

Analysis of Netizens' Social Mentality of Sina Weibo Hot Network Security Incidents



Meihui Ren

School of Marxism, Ji'nan University, Guangzhou, Guangdong, China

Email: renmeihui518@163.com

Abstract:

This study takes the network public opinion of Sina Weibo as the starting point, and uses big data mining technology to technically capture three types of network security events and user comments on Sina Weibo in the new era. LIWC (Linguistic Inquiry and Word Count) software conducts machine-assisted text analysis on the main viewpoints in the progress of network security incidents, aiming to explore the social mentality of netizens such as emotions and rationality in different types of network security incidents.

Keywords: *LIWC, hot cybersecurity incidents, emotion, rationality, social mentality*

1 INTRODUCTION

In the information age, the Internet continues to accumulate massive data on people's behaviors and texts, bringing unprecedented opportunities to psychological science [1] [2] [3]. The current Internet space is roughly divided into five types of information fields: news websites, forums, Weibo, WeChat and news clients [8]. The 49th "Statistical Report on Internet Development in China" [5] by the China Internet Network Information Center shows that as of December 2021, the number of Internet users in my country has reached 1.032 billion, an increase of 42.96 million compared with December 2020, and the Internet penetration rate has reached 73.0%. General Secretary Xi Jinping has also repeatedly emphasized the dialectical and unified relationship between network security and informatization development. [6] In recent years, a situation of "two microblogs and one end" (Weibo, WeChat and Toutiao client) has gradually formed. Although WeChat is the most active information dissemination platform, it is not a space for public opinion in society. Netizens make immediate judgments of online incidents, often based on emotions and intuition, as well as the deeper social psychology behind them. Language is an important medium for understanding individual emotional states, social relationships, thinking styles and even distinguishing individuals [4]. Since the 1980s, Pennebaker et al. have invented a text analysis tool based on a computer software program, "Language Exploration and Word Counting" (LIWC), when studying the therapeutic effects of emotional writing. Software for quantitative analysis of word categories (especially psychological words). After more than two decades of development, due to the good validity of LIWC, the tool has been widely used by psychological researchers in various aspects [7].

To sum up, this study uses Sina Weibo as a sample to investigate typical events affecting network security, and conducts technical capture of network security events and user comments. The main viewpoints in the progress of network security incidents are machine-assisted text analysis, and reflect the social mentality of netizens in different types of network security incidents.

2 RESEARCH METHODS

2.1 Identify alternative hotspot cybersecurity incidents

This study mainly identifies three types of cyber incidents: political security, economic security and technical security. Two representative cases are captured for each type: Snowden's "Prism" incident (2013) and Hillary's "Email". The incident (2016) involved national sovereignty security and the construction of a community with a shared future on the international network. The "Bitcoin" ransomware (2015) and the "Eternal Blue" incident (2017) involved the economic security of the combination of Internet economic fraud and ransomware. The Apple users information theft incident (2015) and the Yahoo leak incident (2016) involved the security of network technology, and the leakage of user information greatly threatened the security of citizens' personal information.

2.2 Hotspot network security event preprocessing

First, the keywords for time search are extracted, and by analyzing the content of the event, two coders summarize the keywords separately, and then return to Weibo to search, and select the keyword combination with a large and accurate return value. On average, 2-3

keywords are extracted per event. The main purpose is to ensure that event-related texts are obtained as accurately and comprehensively as possible when scraping Weibo.

Second, determine the time event range. For example, the "prism gate event" is temporally sorted to determine the time range in which the time occurs. The start time is the time of the first media report, and the end time of the event is determined according to the amount of discussion on Weibo. If the amount of discussion on Weibo is 0 for 3 consecutive days, the event is considered to be over. If there is a discussion caused by subsequent sub-events, it will be dealt with separately. The searched event data can be accurately located on the network platform through a combined search of keywords and the time when the practice occurred.

2.3 Sina Weibo Crawl and Sort

Firstly, we use web crawlers to crawl Weibo text content related to popular event reports on Weibo based on keywords.

Secondly, combined with the start and end time of the time as the limit, the event keywords are searched to obtain the number of related Weibo texts. This number indicates the popularity of the event on the Weibo platform. The more the number, the more attention it is on the Weibo platform.

Thirdly, the work of deduplication and noise removal is carried out on the correlation number of microblogs, and the correlation degree is guaranteed to be increased to more than 70% after sampling inspection, and the irrelevant microblogs are deleted.

2.4 Text analysis of the investigation content of hot network security incidents

Psychological features of the text were extracted using the simplified Chinese version of LIWC. LIWC is a kind of natural language processing technology, which can quantitatively analyze the text content and calculate the different categories of words in the imported text file. This report mainly uses the "emotional word frequency" indicator.

3 RESEARCH RESULTS

3.1 Analysis of Netizens' Popularity of Hot Network Security Events

Comparing the three types of hot network security incidents initially screened, it can be found that the concerns of netizens are shown in Table 1:

Table 1: Analysis of the popularity of different types of network security incidents

Event name	Number of comments	Number of people forwarded	Total number of participants	Number of media participation in the first layer	Number of propagation layers
Bitcoin blackmail virus event	25467	29958	55425	28	2
Hillary Clinton's "mail gate" incident	20263	21932	42195	21	2
Eternal Blue event	15846	23902	39748	18	2
Snowden "prism gate" incident	1917	47237	49154	26	2
Apple user information theft	1715	936	2651	5	2
Yahoo leaks	650	687	1337	12	2

First of all, in terms of the number of comments, the economic security incident received the highest attention, and the number of comments on bitcoin blackmail virus was 25467; The second is political security events. The number of comments on Hillary Clinton's "mail gate" event is 20263; Finally, there are technical security incidents. The number of comments on Yahoo leaks is 650.

Secondly, in terms of the number of people forwarded, Snowden forwarded 47237 political security incidents; The second is the largest number of economic security events forwarded, with 29958 people forwarding bitcoin extortion virus; Finally, there are technical security incidents. The number of comments on Yahoo leaks is 687.

To sum up, from the total number of participants (comments + forwarding), economic and security events are the most concerned, with the total number of bitcoin blackmail virus participants reaching 55425; The second is the political security events. The total number of Snowden prism gate events involved reached 49154; Finally, there are technical security incidents. The total number of Yahoo leaks involved is 1337 respectively.

Finally, in terms of the number of media participants in the first layer, economic security events are the most concerned. The number of media participants in the first layer of bitcoin blackmail virus is 28; The second is political security events. The number of media participation at the first level of Snowden prism gate event and Hillary Clinton mail gate event is 26 and 21 respectively; Finally, there are technical security incidents. The number of media participants in the first

layer of Yahoo leaks is 12, and the number of Apple users' information theft is 5.

Based on the above data analysis indicators, it can be seen that economic security events are the most concerned by the media and Internet users, followed by political security events, and finally technical security events. Netizens are more intuitively concerned about events related to their vital interests. Economic and political security are closely related to everyone's personal safety, property safety and national sovereignty. However, the attention to the security in the field of technology is relatively low, mainly because the correlation between information technology and network facilities and people's daily life is relatively weak, and it is related to the individual scientific, technological and cultural quality of citizens.

3.2 Analysis of Internet users' involvement in hot network security events

Netizens' involvement in hot network security events is mainly analyzed from the aspects of positive emotion, negative emotion, anxiety, anger and sadness. See Table 2.

Table 2: Emotional index of Internet users on network security events

Event name	Positive emotions	Negative emotions	anxiety	Feeling of anger	Sadness
Bitcoin blackmail virus	0.024	0.019	0.006	0.005	0.001
Apple user information theft	0.183	0.275	0.008	0.014	0.004
Snowden "prism gate" incident	0.032	0.024	0.004	0.009	0.001
Hillary Clinton's "mail gate" incident	0.205	0.21	0.027	0.088	0.0278
Yahoo leaks	0.027	0.013	0.004	0.004	
Eternal Blue	0.133	0.021	0.004	0.011	

First of all, from the perspective of positive emotions, the political security event had the highest positive emotions. The positive emotions of Hillary Clinton e-mail event was 0.205, followed by technical security events. The positive emotions of Apple user information theft event was 0.183, and finally economic security

events. The positive emotions of eternal blue event was 0.133. From the perspective of negative emotion, the negative emotion of technology security events is the most obvious. The negative emotion of Apple user information theft event is 0.275, followed by political security events. The negative emotion of Hillary Clinton e-mail event is 0.21, followed by economic security events. The negative emotion of eternal blue event is 0.03.

Secondly, from the perspective of the anxiety of Internet users, the anxiety of political security events is the strongest. The anxiety of Hillary Clinton e-mail door event is 0.027, followed by the technical security event. The anxiety of Apple users' information theft event is 0.008, and finally the economic security event. The anxiety of bitcoin extortion virus event is 0.006.

Finally, from the perspective of netizens' sadness, political security events have the strongest sadness, Hillary Clinton's email door event has a sadness of 0.0278, followed by technical security events, Apple user information theft event has a sadness of 0.004, and finally economic security events, and bitcoin blackmail virus event has an anxiety of 0.001.

To sum up, it can be seen from the positive and negative emotions that, on the whole, the overall emotions of Internet users are mainly concentrated in political security events and technical security events, and the word frequency of positive and negative emotions is relatively high. The word frequency of anxiety and sadness is high in political security events.

3.3 Analysis on attitude index of Internet users towards network security events

Based on the emotional analysis of Internet users, the attitude index of Internet users can better reflect the mentality of Internet users. As shown in Table 3.

Table 3: Analysis of netizen attitude index

Event name	Sensibility	anger	sadness	Rationality	Degree of hope	NonPart
Bitcoin blackmail virus event	0.049	0.005	0.001	0.203	0	0
Apple user information theft	0.054	0.014	0.004	0.178	0	0.001
Snowden "prism gate" incident	0.075	0.009	0.001	0.2	0.006	0.003
Hillary Clinton's "mail gate" incident	0.05	0.088	0.028	0.185	0.001	0.001
Yahoo leaks	0.045	0.004	0	0.199	0	0
Eternal Blue event	0.042	0.011	0	0.213	0	0

Netizens' attitude index is the internal attribution mechanism and final attitude type of netizens towards hot events. Netizens' attitudes towards network security incidents can be divided into positive attribution and negative attribution, specifically divided into hope, expectation of problem solving, pessimism and indifference. Hope is a collection of positive emotions + high sense of control, such as awesome, happiness, etc; Expected problem solving degree is a combination of negative emotion and high sense of control, anger and rolling; Pessimism is the combination of negative emotion and low sense of control, such as pit dad and sadness; Apathy is a combination of positive emotion and low sense of control, such as floating clouds, onlookers, etc. When the optimism and expectation of hot network security events exceed the pessimism and indifference, the attitude index is >0 ; On the contrary, the attitude is <0 . It can be seen from the attitude index of Internet users on various types of network security leave that the highest sensitivity is political security incidents, Snowden's "prism gate incident" is 0.075, followed by technical security incidents; The highest degree of rationality is the economic security event. The eternal blue event and bitcoin blackmail virus are 0.213 and 0.203 respectively; The highest degree of hope is the political security event, and the Snowden prism gate event has a degree of hope of 0.006.

Netizens' attitudes towards political security incidents, such as Hillary Clinton's "e-mail gate" incident, were slightly more negative than positive, and their sense of anger was higher than their sense of anxiety and sadness. Netizens' high-frequency words in their comments and replies, such as Trump's victory and Google's shame, showed that netizens' attitude towards political democracy tended to be rational, and citizens' awareness of democratic politics was constantly enhanced; Netizens' attitude towards economic security events, taking eternal blue as an example, positive emotions are higher than negative emotions, and feelings of anger and anxiety are higher than feelings of sadness. Netizens' high-frequency words in comments and replies, such as "watch out" and "turn off the port", indicate that Internet names are beginning to be more vigilant about personal information security, and their awareness of information security is enhanced; Netizens' attitudes towards technology security incidents take the theft of Apple users' information as an example. Negative emotions are higher than positive emotions, and feelings of anger and anxiety are higher than feelings of sadness. Netizens' high-frequency words in comments and replies, such as "atmosphere", "terrible" and "also stolen", indicate netizens' hidden worries about information technology, and expressed strong dissatisfaction with apple and other large companies' violations of citizens' personal information privacy.

4 CONCLUSION

By sorting out and analyzing the typical network security events that affect the process of network security in China in the new era, we can roughly draw the following conclusions: Internet users pay more attention to social security topics related to their immediate interests, show high positive emotions, and have a strong sense of responsibility and mission; The emotion of netizens is easily infected by the network emotion, and the herd mentality is obvious; On the Internet platform, Internet users have an international perspective and pay high attention to international political events; The overall attitude of netizens is positive and stable, showing strong dissatisfaction and anxiety in major security incidents.

ACKNOWLEDGMENT

Thank you for the big data analysis of this article by Professor He Lingnan of the communication laboratory, School of communication and design, Sun Yat sen University.

FUDING SOURCES

This article is the national outstanding young and middle-aged ideological and political theory course teaching Teacher-selected funding plan "Socialist core values lead the Internet popularity" Research on the Mechanism of Culture" (18JDSZK128) Phased Achievements.

REFERENCES

- [1] Adjerid, I., & Kelley, K. (2018) Big data in psychology: a framework for research advancement *American psychologist*, 73 (7), 899-917.
- [2] Le Guo an, Dong yinghong, Chen hao, Lai Kai sheng (2013) Online text emotion analysis technology and application heart. *Advances in science*, 21 (10), 1711-1719.
- [3] Park, g., Schwartz, H. A., eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., & Seligman, M. E. P. (2014) Automatic personality assessment through social media language *Journal of personality and social psychology*, 108 (6), 934-952.
- [4] Tausczik, Y.R., & Pennebaker, j.w.the psychological meaning of words: LIWC and computerized text analysis methods [J]. *Journal of language and social psychology*, 2010, (29)
- [5] The 49th statistical report on the development of Internet in China.

- [6] Xi Jinping: Speech at the Symposium on Network Security and Informatization (April 19, 2016) Xinhuanet.http://www.xinhuanet.com/politics/2016-04/25/c_1118731175.htm
- [7] Zhang xin yong LIWC: a text analysis tool based on word measurement [J]. Journal of Southwest University for Nationalities (Humanities and Social Sciences Edition) 2015 (4).
- [8] Zhang Zhian, He ling nan, Tan lini, Li Yu: analysis of positive energy situation and Countermeasures in the Internet public opinion field, legalization of cyberspace - annual report on Internet and national governance (2015), commercial press, 2015 edition.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

