Information Dissemination Model of Microblogging with Internet Marketers

Dongliang Xu*, Jingchang Pan*, Bailing Wang**, Meng Liu*, and Qinma Kang*

Abstract

Microblogging services (such as Twitter) are the representative information communication networks during the Web 2.0 era, which have gained remarkable popularity. Weibo has become a popular platform for information dissemination in online social networks due to its large number of users. In this study, a microblog information dissemination model is presented. Related concepts are introduced and analyzed based on the dynamic model of infectious disease, and new influencing factors are proposed to improve the susceptible-infective-removal (SIR) information dissemination model. Correlation analysis is conducted on the existing information dissemination risk and the rumor dissemination model of microblog. In this study, web hyper is used to model rumor dissemination. Finally, the experimental results illustrate the effectiveness of the method in reducing the rumor dissemination of microblogs.

Keywords

Information Dissemination, Microblog, Rumor Spread, Web Hyper

1. Introduction

With the development of information technology, the method of communication among people is no longer confined to traditional letters, radio, and telephone. The introduction of the Web 2.0 era has brought a new method of information exchange and encouraged people to use social network to transfer and share information [1]. Compared with the traditional network, the microblogging network is the representative for the social network due to its strong information dissemination, diffusion capabilities, and organizational skills of users, which contribute to the spread of information on the platform. Given these advantages, Sina Weibo has become one of the major social media sites today. A large amount of false information can spread rapidly in microblogging networks because microblogging sites feature indirect [2], short path [3], and timely information dissemination coupled with user complexity and insufficient existing review of network information.

Large amounts of information among users are transferred in different forms of communication throughout the microblogging platform. However, a direct and effective way to transfer information is via the forwarding process, in which information can spread rapidly among different groups of users. Yu
et al. [4] used the Sina microblogging site and Bowen sample analysis, and concluded that more than 50% of the blog posts are forwarded in Bowen. In [5], seven major forms of information dissemination, namely, the corrugated model, colony model, fireworks model, dandelion model, Gemini model, cellular model, and explosive model, are used for the spread analysis of 500 hotspot information in Sina microblogging. Different types of information in the microblogging network have varying dissemination forms. However, the popular forwarding process of microblogging usually has two peaks, which are a large number of forwarding peaks in a short time, or a decline that will lead to a new round of forwarding peaks [6,7]. These conditions revealed that the information will not be transmitted continuously to some extent [8].

The dissemination of information has a certain degree of uncertainty and randomness. Different rules are applied for various forms of information dissemination. Thus, the propagation path of information is difficult to accurately predict [9]. However, scientific guidance and control of information dissemination are necessary to create a better network environment. The mode of information transmission should be investigated to avoid rumors and other false information that can damage and disrupt network order. In addition, the rules of information dissemination in the microblogging networks should be analyzed.

2. Related Work

Information dissemination in microblogging platforms has become a research hotspot in recent years. Scholars have proposed and established a variety of information dissemination models and verified them according to different entry points and information dissemination characteristics. Kermack and McKendrick [10] established an susceptible-infective-removal (SIR) model of classic infectious disease using differential dynamics, in which the basic content is \( N(t) = S(t) + I(t) + R(t) \), and the differential equation is as follows:

\[
\frac{ds(t)}{dt} = -\lambda s(t)i(t) \\
\frac{di(t)}{dt} = \lambda s(t)i(t) - \mu i(t) \\
\frac{dr(t)}{dt} = ui(t) \\
S(t) + I(t) + R(t) = 1
\]

where \( N(t) \) is the total population, which is divided into susceptible people, infected people, and removers. \( S(t) \) is the proportion of susceptible people, which represents the number of people who have not been infected but are likely to be infected. \( I(t) \) is the proportion of infected people, which represents the number of people who have been infected and are capable of infecting others at time \( t \). \( R \) is the number of people who have been removed from the infected category at time \( t \) (removers include two categories: the first category comprises people who will be infected again, and the second category is composed of people who will never be infected).

A previous study [11] was conducted based on the dynamics of infectious disease. The dynamic evolution equations of information dissemination based on online social network have been constructed. The information transmission model based on the SIR model is proposed and established. In [12], the
entire process of information dissemination in social networks is analyzed. The process is divided into known network topology and unknown network topology. The former classifies users according to the degree of nodes in the social network and designs an improved SI model. The latter uses differential equations and probabilistic methods to describe the propagation process of information in detail and establishes an improved SIR model. The propagation process of information is described by the SI model, and the improved SIR model is established. In [13], an information dissemination model based on SIS is proposed by analyzing the composition and analysis of information dissemination rules for online social networks. In [14-16], the information dissemination model is analyzed using a microblogging site, and the social network, influence, and rules, and information dissemination from the micro-perspective of a small world network behavior are investigated. The behavior of information in the small world network communication is assumed to be equivalent to some kind of categorization problem.

Fig. 1. SIR model of infectious disease.

By contrast, for the study of rumors, such as a study by Liu and Chen [17], the SIR model showed in Fig. 1 is used as a reference, and NetLogo is used to construct a rumor spread model based on the subject. In this model, the masses are divided into communicators, unknowns, and non-affected persons, and the proportion of node access as an alternative for the influence of a user is utilized. In addition, the Twitter-like microblogging network structure without scale, the characteristics of information dissemination, and rumors of maximum spread time are used. Xu et al. [18] implemented the SIR model to describe the spread of rumors on the basis of the characteristics of microblogging community information communication network, and they studied the influence of infection rate and network topology on the spread of rumors by computer simulation. They concluded that the reduction in effective transmission rate and network degree distribution entropy can help reduce the spread of rumors. Duan [19] enhanced the rumor spread model proposed by Tian and Liu [20], which is based on the SIR model, in which the population is divided into ignorants, spreaders (spread rumors), and stiflers (who heard rumors but did not spread rumors), and they improved the model by adding several changes.

The abovementioned research on the information and rumor transmission process in the social network from a micro-perspective presents a certain degree of difference in the rumors and information dissemination model. The characteristics of information dissemination and users are not analyzed because the information dissemination platform of microblogging networks is different from others. For example, Wang et al. [21] did not consider the forwarding process in which the user will forward microblogging information and the forwarders will forward the same information multiple times. The two types of forwarding behavior can describe the dynamic changes in participating information based on the SIS model.

A previous study [22-24] described the speed of information propagation in the network. The speed of transmission refers to the microblogging information initially transferred from the forwarding behavior. Other research [25-31] provided a detailed description of the life cycle of microblogging information. The life cycle of microblogging information refers to the released information that does not require any attention, including forwarding or commenting at time intervals.
For all the abovementioned studies, scholars have not developed a systematic way to prevent the spread of rumors in the information dissemination network, but they conducted considerable research on rumors and microblogging information dissemination. Thus, the information dissemination model for the microblogging network and the model that has been changed with the involvement of the network pushing hands are analyzed. Then, the influence of an Internet marketer on the spread of rumors can be compared and analyzed. Finally, the control of rumors spread by the network pushing hands is verified through experiments.

3. Our Method

In this section, the influence model of network pushing hands on rumor transmission is proposed.

3.1 Microblog Information Dissemination Model with the Involvement of an Internet Marketer

With the increase in understanding and dependence of people in networks and the rapid development of network technology in recent years, a large number of users have focused on network media. Such users are called Internet marketers. An Internet marketer is a person who organizes, implements, and promotes a particular object, such as an enterprise, a brand, an event, and an individual, via a network medium. Internet marketers have a strong ability to disseminate information. Thus, the intervention of an Internet marketer can effectively change the degree and speed of information dissemination. The main difference between an Internet marketer and ordinary user is that the contact rate with other nodes differs. An Internet marketer can push node P and transform the original SIR information propagation model into an SPIR model. The structure is shown in Fig. 2.

![Fig. 2. Room map of microblogging information dissemination with the involvement of an Internet.](image)

Where \( p(t) \) indicates the number of Internet marketers in the microblogging network, \( \lambda_2 \) is the contact rate of an Internet marketer in contact with other nodes in the microblogging social network, and \( \lambda_1 \) is the contact rate of an ordinary infected node in contact with other nodes. The new ordinary communication nodes are \( i_2(t) \) and \( i_1(t) \). From the above room map, the speed of information dissemination can be effectively improved with the involvement of an Internet marketer.

According to Fig. 2, the differential equation for the new propagation model of microblogging information can be introduced:

\[
\frac{dS(t)}{dt} = -\lambda_1 S(t) I(t) - \lambda_2 P_S(t) - \gamma S(t)
\]
where $\gamma$ and $\delta$ are the immune coefficients transformed from the infected nodes spontaneously under the influence of the environment, $\mu$ is the immune coefficient, $\lambda_1$ is the contact rate for ordinary propagation nodes, and $\lambda_2$ is the promoting contact rate of nodes with other nodes. Given that the number of Internet marketers is generally fixed, the proportion of network pushing hands in the system is also fixed (i.e., $P_0$).

### 3.2 Transmission Mode of Microblog Rumors

A registered user ID is a node in the Sina microblogging site, which is between the side of concern and attention in a relationship. If node A is concerned with node B, then A will become B fans. The relationship between the two nodes is concern and attention. If node A is concerned with node B, then A will be the fan of B. The A node in microblogging publishes rumors. Fans can view and forward the rumors in the home page, and the probability of forwarding also depends on the interest of fans in rumors. If the fans think that the rumor is true and interesting, then the possibility of the rumor to be forwarded and spread is high. Thus, the spread of rumors can also be divided into several nodes:

- Suspected node $S$: never heard of a rumor message node.
- Infectious node $I$: a node that saw rumors but did not forward rumors for some reason.
- Forward node $F$: saw the rumors, agreed with the authenticity of rumors, and forwarded the rumor to the node.
- Immune node $R$: immune node can be divided into temporary immune and permanent immune nodes. The temporary immune node is an infected node that does not see a rumor or is not interested in rumors, whereas the permanent immune node is a node that forwards a rumor or determines that the message is a rumor. According to the characteristics of each node, Fig. 3 explains the state transition between the nodes.

\[
\begin{align*}
\frac{dI(t)}{dt} &= \lambda_1 S(t) I(t) + \lambda_2 P_0 S(t) - (\mu + \delta) I(t) \\
\frac{dR(t)}{dt} &= \gamma S(t) + (\mu + \delta) I(t) \\
S(t) + I(t) + R(t) + P_0 &= 1
\end{align*}
\]

where $\gamma$ and $\delta$ are the immune coefficients transformed from the infected nodes spontaneously under the influence of the environment, $\mu$ is the immune coefficient, $\lambda_1$ is the contact rate for ordinary propagation nodes, and $\lambda_2$ is the promoting contact rate of nodes with other nodes. Given that the number of Internet marketers is generally fixed, the proportion of network pushing hands in the system is also fixed (i.e., $P_0$).

Fig. 3. Diagram of node state transition.

1) Assuming that the contact rate between the suspected node $S$ and the forwarding node $F$ is $p_1$ in the event of contact, the suspected node $S$ becomes infected node $I$.

2) Infected node $I$ has a probability of $p_2$ forwarding a rumor message and becomes forwarding node $F$. If $I$ do not see rumors or forwarding, the infected node becomes the temporary immune node $R_t$, and the probability that the infected node becomes an immune node is $p_3$. 

J Inf Process Syst, Vol.15, No.4, pp.853~864, August 2019 | 857
3) To facilitate modeling, we assume that the forwarding node F will only forward the rumor information once and then become the permanent immune node with the probability of 1.

4) If the temporary immune node becomes interested in rumors and transforms, the temporary immune node is converted into the forwarding node and the probability is \( p_4 \).

5) At time \( t \), the number of all types of nodes in the network is \( s(t), i(t), f(t), \) and \( r(t) \). According to the state of node change and the rate of change, we can conclude:

\[
\frac{ds(t)}{dt} = -p_3s(t)f(t) \quad (9)
\]

\[
\frac{di(t)}{dt} = p_3s(t)f(t) - (p_2 + p_3)i(t) \quad (10)
\]

\[
\frac{df(t)}{dt} = p_2i(t) + p_4r(t) - f(t) \quad (11)
\]

\[
\frac{dr(t)}{dt} = p_3i(t) - p_4r(t) + f(t) \quad (12)
\]

Among the nodes, the temporary immune and permanent immune nodes belong to the immune node. \( s(t) + i(t) + f(t) + r(t) = N \), where \( N \) is the total number of nodes.

3.3 Analysis of the Effectiveness to Stop Rumors with Internet Marketers

According to the above analysis, the transmission path and mode of rumors are similar to those of general information. Rumors for the network environment influence the reality of people’s lives. To reduce the degree and speed of rumors, the power of Internet marketers in the microblogging network can be utilized to provide accurate information and maintain a good environment.

Rumor creators usually know the rules of network communication. An example of a network transmission rule includes the network public opinion of heat law, that is, raise concern, “ignite” public opinion, and make public opinion warm or constant, which can promote an event and become the topic of hot news network. Rumor creators know the elements and processes to cause waves, but they lack knowledge on the characteristics and rules of microblogs and other media communication sites. The speculation behavior of Internet marketers is crucial to the rumors in the network. As the two conditions are similar, we can use Internet marketers to push the spread of rumors and block control.

According to the behavior of people in sharing information and the dissemination of information rules and patterns, Internet marketers can be utilized in two ways to reduce the spread and influence of rumors. First, microblogging information is complex and extensive due to the scope of microblogging networks and low-level platforms. When a rumor starts to spread and the intensity of spread is shown, Internet marketers can publish other news or information to divert the attention of the people. Second, rumor microblogging is also the result of materials processing by rumor creators according to rumor characteristics, which are the same as general rumors. Specifically, rumor creators will add false information in the messages and transform real and plain elements to become odd and consequently trend.

Internet marketers can increase the rapid publication of relevant and actual information and enhance the dissemination of real information in the network to guide people and prevent others from spreading the rumors.
Eqs. (9), (10), (11), and (12) show that the state of rumors spread in the network and the number of changes in each node are associated with the number of forwarding nodes. The number of people who have never heard of rumors in the system will decrease with time, and it is inversely proportional to the number of people who saw rumors and forwarded them in the system. The people who saw rumors but did not forward them are negatively correlated with F. The permanent immune nodes in the system are positively related to F because each node that has been rumored will be transformed into a permanent immune node.

Therefore, the number of people who forward rumors first should be controlled to reduce the spread of rumors. The factors $\lambda_2 P_0 S(t)$ related to Internet marketers in Eqs. (5) and (6) are listed in Eqs. (9) and (10) because Internet marketers interrupt the spread of rumors in the form of negative factors:

$$\frac{ds(t)}{dt} = -p_1 s(t) f(t) + \lambda_2 P_0 s(t)$$

$$\frac{di(t)}{dt} = p_1 s(t) f(t) - (p_2 + p_3)i(t) - \lambda_2 P_0 s(t)$$

On the basis of rumors from the original rumor, the reduction in susceptible nodes in the system is suppressed at some point and the increase in the infected nodes is inhibited. The forwarding node is positively related to the infected node. The slow increase in the infected node is bound to increase the speed of the forwarding node through Eq. (11).

The infected node plays a control system to forward the number of nodes to reduce the degree of rumor spread in the microblogging network system.

4. Performance Analysis

We conducted experiments to evaluate the performance of our method.

4.1 Experiment and Analysis of Rumor Transmission

For example, on March 18, 2016, “gold python swallow puppy incident” in the microblogging information dissemination data was used for data simulation and analysis. The origin of March 18 in the microblogging spread took 6 seconds in the screen. The video content was a gold python devouring a live puppy, and the background sound was a dialogue of a man and a woman. The video screen in the Internet rapidly became a hot topic. Friends released the “human flesh” to raise the pictures of gold python. One of them was called “Wang Tao.” Users were human, and microblogging released contact information and other information. However, the late parties issued a video statement to clarify the truth. Rumors were stopped by clarification in the screen of the network with the aid of a promoter to spread in the entire microblogging network.

A crawler program is used to capture the required data through the Sina microblog open interface. From March 18 to March 20, the number of participants in the discussion or forwarding was 53,000 before the parties clarified the fact that the microblogging had 269,000 searches and received 821,000 readings. On March 20, an Internet marketer with 300 million fans released a clarification in the screen. About 3.8 million fans replied in a day, of which 60% of the information was forwarded. After calculation,
the probability of transmission was 0.78% and the probability of immunity was 99.22%. Ordinary users contact about 30 people daily. However, the average daily touch rate of Internet marketers is 600 people. Each Internet marketer has an initial number of 10 with no general information promoters.

According to the above information, the model parameters and the initial values can be determined as follows: \( \lambda_1 = 30, \lambda_2 = 600, \mu = 99.22\%, p_1 = 5 \times 10^{-8}, i(0) = 0, s(0) = 1, \) and \( r(0) = 0. \) The change in proportion for different nodes in the transmission of microblogging information can be determined through the numerical solution of the equations.

From Fig. 4, the following conditions can be provided: (1) more than 98% of users in the microblogging network can obtain the information for about 8 days or more. The proportion in the number of nodes will be controlled in a stable value and remain unchanged at the end of information dissemination in the microblogging network. The number of nodes is stabilized, and the basic information will not be spread when the cycle of information dissemination is 8 days. (2) In the first day, the speed of information transmission is relatively slow. The speed of information dissemination increases over time, but the speed is relatively slow from 3 days to 6 days. After 6 days, the information disseminates rapidly and the spread of information becomes slow. (3) The proportion of transmission nodes remains the same, the proportion of susceptible nodes decreases, and the proportion of removed nodes increases. All nodes are removed nodes at the end of information transmission.

Fig. 4. Three categories of changes in the proportion of nodemap in the microblogging network.

4.2 Influence of Daily Contact Rate by Internet Marketers on the Information Communication Process

The main difference between an Internet marketer and an average user is that their daily contact rate differs. In theory, a high daily contact rate will result in rapid dissemination of information. We conducted an experiment to prove this theory.

By comparing the information in Figs. 4 and 5, the corresponding periods of information transmission are 8, 7, and 6 days when the daily contact rates are 200 and 300. The daily contact rate of Internet marketers can directly affect the speed of information transmission. A high daily contact rate will lead to rapid dissemination of information, short cycle of information dissemination, and stable number of nodes in the microblogging network.
Fig. 5. Influence of the change in contact rate on the changing state of nodes in the network: (a) $\lambda_1 = 20$, $\lambda_2 = 200$, $\mu = 99.22\%$ and (b) $\lambda_1 = 20$, $\lambda_2 = 300$, $\mu = 99.22\%$.

5. Conclusion

People are affected by the information that they rely on. The technology of network information is developing rapidly. People are increasingly using microblogging networks to transmit information. The review system is not comprehensive, and rumors are inevitably spread through the microblogging network due to low-level platforms, user diversity, and rapid spread of information. Rumors can affect a large number of microblogging users in a short time. By analyzing the information dissemination mode of microblogging networks, the information dissemination model with the intervention of an Internet marketer is introduced. Moreover, the information dissemination model of the original online social network based on infectious disease is improved. The relevant data of microblogging networks are used to simulate the proposed model. The trends of susceptible, propagation, and removed nodes in the microblogging network are analyzed with the involvement of Internet marketers. The experimental results showed that the main difference between the network promoter and ordinary user is that the daily contact rate influences the propagation period of information. An Internet marketer contacts many users within a certain time. The faster the information disseminates, the shorter the cycle of information dissemination is. Removed nodes increase slowly and then rapidly, followed by another gradual increase. The period of information communication is completed when the number of nodes is stable.

In this study, we investigated the influence of Internet marketers in controlling the spread of rumors in a network and considered the contact rate as the parameter on the propagation of information. The immune coefficient and initial transmission were determined, which contribute to the dissemination of information in our model. The immune coefficient can also prevent the spread of rumors to some extent. Other effective methods that prevent the large-scale dissemination of rumors in the network environment and do not depend on Internet marketers can be used.

Acknowledgement

This research was partially supported by the National Natural Science of China (Grant No. 61371177...
and U1431102), Shandong Provincial Natural Science Foundation (No. ZR2015AM015), Science and Technology Major Project of Shandong (No. 2015ZDXX0201B04), Science and Technology Development Program of Shandong Province (No. 2014GGX101053), and Science and Technology Development Program of Weihai Province (No. 2014GGX101053).

References


---

**Dongliang Xu**  
https://orcid.org/0000-0002-6908-253X

He is a lecturer, Shandong University (Weihai). He received Ph.D. degrees from Harbin Institute of Technology, Harbin, China, in 2015. His research interests include Network Intrusion Detection Systems (NIDS), multi-pattern matching, data mining, and information security.
Jingchang Pan  https://orcid.org/0000-0003-3836-665X

He is Professor, Ph.D. adviser, was born in Harbin, in 1963. He received the B.S. degree in Peking University, M.S. degree in Harbin Institute of Technology, and Ph.D. degree in Shandong University. His research interests include intelligent computing and data mining.

Bailing Wang  https://orcid.org/0000-0001-6393-5418

He received the M.Sc. and Ph.D. degrees from Harbin Institute of Technology, Harbin, China. He is the Vice Dean and as a Professor with the School of Computer Science and Technology, Harbin Institute of Technology, China. His interests include network & information security, network measurement & modeling, network computing, and parallel processing.

Meng Liu  https://orcid.org/0000-0003-2442-2841

He received his Ph.D. degree from the Computer Application Research Center, Harbin Institute of Technology Shenzhen Graduate School, Shenzhen, in 2017 and M.S. degree from the School of Computer Science and Technology, Shandong University, China, in 2004. Currently, he is a lecturer in the School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai, China. His main research interests include network and information security.

Qinma Kang  https://orcid.org/0000-0002-8628-6403

He received his Ph.D. degree in computer software and theory from Tongji University, Shanghai, China, in 2011. He is currently an associate professor with School of Information Engineering, Shandong University at Weihai, China. His research interests include data mining, optimization algorithms and application, etc.