Journal of Scientific Research & Reports
3(19): 2486-2500, 2014; Article no. JSRR.2014.19.002

SCIENCEDOMAIN *international*
www.sciencedomain.org

The Study of Information Dissemination Behavior in Online Social Network with Propagation Delay

Zhang Wei[1], Wang Lin[1], Dai Jing[1] and Liu Yicheng[2*]

[1]College of Mechatronic Engineering and Automation, National University of Defense Technology, P.R. China.
[2]College of Science, National University of Defense Technology, P.R. China.

Authors' contributions

This work was carried out in collaboration between all authors. Author ZW designed the study and managed the literature searches. Authors WL and DJ performed the statistical analysis. Author LY as supervisor and corresponding author for writing papers provide a lot of help and many useful suggestions. All authors read and approved the final manuscript.

Original Research Article

Received 30th October 2013
Accepted 17th January 2014
Published 7th August 2014

ABSTRACT

In this paper, the information dissemination model involving propagation delay was studied. The effect of delay in the Social Network Service (SNS) was found by considering the theory of complex network and infectious disease dynamics. Also, the authors simulate the online social network information dissemination process and classify the different types of nodes. The simulation results show that: as the high connectivity of online social network, there is less threshold of information transmission in network; the propagation delay makes the threshold lower and the information is easier to be spread, but the propagation delay speeds down the information dissemination meanwhile it decreases the possibility of the information out breaking. This study posts a well understanding of the dissemination behavior with propagation delay in the SNS network, which is fundamental research on consensus.

Keywords: Propagation delay; online social network; homogeneity; delay differential equations.

**Corresponding author: E-mail: liuyc2001@hotmail.com;*

1. INTRODUCTION

With the rapid growth of the Social Network Service (SNS, building online communities of people who share interests and activities or who are interested in exploring the interests and activities of others), there are various online social networking platforms emerged widely for examples, web chatting, blog, podcasts, community sharing and so on. Also, the expansion and interaction of individual platform will form a large network social platform with the characteristics of Six Degrees of Separation (the theory that everyone and everything is six or fewer steps away) in human society. Recently, people have become increasingly dependent on website and real-time communication networks to receive information and news. Facebook, Google+, Renren, Kaixin are the typical representatives of real-time communication networks. Customers can post the latest and novel information via logs, videos and photos to share news and develop friendships. As a popular social network platform, online social networks show some unprecedented advantages and enormous influences and become an important communication tool with a powerful social impact and public opinion mobilization.

The methodology in complex network provides an effective method to study of topology relationship and online information dissemination characteristics in online social network. Also, rapid development of online social networks provides abundant data for understanding the complex networks. Hu and Wang [1], Kumar, Novak and Tomkins [2] and Mislove et al. [3] studied the evolution of the network topology structure, including the degree distribution, clustering coefficient (a measure of the degree to which nodes in a graph tend to cluster together), vertex degree correlation coefficients and the other attributes. For more, Chun etal [4] and Ahn et al. [5] considered the comparison analysis for several SNS networks. However, there results are restricted to the qualitative or semi-quantitative statistical analysis and lack of theoretical findings and mechanism analysis.

There are significant differences and close internal relations between real-life social networks and online communication networks. Previous researchers are more focused on the real social networks in their study. Online social network is human beings-centered and the information spread based on person to person relationship. Moreover, people emphasize the credibility of the information. Therefore, it would be important to know that how people receive the information and whom will be shared with. In the information propagation process, whether the information will be spread out from one to other or not, depends on the credibility and novelty of the information. Thus there is a time delay during the propagation process. The amount of information delay reflects the interest level to customers and it would decrease the spread of the information. However, after a certain time gape (T), customers' interest increase and they will spread out the information finally. As the individuals have different judgment to the importance of the information, the propagation delay is variable.

The previous propagation model is based on regular network, but fundamental are great differences between the online communication networks and the regular network [6,7]. In this paper by using SIR (susceptible-infected-removed) information dissemination model and complex network theory, the theoretical and numerical simulation research for dissemination behavior with propagation delay in the SNS network is carried out, meantime, which is compared with the dissemination model without consider propagation delay for differences. And study influence of propagation delay to the information dissemination. This helps us to reveal the complex evolution behaviors of online communication network.

Provide a factual background, clearly defined problem, proposed solution, a brief literature survey and the scope and justification of the work done.

2. MODEL FORMULATION

2.1 Information Transmission Network

In the information transmission networks, there are three kinds of nodes: one is the node that receives information and will spread it out to other node if the information is enough interested. This node is called infected node. Second is the node that has no idea of the information while it has the opportunity to receive the information from the infected nodes. It is called susceptible node. Third is the node that received information and would not spread it out to other nodes. It is called recovered node [8,9].

The motivation of this paper is to extend the node classification and discover the complex evolution behaviors of online communication network. To this goal, we focus on the significant period of time in the information transmission, during which the node has been infected but not yet infectious himself. The node within this period is classified as compartment E (Abbreviation for Exposed) and it is called exposed node. Also, when exposed node received the information, it may not spread the information immediately. As time goes on, the interest of the information will rise or fall. And after propagation delay, the exposed nodes develop into the infected or recovered nodes. In general, different nodes would have different delays. In order to formulate the information transmission process, we consider the following propagation rules [10]:

(1) If an infected node contact with a susceptible node, the susceptible node has a probability of p_1 to become an exposed node.
(2) If an infected node contact with a recovered node, the infected node has a probability p_2 to become a recovered node. After propagation delay, the susceptible nodes have a probability γ to become infected nodes and $1-\gamma$ to become recovered nodes.
(3) The infected node will not spread the information infinitely. They will stop spreading with a certain velocity v and without contacting with other nodes.

2.2 The Information Transmission Model Based on SIR

2.2.1 Model establishment

In the field of mathematics, delay differential equations (DDE) is one kind of differential equations, there into the unknown function in the determining time's derivative is decided by the previous moment. From the propagation mechanism, the information transmission model is the delay differential equations model [11-13]. In the epidemiology, the delay can reflect the incubation period (Incubation period is the time elapsed between exposures to a pathogenic organism). In the information transmission, the delay can reflect the propagation delay. Therefore, the delay differential equation is more practicable.

In Fig. 1, the direction of the arrow in pointing can show all kinds of nodes. $\tau>0$ is the latent period. Different interest of the information brings different latent period. So τ is the average latent period. After τ, the susceptible nodes have a probability of γ to become infected nodes and $1-\gamma$ to become recovered nodes.

$$\alpha \bullet \rho_S(t-\tau)\rho_I(t-\tau) = \gamma \bullet \alpha \bullet \rho_S(t-\tau)\rho_I(t-\tau) + (1-\gamma) \bullet \alpha \bullet \rho_S(t-\tau)\rho_I(t-\tau) \quad (1)$$

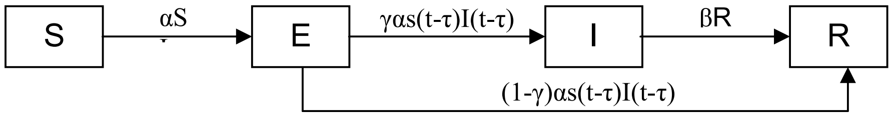

Fig. 1. SEIR block diagram

The probability of infected nodes spread into recovered nodes associated with the recovered nodes' density. Therefore, the dynamic differential equation model is built like this:

$$\begin{cases} \dfrac{d\rho_S(t)}{dt} = -\alpha \bullet \rho_S(t)\rho_I(t) \\ \dfrac{d\rho_E(t)}{dt} = \alpha \bullet \rho_S(t)\rho_I(t) - \alpha \bullet \rho_S(t-\tau)\rho_I(t-\tau) \\ \dfrac{d\rho_I(t)}{dt} = \gamma \bullet \alpha \bullet \rho_S(t-\tau)\rho_I(t-\tau) - \beta \bullet \rho_I(t)\rho_R(t) \\ \dfrac{d\rho_R(t)}{dt} = \beta \bullet \rho_I(t)\rho_R(t) + (1-\gamma) \bullet \alpha \bullet \rho_S(t-\tau)\rho_I(t-\tau) \end{cases} \quad (2)$$

According to propagation rule, the infected nodes will stop spreading with a certain velocity v. As a result, the density of the infected nodes and the recovered nodes should be modified as following:

$$\begin{cases} \dfrac{d\rho_S(t)}{dt} = -\alpha \bullet \rho_S(t)\rho_I(t) \\ \dfrac{d\rho_E(t)}{dt} = \alpha \bullet \rho_S(t)\rho_I(t) - \alpha \bullet \rho_S(t-\tau)\rho_I(t-\tau) \\ \dfrac{d\rho_I(t)}{dt} = \gamma \bullet \alpha \bullet \rho_S(t-\tau)\rho_I(t-\tau) - \beta \bullet \rho_I(t)\rho_R(t) - v\rho_I(t) \\ \dfrac{d\rho_R(t)}{dt} = \beta \bullet \rho_I(t)\rho_R(t) + (1-\gamma) \bullet \alpha \bullet \rho_S(t-\tau)\rho_I(t-\tau) + v\rho_I(t) \end{cases} \quad (3)$$

Combined with the diagram and the equations, the susceptible nodes have a certain probability to become the exposed nodes. Some of the exposed nodes become infected and some become recovered. The goal of this study is to derive α and β.

2.2.2 Homogeneous network

The so-called homogeneous network refers to the network, the nodes in which have similar property, propagation mechanism and propagation process.

In the homogeneous network, the nodes' degree changes smoothly. The node's degree is similar to the average degree <k>. Using means field theory (MFT also known as self-consistent field theory, studies the behavior of large and complex stochastic models by studying a simpler model), α and β can be expressed as <k>:

$$\alpha = p_1(k) < k >, \beta = p_2(k) < k > \tag{4}$$

Substitute for α and β, the dynamic differential equations will be:

$$\begin{cases} \dfrac{d\rho_S(t)}{dt} = -p_1(k) < k > \rho_S(t)\rho_I(t) \\ \dfrac{d\rho_E(t)}{dt} = p_1(k) < k > \rho_S(t)\rho_I(t) - p_1(k) < k > \rho_S(t-\tau)\rho_I(t-\tau) \\ \dfrac{d\rho_I(t)}{dt} = \gamma \bullet p_1(k) < k > \rho_S(t-\tau)\rho_I(t-\tau) - p_2(k) < k > \rho_I(t)\rho_R(t) - v\rho_I(t) \\ \dfrac{d\rho_R(t)}{dt} = p_2(k) < k > \rho_I(t)\rho_R(t) + (1-\gamma) \bullet p_1(k) < k > \bullet \rho_S(t-\tau)\rho_I(t-\tau) + v\rho_I(t) \end{cases} \tag{5}$$

Where $p_1(k)$ is a susceptible node becomes infected through one edge. $\rho_S(t)$, $\rho_E(t)$ and $\rho_I(t)$ are the percent of susceptible, exposed, infected and recovered nodes correspondingly. With the use of the assumption of the probability of the begin or the end of the information transmission is in proportion to $<k>^{\alpha}(0<\alpha<1)$.

$$p_1(k) = \frac{p_1 < k >^{\alpha}}{< k >}, p_2(k) = \frac{p_2 < k >^{\alpha}}{< k >} \tag{6}$$

Let p_1 be the infection factor and p_2 the stop factor. Substitute for p_1 and p_2, we get the delayed SEIR model described as following equations:

$$\begin{cases} \dfrac{d\rho_S(t)}{dt} = -p_1 < k >^{\alpha} \rho_S(t)\rho_I(t) \\ \dfrac{d\rho_E(t)}{dt} = p_1 < k >^{\alpha} \rho_S(t)\rho_I(t) - p_1 < k >^{\alpha} \rho_S(t-\tau)\rho_I(t-\tau) \\ \dfrac{d\rho_I(t)}{dt} = r \bullet p_1 < k >^{\alpha} \rho_S(t-\tau)\rho_I(t-\tau) - p_2 < k >^{\alpha} \rho_I(t)\rho_R(t) - v\rho_I(t) \\ \dfrac{d\rho_R(t)}{dt} = p_2 < k >^{\alpha} \rho_I(t)\rho_R(t) + (1-r) \bullet p_1 < k >^{\alpha} \rho_S(t-\tau)\rho_I(t-\tau) + v\rho_I(t) \end{cases} \tag{7}$$

2.2.3 Disease-free equilibrium

In order to study the model, it is assumed that there is disease-free equilibrium. Let p_1 and p_2 be zero. Then the equilibrium equation is

$$\begin{cases} r \cdot p_1 < k >^{\alpha} \ \rho_S - p_2 < k >^{\alpha} \ \rho_R - v = 0 \\ p_2 < k >^{\alpha} \ \rho_R + (1-r) \cdot p_1 < k >^{\alpha} \ \rho_S + v = 0 \end{cases} \quad (8)$$

Thus $\rho_S(t)=0$, then $\rho_R(t)=1$ since $\rho_S(t)+\rho_S(t)+\rho_I(t)+\rho_R(t)=1$. In this case, all the susceptible nodes will transfer into the infected nodes ultimately.

3. NUMERICAL SIMULATION AND RESULTS

(1) In SNS network, the node with high degree is called the centre node, which will accelerate the speed of information transmission in network. In order to validate the model and study the online social network's structure, dissemination with considering delay propagation and the central node more vividly, the experiment is conducted with the real Data of Last. Fm online social network. Gephi is used to visualize social network structure by colour classification [14]. The node size in figure is proportional with the node degree; the nodes with large degree are annotated by ID that shown in Fig. 2 and community network of Last. FM is scale-free network. The characteristic of acquired data is: the node number for N = 2100, the number of edges between nodes is 25434, the average degree is 13.443, and the average clustering coefficient is 0.187, the average path length: 3.519 and the network diameter is 9. Fig. 3 describes the distributions of network initial degree; it is shown that this network is power-law distribution network.

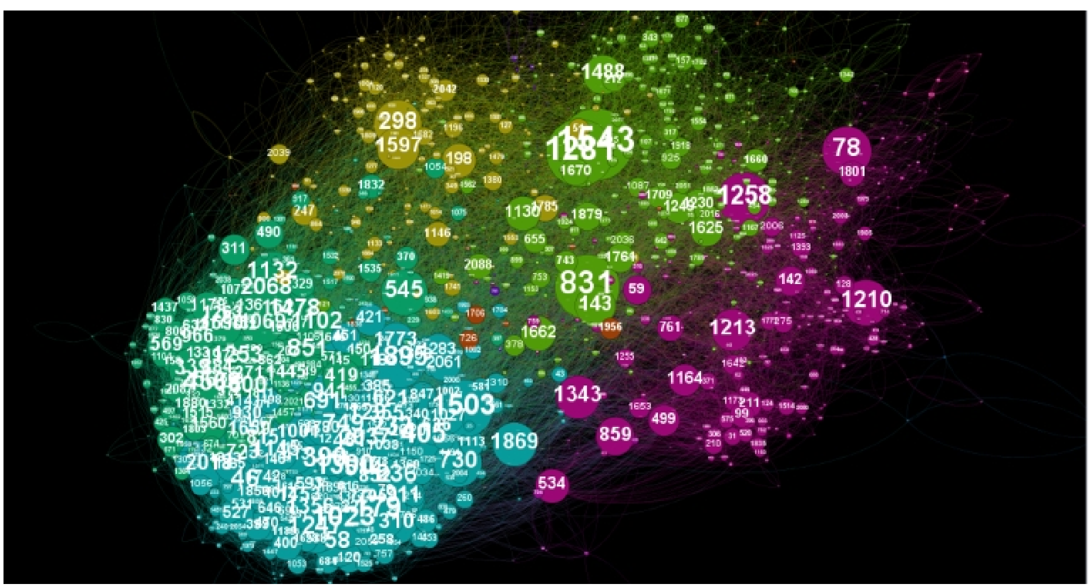

Fig. 2. Community network structure

Initial data includes an only infected node and all others are susceptible nodes in the network. Then, we set up the model parameters as follow: p1=0.3, p2=0.1, υ=0.05, α=0.5, γ=0.7~0.99, the number of iterations T=100. Then the evolution of the density of infected node, susceptible node, exposed node and recovered node as time prolonging, both with and without delay are shown in Figs. 4 and 5.

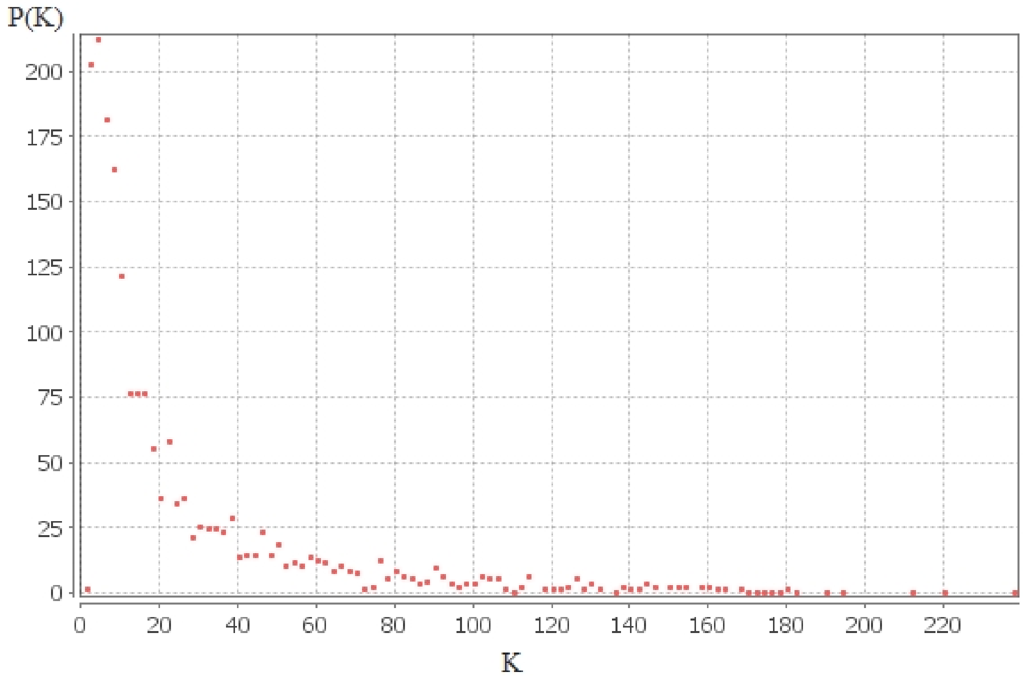

Fig. 3. Distribution of network initial degree

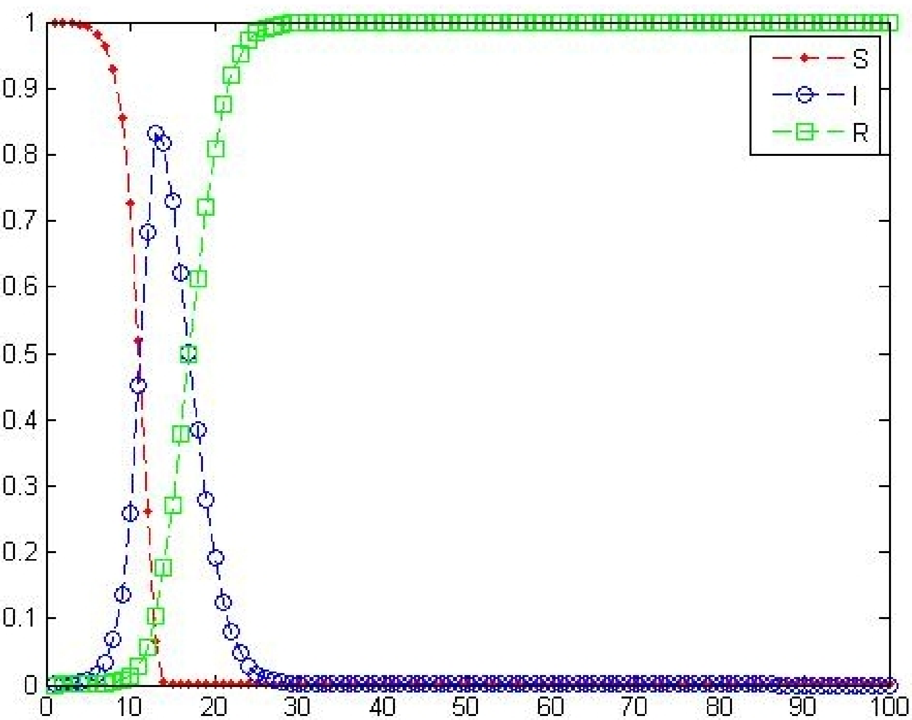

Fig. 4. Evolution of different nodes density (without delay) with time

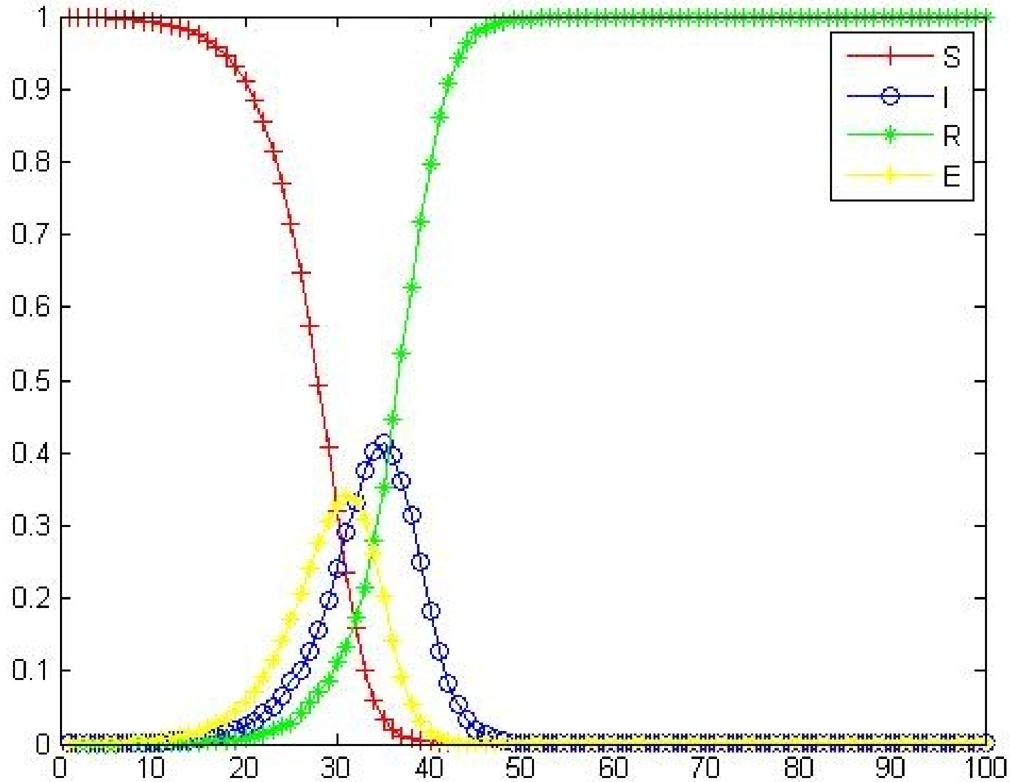

Fig. 5. Evolution of different nodes density (with delay) with time

As shown in Figs. 4 and 5, at the early stage, the density of the recovered node R(t) presents a trend of fast increase, then levels off after a certain of period, till tends to 1; the infected node density I(t) increases rapidly at the start phase, and decreases dramatically after reaching the maximum, until tends to 0; the susceptible node density S(t) diminishes till tends to 0.

Comparing Fig. 4 with Fig. 5, it shows that when considering the effect of social network information propagation delay, the recovered node density R(t) and the infected node density I(t) at the initial stage have the same change trend as without delay dose but their rates of changing get slow down and encounter a delay. The change trend of the exposed node density E(t) and the infected node density I(t) is: rapid increase at the initial stage, then decrease sharply after reaching the maximum till tend to 0. Meanwhile, since there is a transformation relation between the exposed node and the infected node, the exposed node density occurs earlier than the infected node density.

(2) Study the influence of time delay on propagation behavior. Figs. 6, 7 and 8 show how the susceptible node, the exposed node, the infected node perform as time changing with 2, 5 and 10 delay accordingly.

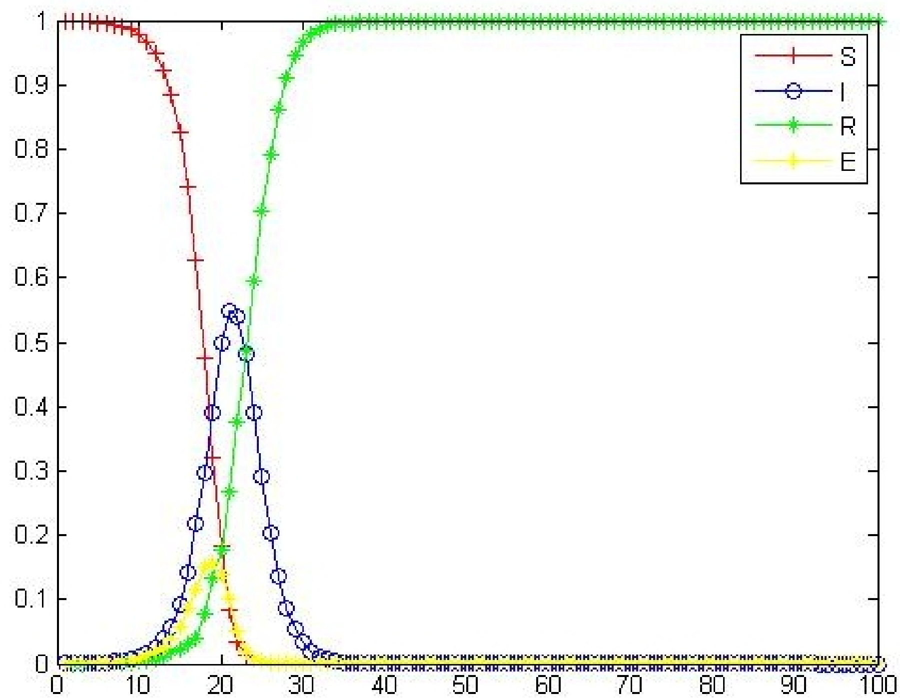

Fig. 6. Change of different nodes density (τ=2) with time

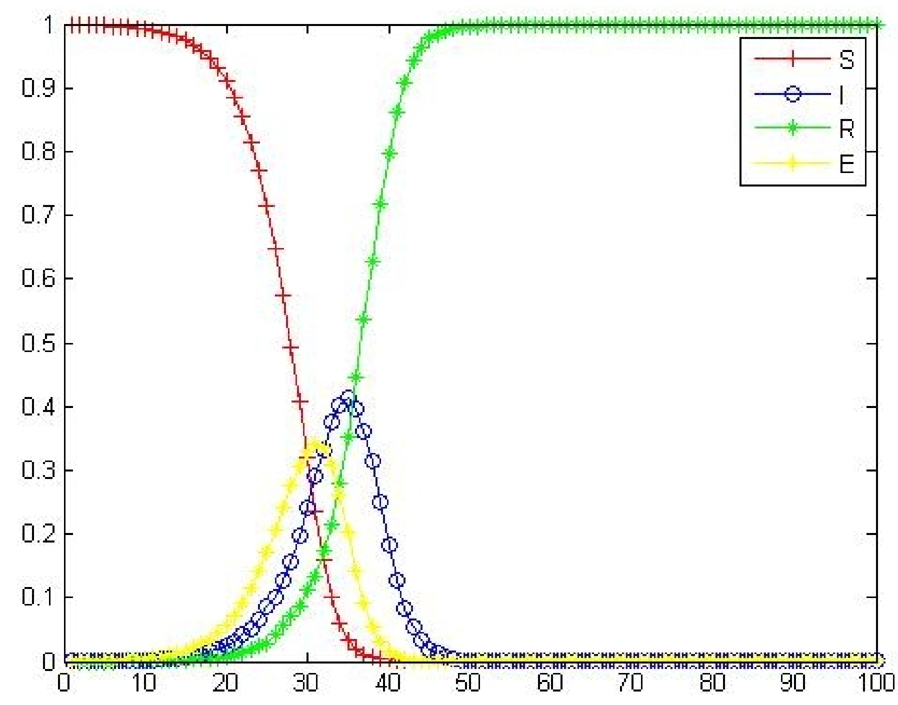

Fig. 7. Change of different nodes density (τ=5) with time

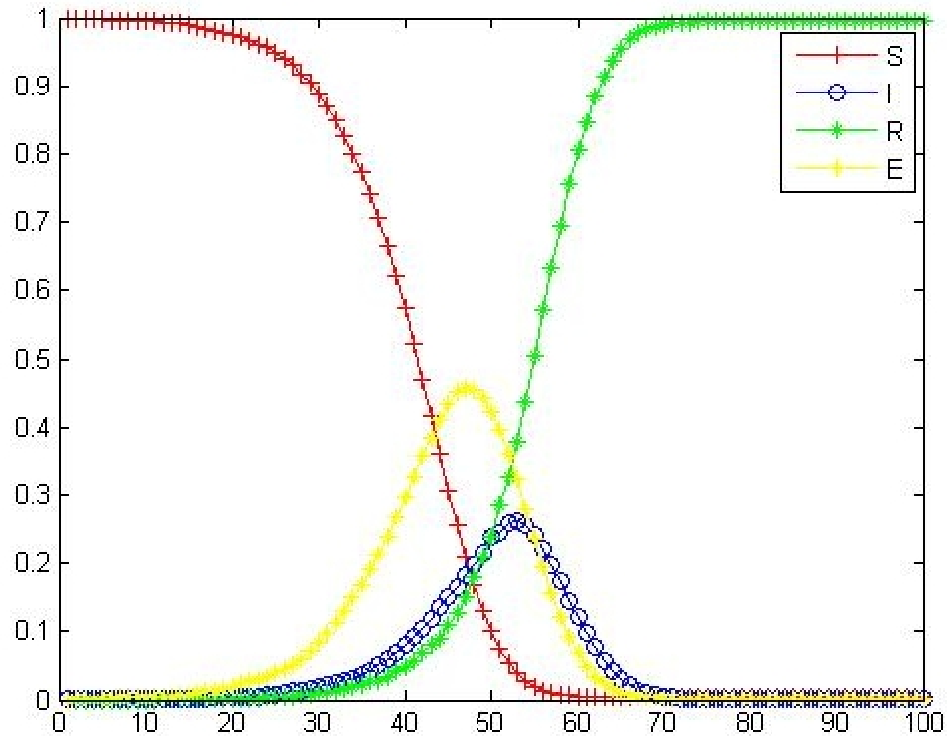

Fig. 8. Change of different nodes density (τ=10) with time

It can be seen that the longer the propagation delay is the slower the information transmission speed in network gets and the transmission of information possesses posteriority. When time approaches infinity, since online social network appears high connectivity, no matter how much the speed of the initial infected node is in the end, the removed node density will tend to 1.

The research focus on how the amount of propagation delay effect the dynamic change process of E(t) and I(t). Change of E(t) and I(t) with different τ values are showed in Figs. 9 and 10. When τ takes a very small value, the exposed node density will be smaller meanwhile the infected node plays a major role within the whole social network dissemination, namely, initial speed of information transmission gets faster in network; With the further increase of τ value, the exposed node density increase gradually at the beginning of information transmission, on the contrary, the infected node density I(t) reduces accordingly. In other words, information wouldn't spread extensively at the beginning of transmission that there is a "buffer", Anyhow, whether final information widely spread is very much depend on depends on the topological structure of network itself and information transmission threshold.

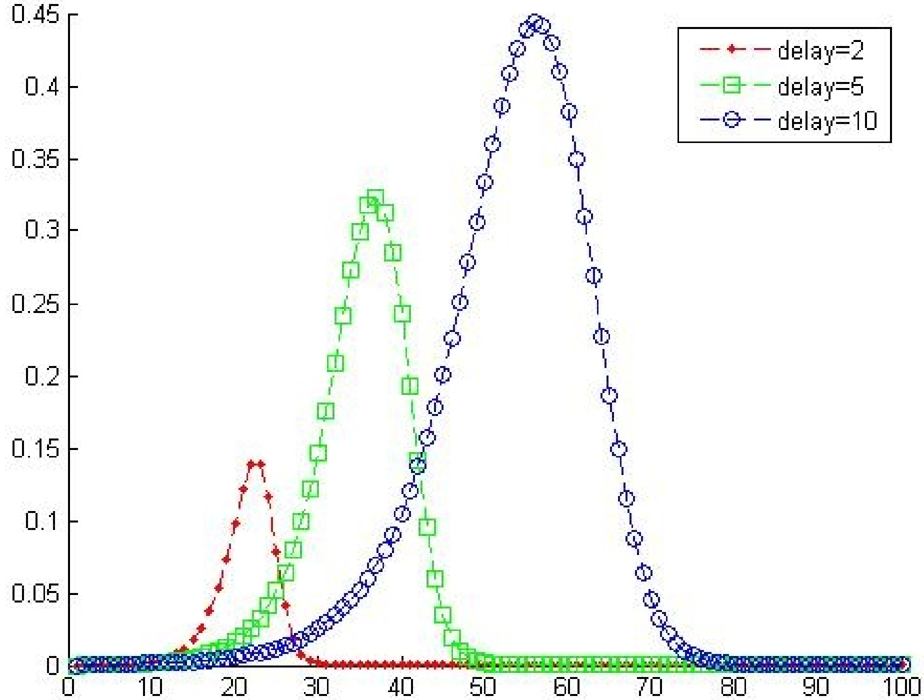

Fig. 9. Change of exposed node density E with different delay value τ

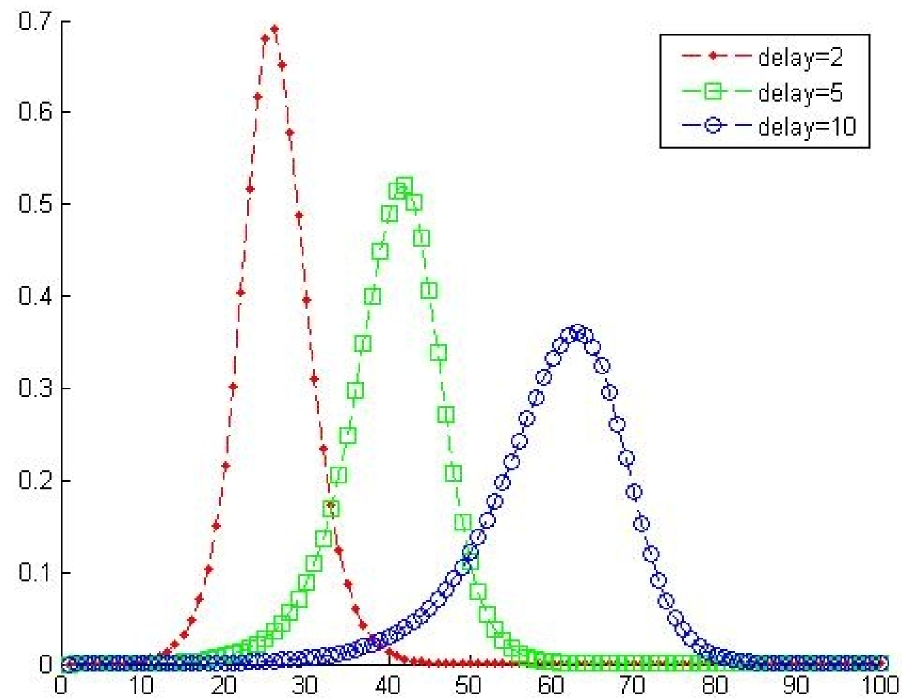

Fig. 10. Change of infected node density I with different delay value τ

(3) Define $R = \rho_R(t_\infty)$, R represents the final value of the recovered node density. When t approaches infinity, transmission process appears equilibrium state, number and density of all kinds of nodes will not change any more. Then we study the impact of the parameters on the R in the model, Fig. 11 shows the change relation of the R with the p1 in different p2. It is observed that with small value of p1 information still can be spread in network, in other words, there almost do not have any threshold of information transmission in network; while based on the regular network the traditional transmission behavior can only spread the transmission probability exceed a certain threshold, which is determined by the topologic structure of network itself. From the figure we can have better understanding of how p1 and p2 affect R respectively: with the same value of p1, R decreases with the increase of p2; with the same value of p2, R increases with the increase of p1.

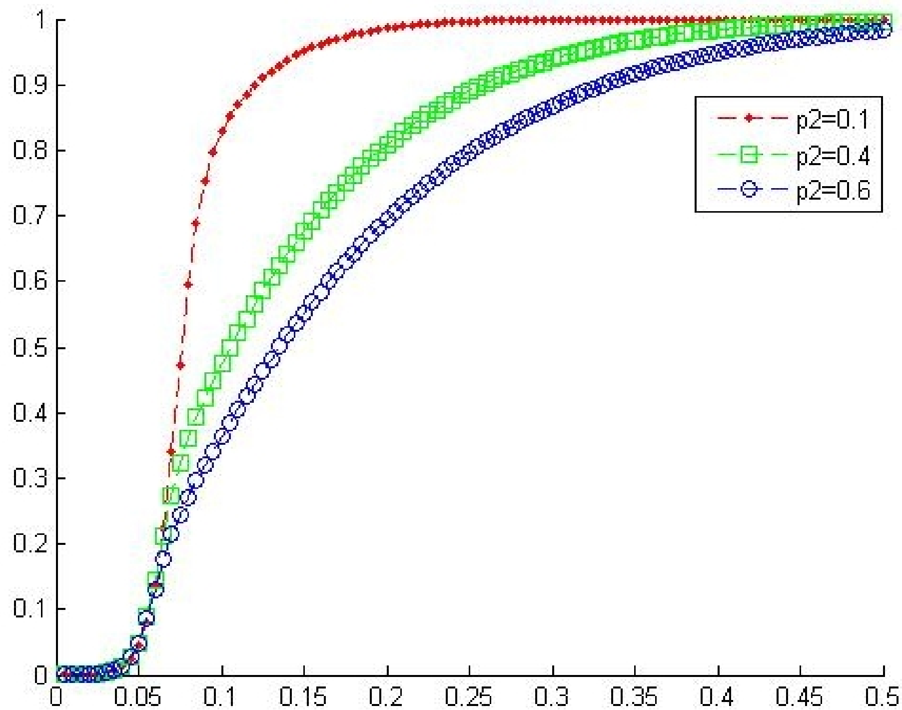

Fig. 11. Change of the infected node final density R with p1 in different values of p2

In same principle, S represents the final value of the susceptible node density. Fig. 12 shows the change relation of S with p1 in different p2. With the same value of p2, S decreases with increase of p1; with the same value of p1, S increases with increase of p2.

(4) Referring to a variety of propagation models which include the SIR model and SIS model in homogeneous network, there exist a positive transmission threshold λc. When the effective transmission rate λ>λc, information will be outbreak and spread in social network. On the contrary, information will be extinguished exponentially.

Without considering inhomogeneous transmission, Fig. 13 considers the effect of delay on information transmission process, it shows that the transmission threshold value constantly decreases as time delay increases, the thresholds are 0.08527, 0.06614 and 0.03909. Therefore by considering the case with propagation delay, as the delay increases, the transmission threshold decreases and information is easier for dissemination.

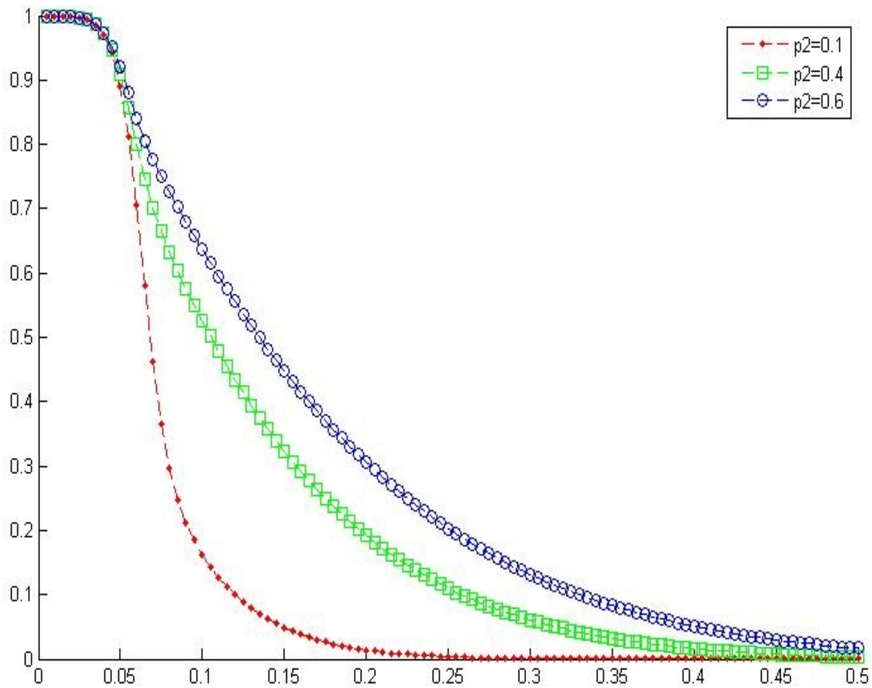

Fig. 12. Change of the infected node final density S with p2 in different values of p1

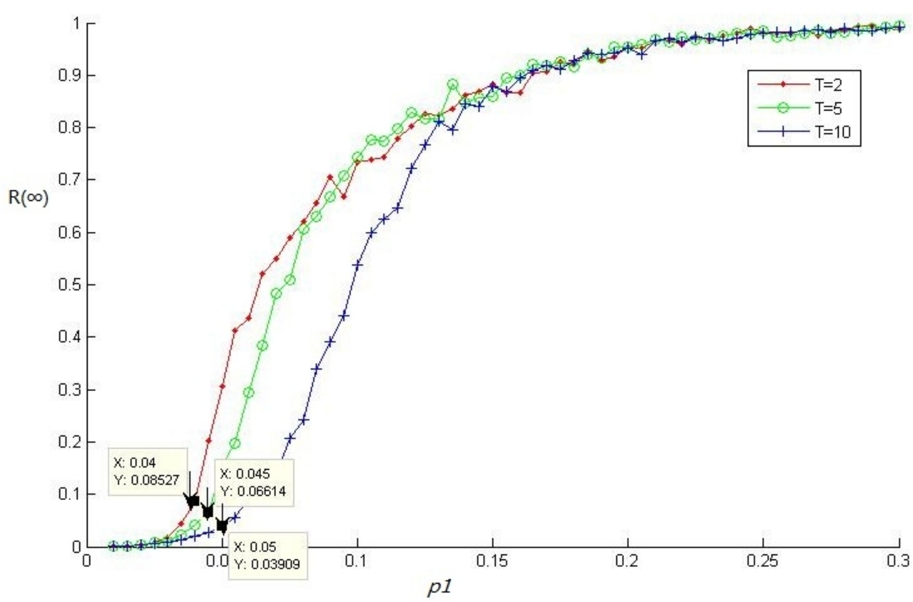

Fig. 13. Time delay on the impact of information dissemination

The speed of information transmission in network will be accelerated by centre node, but once the centre node becomes recovered, it is able to hinder information dissemination effectively as well. It shows that the centre node has large "social influence", which is similar with the basic situation in real life as an important feature of online social network. Meanwhile, the impact of propagation delay in social networks will greatly affect the capability of the centre node. Actually, the views are expressed by celebrities can spread rapidly in network and arouse great social impact in the real world. On the contrary, the views are expressed by ordinary people who can merely obtain much less attention. It is consistent with the simulation results of this paper.

4. CONCLUSION

This paper posts the information dissemination behavior in online social network with propagation delay. It proposes an information dissemination model involving the delay effect of SNS network. Three propagation rules are defined, and the SEIR model is combined with the theory of complex network and infectious disease dynamics. The delay differential equations are established to make the model more consistent with the practical features of SNS. The real data of Last. Fm is used to simulate the dissemination behaviors of information infected node, exposed node, recovered node in online social network. By analysing of mean-field theory and numerical simulation shows that: as the high connectivity of online social network, there isn't exist thresholds of information transmission in network. The propagation delay makes the threshold value lower and the information is much easier to spread but the delay will slow down the speed of information dissemination, therefore, the possibility of the information outbreak is lower. This paper gain a well understanding of the dissemination behaviors with propagation delay in the SNS network, which is the foundation of research on consensus dissemination in network. In order to facilitate analysis of the problem, only the homogeneous network with the same degree of nodes is considered. But in reality the degrees of different individuals are different, and different degrees have different effect on information dissemination. This will be considered in our future research.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Hu HB, Wang XF. Evolution of a large online social network, Physics Letters A. 2009;373:1105-1110.
2. Kumar R, Novak J, Tomkins A. International Conference on Knowledge Discovery and Data Mining, Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, Philadelphia, PA, USA. 2006;611.
3. Mislove A, Marcon M, Gummad KP. Internet measurement conference, proceedings of the 7th ACM SIGCOMM conference on Internet measurement, San Diego, California, USA. 2007;29.
4. Chun H, Kwak H, Eom YH, Ahn Y, Moon S, Jeong H. Internet measurement conference, proceedings of the 8th ACMSIGCOMM conference on Internet measurement, Vouliagmeni, Greece. 200;57.
5. Ahn YY, Han S, Kwak H, Moon S, Jeong H. International world wide web conference，Proceedings of the 16th international conference on World Wide Web, Banff, Alberta, Canada. 2007;835.

6. Wang Ya-Qi, Jiang Guo-Ping. Spreading of epidemics in complex networks with infective medium and spreading delay, Acta Phys Sin. 2010;59.
7. Zhao Jing, Xia Chengyi, Sun Shiwen, Wang Li. A novel SIR model with infection delay and non-uniform transmission in complex networks, CAAI Transactions on Intelligent Systems. 2012;8.
8. Tailei Zhang, Zhidong Teng. Global behavior and permanence of SIRS epidemic model withtime delay, Nonlinear Analysis: Real World Applications. 2008;9:1409–1424.
9. Tongqian Zhang, Xinzhu Meng, Tonghua Zhang, Yi Song. Global dynamics for a new high-dimensional SIR model with distributed delay, Applied Mathematics and Computation. 2012;218:11806-11819.
10. Zhang Yan-Chao, Liu Yun, Zhang Hai-Feng, Cheng Hui, Xiong Fei, The research of information dissemination model on online social network. Acta Phys Sin. 2011;60:550-510.
11. Wanbiao Ma, Mei Song, Takeuchi Y. Global stability of an SIR epidemic model with time delay. Applied Mathematics Letters. 2004;17:1141-1145.
12. John Mallet-Paret, Roger D. Nussbaum Super stability and rigorous asymptotes in singularly perturbed state-dependent delay-differential equations. Journal of Differential Equations. 2011;250:4037-4084.
13. John Mallet-Paret, Roger D. Nussbaum, Stability of periodic solutions of state-dependent delay-differential equations. Journal of Differential Equations. 2011;250:4085-4103.
14. Yuan Wei-Guo, Liu Yun, Cheng Jun-Jun, Xiong Fei. Empirical analysis of microblog centrality and spread influence based on Bi-directional connection. Acta Phys Sin. 2013;62:038-901.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=622&id=22&aid=5654