2022 15th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI) | 978-1-6654-8887-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/CISP-BMEI56279.2022.9980251

Public Opinion Evolution and Communication Stages in Complex Network

Shan Liu, Xiaoqing Wu
Faculty of Information Science and Technology
Communication University of China
Beijing, China
Email: liushan@cuc.edu.cn

Abstract—**Nowadays, online public opinion is developing rapidly and has a great influence. If it is not guided in time, it will easily endanger the stability of society. Therefore, this paper focused on building a network public opinion dissemination model, and analyzed the evolution law of public opinion dissemination at different stages, so as to help the government do a good job of public opinion prevention and intervention to build a good network environment. This paper took the hot event of "the Fukushima nuclear wastewater" as an example, drawn the evolution curve of public opinion by fitting the collected data, and divided the development of public opinion into three stages, and researched their network structure and public opinion diffusion law respectively. Research shows that the process of public opinion dissemination can be divided into three stages: incubation period, diffusion period, and recession period. In different stages, the network structure has different characteristics.**

Keywords-public opinion, communication stages, complex network

I. Introduction

The "2020 Global Digital Report" shows that the number of people using the Internet in the world has grown to 4.54 billion, and social media users have exceeded 3.8 billion [1]. With the rapid development of communication technology and the popularization of mobile terminals, online social platforms have become the most active positions for the generation, fermentation and dissemination of network public opinion, which have an important impact on economic development and social stability [2].

By far, Twitter is one of the largest social platforms in the world. As of March 2018, Twitter had 336 million active users who posted about 340 million tweets a day, the CEO of Twitter Dick Costello said. What's more, Twitter also handles about 1.6 billion web search requests every day. Therefore, the data in this paper are all derived from Twitter.

The social network is a common application scenario for complex network, which is usually considered to have the characteristics of small-world and scale-free network models. Using the theoretical knowledge of the complex network, the influence of social network structure on the information dissemination process can be revealed, which provides a scientific and effective method for the study of network public opinion dissemination [3-5].

In recent years, emergencies have occurred from time to time, and they have entered the public eye through the Internet, become the focus of heated discussions among netizens and media, and finally form the Internet public opinion. The Internet public opinion develops rapidly and periodicity, sometimes it may be emotional and harmful. Hence, it is necessary to strengthen prevention and control to make public opinion develop in a correct and benign direction [6-7]. If bad public opinion (extreme, provocation, rumors, etc.) is not dealt with timely and guided properly, it is easy to magnify social conflicts, and even be used by some people with ulterior motives, thereby disrupting government decision-making and social stability. Therefore, analyzing the evolution law of public opinion dissemination in different stages and constructing the dissemination model of network public opinion is the core issue of doing a good job in the prevention and control of public opinion.

II. Related Work

The research on public opinion is based on the research on the model of information dissemination. Hagerstrand was the first to study the traditional information dissemination mode of geographic entity space. He pointed out that the accumulation and change of information adoption over time fits an S-shaped logistic curve. He also pointed out that the information dissemination in traditional geographic entity space has multiple effects, which will be affected by dissemination and interaction [8]. Rogers elaborated on the "S-curve" theory in his book "The Diffusion of Innovation", and combined with this theory, he pointed out that the innovation and diffusion of new things and new ideas need to be realized with the help of a certain social network [9]. Fiona Duggan proposed a "crisis information dissemination model" combined with the "S-curve" theory [10]. Xiangdong Liu et al. [11] established three public opinion prediction models based on the traditional curve fitting method for data fitting, and on this basis, by adding the factors that occurred when the objective environment suddenly changed in the network space, they further explored the network public opinion Initial and subsequent information dissemination models. Li Jinze et al. [12] constructed a model of the influence mechanism of microblog public opinion dissemination of emergencies through the λ-Logistic model and SEM (structural equation model), combined with the actual data of "red, yellow and blue" events, and analyzed the impact of factors and indicators on microblog public opinion at each stage. The direction and extent of the spread of influence.

978-1-6654-8887-7/22/$31.00 ©2022 IEEE

The above studies have proved that the process of public opinion dissemination is in line with the "S-curve" trend. However, these papers currently only consider the timing characteristics, focus on the fitting of the curve, and completely ignore the network structure behind the propagation. Therefore, this paper will establish a model based on the traditional curve fitting method for data fitting, and preliminarily determine each stage of the communication process.

III. MODEL

A. The Network Model Construction

In order to study the law of event development through complex network theory, it is first necessary to construct a public opinion diffusion network. Here's how to build the network:

Step 1: Pick out the most influential tweets from the collected data (determined here as tweets with more than 5 citations), and find the users who cited their tweets.

Step 2: Collect the tweets about the event published by the user who cited the tweet after that, go back to the first step and execute it again until the required network size is reached.

Step 3: Define the user as a node, define the reference relationship as an edge, and construct a network by pointing the user who sent the original tweet to the user who referred to the tweet.

B. "S-curve" Fitting

The process of public opinion dissemination is usually divided into three stages: latent, diffusion, and decline. We describe the process of public opinion dissemination by the change of public opinion heat (the number of texts discussing the event) over time. This process is considered to approximate an "S"-shaped curve. Let the cumulative growth value be x and the daily growth amount be Δx.

The "S-curve" is a logistic curve here, and its formula is as follows:

$$\begin{cases} \frac{d_x}{d_t} = rx\left(1 - \frac{x}{K}\right) \\ x(0) = x_0 \end{cases} \quad (1)$$

$$x(t) = \frac{K}{1+\left(\frac{K}{x_0}-1\right)e^{-rt}} \quad (2)$$

$$\Delta x(t) = x(t+1) - x(t) \quad (3)$$

where r, x_0 and K are parameters that need to be determined by fitting.

IV. SIMULATION

A. The Public Opinion Diffusion Network

Tweets about the Fukushima nuclear wastewater incident from March 1st to April 30th were collected from www.twitter.com, and based on the network construction method mentioned in section III, this paper constructed a public opinion diffusion network. Due to time constraints, the network is only built to one level, that is, only run the steps once. Fig. 1 shows the graph.

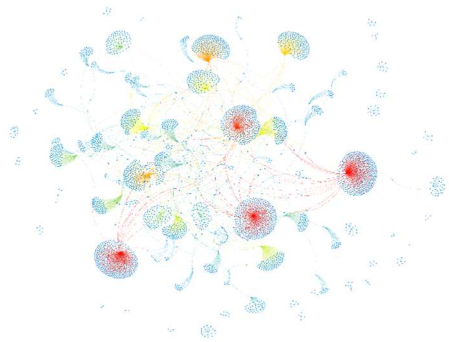


Figure 1. the Public Opinion Diffusion Network for the " Fukushima nuclear wastewater" incident

The color of nodes is related to the size of the node degree, from blue to red, the node with a smaller degree is blue, and the node with a larger degree is red. As Fig. 1 shows, it can be seen that there are many small-scale clusters in the network, and the centers of these clusters are the nodes with a larger degree, which represent the users who interact more frequently in the entire network, that is, the key users. The node with the largest out-degree is the user "@ReutersScience", who has 374 users citing his tweets; the node with the largest in-degree is "@jingKlaus", who cites the tweets of 18 other users in the network.

The degree distribution of the network basically conforms to the characteristics of the power-law distribution, that is, its line approximates a straight line, as shown in Fig. 2. We can roughly think of this network as a scale-free network. The phenomenon of high truncation in the second half of the curve is because the data collected in reality will be affected by social platforms or other factors, resulting in the data being less than ideal.

The in-degree and out-degree distributions of this network are shown in Fig. 3. It can be found that in this network, the average out-degree of nodes is much higher than the average in-degree, indicating that this is a network where users spread information they have known to others so as to make more people know it.

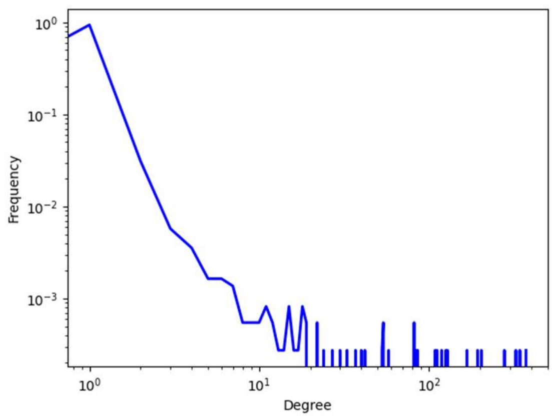


Figure 2. Degree Distribution of the Public Opinion Diffusion Network

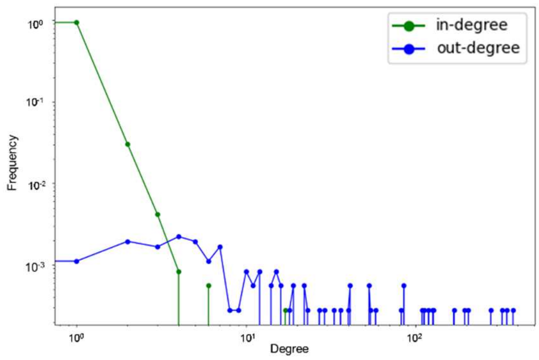


Figure 3. In-degree and Out-degree Distribution of Public Opinion Diffusion Network

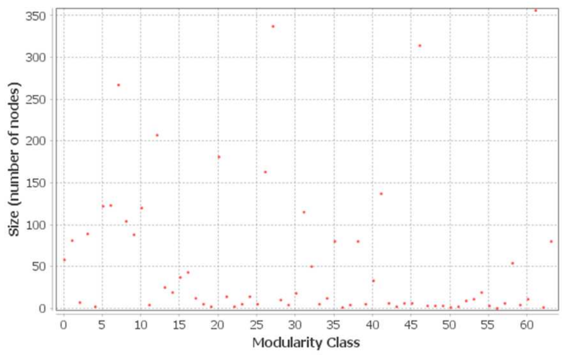


Figure 4. Community Distribution of Public Opinion Diffusion Network

Fig. 4 shows the community distribution of the network., the maximum modularity is 0.91, and the number of communities is 64. Most of the community nodes are below 50, and the largest community node number is close to 350. Dividing each community with different colors, Fig. 1 changes to Fig. 5.

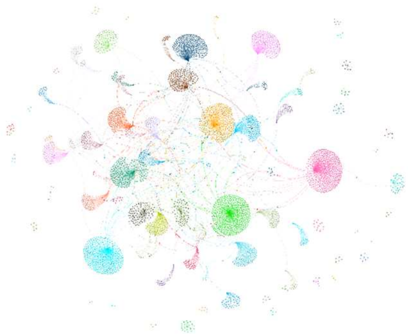


Figure 5. Communities in the network

B. "S-curve" fitting results

After preliminary analysis of the data results, the data from April 1, 2021 to April 30, 2021 (seen in Table I) was selected for fitting.

TABLE I. DATA FRAME

t	date	Δx	x
0	2021-4-1	0	85
1	2021-4-2	0	85
2	2021-4-3	0	85
3	2021-4-4	0	85
4	2021-4-5	0	85
5	2021-4-6	4	89
6	2021-4-7	1	90
7	2021-4-8	0	90
8	2021-4-9	57	147
9	2021-4-10	45	192
10	2021-4-11	98	290
11	2021-4-12	179	469
12	2021-4-13	2051	2520
13	2021-4-14	444	2964
14	2021-4-15	139	3103
15	2021-4-16	93	3196
16	2021-4-17	120	3316
17	2021-4-18	136	3452
18	2021-4-19	32	3484
19	2021-4-20	88	3572
20	2021-4-21	168	3740
21	2021-4-22	28	3768
22	2021-4-23	48	3816
23	2021-4-24	55	3871
24	2021-4-25	2	3873
25	2021-4-26	7	3880
26	2021-4-27	25	3905
27	2021-4-28	12	3917
28	2021-4-29	32	3949
29	2021-4-30	16	3965

Fig. 6 shows the actual total cumulative growth curve and in Fig.7, there is the daily growth curve. Finally, the result obtained by fitting is shown in Fig. 8. The curve fits well at [0,13], with a large deviation from t=14. The obtained parameter result is shown in Table II.

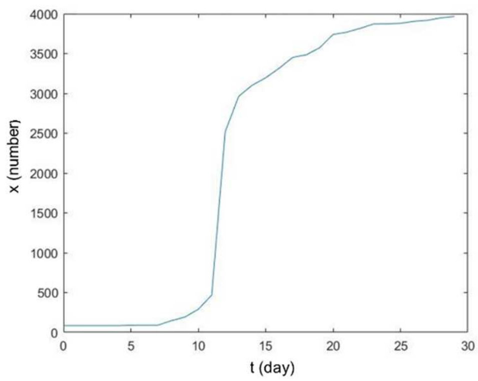


Figure 6. the Actual Total Cumulative Growth Curve

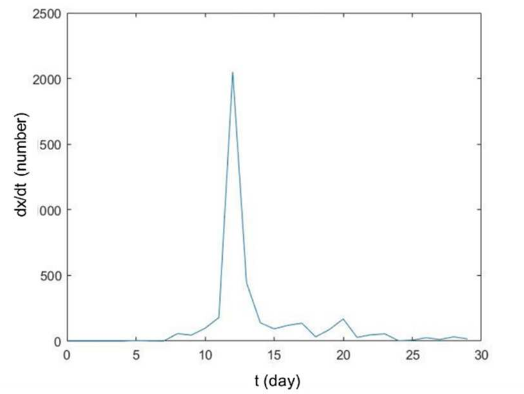


Figure 7. the Actual Daily Cumulative Growth Curve

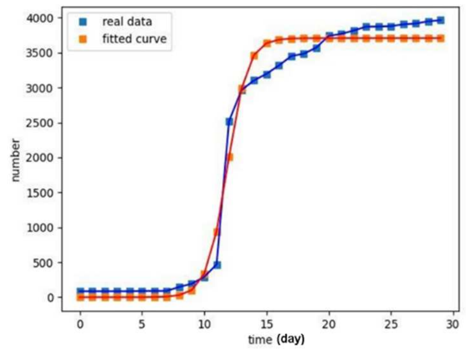


Figure 8. the Fitted Total Cumulative Growth Curve

TABLE II. RESULTS OF FITTING PARAMETERS

Parameter	Value
K	3.700772880e+03
x_0	1.39250682e-03
r	1.24717004e+00

Calculate and plot the derivative of this fitted function (seen in Fig. 9), which should approximate the daily growth curve.

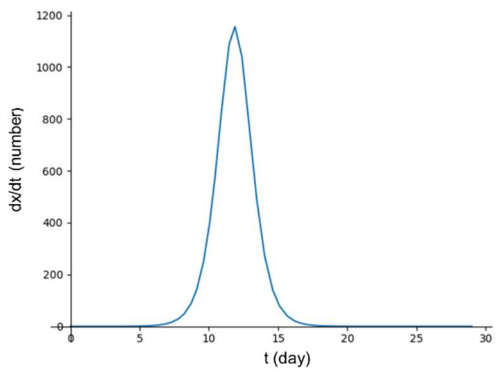


Figure 9. the Derivative function Curve of the Fitted function

It can be seen that the maximum value of the function is around 1200 and the time is located at [12, 13]. According to the data, the Japanese government announced the discharge of nuclear sewage on April 13, which triggered intense discussions that day, so this result is in line with the actual situation. However, compared with the actual daily growth curve, as shown in Fig. 7, the maximum value is still slightly lower than the maximum value in the real scene, indicating that the actual outbreak of public opinion is faster than the fitting.

From the following equation:

$$\frac{d^3x}{d^3t} = 0 \tag{4}$$

we can get 2 key points t_1, t_2, shown in Fig. 10. The S-curve is divided into three stages, namely the incubation period, the diffusion period, and the recession period.

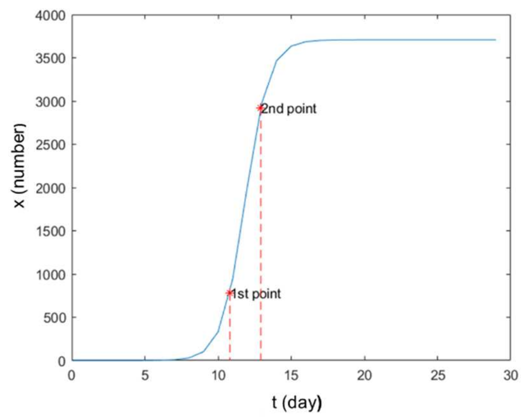


Figure 10. the S-curve Divided into 3 Stages

TABLE III. DIVISION OF PROPAGATION PHASES

t	date	dissemination stage	period
t_1	$\frac{1}{r}\ln\frac{\left(\frac{K}{x_0}\right)-1}{2+\sqrt{3}}$	$(0,t_1]$	the incubation period
		$(t_1, t_2]$	the diffusion period
t_2	$\frac{1}{r}\ln\frac{\left(\frac{K}{x_0}\right)-1}{2-\sqrt{3}}$	$(t_2, +\infty)$	the recession period

Assuming the time series t=[0,1,2,⋯,29], t_1 and t_2 can be calculated according to Table III. Their values are 10.8 and 12.91. Hence, this time period from April 1 to April 12 is called by the incubation period, from April 13 to April 14 is the diffusion period, and the rest is the recession period.

C. Characteristics of network structure at different stages

Taking the tweet time as a timestamp, the public opinion diffusion network is actually a dynamic network. Through Gephi, set the time window to 1 day and the scale to 1 hour. The time window is used to delineate the time period in which we count how many nodes and edges there are currently, and the scale is the length of the time window movement.

Figure 11. Degree Time Series

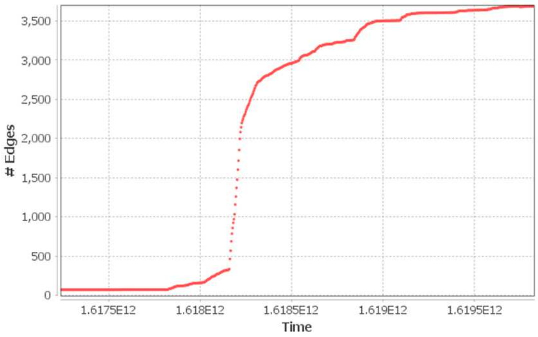


Figure 12. Edges Time Series

In Fig. 11, it can be seen that the average degree first increases slowly over time (latency period), suddenly increases rapidly (diffusion period), and remains stable in the later period (recession period).

Fig. 12 shows the change of the edge over time. Since the establishment of an edge represents the occurrence of propagation, this curve is consistent with the cumulative growth curve (see Fig. 6). In this way, the intrinsic connection between the network structure and the "S" curve is found.

Next, according to the divided three stages, the structural characteristics of the network in each stage are analyzed in detail.

1) The Incubation period

Figure 13. the Network in the Incubation period

In the early stage of the incident, after people got the news of the incident through various media, some groups of netizens will spread the incident information to each other, and pass their views and opinions to the outside world through various platforms.

It can be seen in Fig. 13 that during the incubation period, the event is mainly propagated in a small number of clusters, but edges pointing to nodes in other clusters can also be seen, indicating that there is a trend of information dissemination during the incubation period.

At this time, there are only 542 edges in the graph, accounting for 14.45% of the final number of edges. The average degree is 0.119 and the average path length is 1.002.

2) The Diffusion period

Figure 14. the Network in the Diffusion period

The diffusion period is the stage with the greatest influence and the most serious social harm during the process of network public opinion dissemination. After the formation of online public opinion, its popularity began to grow exponentially. At first, most of the public netizens did not know the detail. Out of concern for the incident, they reprinted the known information and expressed their personal opinions, forming a circle of onlookers, gathering a large number of remarks, and prompting the outbreak of public opinion.

In Fig.14, it can be clearly seen that during the diffusion period, the number of clusters added to the discussion has increased significantly, and the size of the clusters has also increased significantly. During this period, the amount of discussion about the event exploded.

At this time, the network has 2769 edges (the edges will not disappear once they are established), accounting for 73.77% of the total number of edges. The average degree is 0.758 and the average path length is 1.02.

3) The Recession period

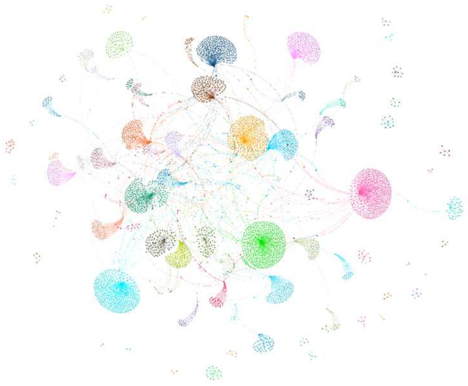
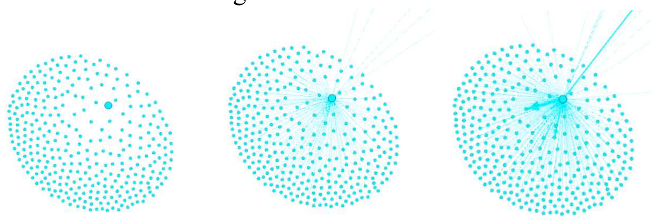


Figure 15. the Network in the Recession period

As time goes by, the public opinion of the event gradually saturates. Fig.15 shows that the structure of the network changes little during the recession period. The edge relationship within the cluster is basically unchanged, and the edge between the clusters has increased. This shows that in the recession period, most users have known about the event, and the diffusion occurring at this time is mostly between users belonging to different clusters, or repeated by users within a cluster.

The average path length is 1.028 higher than before.

Fix a cluster to observe the dynamic changes of the network structure. as Fig.16 shows.



(a) the Incubation period (b) the Diffusion period (c) the Recession period

Figure 16. the Dynamic Changes of Network Structure in a Cluster

It can be clearly seen that during the incubation period, the users of this cluster did not yet join the discussion (only in a few clusters, users began discussions). Since entering the diffusion period, more and more people are participating in the discussion, and the cluster began to spread event-related information. After some time, the Recession period came, and users repeatedly spread the tweets of users who have posted information in the cluster, or spread it to users in other clusters, breaking through the limitations within the cluster.

V. CONCLUSION

According to the "S-curve", the whole process of public opinion evolution can be divided into three stages: incubation period, diffusion period, and recession period. At different stages, the network structure has different characteristics. During the incubation period, the event is mainly spread among a small number of user groups, but sometimes individual users will spill information to users in other groups, indicating that there is a tendency to disseminate information externally. The number of clusters joining the discussion increases significantly, as does the size of the clusters during the diffusion period. During this period, the amount of discussion about the event will explode. In the recession period, there are few changes from before, the edge relationship within the cluster is basically unchanged, and the edge between the clusters increases.

ACKNOWLEDGMENT

The work of this paper was supported by the National Natural Science Fund of China (Project 61901421), the National Social Science Fund of China (Project 19ZDA272), the Fundamental Research Funds for the Central Universities, State Key Laboratory of Media Convergence and Communication, Communication University of China.

REFERENCES

[1] Kemp S. Digital 2020: Global digital overvie. Retrieved from Datareportal: https://datareportalcom/reports/digital-2020-global-digital-overview, 2020.

[2] Wang J. Research on Internet Public Opinion Propagation Considering Symmetric Game and Asymmetric Game Simultaneously. Journal of Modern Information,2022,42(03):109-120+156.

[3] Grabowski A, Kosiński R A. Ising-based model of opinion formation in a complex network of interpersonal interaction. Physica A: Statistical Mechanics and its Applications, 2006, 361(2): 651-664.

[4] Jin X, Wang Y. Research on social network structure and public opinions dissemination of micro-blog based on complex network analysis. Journal of Networks, 2013, 8(7): 1543.

[5] Yang A, Huang X, Cai X, et al. ILSR rumor spreading model with degree in complex network. Physica A: Statistical Mechanics and Its Applications, 2019, 531: 121807.

[6] Wang J, Wu P, Shen S. Modeling Research on the Emotional Influence of Opinion Leaders on Netizens in Emergencies.Journal of Intelligence,2018,37(09):120-126.

[7] Hou X. Discussion on the evolution law of network public opinion and early warning model in big data environment. Journal of Sichuan Police College,2019,31(05):104-110.DOI:10.16022/j.cnki.cn51-1716/d.2019.05.013.

[8] Hagerstrand T. Innovation diffusion as a spatial process. 1967.

[9] Rogers E M, Singhal A, Quinlan M M. Diffusion of innovations//An integrated approach to communication theory and research. Routledge, 2014: 432-448.

[10] Zhu W. New characteristic of crisis information communication model and its management strategies in digital age. Journal of Modern Information, 2009, 29(2): 60-63.

[11] Liu X, Cao A, Li C. Novel Network Public Opinion Prediction and Guidance Model Based on "S-Curve": Taking the Loss of Contact with "Malaysia Airlines". Mathematical Problems in Engineering, 2021, 2021.

[12] Li J, Xia Y, Zhang P, Lan Y. Research on the influence mechanism of microblog public opinion dissemination of emergencies based on Logistic-SEM.Journal of Modern Information,2021,41(09):115-1