Role of Tie Strength in Information Dissemination in Mobile Social Networks

Wang Weiguo[1], Chu Chen[1], Liu Jinzhuo[2,3] and Li Tairan[2]

[1]*School of Mathematics and Quantitative Economics, Dongbei University of Finance and Economics, Dalian, China*
[2]*School of Software, Yunnan University, Kunming, China*
[3]*Key Laboratory for Software Engineering of Yunnan Province(Yunnan University), Kunming, China*

wwguo@dufe.edu.cn,chuchen@mail.ustc.edu.cn

Abstract

In this paper, we propose a new model for information dissemination in mobile social networks, in which the infectious probability is defined as a function of the homogeneity and heterogeneity between nodes. Moreover, we investigate numerically the behavior of the model on a real scale-free social site with the exponent $\gamma =1.08$. We find that weak ties between nodes play an important role in information dissemination process. Specially, selecting weak ties preferentially can make information spread faster and wider, and the efficiency of diffusion will be greatly affected after removing them.

Keywords: *Complex networks, Mobile social network, Information dissemination, Hypothesis of ties strength.*

1. Introduction

In 1967, Stanley Milgram proposed six degrees of separation, that everyone and everything is six or fewer steps away, from any other person in the world, so that a chain of "a friend of a friend" statements can be made to connect any two people in a maximum of six steps.[1] A social Network Service(SNS) is a platform to build social networks among people who share interests, activities or group-centered online community such as Facebook and Twitter. With the development of mobile Internet, SNS face mobile terminal users gradually .Mobile SNS developed rapidly in recent years, which has a tendency to replace traditional online SNS. Monthly active users (MAU) of Facebook on mobile terminal have exceeded 1 billion mark. Twitter users use Twitter mobile client for 3 minutes and 7 seconds per month, while using the PC login the Twitter page is only 36 seconds long. Mobile SNS belongs to complex network. With the development of the theory of complex networks, small-world networks [2] and scale-free networks [3] model can describe the characteristics of many real networks. As the mobile social network has many characteristics that traditional social network does not have, the study of the mobile social network model and transmission dynamics on complex networks has received a lot of attention. D.J.Watts and S.H.Strogatz[4] simulated the simple epidemic spreading on small-world networks; they found that the disease spread on small-world networks than regular network.

Most studies considered the infectious probability as a constant, so that it could facilitate to solve the equations but cannot restitute the realistic interaction process in large-scale online social networks.

In order to overcome this drawback, we want to get help from the hypothesis of ties strength of ties was proposed by Granovetter firstly in his representative work in 1973 [5], which he thought the strength of ties could be measured by the relative overlap of the neighborhood of two nodes in the network. He found that weak ties were helpful in

finding a new job [6]. This novel finding of research has become a hotspot for decades. A predictive model [7] was proposed to map social media data to the tie strength. Weak ties theory seems also to work for stability of biological systems [8] and information filtering [9].

In brief, based on weak tie theory, a new transition probability is given which is positively correlated with the homogeneity and heterogeneity between nodes, and information propagation path.

Information propagation process is simulated and analyzed numerically on the real online social site Twitter Data Set [10]. Through the simulations on large-scale real-world data set, we find that propagation path, which more rely on heterogeneity, can expand the scope of the information dissemination. Similarly, through the simulations of removing all sorts of strong or weak ties, we find that removing few fractions of weak ties has greater influence on the scope of the information dissemination.

2. Model

Consider a model for information dissemination in mobile social networks consisting of N individuals who can be subdivided into three groups including ignorant, spreaders and stiflers. We assume that the information disseminates by directed contact of the spreaders with others in the population. The contacts between the spreaders and the rest of the population are governed by the following set of rules.

Whenever a spreader contacts an ignorant, the ignorant becomes a spreader at a rate λ, or becomes a stifler at a rate $1-\lambda$.

After a spreader contacting all neighbors, it becomes a stifler.

In summary, there are only three status transitions in the model: from ignorant to spreaders, from spreaders to stiflers after contacting, and from ignorant to stiflers spontaneously. For more clear descriptions, we give the schematic diagram of transition mechanism among three statuses in Figure 1.
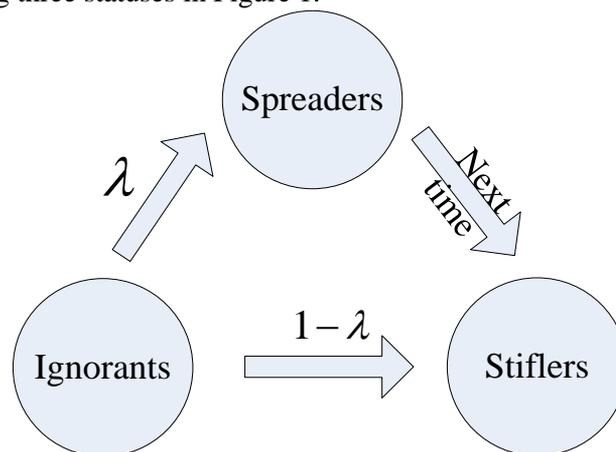

Figure1. Schematic Diagram of Statuses Transition Mechanism in the Model

The mean-field equations that describe the dynamics of the model on mobile social networks are shown as follows:

$$\frac{\partial \rho^I(k,t)}{\partial t} = \rho^I(k,t)(\frac{1}{N_k}\sum_j (1-\omega(k,t))^k - 1),$$

$$\frac{\partial \rho^S(k,t)}{\partial t}$$

$$= -\rho^S(k,t) + \rho^i(k,t)(1 - \frac{1}{N_k}\sum_j (1-\omega(k,t))^k - \frac{1}{N_k}\sum_j \sum_{g=1}^{k} \binom{k}{g} \sum_{m=0}^{g} (-\lambda^{mj})\phi(k,g,t)),$$

$$\frac{\partial \rho^R(k,t)}{\partial t}$$

$$= \rho^S(k,t) + \frac{\rho^I(k,t)}{N_k}\sum_j \sum_{g=1}^{k} \binom{k}{g} \sum_{m=0}^{g} (-\lambda^{mj})\phi(k,g,t)),$$

Where $\rho^I(k,t)$, $\rho^S(k,t)$, $\rho^R(k,t)$ are densities of ignorant, spreaders and stiflers with indegree k at time t, $\phi(k,g,t) = \binom{k}{g}\omega^g(k,t)(1-\omega(k,t))^{k-g}$, and $\omega(k,t)$ is the probability at time t that an ignorant node with in-degree is k follows a spreader node. This quantity can be written as

$$\omega(k,t) = \sum_{k'} p(k'|k)p(s_{k'}|i) \approx \sum_{k'} p(k'|k)\rho^s(k',t),$$

where $p(k'|k)$ is the indegree-indegree correlation function, $p(s_{k'}|i)$ the conditional probability that a node with indegree k' is in the spreader state, given that it is connected to an ignorant node with indegree k, and $\rho^s(k',t)$ is the density of spreader nodes with indegree k' at time t.

3. The Infectious Probability

The infectious probability λ has a great influence on the transmission results. Before proposing the definition of λ, we firstly give the yardstick of strength of the tie between two nodes. Because weak ties are closely related with heterogeneity between two nodes, the strength of weak ties can be defined as:

$$W_{ij}^w = \frac{1}{\left|k_i - k_j + 1\right|},$$

Where k_i and k_j are in-degree of node i and j.

To demonstrate the difference between the weakest strong tie(no common friends) and the weak tie, the strength of strong ties can be defined as $w_{ij} = \frac{o_{ij}+1}{k_i - 1 + k_j + o_{ij}}$ instead of

$w_{ij} = \frac{o_{ij}}{k_i - 1 + k_j - 1 - o_{ij}}$ in[19], where o_{ij} is the number of common friends of node i and node j. Friends denotes two nodes in mobile SNS that follow each other.

The infectious probability of information diffusion through weak ties is affected by heterogeneity between two nodes, while the infectious probability of information diffusion through strong ties is affected not only by homogeneity but also heterogeneity. So, the infectious probability that node i (ignorant) trusts the information under the influence of node j (spreader) can be defined as:

$$\lambda_{ij} = 1 - e^{-((1-r)w_{ij} + rH_{ij})}$$

Where H_{ij} is the normalized heterogeneity index defined as:

$$H_{ij} = \frac{\ln|k_i - k_j + 1|}{Max \ln|k_i - k_j + 1|},$$

for $Max \ln|k_i - k_j + 1|$ is the maximum of heterogeneity in the mobile SNS. r is an adjustable parameter characterizing the preference of information dissemination. The bigger (smaller) r, the more the information diffuse along with heterogeneity (homogeneity). Thus, the infectious probability can restitute the realistic interaction process in mobile online social networks: information can diffuse more easily between tight acquaintances or cewebrities and their fans.

4. Data Set

The data used in this paper is detected from the twitter. Twitter is an online social networking service that enables users to send and read short 140-character messages called "tweets". In this data set, each node represents a Registered user, while a tie from a to b indicates that a follows b. In general, creating a tie from a follower to a followee does not need mutual permission. So we can formalize the data set as a directed graph G(V, E), where V is the set of nodes and E is the set of ties. The size of the network |V| is 14897, and the size of ties |E| is 364892. The max degree in this data set is 8617, while the max out-degree is 4062. The average degree is close to 24. The detailed distribution of out-degree of weak ties is shown in Figure2. The cumulative distribution function (CDF) of the strength of strong ties is shown in Figure3.
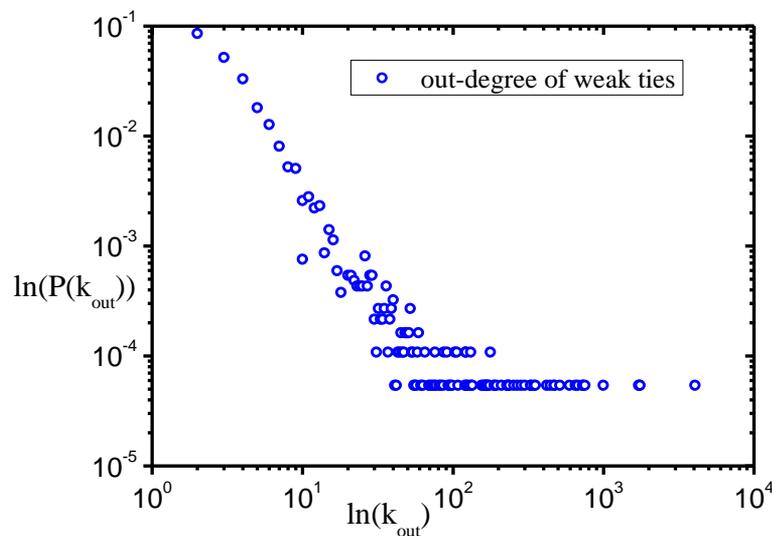

Figure 2. Out-degree Distribution of Weak Ties

Figure2 shows that the data set has a behavior consistent with a power-law network; the majority of the nodes have smaller in-degree, and a few nodes have significantly higher in-degree.

According to the definition of the strength of strong ties in Section2, we can obtain the CDF of the strength of strong ties shown in Figure3. The CDF curve grows sharply before the strength equals to 0.05. So we can draw the conclusion that the great part of ties has the relatively weaker strength, and a small amount of ties are stronger.

Based on the data set of real social network, we will analyze numerically the dynamic behavior of the model in next section.
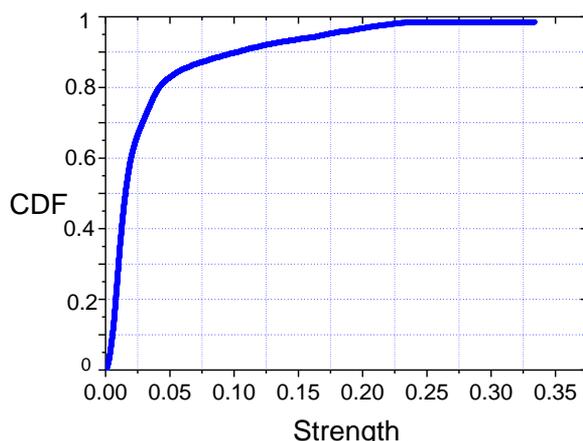
Figure 3. CDF of the Strength of Strong Ties

5. Result and Analysis

5.1 Diffusion Influence of Parameter *r*

Based on the description of Section 2, the adjustable parameter r determines the rumor diffusion paths. When r is big (small), the information will be more likely to be disseminated along with heterogeneity (homogeneity). Consider extreme conditions, when r=1, we have $\lambda_{ij} = 1 - e^{-H_{ij}}$ that means information dissemination between all nodes is only related to heterogeneity. From another point of view, all strong ties in the social networks degenerate to weak ties. On the contrary, when r=0, we have $\lambda_{ij} = 1 - e^{-w_{ij}}$ that means information dissemination is only related to homogeneity. Similarly, the social network from another point of view is equivalent to the networks that all weak ties are removed.

Assume that there is only single spreader at the very beginning, and all other nodes belong to the ignorant class. For the purpose of reducing the impact of the different initial spreader as far as possible, a set of nodes with the same out-degree (in this simulation, the out-degree is chosen as 24, average values of out-degree, and there are a total of 93 nodes in this set) are selected, which can act as the initial spreader, respectively.

We can find that the maximum fraction of stiflers grows with increasing r. It means that selecting disseminating path along with heterogeneity preferentially will make the information more powerful than using the homogeneity strategy. Possible reason may be explained as follows, the nodes with lots of followers would be the core position of the network more often, and ties between theses nodes and followers are weak ties with high out-degree. If information passes by heterogeneity, the nodes in the core position with higher out-degree may have greater chances to spread it to their followers. In the next time step, more nodes may change their statuses from ignorants to spreaders. On the contrary, if we choose the homogeneity path, the nodes in the core position can't spread information to their followers easily.
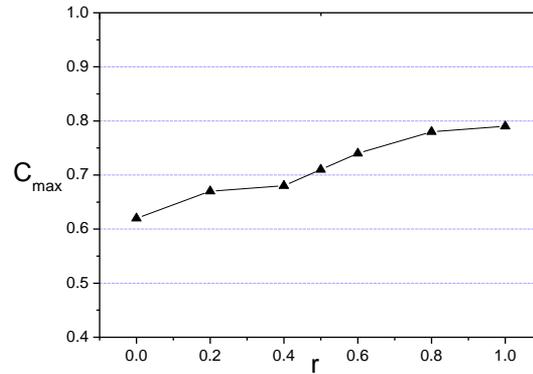
Figure 4. The Diffusion Range (C) of the Information Dissemination for Different *r*

4.2 Diffusion Influence of Weak Ties

We remove several weak ties or strong ties from the real network, and then observe the diffusion range of the information (C) so that we can investigate diffusion roles of weak ties or strong ties in the process of information dissemination. In order to adhere to the principle of single variance, 20% of weak (strong) ties are removed from initial network in descending order by strength. Let r=0.5 and other conditions keep constant. Details are shown in Fig 5.
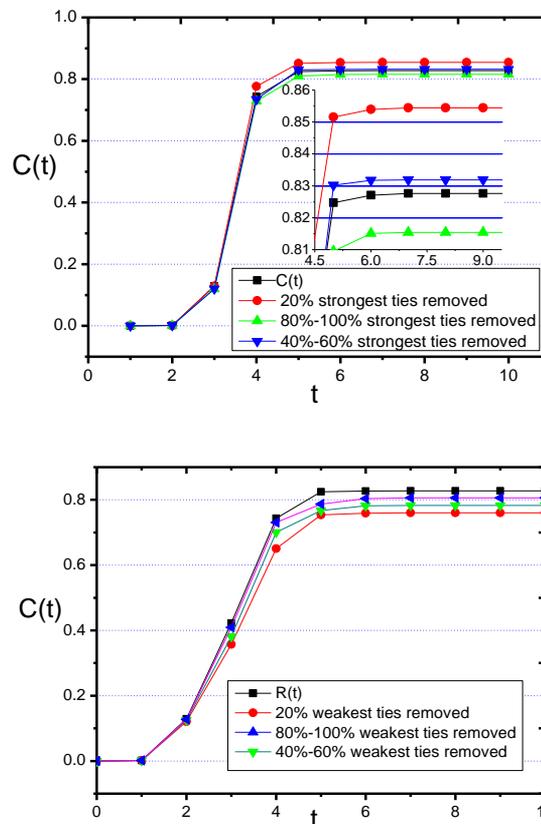

Figure 5. The Diffusion Range(C) of the Information Dissemination During the Removal of Different Fractions of Weak Ties and Strong Ties

As shown in Fig 5, although the C (range of diffusion) decreases in the case of removing both weak ties and strong ties in descending order by strength, weak ties play a more crucial role in information dissemination process. An unusual phenomenon is that the C in the case of removing the strongest 20% of strong ties is larger than the C in initial network. A reasonable explanation for this phenomenon is that the heterogeneity of tie strengths becomes smaller by removing the strongest 20% of strong ties. This agrees with previous works [11].

5. Conclusions

In this paper, we give a new transition probability which is positively correlated with the homogeneity and heterogeneity between nodes, and information propagation path.

Information propagation process is simulated and analyzed numerically on the real online social site Twitter Data Set.[10] Besides, we find that propagation path, which rely more on heterogeneity, can expand the scope of the information dissemination. Similarly, through the simulations of removing all sorts of strong or weak ties, we find that removing few fraction of weak ties has greater influence on the scope of the information dissemination. But how to choose an appropriate r in real social network needs further study. Moreover, we use Jaccard-like weight when considering the homogeneity between two nodes. Next time, we can valid the main results by using AA-index or RA-index. Similarly, other index should be studied in the case of considering the heterogeneity between two nodes.

Acknowledgements

References

[1] Wikipedia, http://en.wikipedia.Org//wiki.
[2] H. B. Hu, D. Y. Han and X. F. Wang, Physica A, vol. 389, (2010), p.1065.
[3] R. Kumar, J. Novak and A. Tomkins, "International Conference on Knowledge Discovery and Data Mining", Proceeding of the eleventh ACM SIGKDD international conference on Knowledge discovery and data mining, (2006); Philadelphia, PA, USA.
[4] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks", Nature, vol. 393, no. 6684, (1998), pp. 440-442.
[5] M.S. Granovetter and J. Am, Sociol, vol. 78, (1973), pp. 1360.
[6] M. S. Granovetter, "The Strength of Weak Ties", University of Chicago Press, (1974).
[7] E. Gilbert and K. Karahalios, CHI, ACM, (2009), pp. 211-220.
[8] P. Csermely, Trends Biochem. Sci., vol. 29, no. 331, (2004).
[9] L.Y. Lv and T. Zhou, Europhys. Lett., vol. 89, no. 18001, (2010).
[10] R. Zafarani and H. Liu, "Social computing data repository at ASU", School of Computing, Informatics and Decision Systems Engineering, Arizona State University, (2009).
[11] J. J. Cheng, Y. Liu, B. Shen and W. G. Yuan, "An epidemic model of rumor diffusion in online social networks", The European Physical Journal B, vol. 86, no. 1, (2013), pp. 1-7.

Authors

Wang Weiguo is a professor of School of Mathematics and Quantitative Economics in Dongbei University of Finance and Economics of China. He received his M.Sc. Degree in Zhongnan University of Economics and law in 1986, and Ph.D. Degree in Dongbei University of Finance and Economics of China in 1998. His current research interests include Econometrics and Management Science.
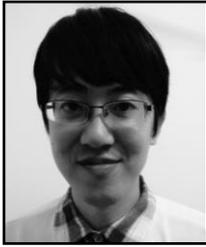
Chu Chen. He received his M.Sc. in Nuclear and its Applications (2008) from University of Science and Technology of China. His current research interests include different aspects of Econometric.

Liu Jinzhuo. She received her B.Sc. in Software Engineering (2007) and Ph.D. in System Analysis and Integration (2013) from Yunnan University. Now she is lecturer at School of Software, Yunnan University. Her current research interests include different aspects of Software Engineering and Formal Methods.
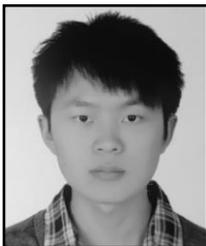
Li Tairan. He is currently an undergraduate student major in Digital Media Technology from School of Software in Yunnan University since 2013. His current research interests include different aspects of Digital Media Technology and Software Engineering.