

## Research Article

# Model Study of Information Dissemination in Microblog Community Networks

Bo Wang,<sup>1</sup> Jianhua Zhang,<sup>2</sup> Haifeng Guo,<sup>3</sup> Yuwang Zhang,<sup>1</sup> and Xiaotuo Qiao<sup>3</sup>

<sup>1</sup>School of Humanities and Social Sciences, Harbin Institute of Technology, Harbin 150001, China

<sup>2</sup>Northeast Petroleum University, Daqing 163318, China

<sup>3</sup>School of Management, Harbin Institute of Technology, Harbin 150001, China

Correspondence should be addressed to Haifeng Guo; guohaifenghit@126.com

Received 15 August 2016; Accepted 20 September 2016

Academic Editor: Allan C. Peterson

Copyright © 2016 Bo Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We build an information dissemination model based on SIR model to study information dissemination in microblog networks. We consider different influence factors of information dissemination such as activity, credibility, and weight of network and construct calculation methods of various parameters, for instance, direct immune rate, indirect immune rate, and information dissemination rate. Meanwhile, by collecting data from API in Weibo and using the result of microblog information dissemination life cycle analysis, we utilize the model to conduct simulation and get the change trend of proportion in Stages S, I, and R. After comparing with the actual situation, this model is proved to be effective in predicting the trend of information dissemination.

## 1. Introduction

Microblog, as an important part in our modern life, is a platform for information sharing and dissemination on the basis of web technology [1]. The first microblog community, Twitter, was founded in 2006 by Evan Williams. Tweet dissemination has the characteristics of fission and polymerization, which are different from traditional media [2–5]. Twitter mainly concerns the fields of e-learning, cultural exchange and information feedback, and so forth [6, 7]. Founded in 2010, Weibo develops rapidly and is becoming a representative of new social media in China. In 2014, it had more than 500 million users and became China's largest virtual community. In China, Weibo community is a popular research object. In most researches, Chinese scholars tend to select model study since it simulates the trend of information dissemination concisely.

There are mainly four types of dissemination model: independent waterfall model [8], linear threshold model [9], game theoretical model [10], and epidemic model. Among them the epidemic model is widely applied to empirical research on information dissemination because of the similarity between information dissemination and epidemic disease spreading

[11–13]. Classical epidemic model mainly includes SI model, SIS model, and SIR model. SIR model is the most optimal one among the three models. SI model and SIS model only include infected stage (i.e., Stage I) in which members who have been infected will continue to spread the disease. SIS model assumes that the susceptible population can be cured and they infect and spread the viruses at the same time. The hypotheses and transmission mechanism of SI model and SIS model are different from the general rule of information dissemination, while SIR model well maps the objective conditions and the process of information dissemination.

The SIR model is one of the most widely used epidemic models; many scholars studied the information dissemination of microblog by optimizing and improving the SIR model. Moreno and Vázquez [14] profoundly observed and studied dissemination behaviors from different scale-free networks and drew a conclusion that different information dissemination models and initialization conditions will produce different effects. Li and Sun [15] established information dissemination evolution model based on the SIR model and studied the features of social populations when two pieces of conflicting information spread simultaneously in social network. Different from diseases spreading, containment

system and forgetting mechanism will affect information dissemination [16–18]. Based on what was mentioned, Wang et al. [19] improved the SIR model considering containment system and forgetting mechanism and then analyzed their influence on information dissemination process using the improved model. The SIR model study of May [20] shows that in scale-free network, when scale of network is infinite, the threshold value of information dissemination will drop to zero. These studies made certain theoretical contribution by introducing the epidemic models to the study of information dissemination and became the fundamental of microblog information dissemination study.

In recent years, microblog information dissemination studies have achieved appreciable progress [21–23]. Since microblog information dissemination belongs to a kind of information dissemination, SIR model has been gradually introduced to the study of microblog information dissemination. Xu et al. [24] built a microblog community information dissemination model considering the multisubject feature of microblog community, constructed a microblog community information dissemination model, and studied factors that influence the popularity of microblog. Ding et al. [25] built a modified SIR model based on the direct immunization situation. After investigating the features such as the existing memory efficiency, social consolidation, and nonredundancy of interrelationship of the microblog information dissemination progress, Guo [26] explored the relations between microblog's coverage rate and relevant factors. Taking low activity character and reliability factors into account, Wang et al. [27] established microblog information dissemination SIR model and performed simulation and prediction. In consideration of the network weight's influence on information dissemination rate, Yan et al. [28] redefined information dissemination on the basis of the epidemic disease model. The scholars above studied SIR model in the areas of microblog information dissemination process and provided reference meaning for this paper.

According to the current studies, SIR model has been widely used in information dissemination model and gradually introduced to microblog information dissemination studies. However, the life cycle feature of microblog information dissemination and the differences between microblog information dissemination and epidemic diseases spreading are rarely studied. In addition, seldom scholars have tried to collect and calculate data directly and conduct the stepwise simulation. Therefore, this research aims to construct microblog information dissemination model based on SIR model after comparing the differences between microblog information dissemination and epidemic diseases spreading. This model combines factors such as activity, reliability, and life cycle, describes the information dissemination progress by collecting data, calculating parameters, and stepwise simulation, and provides theoretical foundation and reality tool for the following researches.

The structure of this paper is arranged as follows. Section 2 introduces SIR model to the study of microblog information. Section 3 constructs the initial information dissemination model after comparing the differences between the direct immune rate and indirect immune rate of microblog

information dissemination. Section 4 optimizes the SIR model considering the low activity feature of microblog and reliability. Section 5 conducts simulation according to the established model. Section 6 is the conclusion.

## 2. Methodology and the SIR Model

*2.1. Applicability of the SIR Model.* In SIR model, participants are divided into three parts: Stage S is the susceptible population that have probability of infecting virus, Stage I is the infected population that transmits virus to Stage S with the probability of  $\lambda$ , and Stage R is the recovered population who will not transmit virus any more that transformed from Stage I with the probability of  $\mu$ . We use  $s(t)$ ,  $i(t)$ , and  $r(t)$  to represent the proportion of Stage S, Stage I, and Stage R, and the model is shown as formula (1). Consider

$$\begin{aligned}\frac{ds(t)}{dt} &= -\lambda i(t) s(t), \\ \frac{di(t)}{dt} &= \lambda i(t) s(t) - \mu i(t), \\ \frac{dr(t)}{dt} &= \mu i(t).\end{aligned}\tag{1}$$

Then, we introduce the three stages into the information dissemination model. We apply the SIR model to the informational dissemination network because the information dissemination is similar to the virus spreading and its effect of dissemination from both network structures is similar. Network structure and the evolution of the SIR epidemic model are similar to the information dissemination model. Because user types of the SIR model have a great similarity with the microblog community's and they are subjects to information dissemination characteristics of scale-free networks, SIR model can be applied to the microblog information dissemination study.

In order to facilitate the research and differ from diseases spreading, we redefine the three stages. Stage S is the spread population who are information publishers; Stage I is ignorant population who have never received the information, and Stage R is the recovered population who will no longer disseminate the information. We still use  $s(t)$ ,  $i(t)$ , and  $r(t)$  to denote the proportion of Stages S, I, and R, and the model is shown as formula (2). Consider

$$\begin{aligned}\frac{di(t)}{dt} &= -\lambda i(t) s(t), \\ \frac{ds(t)}{dt} &= \lambda i(t) s(t) - \mu s(t), \\ \frac{dr(t)}{dt} &= \mu s(t).\end{aligned}\tag{2}$$

We describe the information dissemination process in microblog community as follows: Stage S is information publishers, which accounts for a small proportion at initial time step. Stage I is followers of Stage S who never received the information. When they received information, they transform into Stage S with the proportion of  $\lambda$  and start

to disseminate. The population in this stage take the largest proportion in initial time step. Stage R is the users who have received the information, but they will no longer disseminate it. They transform from Stage S with the probability  $\mu$ , and their values are 0 in the initial time step.

**2.2. Improvement of the SIR Model.** The dissemination of microblog information relies on the behaviors of followers such as forward, comment, and like; hence when we consider microblog information dissemination mechanism, we should figure out the differences between epidemic disease spreading and microblog dissemination. In microblog information dissemination, a piece of information will be available to all the followers after posting or forwarding, but not all the followers will accept and forward the information, and part of the followers will receive but do not accept it and transform into immune population directly. Since information is valued by only certain population at certain time, part of the users who accept and disseminate the information will transform into immune population according to a certain proportion with the time step. Therefore, when we construct microblog information dissemination model, we should notice two situations: one is the susceptible population who no longer disseminate the information and transform directly into immune population; the other is the dissemination population who have disseminated the information and transform indirectly into immune population.

We describe the dissemination of microblog information as follows according to the introduced double immunization rate: part of the population in Stage S disseminate the information to susceptible population with the probability  $\lambda$ , and susceptible population start to disseminate the information. The remaining part in Stage S no longer disseminate the information and transform into immune population with the probability  $\mu$ . In Stage I, when the population receive the microblog information, part of them will choose to accept and disseminate it. The remaining part will not disseminate it and transform into immune population with the probability  $v$ . So, in the dissemination progress, part of the population in Stage I will transform into Stage S with the probability  $\lambda$ , and the remaining part of it will transform into Stage R with the probability  $v$ . We optimize the model according to these principles.

$$\begin{aligned}\frac{di(t)}{dt} &= -\lambda i(t) s(t) - \mu s(t), \\ \frac{ds(t)}{dt} &= \lambda i(t) s(t) - \mu s(t), \\ \frac{dr(t)}{dt} &= \mu s(t) + v i(t) s(t).\end{aligned}\quad (3)$$

In formula (3), we use  $s(t)$ ,  $i(t)$ , and  $r(t)$  to represent the proportion of Stage S, Stage I, and Stage R at time  $t$ .  $\lambda$ ,  $\mu$ , and  $v$  represent the dissemination rate of microblog information, the indirect immunization rate, and the direct immunization rate, respectively.

So, we describe the microblog information dissemination as follows: when Stage S contacts Stage I, part of Stage S will

disseminate information to Stage I and make them change into Stage S with the probability  $\lambda$ . The remaining part of Stage S will lose the interest in dissemination when they encounter Stage S or R during the dissemination; then they change into Stage R with the probability  $\mu$ . When Stage I receives information, part of them will choose to accept and disseminate it and then turn into Stage S with the probability  $\lambda$ . The remaining part of Stage I will change into Stage R with the probability  $v$ .

### 3. Analysis Process

**3.1. Network Weights.** The microblog community is a complex network, so the information dissemination models have some network characters. As for the uniform network, to make it clear, the mean field was introduced to the model formulated in equation. In short, mean field theory is unified handling the role of the research subject in the network environment, replacing the totaling single effect with the average effect. And the model is shown in

$$\begin{aligned}\frac{di(t)}{dt} &= -\lambda \langle k \rangle i(t) s(t) - v \langle k \rangle i(t) s(t), \\ \frac{ds(t)}{dt} &= \lambda \langle k \rangle i(t) s(t) - \alpha \langle k \rangle s(t) [s(t) + r(t)], \\ \frac{dr(t)}{dt} &= \alpha \langle k \rangle s(t) [s(t) + r(t)] + v \langle k \rangle i(t) s(t).\end{aligned}\quad (4)$$

In formula (4),  $\langle k \rangle$  is the average degree of the network,  $\lambda \langle k \rangle$  is the probability that Stage I is infected, and  $\alpha \langle k \rangle$  is the probability that Stage S meet Stage S or R and change into Stage R.

When we generalize the above to heterogeneous network, we get the informational dissemination differential equations in the corresponding nonuniform network [23].

$$\begin{aligned}\frac{di_k(t)}{dt} &= -\lambda k i_k(t) \sum_{k'} \frac{k' P(k') s_{k'}(t)}{\langle k \rangle} - v k i_k(t) \\ &\quad \cdot \sum_{k'} \frac{k' P(k') r_{k'}(t)}{\langle k \rangle}, \\ \frac{ds_k(t)}{dt} &= \lambda k i_k(t) \sum_{k'} \frac{k' P(k') s_{k'}(t)}{\langle k \rangle} - \alpha k s_k(t) \\ &\quad \cdot \sum_{k'} \frac{k' P(k') [s_{k'}(t) + r_{k'}(t)]}{\langle k \rangle}, \\ \frac{dr_k(t)}{dt} &= \alpha k s_k(t) \sum_{k'} \frac{k' P(k') [s_{k'}(t) + r_{k'}(t)]}{\langle k \rangle} \\ &\quad + v k i_k(t) \sum_{k'} \frac{k' P(k') s_{k'}(t)}{\langle k \rangle}.\end{aligned}\quad (5)$$

In formula (5),  $i_k(t)$ ,  $s_k(t)$ , and  $r_k(t)$  are the proportion of the population with connection degree of  $k$  in Stages I, S,

and R respectively.  $i_{k'}(t)$ ,  $s_{k'}(t)$ , and  $r_{k'}(t)$  are proportion of the population with connection degree of  $k'$  in Stages I, S, and R separately.  $t$  is the time step.  $P(k)$  is the conditional probability of connecting.

**3.2. Activity Factor.** As microblog community experiences a high-speed development, it faces a problem of low activity. Activity refers to the login proportion of all the registered users during the unit time in microblog community. There are two frequently used statistics indicators in calculation of the activity, one is daily activity users (DAU), and the other is month activity users (MAU). In this paper, we use daily activity users (DAU) as a statistics indicator. Take Weibo as an example: in the fourth quarter of 2012, its registered users exceeded 500 million, while its daily active users were only 46.29 million, with 0.1 activity level. It suggests the user activity in microblog community is rather low. Thus, we should consider the influence of low activity when we research microblog information dissemination, and the first consideration is the parameters of dissemination rate and immunization rate. There are a large number of users who never or seldom log in microblog account for a long period. If we employ all registered users in microblog community to calculate parameter values of the information dissemination rate and immunization rate, a great deviation from actual value will be produced and directly affect the simulation result. So, we design the calculation method of microblog information dissemination rate and immunization rate on the basis of low activity.

**3.3. Information Dissemination Rate.** After posting the information, the users in microblog community are involved in the microblog information dissemination by forward, comment, and like. So we consider influence of these behaviors first when we calculate the information dissemination rate.

According to the research of microblog information dissemination, we find that microblog information has different dissemination rate at different time step. Thus we use  $\lambda_t$  to represent the dissemination rate of certain information at time step  $t$ . Generally, the higher microblog information dissemination rate  $\lambda_t$  suggests the better result of dissemination effect. In each time step, microblog users participate in information dissemination by comment, forward, and like, in which forward belongs to direct information dissemination since, after forward, microblog information will enlarge its dissemination range into the follower's network community, which prolongs the life cycle of information dissemination. While comment and like belong to indirect dissemination information, since these behaviors can only influence other users' dissemination behaviors and cannot form information dissemination.

In consideration of low microblog activity, we should focus on microblog active users rather than all the registered users in microblog community. In this case, we calculate  $\lambda_t$  as

$$\lambda_t = \frac{aX_t + bY_t + cZ_t}{df_i}. \quad (6)$$

In formula (6),  $\lambda_t$  is the information dissemination rate denoted as specific information at the time step  $t$ ,  $f_i$  represents microblog user numbers among the specific community, and  $d$  represents the degree of user activity, which can be processed through the ratio of the number of daily active users and registered users in microblog community.  $X_t$ ,  $Y_t$ , and  $Z_t$  denote times that microblog information behaviors occur at specific time  $t$  including comment, forward, and like.  $a$ ,  $b$ , and  $c$  are three introduced parameters representing the microblog information dissemination rates of weight affected by the three behaviors above. In the actual dissemination, the forward will enlarge the dissemination range and extend the life cycle of microblog information as stated. Therefore, the forward affects the dissemination rate most. Although the comment and like will not enlarge the information dissemination range directly, they enhance the dissemination effect in the original range. Therefore, their effect is less than that of forward. Among the two behaviors, comment's influence is higher than that of like, because it attracts more information in the original message. Thus, among the three behaviors, forward has the maximum affecting weight, comment follows, and like has the minimal affecting weight. So the weights are in decreasing order of  $b$ ,  $a$ , and  $c$ . Since, in the actual microblog information dissemination, the affecting weight of forward, comment, and like cannot be measured by the existing technical conditions, the assignment calculation method is used in this paper, with the distribution which sets their parameter values as 3, 2, and 1.

**3.4. Information Immunization Rate.** In the dissemination process of microblog information, Stage R transforms from Stage I directly and Stage S indirectly. Hence, immunization rate can be divided into direct immunization rate and indirect immunization rate during the information dissemination.

**3.4.1. The Direct Immunization Rate.** The direct immunization rate is the degree of susceptible population who do not disseminate the information and transform into immune population directly. This situation in the actual dissemination shows that forward, comment, and like will not emerge with the probability of the direct immunization rate. We also need to consider activity factors when calculating the rate, so the calculation method of direct immunization rate can be constructed as formula (7). Consider

$$\varepsilon_t = \frac{df_i - X_t - Y_t - Z_t}{df_i}. \quad (7)$$

In formula (7),  $f_i$  represents microblog user numbers among the specific community,  $d$  represents the degree of user activity, and  $X_t$ ,  $Y_t$ , and  $Z_t$  denote times that behaviors including comment, forward, and like occur at specific time  $t$ .

**3.4.2. The Indirect Immunization Rate.** The indirect immunization rate represents the immunization degree of the users who will not continue spreading the information in Stage S. When we construct this calculation method, we should

consider two aspects. One is the change of spreading population. During the process of information dissemination, information dissemination rate and the number of Stage S members are different at each time step. The other is the number of immune population members who transformed from dissemination population that changes all the time. The number of dissemination population members is changing at each time step; it is the same as the number of immune population members in Stage S. Integrating these factors, we construct the calculation method of indirect immunization rate of microblog information dissemination as

$$\alpha_t = \lambda_t \varepsilon_t = \frac{(aX_t + bY_t + cZ_t)(df_i - X_t - Y_t - Z_t)}{(df_i)^2}. \quad (8)$$

In formula (8),  $\alpha_t$  represents the indirect immunization rate of microblog information and  $\varepsilon_t$  is the direct immunization rate in Stage S.  $\lambda_t$  is the information dissemination rate,  $f_i$  represents microblog user numbers in the specific community,  $d$  represents the degree of user activity, and  $X_t$ ,  $Y_t$ , and  $Z_t$  denote times that microblog information behaviors including comment, forward, and like occur at specific time  $t$ .  $a$ ,  $b$ , and  $c$  are three introduced parameters representing the microblog information dissemination rates of weight affected by the three behaviors above.

So far we have completed the calculation method of microblog information dissemination rate and immunization rate considering low activity. Activity factor is also taken into account in other parameter calculations.

**3.5. Credibility Factor.** The credibility factor is so important that it cannot be ignored when studying information dissemination in microblog community. According to the reality of microblog community, the credibility is mainly affected by intimate level and microblog user's attributes and reputation. Because of the credibility factor, it is possible that certain microblog information experiences explosive growth in a short time when emergencies happen. Therefore, we study microblog information credibility specifically.

We introduce the concept of edge weight to describe the credibility relationship between microblog users. Referring to the foregoing study, we set edge weights in microblog community as similar weights, denoted as  $\omega_{ij}$ . Weight  $\omega_{ij}$  can be explained as the degree of intimacy between users. Meanwhile, we believe there is a positive correlation between users' edge weight and credibility. The greater the weight of edge between two random users, the higher the credibility between them. The greater the numeric value  $\omega_{ij}$  is, the greater the likelihood microblog information is to be accepted and disseminated. The calculation method of  $\omega_{ij}$  is represented as

$$\omega_{ij} = \frac{\theta_i f_i}{\sum \theta_i f_i}. \quad (9)$$

In formula (9),  $\theta_i$  is a random parameter given for each user node. Upon setting this parameter, network random shocks and other factors that might affect connection rules should be taken into consideration to improve the network

and choose the best characteristics.  $f_i$  represents the whole user numbers in specific microblog community.

On the basis of the above, we introduce credibility calculation method of microblog information dissemination.

$$\phi(k) = \left( \frac{\omega_{ij}}{\omega_M} \right)^{\varepsilon_t}. \quad (10)$$

In formula (10),  $\phi(k)$  represents the credibility of microblog information in consideration of weighted network. It includes the weight effect on the microblog information dissemination.  $\omega_{ij}$  denotes the edge weight among microblog users,  $\omega_M$  represents the maximum of edge weight in microblog community, and  $\varepsilon_t$  represents the direct immune probability of microblog dissemination.  $\theta_i$  is the fixed reliability parameters set for each node randomly. It shows the bigger parameter means the higher credibility. In the calculation, we use values within (0, 1) which obey normal distribution.

So far we have completed the credibility study of microblog information.

In consideration of activity and credibility of the microblog community, we optimize the kinetic equations of information dissemination in microblog community.

$$\begin{aligned} \frac{di_k(t)}{dt} &= -\lambda_t ki_k(t) \sum_{k'} \frac{k' P(k' | k) (\omega_{ij}/\omega_M)^{\varepsilon_t}}{\langle k \rangle} s'_k(t) \\ &\quad - \varepsilon_t ki_k(t) \sum_{k'} \frac{k' P(k' | k) (\omega_{ij}/\omega_M)^{\varepsilon_t}}{\langle k \rangle} s'_k(t), \\ \frac{ds_k(t)}{dt} &= \lambda_t ki_k(t) \sum_{k'} \frac{k' P(k' | k) (\omega_{ij}/\omega_M)^{\varepsilon_t}}{\langle k \rangle} s'_k(t) \\ &\quad - \alpha_t ks_k(t) \\ &\quad \cdot \sum_{k'} \frac{k' P(k' | k) (\omega_{ij}/\omega_M)^{\varepsilon_t}}{\langle k \rangle} [s'_k(t) + r'_k(t)], \\ \frac{dr_k(t)}{dt} &= \alpha_t ks_k(t) \\ &\quad \cdot \sum_{k'} \frac{k' P(k' | k) (\omega_{ij}/\omega_M)^{\varepsilon_t}}{\langle k \rangle} [s'_k(t) + r'_k(t)] \\ &\quad + \varepsilon_t ki_k(t) \sum_{k'} \frac{k' P(k' | k) (\omega_{ij}/\omega_M)^{\varepsilon_t}}{\langle k \rangle} s'_k(t). \end{aligned} \quad (11)$$

In the differential formula (11), parameters are in accordance with the meaning described above. It is important to introduce two parameters:  $P(k' | k)$  and  $\langle k \rangle$ .  $P(k' | k)$  represents a node (the newly added one) with the degree of  $k$  to be connected node (the original one) with dimension  $k'$ . According to the practice, when a new user is selecting the users he or she wants to follow, he or she will select users they are familiar with or interested in. We use DMS with initial attraction as the calculation method, and the random variable

$\mu_i$  in the formula is adapted to the Poisson distribution. So  $P(k' | k)$  calculation method can be described as follows:

$$P(k' | k) = \frac{k_i + \mu_i}{\sum_j (k_j + \mu_i)}. \quad (12)$$

In formula (12),  $\langle k \rangle$  is the average degree of the network and it measures the average speed of information dissemination in the network. If  $M$  represents connecting edges among users in microblog community and  $N$  denotes the number of microblog users, then the average calculation method can be represented as

$$\langle k \rangle = \frac{2M}{N}. \quad (13)$$

We set the initial conditions of dynamic equations of information dissemination in microblogging community as follows:

When  $t = 0$ ,  $s(0) > 0$ ,  $i(0) = 1 - s(0)$ , and  $r(0) = 0$ .

So far, we have got complete dynamic equations for the dissemination of information in the microblog community.

#### 4. Simulation of Information Dissemination Model

**4.1. Data.** We get the data of Weibo from the API (Application Program Interface, <https://api.weibo.com/>). This part mainly uses Java encapsulation; the API of Java package sends and receives link information as the interface to support the users. To make the study more representative, this paper adopts ordinary users as the research data, with six hours for a time step, from June 11, 2015, at 0 o'clock to June 12, 2015, at 24 o'clock, consecutive data of 48 hours. According to the nodes ID, we get the number of comments, forwards, and likes of every microblog user in every time step. Based on that, we process the data and get the related data needed in the parameters calculation. As shown in Tables 1–8, we finally keep 14 valid users' data in 8 time steps after data filter. In order to protect users' privacy, we use the code instead of the name of the users. The data are shown in Tables 1–8.

**4.2. Simulation.** Based on data processing, we calculate the information dissemination rate and immunization rate in 8 time steps as in Table 9.

We get the average degree  $\langle k \rangle$  based on the crawled  $M$  and  $N$  and calculate  $\varphi(k)$  and  $P(k' | k)$ . Considering the low activity characteristics of the users in microblog community and the practice observation, we set the active degree as 0.1. As the information dissemination rate and immunization rate are different at different time step, we conduct the stepwise simulation. The initial value of Stage I and Stage S can be calculated by average of the eight time steps' information dissemination rates and immunization rates, which is 0.93 and 0.07, respectively. Correspondingly, the initial value of Stage R is 0. In the next seven time steps, we set the simulation result of each time step as the initial value of the next time step and conduct simulation in turn.

Then, we use MATLAB to solve the equations of three stages and illustrate the value of Stages S, I, and R in every

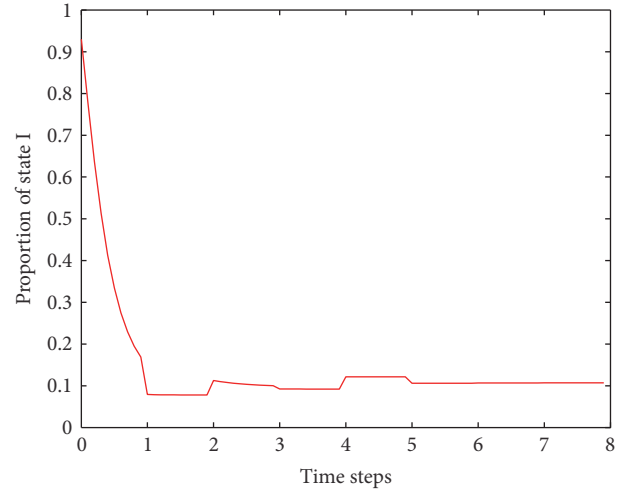


FIGURE 1: The variation trend of state I.

time step. Finally, we make the graph to describe the variation trend of Stages S, I, and R.

#### 5. Result Discussion

In Figure 1, the variation trend of Stage I has three stages. First it decreases quickly, and then it decreases slowly with fluctuating. Finally, it becomes a horizontal line approaching zero. This graph explains the susceptible population decreases quickly in the initial stage. The users of microblog who can receive the information are increasing rapidly after the information is posted. When they received the information, some of them join the people who disseminate the information, and some of them refuse to accept the message and become immune population. Stage I will turn into immune population suddenly especially when the information is rumor and the media clarify rumors. No matter what form it is in, all of the susceptible population will turn to immune population when the life cycle of microblog information is over. Stage I will decrease and approach zero. There are sustained minor fluctuations on point 0.1. It is relative to forgetting mechanism and repeatability of information dissemination. It means part of the microblog users will become information receiver and spreader again because they forget the information.

In Figure 2, Stage S appeared to rise before a sharp drop to 0 in the first time step and then drops approximately to 0 and keeps the level for a long time with a slight rebound. After a piece of microblog information being posted, the users of Stage S will be getting more with the increasing trend of information dissemination rate. However, microblog users will receive too much information in a short time so most people will lose interest quickly. The population who choose to accept and disseminate the information will be sharply reduced, just as the dropping trend of Stage S in Figure 2. A slow growth trend appeared in the third time step due to the forgetting mechanism. After the peak of information dissemination, the information may be forgotten. But a few users who forgot the information may be interested in the information again and disseminate it after several time steps.

TABLE 1: Data of the 1st time step.

User code	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Fans	21734	22451	19623	98129	27533	17525	12887	7816	7816	6979	10747	6216	6669	5466
Messages	3	4	2	10	2	2	2	1	1	1	1	4	3	2
Comments	2	12	7	62	111	55	11	5	5	1	3	5	5	0
Forwards	23	26	242	417	191	89	15	9	9	97	0	117	1	0
Likes	11	18	214	254	218	40	13	28	28	3	0	8	0	5

TABLE 2: Data of the 2nd time step.

User code	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Fans	21733	22451	19635	98143	27536	17523	12888	7820	7820	6979	10747	6217	6669	5466
Messages	3	4	2	10	2	2	2	1	1	1	1	4	3	2
Comments	2	12	7	62	132	55	11	5	5	1	3	5	5	0
Forward	25	26	277	444	230	91	15	9	9	97	0	148	1	0
Likes	11	18	256	287	253	42	13	29	29	3	0	8	0	5

TABLE 3: Data of the 3rd time step.

User code	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Fans	21734	22451	19657	98163	27553	17499	12891	7819	7819	6979	10745	6217	6668	5467
Messages	3	4	2	10	2	2	2	1	1	1	1	4	3	2
Comments	2	12	9	66	149	56	10	5	5	1	3	5	5	0
Forward	27	27	349	478	281	103	12	10	10	98	0	563	1	1
Likes	14	18	296	294	344	44	9	32	32	3	0	8	0	6

TABLE 4: Data of the 4th time step.

User code	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Fans	21732	22453	19730	98163	27572	17451	12891	7822	7823	6979	10746	6216	6669	5467
Messages	3	4	2	10	2	2	2	1	1	1	1	4	3	2
Comments	2	12	9	50	156	56	13	0	5	0	3	5	5	0
Forward	18	29	487	419	313	112	16	0	10	0	0	687	1	1
Likes	8	18	352	276	376	46	17	0	33	0	0	8	0	6

TABLE 5: Data of the 5th time step.

User code	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Fans	21722	22434	19761	98247	27621	17443	12901	7838	7839	6979	10746	6213	6670	5466
Messages	3	4	2	10	2	2	2	1	1	1	1	4	3	2
Comments	2	13	9	70	157	56	13	5	5	1	4	5	5	0
Forward	26	29	526	523	319	118	16	10	10	104	0	731	1	1
Likes	14	19	383	334	392	46	18	33	33	3	0	8	0	6

TABLE 6: Data of the 6th time step.

User code	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Fans	21722	22435	19765	98248	27626	17441	12903	7845	7845	6979	10746	6213	6669	5466
Messages	3	4	2	10	2	2	2	1	1	1	1	4	3	2
Comments	2	13	9	70	157	56	14	5	5	1	4	5	5	0
Forward	26	29	531	522	320	119	16	10	10	104	0	746	1	1
Likes	14	19	389	335	392	46	18	33	33	3	0	8	0	6

TABLE 7: Data of the 7th time step.

User code	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Fans	21723	22435	19771	98248	27626	17434	12903	7870	7845	6979	10739	6213	6667	5466
Messages	3	4	2	10	2	2	2	1	1	1	1	4	3	2
Comments	2	13	9	67	159	56	14	5	5	1	4	5	3	0
Forward	26	29	533	488	320	119	16	10	10	104	0	788	0	1
Likes	14	20	389	308	397	46	18	36	36	3	0	8	0	6

TABLE 8: Data of the 8th time step.

User code	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Fans	21723	22446	19771	98248	27683	17434	12902	7927	7928	6979	10739	6213	6667	5466
Messages	3	4	2	10	2	2	2	1	1	1	1	4	3	2
Comments	2	13	11	72	159	56	14	5	5	1	4	5	5	0
Forward	26	28	558	524	320	119	16	10	10	103	0	794	1	1
Likes	14	19	413	337	400	46	18	36	36	3	0	8	0	6

TABLE 9: Information dissemination rate and immunization rate in each time step.

Time Step	$\lambda_t$	$\varepsilon_t$	$\alpha_t$
1	0.098244137	0.956135637	0.093934721
2	0.106580649	0.952292848	0.101495989
3	0.152921797	0.935106804	0.142998213
4	0.137628064	0.941369427	0.129558851
5	0.255016715	0.922637314	0.235287937
6	0.186573434	0.921958661	0.172012993
7	0.190597452	0.920242113	0.175395802
8	0.19286514	0.919175282	0.177276869

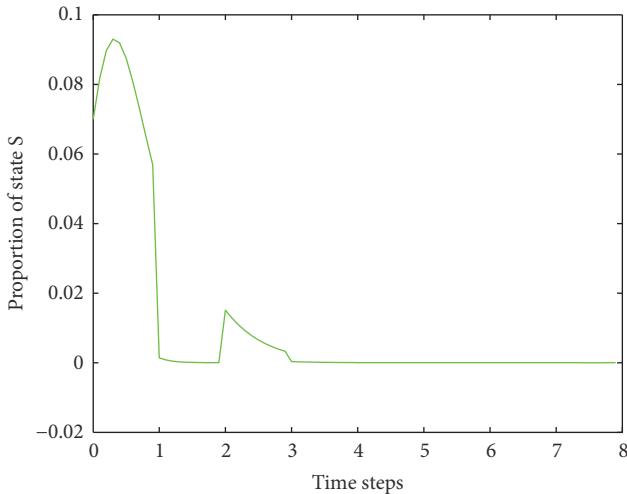


FIGURE 2: The variation trend of state S.

In Figure 3, Stage R shows the rising trend, and capability of the microblog information dissemination is weakened. Because of the rapid increase of Stage S and decrease of Stage I, Stage R, which transformed from dissemination population and susceptible population, increases quickly. Then, the increasing speed of Stage R will slow down, and its

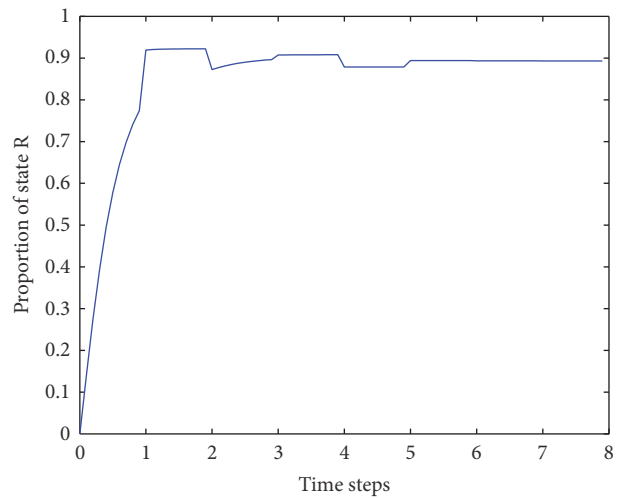


FIGURE 3: The variation trend of state R.

value finally approximates to 1 and keeps steady for a long time. It is consistent with the actual situation that most of the microblog users will change into the immune population in the end of microblog information life cycle. Also, because of the forgetting mechanism, a few microblog users will turn into susceptible population and dissemination population. Therefore, in the short term, Stage R is close to 1 but has not yet reached the level equal to 1. This is the reason why the variation trend of Stage R shows small amplitude range fluctuations.

We find that the microblog information dissemination changes mainly occurred in the first time step, which explains the short life cycle of microblog information dissemination, with an average of six hours or so. So far we have completed the simulation.

After completing the trend analysis of the simulation, we verified the model. We take Stage S as the reference population and then compare the simulation result and the



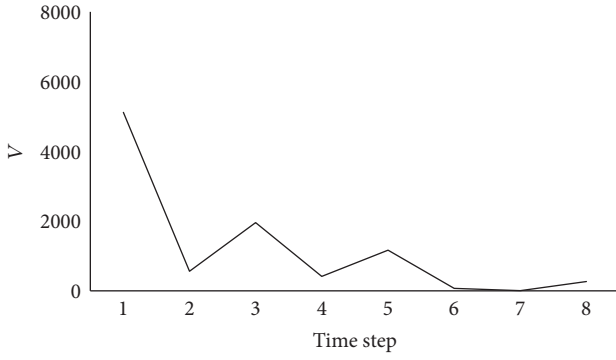


FIGURE 4: The variation trend of information dissemination effect in microblog community.

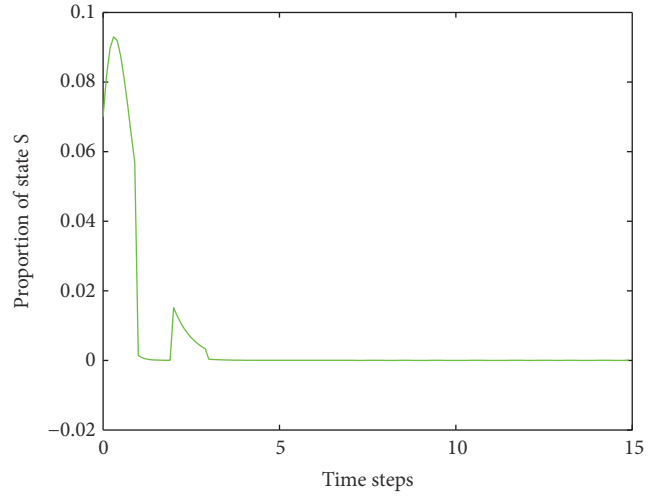


FIGURE 6: The variation trend prediction of state S.

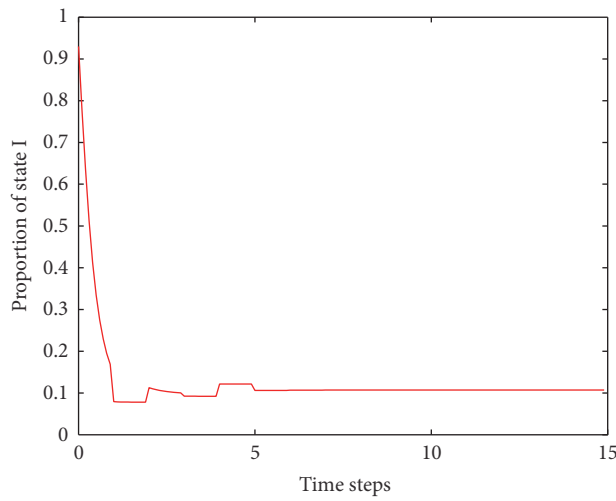


FIGURE 5: The variation trend prediction of state I.

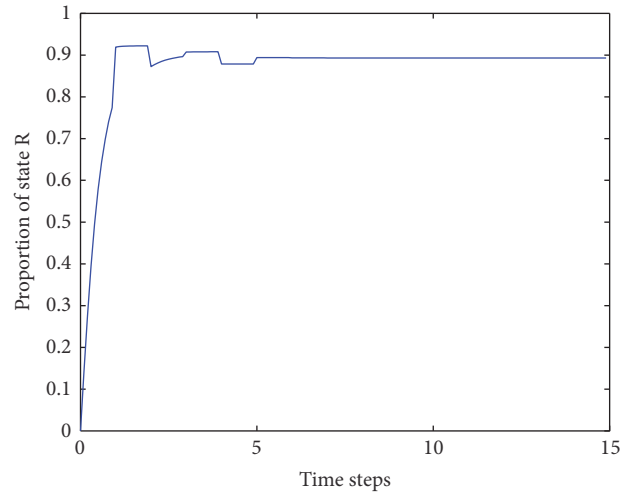


FIGURE 7: The variation trend prediction of state R.

calculation of the crawling data. The calculation method of the crawling data is

$$V = a\Delta X_t + b\Delta Y_t + c\Delta Z_t. \tag{14}$$

In formula (14),  $a = 2$ ,  $b = 3$ ,  $c = 1$ , and  $\Delta X_t$ ,  $\Delta Y_t$ , and  $\Delta Z_t$  separately represent the change amount of the number of comments, forwards, and likes.

As shown in Figure 4, due to the short life cycle, the turning point of dissemination is mainly contained in the first time step. The information dissemination effect drops suddenly in the first time step, slows down in the second time step, and gradually becomes parallel to the horizontal axis, and this trend is basically consistent with the simulated trend.

After the verification, we apply the model to predict the microblog information dissemination trend. According to the simulation result, the changes of microblog information dissemination rate and immunization rate in the end of process are very little, so we take the information dissemination rate and immunization rate of the eighth time step as the initial values of the prediction. Results are shown as in Figures 5–7.

Figures 5, 6, and 7 illustrate that, from the eighth time step, the trend of Stages I, S, and R basically becomes a

straight line. Specifically, Stage I drops to about 0.1 and begins to flat, Stage R increases to about 0.9 and begins to flat, and Stage S decreases close to 0 but still has a slight fluctuation. The phenomenon above is related to the forgetting mechanism and repeatability of information dissemination. A few microblog users will become susceptible population and immune population again after a certain period of time.

## 6. Conclusion

This paper distinguishes the differences between the information dissemination and epidemic disease spreading and discusses the applicability of SIR model for microblog information dissemination study. We optimized the SIR model by introducing the activity, credibility, and network weights into the model and constructed the calculation method of the parameters such as information dissemination rate, information immunization rate, network degree. As the information dissemination rate, immunization rate, and other parameters

are different at different time step, we conduct the stepwise simulation. Through the variation trends of Stages S, I, and R, we found that the model fits the actual situation of the microblog information dissemination well and has good application value. However, this study also has certain deficiency, one is that the active degree of microblog community needs more accurate calculation, and the other is that some more complicated models should be considered to describe the information dissemination more accurately in the future.

## Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

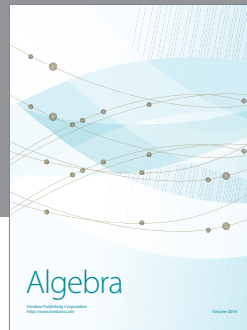
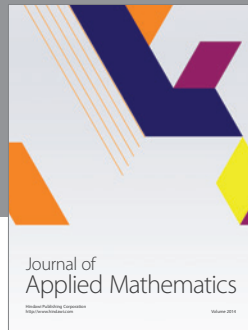
## Acknowledgments

This study was sponsored by the National Natural Science Foundation of China (71532004, 71303069), the New Century Talents of Ministry of Education (NCET-13-0167), the Heilongjiang Postdoctoral Special Fund (LBH-TZ0508), Natural Science Foundation of Heilongjiang Province (F201209, G2016002) and the fundamental Research Funds for the Central Universities (HIT.BRET III.201408), Humanities and Social Science Foundation of Ministry of Education (15YJA630074), Philosophy and Social Sciences Planning Research Project of Daqing (Project nos. DSGB2015054, DSGB2016004), Northeast petroleum University Youth Science Foundation (Project no. 2013QN205), and Heilongjiang Provincial Department of Education Science and Technology research project (12541069).

## References

- [1] F. Li and T. C. Du, "Listen to me—evaluating the influence of micro-blogs," *Decision Support Systems*, vol. 62, pp. 119–130, 2014.
- [2] G. Groszbeck and C. Holotescu, "Can we use Twitter in educational activities?" in *Proceedings of the 4th International Scientific Conference on eLearning and Software for Education (eLSE '08)*, Springer, Bucharest, Romania, April 2008.
- [3] A. Java, X. Song, T. Finin, and B. Tseng, "Why we Twitter: understanding micro-blog usage and communities," in *Proceedings of the 9th ACM WebKDD and 1st SNA-KDD Workshop on Web Mining and Social Network Analysis*, pp. 56–65, August 2007.
- [4] U.-D. Reips and P. Garaizar, "Mining Twitter: a source for psychological wisdom of the crowds," *Behavior Research Methods*, vol. 43, no. 3, pp. 635–642, 2011.
- [5] E. Fischer and A. R. Reuber, "Social interaction via new social media: (how) can interactions on Twitter affect effectual thinking and behavior?" *Journal of Business Venturing*, vol. 26, no. 1, pp. 1–18, 2011.
- [6] M. Ebner and M. Schiefner, "Microblogging—more than fun?" in *Proceedings of the IADIS International Conference on Mobile Learning*, pp. 155–159, Algarve, Portugal, 2008.
- [7] K. Borau, C. Ullrich, J. Feng, and R. Shen, "Microblogging for language learning: using Twitter to train communicative and cultural competence," in *Advances in Web Based Learning—ICWL 2009*, vol. 5686 of *Lecture Notes in Computer Science*, pp. 78–87, Springer, 2009.
- [8] V. Bevilacqua, P. Chiarappa, G. Mastronardi, F. Menolascina, A. Paradiso, and S. Tommasi, "Identification of tumor evolution patterns by means of inductive logic programming," *Genomics, Proteomics & Bioinformatics*, vol. 6, no. 2, pp. 91–97, 2008.
- [9] M. Granovetter, "Threshold models of collective behavior," *American Journal of Sociology*, vol. 83, no. 6, pp. 1420–1443, 1978.
- [10] S. H. Bae, C. S. Yoo, and J. Sarkis, "Outsourcing with quality competition: insights from a three-stage game-theoretic model," *International Journal of Production Research*, vol. 48, pp. 327–342, 2012.
- [11] F. Brauer, "Epidemic models with heterogeneous mixing and treatment," *Bulletin of Mathematical Biology*, vol. 70, no. 7, pp. 1869–1885, 2008.
- [12] B. K. Mishra and D. K. Saini, "SEIRS epidemic model with delay for transmission of malicious objects in computer network," *Applied Mathematics and Computation*, vol. 188, no. 2, pp. 1476–1482, 2007.
- [13] F. Rao, "Dynamics analysis of a stochastic SIR epidemic model," *Abstract and Applied Analysis*, vol. 2014, Article ID 356013, 9 pages, 2014.
- [14] Y. Moreno and A. Vázquez, "Disease spreading in structured scale-free networks," *European Physical Journal B*, vol. 31, no. 2, pp. 265–271, 2003.
- [15] L. Li and J.-H. Sun, "Population characteristics based on conflict information spreading on social network," *System Engineering Theory & Practice*, vol. 34, no. 1, pp. 207–214, 2014.
- [16] L. J. Zhao, Q. Wang, J. Cheng, Y. Chen, J. Wang, and W. Huang, "Rumor spreading model with consideration of forgetting mechanism: a case of online blogging LiveJournal," *Physica A: Statistical Mechanics and Its Applications*, vol. 390, no. 13, pp. 2619–2625, 2011.
- [17] J. Gu, W. Li, and X. Cai, "The effect of the forget-remember mechanism on spreading," *European Physical Journal B*, vol. 62, no. 2, pp. 247–255, 2008.
- [18] H. Ebbinghaus, "Memory: a contribution to experimental psychology," *Annals of Neurosciences*, vol. 20, no. 4, pp. 155–156, 2013.
- [19] C. Wang, C.-Y. Liu, Y.-P. Hu, Z.-H. Liu, and J.-F. Ma, "Stability of information spreading over social network," *Acta Physica Sinica*, vol. 63, no. 18, Article ID 180501, 2014.
- [20] R. M. May and A. L. Lloyd, "Infection dynamics on scale-free networks," *Physical Review E*, vol. 64, no. 6, Article ID 066112, 2001.
- [21] J.-X. Cao, J.-L. Wu, W. Shi, B. Liu, X. Zheng, and J.-Z. Luo, "Sina microblog information diffusion analysis and prediction," *Chinese Journal of Computers*, vol. 37, no. 4, pp. 779–790, 2014.
- [22] H.-F. Lu, Q.-M. Zhang, Y. Zhou, X. Yu, T. Zhou, and Z.-K. Zhang, "Information spreading in microblogging systems: media effect versus social impact," *Journal of the University of Electronic Science and Technology of China*, vol. 43, no. 2, pp. 167–173, 2014.
- [23] W.-G. Yuan, Y. Liu, J.-J. Cheng, and F. Xiong, "Empirical analysis of microblog centrality and spread influence based on Bi-directional connection," *Acta Physica Sinica*, vol. 62, no. 3, Article ID 038901, pp. 494–503, 2013.
- [24] Y. Xu, C. Xiao, and Y. J. He, "A study of the hotness and sustainability of micro-blogging information dissemination," *Journal of Intelligence*, vol. 33, no. 3, pp. 81–87, 2014.

- [25] X. Ding, Q. C. Liu, and W. Zhang, "An improved model for information dissemination and prediction on microblog networks," *Journal of University of Science and Technology of China*, vol. 44, pp. 582–598, 2014.
- [26] H. X. Guo, "Research on micro-blogging information influential based on the small-world network," *Information Science*, vol. 32, no. 9, pp. 60–63, 2014.
- [27] B. Wang, G.-H. Fang, L. Shao, and Y.-W. Zhang, "Information dissemination model of micro-blog community in condition of low activity," *Journal of Harbin University of Commerce (Natural Sciences Edition)*, vol. 31, pp. 215–222, 2015.
- [28] G. Yan, T. Zhou, J. Wang, Z.-Q. Fu, and B.-H. Wang, "Epidemic spread in weighted scale-free networks," *Chinese Physics Letters*, vol. 22, no. 2, pp. 510–513, 2005.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

