Research on False Public Opinion Recognition Method of Social Network Based on Fuzzy Cluster Analysis

Gang Qiu[1](✉) and Jun Xie[2]

[1] School of Information Engineering, Changji University, Changji 831100, China
coral1001@163.com
[2] School of Physical Education and Health, A'ba Teachers' University, Wenchuan 623002, China

Abstract. In order to solve the problem that the classification of network pseudo public opinion events is too subjective, a social network pseudo public opinion recognition method based on fuzzy cluster analysis is proposed. Combined with the principle of fuzzy cluster analysis, a new index system for identifying false and emotional events is established. On this basis, the relevant index data of network false and emotional events are collected, and the classical fuzzy cluster analysis algorithm is used to cluster and analyze the network false and emotional events, so as to obtain different types of network false and emotional event sets, analyze and summarize the characteristics of all kinds of false and emotional events, and finally confirm through experiments, The social network false public opinion identification method based on fuzzy cluster analysis has high practicability, provides a new method for the identification and classification of network false and sentiment, and provides a reference for relevant departments to accurately control all kinds of false and sentiment by using network big data.

Keywords: Fuzzy cluster analysis · Social networks · Identification of false public opinion

1 Introduction

With the continuous development of computer technology and artificial intelligence, network public opinion analysis in the era of big data has attracted much attention. At present, the number of Internet users in China has exceeded 900 million, and the Internet has gradually become a distribution center of Ideological and cultural information and an amplifier of public opinion. With the gradual integration of network public opinion into people's life, some fictional, untrue and misguided remarks or events, that is, network pseudo public opinion, are disturbing people's life, enterprise operation and government decision-making. Therefore, how to use big data technology to deeply mine historical data, clarify network pseudo public opinion events, improve the pertinence of pseudo public opinion management, and provide reference basis for relevant departments to

W. Fu and L. Yun (Eds.): ADHIP 2022, LNICST 469, pp. 703–719, 2023.
https://doi.org/10.1007/978-3-031-28867-8_51

accurately control network pseudo public opinion [1]. The decision tree method of network pseudo public opinion collects and preprocesses network public opinion, classifies the processed data through the decision tree method, and finally completes the identification of social network pseudo public opinion on the classified data. However, this method is relatively simple and has the problem of low classification accuracy. In this regard, this paper proposes a method for identifying network pseudo public opinion based on the combination of fuzzy clustering analysis and k-nearest neighbor. To sum up, the existing research on the classification of network pseudo public opinion is only limited to text analysis and manual classification of features, and most of them refer to the existing index system of network public opinion classification. In view of the particularity and complexity of pseudo public opinion, it has many features different from real public opinion [2]. This paper will comprehensively consider the characteristics of network pseudo public opinion, build a brand-new pseudo public opinion index system, use fuzzy clustering analysis algorithm to classify and identify network pseudo public opinion, and combine network big data to provide a new idea for the classification of pseudo public opinion.

2 Identification Method of False Public Opinion in Social Network

2.1 Collection of Pseudo Public Opinion Characteristics of Social Networks

Fuzzy clustering analysis algorithms aim to divide existing datasets into multiple mutually exclusive groups or clusters, so that objects in a cluster are as similar as possible and different from objects in other clusters. Fuzzy cluster analysis algorithm is by far the most widely used and mature data mining cluster analysis algorithm. It has the advantages of being simple, fast and suitable for processing large-scale data. Its basic idea is to randomly select from a dataset containing a large number of data objects. Select k data objects as the initial centroids, calculate the distance between each data object and d centroids, and divide all sample data into the clusters represented by the centroids closest to it. The mean value updates the C_i centroids until the x centroids no longer change in the iterative process, then the clustering result is output [3]. The traditional fuzzy clustering analysis algorithm is adopted, and the specific steps are as follows: select the appropriate k value, and use the sum of squares of errors as the objective function to measure the clustering quality, as shown in the formula:

$$\mathrm{SSE} = g \sum_{i=1}^{K} k \sum_{x \in C_i} d(C_i, x)^2 \tag{1}$$

where g is the standard Euclidean distance between two objects in Euclidean space, select the most suitable k value according to the fuzzy cluster analysis rule, and randomly select S centroids from the data set, usually the selected centroids are The n actual data points already in the sample dataset. Calculate the distance between each data in the data set and S centroids respectively, and divide the data objects into the classes represented by the nearest centroids. This article will use Euclidean distance to measure and calculate the distance between two data points. The centroid is updated with the mean of all data

objects contained in each cluster, and the centroid of the i cluster is defined, where e is the object cluster, and m is the number of objects in the i cluster:

$$|S| = \frac{1}{m_i} \sum_{X \in c_i} \mathrm{SSE} - Sen \tag{2}$$

The discretization algorithm based on normal gain generally calculates the p_i of each attribute, and then selects the attribute value corresponding to the extreme value of $\mathrm{NG}(A, S)$ as the discrete segmentation point. The normal gain calculation formula is as follows:

$$\mathrm{NG}(A, S) = \frac{\sum\limits_{j \in value(A)} \frac{1}{|S|} \sum\limits_{i=1}^{k} -p_i \log_2 p_i}{\log_2 kn} \tag{3}$$

where p_i is the proportion of examples belonging to category i in training set S, and k is the total number of categories. $x(i, k)$ is an attribute, and k is the set of values. $y(i)$ is the total number of samples, and $y(j)$ is the number of samples whose attribute y is j. The normalized input-output correlation value of each sample attribute is calculated as an index for evaluating the importance of the attribute. This paper proposes to use the following formula to calculate the normalized input-output correlation value of the attribute:

$$IOC(k) = \mathrm{NG}(A, S) \sum |x(i, k) - x(j, k)| \times \mathrm{sign}|y(i) - y(j)| \tag{4}$$

Confirm whether the iterative process is terminated according to whether the centroids of various types have changed. If the centroids have not changed, the clustering results will be output. The main flow of the fuzzy clustering analysis algorithm in this paper is shown in Fig. 1.

First, input the data into the data set, randomly select k data objects as the initial centroids, calculate the distance between each data object and the S centroids, divide all sample data into the clusters represented by the nearest centroid, update the centroids according to the average value of all data objects in the newly generated clusters, confirm whether the iterative process ends according to whether the centroids change, and return to recalculate the new centroids if the centroids change, If there is no transformation, each point in the data set is assigned to the nearest centroid to form a cluster, and the cluster result is output.

Online public opinion includes real online public opinion and online fake public opinion. The essence of identifying online fake public opinion is to classify the real online public opinion and online fake public opinion[4]. The network pseudo-public opinion recognition model based on fuzzy clustering analysis is shown in Fig. 2.

Combined with the principle of fuzzy clustering analysis, a brand-new pseudo intelligence identification index system is established. On this basis, the relevant index data of network pseudo intelligence events are collected, and the collected data are preprocessed. Randomly generate a training set and a test set from the processed data. If the proportion of the training set data to the support vector machine is $> 80\%$, the parameter

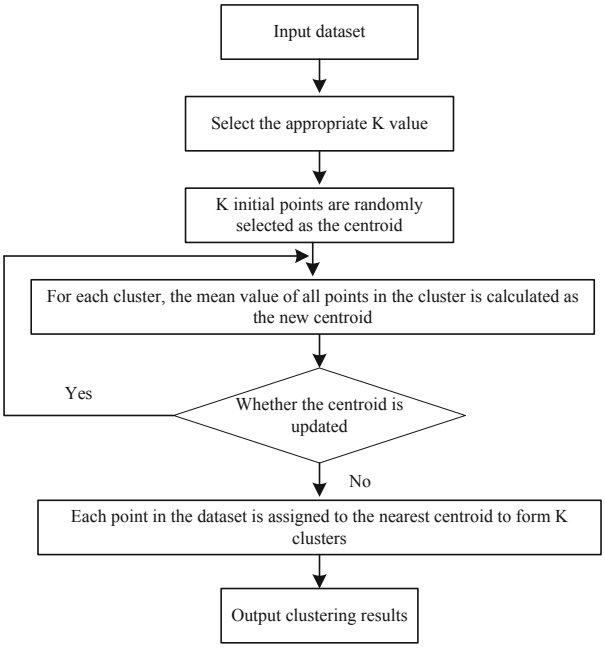

Fig. 1. Flowchart of fuzzy clustering algorithm

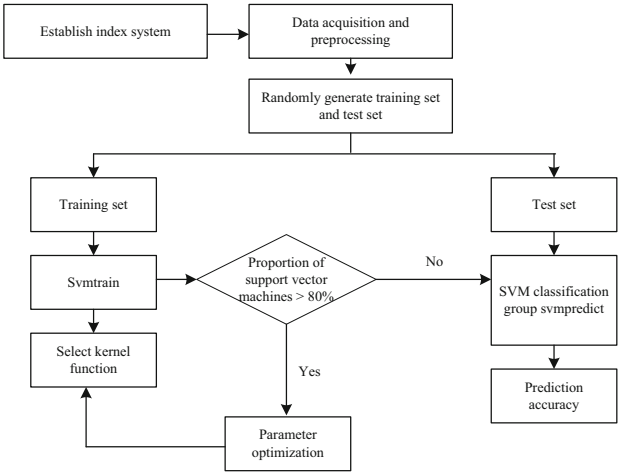

Fig. 2. Network fake and emotion recognition model

optimization will continue, and if it is not $> 80\%$, the iterative process will be terminated. Then the accuracy is tested by SVM prediction of SVM classification group. And the test set is directly tested by SVM prediction of SVM classification group.

Establish a discriminant index for online pseudo-public opinion. Network public opinion can be described by some main index data. There is a relationship between entities and attributes between these evaluation indicators. The selection of reasonable and effective evaluation indicators is the basis and premise of effectively identifying online pseudo-public opinion. Secondly, Collect data based on public opinion indicators [5]. The collected public opinion data includes both real public opinion and false public opinion, and at the same time pay attention to the quantitative balance of the two types of data. The collected data needs to be preprocessed, mainly including the determination of the weight of the information source data, the calculation of the direct information coefficient of the public opinion information from different channels, the supplementation of missing data, and the normalization of data of different dimensions, etc. [6].

2.2 Construction of the Evaluation Index System of Pseudo-public Opinion Features

An important purpose of false public opinion features is to provide a theoretical basis for identifying and monitoring false public opinion. However, not all pseudo-public opinions are both false and artificial [7]. Modeling is carried out according to partial differential equations. First, the relevant symbols are given as shown in Table 1.

Table 1. Description of symbols of pseudo-public opinion dissemination model

Symbol	Description (probability)	Symbol	Description (probability)	Symbol	Description (probability)
a1	P → F	a5	F → R	a9	B → R
a2	R → F	a6	AS → R	a10	Equilibrium value
a3	F → AS	a7	IS → R	a11	Node transition state time
a4	F → IS	a8	B → AS		

The conversion rate in the social network public opinion crisis information dissemination model can be determined by the multi-subject attribute. The subject attribute is usually statistically analyzed by the expert scoring method and the survey method. Because the subject attribute statistics are carried out by means of dimensionless quantity statistics or dimensionless expert scoring values. Given, the conversion rate and the influence rate should be controlled within the range of [0,1], so all conversion rate formulas calculated by using the main attributes are expressed as ratios. The main attributes are shown in Table 2.

The three types of pseudo-public opinions A, B, and C are defined as follows: if some public opinions are false, that is, the events or opinions that caused the public opinions are not in line with the facts, then no matter whether there is human intervention or whether the publishers have a tendency to the opinions, The resulting public opinion discussions are also false public opinion. Such pseudo-public opinions are category A pseudo-public opinions, such as false information such as “radish can cure cancer” [8].

Table 2. Attributes of pseudo-public opinion subjects

Subject	Main attributes
government	Authority, credibility, processing speed, information disclosure and transparency
netizen	Participation, attention, conformity and polarization
media	Attitude, media influence, communication intensity
Internet media	Credibility, activity and reporting frequency
event	Hazard, fuzziness, explosiveness and sensitivity
Mobile social platform	Social influence, communication speed, opinion leader force

Falsehood is a necessary condition for Class A false public opinion. If the pseudo-public opinion itself cites real events or a certain personal opinion, likes and dislikes, but uses human intervention to create and guide public opinion, the final public opinion that violates the objective development law of online public opinion is also a kind of pseudo-public opinion, which is called Type B pseudo public opinion, that is, the opinion or event of public opinion is not false or does not involve authenticity but the process of formation is false. The creation of such pseudo-public opinion often has a clear purpose, and the views held have a clear tendency, otherwise it will be meaningless to artificially intervene in the dissemination of topical events or speeches. Fundamentally speaking, the reason why the B type pseudo-public opinion is called pseudo-public opinion is that it adopts the means of artificially intervening in the formation and development of public opinion, that is, the artificial characteristic is the necessary condition for the B type pseudo-public opinion. A considerable part of pseudo-public opinion may have both false and artificial characteristics, which is called C type pseudo-public opinion [9]. The use of false, forged, tampered, and exaggerated events and the use of various means to induce or mislead netizens in order to achieve their own goals is often the most serious harm of such false public opinion, and it is also the focus of online public opinion supervision. Falsehood and artificiality are the necessary and sufficient conditions for Class C false public opinion. Therefore, the relationship between pseudo-public opinion features and types is shown in Fig. 3.

In order to effectively warn the public opinion crisis of mobile social networks, this paper uses the fuzzy analytic hierarchy process to construct the crisis early warning index system, and uses the triangular fuzzy number to determine the index weight, so as to determine the risk index of the network public opinion crisis early warning index. The network public opinion crisis early warning index system is shown in Fig. 4.

From the perspective of six main factors, the popularity of online public opinion is mainly affected by social networks and public opinion; for online public opinion attitude, it is mainly affected by the government and netizens; for online public opinion behavior, it is mainly affected by events, media, online media and social influence, therefore, for different indicators, they correspond to different main influencing factors, this study will give each main factor and its internal influencing factors in the simulation of crisis early warning effect on the network public opinion crisis early warning. Overall effect [10].

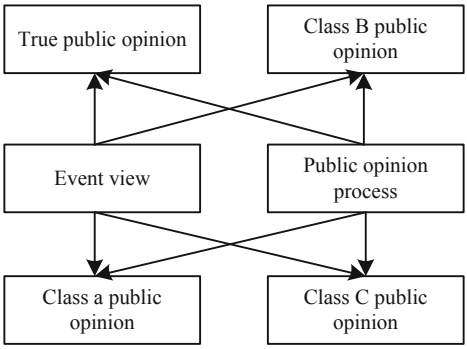

Fig. 3. Relationship between characteristics and types of pseudo public opinion

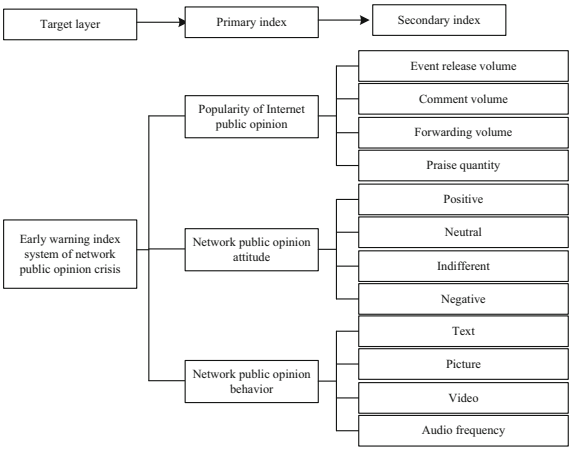

Fig. 4. Network pseudo-public opinion crisis early warning indicator system

For the subjective influence level, the triangular fuzzy comment value given by experts is used to determine the importance of the index, and the evaluation judgment matrix and the fuzzy positive and negative value matrix are determined. Combined with the above-mentioned indicators obtained by intuitive observation, some indicators that have no hysteresis defects in the past network pseudo-public opinion research are used, and the final indicator system is constructed as shown in Table 3.

For each event, the propagation trend data of three channels are collected, and the raw data of the kth network public opinion event is

$$X_k = \left\{x_{t_{k1}}, x_{t_{k2}}, x_{t_{k3}}, \cdots, x_{t_{ki}}, y_k\right\} \tag{5}$$

Among them: x is the propagation times of the k event t_k time, and its expression is

$$\overrightarrow{x_{t_{kl}}} = \left(t_{ki}, w_{t_{ki}i}, e_{t_{ki}i}, n_{t_{ki}}, s_{t_{ki}i}\right) \tag{6}$$

Among them: t_{ki} is the i hour after the first occurrence of the k event, in hours, $w_{t_{ki}i}$ is the number of social communication within the ki hour after the first occurrence of

Table 3. The index system of the network pseudo-public opinion recognition model

Category	Index name	Remarks
degree of heat	Number of hot discussions	The total number of peaks in the transmission trend after the event
	Average propagation speed	The value of the ratio of the total propagation to the duration up to now
Communication channels	Proportion of microblog contribution	The spread rate of microblog to change time
	Proportion of wechat contribution	Wechat's spread rate of the event
	Proportion of online media contribution	The spread rate of the event by online media
	Dispersion of communication channels	Describe the channel concentration of events in the communication trend
information source	channel	Channels of initial information sources
	User influence	H index of the first user within 30 days before the event

the t event, $e_{t_{ki}i}$ is the first occurrence of the k event. The number of spreads in WeChat within the i hour, $n_{t_{ki}}$ the number of spreads in the Internet media within the i hour after the k event first appeared, $s_{t_{ki}i}$ is the total number of spreads within the ith hour after the k event first appeared y_k is The authenticity of the artificially marked k public opinion event, 0 represents the real public opinion, and 1 represents the network fake public opinion. The degree of early warning of online public opinion crisis is analyzed through the fuzzy comprehensive evaluation results. The delineation of the police limit can fully describe the effect of early warning of online public opinion crisis, provide a scientific and accurate judgment basis for the dynamic simulation of the online public opinion crisis system, and ensure the early warning of social network public opinion crisis. The sensitivity of the model is ensured, and the government and relevant departments can timely detect changes in the police situation and make corresponding responses. The delineation of the police limit is shown in Table 4.

In order to effectively control the occurrence and development of the crisis of network public opinion, it is necessary to enhance the awareness of the crisis of network public opinion in relevant government departments, and the managers of network public opinion should unify the opinions raised by netizens when solving emergencies and network public opinion problems before making decisions., Comprehensively consider the real situation of the incident and compare it with the public feedback information, so as to avoid the occurrence of mass incidents. After emergencies occur, the government

Table 4. Delineation of early warning limits for online public opinion crisis

Indicator color	Alarm degree	Warning limit
Yellow	D	[0,0.2]
Orange	C	[0.2,0.3]
Powder	B	[0.3,0.7]
Purple	A	[0.7,0.9]
Crimson	S	[0.9,1.0]

should publish authoritative and authoritative information in a timely manner, and open communication channels between the public and the government, so that the public's right to speak is not limited to social networking, so as to avoid online public opinion caused by the spread of rumors deterioration occurs. The media must be guided by the mainstream opinions of the real-time information released by the media, and explain in a rational and objective way. In the process of information dissemination of online public opinion crisis, the media mainly takes the role of "gatekeeper" to screen and transmit information. Effective control, and the authenticity and reliability of media information dissemination determine the success of social network public opinion crisis early warning. Combined with the multi-agent interactive analysis of the public opinion crisis information dissemination mode, and based on the SIR model to explain the network public opinion crisis information dissemination, the network public opinion crisis information dissemination mode is shown in Fig. 5.

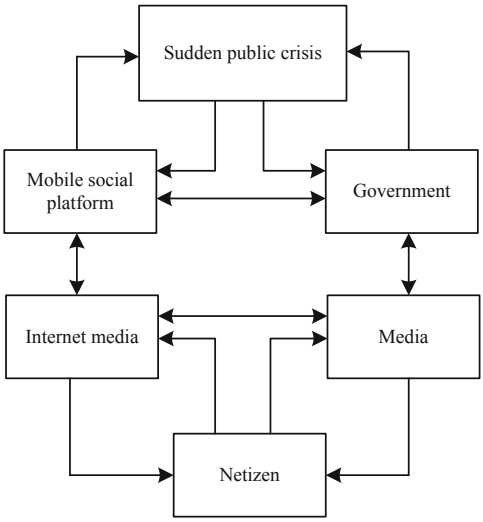

Fig. 5. Network public opinion crisis information dissemination mode

The development law of network public opinion is contained in various public opinion index data, and collecting effective indicators can improve the rationality and accuracy of pseudo-public opinion classification. At present, the research on fake public opinion on the Internet mainly focuses on improving the efficiency and accuracy of identifying fake public opinion. Therefore, the public opinion indicators selected by these studies focus on the characteristics that distinguish fake public opinion on the Internet from real public opinion. In order to achieve a reasonable classification of fake public opinion on the Internet On the basis of relevant research, this paper adds relevant indicators for the overall monitoring of public opinion, and on the basis of following the principles of scientificity, practicability, flexibility and clarity of division, the following 9 main indicators are selected as the identification of false public opinion on the Internet Indicator: Influence Index. Based on the social media and online media data of the whole network, weighted summation is made according to the dissemination effect of public opinion events on social media (mainly social media, WeChat) and online media, and the event influence after the summation Then, the event influence index ranging from 0 to 100 is obtained through normalization operation, which is an authoritative index used to reflect the effect of a public opinion event on the Internet.

2.3 Realization of Network Fake Public Opinion Recognition

The connotation of network public opinion Since it was put forward, public opinion usually refers to a collection of attitudes and opinions expressed by a certain number of people towards a certain event. The Internet public opinion originates from life but is formed on the Internet, and it is spread through the Internet. The relevant analysis of network public opinion is given below, as shown in Fig. 6.

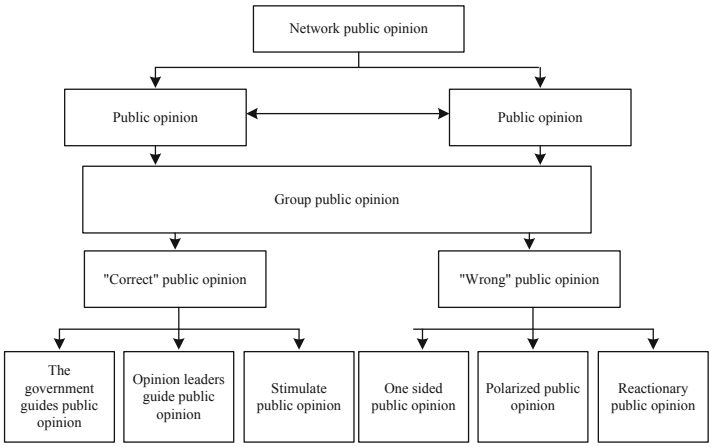

Fig. 6. Correlation analysis of network examples

Public opinion in a broad sense mainly includes three types, one is social public opinion, which can reflect public opinion at different levels from the national level,

social level and class level, and highlights the social problems existing at different levels. The second is public opinion, also known as public opinion, which can reflect the perceptions, views, wishes and demands of people of different classes and strata towards the real society. The third is Internet public opinion. People use the Internet as a public platform to express their opinions and opinions, and use Internet language, video, pictures, comments, likes, forwarding and other forms to evaluate public events. The determination of the overall risk index of the social network public opinion crisis warning is also calculated. Through the calculation, the risk index of the social network public opinion crisis warning is equal to 0.612. The state of the crisis warning in this period is a severe level, and the indicator color is orange. At the same time, the government and relevant public opinion management departments should be reminded to increase the management of online public opinion Table 5. The crisis in which the risk index of all levels of indicators is in:

Table 5. Hazard index of indicators at each level

Index hierarchy	Variable	Hazard index	Index hierarchy	variable	Hazard index
Primary index	S1	0.358	Secondary index	S11	0.591
				S12	0.716
				S13	0.429
				S14	0.218
	S2	0.462		S21	0.0258
				S22	0.148
				S23	0.0035
				S24	0.796
	S3	0.629		S31	0.832
				S32	0.593
				S33	0.133
				S34	0.723
				S35	0.599

Therefore, from the perspective of practical analysis, it can be seen that the essence of network public opinion in a broad sense is still a form of public opinion that condemns and resists social events in the real world. Network public opinion in a narrow sense refers to an act of using network media to discuss emotions and attitudes. The study uses the broad form of network public opinion to extract social hot events, so as to carry out statistics of relevant information and data, and analyze the narrow network public opinion from different angles, in order to determine the development trend of network public opinion in emergencies, so as to correctly guide the development of public opinion. Provide theoretical support. Constant parameters are given as Table 6.

Table 6. Constant parameters of pseudo public opinion recognition model

Variable name	Value	Variable name	Value
Event publicity degree	36.85	Platform audience	1000
Event sensitivity	74.65	Platform activity	68.15
Hazard degree of the event	79.33	Media Authority	42.15
Emergency degree	46.85	Media activity	46.98
Polarization degree of Internet users	66.85	Online entertainment authority	46.86
Internet users' psychological intensity of curiosity	43.19	Online entertainment activity	52.28
Internet users' emotional intensity	75.26	Information disclosure and transparency	72.6
Role of opinion leaders	45.83	Authority of press spokesman	62.6
Internet users' forwarding intention	58.65	Government response speed	35.32
Content audit degree	33.65	Satisfaction with event handling	21.36

Based on social public opinion activities, from the perspective of multiple subjects, it determines the relevant factors that affect the formation of online public opinion crisis, and identifies the role of public opinion content in netizens, the government and society. It fully describes the size and judgment process of the crisis that may be caused by the public opinion formed by the event, and clarifies the judgment basis and process of the early warning signal. Social Network Public Opinion Crisis Early Warning is a universal public opinion crisis early warning system that can strengthen the protection of online public opinion. The specific framework is described in Fig. 7.

The network public opinion crisis early warning system used in social networking can accept comments from netizens at any time, and can timely, refine and deeply understand the trend of group intentions, and conduct targeted online education for netizens to avoid netizens because of group conformity. Psychology tends to polarize public opinion; scandal revelations, etc., in order to prevent public opinion from being pushed to the brink of group polarization and let the irrational behavior of netizens prevail, it is possible to effectively and timely respond to the crisis that may be about to arise in network public opinion. Forecast, minimize the degree of harm suffered by netizens, and correct the role of the media in responding to the crisis of online public opinion. The media, as the main medium for information dissemination, should supervise the positive development of online public opinion with a benign orientation in the development process of the crisis of online public opinion. The media has a certain "information buffer" effect on the network public opinion formed by sudden events. The media has different influences on online public opinion in the early, middle and late stages of the incident. In the early stage, because netizens reported the emergency through WeChat, social networking and other social networks, the public had an intuitive view of the malignant impact of the

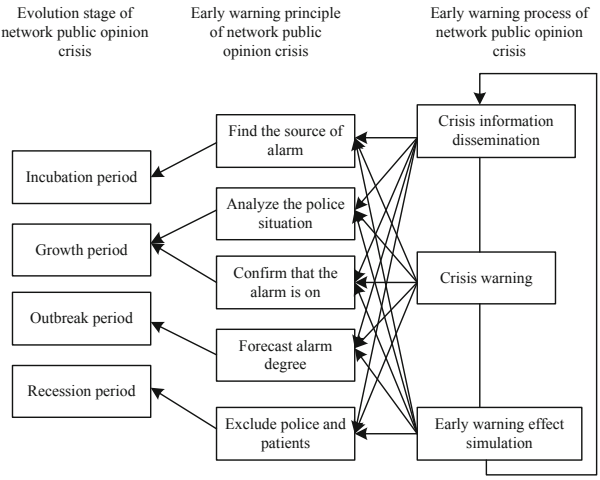

Fig. 7. Designing the network as a framework for public opinion identification

emergency. The feeling of the Internet public opinion has transformed from the gestation period to the growth period, which shows the backwardness of traditional media and Internet media and the mobile social network platform. In the middle and later stages of emergencies, the real-time reports on the extent of the impact of emergencies on the country and society, the extent of emergency handling, and the government's handling of such emergencies are unclear. Increased guesswork.

3 Analysis of Results

Using the scikit-learn library in Python to realize the construction of Logistic Regression and the recognition model in this paper, in order to verify the application effect of using the propagation trend data of public opinion events in identifying fake public opinion on the Internet, according to the authenticity of public opinion events and the prediction of the model As a result, the classification results can be obtained as shown in Table 7.

Table 7. Identification result confusion matrix

		Forecast category	
Real category	yes	yes	no
		YP (real)	WN(false negative)
	no	WP (false positive)	YN (true negative)

In order to further verify the effectiveness of the method in this paper, the four types of public opinion A, B, C, and D are unified into pseudo public opinion, and the four sets of experimental data are respectively applied to the proposed fuzzy clustering analysis

model, in which the kernel function adopts the radial basis kernel. The parameters are optimized by the particle swarm parameter optimization method. The recognition and prediction accuracy of the 4 sets of validation sets are shown in Table 8.

Table 8. Pseudo-public opinion recognition and prediction accuracy

Group	A	B	C	D
Accuracy of true public opinion identification	100%	96.8%	100%	96.5%
Accuracy of false public opinion identification	90.8%	90.5%	90.5%	91.3%

Divide the data into the training set and the test set, use the method in this paper and the Logistic Regression model to fit the test set, and finally use the ten-fold cross-check to identify the public opinion events in the test set, and calculate the accuracy, precision and recall of the two models respectively. The average of the rate scores, as shown in Table 9.

Table 9. SVM and Logistic Regression model multi-index evaluation table

	Accuracy	Accuracy rate	recall	F1 score
Paper method	0.8335	0.8223	0.8335	0.7667
Logistic Regression	0.6811	0.6	0.69	0.6155

Compared with the data in the table, it is found that the prediction accuracy of the method in this paper is higher than that of the fuzzy cluster analysis method in the two models. The final cluster centers of the three categories obtained by mean clustering are shown in Table 10.

It is worth noting that the model only divides the authenticity of public opinion, while the method proposed in this paper needs to further divide the specific categories of false public opinion. The specific impact of online false public opinion events is shown in Fig. 8.

The impact on the formation degree of online public opinion crisis is shown in the curve in the figure above. The crisis formation degree shows decreasing changes, especially the decrease in the sensitivity of the event leads to a greater reduction in the formation degree of online public opinion crisis. For any pseudo-public opinion example, If it is only classified as false public opinion, the judgment is correct. Because the training set is randomly selected, there are differences in the dividing points and the number of intervals after discretization, but they are roughly stable at 4 or 5 intervals from high to low. The rules for each experiment were not exactly the same. Based on this comparison of the actual dissemination of fake public opinion, the public opinion tracking and identification of the method in this paper and the traditional method are compared and recorded, as shown in Fig. 9.

Table 10. Cluster center result table

Category	Influence index	Total number of participating media	Central level media participation	Proportion of public opinion field	Peak event heat
A	55.9755	28.9509	0.2425	0.0425	316.8565
B	59.377	32.9	0.2758	0.0723	1099.856
C	62.5	41.42859	0.385	0.1358	2392.652
Category	Peak propagation velocity	Duration	Number of communication channels	information source	Number of events
A	49.4985	182.6528	96.785	1.985	62
B	113.9	129.6	149.9856	1.8668	16
C	160.2859	195.1652	232.7658	1.7256	8

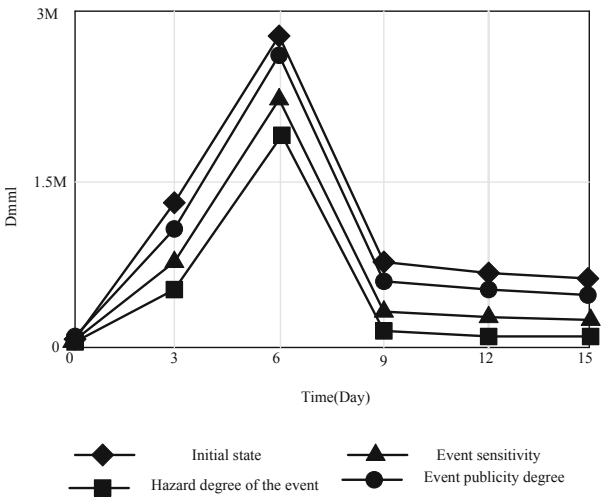

Fig. 8. Influence of online pseudo-public opinion events

In the process of establishing the fuzzy clustering identification model, the number of posts per hour, the duration of the first media, and the proportion of the largest opinion are the most important attributes of the root node, which are basically consistent with the identification of actual public opinion. The accuracy is relatively high.

In order to verify the effectiveness of the method in this paper, taking the accuracy of pseudo public opinion recognition as the experimental index, the method in this paper and the decision tree method are used for experimental testing. The test results are shown in Table 11.

It can be seen from the results shown in Table 11 that the accuracy of the method in this paper is up to 100%, and that of the decision tree method is up to 82%. It can be seen

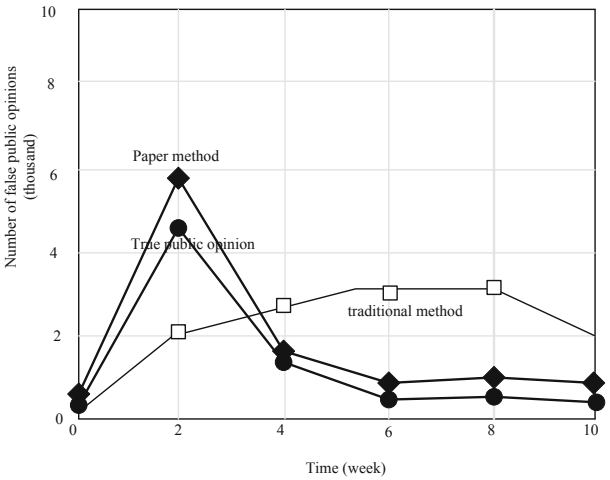

Fig. 9. Shows the situation of public opinion recognition

Table 11. Comparison table of pseudo public opinion identification and prediction accuracy

Number of experiments / time	Methods in this paper	Decision tree method
10	100%	82%
20	99%	80%
30	99%	81%
40	98%	82%
50	100%	80%

that the accuracy of public opinion recognition of this method is significantly higher than that of decision tree method. It shows that the method in this paper has high accuracy in identifying false public opinion. It proves that the technology level and application value of the method proposed in this paper are high.

4 Concluding Remarks

A set of new index system is established for the identification of online pseudo-public opinion, which covers the macro-index data characteristics of most pseudo-public opinion events, and further uses the classical fuzzy clustering analysis algorithm to divide the network pseudo-public opinion events into three categories, and analyzes them in detail. The characteristics of various kinds of pseudo-public opinion events are enriched, and the classification research of online pseudo-public opinion is enriched, which provides a reference for the relevant departments to accurately control and guide all kinds of pseudo-public opinion events, and contributes to the harmonious and stable development of the network society.

5 Fund Project

The work was supported by the 2021 Autonomous Region Innovation Environment (Talents, Bases) Construction Special-Natural Science Program (Natural Science Foundation) Joint Fund Project(2021D01C004) and 2020 Changji University Teaching and Research Project (20jyyb002).

References

1. Li, A., Fu, J., Shen, H., et al.: Indoor positioning algorithm based on fuzzy clustering and cat swarm optimization. Chinese J. Sci. Instr. **41**(01), 185–194 (2020)
2. Huang, L., Yao, B., Chen, P., et al.: Superpixel segmentation method of high-resolution remote sensing image based on fuzzy clustering. Acta Geodaetica et Cartographica Sinica **49**(05), 589–597 (2020)
3. Qz, A., Zw, A., Aha, C., et al.: A two-step communication opinion dynamics model with self-persistence and influence index for social networks based on the DeGroot model. Inf. Sci. **519**, 363–381 (2020)
4. Yin, J., Chen, H., Wang, J.: Negative public opinion propagation mechanism and evolutionary game analysis in online social networks. Inf. Sci. **38**(04), 153–162 (2020)
5. Ma, H., Wang, J., Lin, H., et al.: A multi-view network for real-time emotion recognition in conversations. Knowl.-Based Syst. 107751.1–107751.10 (2022)
6. Han, P., Gao, S., Yu, Z., et al.: Case-involved public opinion news summarization with case elements guidance. J. Chinese Inf. Process. **34**(05), 56–63+73 (2020)
7. Chen, L., Zheng, J., Cao, J., et al.: Intelligent traffic sign recognition method based on capsule network. J. Comput. Appl. **40**(4), 1045–1049 (2020)
8. Wang, L., Gan, S., Lin, N., et al.: Construction of collective system and analysis framework for china-related public opinion. J. Intell. **39**(06), 135–142 (2020)
9. Zhang, X., Wang, Y.: Information dissemination efficiency evaluation of network public opinion of the government microblogs based on input-output analysis. Inf. Sci. **38**(05), 43–48 (2020)
10. Liu, X., Nan, Y., Xie, R., et al.: DDPG optimization based on dynamic inverse of aircraft attitde control. Comput. Simul. **37**(07), 37–43 (2020)