

# Semantic Fusion and Propagation Model for Internet Public Opinion Data in Big Data Environment

Pengju Wang<sup>a,\*</sup>, Huifeng Xue<sup>a</sup>, Zhe Yu<sup>b</sup>, and Feng Zhang<sup>c</sup>

<sup>a</sup>*School of Automation, Northwestern Polytechnical University, Xi'an, 710072, China*

<sup>b</sup>*Information Comprehensive Office, Yan'an Municipal Committee, Yan'an, 719000, China*

<sup>c</sup>*School of Information Engineering, Yulin University, Yulin, 719000, China*

---

## Abstract

In order to improve the monitoring and early warning efficiency of network public opinion, and to reveal the spread of network public opinion, an evolution process model of public opinion in the Internet is proposed. It tries to analyze the evolution mechanism of network public opinion in the Internet, and provides certain theoretical support and method guidance for network public opinion monitoring and prediction. At the semantic level, the implementation method of knowledge fusion of different levels of public opinion information resources is proposed, which is supported by semantic technologies such as the semantic web in the big data environment. Then the multi-agent modeling and simulation method is used to establish the network public opinion information communication simulation model. The attributes of each participating entity were constructed in the model, and the influence of various factors on the network crisis information transmission was analyzed. The experimental results show that the proposed verification simulation model has high credibility. By analyzing the number of "opinion leaders" in the model, the participation of netizens, the credibility of the government, the speed of information disclosure and the transparency of the public, it can improve the monitoring and early warning efficiency of network public opinion, and better reveal the propagation law of network public opinion.

**Keywords:** internet public opinion; semantic fusion; propagation model; information dissemination; complex social networks

(Submitted on May 22, 2019; Revised on October 12, 2019; Accepted on November 22, 2019)

© 2019 Totem Publisher, Inc. All rights reserved.

---

## 1. Introduction

Internet public opinion is the product of the development of the Internet to a certain stage. It is the mapping of social public opinion in the Internet space [1]. With the diversification of complex social systems and the rapid development of information technology, the way of information dissemination has changed greatly; it is different from the traditional model of interpersonal relationship and mass media communication. It not only has a greater concealment and diversity of information dissemination on the Internet, but also the new social relationships and mutual influences formed by netizens in the network. It is also more complex and diverse than the real world. Therefore, once the lyrics are formed and focused, it will have greater influence [2].

The elements of online public opinion communication generally include five elements: lyric subject, lyric body, lyric media, lyric space, and lyric object [3-4]. In a big data environment, a single information object cannot reflect knowledge. The knowledge acquired by users should be refined and validated high-quality knowledge objects. They are often not ready-made, single-source data. It can be different from a perspective to observe, compare, and understand the multi-source knowledge of the same problem [5]. We can be real-time, accurate, intelligent analysis of multi-source, cross-cutting, large-scale information by extracting resources and knowledge integration. We can process and utilize information from the perspective of machine understanding, then the knowledge analysis and processing technologies based on the semantic web, knowledge ontology, data mining, and knowledge discovery are realized through the conversion and integration of knowledge. We can obtain new knowledge or new solutions to achieve knowledge innovation. In the big data environment, facing the challenge of knowledge sharing and knowledge service of network public opinion brought by a large number of

---

\* Corresponding author.

E-mail address: 416848742@qq.com

multi-source data [6], as well as the dual needs of knowledge management and knowledge innovation of network public opinion, through the processing of big data through knowledge fusion, the dissemination mode and mechanism of network public opinion can be analyzed, which is conducive to the realization of precise knowledge services and knowledge value-added functions, and improve the monitoring level of network public opinion in large data environment. Therefore, in the field of network public opinion, there is an urgent need for the dissemination mechanism of network public opinion and the practical application of integration methods.

In this paper, big data semantic analysis method, social network analysis method and complex system simulation method is used to comprehensively study the propagation law of network crisis information and the semantic association model of public opinion data, and further reveal the influence factors of various influencing factors in the information dissemination of network public opinion crisis, and thus Enterprises and governments provide scientific evidence for making decisions about the crisis of online public opinion. By rationally abstracting the behavior characteristics, behavioral motives and network relationships of network users, the network public opinion evolution model in complex social networks is constructed. From the network relationship between the government, the media and the netizens, we have conducted in-depth research on the mechanism of the dynamic evolution for online public opinion.

## 2. Network Public Opinion Evolution Modeling Method

### 2.1. Complex Social Network Theory

A large number of complex systems in nature and the human world can be represented in the form of networks. According to the difference of attribute characteristics, the network can be divided into different types, such as biological network, technology network, information network, social network, etc. [7-8]. From the perspective of graph theory, a specific network can be abstracted into a two-group  $G=(V, E)$ , where the set  $V=\{v_1, v_2, \dots, v_N\}$  is called a point set, and set  $E=\{e_1, e_2, \dots, e_M\}$  is called edge set. Each edge in set  $E$  has a point pair in set  $V$  corresponding to it. In complex network research, the distance  $d_{ij}$  between node  $i$  and  $j$  is usually defined as the number of edges on the shortest path connecting the two nodes, the average path length  $L$  of the network is the average distance between any two nodes. In social networks, it represents the number of individuals in the shortest chain of relationships between two individuals, and its calculation method is shown in Equation (1).

$$L = \frac{1}{N^2} \sum_{j=1}^N \sum_{i=1}^N d_{ij} \quad (1)$$

$N$  is the total number for the nodes in the network. The agglomeration coefficient is to reveal the agglomeration characteristics of the network. The so-called agglomeration characteristics mean that two nodes in the network that are connected to a node are likely to be connected to each other [9]. Suppose node  $i$  in the network has  $k_i$  edges to connect it to other nodes, these  $k_i$  nodes are the neighbor nodes of node  $i$ . After a simple calculation, it can be concluded that there are up to  $k_i(k_i-1)/2$  edges between the neighbor nodes of node  $i$ . If the number of edges actually existing between the neighbor nodes of node  $i$  is  $E_i$ , then the aggregation coefficient of node  $i$  can be expressed as:

$$C_i = \frac{2E_i}{k_i(k_i-1)} \quad (2)$$

The aggregation coefficient of the entire network is the mean of the aggregation coefficients of all nodes, which can be written as:

$$C = \frac{1}{N} \sum_{i=1}^N C_i \quad (3)$$

The value of  $C$  is between 0 and 1. When  $C=0$ , all nodes in the network are isolated nodes, there are no connected edges; when  $C=1$ , the network is fully coupled, and any two nodes are directly connected.

## 2.2. Public Opinion Propagation Model in Complex Networks

The discovery of complex network characteristics has injected new vitality into the study of public opinion communication. Whether it is a regular network or a random network, there is a huge gap between our social network and the reality. These two networks are fundamentally incapable of being close to the high similarity of the real network. In this sense, it is obviously difficult to use these two network models to simulate the reality network and to study the public opinion communication. However, some features of complex networks are more tailored to the theory of public opinion communication [10]. It is easy to understand that complex network theory itself is derived from the real social network environment, especially the interpersonal network in society. In the process of studying public opinion communication, an inevitable problem is modeling. The small world network and the scale-free network reflect the decisive role of the network structure to the network attributes from one aspect. The great role brought by this structure also continues to exist and inherit in the public opinion communication [11-12]. As shown in Figure 1, in the evolution of network public opinion, the autonomous network behavior main body mainly includes grassroots netizens, opinion leaders and network communication media, which together constitute the main force in the Internet space to promote the development of network public opinion.

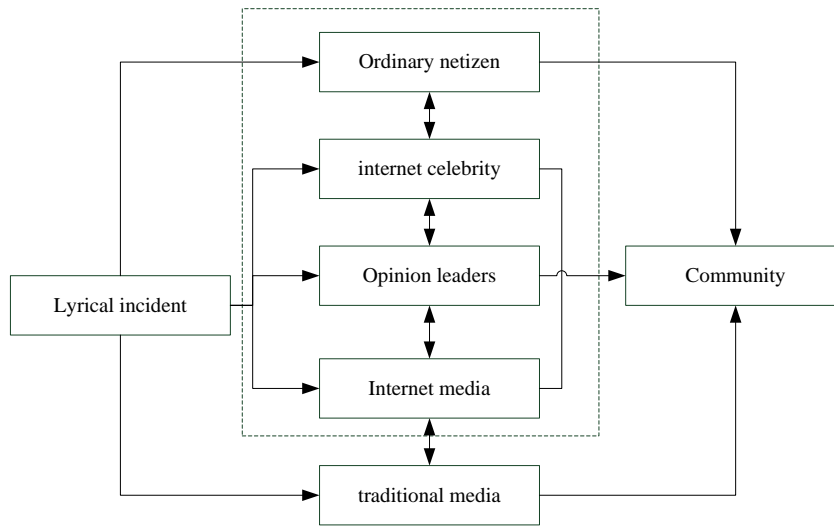


Figure 1. The evolution of network public opinion

In the process of network public opinion evolution, the synergy between different actors has promoted the development and evolution of network public opinion. Therefore, the multi-agent interaction model in the evolution of network public opinion is essentially a self-organizing collaborative model [13]. In the description of the Sznajd model, the opinions of one person in the group are influenced by others. The larger the size of the group, the greater the impact. It is the ratio of the difficulty of one's own opinions and the number of people who influence his opinions. Sznajd model is a two-dimensional discrete viewpoint value exchange model on a one-dimensional straight line [14-15]. The design is described as follows: Each node occupies a position in a one-dimensional data chain and has a binary view, where 1 means that the node agrees with the matter, and -1 means that the node opposes the matter. Two neighbors, Node  $i$  and Node  $(i + 1)$ , determine the opinions of their neighbors  $(i - 1, i + 2)$ . The specific evolution rules are as follows:

$$\begin{cases} \text{if } N_i = N_{i+1}, \text{ then } N_{i-1} = N_i = N_{i+1} = N_{i+2} \\ \text{if } N_i \neq N_{i+1}, \text{ then } N_{i-1} = N_{i+1}; N_i = N_{i+2} \end{cases} \quad (4)$$

If  $N_i$  and  $N_{i+1}$  have the same opinion, the opinions that affect their neighbors are evolved by Equation (4). Conversely, if the two nodes have different opinions, only the neighbors of the other node are affected.

The Sznajd model is a typical discrete opinion model, but the opinion model like this type is not very applicable, because not all of its views are described in terms of 0 and 1 representations [16]. The attitude of the wall in the middle cannot be accurately and properly described. Therefore, a continuous opinion model emerges, and the node attitude values in this model usually take continuous values in a value range [17]. Assuming that  $M$  nodes have random opinion values, they can take any possible rational numbers in  $(-1, +1)$ , which causes the initial opinions of all nodes to be different, and the picture of the evolution process is also extremely complicated, and the result of evolution. It is the end of some opinion clusters.

In the Deffuant model, the node is given a value of  $x_a$  between 0 and 1, that is  $0 \leq x_a \leq 1$ . At a certain time  $t$ , if node  $a$  and node  $b$  are two nodes that need to be balanced, then their opinion values are  $x_a(t), x_b(t)$ ; If the absolute value of the difference in  $x_a(t), x_b(t)$  exceeds a certain threshold  $A$ , then there is no communication between them. This situation is also easy to understand on the real network. If the opinions of two people, A and B, are too big to reach a certain threshold, then they are not willing to communicate, that is, they will not affect each other's views, that is, they do nothing with each other. If and only if the difference is less than  $a$ , then:

$$\begin{cases} x_a(t+1) = x_a(t) - \beta |x_b(t) - x_a(t)| \\ x_b(t+1) = x_b(t) - \beta |x_a(t) - x_b(t)| \end{cases} \quad (5)$$

Where the parameter  $\beta$  is the convergence parameter, and  $\beta \in [0, 1/2]$ . An important attribute of Deffuanat model is that no matter what the network structure is, the factors can form a consensus. However, if the value of  $\beta$  is small, its many clusters will appear in the system.

### 3. Construction of Network Public Opinion Communication Simulation Model

#### 3.1. Subject Classification and Attribute Description

According to the interaction of subjects in the network public opinion information dissemination, this study sets the attributes of five categories of subjects: Internet users, government, network media, self-media and network environment. The main subject of the network environment is to record the changes in the number of subjects and the value of the attributes in the process of public opinion information dissemination. The properties of various types of subjects are shown in Table 1.

Table 1. Subject attribute

| Subject   | Internet users                                                                            | Self-media                                                                                   | Network media                                                     | Network environment                                                                                                                                         | Government                                                                                                                                |
|-----------|-------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|-------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|
| Attribute | 1. Attitude<br>2. Influence<br>3. Credibility<br>4. From the crowd<br>5. Desire to spread | 1. Importance<br>2. Ambiguity<br>3. Timeliness of public opinion<br>4. Propagation Intensity | 1. Attitude<br>2. Authoritative<br>3. Credibility<br>4. Influence | 1. Number of subjects<br>2. Changes in the status of each subject<br>3. Changes in the credibility of each subject<br>4. Authoritative changes in the media | 1. Attitude<br>2. Credibility<br>3. Information disclosure speed<br>4. Information disclosure transparency<br>5. Supervision and guidance |

In order to facilitate the research, this paper uses the form of function to express the attributes of the netizens, in the form of the following:

$$N[Q_i(t), W_i(t), T_i(t), U_i(t), R_i(t)] \quad (6)$$

In Equation (6), from left to right, the function values of the five attributes of the state, influence, credibility, heredity and propagating desire of the netizen at time  $t$  are sequentially represented.

(1) Attitude: In the process of network public opinion information dissemination, netizens share three attitudes of spread, receive, and quit.  $Q_i(t)=1$  indicates that netizens  $i$  have a propagating attitude at  $t$ ;  $Q_i(t)=0$  indicates that the netizen  $i$  pays attention to the crisis event at time  $t$ , and is in the receiving state;  $Q$  indicates that the netizen  $i$  is not concerned about the crisis event at the time  $t$ , and is in the exit state.  $Q_i(t)=-1$  indicates that the netizen  $i$  is not concerned about the crisis event at the time  $t$ , and is in the exit state.

(2) Influence:  $W_i(t)$  is refers to the strength of Internet users' communication ability at  $t$  time. In this model, netizens are divided into three levels: small, medium and large according to the size of influence, when the value of  $W_i(t)$  is between  $[0, 0.4]$ , it means that the influence of netizen  $i$  is small; when the value of  $W_i(t)$  is between  $[0.5, 0.8]$ , it means that the influence of netizen  $i$  is medium; when the value of  $W_i(t)$  is between  $[0.8, 1]$ , it means that the influence of netizen  $i$  is large. Different types of netizens account for a different proportion, and netizens with the least influence account for the largest proportion, followed by the largest and the smallest, with the proportions of  $P(s) = 0.7$ ,  $P(m) = 0.2$ ,  $P(b) = 0.12$ .

In the simulation model, the initial energy level of the subject is determined by the level of its influence, and the energy of the subject decreases with time. Use the formula  $Ne_i(t) = Ne_i(t_0) - (t - t_0)$  to represent the energy value of the netizen  $i$  at time  $t$ , where  $Ne_i(t_0)$  is the initial value, and  $T$  is the moment when the netizen begins to participate in the network public opinion event.

(3) Credibility:  $T_i(t)$  indicates the credibility of the crisis information transmitted by Internet users at time  $t$ , that is, how many other netizens believe the information after seeing the information, and its value is randomly selected in consecutive intervals  $[0, 1]$ ,  $T_i(t)=0$  indicates that the netizen completely distrusts the information he spreads. After seeing it, he will not spread it. The probability of withdrawing attention is very high;  $T_i(t)=1$  indicates that netizens fully trust and believe that the information is true and there is a high probability of spreading after seeing it.

(4) Conformity:  $U_i(t)$  denotes the conformity of Internet democracy at  $t$ -time. Its value is selected randomly in the continuous interval  $[0, 1]$ . 0 denotes that the status of Internet users will not change because of the influence of other netizens. 1 denotes that Internet users have no opinions, and their status will change because of the influence of other netizens.

(5) Spread desires: Because different netizens have different personalities, their preference for information is different, and the desire of netizens to spread will decline with time.  $R_i(t)$  is used to express the desire of Internet users at  $t$  time. Its initial value is randomly selected in the continuous interval  $[0, 1]$ . 0 means that the netizens have no desire to spread the crisis information. It does not spread after receiving the information. The desire to spread is strong, and it will spread after accepting the crisis information. The formula  $R_i(t) = r_i - a(t - t_0)$  indicates that the desire of the netizens to communicate is a change with time, where  $R_i(t) = r_i - a(t - t_0)$  is the initial test value of the desire of the netizens to spread,  $a$  is a constant, and the value of  $a$  can be adjusted according to the actual situation.

### 3.2. Semantic Relevance Model of Public Opinion Communication

Semantic relevance analysis of public opinion dissemination is mainly used to identify the relevance of different resources with different order. According to the links between public opinion, publishing time and location, we can find the order of information dissemination among them and trace the source of information [18].

We can record the public opinion of the lyric information resource as  $doc_i (i \in [1, n])$ , record the time parameter text  $doc_i$  as  $time_i (i \in [1, n])$ .  $DocTim = \{doct_1 = (doc_1, time_1), doct_2 = (doc_2, time_2), \dots, doct_n = (doc_n, time_n)\}$  is the timing.

If the timing is  $DocTim = \{doct_1 = (doc_1, time_1), doct_2 = (doc_2, time_2), \dots, doct_n = (doc_n, time_n)\}$ , the timing set is recorded as  $L = \{DocTim_1, DocTim_2, \dots, DocTim_m\}$ . For any timing  $DocTim_i \in L$  and  $DocTim_j \in L$  satisfy  $DocTim_j \cap DocTim_i = \Phi$ , then the timings  $DocTim_i$  and  $DocTim_j$  are time series associations. If the timing  $DocTim_j$  occurs after the timing  $DocTim_i$  time interval  $\Delta t$ , then the timing relationship between the two is:

$$DocTim_i \xrightarrow{\Delta t} DocTim_j, \forall i, j \in [1, n] \quad (7)$$

We refer to the sensation information resource set  $I = \{r_1, r_2, \dots, r_n\}$  as a sequence of resources, so the resource sequence  $S$  is composed of certain web page texts according to the corresponding chronological sequence. Therefore  $S$  can be written as the following expression:  $S = \{r_{i_1}, r_{i_2}, \dots, r_{i_m}\}$ ,  $r_{i_m} \in I$ ,  $m \in [1, k]$ .

For a time series library, the sequence support of a time series  $S$  refers to the proportion of all information sequences containing  $S$ .

$$S(doc) = \frac{|DocTim_i|}{|DocTim|}, S \subset DocTim_i \quad (8)$$

Where  $|DocTim_i|$  represents the number of information sequences of the government website information resource library  $DocTim$ .

Causal association is a manifestation of the connection and interaction of the universal existence of objective things. If two events have a causal relationship, where event A is "cause" and event B is "fruit", it indicates that the occurrence of event A necessarily leads to the occurrence of event B [19]. The target resource relationship in which  $[A_i] \rightarrow [S_i]$  exists in the form of relationship is called causal association, where the variables  $[A_i]$  and  $[S_i]$  are the expressions of the target resource values of the cause and the result state, respectively, the relationship symbol  $\rightarrow$  represents a causal relationship [20]. The causal association structure of the target resources in the government website information resource classification knowledge organization architecture is shown in Figure 2.

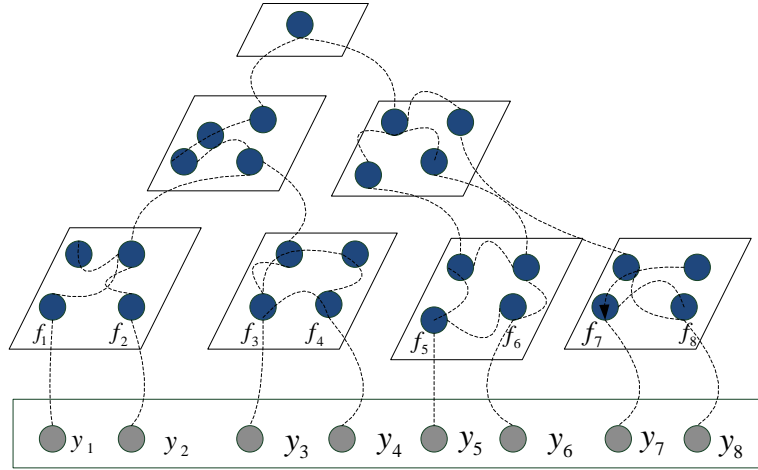


Figure 2. Causal correlation structure and causal correlation function graph of target resources

For the causal connection problem for the target resource can be expressed as a set of quaternary variables:  $FD = \{D, M, C, R\}$ , where the variable  $D = \{d_1, d_2, \dots, d_n\}$  represents a finite and non-empty target resource "cause" set,  $M = \{m_1, m_2, \dots, m_n\}$  is a finite and non-empty target resource "joined" set, which represents all associated attribute values of possible target resources,  $C \subseteq D \times M$  represents the mutual causal connection between the "cause" set and the "joined" set, and  $R$  represents the set of joins determined to occur in the join set  $M = \{m_1, m_2, \dots, m_n\}$ .

The target resource connection set is  $Y = \{y_1, y_2, \dots, y_n\}$ , and the target resource cause set is  $F = \{f_1, f_2, \dots, f_n\}$ , then the target resource causal association "factor function" is:

$$object(m_j) = \{d_i \mid (d_i, m_j) \in C\} \quad (9)$$

The corresponding target resource causal association "fruit function" is:

$$result(d_i) = \{m_i \mid (d_i, m_j) \in C\} \quad (10)$$

#### 4. Experimental Analysis

This paper studies the evolution of network public opinion information over time when each subject interacts with each other, so this paper chose NetLogo simulation platform. The virtual world of NetLogo is composed of subjects. The subject is the individual who executes the instructions, and the behavior of all subjects occurs in parallel. There are three main types of entities in NetLogo, namely turtles, patches, and observers. Sea turtles refer to the main body that can move in the world. The sea turtles involved in this article mainly include netizens, media and government, and patches means that the world is a two-dimensional grid of tiles; each tile occupies a small rectangular block on which the turtle can move. Tiles also have their own properties and behavior, but patches cannot move. The observer is a global subject that observes the virtual world of turtles and tiles and is able to execute commands to control the world.

In this paper, the dissemination data of network public opinion information in the case of "Yan'an City Management Violence Law Enforcement" in Shaanxi Province is taken as sample data, and compared with the simulation data provided by NetLogo simulation platform to verify the credibility of the simulation model. In the simulation model of interaction between netizens, three simulation algorithms are mainly included. In the advancement of the simulation process, the three algorithms are executed cyclically and in parallel. The algorithms are mainly included in the three processes of communicator and new participation netizens, communicators and recipients, and netizens' own state changes.

(1) Parameter setting: According to the size of influence, netizens are divided into three intervals:  $[1, 10)$ ,  $[10, 40)$ ,  $[40, 50]$ . According to the interval, "receiver", "total" and "to" are used in the simulation model. The "quit" is represented by three different shapes, the shape size is 0.2, 0.6, 1, respectively, and the initial influence of the propagation source is set to 2. The initial timeliness, importance and initial ambiguity of the network public opinion information are set to 0.6, 0.66, and 0.71, respectively; in the calculation of the intensity of public opinion information, the degree of ambiguity of information and the timeliness of information account for 0.5 and 0.3, respectively.

(2) Simulation process: It is clearly evident from the Figure 3 that our proposed in red indicates the communicator, green is the recipient, and blue indicates the exit of the follower. The left view shows that the crisis information has just appeared on the network. From left to right, the communicator enables more netizens to participate in the crisis event. In the simulation model, process 1 is first run to determine whether the new participant is a communicator or a recipient. Figure 3 is the early stage of online public opinion communication. Then the communicator interacts with the recipient, running process 2. Figure 4 is the peak period of Internet public opinion dissemination. At the same time, the netizens themselves run the process 3, judging their own state changes. Figure 5 is the mitigation period of Internet public opinion dissemination.

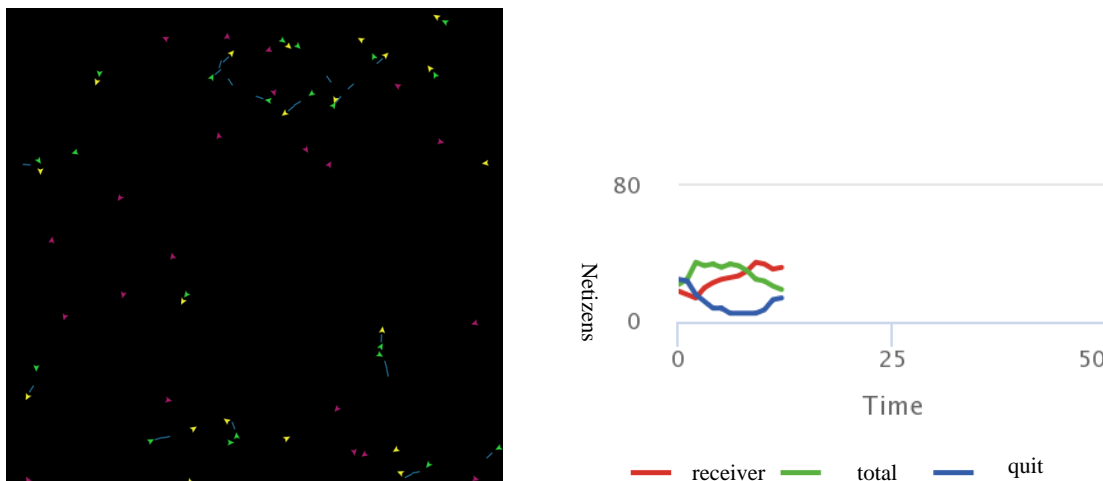


Figure 3. Early stage of online public opinion communication

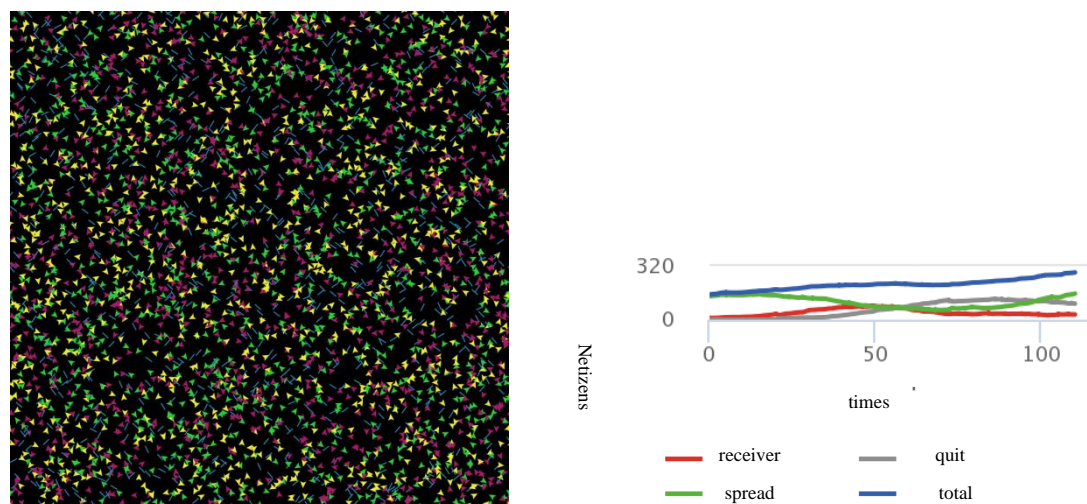


Figure 4. The Peak period of Internet public opinion dissemination



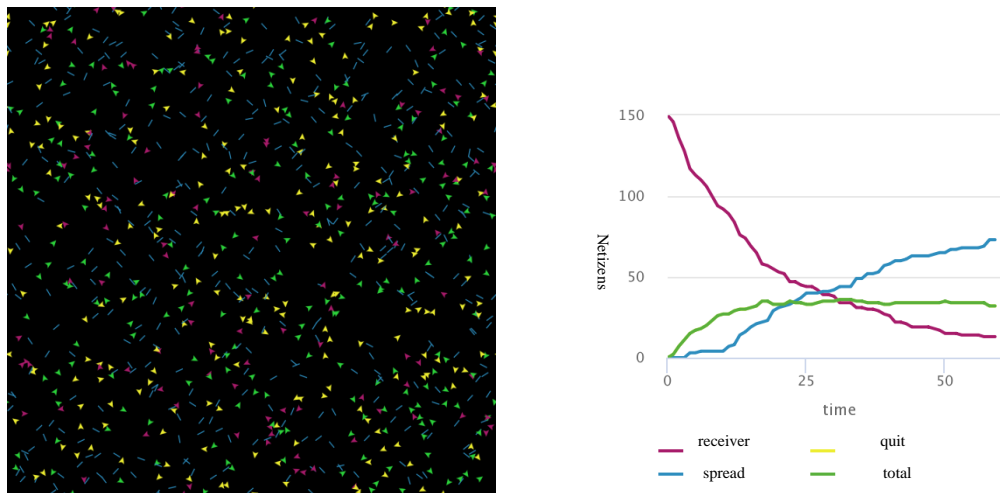


Figure 5. The mitigation period of Internet public opinion dissemination

(3) Analysis of simulation results: As shown in Figure 4, we proposed that the highest value of the communicator is about 1,200 and the number of netizens participating in the entire public opinion event is as high as 20,000 and the proportion of communicators is only 11%. It can be seen that most of the netizens participating in the online public opinion incident are onlookers. If the grievances are not handled properly, these onlookers may turn into communicators, and the destructiveness brought about at that time will be unimaginable. In the real world, we call the more influential netizens the "opinion leaders" who are 5% of the network affecting 85% of netizens. The "Opinion Leaders" in the network are quite attractive and play an important intermediary or filtering role in the dissemination of information. In the event of a network crisis, "opinion leaders" often evoke a large number of topic followers with lyrical words, pertinent comments, and unique insights, which can lead to a wave of public opinion in the development of the crisis.

## 5. Conclusions

This paper takes the information audience's point of view analysis and monitoring and control in the network's sentiment field as the goal. The research dialectically absorbs the field theory in the natural sciences and social sciences, and deconstructs the concept of the network sentiment field from the aspects of connotation and structure on the basis of fully grasping the characteristics of the network public opinion, the movement process and the development trend. Then the elements of the network public opinion field are defined as five parts: subject, object, ontology, carrier and environment. It is clearly put forward that the network public opinion information as the body element is the explicit manifestation of the movement law of the network public opinion field. Based on the comparative analysis of existing domain knowledge association clustering methods, this paper proposes a public opinion fusion method based on association mining and semantic clustering. In the simulation stage of public opinion dissemination, the long-term dissemination stage of network crisis information is divided into five stages: latent period, outbreak period, spread period, mitigation period and long-tail period. This study has made a useful exploration for the analysis, monitoring and control of public opinion in the field of network public opinion. Although the results achieved preliminarily achieve the expected research objectives, but due to the limitations of time and energy and many other conditions, in the future, we need to further study the multi-dimensional knowledge fusion optimization and decision-making methods of public opinion information resources in the large data environment.

## Acknowledgements

This work was partially supported by the Natural Science and Technology Project Plan in Yan'an of China (2016KG-07); Natural Science Basic Research Plan in Shaanxi Province of China (2017NY-134, 2016NY141, 2016KJXX-62, 2017NY132), Funding Project for Department of Yulin University (16GK24).

## References

1. M. Ghiassi, J. Skinner, and D. Zimbra, "Twitter Brand Sentiment Analysis: A Hybrid System using n-Gram Analysis and Dynamic Artificial Neural Network," *Expert Systems with Applications*, Vol. 16, No. 1, pp. 6266-6282, February 2013
2. F. Zhang, H. F. Xue, and D. S. Xu, "Big Data Cleaning Algorithms in Cloud Computing," *International Journal of Online Engineering*, Vol. 9, No. 3, pp. 77-81, March 2013
3. J. Zhang and Y. Hong, "Opinion Evolution Analysis for Short-Range and Long-Range Deffuant-Weisbuch Models," *Physica A Statistical Mechanics and its Applications*, Vol. 392, No. 21, pp. 5289-5297, April 2013



4. A. C. R. Martins, "Continuous Opinions and Discrete Actions in Opinion Dynamics Problems," *International Journal of Modern Physics C*, Vol. 19, No. 4, pp. 617-624, April 2008
5. N. Anstead and B. O'Loughlin, "Social Media Analysis and Public Opinion: The 2010 UK General Election," *Journal of Computer-Mediated Communication*, Vol. 20, No. 2, pp. 204-220, March 2015
6. J. Berndtsson, C. Dandeker, and K. Ydén, "Swedish and British Public Opinion of the Armed Forces after a Decade of War," *Armed Forces and Society*, Vol. 41, No. 2, pp. 307-328, February 2015
7. J. Bolstad, "Dynamics of European Integration: Public Opinion in the Core and Periphery," *European Union Politics*, Vol. 16, No. 1, pp. 23-44, February 2015
8. V. G. Donkal and K. Gyanendra, "A Multimodal Fusion based Framework to Reinforce IDS for Securing Big Data Environment using Spark," *Journal of Information Security and Applications*, Vol. 43, No. 1, pp. 1-12, January 2018
9. F. Zhang, H. F. Xue, and J. -C. Zhang, "Multi-Source Big Data Dynamic Compressive Sensing and Optimization Method for Water Resources based on IoT," *Sustainable Computing: Informatics and Systems*, Vol. 20, No. 6, pp. 210-219, July 2018
10. J. Chae, D. Thom, Y. Jang, S. Y. Kim, T. Ertl, and D. Ebert, "Public Behavior Response Analysis in Disaster Events Utilizing Visual Analytics of Microblog Data," *Computers and Graphics*, Vol. 38, No. 2, pp. 51-60, February 2014
11. S. P. Chang and B. K. Kaye, "The Tweet goes on: Interconnection of Twitter Opinion Leadership, Network Size, and Civic Engagement," *Computers in Human Behavior*, Vol. 69, No. 6, pp. 174-180, June 2016
12. J. D. Clinton and J. A. Grissom, "Public Information, Public Learning, and Public Opinion: Democratic Accountability in Education Policy," *Journal of Public Policy*, Vol. 35, No. 3, pp. 355-385, April 2015
13. S. Counts and K. E. Fisher, "Mobile Social Net Working as Information Ground: A Case Study," *Library and Information Science Research*, Vol. 32, No. 2, pp. 98-115, February 2010
14. E. Ellis, N. James, and R. L. Jacobs, "Who Governs? Presidents, Public Opinion, and Manipulation," *Public Opinion Quarterly*, Vol. 80, No. 1, pp. 212-215, February 2016
15. J. A. Torkestani, "An Adaptive Focused Web Crawling Algorithm based on Learning Automata," *Applied Intelligence*, Vol. 4, No. 1, pp. 586-601, January 2012
16. R. Kontchakov, M. Rodríguez-Muro, and M. Zakharyashev, "Ontology-based Data Access with Databases: A Short Course," *Lecture Notes in Computer Science*, Vol. 2, No. 2, pp. 194-229, February 2013
17. H. Liu, K. H. Chai, and J. F. Nebus, "Balancing Codification and Personalization for Knowledge Reuse: A Markov Decision Process Approach," *Journal of Knowledge Management*, Vol. 17, No. 5, pp. 755-772, June 2013
18. N. Ur-Rahman and J. A. Harding, "Textual Data Mining for Industrial Knowledge Management and Text Classification: A Business Oriented Approach," *Expert Systems with Applications*, Vol. 39, No. 5, pp. 4729-4739, June 2012
19. K. Sznajd-Weron and J. Sznajd, "Opinion Evolution in Closed Community," *Expert Systems with Applications*, Vol. 11, No. 6, pp. 1157-1165, June 2000
20. R. Hegselmann and U. Krause, "Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulations," *Journal of Artificial Societies and Social Simulation (JASSS)*, Vol. 5, No. 3, pp. 1-33, April 2002

**Pengju Wang** is received the MS degree in Computer science from Northwestern Polytechnical University in 2006. Now he is a PhD of Northwestern Polytechnical University. His research interests are in the areas of network public opinion analysis and data mining, the modeling of complex systems.

**Huifeng Xue**, he was born in Shanxi Yuncheng on June 16, 1964, received the PhD degree in Water resource economics from Xi'an Polytechnic in 1995. He is currently a professor in Northwestern Polytechnic University. His research interests are in the areas of the modeling of complex systems, Simulation and performance evaluation, management, systems engineering, energy and environmental systems engineering, computer control, intelligent control, network control.

**Zhe Yu** is received the MS degree in Chinese Language and Literature from Northwestern University in 2005. She is currently a deputy director of information comprehensive office of Yan'an municipal committee. Her research interests are in the areas of Public Opinion Data Analysis, Data Mining, and Complex System Modeling.

**Feng Zhang** is received the MS degree in Computer science from Xidian University in 2009. Now he is a PhD of Northwestern Polytechnical University. He is currently a professor in Yulin University. His research interests are in the areas of Cloud integrated manufacturing technology, the modeling of complex systems, the Internet of Things applications.