

A Study of Netizens' Attention Preferences for Hotspots in Weibo Public Opinion on Public Emergencies: Postprint

Authors: Wang Linwei

Date: 2023-10-08T00:00:00+00:00

Abstract

[Purpose/Significance] This study aims to quantify the vocabulary frequency of hot-topic netizens' Weibo posts, summarize their concerns regarding sudden public incidents, and thereby identify new breakthroughs in public opinion prevention, control, and governance. [Method/Process] By crawling Weibo data from hot-topic netizens, this study employs the Word2vec vector model and TF-IDF weighting algorithm to obtain indicators such as word categories, part-of-speech proportions, and lexical similarity ratios in the word cloud diagrams of hot-topic netizens' Weibo posts, and finally conducts a comprehensive analysis of these indicators across natural disasters, accidental disasters, public health incidents, and social security incidents. [Results/Conclusion] In public opinion regarding sudden incidents, hot-topic netizens demonstrate considerable attention to the basic components of incidents, including incident figures, incident names, and incident processes. The directional focus of hot-topic netizens across the four types of sudden incident public opinion exhibits differences and is relatively dispersed, yet commonalities still exist. Simultaneously, through analyzing the word cloud diagrams of hot-topic netizens, communication phenomena such as celebrity super topics, hierarchical propagation energy, and similarity polarization were discovered.

Full Text

ChinaXiv Collaborative Journal [Academic Exploration]
Hot Netizen Attention in Weibo Public Opinion on Sudden Public Events
School of Journalism and Communication, Guangxi University, Nanning 530004

Abstract

[Purpose/Significance] This study aims to quantify the vocabulary frequency in hot netizens' Weibo posts, summarize their concerns in sudden public events, and identify new breakthroughs for public opinion prevention and governance. **[Method/Process]** By crawling Weibo data from hot netizens, this study employs the Word2vec vector model and TF-IDF weighting algorithm to obtain indicators such as word class proportion, part-of-speech ratio, and vocabulary similarity ratio in hot netizens' Weibo word clouds, and finally conducts a comprehensive analysis of these indicators across natural disasters, accidental disasters, public health incidents, and social security events. **[Results/Conclusion]** In emergency public opinion, hot netizens show considerable attention to the basic components of events, including event figures, event names, and event processes. Moreover, the attention directivity of hot netizens in the four types of sudden event public opinion exhibits both differences and dispersion, yet commonalities remain. Simultaneously, through analyzing hot netizens' word cloud diagrams, dissemination phenomena such as celebrity hypertalk, hierarchical dissemination energy, and extreme similarity were discovered.

Keywords: sudden public events; Weibo public opinion; hot netizens; focus preferences

Classification Number: G210.7

DOI: 10.13266/j.issn.2095-5472.2020.026

Citation Format: Wang Linwei. A Study on the Focus Preferences of Hot Netizens in Public Opinion on Weibo of Unexpected Public Events [J/OL]. Knowledge Management Forum, 2020, 5(5): 271-282 [citation date]. <http://www.kmf.ac.cn/p/222/>.

1 Introduction

According to the Emergency Response Law of the People's Republic of China implemented on November 1, 2007, sudden incidents refer to natural disasters, accidental disasters, public health incidents, and social security incidents that occur unexpectedly, cause or may cause serious social harm, and require emergency response measures [1]. The U.S. crisis management agency ICM further categorizes crises into sudden crises and smoldering crises [2]. Due to their volatile nature and unpredictable consequences, sudden public events have attracted increasing attention from governments and emergency departments. The release of public policies by government departments must be based on the prevention and emergency response perspectives of sudden public events, with corresponding adjustments to emergency measures, reflecting the unified purpose of state institutions and social public organizations. When selecting channels for social opinion monitoring and public policy release, implementation, adjustment, and revision, Sina Weibo has become an inevitable choice for governments due to its unique open platform and public resources. China has currently formed a network of interactive platforms represented by Sina Weibo, WeChat Moments, and Zhihu, among which Sina Weibo has 350 million users, reaching a usage rate

of 42.3% among netizens [3]. In Weibo public opinion on sudden public events, the focus preferences of hot netizens have become an invisible driving force for online public opinion, and public events are resolved through the promotion of hot netizens, making their research value increasingly recognized [5]. Given the vastness of online information, the dissemination effectiveness of online public opinion events can help government departments continuously formulate corresponding measures and revise and improve relevant policies based on actual conditions.

2 Literature Review

2.1 Current Research Status

First and foremost, the supporting technical issues have attracted attention. Network public opinion research technologies include various information collection, feature extraction, data mining, text analysis, automatic classification, automatic clustering, automatic summarization, intelligent retrieval, etc. Among them, clustering analysis methods include conceptual semantic space and similarity measurement, webpage classification based on the combination of support vector machines and unsupervised clustering [6]. Sudden public events have become a research hotspot in academia due to their unique characteristics and serious harm to society. By analyzing research results on network public opinion, some scholars have conducted concentrated studies on the basic theories of network public opinion, specifically including the concept, constituent elements, and evolution mechanisms of network public opinion. Domestic Weibo research has involved more than 70 sudden incidents. In the ranking of main keywords in domestic Weibo research from 2017 to 2019, the rankings of user, word frequency, and netizen were 4, 13, and 48 respectively, with corresponding word frequencies of 573, 230, and 88 [4]. Research on Weibo sudden incidents in the library and information science field focused on crisis public relations and responses in 2010 and 2011. With the frequent occurrence and expanding types of sudden incidents, research hotspots have shifted to the dissemination and diffusion patterns of sudden incident Weibo public opinion. In the ranking of network public opinion-related research topics from 2007 to 2017, sudden incidents ranked first, with 2014 being the hottest year. After 2016, the library and information science field gradually focused on the dissemination of Weibo public opinion as a key research direction, obtaining numerous innovative research results through the organic combination of basic communication laws with mathematical, statistical, and computer science models and algorithms. Research on this topic involves multiple disciplines using different research methods: library and information science and computer science tend to solve problems through technology, modeling, and mining for sudden incidents, while journalism tends to discuss and raise questions. Through research on relevant literature, numerous scholars have studied the viewpoints and positions of the public and netizens. Domestic scholars generally believe that netizens are extremely sensitive to the social identity of event protagonists, that public opinion trials precede judicial

trials, and that public opinion can even affect judicial independence [4]. This study takes netizen attention, which is more valuable for actual public opinion governance, as the entry point to maximize practical application value and provide more novel and effective approaches for network public opinion governance of different types of events.

2.2 Research on Public Opinion Dissemination

In studies focusing on Weibo netizens, netizens are first identified and classified according to their roles in network public opinion dissemination, followed by analysis and research on the dissemination value of various types of netizens in the public opinion dissemination mechanism. As an important category in identity classification, the dissemination value of hot netizens has attracted increasing attention from domestic scholars. Moreover, as many public events fermented on Weibo have been resolved through the promotion of hot netizens, their research value has become increasingly recognized [5].

2.3 Research on Netizen Focus Points and Attention Levels

Netizen focus points exhibit broad and diverse characteristics and play a very important role in the field of news communication. Currently, domestic scholars have conducted limited research on netizen focus points and attention levels as objects of study. By organizing and analyzing existing research, the main aspects include the following three: Using netizen attention level as a quantitative value in the public opinion dissemination process to obtain the influence of various participating factors on public opinion attention through relevant calculations. Aiming to analyze the regional distribution of netizen attention in network public opinion on sudden public health incidents, methods such as regression equations and cluster analysis are used to ultimately obtain the relationship between netizen quantity, urban economic development level, and public opinion development level. Analyzing network construction composition and netizens' internal needs from the perspective of netizen focus points. By analyzing network development reports and hot issues of netizen concern in recent years, netizens' focus points and their internal needs are profoundly analyzed. Treating netizen focus points as a node in network dissemination, with the phenomenon of public attention shift belonging to a part of network public opinion, analyzing its transfer and change patterns in different types of events. Through summarizing current research, numerous scholars agree that netizen focus points and attention levels have become very important influencing factors in network public opinion dissemination. Netizen focus points can not only reflect netizens' real internal needs for online life but also reflect the composition of network structure and the changing patterns of network public opinion. Research in this field includes quantitative and qualitative analysis. Studying the changing patterns of netizen focus points in various types of events is beneficial for governments to correctly manage and guide network public opinion, creating a healthy and solid network environment.

3 Methodology

3.1 Data Collection

The selection principles for research materials mainly include: In terms of event nature, the event development process should be tortuous and fluctuating, with significant social repercussions and substantial influence on stimulating netizens' emotions and viewpoints; In terms of dissemination effect, the public opinion fermentation degree should be high and capable of stimulating a broad audience to participate; In terms of dissemination scope, Weibo should be the main dissemination platform with wide coverage; In terms of public opinion dissemination, the development cycle should be long, containing stage-based fluctuations and changes. Simultaneously, according to the classification of sudden public events, natural disasters, accidental disasters, public health incidents, and social security events should be evenly covered. Based on the above material selection principles, a comprehensive analysis of hot netizens' dissemination and focus preferences in sudden incident public opinion can be conducted. The case data source was obtained using the Sina Weibo Public Opinion Monitoring System, creating event schemes with "event name + location + figures" as keywords. This scheme includes 20 case events. Within the time window from event occurrence to conclusion, data were crawled including relevant hot netizen names, number of posts by hot netizens, ranking of hot netizens by post quantity, word cloud diagrams of each netizen's Weibo, and vocabulary word frequency counts.

3.2 Operational Definitions

Based on the research material content, specific definitions are provided for hot netizens, vocabulary density, and hot netizen Weibo similarity: **Hot netizens:** Based on the Sina Public Opinion Monitoring System, the top 10 Weibo users in terms of number of posts related to the public opinion event throughout the entire event development process. **Vocabulary density:** The sum of the products of all font size weights involved for a certain vocabulary (target vocabulary) in the word cloud diagram and the corresponding number of occurrences of that font size. The sum result represents the mention degree of the target vocabulary in hot netizens' Weibo posts about the public opinion event and can also reflect the focus of hot netizens' attention. **Hot netizen Weibo similarity:** The weighted value of the number and occurrence frequency of identical vocabulary appearing in the word cloud diagrams of relevant Weibo posts published by hot netizens.

3.3 Research Variables and Their Operationalization

In the empirical study of hot netizens' focus preferences, there are four key variables: word cloud word class, word cloud diagram similarity, vocabulary density, and Weibo user type. This section will analyze the calculation of these key variables.

3.3.1 Classification of Hot Netizen Weibo Word Cloud Word Classes

The LDA model (Latent Dirichlet Allocation, a document topic generation model also known as a three-layer Bayesian probability model) is a commonly used model in the field of natural language processing. It can extract text themes from research materials and present them in the form of probability distributions, with clear themes and descriptive words. However, in practical applications, the LDA model cannot accurately calculate changes in positional relationships between vocabulary, which is a common scenario, and the method has defects. The model also limits text length, content, and format, reducing ease of use. In contrast, the Word2vec model converts words into word vectors, achieving the function of representing text similarity with word vectors without limiting text length and format, and with high computational accuracy. This study used the Word2vec model in the process of constructing vector space, converting the word cloud vocabulary of hot netizens' Weibo into word vectors. Based on the vector space calculation results, the semantic similarity of vocabulary was represented. In multi-target event scenarios, a large amount of interference information appears, affecting clustering effectiveness. Therefore, after constructing the vector space, the TF-IDF algorithm was adopted to quantify the importance of word cloud vocabulary, and the correlation between vocabulary and public opinion events was obtained based on the quantification results. After obtaining clustering results, word cloud word classes were distinguished by part of speech, such as unit names, locations, times, etc., and combined with the dissemination elements of public opinion events: public opinion subject, public opinion object, public opinion noumenon, public opinion media, and public opinion space. Through artificial fine-grained division, 11 categories of word cloud word classes for sudden public event public opinion Weibo were obtained. For example, in the Shandong Pingdu veteran gathering incident, "criminals" and "retired soldiers" in hot netizens' Weibo word clouds belong to the public opinion subject part, so these two vocabularies were classified into the category of "event figures/enterprises/units." The classification method for word cloud word classes first uses the Word2vec model to transform the word cloud diagram into word vectors. After generating the hot netizen Weibo word cloud corpus, the feature word weight calculation method TF-IDF is used to extract word vectors from the initial data. Finally, Word2vec is used to compose the extracted word vectors into the final target vector set, which is imported into SPSS 26.0 for clustering processing, thereby obtaining the distribution and frequency quantification values of hot netizen Weibo word cloud vocabulary for each incident.

3.3.2 Word Cloud Diagram Similarity First, analyze the hot Weibo word corpus for each public opinion event, count all vocabulary in the corpus, then number each vocabulary and count the frequency of numbered vocabulary. Second, since the vocabulary in hot netizens' word cloud diagrams consists entirely of Chinese characters with very few numerical characters, no English characters, punctuation marks, or other stop words, and all byte counts are greater than or

equal to 2, the filtering operation needs to be repeated twice. The first filtering corresponds to the Chinese stop word list to filter stop words, and the second filtering extracts vocabulary with occurrence counts greater than or equal to 2 and processes the frequency of these vocabularies. Finally, the similarity of hot netizens' Weibo for each public opinion event is summarized. Vocabulary with occurrence counts greater than or equal to 2 is marked as repeated vocabulary, with occurrence intervals of [2, 10]. For vocabulary appearing with different frequencies, unified quantification processing is conducted to derive the vocabulary similarity formula in hot netizen Weibo word cloud diagrams, as shown in Formula (1):

[Formula (1) is referenced but not displayed in the original text]

Where U represents the repetition rate of a single vocabulary in the word cloud diagram, and n represents the number of vocabulary in the word cloud diagram. W represents the number of times a single vocabulary appears repeatedly; $T=10$ represents the maximum number of times a single vocabulary can appear in the word cloud diagram of hot netizens' Weibo for each event. After counting the repetition rate of individual vocabulary, the similarity of the word cloud diagrams of 10 hot netizens' Weibo can be obtained by summing up the repetition rates of all repeated vocabulary, as shown in Formula (2):

$$U = U_1 + U_2 + U_3 \dots U_n \quad \text{Formula (2)}$$

Based on the obtained Weibo word cloud diagrams, vocabulary in the diagrams is presented through five hierarchical font sizes according to different word frequency levels. Synthesizing the word frequency levels of hot netizens' Weibo word cloud diagrams across 20 event cases, this study calculates the weights represented by the five different font sizes, as shown in Formula (3):

$$TF(t_i, d) \times \log \frac{a}{\text{weight}(t_i, d)} = \quad \text{Formula (3)}$$

Where $TF(t_i, d)$ represents the term frequency of feature t_i in text d , that is, the number of texts containing t_i in the text collection; a represents a constant, usually valued at 0.01; the denominator in the formula is a normalization factor. The occurrence interval of target vocabulary in hot netizens' Weibo is [0,10]. Vocabulary density calculation is shown in Formula (4):

[Formula (4) is referenced but not fully displayed in the original text]

Where w (0.1, 0.2, 0.3, 0.4, 0.5) represents the word frequency weights represented by the five vocabulary font sizes in the word cloud diagram, and $k[0, 10]$ represents the number of times the target vocabulary appears in the word cloud diagram of hot netizens' Weibo for each public opinion event.

3.3.4 Types of Hot Netizens Public opinion in sudden public events features uncertainty, mobility, and fragmentation. Combined with the openness, diversity, and interactivity of the Weibo platform, Weibo users participating in public opinion dissemination maintain the right to speak freely. As the main body participating in the important link of public opinion evolution, different types and identities of Weibo users play different roles in public opinion evolution. Therefore, this study classifies and statistically analyzes Weibo users participating in natural disasters, accidental disasters, public health incidents, and social security events, and summarizes the evolution mechanism characteristics of different types of Weibo users participating in public opinion, serving as an important breakthrough for emergency and management of sudden incidents. First, users involved in Weibo posts related to sudden incidents are categorized into six Weibo user types based on their different positions in the public opinion event chain, including celebrities, media, involved enterprises, government departments, parties involved, and ordinary netizens. Second, statistics are conducted on the six user types of hot Weibo netizens, and then the category data of Weibo users for the four major types of events are aggregated. Finally, comparative analysis of categories is performed.

4 Empirical Analysis

4.1 Case Sources

This study selected 20 social events that occurred in 2018 according to the four major categories of sudden public events. These 20 case events feature fast dissemination speed and wide dissemination scope on the Weibo platform, meeting the selection principles for research materials. Based on the principles of involving multiple objects, dense public opinion, and long event development cycles, selecting these 20 events as research cases can produce more valuable and persuasive research results. The specific case names are shown in Table 1 .

4.2 Word Cloud Word Class Classification

The 11 word cloud word classes are: event type; event figures/enterprises/units; event process; event results; event location; event-related data; event-related nouns; government units/departments/leaders involved in the event; netizen viewpoints; celebrity hypertalk; irrelevant word classes. By aggregating the 20 events according to event type, the distribution characteristics of word cloud part-of-speech features for hot netizen Weibo in the four types of sudden events can be obtained, with results shown in Table 3. In different types of sudden public events, Weibo users convey different focus points through their Weibo posts. Based on part-of-speech classification and word vector calculation of vocabulary in hot netizens' Weibo word cloud diagrams, the top three word classes in each event type were counted and presented in Figure 1 [Figure 1: see original paper]. In the four event types, the top three word classes (event figures/enterprises/units, event process, netizen viewpoints) are relatively concentrated, and the top three word classes across the four event

types can intersect. Overall, in Weibo public opinion on the four types of sudden events, netizens pay considerable attention to the basic components of public opinion events (event figures/enterprises/units + event name + event process). Simultaneously, netizens' word cloud diagrams can also reflect the roles netizens play in participating in public opinion dissemination and the unified operation patterns that have currently formed. As shown in Table 3, in public opinion on public health and social security events, "netizen viewpoints" ranks third, while this word class does not appear in the top three for accidental disaster and natural disaster Weibo public opinion, forming sharp contrasts between paired event types.

From a horizontal comparative analysis perspective, it is found that in Weibo public opinion on social security, public health, and accidental disaster events, netizens' focus points all include the figures/enterprises/units involved in the event, the event process, and relevant noun details in the event. These three event types all radiate from people as the main subject source, with figures, processes, and objects becoming important components that constitute events and important elements of event public opinion. The suddenness of natural disaster events originates from natural forces, and through netizens' focus points, it is evident that they pay more attention to event location and event type. The difference in netizens' focus points between this event type and the other three types reflects the distinction in event nature and simultaneously the difference in the source of event suddenness. Targeting this characteristic can provide government departments with emergency exits for public opinion governance, using the netizen focus points that distinguish this event type from other types as entry points to comprehensively grasp netizens' psychological and emotional states of attention. Based on hot, focal, and headline events that netizens focus on, top reports can be pinned, while simultaneously conducting agenda setting regarding the frequency, number, and layout of related event reports, thereby reversing public opinion trends and weakening the diffusion intensity of negative public opinion.

In sudden public opinion incidents, identifying the source of suddenness is the primary task for judging public opinion development and adopting governance measures. The so-called source of suddenness refers to certain human or natural factors that induce the occurrence of sudden incidents. The internal factors of network public opinion in sudden public events include intentional disseminators, network media, and audiences. First, it must be clarified that identifying the source of suddenness does not mean that all sources of suddenness in sudden incidents are non-human factors. Second, the driving force of the source of suddenness must be evaluated.

- (1) **Weibo Public Opinion on Social Security Events.** These are generally group incidents involving large numbers of people, with characteristics of human intention, premeditation, violence, and political sensitivity. Compared with other types of sudden incidents, netizens show relatively high attention to the development process of such events, which is closely

related to the generally vicious nature of these events. The suddenness of such events has dual aspects: On one hand, event organizers or main participants plan the initiation process and social impact before the event occurs, and the event happens to achieve preset objectives. On the other hand, the social impact produced by such events is vicious, and event organizers or participants often receive legal sanctions, generating negative effects on the public. The motives of event organizers can be divided into two categories: Attempting to generate certain social impacts through the outbreak of the event, expanding the publicity scope of their objectives, and then driving more social public members to invest emotions and reach consensus through group mobilization, ultimately achieving their goals. Based on event presets, not expecting the matter to be publicized and generate social effects, but the harm caused to society has concealment and latency. Once the event breaks out, its sudden harm is extremely severe; if the event enters a latent period until stabilization, its concealed social security harm becomes more serious, increasing the probability of evolving into hidden motives for such event outbreaks.

- (2) **Weibo Public Opinion on Public Health Incidents.** These typically involve health departments, related enterprises, and medical staff. The nature of events is diverse and flexible. In such events, netizens' attention to "event figures/units/departments" forms obvious differences with netizen viewpoints and event processes. From this, it can be inferred that in public health sudden incidents, stakeholders in the event have the strongest restrictive power among various public opinion elements. Simultaneously, public opinion development also revolves around event stakeholders. The so-called "stakeholder" is a concept first introduced in academic research methods by the Stanford Research Institute, after which dozens of definitions for this concept gradually emerged in academia. These definitions all recognize from different angles the individuals and groups that have stake relationships diffusing from an enterprise center. From the essence of public opinion events, the results produced by events involve the interests of multiple behavioral subjects. These behavioral subjects will influence the occurrence of such results through their own behaviors, and thus may produce different results due to the different behaviors and combinations of behavioral subjects. These behavioral subjects are called stakeholders [7].

The word class "netizen viewpoints" demonstrates obvious differences in the vocabulary ranking of hot netizens' word cloud diagrams across the four types of event public opinion. It ranks third in both public health and social security events, but does not appear in the top three for natural disaster and accidental disaster events. Analyzing this phenomenon, the four event types can be paired into groups with commonalities. Public health incidents and social security events are now classified as Group A, while natural disaster events and accidental disaster events are classified as Group B. Group A event public opinion has typical dissemination characteristics, as the event subject is accompanied by

derivative topics after the initial outbreak, forming derivative public opinion, such as topics on infectious disease prevention, medical treatment, riots, drug use, unemployment, vaccines, and casualties. As derivative topics impact the heat of the outbreak event issues, media also begin to join the reporting of emerging topics. Therefore, the output field of netizen viewpoints shifts from one topic to multiple topics, and network media reporting on events shows a continuous trend. Consequently, Group A events have the longest survival time in network media, and the accumulation of netizen viewpoints is naturally the largest. In Group B events, disasters and accidents are not based on human will, and the sources of suddenness basically belong to uncontrollable natural forces. Typically, the period from event outbreak to the announcement of handling results is a dormant period for media.

From a vertical comparative analysis perspective, accidental disasters and natural disaster sudden incidents rank first in terms of suddenness. By comparing and analyzing natural disaster events among the case events, it can be found that event type and location are representative attribute labels for natural disaster sudden incidents. Therefore, netizens' focus points on natural disaster event public opinion concentrate on two aspects: event location and event type. Netizens' attention to these two aspects highlights their need to understand detailed event information. Their attention to event public opinion has two starting points: Paying attention to the latest development of the disaster. Paying attention to the control and treatment of the disaster by the government and relevant departments.

4.3 Word Cloud Diagram Similarity Analysis

Word cloud diagrams are composed of numerous vocabulary words. The first step in calculating word cloud diagram similarity is to calculate the individual repetition rate of all vocabulary that appears repeatedly in the word cloud diagram. Next, weighting operations are performed on the similarity of each hot netizen's Weibo word cloud diagram. Finally, the word cloud diagram similarity is obtained. Hot netizen word cloud similarity can, to a certain extent, reflect whether netizens can generate their own unique insights while following the development dynamics of public opinion. If similarity is high, indicating many identical vocabulary words in the word cloud diagram and similar Weibo content, it can basically be considered that these netizens post Weibo from the same standpoint and perspective, with individual unique insights approximating zero. Within the scope of public opinion attention, it can also demonstrate that netizens' attention intensity to event public opinion varies. Research findings show that hot netizens' word cloud diagram similarity exhibits certain regular patterns, specifically manifested in the following two points: In public opinion events where word cloud similarity phenomena appear, the word cloud similarity of 10 hot netizens shows extreme characteristics. The so-called extremization refers to word cloud diagrams of hot netizens where vocabulary presents either extremely low or extremely high similarity. In other words, hot netizens'

attention intensity to target event public opinion also shows a unidirectional extremization phenomenon. After calculating the mean similarity of the four event types, it can be found that the source of event suddenness directly affects hot netizens' word cloud similarity.

By counting the number of similar vocabulary in word clouds, it was found that the number of similar vocabulary in word cloud diagrams with similar vocabulary basically exceeds half of the total. Therefore, the ratio of completely identical vocabulary number to total vocabulary number in a single hot netizen's word cloud diagram is set at 0.5 as the critical threshold. In target event public opinion, the similarity ratio of hot netizens' word cloud vocabulary is basically distributed in the interval $[0.5, +\infty)$. That is to say, once similarity phenomena appear in hot netizens' word cloud diagrams, the similarity ratio will exceed 0.5, with extremely high frequency of completely identical vocabulary states presenting nearly 70%-80% of the vocabulary in the entire word cloud diagram. Simultaneously, in public opinion events where similarity phenomena appear, more than 5 of the 10 hot netizens hold similar word cloud diagrams, and in some events, nearly 90% of hot netizens show completely similar word cloud diagrams. Through analysis, it can be determined that hot netizen Weibo posts with word cloud vocabulary similarity ratios above 0.5 are forwarding or copying Weibo posts with high forwarding rates in that event's public opinion, without any personal viewpoint output.

According to Table 4, aggregating Weibo word cloud vocabulary similarity ratios by event type yields: In Weibo public opinion on natural disasters and accidental disasters, the mean word cloud similarity of hot netizens is below 1; in Weibo public opinion on public health incidents and social security events, the mean word cloud similarity of hot netizens is above 1. The sources of suddenness for the first two categories belong to non-human factors, and people have no anticipation or preparation for the occurrence of events. In this unknown situation, netizens subconsciously conduct divergent thinking and associations based on event uncertainty, capable of generating different viewpoints from various factors involved in the event. Simultaneously, these two event types have stronger influence than the other two types in terms of spatial geographic location and involved interest groups, capable of lengthening the time front of public opinion development under intersecting influence relationships. The longer time front provides netizens with sufficient objective conditions, promoting the event to maintain high public opinion productivity across five stages: latent period, incubation period, acceleration period, maturity period, and decline period. The mean word cloud similarity of hot netizens in public health incidents and social security events being above 1 is a quantitative representation of lower public opinion generation rates and higher repetition rates in the public opinion of these two event types. Combined with analysis of vocabulary types in hot netizens' word cloud diagrams, the top three vocabulary types in the two event types are completely identical, and netizens' focus points on public opinion of these two event types remain on specific event detail information and social public opinion.

4.4 Analysis of Weibo Hot Netizen Types

Based on an analysis and classification of hot netizens' identities according to their roles in Weibo public opinion events, six major identity types are identified: celebrities, media, involved enterprises, government departments, parties involved, and ordinary netizens, as shown in Table 5. Before hot event public opinion reaches a certain level of discussion heat and dissemination breadth, media reports serve as the opportunity for the beginning of public opinion evolution, transforming event information from news form into Weibo form. The main driving force for event information dissemination in Weibo comes from the attention, forwarding, and commenting on events by ordinary identity Weibo users, thus generating hot netizens. When information, public opinion viewpoints, and commentaries gather and flood into the Weibo platform in large quantities, event public opinion development enters a white-hot stage, with public opinion development speed showing a positive linear relationship with Weibo user attention levels. After this stage, the first energy level of event public opinion reaches saturation, and public opinion dissemination energy will rise from ordinary Weibo users to Weibo influencers, such as public figures and famous media. Thereafter, Weibo users at different energy levels quickly gather and resonate, continuously expanding the scope of event public opinion. Therefore, according to research data, it can be concluded that the energy followed in the formation, evolution, and dissemination of Weibo public opinion is gradually stratified from low to high according to Weibo user influence. Moreover, Weibo user influence has an inverse linear relationship with energy saturation points and a positive linear relationship with energy levels. The lower the energy level, the weaker the promoting effect on public opinion dissemination, and the more energy needs to be stored.

4.5 Regression Analysis of Hot Netizen Word Cloud Phenomena

4.5.1 Dissemination Energy Hierarchy As shown in Table 5, the majority of Weibo users participating in public opinion dissemination and driving public opinion development are ordinary netizens, followed by media Weibo and government affairs Weibo. The dissemination of hot event Weibo public opinion requires the joint participation of netizens of various identities, and their roles in the public opinion dissemination process are not the same. As the dissemination object of Weibo public opinion, hot events receive dispersed energy dissemination from multiple dissemination subjects. According to the dissemination motivations of these subjects, they can be divided into four categories [8]: Free subjects that emerged in the self-media era, having no direct or indirect association with public opinion events; Leader subjects spawned by personal influence, composed of journalists, commentators, writers, Weibo big Vs, etc., expressing both media viewpoints and reflecting public opinions; Control subjects—voices of government regulations, playing a guiding role in public opinion trends; Stakeholder subjects closely related to public opinion events, which are key factors in public opinion development [9].

Hot event public opinion is a collective reflection of various social events and contradictions. Weibo users express specific attitudes or emotions through the Weibo platform. When participating Weibo users reach a certain scale, the event constitutes a hot topic in Weibo. When hot event public opinion diffuses to a certain scope, the main guiding role belongs to Weibo opinion leaders with high attention, high traffic, and greater influence, such as Weibo big Vs, celebrities, marketing accounts, etc. Ordinary Weibo users, due to their small individual energy, only manifest their dissemination effect when a certain number of users aggregate for dissemination.

4.5.2 Celebrity Hypertalk Dissemination Phenomenon Hypertalk, as a popular online term, is the abbreviation for super topic, a function launched by Sina Weibo that refers to circles formed by people with common interests, similar to interest tribes on Tencent QQ, mostly centered on celebrity idols. However, in Weibo's environment, fans can communicate with celebrity idols, and this function aims to strengthen communication between celebrities and fans [10]. Through data statistics, among the 20 case events, 4 cases involved celebrity hypertalk vocabulary in word clouds, with relatively high word cloud similarity in two cases. The celebrity hypertalk phenomenon in hot netizens' word cloud diagrams indicates that sudden incidents can provide sufficient space for celebrity cross-border image shaping and fan agency behavior. Because sudden incidents can focus network users' attention in a short time, celebrity participation can guarantee their exposure and fan traffic. Celebrities can carry out public welfare actions around donating materials, participating in actual rescue and consolation, contributing to the logistical support of sudden incidents, enhancing public favorability, and also providing a free field for their fans to seek pleasure and contribute strength. In case events, the appearance of the celebrity hypertalk phenomenon represents that the focus of hot netizens' attention shifts due to celebrity effects, centering on all related activities of the concerned celebrity. For celebrity fans, they pay attention to every move of the celebrity. Compared with entertainment news, work announcements, and negative news, fans' response to celebrities' positive public welfare actions will be greater.

Natural disaster events belong to the event type with relatively high hazard degree among the four types of sudden incidents, and the celebrity hypertalk phenomenon also appears more concentrated in natural disaster sudden incident public opinion. This indicates that natural disaster sudden incidents have a greater impact on publics of all social strata and easily generate various social effects. Such incidents cause less noise to the communication system and will not fundamentally challenge or change the core institutions, values, and structures of society. At the same time, for public figures like celebrities, it is easy to find breakthroughs for transmitting their positive energy. The unlimited dissemination potential of event public opinion infinitely amplifies the social public welfare energy that originally held a subsidiary position. Netizens' attention to event public opinion expands the boundaries of public opinion with-

out limit, and topics triggered by celebrity participation in the event itself raise the information volume of event public opinion to an abnormal level. Natural disaster events become the main event type for celebrity hypertalk attention, primarily for the following reasons: This event type has unpredictability and hazard, and sudden incidents completely determined by natural factors often cause serious personal injury and property loss. From the perspective of human emotional proximity, such events can most easily trigger social attention and sympathy. This event type does not involve human nature and other complex social relationship issues. Although event public opinion involves a wide range, each stage cycle can switch and quickly fade within a relatively short time. The “simplicity” of such events is the main reason why public figures represented by celebrities choose them for public welfare activities. On one hand, they hope to contribute to public welfare; on the other hand, they hope to establish a positive public image through public welfare. At the same time, unnecessary negative public opinion triggered in a series of participation activities should be avoided as much as possible. The social harm caused by such events basically belongs to “physical damage,” basically not causing psychological and spiritual harm to humans, making it easy to find entry points for social public welfare activities and easy for celebrity persona to find breakthroughs.

5 Conclusion

Research results indicate that in Weibo public opinion on the four types of sudden public events, hot netizens’ focus points exhibit distribution changes according to the characteristics of different event types. As an important output and dissemination link in Weibo public opinion dissemination, hot netizens’ attention level is an important indicator that can reflect public opinion development trends and public opinion governance directions. The directivity generated by different types of sudden events at this link shows mutual differences while also demonstrating commonalities. The differences in focus points and attention levels of hot netizens in public opinion on the four types of sudden events reflect the differences in the diffusion forms and dissemination patterns of the events themselves in the Weibo network. Research results show that hot netizens pay the most attention to basic information about the four event types: natural disasters, accidental disasters, public health incidents, and social security events. This simultaneously proves that the four types of sudden incidents easily generate network public opinion and strongly drive netizen participation. Various phenomena in the dissemination process demonstrate that hot netizens’ attention power lays a solid foundation for public opinion dissemination, and different focus points on events are important bases for distinguishing the sudden nature of the four event types. The directivity of hot netizens’ attention levels indicates that network public opinion on sudden public events presents a trend of internal differentiation and stratification among hot netizens. With the development of the Internet, the netizen group that could be viewed as a whole has gradually evolved into a diversified composite group with different viewpoints and dispersed dissemination paths. Netizen emotions easily arise at

focus point gathering places, which may further stimulate netizens to make extreme behaviors. Therefore, netizen focus points can be used as the entry point for sudden incident public opinion governance, and dispelling negative emotions generated by netizens at focus point gathering places is a powerful method for governing public opinion.

References

- [1] State Council of the People's Republic of China. Emergency Response Law of the People's Republic of China [Z]. Beijing: State Council, 2007.
- [2] Zhang Yiwen. Research on the Mechanism of Interaction Between Sudden Public Crisis Events and Network Public Opinion [D]. Beijing: Beijing University of Posts and Telecommunications, 2012.
- [3] China Internet Network Information Center. The 43rd Statistical Report on China's Internet Development [EB/OL]. [2019-04-20]. http://www.cnnic.net.cn/hlwfzyj/hlwxzbg/hlwtjbg/201902/t20190228_70645.htm.
- [4] Zhou Jinyuan, Zhang Shasha, Liu Guifeng, et al. A Review of Domestic Weibo Research [J]. Journal of Intelligence, 2013, 32(9): 46-51.
- [5] Wang Yan. Public Opinion Expression and Public Participation: Research on Weibo Opinion Leaders [D]. Beijing: Chinese Academy of Social Sciences, 2014.
- [6] Li Xiaoli. Research on Data Mining in Web Information Retrieval and Classification [D]. Beijing: Institute of Computing Technology, Chinese Academy of Sciences, 2001.
- [7] Fang Jie. Research on Classification and Behavioral Motivation of Stakeholders in Weibo Public Opinion [D]. Nanjing: Nanjing University, 2014.
- [8] Yang Yansheng. Analysis of Micro-Circles and Micro-Circle Public Opinion [J]. Journal of Intelligence, 2017(4):
- [9] Duan Fengfeng, Zhu Juan. Spatial Reconstruction and Interactive Relationship of Weibo Public Opinion Dissemination Elements [J]. News Dissemination, 2018(23): 4-6.
- [10] Hypertalk. Xiumu Network [EB/OL]. [2020-04-20]. <http://www.xiumu.cn/question/6480.html>.
- [11] Li Dan. Development and Trend of China's Weibo Public Opinion from the Perspective of Civil Society [J]. Southeast Communication, 2011, 81(5): 6-8.
- [12] Rao Yuan. Public Opinion Calculation Methods and Technology [M]. Beijing: Electronics Industry Publishing Group, 2016.
- [13] Wang Guohua. Dynamic Elements and Governance of Network Public Opinion in Sudden Incidents [M]. Wuhan: Huazhong University of Science and Technology Press, 2017.
- [14] Wen Qing. Research on Dissemination Elements and Network Structure of Weibo Natural Disaster Public Opinion [D]. Changchun: Jilin University, 2018.

A Study on the Focus Preferences of Hot Netizens in Public Opinion on Weibo of Unexpected Public Events

Wang Linwei

School of Journalism and Communication, Guangxi University, Nanning 530004

Abstract: [Purpose/significance] The purpose of this paper is to find a new breakthrough in the prevention and control of public opinion by quantifying the vocabulary frequency of hot netizens' microblog and summarizing their concerns in public emergencies. [Method/process] In this study, by crawling the microblog data of hot Netizens, Word2vec vector model and TF-IDF weight algorithm were used to obtain the indexes such as the proportion of part of speech and the similarity ratio of words in the cloud map of hot Netizens' microblog words, and finally the indexes of the four kinds of events were summarized and analyzed. [Results/conclusion] In the public opinions of emergencies, the hot netizens pay more attention to the basic components of events, including the event person, event name and event process. Moreover, in the public opinions of the four types of emergencies, the hot netizens show different and dispersed attention directivity, but there are still some commonalities. At the same time, through the analysis of the hot internet users' word cloud map, we found that the star hypertalk, the dissemination of energy hierarchy and extreme similarity and other dissemination phenomena.

Keywords: sudden public events; public opinion on Weibo; hot internet users; focus on preferences

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.