J. Electrical Systems 20-2 (2024): 1961-1972

[1]Jiaqi Mo
[2]Ziqi Wang
[3]Yuhan Jiang
[4]Mingying Shi
[5]Mengru Wang
[6,*]Yongcong Luo

Research on the Sentiment Analysis and Evolutionary Mechanism of Sudden Network Public Opinion Based on SNA-ARIMA Model with Text Mining

Abstract: - This article is based on the massive text data in the digital era, using text mining methods to explore the evolution mechanism of online public opinion, in order to achieve the control of public opinion trends in sudden network events and regulate the stability of online and offline public opinion. We introduced time series for decision-making assistance, combined with time series autoregressive integrated moving average model (ARIMA) for prediction, and comprehensively evaluates the evolution of public opinion from two dimensions: sentimental orientation and social network analysis (SNA). Based on this, a new exploration model for public opinion evolution, SNA-ARIMA, is constructed. The research results indicate that the process of public opinion dissemination exhibits distinct phased characteristics, and each stage is comprehensively influenced by different key nodes or factors. This study provides decision support for managing public opinion crises, making the identification of key nodes in online public opinion events, the evolution mechanism of public opinion, and the guidance of public opinion crises more systematic, forward-looking, and scientific.

Keywords: Evolution of Online Public Opinion (EOPO), Key Nodes (KN), Text Mining (TM), Time Series Analysis (TSA), Sentiment Analysis (SA), Social Network Analysis (SNA).

I. INTRODUCTION

With the popularization and development of digital technology, digital resources continue to enrich, and social platforms continue to play a bridging role between virtual space and the real world, gradually becoming one of the mainstream tools for netizens to express their opinions and sentiments. The emergence of the Internet and social media has changed the mode of information dissemination and interaction. All kinds of emergencies rely on social media platforms to enter the public view and become the focus of online public opinion dissemination [1].

Social media, as a comprehensive information platform that combines social attributes and communication functions, has risen to become a key carrier of information dissemination and communication in modern society [2]. Its unique dissemination characteristics make the re dissemination of information the mainstream way of information dissemination, thus profoundly affecting people's communication and exchange methods. With the fermentation of public opinion, the research object of online public opinion gradually turns to sentimental flow, and users gather independent individual users into a network of relationships through forwarding, commenting, posting, and other forms [3].

Therefore, in the era of big data and the increasingly developed artificial intelligence technology, this article is based on the text data of social media and uses text mining methods to deeply study the evolution process of online public opinion and the key nodes that affect this process. Among them, the TSA and ARIMA was combined for prediction, and the evolution of public opinion was comprehensively evaluated from two dimensions: sentimental orientation and SNA. On this basis, a new model for exploring the evolution of public opinion, SNA-ARIMA, was constructed. This study has crucial practical significance for controlling public opinion and maintaining people's sentimental stability.

The sudden public incident of "A company cutting off its knife" has caused widespread discussion among Sina Weibo users, leading to a brand crisis that has caused the company's market value to evaporate by 98 million yuan. This event has been selected as one of the "Top 10 Consumer Rights Protection Public Opinion Hotspots in 2022" jointly announced by the China Consumers Association and the People's Daily Public Opinion Data Center. In this context, this article takes the "A company's knife cutting incident" as an example, uses time series prediction

[1] School of Computer and Computing Science, Hangzhou City University, Hangzhou 310015, China
[2] School of Business, Hangzhou City University, Hangzhou 310015, China
[3] School of Computer and Computing Science, Hangzhou City University, Hangzhou 310015, China
[4] School of Business, Hangzhou City University, Hangzhou 310015, China
[5] School of Medicine, Hangzhou City University, Hangzhou 310015, China
[6] School of Business, Hangzhou City University, Hangzhou 310015, China
*Corresponding author: Yongcong Luo

methods, combines sentimental tendencies and social network indicators, identifies key nodes in the evolution process of online public opinion, clarifies the internal mechanism of public opinion evolution, and achieves the role of timely analysis and scientific management of public opinion.

In summary, after considering and analyzing the characteristics of previous research, this paper constructs a model called SNA-ARIMA to analyze the sentimental characteristics and evolution process of public opinion. Our main contributions are summarized as follows:

- This article innovatively introduces time series analysis, enriching the methodological system of online public opinion research, making the research more comprehensive and in-depth. The comprehensive evaluation of sentimental tendencies and social network attributes not only deepens the understanding of key nodes, but also provides strong support for public opinion prediction and trend grasping.
- This article aims to gain a deeper understanding of the dynamic evolution of online public opinion, with a particular focus on identifying hidden key nodes: nodes that may appear inconspicuous on the surface but may play a significant role in specific contexts, such as potential bridging roles or suddenly active "dormant" accounts. This helps decision-makers grasp certain unexpected public opinion events and take targeted measures to guide or calm public opinion.
- This article proposes a construction plan for a real-time network public opinion monitoring system and intervention strategies for key nodes, providing a practical and feasible path for rapid response and effective management of public opinion crises.

The following parts of this article are organized as follows: In Section II, we conducted a comprehensive review of previous research on sentiment analysis of network public opinion in the field of emergency events and related fields, i.e., the related work. Section III discusses model construction and algorithm design, i.e., the methodology. Section IV introduces the experimental results and related analysis, i.e., the results and discussion. Finally, Section V presents the overall conclusion of this article and some suggestions for future development direction, i.e., the conclusion.

II. RELATED WORKS

A. Public Opinion and Sentiment Analysis

Online public opinion reflects the public's cognition, attitude, sentiment, and inclination towards social events and immediate interests [4]. The current research on sentiment analysis and mining of social media texts by scholars is also known as opinion analysis and mining. It integrates multiple disciplines such as sociology, psychology, and computer science. By preprocessing subjective texts and objective texts containing implicit subjective sentiments, sentiment classification is carried out to analyze people's attitudes towards various events, thereby achieving the goal of finding hot topics of public concern, understanding public psychology, and analyzing sentimental trends [5]. The sentimental expression of users not only affects the speed of information dissemination, but also to some extent infects the sentiments of other users, leading to the spread of public opinion [6]. Taking “Aspect based Sentiment Analysis” as the research object, Nazir et al. [7] addressed issues and challenges in extracting different aspects and their related sentiments, mapping relationships between aspects, interactions, dependencies, and contextual semantic relationships between different data objects, in order to improve sentimental accuracy and predict sentimental evolution dynamics. At present, supervised sentiment analysis methods require dividing a large amount of data in specific fields into training and testing sets, so that the model can learn to distinguish the polarity expression of emotions in sufficient training, which is not applicable in other fields [8].

As an important tool for sentiment analysis, sentiment lexicons are constructed by obtaining texts posted by users on social media, and using programming languages to construct models to analyze users' sentimental attitudes [9]. In order to analyze the public sentiment of popular topics on the Internet, Wang et al. [10] mined the changes in the public sentiment of topics through the search and association of frequent words in text data. The sentiment analysis method based on sentiment lexicons assigns values to sentimental words in the text, calculates the overall sentiment score of the text, and judges the overall sentiment polarity of the text. This has been well applied in sentiment analysis of various types of network texts [11]. Once the sentiment dictionary is successfully constructed, adjustments only need to be made with a small amount of corpus data to achieve efficient and accurate analysis results [12]. But in the analysis of public opinion in sudden events involving public interest, this method is still being explored. Based on previous research and the sentimental needs of unexpected events in online public opinion, this article adopts a data augmentation method to reduce the negative impact of data imbalance on learning models in sentiment analysis. The sentiment lexicon method is used for sentimental research of social network public opinion.

B. Public Opinion and Time Series Analysis

Time series refers to arranging a series of data in chronological order. Time series prediction analysis is commonly used to study the trend of events over time, including periodic, seasonal, and random changes [13]. With the advent of the big data era, users constantly engage in interactive behavior with the outside world, generating and accumulating a large amount of event data. More and more data sources and research objects are emerging, making time series analysis increasingly important. Time series data has become a data of interest for many researchers and practitioners [14]. The application fields of time series analysis are constantly expanding. For example, in finance, historical time series data and related pattern information can be used to predict future volatility trends, which can help enterprises or individuals make more accurate investment decisions and reduce decision-making risks [15]. Time series is ubiquitous in people's lives, and making good use of time series can comprehensively monitor the process of public opinion, analyze the sentimental evolution trend of network users in public opinion events, grasp the development status of public opinion, and help with public opinion monitoring [16], [17]. Li et al. combined text information with emotional time series to comprehensively analyze multi user and multi document sentiment within a unit time of the time series, in order to achieve multi document sentiment prediction [18]. The current mainstream heat prediction method is a time series analysis method based on quantitative models. Time series analysis is not only suitable for short-term prediction but also for long-term prediction, with advantages such as strong interpretability, and is widely used in the field of heat prediction [19].

At present, the main methods for time series prediction are statistical regression, and commonly used methods include autoregressive moving average (ARMA) and autoregressive integrated moving average model (ARIMA) [20]. ARIMA uses differential operations to make the time series tend to be stationary, thereby finding patterns in the data. In order to accurately predict the degree of online public opinion hotspots, Su et al. [21] proposed an improved seasonal public opinion prediction model. It uses a model combining SMEGBM and ARIMA to predict seasonal and trend sequences. Given the complexity of the relationship between sudden public opinion event data, this article adopts the ARIMA model for sentiment analysis of public opinion events.

C. Public Opinion and Social Network Analysis

Social network analysis, as an interdisciplinary research method, aims to reveal the relationship patterns and dynamics in social structures. The main body of information dissemination is "points", and the relationship between each subject is "lines". A complex social network relationship domain is formed by points and lines, in which there are a few nodes with a wide range of influence, high activity, and strong importance, called key nodes, also known as opinion leaders, network elites, etc. [22]. With the rapid development of network technology and the frequent occurrence of emergencies, the academic community has begun to combine the analytical paradigm of networks with the dissemination of crisis information. In emergency situations, network-based emergency response has more advantages than traditional hierarchical response [23]. Wei et al. [24] constructed a cost function for public opinion change and studied the evolution law of public opinion based on the scope and amount of information dissemination by public figures. They found that key figures play a guiding role in public opinion, and they bear group pressure from the amount of information. The relevant research on key nodes in online public opinion mainly elaborates and analyzes in detail from the aspects of social network topology and comprehensive evaluation based on content attributes.

The network topology analysis method considers the contribution of inter node interconnection relationships to influence based on the social network relationships of nodes. Riquelme et al. [25] proposed a generalized, parameterizable centrality measure called Generalized Influence Diffusion Rank to investigate how these parameters interfere with the centrality of public opinion participants. The research results indicate that active public opinion individuals must rely on some neighboring nodes to increase their initial influence on public opinion dissemination. Rehman et al. studied the impact of key nodes (opinion leaders) of public opinion on followers through various central indicators such as central position, structural pores, and personal penetration [26]. Kumar et al. proposed a voting-based method and neighborhood kernel method to identify influential nodes [27].

The content attribute based comprehensive index evaluation method focuses on specific platform environments, mainly using methods such as Analytic Hierarchy Process, Principal Component Analysis, and Comprehensive Evaluation to analyze the influencing factors from the perspective of topic dissemination content, such as the number of fans, likes, forwarded comments, and other indicators for mixed weighted modeling, in order to evaluate the importance of key nodes in online public opinion [28]. Hou explores whether the network or content characteristics of users are more decisive in identifying opinion leaders through the Yelp platform, where users are voted by others [29]. Zhan et al. [30] proposed a social network opinion and behavior evolution model to study the bounded confidence evolution of opinions and behaviors in social networks. The results showed that the probability of social network connection, bounded confidence, and opinion threshold of behavior selection parameters have a strong impact on the evolution of opinions and behaviors.

The above studies have used different methods to provide a basis for identifying key nodes, but there is a common problem of limitations in identifying key nodes due to the single analysis dimension, and very few have combined social network indicators, sentiment analysis, and time series prediction of public opinion information to study the identification and evolution mechanism of key nodes in public opinion. Therefore, this article optimizes the multi weighting of indicators that integrate sentimental tendencies, as well as the identification and classification of key nodes.

III. METHODOLOGY

When dealing with unexpected incidents of online public opinion, public opinion emergency managers face complex issues of dynamic information changes. This study is mainly divided into three modules: the sentiment analysis module for unexpected events, which executes data on initial network event public opinion, and constructs an sentiment analysis model for unexpected events by combining sentiment dictionaries; In the social network analysis module, proximity centrality is an important indicator reflecting the degree of control of communicators over the network, while betweenness centrality is a key indicator in social network analysis, used to measure the ability of a node to control information flow in the network; Time series analysis module, based on ARIMA model for time series analysis, tracks the evolution process of public opinion over time. Combining social network analysis and time series analysis, further classify the calculation results of the filtering model, and identify the key nodes and turning points that cause public opinion mutations. To gain a deeper understanding of the dynamic evolution of online public opinion. The specific framework is shown in Figure 1.

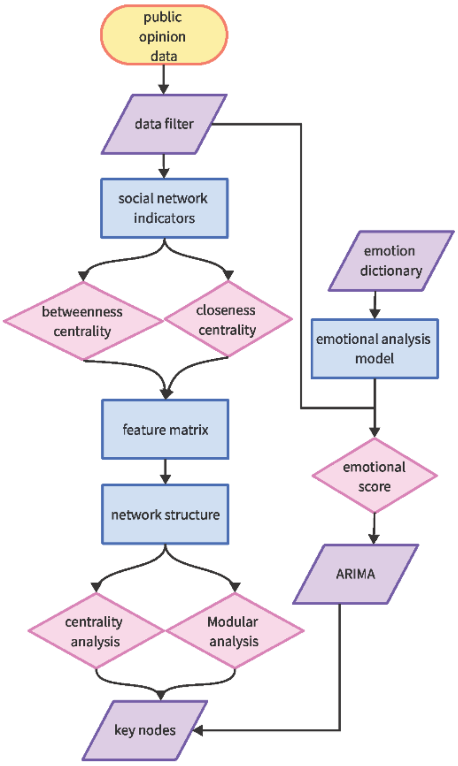


Figure 1: SNA-ARIMA Model Framework Diagram

A. *Sentimental Analysis Module for Unexpected Events*

This article aims to establish a rigorous sentiment analysis model based on the Posen sentiment dictionary, negative word list, and degree level words. This module can accurately identify sentimental words, negative words, and degree adverbs in the text, and calculate the sentimental score of the text based on their weights. This will help this article to better understand and analyze the sentimental tendencies of online public opinion. The specific operation steps are as follows:

1) Data preparation

Select the Posen Sentiment Dictionary as the basic sentiment lexicon and integrate the negative word list. In order to build a more comprehensive stop word list, this article integrates resources such as the Chinese stop word list, the Harbin Institute of Technology stop word list, Baidu stop word list, and the Machine Intelligence Laboratory stop word library. After referring to a large number of literatures, this article briefly marked the dictionary with degree adverbs, where greater than 1 indicates sentimental strengthening and less than 1 indicates sentimental weakening. Classify and assign weights to words based on specific levels of hierarchy.

2) Data preprocessing

Filter negative words or degree adverbs in the stop word list. It should be noted that the negative words or degree adverbs included in the dictionary should be filtered out from the stop word list, otherwise the sentimental score obtained may have errors.

3) Model construction

Identify sentimental words, negative words, and degree adverbs in the text; calculate single sentence score: traverse all sentimental words to see if there are negative words and degree adverbs before the current sentimental word. If there are negative words or multiple negative words, multiply the sentimental score by (-1) to the power of the number of negative words; If there are degree adverbs, multiply the current sentimental word score by the degree level of the degree adverb. Finally, segment the text based on punctuation and calculate the average sentimental score of the text. On this basis, this article constructs an indicator filtering model to perform the following calculations:

- Calculation of indicator attribute eigenvalues: Determine the performance and importance of each indicator in the overall network.
- Key node calculation: Based on the characteristic values of indicators, the SNA algorithm is used to determine the key nodes in the network, namely those individuals or groups that play a core role in information dissemination and opinion formation.

B. Social Network Analysis Module

On the basis of data analysis in the sentimental analysis module of early unexpected events, we draw on complex network theory algorithms to filter all nodes, laying a solid foundation for identifying key nodes in subsequent time series modules.

Betweenness centrality is a key indicator in social network analysis, used to measure the ability of a node to control information flow in the network. Specifically, the centrality of a node in the middle refers to the proportion of all shortest paths passing through that node. The equation for the betweenness centrality is:

$$C = \frac{\sum_{j<k} b_{jk}(n_i)}{b_{jk}} \quad (1)$$

In the equation, b_{jk} is the shortest number of paths between points j and k, and $b_{jk}(n_i)$ is the shortest number of paths between two nodes containing node n_i.

If a node is located on the shortest path between many other nodes, its centrality in the middle is higher, which means that the node plays an important "bridge" or "intermediary" role in the process of information dissemination, that is, a key node. Through betweenness centrality analysis, this article not only quantifies the influence of each communicator, but also reveals which communicators play the most critical role in specific public opinion events, providing scientific basis for public opinion management and intervention.

Closeness centrality is an important indicator reflecting the degree of control of communicators over the network. The equation for closeness centrality is:

$$C_C^{-1} = \sum_{j=1}^{n} d_{ij} \quad (2)$$

In the equation, d_{ij} is the shortcut distance between node i and node j.

The smaller the closeness centrality, the more prominent the disseminator's core position in the network, and the lower their dependence on information, which brings them advantages in the process of information acquisition and dissemination. Specifically, disseminators with lower closeness centrality have a higher reputation for power in the network, which means that their influence is more significant throughout the entire network structure and they are less susceptible to the influence or control of other disseminators. Through in-depth analysis of the distribution of closeness centrality among disseminators in specific network environments, this article analyzed network data and obtained the Top 15 results of proximity centrality, which were presented through tabulation and visualization methods. This process not only reveals which disseminators are located at the core of the network, but also provides intuitive basis for further research on the roles and influence of these disseminators.

C. *Time Series Analysis Module*

By classifying the calculation results of the filtering model, this article can gain a deeper understanding of the dynamic evolution of online public opinion. This article will focus on the effective classification of the following types of nodes:

- Key nodes: nodes that occupy an important position in the network structure and have a significant impact on information flow.
- Key active nodes: nodes that are not only structurally important, but also highly active, and may be public opinion leaders or influential individuals.
- Hidden key nodes: nodes that may appear inconspicuous on the surface but may play a significant role in specific contexts, such as potential bridging characters or suddenly active "dormant" accounts.

In addition, by constructing time series analysis, this article can track the evolution process of public opinion over time, identify key time points and turning points that cause public opinion mutations. This helps decision-makers grasp the optimal timing for intervention and take targeted measures to guide or calm public opinion.

$$Y_t = c + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varphi_p Y_{t-p} \dots + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{3}$$

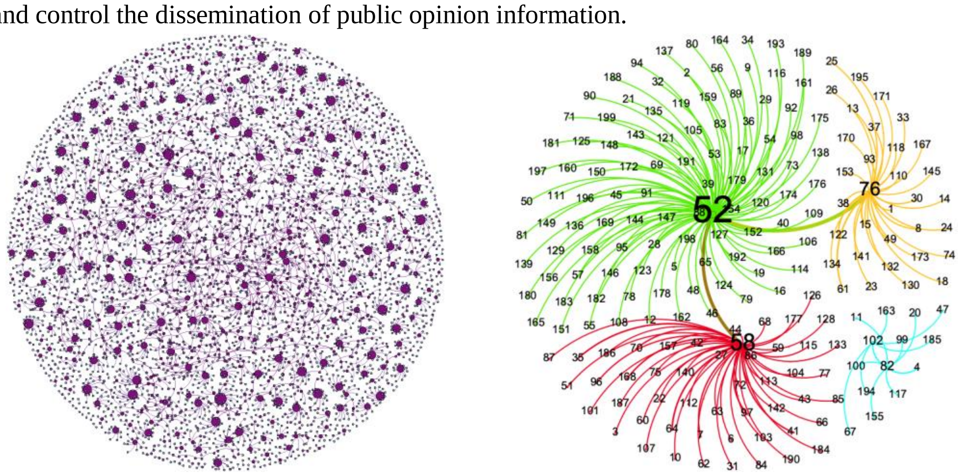
Y_t is the time series data under consideration. φ_1 to φ_p is a parameter of the AR (Autoregressive) model, which is used to describe the relationship between the current value and the past p time points. θ_1 to θ_q is a parameter of the MA (moving average) model, which is used to describe the relationship between the current value and the error at q past time points. ε_t is the error term at time t; c is a constant term.

Through analysis, it was found that the disseminator with the lowest closeness centrality is closer to the core position of online public opinion dissemination, and has the lowest dependence on information from other disseminators, which gives them a clear advantage in obtaining information and enjoys higher power and reputation in the dissemination process. These types of disseminators are usually not easily controlled by other disseminators in the network, and therefore play a crucial role in the flow of information and the formation of opinions. Based on the analysis of closeness centrality, this article can further infer that in the dissemination of online public opinion in sudden events, those disseminators with lower closeness centrality are likely opinion leaders. They can not only quickly obtain key information, but also exert significant influence on other disseminators during the dissemination process.

IV. Results and Discussion

A. *Results and Discussion of Social Network Analysis Module*

Process the collected data and analyze it to obtain the initial node relationship diagram, as shown in Figure 2. There are 1328 nodes and 26445 edges in the Weibo network. According to Figure 2(b), different colors represent different modules, which are communities composed of associated nodes. There is a central node in the middle of each color module, which has a high degree of connectivity and a low degree of connectivity. There are two modules with more associated nodes (i.e. the modules where nodes 52 and 58 are located), while the other module has relatively fewer associated nodes, indicating that the connection status (degree) between nodes has a certain degree of uneven distribution, and there is heterogeneity in the network, that is, scale-free network. In scale-free networks, several key nodes play a dominant role in the network. If the key nodes maintain monitoring, they can intervene and control the dissemination of public opinion information.



(a) Overall network connection (b) Modular distribution structure

Figure 2: Topological Structure Diagram of Network Public Opinion Dissemination Nodes

By performing degree operations on the data, the average degree of the network is 1.98, and the proportional distribution of node degrees is shown in Figure 3.

According to Figure 3, 97.49% of nodes have a degree of 1, indicating that there is only one node connected to it in the network. Therefore, the vast majority of nodes in this Weibo social network have very low degrees, which means the probability of being connected to other nodes is very low. Only a few nodes have high degrees, so these high degree nodes are the core nodes in the network's public opinion dissemination.

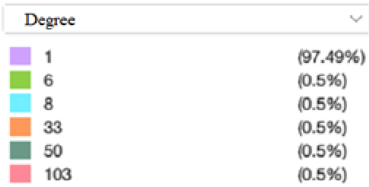


Figure 3: Proportional Distribution of Node Degrees

The measurement results of the centrality of each network node are shown in Table 1 (only the first 15 are shown due to space limitations), which measures the degree of proximity between a node in the network and all other nodes. According to the data shown in Table 1, the proximity centrality of all nodes in the network is greater than or equal to 1. Therefore, it is difficult for nodes in this network to reach other nodes in the first time, and only monitoring these nodes with a proximity centrality greater than 0 can affect the process of information dissemination.

Table 1: Top 15 Measurement Results of Point Centrality

Label	In-degree	Out-degree	degree
52	103	1	104
58	50	1	51
76	32	1	33
82	8	1	9
102	6	1	7
15	0	1	1
61	0	1	1
64	0	1	1
35	0	1	1
168	0	1	1
4	0	1	1
41	0	1	1
70	0	1	1
97	0	1	1
122	0	1	1

Table 2: Closeness Centrality and Betweenness Centrality Data (Partial)

Label	Closeness Centrality	Betweenness Centrality
52	0.72222222	63
58	0.69433962	15083
76	0.65	50
82	0.52571429	7791
102	0.47916667	5360
15	0.43333333	0
61	0.43333333	0
64	0.43333333	0
35	0.43333333	0
168	0.43333333	0
4	0.43333333	0
41	0.41071429	0
70	0.41071429	0
97	0.41071429	0
122	0.41071429	0

The betweenness centrality represents the frequency of nodes appearing on the shortest path in the network. According to Table 2, only 5 nodes have a betweenness centrality greater than 0, indicating that only 5 nodes are on the shortest path of propagation among other nodes in the network, indicating that these nodes play a crucial role and have the greatest impact on information dissemination. Points with high degree and in-degree have low out-degree, indicating that these nodes have high importance or influence in the network in Table 1. And due to more nodes being connected to that node, these may be more susceptible to the influence of other nodes, or have a greater impact on other nodes.

Based on the analysis results of closeness centrality and betweenness centrality, it can be concluded that the top 5 nodes have higher closeness centrality and betweenness centrality, as well as the highest influence and interactivity. Therefore, when guiding and controlling online public opinion, special attention should be paid to these nodes, namely key nodes or hidden key nodes. This type of node plays a decisive guiding role in the entire evolution process of public opinion events. In other words, controlling these key nodes will effectively guide and control the evolution process and future direction of public opinion.

B. Results and Discussion of Time Series Analysis Module

According to the constructed emergency sentiment analysis module, sentiment analysis was conducted on public opinion events, and the distribution of sentiment states obtained is shown in Table 3.

Table 3: Sentiment Attitude Statistics for Review Data

Sentimental Dispositions	Positive	Negative	Neutral
proportion	39.39%	55.26%	5.35%

According to Table 3, 94.65% of the comments have a clear sentimental tendency, with over half (55.26%) of the comments having negative sentiments. This indicates that the comment data has a clear imbalance, which is also in line with the real world situation

On the basis of sentimental state analysis, time slices are performed on a "day" basis, and the missing daily average sentimental values are calculated using segmented cubic Hermite interpolation method to understand the trend of daily average sentimental values over time. This article uses the time series analysis method of Eviews software and ARIMA model for numerical estimation. Subsequently, a stationarity test was performed on the sequence composed of the initial data, and the results obtained are shown in Figure 4. It was observed that the significance result of the ADF (Augmented Dickey Fuller) test was 0.0004, indicating that the sequence does not have a unit root and is a stationary sequence.

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.623850	0.0004
Test critical values:	1% level	-3.555023	
	5% level	-2.915522	
	10% level	-2.595565	

Figure 4: Augmented Dickey Fuller Test of the Sequence

Based on the observation of the autocorrelation coefficient and partial autocorrelation coefficient of the sequence, as well as the comparison of the AIC (Akaike Information Criterion), this paper estimates parameters through various combinations of autoregressive or moving average terms. The comparison shows that ARIMA (1, 0, 1) is the model with the best fitting effect, and its specific parameter values are shown in Figure 5. Among them, R-squared represents the overall fitting degree of the model, with a larger value indicating better fitting effect. AIC is the minimum information criterion, and for this model, a smaller value is better.

Variable	Coefficien...	Std. Error	t-Statistic	Prob.
AR(1)	0.747454	0.087406	8.551535	0.0000
MA(1)	0.467203	0.135151	3.456894	0.0011
SIGMASQ	0.008573	0.001028	8.342213	0.0000

R-squared	0.717438	Mean dependent var	-0.087479
Adjusted R-squared	0.706973	S.D. dependent var	0.175731
S.E. of regression	0.095127	Akaike info criterion	-1.786839
Sum squared resid	0.488650	Schwarz criterion	-1.679310
Log likelihood	53.92490	Hannan-Quinn criter.	-1.745049
Durbin-Watson stat	1.801219		

Figure 5: Parameter Estimation Results of the ARIMA Model

Observations have shown that the AR (1) and MA (1) coefficients of the ARIMA (1, 0, 1) model are significantly lower than the significance level, and the AIC value is less than 0. The D-W (Durbin Watson) value indicates that the model has a significant positive correlation with residual sequences. At the same time, the ARIMA (1, 0, 1) model satisfies that all parameters pass the t-test, the values of feature roots are all outside the unit circle, and there is no autocorrelation in the error sequence, which verifies the rationality of the model. The fitting model of this model is shown in Figure 6, and the estimated equation obtained is:

$$X = 0 + [AR(1) = C(1), MA(1) = C(2), UNCOND, ESTSMPL = "7/14/2022\ 9/08/2022"] \quad (4)$$

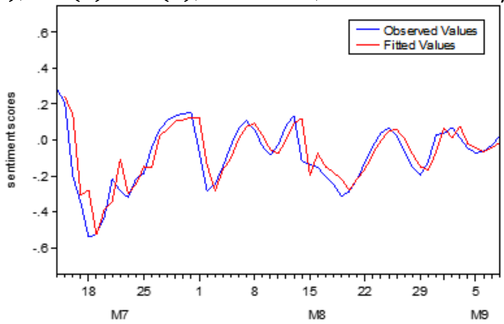


Figure 6: The Fitted Value Curve of the ARIMA Model

As shown in Figure 6, the blue curve represents the actual observation value, the red curve represents the fitted value, and the horizontal axis represents the time series of the observation interval. Among them, fitting values are obtained through static predictions in Eviews' Forecast function. As shown in Figure 6, the data simulation effect based on ARIMA time series model is good and basically reaches a consistent level. After verifying the good fit of the time series, this article analyzes the important turning points of sentiments in the trend graph of the time series by combining the key nodes of public opinion mined through SNA. Looking back at the evolution of the public opinion event, we can analyze the drastic changes in sentiments, as shown in Figure 7.

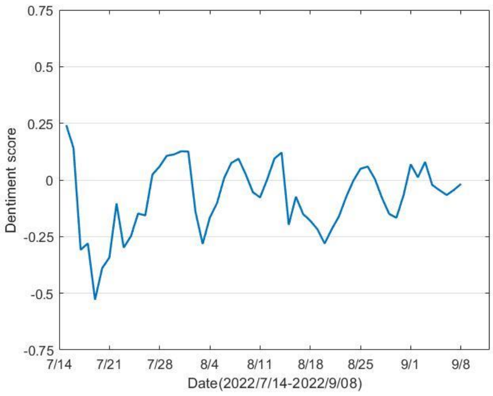


Figure 7: Plot of Trends in Average Daily Sentiment Score

As shown in Figure 7, it can be observed that when the parties involved in the event raise doubts about the product quality and public relations content of Company A or release the latest updates, such as the after-sales records uploaded by consumers on social media platforms with Company A's customer service, the video circulating online that the General Manager of Company A claimed that the buyer's cutting method was incorrect, and the subsequent reports on the case by news media, all have a negative impact on the sentimental state of netizens, This leads to significant fluctuations in daily sentimental values and a rapid decline. By analyzing the intensity of sentimental polarity and the trend of changes in the number of comments, the evolution of public opinion events is divided into multiple stages based on the life cycle theory. Its specific manifestations are as follows:

(1) On July 12, 2022, the entry related to "A company's customer service claiming that the kitchen knife cannot touch garlic" became a hot search, marking the beginning of public opinion on the incident. Subsequently, after continuous coverage by multiple media outlets and extensive discussions among netizens about the incident, it entered a period of public opinion outbreak on July 16th. With the outbreak of subsequent public opinion hot topics such as the "Haitian soy sauce double standard incident", the incident has entered a long tail period of public opinion, and public sentiment is still generally negative. In the initial stage of public opinion, the comments of online users on events tend to be rational, diluting the impact of negative sentiments on the overall sentimental state. With the continuous fermentation of public opinion, online users have a strong sense of role involvement in cases involving consumer rights, and their sentimental inclination towards the event has been in a negative state, leading to a continuous decline in daily sentimental values.

(2) With the continuous interpretation and reporting of Weibo's big V (key node) and online and offline media, the focus of public opinion events has shifted from "kitchen knife quality" to "corporate attitude", and the brand

reputation of Company A has shown a continuous decline. On July 15th, the official response of Company A to the incident did not resolve the negative sentiments of netizens, but instead intensified their dissatisfaction and questioning, leading to the secondary spread of public opinion. In addition, on July 18th, a video of an interview with the general manager of Company A claiming that the buyer's chopping method was incorrect was leaked, leading to a large accumulation of negative sentiments among netizens. The hot search for related topics lasted for 18.4 hours, causing public opinion to enter a secondary outbreak stage and forming fluctuations in public opinion. The formation of this stage is mainly due to the lack of official responses and inappropriate wording from the company. A company's attitude of requiring consumers to adapt to the product has led to the abdication of discourse dominance and unfavorable interaction between online and offline media, resulting in an increasing dissatisfaction and negative state of netizens towards it.

(3) After the event entered a period of public opinion decline, the daily average sentimental value gradually increased, and online comment texts returned to rational sentiments. The attention of online users to cases is gradually decreasing, and this stage lasts for a long time with some recurrence. The main reason for sentimental fluctuations is a decrease in the number of comments, an increase in the proportion of long-term followers of the event, and an increase in the impact of extreme comment texts on the overall sentimental tendency of comments.

C. Analysis of Response Strategies for Network Public Opinion in Emergencies

(1) In the initial stage of public opinion, it is necessary to quickly adopt public opinion response strategies. From July 12 to July 14, 2022, the number of comments posted by netizens increased slowly, and the overall intensity of sentimental polarity was in a positive state. At this time, the comprehensive attention to public opinion events is low and there are fewer negative sentiments. A company should actively release clarifications on public opinion events in mainstream media, guide the stable development of public opinion with a positive response and active voice, seize the dominant position of discourse power, and minimize the possibility of public opinion outbreaks.

(2) Entering the period of public opinion outbreak, it is necessary to respond to public opinion in a targeted manner. From July 15th to July 18th, the intensity of sentimental polarity among netizens showed a rapid decline trend, and reached the lowest point of sentimental value on July 18th. The reason for this phenomenon is that the video of the General Manager of Company a questioning the way buyers cut vegetables has leaked, leading to further intensification of public opinion. Therefore, the public relations department of Company A should first sort out the topics of concern for netizens. The focus of attention on this public opinion event has shifted to the attitude of the company at this stage. Therefore, it is necessary to be sincere when issuing statements and respond to doubts with more reasonable solutions from the perspective of consumers. On this basis, attention should also be paid to downplaying the malignant factors in this incident, shifting the attention of netizens reasonably, in order to accurately and timely resolve the contradictions in public opinion, and accelerate the decline of public opinion.

(3) As public opinion fluctuates and falls, public opinion response strategies should focus on being comprehensive. Starting from August 4, 2022, due to the lack of follow-up hot information on the event, the sentimental value gradually increased and the number of comments correspondingly decreased. At this stage, the public relations department of Company A should continue to monitor online comment texts, closely monitor the interpretation and comments of events by Weibo influencers and key nodes, and focus on reshaping the company's reputation with mainstream online and offline media to restore consumer trust.

V. CONCLUSIONS

This article deeply analyzes the evolution laws of online public opinion, and comprehensively applies various research methods such as social network analysis, time series analysis, and sentiment analysis to draw a series of conclusions with profound theoretical value and practical guidance significance. The research results indicate that the process of public opinion dissemination exhibits distinct phased characteristics, and each stage is comprehensively influenced by different factors, providing important theoretical support and practical guidance for public opinion management. Meanwhile, the precise identification and effective utilization of key nodes have become a new practical tool for public opinion management, especially in the areas of opinion leaders and mainstream media, whose influence has a significant impact on the direction of public opinion and cannot be ignored.

This study innovatively introduces time series analysis, enriching the methodological system of online public opinion research, making the research more comprehensive and in-depth. The comprehensive evaluation of sentimental tendencies and social network attributes not only deepens the understanding of key nodes, but also provides strong support for public opinion prediction and trend grasping. In addition, this study also proposes a construction plan for a real-time network public opinion monitoring system and intervention strategies for key nodes, providing a practical and feasible path for rapid response and effective management of public opinion crises.

This study has not only made significant progress in the field of online public opinion analysis, but also provided important scientific basis for social management and public opinion guidance. Its rigorous research methods and in-depth analysis not only contribute to promoting theoretical development in related fields, but also provide strong guidance for practical applications. At the same time, the results of this study have important reference value for government or enterprise decision-making, helping to improve the efficiency and effectiveness of public opinion management, and thereby promoting social harmony, stability, and development progress.

ACKNOWLEDGMENT

The authors would like to express their gratitude to the editors and anonymous reviewers for their comments and suggestions. Jia-Qi Mo and Ziqi Wang contributed equally to this work and are co-first authors. This research was funded by the Hangzhou City University Student Innovation Training Program Project (grant No. 202313021040) and the National Natural Science Foundation of China (72204212).

REFERENCES

[1] Y. Yuan, N. Pang, Y. Zhang, et al. Which cascade is more decisive in rumor detection on social media: Based on comparison between repost and reply sequences. Knowledge-Based Systems, vol. 278, pp. 110857, 2023.

[2] G. Wang, Y. Wang, K. Liu, et al. Multidimensional influencing factors of public opinion information dissemination in social media: Evidence from Weibo dataset. International Journal of Modern Physics B, vol. 33, no. 31, pp. 1950375, 2020.

[3] Y. Luo, J. Ma and C.K. Yeo. BCMM: A novel post-based augmentation representation for early rumour detection on social media. Pattern Recognition, vol. 113, Article No. 107818, 2021.

[4] Z. Hao et al., "A Novel Public Sentiment Analysis Method Based on an Isomerism Learning Model via Multiphase Processing," in IEEE Transactions on Neural Networks and Learning Systems, doi: 10.1109/TNNLS.2023.3274912. in press.

[5] R. Arya, J. Singh and A. Kumar. A survey of multidisciplinary domains contributing to affective computing. Computer Science Review, vol. 40, Article No. 100399, 2021.

[6] H. Luo, X. Meng, Y. Zhao and M. Cai. Exploring the impact of sentiment on multi-dimensional information dissemination using COVID-19 data in China. Computers in Human Behavior, vol. 144, Article No. 107733, 2023.

[7] A. Nazir, Y. Rao, L. Wu and L. Sun, "Issues and Challenges of Aspect-based Sentiment Analysis: A Comprehensive Survey," in IEEE Transactions on Affective Computing, vol. 13, no. 2, pp. 845-863, 1 April-June 2022.

[8] Z. Wu, G. Cao and W. Mo, "Multi-Tasking for Aspect-Based Sentiment Analysis via Constructing Auxiliary Self-Supervision ACOP Task," in IEEE Access, vol. 11, pp. 82924-82932, 2023.

[9] C. Messaoudi, Z. Guessoum and L. Ben Romdhane, Opinion mining in online social media: a survey. Social Network Analysis and Mining. vol. 12, no. 1, pp. 25, 2022.

[10] C. Wang, Z. Xiao, Y. Liu, Y. Xu, A. Zhou and K. Zhang, "SentiView: Sentiment Analysis and Visualization for Internet Popular Topics," in IEEE Transactions on Human-Machine Systems, vol. 43, no. 6, pp. 620-630, Nov. 2013.

[11] Y. Cai, K. Yang and D. Huang, et al. A hybrid model for opinion mining based on domain sentiment dictionary. International Journal of Machine Learning and Cybernetics, vol. 10, no. 8, pp. 2131-2142, 2019.

[12] F. Hui. Research on the Construction of Emergency Network Public Opinion Emotional Dictionary Based on Emotional Feature Extraction Algorithm. Frontiers in Psychology, vol 13, article no. 857769, 2022.

[13] F. Kokoszka, Baptiste Le Roux, Daniele Iudicone, et al. Long-term variability of the coastal ocean stratification in the Gulf of Naples: Two decades of monitoring the marine ecosystem at the LTER–MC site, between land and open Mediterranean Sea. in Marine Ecology-an Evolutionary Perspective, vol. 44, no. 3, pp. e12725, 2023.

[14] A. Blázquez-García, A. Conde and U. Mori, et al. A Review on Outlier/Anomaly Detection in Time Series Data. ACM Computing Surveys, vol. 54, no.3, pp. 56, 2022.

[15] Z. Liu, Z. Li, W. Li, et al. Deep graph tensor learning for temporal link prediction. Information Sciences, vol. 660, pp. 120085, 2024.

[16] Z. Weng, "Application Analysis of Emotional Learning Model Based on Improved Text in Campus Review and Student Public Opinion Management," in Mathematical Problems in Engineering, vol. 2022, Article No. 5135200, Apr 2022.

[17] R. Kobayashi, Y. Takedomi, Y. Nakayama, et al. Evolution of Public Opinion on COVID-19 Vaccination in Japan: Large-Scale Twitter Data Analysis. Journal of Medical Internet Research, vol. 24, no. 12, pp. e41928, 2023.

[18] L. Li, Y. Wu, Y. Zhang and T. Zhao, "Time+User Dual Attention Based Sentiment Prediction for Multiple Social Network Texts with Time Series," in IEEE Access, vol. 7, pp. 17644-17653, 2019.

[19] W. Hilal, S. A. Gadsden and J. Yawney. Financial Fraud: A Review of Anomaly Detection Techniques and Recent Advances. Expert Systems with Applications, vol. 193, pp. 116429, 2022.

[20] X. Feng, X. Wang, Y. Zhang, “Research on public emotional polarization and public opinion evolution of OTC and learning during the COVID-19 epidemic: taking the topic of OTC on Zhihu as an example,” in Library Hi Tech, vol. 40, no. 2, pp. 286-303, 2022.

[21] Q. Su, S. Yan, L. Wu and Xiangyan Zeng, “Online public opinion prediction based on a novel seasonal grey decomposition and ensemble model,” in Expert Systems with Applications, Vol. 210, pp. 30, Dec. 2022.

[22] X. Liu, L. Zheng, X. Jia, H. Qi, S. Yu and X. Wang, "Public Opinion Analysis on Novel Coronavirus Pneumonia and Interaction With Event Evolution in Real World," in IEEE Transactions on Computational Social Systems, vol. 8, no. 4, pp. 1042-1051, Aug. 2021.

[23] X. An, Zhiming Yin, Qi Tong, et al. An integrated resilience assessment methodology for emergency response systems based on multi-stage STAMP and dynamic Bayesian networks. Reliability Engineering & System Safety, vol. 238, pp. 109445, 2023.

[24] J. Wei, Y. Jia, Y. Zhang, et al. The Public Opinion Evolution under Group Interaction in Different Information Features, in Complexity, vol. 2022, article no. 1016692, May 2022.

[25] R. Riquelme, J.A. Vera. A parameterizable influence spread-based centrality measure for influential users detection in social networks. Knowledge-Based Systems, vol. 257, pp. 109922, 2023.

[26] A. Rehman, A. Jiang, A. Rehman, et al. Identification and role of opinion leaders in information diffusion for online discussion network. Journal of Ambient Intelligence and Humanized Computing, vol .14, pp. 15301–15313, 2023.

[27] S. Kumar and B.S. Panda. Identifying influential nodes in Social Networks: Neighborhood Coreness based voting approach. Physica A-statistical Mechanics and Its Applications, vol. 553, article. No. 124215, Sep. 2020.

[28] F. Meng, Xixi Xiao, Jingdong Wang. Rating the Crisis of Online Public Opinion Using a Multi-Level Index System. International Arab Journal of Information Technology, vol. 19, no. 4, pp. 597-608, 2022.

[29] L. Hou. Network versus content: The effectiveness in identifying opinion leaders in an online social network with empirical evaluation. Physica A: Statistical Mechanics and its Applications, vol. 592, pp. 126879, 2022.

[30] M. Zhan, G. Kou, Y. Dong, F. Chiclana and E. Herrera-Viedma, "Bounded Confidence Evolution of Opinions and Actions in Social Networks," in IEEE Transactions on Cybernetics, vol. 52, no. 7, pp. 7017-7028, July 2022.