Application research on complex network theory in the model of public opinions monitoring on weibo netspeak

Haiyan Gao
Jilin Sport University, Changchun, 130000, (CHINA)

ABSTRACT

The monitoring and warning of public opinions on the Internet has been the targets of experts and scholars in this industry since it posed a serious challenge for the Internet as the third media. Currently, the research in this field has scored some achievements. This thesis adopted complex network theory as the basic theory and arranged modeling with the target of the evolution of public opinions on the Internet in the hope to discover the evolution patterns of public opinions on the Internet and provide technical support and aids to the monitoring of public opinions on the Internet. This thesis has 3 main parts: the first part is to analyze the subjective initiative of each netizen in the evolution of public opinions on the Internet, which was conducted based on the basic online structure, and come up with the evolution model of public opinions on the Internet through detailed analysis; the second part is come up with the spreading model of public opinions in weibo based on the conventional virus spreading model with the analysis of whether weibo users repost on weibo; the third one is to come up with derivative public opinions spreading model with focus on topics based on public opinions spreading model in weibo and with full considerations of the derivative behaviors during the spreading of public opinions on the Internet. This thesis conducted simulation experiments on all parameters in all the models. By the thorough analysis of experimental data, it shows that the models composed by this thesis can reflect the evolution rules of public opinions on the Internet and is highly practical.

KEYWORDS

Complex network theory; Netspeak; public opinions monitoring; Evolution model.

INTRODUCTION

As the Internet fast develops, netizens have become the creators and carriers of the contents on the Internet. The Internet has become an inseparable tool for people to obtain information, express opinions and exchange thoughts[1]. The netspeak constitutes a kind of unique expressions shown in online communication. Meanwhile, the netspeak is also a special language created by netizens for online communication[2]. It has become an indispensable part in people's life and there is a lengthening queue of Chinese netizens.

If the individual netizen is seen as a node on the Internet, the evolution process of public opinions is finished in a network with a large number of interconnecting nodes on it. As the Internet becomes more open, the number of Chinese netizen is increasing at a surprising rate. Therefore, to address the evolution of public opinions on the Internet from the perspective of complex network theory is of great significance.

THE EVOLUTION MODEL OF PUBLIC OPINIONS ON THE INTERNET BASED ON BA NETWORK

The main model of the evolution of public opinions is the model worked out by Sznajd and Vazquez based on Ising evolution rules[4]. However, in Ising and Galam models, the initial values of nodes discrete, which means to consider 3 stances: neutral, completely negative and completely positive. But on the Internet, netizens' attitude is not so clear in the beginning. The Internet and the media system derived from it, and all linguistic systems all have influences on people's life. The model comprehensively considered the node degree of the Internet and external factors' influences on public opinions[5]. However, the precondition of this model's iterated condition doesn't fit the open Internet.

Since the fact that new Internet users tend to connect to the known nodes that already existed on the Internet is similar to the forming pattern of world-wide-web, w is used in the model to represent the change of trust among the modes caused by similar opinions. This kind of change is the main stimulus of the evolution of public opinions.

During the process of opinion evolution, the iterated conditions of variables of all status are shown as follows:

$$\alpha_i(j,t) = 1 - \frac{\left|o_j(t) - \mu_i\right|}{\max\left|o - \mu_i\right|} \tag{1}$$

Among which,

$$\max\left|o - \mu_i\right| = \begin{cases} 1 - \mu_i & o \le \mu_i < 0.5 \\ \mu_i & 0.5 \le \mu_i \le 1 \end{cases}$$

$$S_i(j,t+1) = \frac{S_i(j,t) + S_i(j,t) * a_i(j,t)}{\sum\limits_{j=k}^{n}(S_i(j,t) + S_i(j,t) * a_i(j,t))} \tag{2}$$

$$o_i(t+1) = \frac{\sum\limits_{j=k}^{n} S_i(j,t) * o_i(t)}{n} \tag{3}$$

On formula (1), the number of adjacent nodes of node i is n. When there is an interaction between nodes $< i, j >$ and if and only if $S_i(i,t) * a_i(i,t) \ge \varepsilon$, other similar researches don't analyze the topological structure of spreading network of the netspeak. Instead, they just put up some new ideas on the control and monitoring of the spreading of the netspeak. This thesis supposes the interactions among all users are symmetrical, so the edge linking the nodes is undirected. The analysis focuses on the topological structure of the spreading network of the netspeak as the starting point and discovered that the topological structure of the structure of social network lead to fast spreading of the netspeak in the complicated social network, which is similar to the spreading of other popular expressions. Besides, during the accelerated process of the spreading of the netspeak, the key is that individual users are very acitive. The degree distribution of nodes on the Internet is shown in Figure 1.

According to the aforementioned evolution rules of opinions, the precision of the original computer network security evaluation system is low, including the conventional fault tree analysis, analytic hierarchy process and grey model. And the so-called expert cyber security evaluation system is a different evaluation system, which required experienced expert

evaluations. The results are subjective and the security evaluations are not scientifically solid, not suitable for evaluating complex systems concerning multiple factors. The results of the experiment are shown in Figure 2.
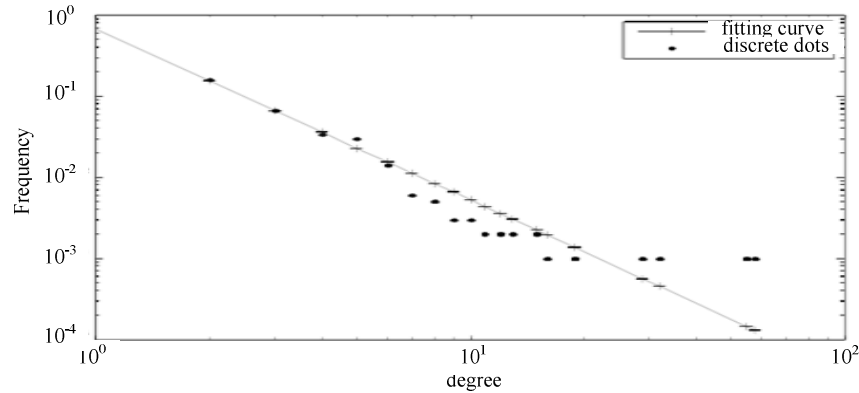

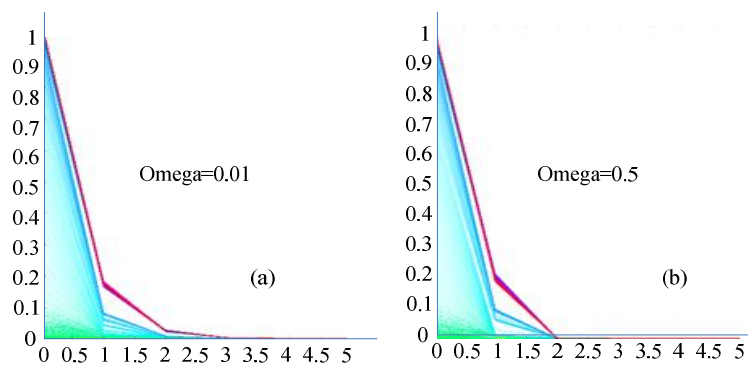**Figure 1 : The distribution of initial network**

Figure 2 : The evolution process of online opinions, when omega=0.01 or 0.5

It can be seen from Figure 2 that online opinions tend to get stable as time passes by. The cyber security evaluation system was founded for the influencing factors of cyber security. This system can comprehensively, objectively and scientifically reflect the main factors of cyber security[7]. The first-class evaluation standards of cyber security evaluation system include physical security, logical security, and management security. Physical security includes the safety of power supply, the safety of computer room, fault tolerant redundancy, line security and electromagnetic leakage-proof measures. Logical security includes anti-virus measures, software security, intrusion defense, data encryption, system audit, digital signature, access control, data backup and data recovery. The rules of cyber security evaluation system include feasibility, independence, completeness, accuracy and conciseness.

From the iterated conditions of the evolution model of opinions, it can be seen that nodes of online structure have influenced the trustworthiness among the nodes. This thesis thoroughly analyzed particle swarm optimization, which is the new search algorithm simulating birds flying over. This algorithm needs few samples but has powerful global search ability, such as optimizing neural network with particle swarm optimization can make up the weaknesses of neural network. This thesis starts from the current cyber security standards and the number of new vulnerabilities reported to CERT/CC is always on the rise. The computer network is very complex and it needs to consider all the factors that influence cyber security, so as to form scientifically solid cyber security evaluation standards. The results of the experiment are shown in Figure 3.
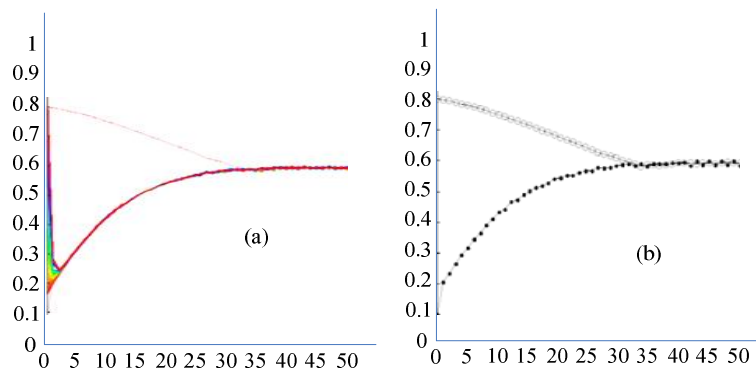
Figure 3 : Iterated structure's influence on opinions evolution

(a) In Figure 3 shows that as time passes, online opinions differ. An opinion value with only one node changes at a low speed. According to the standards of cyber security evaluation, the cyber security evaluation mechanism this thesis first confirms the scope of the target, also with the consideration of the actual situations of cyber security, and make evaluations in a scientific way to come up with the safety factor. Among which, the influential screw warms such as CodeRed, Nimda, Blaster caused 250 thousand computers to shut down and the economic damage was 4.6 billion dollars[6]. (b) shows the processes of changes of the largest node and the smallest node respectively. With the full consideration of management security, logical security and physical security in the cyber security system and the expert system, this thesis scored the cyber security evaluation standards to work out its weights. The cyber security standards reflect the security of computers from various perspectives. However, since each security standard represents different dimensions, it cannot be directly compared. To all data suitable for comparison and calculation, the results of the experiment are shown in Figure 4.
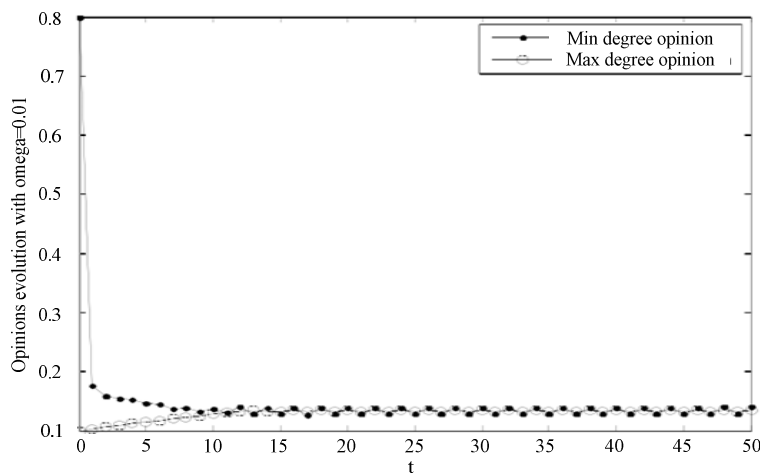

Figure 4 : The largest node-degree's influence on the results of opinions evolution

THE EVOLUTION MODEL OF THE SPREADING OF SINGLE TOPICS ON WEIBO

The spreading of a certain topic on weibo is finished by users' reposts. Select the optimum from all particles as the optimal weights and threshold of BP neural network based on PSO and set up the limit of iteration, and randomly set up the positions of particles. The original BP neural network model is taken as the comparison model of the experiment, and set up the error amount as 0.001 and excitation function being traingdx. This thesis takes the first 45 pairs of data as the training data for the 2 models. The experimental data shows that after the 417th iteration finished, the error amount is 0.000884. In contrast, after the 312nd iteration, the error amount is 0.000751. The construction algorithm of Price directed network mode[8] is:

(a) increase: starting from a network with m_0 I isolated nodes, every time a new node is introduced and comes through m directed lines m existed nodes, m<=m0 ;

(b) cumulating advantages: the probability of a new node edged coming to an existed node i with the indgree as k_i fits the following equation (a is a given constant):

$$\Pi_i = \frac{k_i + a}{\sum_j (k_i + a)}$$ **(4)**

The original network of this model is made up with m_0 isolated nodes. Through analyzing the characteristics of online management, it's possible to put cyber security management system in use. The practice has proved that this system has significant improvement comparing with conventional management methods in the perspectives of cyber security management, resources sharing, saving management costs and flexibility. It's fair to say that the efficient calculating ability and powerful storage ability have been fully displayed in this system, enabling China's cyber security management system to largely advance, shown in Figure 5.

Figure 5 shows that the online indegrees distribute in accordance with power-law distribution, the nodes of most indegrees being small. From this experiment, it can be seen that this research is applicable to large-scale networks or information systems. The second experiment mainly analyzed the changing tendency of this research as the alerts increased with the calculation being updated. The network speed is not fast enough in China and networks and servers cannot stand centralized access. Universities and colleges that have cyber security systems usually have many distribution servers for learning centers outside the campus, and they mostly use the whole structure. Some universities and colleges have servers for learning centers in the main campus and other learning centers have their own servers for learning resources. All the system backup servers have stored learning resources, including documents and video clips for use. All this led to the changing tendency of this system as as the alerts increased with the calculation being updated. The distribution of online node degree is shown in Figure 6.
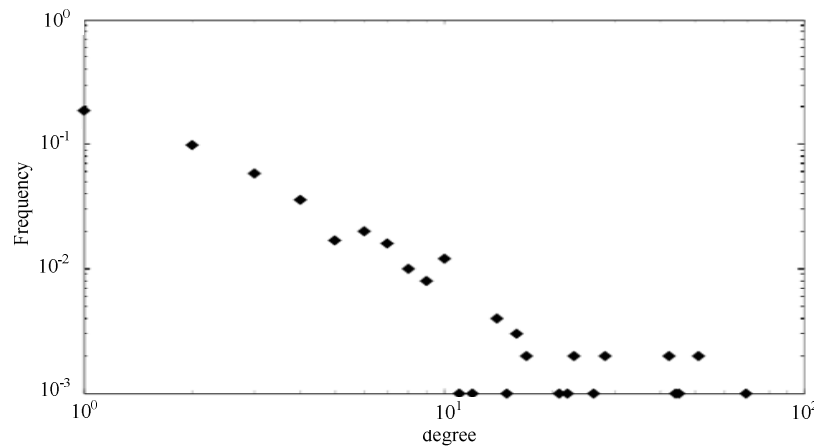

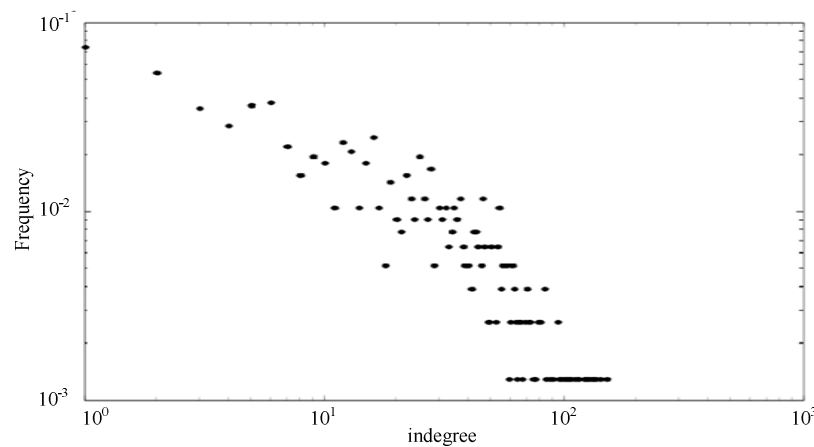**Figure 5 : Distribution of directed scale-free network indegrees**

Figure 6 : Distribution of real website users' access to the Internet

From Figure 5 and 6, it can be seen that there is a large difference of node degrees in Price network and responded by relative appearance model. This kind of request needs Server Scope to show the given information with its service statues

unchanged. Another kind is system configuration request, which requires Server Scope to change relative configuration. Because the design of configuration model comes from service model and this kind of request needs to change service status. The pagerank correction algorithm is used to work out the pr value of each user and the pr value, instead of the indegree of nodes, is used for evaluation standard for the level of attention of users.

CONCLUSION

This thesis conducted simulation experiments on all parameters in all the models. By thoroughly analyzing the experimental data, it shows that the models made by this thesis can reflect the evolution rules of public opinions on the Internet. The monitoring and warning of public opinions on the Internet has been the targets of experts and scholars in this industry since it posed a serious challenge for the Internet as the third media. Currently, the research in this field has scored some achievements. The making of this model has practical guidance for monitoring netspeak and public opinions on Weibo.

REFERENCE

[1] Qiu Chenzi; "A research on the analysis model of the increase of hotspot in online opinions on weibo" [J], *Dalian Maritime University*, **05(10)**, 102-104 **(2013)**.

[2] Ren Wei; "A research on the characteristics of social network based on weibo"[J], *Southwest University*, **05(11)**, 207-209 **(2014)**.

[3] Xiong Huihui; "A research on the spreading mechanism of weibo and blog based on complex network—a case study of sina weibo" [J], *South China University of Technology*, **06(12)**, 70-73 **(2012)**.

[4] Zhang Ting, Wei Hui; "The empirical study of the monitoring of the public sentiment about language on the internet"[J], *Applied Linguistics*, **05(3)**, 151-155 **(2011)**.

[5] Tian Zhanwei; "A study on the spreading of weibo message based on complex network"[J], *Harbin Institute of Technology*, **12(4)**, 51-55 **(2012)**.

[6] Xu Tenglong; "A study on the spreading model of weibo message based on complex network" [J], *East China University*, **01(l)**, 44-47 **(2013)**.

[7] Liang Hong; "A research on the fitness model of weibo's complex network"[J], *Beijing University of Chemical Technology*, **05(6)**, 67-69 **(2013)**.

[8] Wu Zhu, Hou Xiangyang, Xu Teng; "The application of complex network theory in the military domain" [J], *National Defense Science & Technology*, **10(24)**, 7-10 **(2010)**.