Research on Evolution of Public Opinion at Continuous Emergencies

Xueyan Liu[1], Lan Zhang[2]

liuxy@cqupt.edu.cn[1], 1850970822@qq.com[2]

School of Economics and Management Chongqing University of Posts and Telecommunications Chongqing, China

Abstract. The data from the 51st "Statistical Report on the Development Status of China's Internet Network" by CNNIC reveals that China's internet user base has reached 1.067 billion, underscoring its integral role in national life. Sina Weibo, as China's largest social media platform, not only addresses the entertainment needs of users but also amplifies diverse voices, including those of influential mainstream media users. Concurrently, there has been a sharp increase in the frequency and intensity of emergencies over the past few years. These events, due to their unpredictable and sensitive nature, rapidly form and widely disseminate public opinion shortly after their occurrence. Without proper guidance and control, they can have a detrimental impact on society. The accumulation of continuous emergencies further intensifies public emotions, escalating the challenges for governmental control. Consequently, effectively guiding the network public opinion of continuous emergencies and enhancing the government's capacity for network public opinion governance during crises represent crucial tasks. In this paper, from the perspective of continuous emergency network public opinion, starting from text mining, using word frequency analysis, LDA topic model, and other methods, with the help of network public opinion life cycle theory, web crawler, TF-IDF algorithm, Chinese word segmentation algorithm and other technologies, using the blog text from Sina Weibo as the data source, this paper analyzes the specific changes of the network public opinion of continuous emergencies in Chongqing in different periods. The goal is to provide insights for future responses to network public opinion during such events. The paper concludes that different topics emerge in various stages of continuous emergencies in Chongqing.

Keywords: Continuity emergencies, Public Opinion, LDA topic model

1 Introduction

In recent years, whether facing natural disasters such as the significant rainfall in Henan on July 20, Australian bushfires, or addressing public health emergencies like the COVID-19 pandemic and the Ebola virus in the Congo, as well as emergencies like the capsizing of a passenger ship in Guizhou's Liupanshui, China and governments globally have faced formidable challenges. Due to the sudden occurrence, unpredictability, and high uncertainty, these events may impact the fundamental values and behavioral frameworks of societal systems, with the public being the primary subject affected. Against this backdrop, this study will commence with Weibo data, focusing on continuous emergencies and taking Chongqing as a case study. The research aims to enhance public understanding of emotional changes and fluctuations during continuous

PMIS 2024, March 15-17, Changsha, People's Republic of China

DOI 10.4108/eai.15-3-2024.2346399

emergencies, facilitating emergency management and strengthening the government's capability to address such network public opinion.

The mining and extraction of themes is one of the most important contents in the study of network public opinion, and its research is also increasing year by year at home and abroad. Based on the Twitter platform, Su[1]and others explored the public's thematic views and attitudes on the focus of Japan's Fukushima nuclear accident. To study the effect of e-cigarette policy implementation, Lazard[2] et al. conducted theme mining and analysis of the views and opinions held by Twitter users. By grasping Tencent news reports and Zhihu comments, Wang Yuefen[3] used the LDA model to mine their respective topics of concern and compare their similarities and differences, and found related topics. From the perspective of group polarization, Lu Guoqiang[4]established an extreme topic identification model of network public opinion, and verified that the extreme topic view of public opinion group polarization can be effectively identified through the example of emergency public opinion data. This paper will concentrate on a large volume of online user discourse following continuous sudden disaster events. Utilizing the lifecycle theory, the study will analyze the multi-dimensional aspects of the public opinion evolution process, exploring the patterns of netizens' evolving concerns and emotions after the occurrence of continuous emergencies.

2 Theoretical foundation

2 .1 Life cycle theory

The Lifecycle Theory, first proposed by Kamen in 1966, is a standard framework used to analyze the various stages of development, growth, maturity, and decline experienced by markets. Similarly, the Public Opinion Cycle refers to the lifecycle of online public opinion events, encompassing the phases of inception, growth, eruption, decline, and ultimately extinction as opinions disseminate through the internet. Currently, there is no unified standard in academia for the phase delineation of online public opinion dissemination. This lack of consensus stems from the general agreement that online public opinion follows the universal laws of lifecycle phenomena, leading to the division of the evolutionary process into distinct stages. Depending on different criteria for division, these stages may be categorized into three, four, five, six, or more phases. Overall, the three-stage[5] classification is the most widely applied, providing clarity and comprehensibility, supported by a substantial number of examples. However, this model has limitations in explaining fluctuations and repetitions during the intermediate stages of online public opinion events. Four-stage[6], five-stage[7], and six-stage[8] models extend from the three-stage blueprint, offering a more comprehensive reflection of the diverse characteristics in the evolution of online public opinion. These extended models are more detailed but exhibit weaker universality.

2.2 LDA topic model

The Latent Dirichlet Allocation (LDA) topic model, introduced by Blei and Lafferty in 2003, is a probabilistic generative model designed for the extraction and classification of topics within documents. LDA is a three-layer Bayesian generative model structured as a "document-topic-word." The fundamental premise is that a document contains multiple topics, and each word is generated from a specific topic. The distribution of topics within a document follows a Dirichlet

distribution with parameter α, while the distribution of words for each topic follows a Dirichlet distribution with parameter β. Typically, each document is associated with a primary topic representing its core content, and the LDA topic model can provide the distribution of topics and keywords for a given document. In the LDA topic model, the generation of a document proceeds as follows:

1) Sample the topic distribution θ_i for the i-th document from the Dirichlet distribution hyperparameter α.
2) Extract the topic $z_{i,j}$ for the j-th word in document i from the multinomial distribution θ_i.
3) Sample the word distribution $\varphi z_{i,j}$ for the topic $z_{i,j}$ from the Dirichlet distribution hyperparameter β.
4) Sample the final word $w_{i,j}$ from the multinomial distribution $\emptyset z_{i,j}$ of the selected feature words.

Here, α and β are typically manually set constants, w represents the observable parameter (i.e., the j-th word in the i-th document, which is observable), and θ, z, and φ are parameters to be inferred. The most common estimation method is Gibbs sampling, which is used to estimate the parameters in the model. The LDA probabilistic graphical model is illustrated in Figure 1.

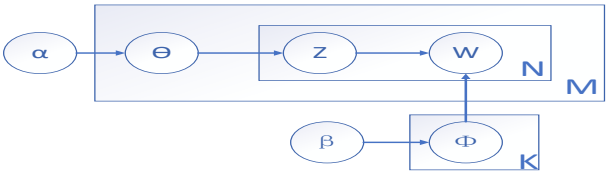

Fig. 1. LDA probability graph model structure

3 Analysis of public opinion evolution of continuous emergencies

3.1 Selection of Continuity Emergencies

Since July, the temperature in Chongqing has been consistently surging towards 40°C almost every day. Meteorological data from the Chongqing Meteorological Observatory reveals that from July 1 to August 18, the range and intensity of high-temperature weather in Chongqing have exceeded those of 2006, ranking as the highest in history. In August, Chongqing suffered the most severe high temperature and drought in more than 60 years. The temperature in many districts was above 40 degrees, even soaring more than 45 degrees! In response to the high temperatures, notices were successively issued in Sichuan and Chongqing to relieve the pressure on residential electricity consumption. On August 16, the Chongqing Economic and Information Commission, in conjunction with State Grid Chongqing Power Company, issued an "Urgent Notice on Expanding the Scope of Industrial Enterprises Providing Electricity to Residents." From August 15 to August 24, industrial enterprises implemented a high-temperature break, providing electricity to residents. Due to continuous extreme clear and hot weather, wildfires broke out successively in Fuling District and Nanchuan District on August 17 and 18; around 22:00 on August 18, a wildfire occurred in Renlong Village, Zhiping Town, Jiangjin District; wildfires occurred in Dazu District and Tongliang District on August 20; and a fire incident took

place in Kaizhou District on August 21. The burning fire directly threatens the safety of this ancient mountain city. In the days of wildfires, the Chongqing urban area in the early morning was filled with the smell of burning firewood. On August 23, from 0:00 to 24:00, Chongqing reported 21 new locally confirmed cases and 19 new locally asymptomatic infections. As of 24:00 on August 23, there are 74 existing locally confirmed cases and 72 locally asymptomatic infections in Chongqing. At the critical moment of firefighting and rescue in Chongqing, more than 1,000 commanders from brother provinces and cities were sent to assist Chongqing. They went deep into Fuling, Beibei, Banan, Jiangjin, Nanchuan, Dazu, Bishan, Tongliang, and other places to make outstanding contributions to extinguishing mountain fires. Some people on the network describe it as follows: before the support force arrives, it is fire chasing people to run; after the support force arrives, it is people chasing fire to run, leaving the fire demon nowhere to run. By the early morning of August 26, less than ten days, the fire in Chongqing was completely extinguished, with no casualties and loss of important facilities! Chongqing people, the Chinese people, created a miracle! This series of events has triggered a public opinion outbreak on social media platforms, attracting considerable attention and discussion from the public. The dissemination and societal impact of this event align with the fundamental characteristics of the network public opinion of continuous emergencies. Therefore, this event topic is selected as the research case for this paper.

3.2 Data acquisition and preprocessing

For this study, we focus on three continuous emergencies: the extreme heatwave in Chongqing on July 1, the wildfires triggered by high temperatures on August 17, and the sudden surge in COVID-19 cases on August 23, all in 2022. These events have garnered significant attention from both the government and the public. Using a Python web crawler, we collected data related to the keywords "Chongqing high temperature," "Chongqing wildfires," and "Chongqing COVID-19" from July 1, 2022, to September 30, 2022, including blogger information, blogger links, publication times, Weibo content, Weibo links, reposts, comments, and likes. In total, 78,367 data points were collected and exported into an Excel spreadsheet. In the data retrieved through keyword crawling, many contents do not meet the requirements. Therefore, to enhance the quality of data analysis, it is necessary to process the original duplicate and invalid data. This paper primarily addresses the duplicate and invalid data in the continuous emergencies in Chongqing. The "Remove Duplicates" feature in Excel is employed to eliminate duplicate entries. Invalid data encompasses blank entries and Weibo posts irrelevant to the designated topics, such as those containing the keyword but lacking relevant content, including donation-related and celebrity-focused data. The "Sort and Filter" function in Excel is used to initially remove blank portions from the raw data. Subsequently, through manual efforts, the remaining invalid data is meticulously deleted and cleaned. After iterative operations, a total of 35,473 valid data entries are ultimately obtained.

3.3 Network public opinion dissemination stage division

Based on the captured Weibo text data, time series line charts were generated with the evolution date of online public opinion as the x-axis and the number of Weibo posts as the y-axis for both the overall chart (Figure 2). Subsequently, the public opinion heat is divided into stages based on the trend of the line chart. From the graphs, it can be observed that the evolution trend of public opinion exhibits multi-peak characteristics with irregularly distributed peaks and significant

fluctuations. This is attributed to the complex and dynamic nature of public opinion dissemination.

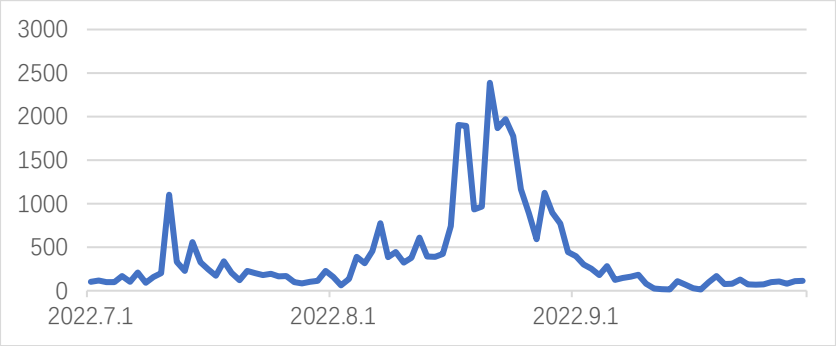

Fig. 2. Overall life cycle diagram

Before July 7, 2022, the number of relevant topics Weibo posts was relatively low. After this date, the count started to increase, reaching the first peak on July 11. Upon analyzing the data, it was found that Chongqing was experiencing a prolonged period of high temperatures and scarce rainfall, with daily maximum temperatures ranging from 35°C to 40°C in most areas, and locally reaching 40°C to 42°C. Simultaneously, the official Sina Chongqing Weibo released the hashtag # Chongqing special severe high-temperature level#, sparking participation from netizens. As the first peak was triggered by the high-temperature event, this period is designated as the formation period. Subsequently, temperatures gradually declined, and the pandemic fluctuated, entering the diffusion period.

Until August 15, when temperatures rose again to 43°C, multiple areas experienced consecutive wildfires. Meanwhile, the severity of the COVID-19 situation increased, leading to another peak in public sentiment. With temperatures above 40°C persisting until the end of August, this stage is classified as the outbreak period. After the wildfires were extinguished, and summer gradually came to an end, the study entered the termination period.

This research categorizes the network public opinion of continuous emergencies into four stages:

1. Formation Period: July 1, 2022, to July 10, 2022

2. Diffusion Period: July 11, 2022, to August 14, 2022

3. Outbreak Period: August 15, 2022, to August 28, 2022

4. Termination Period: August 29, 2022, to September 30, 2022

No analysis is conducted for stages beyond September 30 due to the data collection cutoff date.

3.4 LDA topic model

Through the delineation of public opinion stages in continuous events, the utilization of the LDA model for the analysis of Weibo texts across different periods aids in understanding the varying topics of interest among netizens during distinct phases. In this study, the parameter values yielding the highest coherence score were chosen as the final parameters to determine the number

of topics in each stage. The determined number of topics for each stage is as follows: 5 topics for Stage 1, 4 topics for Stage 2, 4 topics for Stage 3, and 6 topics for Stage 4.

Based on these topic numbers and adjusted parameters, the key terms under each topic were ascertained. The corresponding tables (Tables 1 to 4) present the identified topics for each stage, the key terms under each topic, along with their probabilities, and the assigned names for each topic.

Table 1. Theme distribution in formation period

Themes	Key words
Theme 1: High temperature in summer	stove(0.033), summer(0.029), weather(0.015), isolate(0.014), hot pot(0.011), end(0.008), affect(0.008), travel(0.008), go out(0.008)
Theme 2: Urban epidemic prevention and control	Chengdu(0.046), Guangzhou(0.017), prevention and control(0.015), wish(0.015), plan(0.015), look(0.012), like(0.011), isolate(0.011), travel(0.010), high-temperature vacation(0.010)
Theme 3: Going on a trip	go out(0.027), air conditioner(0.027), travel(0.025), friends(0.016), prevent sunstroke(0.015), child(0.015), temperature(0.015), cool(0.013), school(0.013), early warning(0.013)
Theme 4: Urban epidemic	Nanjing(0.027), Qingdao(0.026), friends(0.021), summer vacation(0.018), go home(0.018), Xinjiang(0.013), plan(0.012), South Central(0.009)
Theme 5: Nucleic acid under high temperature	weather(0.048), nucleic acid(0.023), company(0.017), sun(0.014), summer(0.013), newly increased(0.012), air temperature(0.010), appear(0.009)

Table 2. Theme distribution of diffusion period

Themes	Key words
Theme 1： Newly added cases	nucleic acid(0.050), early warning(0.030), newly increased(0.022), make a definite diagnosis(0.021), heatstroke,(0.018), prevention and control(0.016), Sanya(0.015), cases(0.013), test(0.010), risk(0.010)
Theme 2： Go out at high temperature	night(0.012), go out(0.011), air conditioner(0.010), hotel(0.008), live(0.008), summer(0.008), friends(0.008), everybody(0.007), hot pot(0.007), wish(0.007)
Theme 3： Chengdu tourism	Chengdu(0.057), travel(0.023), plan(0.020), high-temperature vacation(0.018), Changsha(0.017), isolate(0.015), school opens(0.014), wish(0.013), friends(0.013), Beijing(0.010)
Theme 4： High-temperature stove	weather(0.044), temperature(0.027), stove(0.023), air temperature(0.023), highest(0.015), summer(0.013), continuous(0.013), air conditioner(0.012), heatstroke(0.011), exceed(0.009)

Table 3. Theme distribution of outbreak period

Themes	Key words
Theme 1： Mountain fire seeking rain	rain(0.048), nucleic acid(0.040), power rationing(0.028), wish(0.020), too difficult(0.019), hurry up(0.016), unseal(0.014), end(0.014), air conditioner(0.013), sealing and control(0.013)

Theme 2： The impact of regional high temperature weather	region(0.021), weather(0.021), affect(0.014), drought(0.013), newly increased (0.010), air temperature(0.009), rainfall(0.009), electricity consumption(0.009), China(0.009), extreme(0.008)
Theme 3： Cheer up for the city	refueling (0.022), school opens(0.016), live(0.015), assistance(0.013), final(0.013), guard(0.011)
Theme 4： Volunteers put out the fire	volunteer(0.024), fire extinguishing(0.019), wish(0.018), goods and materials(0.018), safe and sound(0.017), stamp out(0.017), China(0.017), fire(0.014), firemen(0.013), rescue(0.011)

Table 4. Theme distribution of the end period

Themes	Key words
Theme 1： Isolation testing	isolation(0.058), test(0.023), personnel(0.023), region(0.021), risk(0.021), policy(0.016), work(0.015), require(0.014), concentrate(0.013), epidemic prevention(0.013)
Theme 2： Go China	China(0.030), cheer up(0.024), earthquake(0.020), hero(0.010), rescue(0.010), country(0.009), stamp out(0.009), spirit(0.009), volunteer(0.009), goods and materials(0.009)
Theme 3： Isolation charges	isolate(0.031), charge(0.028), place(0.021), company(0.020), concentrate(0.020), administrate(0.019), implement(0.018), affect(0.016), Beijing(0.014), examination(0.012)
Theme 4： New cases in cities	newly increased (0.055), Xinjiang(0.035), Guizhou(0.034), make a definite diagnosis(0.032), Shanghai(0.030), Tibet(0.023), cases(0.023), mainland(0.022), Henan(0.017), Zhejiang(0.016)
Theme 5： School opens	school(0.014), school opens(0.014), wish(0.013), end(0.012), National Day(0.011), live(0.010), friends(0.009), go home(0.009), like(0.008), locality(0.008)
Theme 6： Unseal to work	unseal(0.047), start work(0.026), village(0.023), sealing and control(0.022), street(0.021), recover(0.013), stay at home(0.010), rain(0.009)

Based on the probability distribution of topics, the focus of public opinion undergoes shifts across various stages. During the formation stage, discussions center around summer heat (stove accounting for 0.033, weather accounting for 0.048) and regional epidemic policies. As the diffusion stage unfolds, topics encompass the increase in COVID-19 cases, travel during high-temperature weather, and temperature effects (nucleic acid ratio 0.05, weather ratio 0.044), with two out of the four topics related to temperature. With the approach of summer, the focus transitions from the epidemic to the weather. In mid-August, a Chongqing forest fire, triggered by high temperatures, sparks widespread online discussion. Amid the outbreak period, attention shifts to requests for rain (rain proportion 0.048), city refueling (refueling proportion 0.022), and silent volunteers aiding the fire (volunteer proportion 0.024). In the final stage, the wildfires are extinguished, and the city overcomes the crisis. News of school openings (0.014) and the city reopening for work (0.047) raises expectations for normal life. While some topics about the epidemic persist (isolation ratio 0.058, new cases ratio 0.055), they do not significantly impact the overall improvement.

4 Conclusion and discussion

This paper, using continuous emergencies in Chongqing as a case study and applying the four-cycle theory of public opinion life, uncovers the evolution pattern of public opinion dissemination in such scenarios. Unlike previous studies focused on singular cases, this research delves into the evolution process of public opinion topics during continuous emergencies. The findings reveal a rapid dissemination speed within days of the events, with distinct classification types for public opinion topics in different cycles, providing a more comprehensive understanding of netizens' evolving attention. This insight aids emergency management departments in promptly identifying and guiding public opinion events, offering decision support for the government's response to natural disaster events on social media networks. The paper emphasizes the importance of initiating early warning procedures and predicting future event trends based on public opinion development. The government is advised to proactively guide emotions, especially during the initial and concentrated outbreak stages, leveraging media advantages for effective propaganda and timely dissemination of incident facts. In the era of new media, responding to public emergencies should exploit media convenience, ensuring the public's right to information, stabilizing emotions, and influencing the dominant narrative in public opinion. Nevertheless, the research has limitations. For instance, it focuses on continuous emergencies in Chongqing, which may not represent all similar cases. The data is sourced from Weibo, but as just one social media platform, it may not fully depict the emotional evolution and dissemination characteristics across various platforms. To enhance the study's comprehensiveness and draw more universal conclusions, a broader dataset or an extended timeframe may be necessary to cover a wider range of cases.

Acknowledgment. This research was supported by the National Social Science Foundation of China (Grant No.20BSH076).

References

[1] Su Y F.: Analyzing public sentiments online: combining human- and computer-based content analysis. Information Communication & Society. pp. 1-22 (2016)

[2] Lazard A J.: Public reactions to e-cigarette regulations on Twitter: a text mining analysis. Tobacco Control. pp.112-116 (2017)

[3] Wang, Y F.: Topic Identification and Theme Association Analysis of Public Opinion Comments and News Reports. Modern Information, pp. 3-10 (2018)

[4] Lu, G Q.: Research on Identification of Extreme Views in the Network Public Opinion of Emergency Events from the Perspective of Group Polarization. Information and Documentation Services, pp.42-51 (2023)

[5] Burkholder B T.: Evolution of complex disasters. Lancet, pp.112-115 (1995)

[6] Jiang, C B.: Research on a Dynamic Monitoring Model of Network Public Opinion Themes and Sentiments Based on Community Discovery. Information Exploration, pp.56-64 (2023)

[7] Zeng, M.: Research on the Evolution of Weibo Public Opinion Themes in Network Collective Events. Information and Management Research. pp.28-40. (2021)

[8] Mou, X Y.: Analysis of the Evolution of Online Public Opinion Events: A Case Study of the Missing Girl Incident in Hangzhou. New Media Research. pp. 20-23 (2019)