

The Simulation of Rumor's Spreading and Controlling in Micro-Blog Users' Network*

Peng Chen¹, Nan Gao²

¹Security Department, Chinese People's Public Security University, Beijing, China; ²Computer Science and Engineering Department, South University of Carolina, Unite States.
Email: uctzpch@gmail.com

Received 2013

ABSTRACT

In this work, a rumor's spreading and controlling in a directed Micro-blog user network being consisted with 580 000 nodes are simulated. By defining some authority nodes that release anti-rumor information as the prevention strategy, the effect of the nodes' role in network on rumor's suppression is studied. The findings show that rumor will be spread out fast and reach a stable level within limited steps. The suppression of rumor is more predominated by the intervening opportunity, the earlier the intervention strategy was implemented, the better the rumor's controlling could be achieved. The controlling effect is less relevant with the role of the authority nodes in network.

Keywords: Component; Rumor; Micro-Blog; Network; Simulation

1. Introduction

As a fast developed production in the web, Micro-blog is playing a more and more important role in Internet network. Each day billions of blogs are created, commented and transferred by the users. The Micro-blog is so welcomed for its highly convenience in information spreading especially in some disaster emergency responses the people could distribute the instant messages on the web so that public can quickly know what happened in the field and then make decisions of what to do to cope with the problems. However, accompanying the advantages of the Micro-blog also comes out some negative aspects. For example, if someone releases misleading information or rumor on Micro-blog then many people will be influenced through the connections between the users.

As a matter of fact, the study about rumor of information spreading in the social network was researched by some people previously. For example, Sudbury, Zanette and Zhou made simulation about opinion or news spreading in complex social network and they found that as most as 80% of nodes could receive the news in random connected structure [1-3]. Xu investigated how rumor spread through blog and his study indicated that infection rate has significant influences on rumor's spreading [4]. Additionally, by constructing a framework of identification algorithm, Liu studied detection of opinion leaders in Micro-blog and he found that most opinion

leaders are topic dependent [5]. These previous theoretical and simulation work concentrate mainly on the relationship between rumor spreading and topology of the social network, but little was done to study the preventing method or suppression strategy. So, in this work, by collecting data of Micro-blog users' attributes we constructed their social relationship network and made simulations to study the rumor's spreading and controlling with the aim of finding out an effective prevention strategy.

2. Model and Data

The model SIR is the one being used broadly in rumor spreading study. By definition it has three subgroups: spreader, ignorant, stifler. When a spreader contacts with an ignorant the latter becomes a spreader with probability α , and when a spreader contacts another spreader or a stifler the initiating spreader becomes a stifler at a rate β . But considering the rumor's spreading in Micro-blog is such a quick process that many spreaders are not going to become stiflers until they know the truth, we assume that spreaders maintain the infection state for a long time. Therefore, the spreading process is shown in the following way:

$$S(i)+I(j) \rightarrow S(i)+S(j) \quad (1)$$

While the process is totally dominated by parameter α .

But for the spreaders changing into stiflers, we assume they will become the stiflers 100% when they receive the information released from the authority node. Thus this

*This paper is sponsored by Natural Science Project of Chinese People's Public Security University (No.2011LG12)

process could be shown in this way:

$$S(i)+ST(j) \rightarrow ST(i)+ST(j) \quad (2)$$

So in the work two information transferring processes are included, one is rumor spreading originated randomly from a node and the other is the truth spreading originated from an authority node. We study the rumor prevention strategy by examining the interaction between two processes. If the truth spreading is over the rumor spreading we call it success prevention.

The social relationship data between the users was collected from SINA website which is the largest Micro-blog server in China currently. A program was coded and run at background to download the user's personal information automatically. This work started at 01/02/2012 and lasted to 29/02/2012 and more than 580 000 users' recordings were collected. In the data each user has an ID number to represent his identity.

In Micro-blog network the relationship between any pair of nodes are unequal because not all nodes who follow the others with interest can receive equal treatment. So we call the one who following others with interest the 'fan'. The relationships between two nodes are demonstrated in **Figure 1**. From the graph it can be known that when two nodes are fans to each other, the connection is non-directed. However, if only one node is the fan to the other the connection is directed. So using this relationship information a social network between Micro-blog users is constructed.

3. Definition

A sample of 50 000 user's private information was collected to determine the infection probability α . Their visible attributes include gender, age, education and registration. The details of these attributes are listed below:

- Gender: male/female
- Registration: yes/no
- Age: lower than 20/20-30/30-40/larger than 40
- Education: elementary school/junior middle school/senior middle school/university

The percentage of the users by the attributes is calculated and the result is shown in **Table 1**. From the findings we can know that predominant users are male, non-

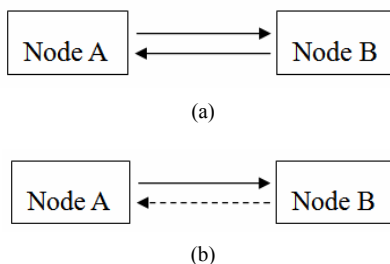


Figure 1. The relationship between two nodes.

registered, higher educated and aged between 20 and 30. So with this distribution of statistic we define a kind of conservative index to represent the users' responding attitudes towards the rumor.

Hypothesis 1: females are more conservative than males, so females' conservative index is 1.0, males is 0.5.

Hypothesis 2: the users who registered has better credit than those don't, so registration ones' index is 1.0, non-registration ones' is 0.5.

Hypothesis 3: the people having better education background are inclined to be more conservative, so the index for people who had been in college is 1.0, senior middle school is 0.75, junior middle school is 0.5, and elementary school is 0.25.

Hypothesis 4: aged people are more conservative than younger ones, so for people who aged more than 40 the index is 1.0, 30-40 is 0.75, 20-30 is 0.5, and lower than 20 is 0.25.

With the proportion of people by their attributes, the infection probability is calculated with formula (3):

$$\alpha = \sum A_i p_{ai} + \sum E_i p_{ei} + \sum G_i p_{gi} + \sum R_i p_{ri} \quad (3)$$

Where A_i, E_i, G_i, R_i are the conservative indexes for age, education, gender and registration attributes. p_{ai}, p_{ei}, p_{gi} and p_{ri} are percentages of the users belonging to different attribute group (see **Table 1**). Based on the formulation the calculated infection probability is 0.35.

4. Simulation and Findings

As the relationships between two nodes are unequal, the Micro-blog user network is a directed connected structure. The node following others with interest is a fan, so its connectivity is out-degree. While for the node that is being followed, its connectivity is in-degree. The distribution of in-degree and out-degree to the nodes are calculated and the relationship between their frequency and scale are demonstrated in **Figure 2**. It can be seen that the out-degree shows a linear trend in log-log plot, while the in-degree shows a non-linear one.

The nodes' in-degree is more important in information transferring than out-degree because the in-degree indicates that the node has so many fans so it could influence its neighbor nodes directly by releasing information. In contrast, out-degree only measures the ability of the node receiving information from the others. In this case, a

Table 1. Percentage Ofusers by Attributes

	Gender	Aged	Registration	Education
		< 20	13.4%	ES
Male	72.7%	20-30	73.1%	Yes
Female	27.3%	30-40	11.8%	No
		> 40	1.7%	80%
				SS
				C
				1.3%

ES- Elementary school; JS- Junior middle school; SS- Senior middle school; C- College

node with maximum fans was selected as the source of the rumor and the spreading process was simulated. The results are displayed in **Figure 3** and each data was the average of 50 repeated simulations. From the picture it can be known that for the network that is consisting with nearly 600 000 nodes the count of infected nodes quickly reach a stable level within 20 time steps, or in other words, the infected nodes reach a maximum level (16 496 nodes, about 2.8% in total,) within limited steps.

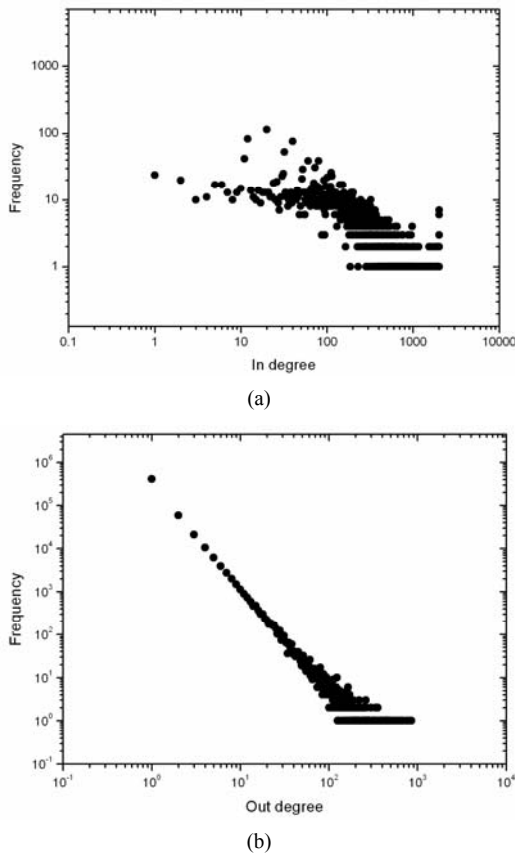


Figure 2. The relationship between node connectivity and frequency in Micro-blog user network. (a) in-degree; (b) out-degree.

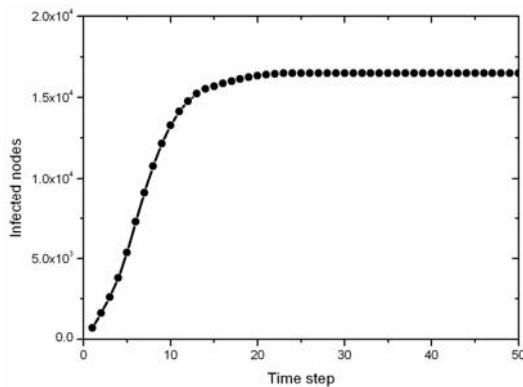


Figure 3. The variation of infected nodes along the timeline.

As a way to prevent the rumor spreading, we introduce the prevention strategy into the simulation. We hypothesize there is an authority node which is playing the role of the information monitoring. Once upon a time a rumor is spreading is discovered in network it will instantly release the truth to combat with the rumor. The infected nodes contacting with the truth will become an ignorant node and never be infected by rumor again. Initially, an authority node with a certain number of fans was selected and rumor's spreading and truth diffusion process are simulated. The results are shown in **Figure 4**. To investigate the effect of the suppression we implement the prevention strategy at different opportunity. It can be seen from the graph that rumor spreading is suppressed successfully, and the earlier the intervention was implemented the better the rumor spreading could be stopped. The advantage of earlier intervention is not only effectively shortening the rumor spreading period but also significantly reduce the infected nodes.

Finally, we further investigate the significance of authority node on rumor prevention. We defined 5 authority nodes with different number of fans that representing their different status. The more fans the node owning the more important the nodes plays in the network. Then simulations were made and results are demonstrated in **Figure 5**. The results showed that no matter how fans count changed the nodes infected by rumor did not vary too much. Combined with the results in **Figure 4** we could find that the significance of authority node in network has little influencing on rumor's controlling, the intervening opportunity is more predominant.

5. Conclusion

In this work, we simulate rumor's spreading and controlling in Micro-blog user network by using a revised SIR model. By collecting user's social relationship and privacy information from online website the topology of the network and spreading rules were defined. The findings of the research demonstrate that in a Micro-blog user

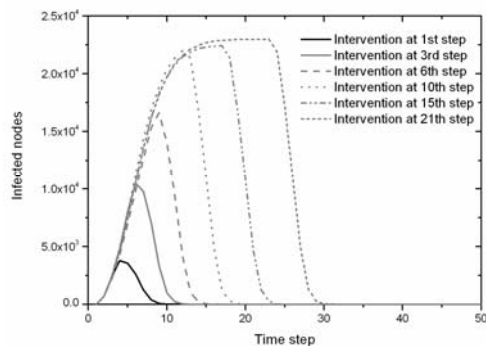


Figure 4. The suppression to the rumor by implementing prevention strategy at different opportunities

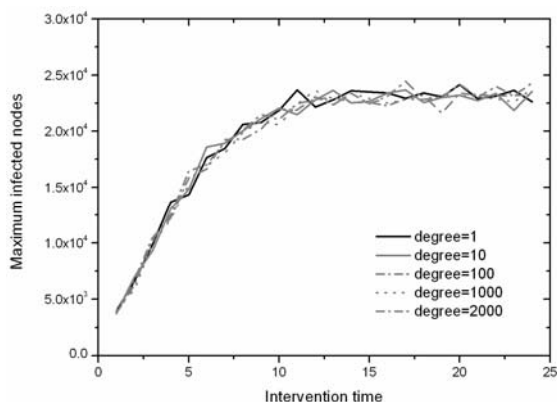


Figure 5. The suppression to the rumor by intervention from the nodes with different in-degree connectivity.

network in which the nodes are directed connected the rumor could quickly spread out and reach a top level within limited steps. Implementing a prevention strategy that releasing truth from an authority node is an effective way to suppress the rumor spreading, but the intervening opportunity is a very significant factor determining the rumor's controlling. The earlier the intervention was im-

plemented, the better the rumor could be suppressed. Additionally the influence of the authority nodes with different number of fans on rumor's spreading and controlling are studied and the results show that the significance of the authority node in network has little relationship with the controlling effect. This study provides a clue that however the authority node's role in network, supervising the information spreading in Micro-blogs and making intervention in time is an effective way to rumor's suppression.

REFERENCES

- [1] A.J.Sudbury, Appl. Prob., 22(1985)443.
- [2] D.H. Zanette, Phys. Rev. E, 64 (2001) 050901
- [3] J. Zhou, Z. Liu and B. Li, Phys. Lett. A, 368 (2007) 458
- [4] Xu Xiaodong, Xiao Yintao, Zhu Shirui. Simulation investigation of rumor propagation in micro blogging community. Computer engineering. 2011, 37(10): 272-274
- [5] Liu Zhiming, Liu Lu. Recognition and analysis of opinion leaders in Microblog public opinions. System Engineering. 2011, 29(6): 8-16