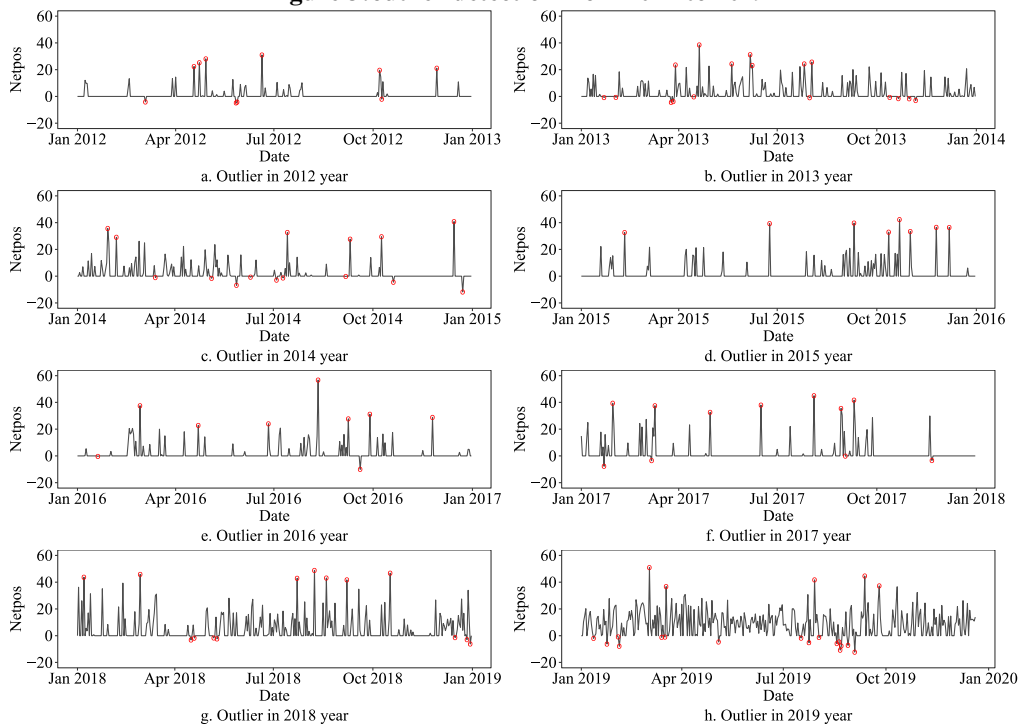
Figure 3:outlier detection from 2012 to 2019

a. Outlier in 2012 year

b. Outlier in 2013 year

c. Outlier in 2014 year

d. Outlier in 2015 year

e. Outlier in 2016 year

f. Outlier in 2017 year

g. Outlier in 2018 year

h. Outlier in 2019 year

Table 3:Abnormalities of corporate public opinion (take 2019 as an example)

News Source	Date	NetPos$_i$	News title	Polarity
The New York Times	2019-01-12	-2	Huawei Fires Employee Arrested in Poland on Spying Charges	Negative
Los Angeles Times	2019-01-24	-6.25	China's arrest of Australian writer is called 'hostage diplomacy' Bing comes back online in China, but there are more questions than answers	Negative
The New York Times	2019-02-03	-0.9	Beating China Is a Team Sport	Negative
The New York Times	2019-02-04	-7.9	How U.S. Criminal Laws Became Weapons in the China Trade War	Negative
USA Today	2019-03-03	50.9	The new phones are coming: The real-world impact of foldable screens and 5G	Positive
…	…	…	…	…
The New York Times	2019-09-12	44.6	With Trade Talks Looming, U.S. and China Move to Relax Tensions	Positive

Enterprises can check the date, related news articles and word cloud maps corresponding to the abnormal points, in order to understand the reasons for the abnormal media emotions, which is not only helpful to assist companies to respond to the crisis of public opinion in time, but also helps companies to take advantage of opportunities to enhance publicity. Next, this article takes negative and positive outliers as examples and analyzes the abnormality of corporate public opinion with Huawei as the main body of the enterprise.

2.4.1   Positive Public Opinion Analysis

Take the positive outlier that appeared on March 3, 2019 in Table 3 as an example. The score of this outlier reached 50.9. The news comes from "USA Today" and reports "Huawei's first 5G folding screen mobile phone- —Huawei MateX was released globally, known asthe world's fastest foldable 5G mobile phone', which shows that the emergence of Huawei's first 5G folding screen mobile phone has driven the upsurge of the media. As can be seen from Figure 4, words such as "5G", "foldable", "smartphone", and "design" that are closely related to 5G folding screen mobile phones have received media attention.

The emergence of positive outliers indicates that the company's new products or new technologies have received strong media attention and promoted the media atmosphere to a climax, which can take this opportunity to strengthen the company's brand promotion to improve consumers' brand image perception.

Figure 4:News keywords on March 3, 2019

2.4.2   NegativePublicOpinion Analysis

Taking a negative abnormal point in December 2018 as an example to analyze the reasons for the negative public opinion of Huawei enterprises, the outlier appeared on December 27, 2018, and the sentiment score reached -3.1. It was discovered that the news came from The New York Post and reported that "President Trump is considering issuing an executive order in the new year, announcing a national emergency and prohibiting American companies from using products made by Huawei and ZTE.". The report implied that the media was not optimistic about the situation that Huawei companies faced. As can be seen from Figure 5, words such as "wireless", "equipment", "order", "government", and "security" have received media attention.

Following up on the incident, it was discovered that on May 15, 2019, the US President Trump signedanexecutive order prohibiting US companies from doing business with foreign telecommunications operators (implying Huawei and ZTE). From December 27, 2018to May 15th, 2019,it took a total of 5 months, and accordingly provided enterprisecertain Coping time. Therefore, when a negative outlier occurs, companies need to increase certain vigilance and do a good job of crisis warning.

Figure 5:News keywords on December 27, 2018

By sorting out the negative anomalies that occurred between 2012 and 2019, this article found that most of the negative public opinion of Huawei companies is related to international political and economic factors such as "national security", "network security", "surveillance", and "government subsidies". In 2012, there were news reports that the United States accused Huawei and ZTE of threatening national security, which affected the introduction and purchase of Huawei technology products by other countries. In the same year, the EU accused the Chinese government of dumping Chinese products overseas by subsidizing companies such as Huawei and ZTE. These two incidents have caused negative sentiment towards Huawei from overseas media. The negative anomalies that appeared intermittently from 2012 to 2018 are mostly related to these two events. As shown in Figure 3, the MengWanzhou incident occurred in 2018 and US-China trade frictions escalated. A large number of negative outliers have gathered along with a large number of international media reports.

2.5 Conclusions

With the development of social media, companies need to understand public opinion in time to make correct decisions, so the management of corporate public opinion is of great value to the company. This paper takesHuaweias a case study object to analyze its overseas public opinion. Finally, it was concluded that the emergence of Huawei's positive public opinion in the international market is related to the release of Huawei's new technology and new products, and the emergence of Huawei's negative public opinion is related to international political and economic factors such as "national security" and "Subsidies". These conclusions are basically consistent with people's cognition, proving the effectiveness of the method used in this article.

This article also has certain limitations. The sentiment analysis method based on machine learning is more accurate than the sentiment analysis method based on basic semantics used here, which can be improved later using the sentiment analysis method based on machine learning. Using technology such as topic tracking

and knowledge graphs to dig deeper into corporate public opinion can also be one of the future research directions.

BIBLIOGRAPHY

[1]. Liyan, X. . (2013). On the evaluation system of the social influence of enterprise public opinion on internet based on microblog. Journal of Intelligence, 32(5), 6-7.

[2]. Xun, L. , Yuan, X. U. , Zhiyu, L. I. , Yuefeng, M. A. , & Yu, L. . (2017). A review on corporate public opinion for social networks. Chinese Journal of Management, 14(06), 925-935.

[3]. Tversky K A. (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica, 47(2), 263-292.

[4]. Cheng, K. , Li, J. , & Liu, H. . (2017). Unsupervised Feature Selection in Signed Social Networks. the 23rd ACM SIGKDD International Conference. ACM, 777-786.

[5]. Jun-Ling, Z. . (2016). A summary of network public opinion information mining in china. Information Science, 34(11), 167-172.

[6]. Korayem M, Aljadda K, Crandall D. (2016). Sentiment/Subjectivity Analysis Survey for Languages other than English. Social Network Analysis & Mining, 6(1), 1-17.

[7]. Chaturvedi I, Ragusa E, Gastaldo P, et al. (2017). Bayesian Network based Extreme Learning Machine for Subjectivity Detection. Journal of the Franklin Institute, 355(4), 1780-1797.

[8]. Giatsoglou M, Vozalis M G, Diamantaras K, et al. (2017). Sentiment Analysis Leveraging Emotions and Word Embeddings. Expert Systems with Applications, 69, 214-224.

[9]. Hichem Rahab, AbdelhafidZitouni, &MahieddineDjoudi. (2017). SIAAC: Sentiment Polarity Identification on Arabic Algerian Newspaper Comments. Computational Methods in Systems & Software. Springer, Cham.

[10]. Sabatovych I. (2019). Use of Sentiment Analysis for Predicting Public Opinion on Referendum: A Feasibility Study[J].The Reference Librarian, 60(3), 202-211.

[11]. Chaturvedi, I. , Satapathy, R. , Cavallari, S. , & Cambria, E. . (2019). Fuzzy commonsense reasoning for multimodal sentiment analysis. Pattern recognition letters, 125(JUL.), 264-270.

[12]. Hochenbaum, J. , Vallis, O. S. , &Kejariwal, A. . (2017). Automatic anomaly detection in the cloud via statistical learning.