RESEARCH ON THE ENTHUSIASM LEVEL OF SELF-MEDIA NETWORK PUBLIC OPINION BASED ON PROJECTION PURSUIT MODEL

Huizi ZHOU [1], Jie ZHANG[2]*

With the development of wireless network technology and a surge in the number of mobile phone users, the self-media network public opinion that is taking WeChat, micro-blog and other new media as the platform gradually has replaced the traditional media public opinion and has become dominant force in the field of public opinion at the same time, it has had an important influence on social stability. Based on this, in order to solve practical problems such as nonlinearity and high dimensionality of sample data in early warning evaluation of self-media network public opinion, we use cultural gene algorithm to determine the optimum projection direction and establishes a projection pursuit model of evaluation of self-media network public opinion which is based on cultural gene algorithm. Simultaneously, we apply the above model to "The event of Gansu, Lanzhou excessive benzene content in tap water" in 2014 and carries out the empirical researches. The empirical researches show that the projection pursuit model which is based on cultural genetic algorithm can effectively solve practical difficulties such as nonlinearity and high dimension of sample data, and it has good practicability.

Keywords: Self-media; network public opinion; coping hierarchy; Projection Pursuit Model

1. Introduction

In recent years, with the unceasing development of Internet, self-media platform which is taking Forum, Tieba, micro-blog, blog, WeChat and others as the principal thing has been emerged in our country. The public speaks their mind freely and delivers their views and comments on the media platform. Self-media network public opinion in China has got rapid growth. Self-media platform can provide a space for the public to freely speak speech. Public opinion is phonated and fermented here, which gradually forms a unique public opinion field. The formation of the private opinion field of self-media platform has a fast speed and high enthusiasm level with opinion leaders appeared one after another. If the check is not strict, information accuracy is low, it is easy to mislead public

[1] School of Economics and Management, Beijing Jiaotong University, Beijing, China
[2]* School of Economic and Management, Shandong University of Science and Technology, Qingdao, China, corresponding author, e-mail: zhangjie0371@163.com

opinion. Such as “8.12” explosion accident in Tianjin Port, "Wei Zexi" incident, “Wang Baoqiang " incident and others, these events form group polarization events on the self-media network platform. However, the rational position of public opinion has been subsided fast, irrational rumors and false contents have added fuel to the flames, which caused social laceration and antagonism, so as to bring about terrible influence, and even they became social problems that were more harmful than the event itself.

2. State of art on self-media network public opinion

For the research on self-media network public opinion, foreign scholars have conducted a large number of researches, and gained some achievements. For example, Ochrombel（2001）adopted network methods of Watts and Newman to create a small world network. In this small world network model, the concepts of leader are reflected prominently. The leader is defined as an individual who is responsible for the dissemination and interpretation of various opinions. It is able to change the views and opinions of other individuals easily and completely, while fully keeping their own views without changing them [1]. Sznajd model created by Stauffer（2002）is the first interactive model of public opinion appeared in academic theory circles. The model is established on the basis of “united we stand divide we fall”, which is thought that the public opinion is expanding from inside to outside [2]. Hegselmalm and others (2002) proposed a new model for self-media network public opinion, based on views distance, they put forward the view of “Limited trust”, and they pointed out that in the process of the dissemination of self-media network public opinion, public opinion not only includes two simple positive and negative viewpoints, but mutual interaction of multiple viewpoints, this is an important research direction in the evolution model of self-media network public opinion [3]. Antioine and others (2007) thought that in the process of supervision of self-media network public opinion, the methods and technologies of text classification based on clustering algorithm can be adopted to conduct the briefing processing of public opinion and propensity score analysis and other operations of self-media public opinion on the Internet, finally to realize the early warning, guidance and monitoring of self-media network public opinion [4]. Vadrevus and others (2007) thought that the analysis of self-media network public opinion is needed to conduct multivariate statistical analysis, and at the same time, it also is adopted multiple classification and discrimination method which includes distance discrimination, Fisher discrimination method and Bayes discrimination [5]. Gil-Garcia and others (2010) thought that, in the process of supervision of self-media network public opinion, it is mainly used text clustering and other methods to find, collect and track hot topics of public opinion, so as to look for the best path to achieve public opinion warning [6].

Jiang Shenghong (2008) divided the evolution of self-media network public opinion into the four stages of generation, rising, fluctuation, and termination [7]. Gu Mingyi (2009) thought that under the “audiences” function of new media with social characteristics and extensive network behaviors; the self-media network public opinion would develop along with the stages of early propagation, the information of sociality, social expression, social action, media commemoration and other stages [8]. Yu Guoming and others (2009) thought that the general process of the evolution of self-media network public opinion is “Event occurrence - Network tip-offs - Media follow-up - Network hype - Public opinion formation - Government intervention - Public opinion down." [9] Cao Jinsong (2010) thought that according to the analysis based on timeline, the development of self-media network public opinion has four stages and three strategic passes that “Dissemination - Gathering - Hot discussion - Fashion" and “Explosion, Sublimation, Continuity” [10]. Zong Liyong (2010) constructed a multi-agent model for the evolution of emergencies of self-media network public opinion, and at the same time, through analogue simulation, he thought that the evolution of public opinion has a state characteristic of the “S” curve [11]. Zhu Hengmin (2012) proposed a SIRS model for the evolution and propagation of self-media network public opinion, which was constructed based on directed BA network. In the results of simulations, he believed that there is a positive correlation in the evolution of self-media network public opinion, and the range and speed of propagation of self-media network public opinion are in proportion to influence [12].

In this paper, we will carry out the intensive study on enthusiasm level of self-media network public opinion and will adopt projection pursuit model to conduct the quantitative evaluation of enthusiasm level to deepen the researches of self-media network public opinion. The paper is structured as follows: The remainder of this introduction explains the context and motivation. Section 2 discusses related work. In Section 3 the analytical model is developed. Section 4 discusses the results of our experimental evaluations. Section 5 presents a round-based perspective of replication and coverage.

3. The construction of evaluation system for the enthusiasm level of self-media network public opinion based on projection pursuit model

3.1 The construction of projection pursuit cluster model

In 1974, Friedman and Tukey formally proposed the concept of projection pursuit. Projection pursuit model is a model established on the basis of projection pursuit theories, which has been widely used, and has obvious advantages in dealing with high-dimensional data. The optimization of the establishment of projection pursuit cluster model based on Memetic Algorithm includes the following steps [13]:

(1) The normalization processing of sample evaluating indicator sets.

To set the sample set of index values as $\{x^*(i,j)|i=1,2,\cdots,n;\ j=1,2,\cdots p\}$, here $x^*(i,j)$ is the j-th index value of i-th sample, n and p are samples respectively. In order to eliminate the dimension and unify the variation range of each index, the Eq. (1) and Eq. (2) can be used to conduct the extremum normalization:

For the index which is the bigger, it will be better:

$$x(i,j)=\frac{x^*(i,j)-x_{\min}(j)}{x_{\max}(j)-x_{\min}(j)} \tag{1}$$

For the index which is the smaller, it will be better:

$$x(i,j)=\frac{x_{\max}(j)-x^*(i,j)}{x_{\max}(j)-x_{\min}(j)} \tag{2}$$

In above formulas, $x_{\max}(j)$ and $x_{\min}(j)$ are the maximum and minimum of j–th index, $x(i,j)$ is the sequence of the normalization of index characteristic values.

(2) Construct the function $Q(a)$ for evaluating indicator

The projection pursuit model method is to synthesize p dimension data $\{x(i,j)|j=1,2,\cdots,p\}$ into one-dimensional projection value $z(i)$ which is taking $a=\{a(1),a(2),\cdots,a(p)\}$ as projection direction, so $z(i)$ can be got through the following formula:

$$z(i)=\sum_{j=1}^{p}a(j)x(i,j), i=1,2,\cdots,n \tag{3}$$

In Eq. (3), a is unit length vector. When we solve composite projection index value, the scatter characteristic of projection value $z(i)$ is required. And the scatter characteristic should be: local projection points are as dense as possible, which are best to agglomerate into a number of dots, but on the whole, the projection points should be scattered as far as possible. Therefore, the function for evaluating indicator can be expressed as:

$$Q(a)=S_Z D_Z \tag{4}$$

In Eq. (4), S_Z is the standard deviation of projection value $z(i)$, D_Z is the local density of projection value $z(i)$, which are defined as follows:

$$S_Z=\sqrt{\frac{\sum_{i=1}^{n}(z(i)-E(z))^2}{n-1}} \tag{5}$$

$$D(z)=\sum_{i=1}^{n}\sum_{j=1}^{n}(R-r(i,j))\bullet u(R-r(i,j)) \tag{6}$$

In Eq. (5) and Eq. (6), $E(z)$ is the average value of sequence $\{z(i)|i=1,2,\cdots,n\}$; R is the window radius of local density. Its selection not only make that the average number of projection points contained in the window is not too small to avoid the deviation of moving average is too big, but also doesn't make it increase too high according to the increasing of n; $r(i,j)$ is the distance between samples, which is defined as follows:

$$r(i,j)=|z(i)-z(j)| \tag{7}$$

In addition, $u(t)$ is unit speed function, when it is the situation that $t\geq 0$, its value is one, and when it is $t<0$, the function value is zero.

(3) Optimize the projection indicator function.

When the sample set of each index value is given, the projection index function $Q(a)$ can only be changed as the changes of the projection direction a, The different projection directions reflect the different features of data structure, the best projection direction is to expose the projection direction of a characteristic structure in high-dimensional data in highest possible. So by solving the maximization problem of the projection index function, it can estimate the optimum projection direction, that is:

$$MaxQ(a)=S_Z\bullet D_z \tag{8}$$

(4) Ensure constraint condition

Fredman and others (1974) set up projection pursuit model, and pointed out that in the projection direction a, each vector should be mutually orthogonal, so the constraint condition is:

$$s.t.\begin{cases}\sum_{j=1}^{p}a^2(j)=1\\ 1\geq a(j)\geq -1\end{cases} \tag{9}$$

(5) To optimize the projection index function by using cultural gene algorithm

Aimed at particle swarm optimization algorithm to solve high-dimensional optimization problems that will be easy to trap into local optima, the sensitivity of simulated annealing algorithm to initial parameter is insufficient, in this paper, we put forward cultural gene algorithm which is taking the improved particle swarm optimization as a global search strategy and taking the improved simulated annealing algorithm as a local search strategy. In the paper, we also select particle swarm optimization algorithm as global search strategy, in the process of search, simulated annealing algorithm is conducted the local search strategy on group extremum generated in each iterative process. This coordination method not only improves the efficiency of search, but also the accuracy of search.

The basic principle of particle swarm optimization is from the simulation of bird predation behaviors. Simulated Annealing, SA, is started at a higher initial temperature, along with the constant decline of temperature parameter, and combined with the probabilistic jumping property which means that at the locally optimal solution, it can jump out of probability and eventually tend to global optimum. Setting population size as N , in the target search space of D dimension (it means: the dimension of the optimum projection direction $a = |a(1), a(2), \cdots, a(p)|$). The i -th ($i = 1, 2, \cdots, N$) particle position in population can be represented as a D vector, $V_i = (V_{i1}, V_{i2}, \cdots, V_{iD})^T$ is indicated the flying speed of i -th particle. $P_i = (P_{i1}, P_{i2}, \cdots, P_{iD})^T$ is shown the best solution searched by i -th particle. $P_g = (P_{g1}, P_{g2}, \cdots, P_{gD})^T$ is shown the best solution searched in this search population, which means the optimum projection direction got in this iteration. A new solution j is got through the perturbation of the iteration. According to principles of Metropolis, the new solution whether to be accepted as the following formula:

$$\exp(-\frac{Q(j) - Q(i)}{t_k}) = \exp(-\frac{\Delta Q}{t_k}) \tag{10}$$

The Simulated Annealing is conducted the cyclic iteration by lowering the temperature, it is:

$$t_{k+1} = \alpha \times t_k, k = 0, 1, \cdots \tag{11}$$

In Eq. (11), α is attenuation factor.

When the temperature reaches the predetermined termination temperature, the optimum solution got at this moment can replace the global optimal solution, which will be conducted the next time iteration until the iteration is over.

In this paper, contraction factor s is introduced [14, 15], which can effectively avoid getting into local optima during the search process, the following velocity and position update formulas are presented:

$$V_{ij}^{k+1} = sV_{ij}^k + c_1 r_{1j} (p_{ij}^k - x_{ij}^k) + c_2 r_{2j} (p_{gj}^k - x_{ij}^k) \tag{12}$$

Here, k is given by:

$$k = \frac{2}{\left| 2 - \phi - \sqrt{\phi^2 - 4\phi} \right|} \tag{13}$$

In Eq. (13), parameter ϕ is:

$$\phi = c_1 + c_2 \tag{14}$$

In Eq.(12), $i = 1, 2, \cdots N$, j is j -th dimension of particulate, k is iterations,

c_1 and c_2 are acceleration constants, whose usual values are between one to two, $r_1 \sim u(0,1)$, $r_2 \sim u(0,1)$ are two independent random functions. In this paper, on the basis of traditional simulated annealing, a storage memory apparatus is set up to preserve the optimal solution of iteration, and at last, the instruction of $\min$ is selected to find the best one of the optimal solution in the storage memory apparatus:

$$H = (a^*, Q(a)) \tag{15}$$

In Eq. (15), a^* is the optimal solution determined by object function in the annealing process, $Q(a)$ is the value of objective function. The specific optimization steps are shown in Fig.1.

(6) To calculate projection value and differentiate grade

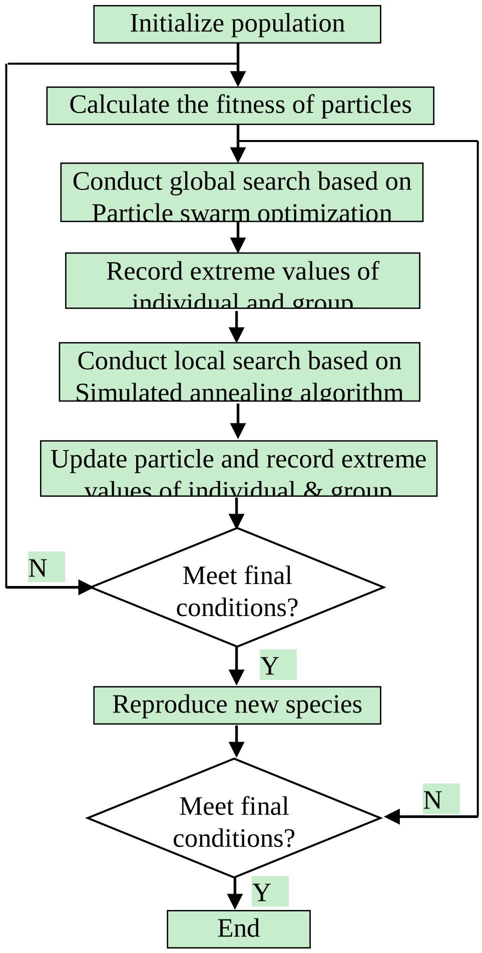

Fig. 1. The optimization process of projection pursuit algorithm

3.2 The construction of evaluation index system of enthusiasm level of self-media network public opinion

In order to construct a scientific, reasonable and effective evaluation index system, the indexes should be scientific, systematic, measurable and hierarchical and other features, and the same time, they must be concerned about the combination of the qualitative and quantitative indicators. According to the influence factors of create procedure and rise and fall of enthusiasm level, starting from the impact index contained in these four aspects: self-media network public opinion, network media, network citizens, government and others, we draws on the experience of relatively mature index of self-media network public opinion contained in literatures [16, 17], at the same time, according to Xue Lan's research results [18] which are mainly reflected the scale and extent of damage of the event, we add these indicators to the index structure and classifies the concerned indexes according to four factors. A team of experts formed by ten experts in the field of emergency management uses Delphi Method to calculate the total score of each factor according to factor score. They come to an agreement to merge repeated indexes, and to abandon low scoring indexes, and they select third level indexes which can constitute the highest display in each second level index to structure an evaluation system which is practical, practical and integrated. In which, there is one first level index, four second level indexes, twelve third level indexes, these indexes can be divided into subjective index and objective index, which can be seen in Table 1:

Table 1

The evaluation index system for enthusiasm level of self-media network public opinion

First index	Second level indexes	Third level index
Enthusiasm level of self-media network public opinion（I）	The intensity of topic（I_1）	Topic sensitivity（i_{11}）
		Attraction degree of topic（i_{12}）
		Influence degree of promulgator（i_{13}）
	The enthusiasm level of topic（I_2）	Focusing degree of topic（i_{21}）
		Posting amount of topic（i_{22}）
		The reply volume for topic（i_{23}）
	The state of public opinion（I_3）	Tendency of public opinion i_{31}）
		Participation degree of public opinion topic（i_{32}）
		The tendency of public opinion（i_{33}）
	The trend of public opinion（I_4）	Change frequency of public opinion（i_{41}）
		Change flexion of public opinion（i_{42}）
		Change activity of public opinion（i_{43}）

4 Empirical simulation

4.1 The setting of calculated parameters

The sample of empirical research in the paper is selected" The event of Gansu, Lanzhou excessive benzene content in tap water" as example to act as a source of experimental data [19]. From April 11, 2016 to April 27, 2016, additional Sina micro-blog public sentiment information about "The event of Gansu, Lanzhou excessive benzene content in tap water" is combed (to accumulate the number of articles, retransmissions, comments, the number of points praise of Micro-blog public opinion about this event.). 17 day data of self-media network public opinion are selected as processing samples, and the sample data are selected average value, at the same time, each phase of time is based on one day, in which $h_1 \sim h_{17}$ are indicated as 17 time phases, they are as shown in Table 2:

Table 2

Sample data

index / phase	I_1			I_2			I_3			I_4		
	i_{11}	i_{12}	i_{13}	i_{21}	i_{22}	i_{23}	i_{31}	i_{32}	i_{33}	i_{41}	i_{42}	i_{43}
h_1	0.6	0.7	0.1	0.5	8975	3572	-0.78	0.2	0.1	0.52	1	0
h_2	0.8	0.3	0.4	0.3	7214	2079	-0.32	0.5	0.3	0.37	1	0
h_3	0.5	0.8	0.2	0.5	14303	8267	-1.0	0.7	0.4	0.28	0	0.48
h_4	0.7	0.5	0.5	0.3	17679	12063	-0.54	0.6	0.4	0.44	1	0.63
h_5	0.4	1.0	0.3	0.2	12702	8985	-0.29	0.8	0.6	0.62	0	0.51
h_6	0.8	0.7	0.1	0.4	9839	5230	-0.86	0.4	0.5	0.49	0	0.39
h_7	0.5	0.6	0.4	0.3	11026	6184	-0.35	0.3	0.4	0.35	1	0.54
h_8	0.7	1.0	0.2	0.3	8967	4305	-0.47	0.5	0.5	0.50	0	0.36
h_9	0.6	0.4	0.3	0.4	7205	3219	-0.65	0.4	0.3	0.38	1	0.42
h_{10}	0.5	0.6	0.5	0.2	10383	5028	-0.33	0.3	0.4	0.29	0	0.55
h_{11}	0.5	0.5	0.2	0.5	8734	3236	-0.71	0.2	0.3	0.31	0	0.43

h_{12}	0.7	0.8	0.3	0.4	6539	2710	-0.45	0.3	0.2	0.42	1	0.34
h_{13}	0.4	1.0	0.5	0.2	7312	2503	-0.32	0.3	0.2	0.36	0	0.40
h_{14}	0.3	0.5	0.2	0.2	5197	2318	-0.38	0.4	0.4	0.29	1	0.37
h_{15}	0.6	0.4	0.2	0.4	4808	1911	-0.29	0.2	0.3	0.30	0	0.28
h_{16}	0.7	0.4	0.3	0.3	3724	1206	-0.25	0.1	0.1	0.28	0	0.33
h_{17}	0.5	0.7	0.4	0.5	1831	327	-0.18	0.2	0.2	0.25	0	0.25

The window radius R is the unique parameter which will determine the optimal projection vector and of its coefficients. The selection criteria of R is that sample points in the window should not be too little, at the same time, when the number of sample points increases, the samples points in the window cannot increase too much. In the paper, the value of R is: $\max(r_{ik})/5 \le R \le \max(r_{ik})/3$, which means when the distance r_{ik} between the two samples is smaller than the above R value, the summing calculation is conducted. In accordance with such principles, samples can usually be divided into 3~5 categories.

Parameter setting for cultural gene algorithm: initial temperature is 100, the coefficient of temperature drop is 0.6, the chain length of Markov is 100, and termination temperature is 0.01, the number of particle is 200, the number of iterations is 1000, the contraction factor is $s = 0.9$, the acceleration factor is $c_1 = c_2 = 2$, and the search ranges of all parameters are initialized to [-1, 1], which are conformed to the ranges of projection vectors.

4.2 Calculation process and result analysis

By using Matlab to conduct the calculation, the optimum projection direction of comprehensive evaluation target is a =(0.4862，0.6154，0.0197，-0.0403，-0.1796，0.4131，0.2538，-0.0497，0.1326，-0.2108) and the projection index function value is 3.4562. The optimum projection directions of sub-evaluation target I_1, I_4 are respectively $a_1 = (0,0,1)$ and $a_4 = (0,1,0)$; and the optimum projection directions of sub-evaluation target I_2, I_3 are respectively $a_2 = (0.5371, 0.7018, 0.4682)$ and $a_3 = (0.8619, 0.4358, 0.2791)$. The negative sign is only indicated that the evaluating indicator is negative index, the projection value is shown in Tab.3, the projection drawing are shown in Fig.2 and Fig.3.

Table 3

Projection value of each index

samples	z_1	z_2	z_3	z_4	Comprehensive indexes z	categories
h_1	0.8725	1.0000	1.3786	0.0000	1.2031	Medium warning
h_2	1.0000	1.6728	1.4297	1.0000	1.7302	Severe warning
h_3	1.0000	1.3067	0.6729	-0.0000	0.5844	Mild warning
h_4	0.0000	0.9325	0.6670	-0.0000	0.0413	security
h_5	0.3879	1.4760	0.5831	1.0000	0.6875	Mild warning
h_6	1.0000	1.5381	0.7382	-0.0000	1.3626	Medium warning
h_7	0.0000	0.8462	0.5633	-0.0000	-0.0437	security
h_8	0.3402	1.3288	0.6951	1.0000	0.7804	Mild warning
h_9	0.4736	0.9835	0.7204	1.0000	0.5870	Mild warning
h_{10}	0.0000	0.8203	0.3527	-0.0000	0.0584	security
h_{11}	0.2670	1.4293	0.7116	1.0000	0.5801	Mild warning
h_{12}	0.4022	1.5628	0.5302	1.0000	0.7249	Mild warning
h_{13}	0.0000	1.2045	0.6075	-0.0000	0.1622	security
h_{14}	0.0000	0.3891	0.2638	1.0000	0.0886	security
h_{15}	0.0000	1.4762	0.8320	-0.0000	0.1417	security
h_{16}	0.0000	1.1063	0.5071	-0.0000	0.2108	security
h_{17}	0.0000	1.2307	0.2375	1.0000	0.0973	security

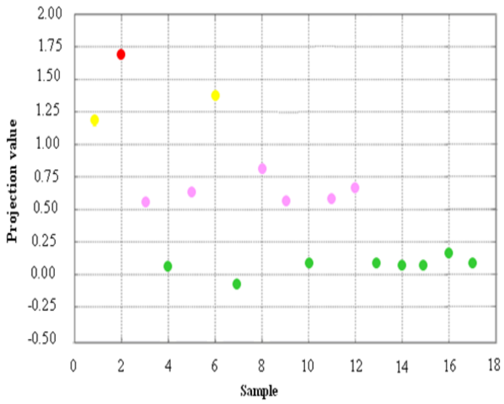

Fig. 2. Projection scatter plot

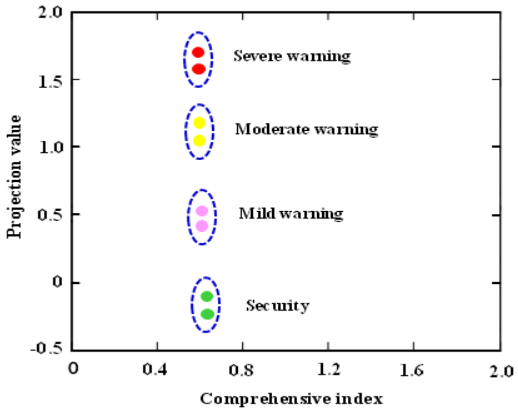

Fig. 3. Projection point mapping

We can obviously see from Fig. 2 and Fig. 3 that, when all the evaluation indicators are integrated, all the samples can be divided into four grades. h_2 is serious warning; h_1 and h_3 are medium warning; h_6, h_8, h_9 and h_{11} are light warning; h_4, h_5, h_7, h_{10} and h_{12} are secure. Through the grade division, we can conduct the analysis of pre-warming of self-media network public opinion commendably, which can provide theory evidences for subsequent network security. From Fig.2, Fig.3 and Tab.3, we can see that projection pursuit clustering model can get classification results without human factor interference, which is obtained only through the projection clustering of each evaluation index that influences self-media network public opinion, so it is relatively objective. In addition, in the evaluation model of the pre-warming of self-media network public opinion, the size of each component of the optimum projection direction actually reflects the influence degree of each evaluation index on the classified evaluation of self-media network public opinion; it is just as shown in Fig.4.

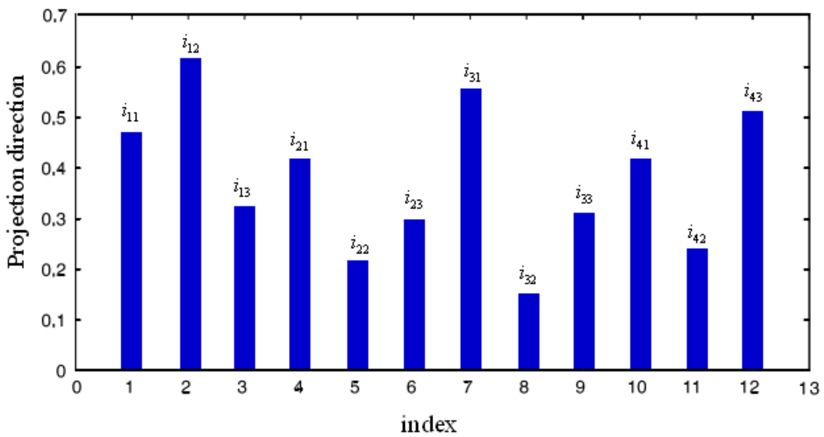

Fig. 4. Projection direction of each evaluation index.

The optimum projection direction shows that the evaluation indexes of influence degree of the classified evaluation of self-media network public opinion are in turn topic sensitivity, the reply volume for topic, participation degree of public opinion topic, tolerance of public opinion topic, change activity of public opinion, tendency of public opinion, change flexion of public opinion, change frequency of public opinion, focusing degree of topic, influence degree of promulgator, attraction degree of topic and posting amount of topic.

5 Conclusions

In the paper, the optimized projection pursuit model based on MA (Memetic Algorithm，MA) can effectively solve modeling problems of high-dimensional nonlinear systems, and the analysis of various evaluation indicators is translated from human subjectivity into objective and quantitative fixed value

analysis. The optimized projection pursuit model is used in pre-warning evaluation of self-media network public opinion, and through the indications of case study, it can evaluate the self-media network public opinion commendably.
MA algorithm is a framework, which is taking the improved particle swarm optimization algorithm as a global search strategy, and the improved simulated algorithm as local search strategy to overcome the shortcomings that particle swarm optimization is easy to trap into local optimum and the sensitivity of simulated annealing algorithm to initial parameter is insufficient, and it has high practicability in the optimization problems.

Acknowledgment

This work was supported by the Natural Science Foundation of Shandong province (Grant No. ZR2015GM013), and by the Key Project of National Statistical Research Program (Grant No. 2015LZ25 and 2016LZ31).

R E F E R E N C E S

[1] Ochrombel R. Simulation of Sznajd socio physices mode with convincing single opinions[J]. Hiternational joumal of modem physics C, 2001,7(7):10-91.
[2] Stauffer D, Oliveira P M C. Persistence of opinion in the Sznajd consensus model: Computer simu-lation [J]. The European physical journal B -condensed matter. 2002, 30 (4):587-592.
[3] Rainer Hegselmann，Ulrich Krause. Opinion dynamics and bounded confidence models, Analysis and Simulation [J]. Journal of artificial societies and social simulation, 2002, 5(3): 324-336.
[4] Antoine Naud, Shiro Usui. Exploration of a text collection and identification of topics by clustering [J]. Lecture notes in computer science, 2007,12(12):115-124.
[5] Vadrevu S, Gelgif, Davulcuh. Information extraction from web pages using presentation regularities and domain knowledge [J]. World wide web, 2007,43 (10): 157-179.
[6] Gil-Garcia R, Pons-Porrata A. Dynamic hierarchical algorithms for document clustering[J]. Pattern, recognition letters ,2010,31(5):469-477.
[7] Yu Guoming, Li Biao. Characteristics and reference significance of government crisis intervention in public opinion hot spot [J]. Journalism and writing, 2009,6(6): 57-59.
[8] Cao Jinsong. The development law of network public opinion [J]. Chinese journal of journalism and writing, 2010,5(5):140-142.
[9] Jiang Shenghong. The formation, development, current situation and public opinion guidance of hot spot of network public opinion [J]. Chinese journal of theory monthly, 2008,4(4):127-129.
[10] Gu Mingyi, Zhou Renwei. Network public opinion and social network information dissemination model [J]. Chinese journal of journalism and communication study, 2009,5(5): 67-73.
[11] Zong Lirong, Gu Baoyan, Sun Shaorong. Research on the evolution of public opinion of network crisis based on the allocation mechanism of attention resources[J]. Chinese journal of intelligence theory and practice, 2010,33(10):143-148.
[12] Zhu Hengmin, Li Qing. Research on topic based micro-blog public opinion communication model[J]. Chinese journal of modern library and information technology, 2012, 5(5):6-9,35.
[13] J H Friedman, J W. A projection pursuit algorithm for exploratory data analysis [J]. IEEE trans on computer, 1974, 23(9):881—890.

[14] J Kennedy, R C Eberhart. Particle swarm optimization// international conference on neural networks[C]. Perth, Australia, IEEE,1995:1942-1948.
[15] Afshar M，Asadpour M. Opinion formation by informed agents[J]. Journal of artificial societies & social simulation, 2010, 13 (4) :5-7.
[16] Zeng Qun, Cheng Xiao etc. Study of public opinion dissemination mechanism on social networks based on dual-pathway model [J]. Chinese journal of information science,2017, 35(6):29-43.
[17] Li Zhe. Comprehensive evaluation of online public opinion based on projection pursuit model [J]. Chinese Journal of computer simulation,2017,34(4):391-395.
[18] Mao Qianren, WANG Chaohin etc. Research on guidance mechanism of internet public opinion based on cellular automaton model [J]. Chinese journal of computer engineering, 2017, 43(6):105-110.
[19] Cao Xue-yan, Zhang Xian etc. Research on Internet Public Opinion Heat based on the Response Level of Emergencies [J]. Chinese Journal of Management Science,2014,22(3):82-89.