

## RESEARCH ARTICLE

# Research on Network Public Opinion in War Damage Incident of Major Water Conservancy Projects

XINYU WANG<sup>1,2</sup>, LIANGHAI JIN<sup>1,2,3</sup>, AND SHUQING WANG<sup>1,2</sup><sup>1</sup>Hubei Key Laboratory of Hydropower Engineering Construction and Management, China Three Gorges University, Yichang 443002, China<sup>2</sup>Department of Hydraulic and Environmental Engineering, China Three Gorges University, Yichang 443002, China<sup>3</sup>Safety and Environment Science and Technology Ltd., Yichang 443002, China

Corresponding author: Lianghai Jin (jinlianghai@ctgu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 52179136, and in part by the Humanities and Social Science Foundation Project of Ministry of Education of China under Grant 21YJA630038.

**ABSTRACT** Major water conservancy projects face the risk of damage from conflict, and the potential consequences of dam failure can lead to widespread public outrage, posing a threat to social stability. This study analyzed the formation model of network public opinion risk, using the Kakhovka Hydropower station damaged during conflict as a case study. Based on data analysis of netizen' comments on Twitter and Weibo (a Chinese social media platform similar to Twitter), we used data mining algorithm to identify the primary topics of interest and examine the evolution of public opinion. The findings indicate that attention to incident resulting in war damage has a negative impact on social stability through a feedback loop. Netizen on Twitter and Weibo were primarily focused on "Flood" "Crisis" "Charge" and "Help" topics. Among them, Twitter was more popular in terms of public opinion and had faster dissemination speed. "Crisis" and "Charge" were more likely to generate public opinion risks. The potential impact of war damage on social stability is noteworthy, and social stability is largely dependent on the attention level of netizen.


**INDEX TERMS** Major water conservancy projects, network public opinion, social stability risk, war damage incident, Twitter and Weibo.

## I. INTRODUCTION

Major water conservancy projects face a high risk of being targeted or utilized as weapons in war [1]. The Ruhr Dam's destruction in Germany during World War II caused significant damage to the local industries. Likewise, the catastrophic collapse of the Huayuankou Dam in China had a disastrous impact on the surrounding region. In recent years, there has been a rise in global instability, leading to frequent incidents of war damage, particularly those involving major water conservancy projects. On June 6, 2023, the Kakhovka Hydropower Station in Ukraine was destroyed, affecting over 16,000 people. Such incident brought an unforeseen risk to residents, who become stakeholders and accelerate the

escalation and dissemination of the situation. The dissemination of potential risks through the interest groups of social networks leads to increased public concern regarding the negative outcomes linked to such incidents. This situation affects the online public sentiment and may even provoke disturbances to social stability [2].

Scholars have extensively studied the potential impact of water conservancy projects on social stability. To evaluate the risk factors associated with these projects, complex systems are often utilized to create evaluation index models at a macro level. For instance, the physical-mature-human theory is applied to land requisition and resettlement in water conservancy projects that determine the index value of risk factors [3]. Furthermore, an evaluation of indicators for hydropower projects was conducted, that simplifies and enhances the indicators' objectivity [4]. However, macro

The associate editor coordinating the review of this manuscript and approving it for publication was Qilian Liang .

models lack a comprehensive analysis of the interactions among social network participants, thus preventing the revelation of complex networks based on social connections, and the exposure of the impact of stakeholder interaction on conflict escalation.

Researchers have directed their focus to micro level studies, uncover macro features of interactional relationships by social networks. Specifically, they analyze pertinent risks and mutual connections among stakeholders, as well as identifying crucial social stability risks [5]. Moreover, the Net-Logo simulation platform was applied to examine the propagation of risks through social networks based on interpersonal connections [6]. Micro level model simulations reveal that stakeholder interactions utilizing social networks accelerate the development and dissemination of potential hazards.

To align simulation results with engineering practice, researchers have implemented various enhancements to social network models. Such as, the behavior of the public in five different risk states was simulated using the Agent-based model [2], and the causality of society stability risk variables was explored through the fault tree analysis [7]. Additionally, the complexity of stakeholder attention and uncertainty arising from stakeholder diversity were addressed using the Bayesian network [8]. However, these updated models limit the stakeholders to network nodes based on interpersonal relationships within the project and region, without considering any subjective evaluations. Consequently, the role of network public opinion was failed to acknowledge in social network models, resulting in disparities between engineering practices and simulation outcomes.

It has been found that network public opinion can amplify conflicts between stakeholders and exacerbate risks to social stability. For example, conflicts over hydropower development in a certain region have sparked an anti-small hydropower movement at the societal level [9]. Also, researchers have found that studying network public opinion can better identify stakeholders and analyze conflicts of interest. Druckman [10] examined the polarization of public opinion during the COVID-19 crisis to understand how American citizens hold different positions on policies and parties. Lyu [11] collected Twitter users' views on hashtags such as #StopAsianHate to identify supportive and opposing groups in the movement. Meanwhile, Lamsal et al. and Khan et al. [12], [13] have used social media platforms such as Twitter and Weibo to mine target groups and analyze public attitudes. Su, Chen [14], [15] have compared the behaviors of Twitter and Weibo user groups to analyze cognitive differences. Obviously, network public opinion plays a vital role in identifying stakeholders, clarifying interaction relationships, and addressing uncertainties.

The aforementioned study outlines the potential for network public opinion in relation to social stability risks. Furthermore, there is a lack of a comprehensive research that assesses the effects of the social stability risks posed by incidents of war damage. However, current research in this area is limited and primarily focuses on evaluating conventional

risks inherent in the project, neglecting sudden risks associated with network public opinion. In particular, Twitter (English) and Weibo (Chinese) netizen' views on the incident, as well as netizen' public opinion patterns and developments in topics of interest.

Therefore, this paper aims to examine war damage incident of major water conservancy projects to better understand their impact on social stability. Section I examines how network public opinion risks are formed. Section II details the process of data collection and research methods employed during the study. Finally, we analyzed the topics of concern and evolution patterns among netizen. The study concludes by presenting some final thoughts and conclusions, and the findings are significant for guiding online public opinion and reducing the risk of social instability.

## II. METHODS

### A. RISK INFLUENCE MECHANISM

During times of war, major water conservancy projects faced a substantial risk of war damage, resulting in catastrophic flooding. Floods represent the most severe kind of natural disaster, with the potential from dam failure or collapse causing the most devastating flooding. Albu [16] examined the dangers of floods caused by breaches in dams, while Wang [1] assessed dams' susceptibility to underwater explosions. For instance, the Kakhovka Hydropower Station is located in a conflict zone shared by Russia and Ukraine. The collapse of the dam, resulting from war damage, inflicted catastrophic downstream damage, which ignited a public outrage on social media.

Objectively, war damage poses a significant risk to public. According to social psychology, the psychological perception of the public regarding natural disasters and war damage varies [17]. The most severe psychological trauma comes from war damage and other "man-made disasters" which brings the public closer to the disaster. This trauma can have impacts on various societal tiers through social media, which can incur risks to the public opinion and lead to unpredictable incident [18].

In 1958, Forrester [19] introduced System Dynamics (SD), which proposes that a system's character and behavior depend substantially on its internal dynamic structure and feedback mechanism. SD is an appropriate approach for investigating the dynamic principles of the propagation of public opinion based on social network ties [20]. Therefore, this paper analyzes the rules that govern the generation of public opinion risk through implementation of the SD model. It views netizen, online media, and society as subsystems and explores the impact of public opinion on social stability, as illustrated in Figure 1.

(a) The proliferation of network public opinion and the increasing focus on netizen imply a favorable feedback cycle within the netizen subsystem. The incident triggered debates on Twitter and Weibo, resulting in a rise in netizen posts and more significant involvement, ultimately intensifying network public opinion.

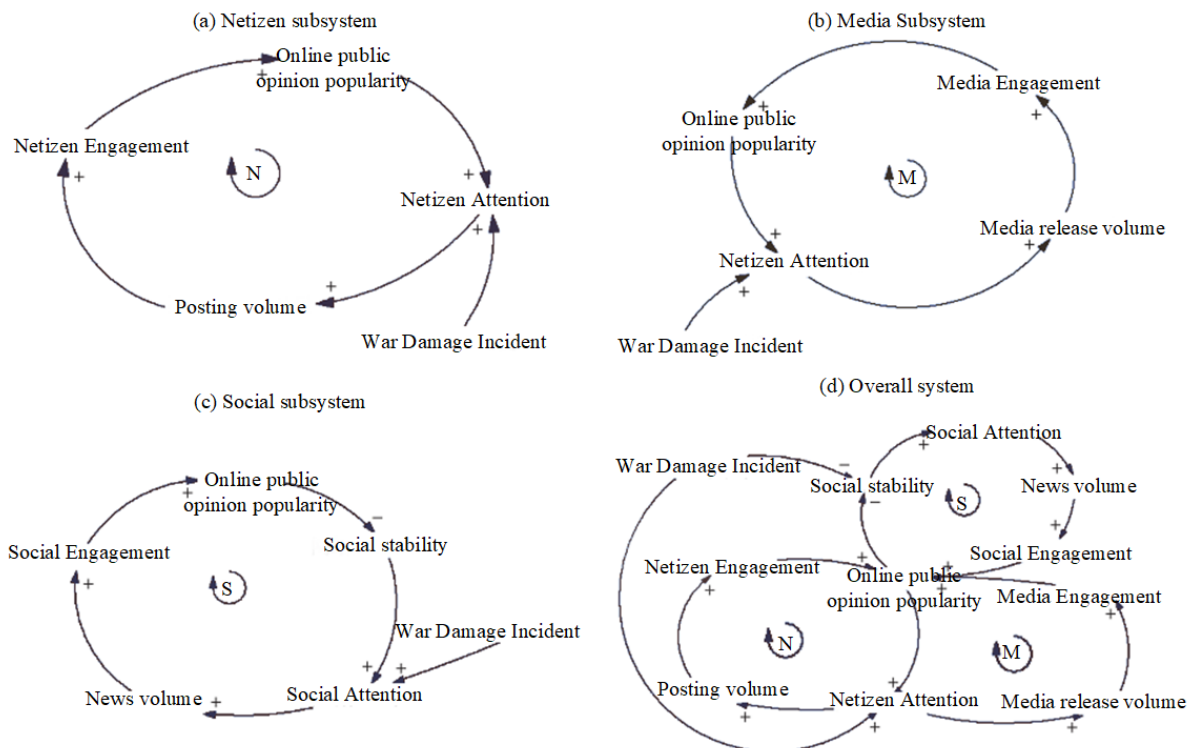


FIGURE 1. Public opinion risk formation system based on netizen, media and social.

(b) The subsystem of online media operates as a positive feedback loop. Online news coverage promotes netizen’s views and behaviors, while netizen’s increased attention to an incident, resulting in more news being published by online media. Consequently, online media participation intensifies, leading to further increased intensity of network public opinion.

(c) The social subsystem forms a negative feedback loop. Increased social attention results in a higher frequency of news releases, demonstrating greater social participation and an increase in network public opinion intensity. However, this surge in public opinion can harm social stability [18]. Consequently, the resulting rise in social instability further elevates social attention.

(d) Overall system consists of three subsystems, two of which are positive feedback systems, while one is a negative feedback system. The attention of netizen serves as the initial factor of the entire system, and the resultant network public opinion carries the potential risks to social stability.

To sum up, the main objective of this study is to analyze the topics that are drawing the attention of netizen, and to examine the evolution in network public opinion.

**B. DATA COLLECTION**

Twitter and Weibo are two notable social media platforms with comparable influence, functionality, and audience target. While sharing similarities, they remain autonomous platforms that offer valuable insights into network public opinion, and provide a comprehensive analysis.

The data were collected from Twitter (English-speaking netizen) and Weibo (Chinese-speaking netizen) during 30 days statistical period from June 7 to July 6, 2023, accounting for time differences. The search term “Kakhovka” was used, and the Python intelligent crawler was employed to capture hourly review text data. Eventually, a total of 16,679 Twitter comments and 4,052 Weibo comments were collected, resulting in output text set D.

To ensure accurate analysis of netizen’s comments, it is crucial to eliminate redundant and duplicative data. This can be achieved through the following steps: eliminating duplicates, filtering, and cleaning. Firstly, repeated comments will be filtered to retain only the text of the first comment to ensure objectivity. Secondly, String filtering is used to remove comments irrelevant to this incident and remove invalid symbols and punctuation and meaningless text. Finally, Repeat the previous steps and modify, add, or delete based on the actual results. For instance, strings such as “coupons” and “welfare” will be removed as they are not related to the topic.

**C. NOISE REDUCTION**

The commenting behavior of netizen on Twitter and Weibo is characterized by high frequency, low volume, and high noise. To mitigate the effects of this noise, the document frequency ( $DF_n$ ) can be reduced using the least common term (LCM) threshold, and the low-frequency feature set can be processed more effectively using the information gain (IG) method. Therefore, this paper integrates the  $DF_n$ -IG function

to perform noise reduction for the text set  $D$ , as below:

$$DF_n - IG = \sum_{T \subseteq D} \frac{DF_n(t, T)}{D} \log \left( \frac{D \times DF_n(t, T)}{T \times DF_n(t, D)} \right) + \sum_{T \subseteq D} \frac{DF_n(\bar{t}, T)}{D} \log \left( \frac{D \times DF_n(\bar{t}, T)}{T \times DF_n(\bar{t}, D)} \right) \quad (1)$$

where  $t$  refers to a single word in the text, while  $D$  represents the text set and  $T$  is a single text within the set.  $DF_n(t, D)$  represents the number of texts in which  $t$  occurs at least  $n$  times in the set. Similarly,  $DF_n(\bar{t}, D)$  represents the number of texts in which  $t$  occurs less than  $n$  times in the set.  $DF_n(t, T)$  is assigned a value of 2 if  $t$  occurs at least  $n$  times in a single text  $T$ , otherwise its value is 1. Similarly,  $DF_n(\bar{t}, T)$  is assigned a value of 2 if  $t$  occurs less than  $n$  times in a single text  $T$ , otherwise its value is 1.

The calculation of  $DF_n-IG$  follows equation (1), and Figure 2 displays the noise reduction outcomes for various thresholds  $n$ . As the frequency threshold increases, the number of terms eliminated in each pass gradually decreases.

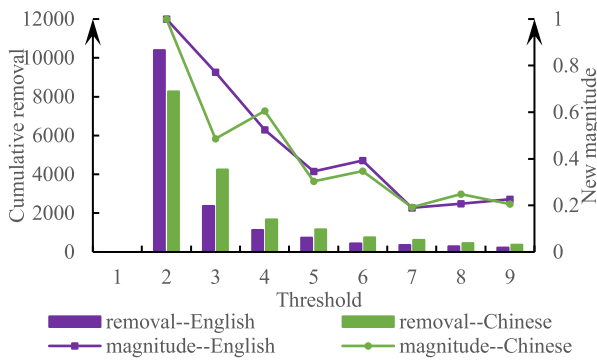


FIGURE 2. Noise reduction results at different thresholds.

The optimal noise reduction effect is achieved when  $n=7$  for both Chinese and English keywords. Therefore, the minimum term frequency threshold is set at the lowest point of a new increase that words with frequencies below 7 are removed from the text set  $D$ . Ultimately, Twitter has 3127 words remaining, while Weibo has 2135.

#### D. KEYWORD EXTRACTION

Public opinion keywords should precisely reflect the event’s characteristics and scope. The term frequency–inverse document frequency (TF-IDF) algorithm identifies significant keywords from high-frequency words in denoised text. As a statistical analysis method for words, it differs from traditional word frequency statistics. TF-IDF algorithm states that a word’s importance is directly proportional to its frequency in the text and inversely proportional to its frequency in the corpus. It effectively filters common words, identifies infrequent feature words, and highlights important words that reveal event characteristics and define event scope [21].

Identifying keywords is fundamental to topic analysis. After conducting several topic cluster analyses, it has been

found that keywords with a weight greater than 0.01 can lead to better convergence and more comprehensive coverage of the core views of public opinion. The equation is as follows:

$$TF_{i,j} = \frac{N_{i,j}}{\sum_k N_{k,j}} \quad (2)$$

$$IDF_{i,j} = \lg \frac{|D|}{|\{D_i \in j:t_i\}| + 1} \quad (3)$$

$$TF_{i,j}IDF_{i,j} = TF_{i,j} \times IDF_{i,j} \quad (4)$$

where  $TF_{i,j}$  denotes the frequency of keyword  $t_i$  in text set  $D$ , while  $N_{i,j}$  signifies the number of occurrences of the  $t_i$  in text  $j$ , and  $\sum_k N_{k,j}$  stands for the total number of text  $j$  containing the  $t_i$  in text set  $D$ . While  $IDF_{i,j}$  represents the probability of occurrence of the  $t_i$  in text set  $D$ ,  $|D|$  denotes the total number of texts in  $D$ , and  $|\{D_i \in j:t_i\}|$  indicates the number of texts containing  $t_i$ , we add 1 that prevent data from not existing.  $TF_{i,j}IDF_{i,j}$  indicates importance of the  $t_i$ .

#### E. PERPLEXITY ANALYSIS

The optimal number of clusters for the extracted keywords is determined using the perplexity function. This function evaluates the probability model distribution, and a lower degree of perplexity improves the clustering effect.

If the perplexity curve reaches its lowest point and the number of topics is relatively small, the algorithm obtains the value of  $k$ . The keywords are then clustered into  $k$  topics using the K-Means algorithm. The calculation equation is as follows:

$$Perplexity = \exp \left\{ - \frac{\sum_{n=1}^N \log (P(w_n))}{\sum_{n=1}^N D_n} \right\} \quad (5)$$

where  $n$  represents the  $n$ th text.  $P(w_n)$  is the occurrence probability of each keyword in the  $n$ th text.  $D_n$  is the total number of words in the  $n$ th text.

The perplexity values of the  $D$  text set are calculated by (5). As the number of topics increases, the perplexity curve initially decreases and subsequently increases, as shown in Figure 3.

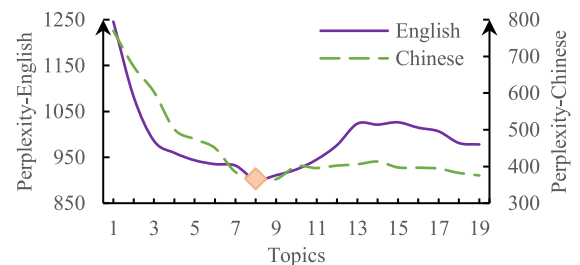


FIGURE 3. Perplexity matrices at different topics.

It is noteworthy that once 8 topics are reached, both the Chinese and English perplexity curves associated hit their lowest point, indicating optimal clustering at this point. Therefore, we set the number of clustering topics  $k=8$ .



TABLE 1. The topics between Twitter (English) and Weibo (Chinese).

Netizen	Topics	Contents
Twitter (English)	E <sub>1</sub>	The Kakhovka was destroyed, Russia and Ukraine accused each other.
	E <sub>2</sub>	Explosion threatened the safety of a nuclear power plant, triggering a nuclear energy crisis.
	E <sub>3</sub>	Thousands of residents were forced to evacuate and relocate due to flooding.
	E <sub>4</sub>	Flooding has halted the Ukrainian army's counterattack.
	E <sub>5</sub>	The lowered water level has disrupted the water supply to the inhabitants.
	E <sub>6</sub>	The Kremlin is lying, the Russian army launched an attack and killed civilians*.
	E <sub>7</sub>	Marine animals in the Black Sea are experiencing significant mortality rates.
	E <sub>8</sub>	In favor of providing assistance to victims, President Zelensky accepts a NATO aid.
Weibo (Chinese)	C <sub>1</sub>	The Kakhovka collapsed, and both Russia and Ukraine have denied any responsibility.
	C <sub>2</sub>	The area was submerged due to flooding. It is uncertain what caused the accident.
	C <sub>3</sub>	The terrorist attack has caused a humanitarian disaster to economy and ecology.
	C <sub>4</sub>	This is a White House conspiracy, similar to destroy the Nord Stream pipeline*.
	C <sub>5</sub>	The nuclear crisis was triggered when the water supply to the power plant was impacted.
	C <sub>6</sub>	The ongoing conflict's impact on the global food crisis persists.
	C <sub>7</sub>	Chinese express concern the conflict that urged a peaceful resolution and provide assistance.
	C <sub>8</sub>	Residents were evacuated as a result of the flood, and the media reported the death toll.

The truth of the incident is unknown, and this paper only elaborates netizen' comments.

produced a co-occurrence network of keywords, as depicted in Figure 4.

Additionally, the keywords are represented as “bubbles” and their co-occurrence relationship as “edges”. The thickness of the “edges” indicates the level of co-occurrence frequency, while the size of the bubbles depicts the frequency of keyword occurrences, with different colors indicating distinct topics. Also, the solid line represents the interaction within the topics, and the dotted line indicates the connection between them.

(a) The English public opinion topics E<sub>1</sub>, E<sub>2</sub>, E<sub>3</sub>, and E<sub>5</sub> have a strong correlation. This indicates that the Kakhovka Hydropower Station War Damage Event (E<sub>1</sub>), resulted in Nuclear Crisis (E<sub>2</sub>), Floods (E<sub>3</sub>), and Water Shortage (E<sub>5</sub>). According to Twitter netizen, there exists a correlation between floods (E<sub>3</sub>) and the obstruction of Ukraine’s counter-offensive (E<sub>4</sub>). They also support providing assistance to victims (E<sub>8</sub>). Additionally, it is indicated that the Kremlin is lying (E<sub>6</sub>), and that a significant number of Black Sea creatures have perished (E<sub>7</sub>).

(b) Chinese public opinion topics Floods (C<sub>2</sub>), Nuclear Crisis (C<sub>5</sub>) and Civilian Deaths (C<sub>8</sub>), are closely linked to the Kakhovka Hydropower Station War Damage Event (C<sub>1</sub>). Additionally, topic C<sub>3</sub> is connected to C<sub>4</sub>, indicating that Weibo netizen suspect US involvement in a similar conspiracy to sabotage the Nord Stream pipeline (C<sub>4</sub>), which is also related to humanitarian disaster (C<sub>3</sub>). Finally, it is indicated that a worldwide food crisis arising from the ongoing conflict (C<sub>6</sub>), and depicts China’s demand for peace (C<sub>7</sub>).

To sum up, the networks of topics reveal semantic connections between keywords and showcase viewpoint context within the plot, enabling exploration of the context and characteristics of netizen perspectives. For instance, the keywords “Kherson” “region” “flooding” “people” and “evacuate” suggest that the flood in Kherson region compelled citizens to evacuate. Moreover, the phrases “China” “concern” “conflict” “urged” and “peace” suggest China’s concern about the conflict and urge for peace, as summarized in Table 1.

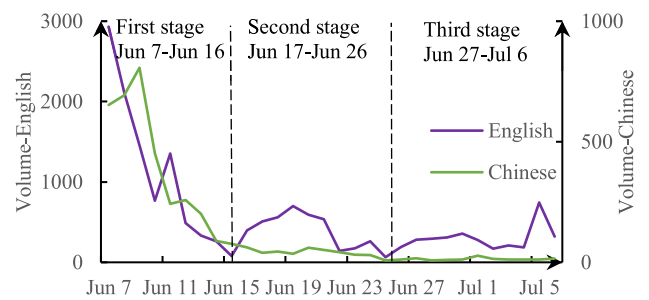


FIGURE 5. The change curves of public opinion popularity.

The truth of the incident is unknown, and this paper only elaborates netizen’ comments.

From the perspective of topic distribution, as shown in Table 1 and Figure 4, Twitter and Weibo netizen’ perceptions regarding this incident are focused on the following topics:

“**Flood**” represents the **direct consequences** and its impacts caused by the incident, including Floods (E<sub>3</sub> and C<sub>2</sub>), as well as their related consequences, like Ukraine’s counter-offensive obstruction (E<sub>4</sub>), Water Shortages (E<sub>5</sub>), and Civilian Deaths (C<sub>8</sub>). The discussion appears to be centered around these topics due to their similarity and relevance.

“**Crisis**” represents the **indirect consequences** caused by the incident, including Nuclear Crisis (E<sub>2</sub> and C<sub>5</sub>), Ecological Crisis (E<sub>7</sub>), and Food Crisis (C<sub>6</sub>). Additionally, Twitter netizen tend to focus more on the Ecological Crisis (E<sub>7</sub>), while Weibo netizen tend to prioritize the Food Crisis (C<sub>6</sub>).

“**Charge**” represents the **allegations** made by netizen against the incident, such as the Kremlin’s lies (E<sub>6</sub>), humanitarian disaster (C<sub>3</sub>), and US conspiracy (C<sub>4</sub>). Additionally, there exists a noticeable disparity in the perspectives of netizen. Twitter netizen tend to support the notion of Russia’s sabotage (E<sub>6</sub>), while Weibo netizen tend to favor the notion of the US’s conspiracy (C<sub>4</sub>).

“**Help**” represents the **appeals** made by netizen in response to the incident, which may involve providing aid (E<sub>8</sub>) or promoting peace (C<sub>7</sub>).

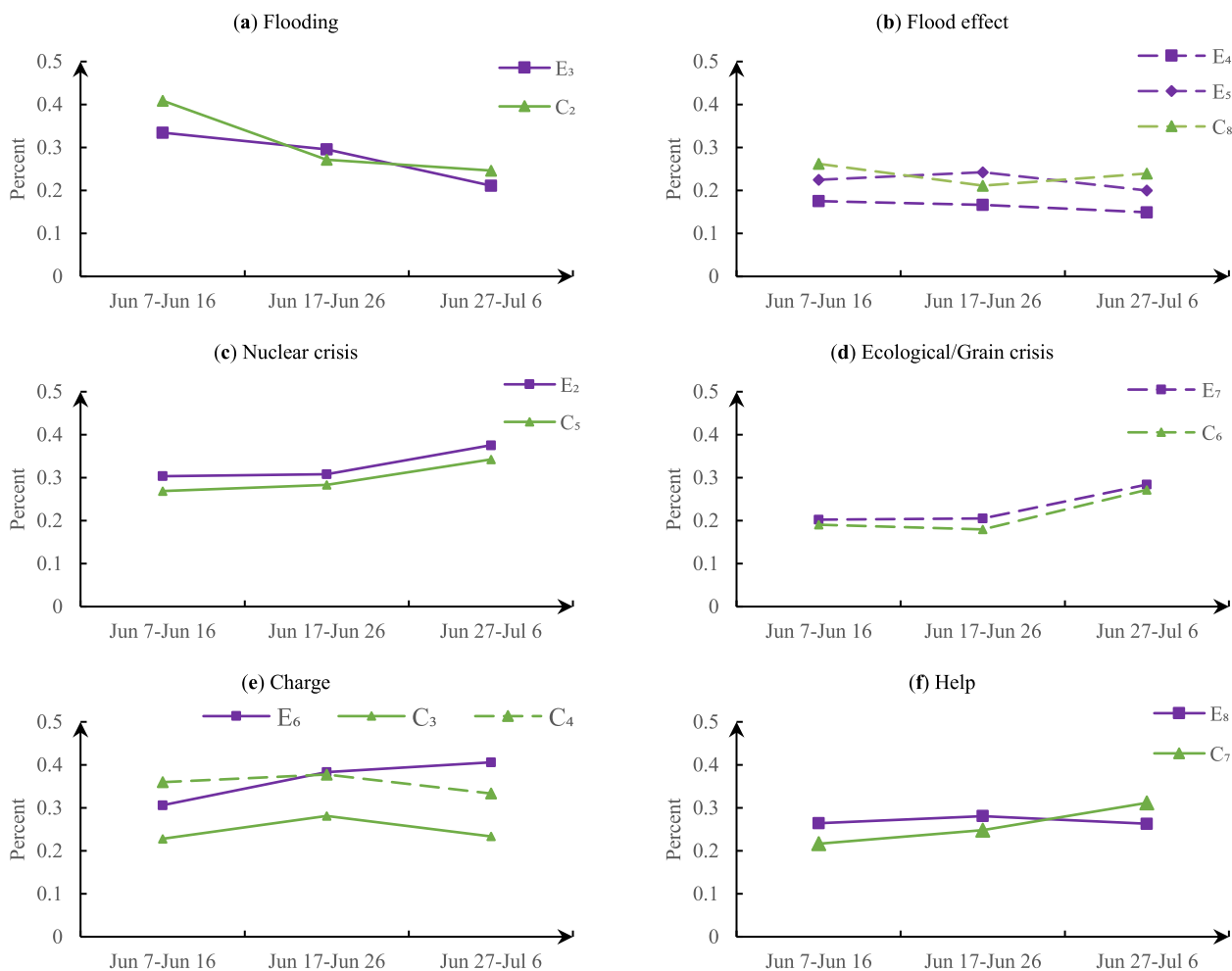


FIGURE 6. The trends of netizen’ attention.

Furthermore, Twitter and Weibo netizen are in agreement that “Flood” and “Help”, however, there is a cognitive differentiation between “Crisis” and “Charge”.

**B. EVOLUTION TREND OF NETWORK PUBLIC OPINION**

1) MACRO TREND

To analyze evolutionary trends, text set D was divided into daily time windows of equal duration. Then, create graphs to track the level of public engagement on Twitter (English) and Weibo (Chinese) based on the daily count of netizen posts over the 30-day period that followed the war damage incident. The results are displayed in Figure 5.

Firstly, Twitter shows faster transmission speed and higher public opinion intensity compared to Weibo. Also, Weibo’s transmission speed lags behind Twitter’s by approximately 1-2 days, as seen in the time it takes for the public opinion curve to reach its peak.

Additionally, Chinese public opinion intensity gradually decreases after the peak. In contrast, English public opinion displays fluctuations with various peaks and valleys during

the periods of June 7-16, June 17-26, and June 27-July 6 that can be categorized into three stages.

2) MICRO TREND

This study presents an improved analysis of netizen’ attention trends by com-paring the percentage of heat for each topic, rather than simply aggregating comment data. Additionally, the data can be independently analyzed within a specific time frame for a static model, or linked to data from other time frames for a dynamic model.

Take the abscissa as the date and the ordinate as the popularity percentage, which represents the proportion of comment text data for a certain topic within a certain time period in the corresponding text set. Based on the topics that netizen pay attention to, such as “Flood” “Crisis” “Charge” and “Help” we construct the change curves of netizen’ attention in three stages (June 7-16, June 17-26, and June 27-July 6). The results are shown in Figure 6.

From the trends of “Flood” as shown in (a) and (b), Twitter and Weibo netizen show considerable interest in Flooding. However, their attention diminishes as the incident continues

to unfold. At the same time, the extent of change in the Flood effect remains stable.

From the trends of “**Crisis**” as shown in (c) and (d), the attention of Twitter and Weibo netizen toward “Crisis” suggests a rising trend, with minor fluctuations in the first and second stages and a rapid surge in the third stage. Additionally, the attention to nuclear crisis is comparatively significant.

From the trends of “**Charge**” as shown in (e), Twitter netizen demonstrates a steady rise in interest in “Charge”. In contrast, Weibo netizen initially display an upward trend of interest in “Charge” followed by a decline. Additionally, topic  $C_4$  generates more attention than  $C_3$ , Weibo netizen associate the US’s conspiracy at  $C_4$  with a humanitarian crisis at  $C_3$ , which aligns with the curve’s findings.

From the trends of “**Help**”, as shown in (f), Weibo netizen is increasingly interested in “Help” over time, whereas Twitter netizen initially exhibit an interest that then decreases.

Analyzing the evolution rules of topics that netizen pay attention to can help trace the evolution of network public opinion at the micro level, and it is beneficial to discover important risk factors based on the evolution curve. To sum up the above analysis, there are the following rules for the attention of Twitter and Weibo netizen:

1) War damage incidents spread faster, generate significant public interest, and are more likely to elicit public opinion reactions on Twitter. Conversely, the public opinion heat on Weibo tends to be short-lived.

2) During the first stage, there was a notable upswing in Twitter and Weibo netizen exhibited high interest in the “Flood” culminating in the initial height of public opinion. In the second and third stages, escalating Twitter netizen’ concentration on “Crisis” and “Charge” became instrumental in creating a subsequent public opinion peak.

3) As per the previous analysis, there is a consensus among netizen regarding “Flood” and “Help” topics, while “Charge” and “Crisis” are controversial. By examining the trends of attention curves, it is evident that “Flood” is the primary consensus and “Help” is the secondary consensus. However, “Charge” is the main point of contradiction, and “Crisis” is the secondary point of contradiction.

#### IV. DISCUSSION

This paper presents new research identifying that major water conservancy projects not only face war damage risk, but also lead to public opinion risk, which exposure public to psychological trauma and brought social instability risks. To analyze this process, we employ SD theory to analyze public opinion risk by three subsystems: netizen, media and social. Positive feedback loops are found to shape network public opinion by the attention of netizen, ultimately impacting social stability through negative feedback loops.

To elaborate further, the  $DF_n$ -IG function is utilized to decrease the noise of comments, and the TF-IDF algorithm is employed to extract high-weight keywords. While the perplexity function is utilized to determine the optimal cluster

number  $K$  of keywords, and the  $K$ -means algorithm is used for clustering. Then, the Jaccard algorithm is utilized to analyze the co-occurrence relationship of keywords, which resulted in the English topics  $E_1 \sim E_8$  and Chinese topics  $C_1 \sim C_8$ , and centered on “Flood” “Crisis” “Charge” and “Help” for both netizens. There are variances exist in opinions between both sides with regards to “Crisis” and “Charge”

Additionally, we examined the macro trend in network public opinion evolution, and the micro trend in the attention of 4 types topics. The results indicate that Twitter exhibits quicker transmission rates, higher public interest, and a greater likelihood of causing public opinion peaks than Weibo, whose public opinion peaks take place a day or two after Twitter’s. Initially, topics “Flood” and “Charge” received significant attention, resulting in the first peak of public opinion. Subsequently, the Twitter netizen’ interest in “Crisis” and “Charge” increased, leading to a subsequent surge in public opinion.

This study presents a novel method for cluster analyzing topics. While conventional methods tend to classify keywords into various clusters, they neglect to examine the correlation among keywords in clusters [25]. Therefore, we included an extra co-occurrence analysis phase, which involves constructing a network of topic-keywords based on their co-occurrence relationships. Furthermore, we divided the evolution of network public opinion into 3 stages, and analyzed the varying proportions of the four types topics of netizen’ concerns within these stages, facilitating the tracking of netizen’ attention trends. This method offers an intuitively understanding of netizen’ public opinion.

As the first study to examine public opinion on incidents of war damage in different cultural backgrounds related to major water conservancy projects, which resulting public opinion pressure can encourage the government to adopt more effective water conservancy project protection policies and measures. Also, some limitations are in this study. Firstly, due to data protection policies, there may be insufficient comment text data from Weibo netizen, which may distort the analysis results from the actual. Conversely, more information is available on open social platforms like Twitter, providing better simulation of reality. Secondly, we examined the attention topics of netizen based on the co-occurrence relationship, which depends on similarity. However, the cognitive differences between Twitter and Weibo netizen are not discussed in depth. Moreover, there may be more efficient ways or different methods, such as deep learning models proposed by Alsaedi for monitoring Twitter data [26], after which we will work on these aspects.

#### V. CONCLUSION

Netizen and media systems create a positive feedback loop, attracting more attention from groups and triggering public opinion. This, in turn, forms a negative feedback loop through the social subsystem, which affects social security and stability.



For the same extreme risk, Twitter and Weibo netizen are in agreement that “Flood” and “Help”, however, there is a cognitive differentiation between “Crisis” and “Charge”. Initially, topics “Flood” and “Charge” received significant attention, resulting in the first peak of public opinion. Subsequently, the Twitter netizen’s interest in “Crisis” and “Charge” increased, leading to a subsequent surge in public opinion.

Twitter exhibits quicker transmission rates, higher public interest, and a greater likelihood of causing public opinion peaks than Weibo, whose public opinion peaks take place a day or two after Twitter’s.

## REFERENCES

- [1] X.-H. Wang, S.-R. Zhang, J.-G. Dai, and C. Wang, “Evaluation of base damage and stability of concrete gravity dam subjected to underwater explosion,” *Structures*, vol. 38, pp. 1502–1514, Apr. 2022.
- [2] W. Li, J. Yuan, C. Ji, S. Wei, and Q. Li, “Agent-based simulation model for investigating the evolution of social risk in infrastructure projects in China: A social network perspective,” *Sustain. Cities Soc.*, vol. 73, Oct. 2021, Art. no. 103112.
- [3] M. Luo, “Establishment of social stability risks evaluation model based on GAHP and IVHFSs,” *J. Intell. Fuzzy Syst.*, vol. 41, no. 1, pp. 917–928, Aug. 2021.
- [4] H. An, W. Yang, J. Huang, A. Huang, Z. Wan, and M. An, “Identify and assess hydropower project’s multidimensional social impacts with rough set and projection pursuit model,” *Complexity*, vol. 2020, pp. 1–16, Nov. 2020.
- [5] W. Wang, X. Guo, Q. Cao, and A. Tang, “A stakeholder perspective on social stability risk of public–private partnerships project for water environmental governance in China: A social network analysis,” *Frontiers Ecol. Evol.*, vol. 10, p. 15, Jan. 2023.
- [6] Z. He, D. Huang, and J. Fang, “Social stability risk diffusion of large complex engineering projects based on an improved SIR model: A simulation research on complex networks,” *Complexity*, vol. 2021, pp. 1–17, Dec. 2021.
- [7] T. Yuan, T. Huo, H. Huo, X. Fang, L. Li, M. Chen, and L. Yu, “Mapping risk strategy of social stability risk considering causal relationships for energy infrastructure projects,” *KSCSE J. Civil Eng.*, vol. 27, no. 11, pp. 4557–4571, Nov. 2023.
- [8] S. Wen and G. Qiang, “Managing stakeholder concerns in green building projects with a view towards achieving social sustainability: A Bayesian-network model,” *Frontiers Environ. Sci.*, vol. 10, p. 18, Mar. 2022.
- [9] N. Magnani, “Beyond NIMBY: Mobilization against mini-hydroelectric power in the Italian Alps,” *Mountain Res. Develop.*, vol. 40, no. 3, pp. R31–R36, Nov. 2020.
- [10] J. N. Druckman, S. Klar, Y. Krupnikov, M. Levendusky, and J. B. Ryan, “Affective polarization, local contexts and public opinion in America,” *Nature Hum. Behav.*, vol. 5, no. 1, pp. 28–38, Nov. 2020.
- [11] H. Lyu, Y. Fan, Z. Xiong, M. Komisarchik, and J. Luo, “Understanding public opinion toward the #StopAsianHate movement and the relation with racially motivated hate crimes in the US,” *IEEE Trans. Computat. Social Syst.*, vol. 10, no. 1, pp. 335–346, Feb. 2023.
- [12] R. Lamsal, A. Harwood, and M. R. Read, “Twitter conversations predict the daily confirmed COVID-19 cases,” *Appl. Soft Comput.*, vol. 129, Nov. 2022, Art. no. 109603.
- [13] A. Khan, H. Zhang, J. Shang, N. Boudjellal, A. Ahmad, A. Ali, and L. Dai, “Predicting politician’s supporters’ network on Twitter using social network analysis and semantic analysis,” *Sci. Program.*, vol. 2020, pp. 1–17, Sep. 2020.
- [14] Y. Su, J. Xue, X. Liu, P. Wu, J. Chen, C. Chen, T. Liu, W. Gong, and T. Zhu, “Examining the impact of COVID-19 lockdown in Wuhan and lombardy: A psycholinguistic analysis on Weibo and Twitter,” *Int. J. Environ. Res. Public Health*, vol. 17, no. 12, p. 4552, Jun. 2020.
- [15] W. Chen, K. K. Lai, and Y. Cai, “Exploring public mood toward commodity markets: A comparative study of user behavior on Sina Weibo and Twitter,” *Internet Res.*, vol. 31, no. 3, pp. 1102–1119, May 2021.
- [16] L. M. Albu, A. Enea, M. Iosub, and I. G. Breaban, “Dam breach size comparison for flood simulations. A HEC-RAS based, GIS approach for Dracăni Lake, Sitna River, Romania,” *Water*, vol. 12, no. 4, p. 27, Apr. 2020.
- [17] K. Cosic, S. Popovic, M. Sarlija, and I. Kesedzic, “Impact of human disasters and COVID-19 pandemic on mental health: Potential of digital psychiatry,” *Psychiatria Danubina*, vol. 32, no. 1, pp. 25–31, Apr. 2020.
- [18] W. Ahmed, J. Vidal-Alaball, J. Downing, and F. López Seguí, “COVID-19 and the 5G conspiracy theory: Social network analysis of Twitter data,” *J. Med. Internet Res.*, vol. 22, no. 5, May 2020, Art. no. e19458.
- [19] J. W. Forrester, “Industrial dynamics: A major breakthrough for decision makers,” *Harvard Bus. Rev.*, vol. 36, no. 4, pp. 37–66, 1958.
- [20] T. Chen, X. Yin, J. Yang, G. Cong, and G. Li, “Modeling multi-dimensional public opinion process based on complex network dynamics model in the context of derived topics,” *Axioms*, vol. 10, no. 4, p. 270, Oct. 2021.
- [21] S. A. Curiskis, B. Drake, T. R. Osborn, and P. J. Kennedy, “An evaluation of document clustering and topic modelling in two online social networks: Twitter and Reddit,” *Inf. Process. Manage.*, vol. 57, no. 2, Mar. 2020, Art. no. 102034.
- [22] K. P. Sinaga and M.-S. Yang, “Unsupervised K-means clustering algorithm,” *IEEE Access*, vol. 8, pp. 80716–80727, 2020.
- [23] D. Huang, C.-D. Wang, H. Peng, J. Lai, and C.-K. Kwoh, “Enhanced ensemble clustering via fast propagation of cluster-wise similarities,” *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 51, no. 1, pp. 508–520, Jan. 2021.
- [24] Z. Zhou, J. Irizarry, and W. Guo, “A network-based approach to modeling safety accidents and causations within the context of subway construction project management,” *Saf. Sci.*, vol. 139, Jul. 2021, Art. no. 105261.
- [25] J. Qiang, Z. Qian, Y. Li, Y. Yuan, and X. Wu, “Short text topic modeling techniques, applications, and performance: A survey,” *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 3, pp. 1427–1445, Mar. 2022.
- [26] A. Alsaeedi and M. Al-Sarem, “Detecting rumors on social media based on a CNN deep learning technique,” *Arabian J. Sci. Eng.*, vol. 45, no. 12, pp. 10813–10844, Dec. 2020.



**XINYU WANG** received the B.E. degree from the School of Water Conservancy and the Environment, China Three Gorges University, in 2022. He is currently pursuing the master’s degree. His research interests include data mining and safety management.



**LIANGHAI JIN** has been a Professor with China Three Gorges University, since 2009. His research interests include safety ergonomics and safety management.



**SHUQING WANG** received the B.E. degree from the School of Water Conservancy and the Environment, China Three Gorges University, in 2022. She is currently pursuing the master’s degree. Her research interests include data mining and machine learning.

...