Design of a Trust Model and Finding Key-Nodes in Rumor Spreading Based on Monte-Carlo Method

Lina Wang*†, Yuntao Yue*, Chi Guo*‡, Xiaoying Zhang*
*Computer School of Wuhan University, Wuhan, P.R.China, 430079
E-mail: guochi[at]mail.whu.edu.cn
†The Key Laboratory of Aerospace Information Security and Trust Computing of Ministry of Education, Wuhan University, Wuhan, P.R.China, 430079
E-mail: lnwang[at]whu.edu.cn

Abstract

This paper combines the results of research on social psychology, and has designed a trust model for rumor spreading. It is considered that when information exchanges between people, the trust of information is related to the interpersonal closeness. In addition, this paper uses Monte-Carlo method to find the key source nodes in rumor spreading by comparing the total number of spread nodes and spreading time. We find that the key nodes that impact the rumor spreading are not necessarily those with higher degree or betweeness. Our work will provide algorithm support for maintaining the credibility of network information and ensuring the security of network information content.

Keywords: *trust model, rumor spreading, key nodes, Monte-Carlo method*

1. Introduction

With the development of computer networks, communication between people is more and more convenient, and the spread of public opinion on computer networks has a bigger impact to the whole society [1]. In such a computer-based public opinion spread network (such as E-mail network, immediate message network, etc.), the function and influence of each node is different from each other [2]. Some nodes are key nodes in the spread of public opinion. Compared with ordinary nodes, they can spread public opinion to more nodes in a short time. In order to maintain the credibility of network information and ensure the security of network information content, it is crucial to find out where these key nodes are.

The spread of public opinion based on computer networks has the following characteristics: (1) The spread network takes on the small world characteristics. Some studies showed that computer communication networks have the small world characteristics, i.e. they have a shorter average path length and at the same time have a higher clustering coefficient [3]. (2) The trust degree between friends varies from each other. In the spread network of public opinion, the credibility of public opinion content is closely related to its source. The closeness between friends is different, so the trust degree between friends is also different, which impacts the spread of public opinion (3) Everyone's personality is different. The packets sending in computer networks are passive, but the spread of public opinion is active. Whether to spread or not depends on the personality of each individual. Some like to spread the news, while others are more cautious, who normally accept the news passively [4]. Therefore, the spread of public opinion is uncertain to some extent, and it can not be simply analyzed from the network topology. (4) The information has the fresh feature. Public opinion, especially those spread on network tend to be fresh, that is, the attraction of public opinion to individuals is decreasing as time passes. This phenomenon has an impact on the spread of public opinion.

Therefore, the spread of public opinion has the feature of randomness, uncertainty and dynamicity. The traditional researches on the spread of public opinion usually analyzed the spread model of public opinion and the trend of spread speed [5][6]; there is rarely a method about the excavation of network key nodes [7]. This paper presents a new method targeted at finding out the key nodes.

Because many characteristics of the spread of public opinion are random, we adopt Monte Carlo

‡ Corresponding author

sampling method. The principle is that the probability of an event can be estimated by its frequency of occurrence in a large number of trials; when the sample size is large enough, the frequency can be taken as its probability of occurrence [7]. We obtain the expectation of the performance of public opinion spread through a large number of simulation experiments samples.

The spread of public opinion is related to the scope of social psychology to some extent. The related research has given us much inspiration, particularly when analyzing the closeness of interpersonal relationships. In the following, we will present a mathematical model portraying the human relationships.

This paper is organized as follows: in the second section, we talk about the design of trust model in public opinion spread; then we discuss the spread simulation algorithm design based on Monte Carlo method; after that, the data of simulation experiments is analyzed; and finally, a brief summary is given in terms of our works and we point out the future work.

2. Trust model design

The model includes four parts: (1) A public opinion spread carrier network with small-world characteristics; (2) The model of closeness between friends, which determines that the same message brings out different influence when spread by different friends; (3) The preferred model of initiative spread of public opinion, it reflects whether the individual would like to spread public opinion actively; (4) The attenuation of public opinion.

2.1 A carrier network of public opinion

Newman et al. have done a wide range of empirical research on a large number of actual network topology characteristics [8]. They pointed out that the relationship network depended on computer network has an obvious small-world effect, such as in the e-mail network, the average degree $\langle k \rangle = 3.38$, the average path length $L = 5.22$, the clustering coefficient $C = 0.13$ .This paper adopts the WS small world network as the spread carrier, whose network topology feature is similar to the email network. The WS small-world network model is proposed by Watts and Strogtz in 1998. It is constructed by re-connecting edges randomly by the probability of p in a regular graph. The WS small-world network has short average path length as well as high clustering coefficient.

Using the WS small-world network as a spread carrier can simulate the spread of public opinion better in the real relationship network, and it can also excavate the key nodes in the spread of public opinion effectively and credibly. Each node in the network topology represents an individual in the spread network.

2.2 The closeness between friends

The sociologists and psychologists have done a lot of research on the closeness of interpersonal relationships. IOS (Inclusion of Others in Self scale) Scale [9] is a good tool to measure the closeness relationships. As shown in Figure 1, IOS scale is constructed by seven pairs of double circle with increasing overlapping level, forming a seven-point equidistant scale. The bigger the overlap between two circles is, the closer the relation between self and the object will be.

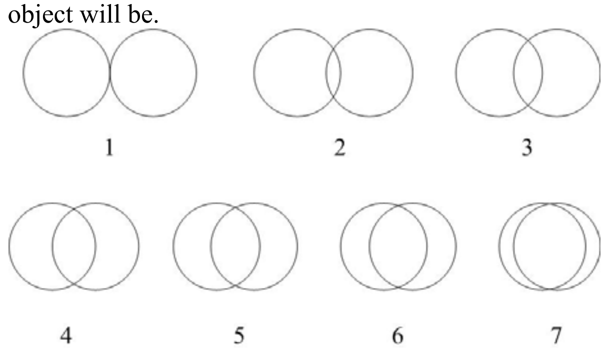


Figure 1 *The relationship between IOS scale and interpersonal closeness.1 indicates the relationship is cold, and 7 indicates closed. The closeness is increasing from 1 to 7.*

Social psychology research showed that the distribution of each person's friends belonging to these seven closeness degrees is non-uniform [10].The closeness distribution can be taken as a $\beta(a,b)$ distribution similarly, as shown in Figure 2.

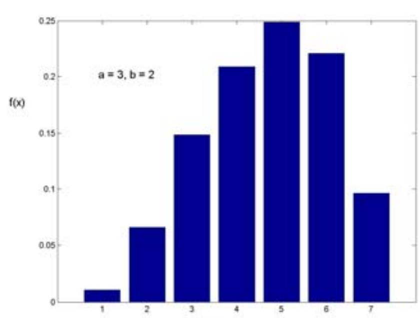


Figure 2 *Number distribution of friends belonging to different degree. After many experiments, the parameters of the β distribution is set to $a=3, b=2$, the distribution matches the results in* [10] *well.*

Assumption 1: In order to explore the effects of closeness on the spread of public opinion, we assign the weight w_{ij} on each edge e_{ij} in the network topology. w_{ij} indicates the closeness of node *i* and node *j*, the relationship will be closer when the weight is greater. We assign w_{ij} as 1-7 positive integer referring to the IOS scale, the distribution of w_{ij} in the entire network meets $\beta(3,2)$

Assumption 2: In the process of public opinion spread, assume that node *i* transmits a message to node *j*, then the trust to the message of node *j* is:

$$m_{ij} = \frac{w_{ij}}{\sum_j w_{ij}} \quad (1)$$

When the node keeps receiving the same message many times, the trust to the message accumulates continually, which is also normal in the spread of public opinion.

2.3 Design of spread critical value distribution

In the spread of public opinion, the personal character is also a major factor. Some people like to spread the news while some people do not. We use a threshold δ to reflect this distinction. For an individual *i*, if its accumulated receiving trust to a message $M_{ji} = \sum_j m_{ji} > \delta_i$, we consider it chooses to spread this message to its own friends; otherwise, it does not.

Considering those who love and those do not like to spread the news very much at two extremes are minority, in our study, we use Poisson distribution $\Pr(\lambda)$ to describe individual threshold distribution of the entire network. Since in the relationship network, even those who do not like to spread the message very much will communicates with people sometimes, we normalize the common Poisson distribution to a set of real number Ω, and define the minimum of Ω is 0 and the maximum is $P\_\max$, $P\_\max < 1$.

Definition 1: Consider the most common situation, those whose spread threshold is the mean value of the Poisson distribution should be act normally in spread of messages, they should receive a certain amount of information and then spread out. We describe this situation with a percentage φ, that is, if everyone has X friends, after he receives the messages of $x = X * \varphi$ friends, he will spread the messages out.

In addition, assume that these people are the most common network nodes, that is the degrees of these nodes equal to the average degree of the network $\langle k \rangle$, and the weights of edges connecting these nodes are average weight $\langle w \rangle$ of the network. Then the mean value P_avg of the normalized Poisson distribution should be

$$P\_avg = P\_\max/2 = \frac{x \cdot \langle w \rangle}{\sum w_i} = \frac{x \cdot \langle w \rangle}{X \cdot \langle w \rangle} = \varphi$$

so $$P\_\max = 2\varphi \quad (2)$$

The spread threshold of each individual is

$$\delta_i = \Pr(\lambda) \cdot k = \Pr(\lambda) \cdot (P\_\max/\max(\Pr(\lambda))) = \frac{2\varphi \cdot \Pr(\lambda)}{\max(\Pr(\lambda))} \quad (3)$$

2.4 Fresh and trust attenuation of public opinion

Definition 2: As the public opinion has the feature of fresh, the attractiveness to the individual is decreasing over time. We introduce an attenuation coefficient γ to describe this trend. In each shot, the cumulative trust $M_i^{(t)}$ of each node which has received the public opinion will attenuate:

$$M_i^{(t)} = M_i^{(t-1)} \cdot \gamma, 0 < \gamma < 1 \quad (4)$$

3. The simulation of rumor spreading based on Monte-Carlo method

Definition 3: For each node in the network, we define a state S to indicate whether the node has received the information. If $S = 1$, we consider the node has received the information, while $S = 0$ stand for the node has not received the information.

Definition 4: For each node in the network, we define a state R to indicate whether the node has spread the information. If $R = 1$, we consider the node has spread the information, while $R = 0$ stand for the node has not spread the information.

Definition 5: For each node in the network, we define a neighbor set Γ, that is, if there is an edge e_{ij}, then join node *j* into set Γ_i.

Assume that information received from the first node is st, and it will certainly spread the information. Then the simulation algorithm of rumor spreading is described as below.

1. Set value for each edge of the network, and make sure the distribution of the weight of all edges meets Poisson distribution.

2. Set spreading threshold for each node in the network according to formula (3), and set S and R to 0 for each node.

3. Define two node sets: A, B

4. Set time $t = 0$, and the total number of the nodes which has received the rumor $Num = 1$

5. Set $S_{st} = 1, R_{st} = 1$, and join st into set A

6. Get element *i* from set A

7. Get element *j* from set Γ_i

8. If $R_j = 0$, then update the trust degree m_i of node *i*, according to formula (1).

9. Set $S_j = 1$

10. If $m_i > \delta_i$ and $R_j = 0$, then join j into set B, and set $R_j = 1, Num = Num + 1$.

11. If $\Gamma_i \neq \varnothing$, then turn to step 7

12. If $A \neq \varnothing$, then turn to step 6

13. Set $A = B, B = \varnothing$

14. Attenuate the trust degree of the node whose $S = 0$ according to formula (4)

15. Set $t = t + 1$.

16. If the number of the nodes whose $S = 0$ is 0, or $A = \varnothing$, then quit.

17. Otherwise, turn to step 6

In this way, after a simulation of the spreading, we would have the total number Num of nodes that can get the rumor and the spread time t when the rumor is spread form node st. We hope to find some key nodes, which the public opinion can be spread to more nodes in less time.

We use Monte-Carlo method to simulate the spreading that begins from all the nodes respectively for many times, and use the average of the total number Num and the time t as the desired outcomes.

4. Data analysis of the simulation result

4.1 Parameter description

We use the topology generators, set the re-connect probability $p = 0.08$, and generate network topology of WS small-world with 100 nodes. The topology characteristics are as below: the average degree $\langle k \rangle = 3.35$, the average path length $L = 5.13$, the clustering coefficient $C = 0.14$, which is basically similar to the real email network in [8].

In our experiment, the parameters of β distribution is fixed as $a = 3, b = 2$. We discuss the key factor impacting the spread of public opinion and find out the key nodes by changing other parameters.

4.2 Experimental results

4.2.1 Identify the key nodes

When the public opinion is spread from a particular node *i*, we use the total number of nodes N received from the public opinion in the network and the spread termination time t to describe the spread performance of the node. We need to find out the key nodes influencing the public opinion spread, which satisfy the bigger N and smaller t relatively.

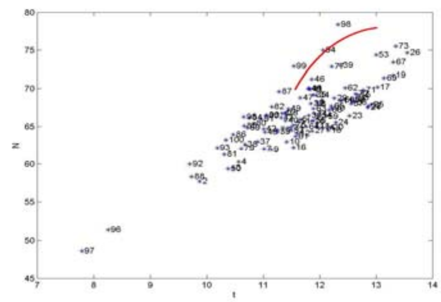


Figure 3 *Comparison of the $N-t$ performance of all nodes in the network. The horizontal coordinate is time t, the vertical coordinate is nodes number N, each node is marked with its number order, and the parameter is set as* $\lambda = 5, \varphi = 0.3, \gamma = 0.9$.

Figure 3 is the simulation results for the experiment. We can see the spread total number of nodes N and the spread time t vary from different network. The nodes covered by the solid line in Fig.3 can be considered as the key nodes in the spread network, such as No.94,98,99.

4.2.2 The impact of the critical value distribution parameter on the spread of public opinion

Figure 5 shows the impact of the changing parameter λ of Poisson distribution on the spread.

We can see that as λ decreases, the public opinion can be spread to more nodes. As shown in Figure 4, this is because when λ decreases, the peaks of Poisson distribution become higher, and less nodes have a high spread threshold. In this way, it is easier for the public opinion to spread in the network.

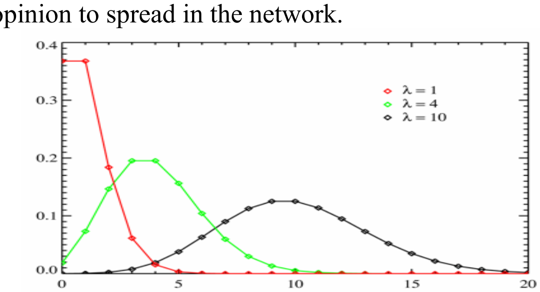
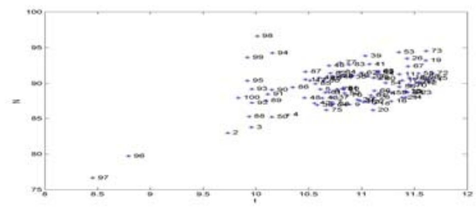
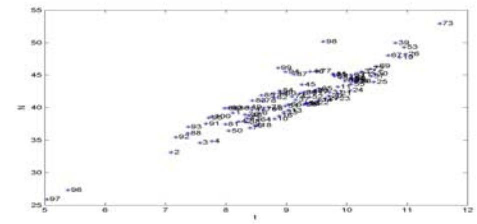


Figure 4 *Poisson distribution*[11]



(a)

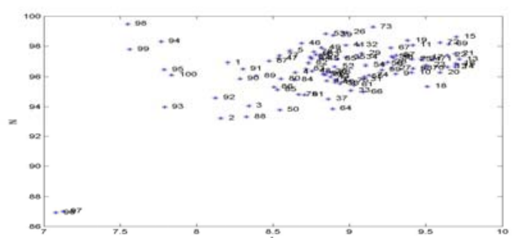


(b)

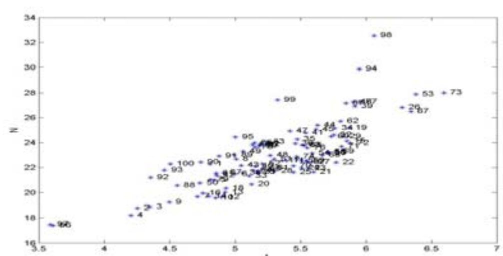
Figure 5 *The impact of Poisson distribution on the spread of public opinion. Parameter settings*: $\varphi = 0.3$, $\gamma = 0.9$ (a) $\lambda = 3$, b) $\lambda = 7$.

4.2.3 The impact of control parameter φ on the spread of public opinion

Figure 6 shows the impact of the change of control parameter φ on the spread of public opinion



(a)



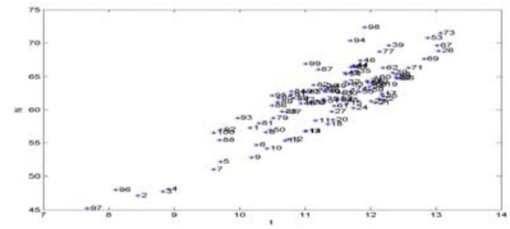
(b)

Figure 6 *The impact of the change of control parameter φ on the spread of public opinion. Parameter settings*: $\lambda = 5$, $\gamma = 0.9$ (a) $\varphi = 0.2$, (b) $\varphi = 0.4$

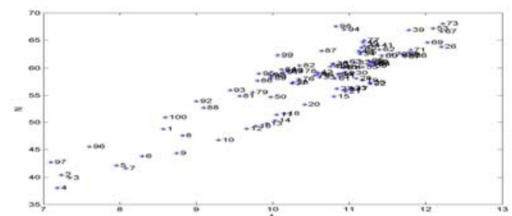
As shown, with the increase of φ , the public opinion can be transmitted to fewer and fewer nodes. This is because when φ becomes greater, the nodes will need to receive more times before spreading the public opinion, so it increases difficulties for public opinion to spread in the network.

4.2.4 The impact of attenuation coefficient on the spread of public opinion

Figure 7 shows the impact of the change of attenuation coefficient γ on the spread of public opinion.



(a)



(b)

Figure 7 *The impact of the change of attenuation coefficient γ on the spread of public opinion. Parameter settings*: $\lambda = 5$, $\varphi = 0.3$ *(a)* $\gamma = 0.7$, *(b)* $\gamma = 0.5$.

We can see that the smaller the attenuation coefficient γ is, the fewer nodes the public opinion can spread to. This is also consistent with the definition of attenuation coefficient, if the trust of individual to the public opinion declines faster, the public opinion is more difficult to spread in the network.

4.2.5 The characteristics of the network topology of key nodes

In the previous study, as the experiment parameters change, the total spread nodes number N and the spread time t of a single node also change. However, the key nodes remains fixed, which proves that our method is universal.

Table 1 *The characteristics of the network topology of key nodes*

Node	degree	Order of degree	betweeness	Order of betweeness
39	8	7	528	36
77	8	7	1412	7
98	12	1	2342	1
99	10	2	1633	3

As shown in table 1, we sort all nodes in the network topology according to degree and betweeness, we find that the degree and betweeness of those key nodes are not the highest. This shows that the key nodes which influence the public opinion spread are not entirely dependent on the static network topology characteristics, which means our approach has certain advantages compared to the traditional method in finding out the key nodes based on network topology.

5. Conclusion

This paper proposes a method to find the key nodes in the rumor spreading of computer networks, We find that the key nodes that impact the rumor spreading are not all have higher degree or betweeness. Our approach combines the research of social psychology, and designs a trust model for rumor spreading. We discuss the impact of each control parameters to the rumor spreading. Next, We will research the game condition of both side of rumor spread in networks when the negative public opinion is led into considered.

6. Acknowledgment

This work was funded by the major project of Chinese National Natural Science Foundation (90718006), the Chinese National Natural Science Foundation of China (60743003), and the Ph.D. Programs Foundation of Ministry of Education of China (20070486107).

7. Reference

[1]. Xide Qin, Dejiang Rao. "Public Opinion Course". *Wuhan: Wuhan University Press*,1998.

[2]. Junyan Su, Lingjiang Kong, Muren liu. "Opinion Dynamic Model Based on Weighted Networks". *Journal of Guangxi Normal University*, 2006, **24**(2), 1-4.

[3]. Watts, D.J, Strogatz, S.H. "Collective dynamics of 'small-world' networks". *Nature*,1998,**393** (6684): 409–10

[4]. Qingfeng Wu, Lingjiang Kong, Muren Liu. "Influence of person's character upon the evolution of the cellular automata model for public opinion".*Journal of Guanxi Normal University*.2004,**22**(4),5-9

[5]. Zhaofeng Pan, Xiaofan Wang, Xiang Li. "Simulation Investigation on Rumor Spreading on Scale-free Network with Tunable Clustering".*Journal of System Simulation*,2006,**18**(8),2346~2348

[6]. Changyu Liu, Xiaofeng Hu. "Public Opinion Propagation Model Based on Small World Networks". *Journal of System Simulation*.2006,**18**(12),3608-3610

[7]. E. Zio, C. Rocco, "Security Assessment in Complex Networks Exposed to Terrorist Hazard: A Methodological approach based on Cellular Automata and Monte Carlo Simulation" *Complex Network and Infrastructure Protection* (CNIP 2006),2006, Rome, Italy

[8]. Newman M E J. "The structure and function of complex networks". *SIAM Review*,2003,**45**:167~256

[9]. Aron A, Aron EN, Smollan D. "Inclusion of others in the Self Scale and the structure of interpersonal closeness". *Journal of Personality and Social Psychology*,1992,**63**:596-612

[10]. Li H Z. "Culture, gender and self-close-other(s) connectedness in Canadian and Chinese samples". *European Journal of Social Psychology*, 2002, **32**(1): 93~104

[11]. Poisson distribution. *Wikipedia*: http://en.wikipedia.org/wiki/Poisson_distribution