The Information Diffusion Model Based on the Different Information Value Evaluated by Social Network Users

Sun Qingsong[1,2*]and Li Ying[1]

[1]School of Business, East China University of Science and Technology, Shanghai, 200237, China
[2]School of Computer and Information, Anqing Normal College, Anhui, 246003, China
*easymarshall@163.com

Abstract

We subdivide social network user's state to four kinds, Unknown state, Known state, Informed state, refractory state. Different users may have different judgment for the same information, they transmit the information to their friends in social network lie on the information value and compare with the information threshold, we think unknown users in social network have different thresholds of an information value and subdivide these users to i class, and assume accumulated information value of users follow uniform distribution. Apply mean field theory to build the diffusion model, from the number of refractory nodes, we can know the diffusion effect of an information in a social network.

Keywords: information value, information threshold, U_iKIR, social network

1. Introduction

There are many researchers concentrated on study information diffusion based on social network [1-7,13], it has the similar process with disease, rumors, vial marketing, and so on. William Goffman and Newill [8-9] describe the fundamental mechanism of disease spreading by SIS model or SIR model, each node represent individual and edge represent their links between different individuals, for the SIS model, each individual can exist in two states [10]: S-susceptible and I-infected. At each time, the susceptible individual which is connected to an infected neighbor will be infected with rate λ. Meanwhile, the infected individuals may be recovered and become susceptible individuals at a rate γ, for the SIR model, once an infected individual becomes recovered (R), then the individual will not be infected again. Based on these results, other researchers studied the spreading process of disease, rumor and information from different aspects. Fu, *et al.* [11], examine epidemic thresholds for disease spread using SIS models on scale-free networks with variable infectivity. Cheng, *et al.* [12] verify that the strength of ties plays a critical role in the rumor diffusion process, they introduce a stochastic epidemic model of the rumor diffusion, in which the infectious probability is defined as a function of the strength of ties. Zhao, *et al.* [13] from the spreader's point of view , analyze a spreader became a stifler by three kinds of reasons, spontaneously loses interest in or forgets about a rumor, contacts a stifler, and may doubt the credibility of the rumor. Li, *et al.* find that [14] different users' characteristics, preference, relationships, and actions in the network lead to individuals' various infection probability. Liu, *et al.* [15] find that the final size of information diffusion is affected by affinity threshold of the information. Zhang, *et al.* [10] thought that the susceptible individuals immunity are different, not each one has the same immunity when they affected by infected nodes, so the susceptible individuals can be classified into different cases according to their immunity. There are many nodes on usual network, Tian [16] classified micro-blog users (nodes) to ordinary users ,famous users and corporation users, charity organization users and government users. Famous users such as

stars may are careful to decide whether transmit an information or not, because that if the information is not real, it may cause many ordinary users generate negative impress about them, compared this point, ordinary users may are not worry about personal impress more than stars, so famous users have a higher information value threshold than ordinary users to decide transmit information, In other words, different users accept an information from their neighbors in social network and have different value threshold. Based on this opinion, we introduced a more realistic model to study information diffusion on social network.

The remainder of this paper is as follows. In Section 2, we discuss the information diffusion mechanism, build the model and analyze the result. In Section 3, we present the numerical analysis results and analyze the different factors how to affect information diffusion. Conclusion and discussion will present in Section 4.

2. Model

We consider there are N nodes and links represent the individuals and their interactions on a social network, respectively, each node adopts one of four states at each one time step [15,17]: (1)Unknown(U), the individual has not receive the information, analogous to the susceptible state of the SIR model. (2) Known(K), the individual has receive and read the information but don't decide transmit it yet, maybe he is not sure about the authenticity or the information value is too low. (3) Informed(I), the individual has known the information and transmit it to his neighbors. (4) Refractory (R), the individual will lose interest and has a probability change into informed state again, analogous to the recovered state in the SIR model.

Whether a user transmit an information or not may be decided by him to compare accumulated information value with value threshold, when a user accept information from one neighbor, he will obtain some value, and each user's tendency to become active to transmit information increase as more of his neighbors to transmit information to him and he will obtain accumulated information value [18-19], if he think the accumulated information value is bigger than his threshold and then to transmit it else not. Because of each node has different age, knowledge level, judgment and so on, so they are not identical, so we think that there are different thresholds of an information value for nodes in social network and subdivide unknown nodes to i class[10],U={U$_1$,U$_2$,....U$_i$}, $U = \sum U_i$,each class has a different value threshold and the corresponding value threshold set is θ={θ$_1$,θ$_2$......θ$_i$}, $\theta \in$ [0] ,node v in class i has a threshold θ$_i$, $\theta_i \in \theta$,for the node v, w is the neighbor of v, and v affected by w, we represent the information value with $b_{wv}^i \in [0,1]$,for all the neighbors of v, we have accumulated value $\sum_{w \in in(v)} b_{wv}^i \in [0,1]$,$in(v)$ is the set of neighbors of v, and assume that $\sum_{w \in in(v)} b_{wv}^i$ follows uniform distribution, the U$_i$KIR model for the diffusion process is shown in Figure1.
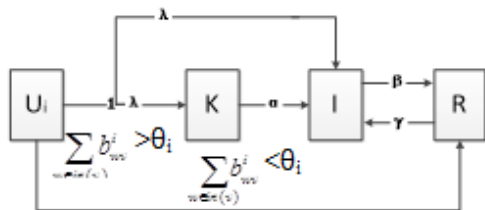
Error! Reference source not found.Error! Reference source not found.

Figure 1. Information Diffusion Model

As shown in Figure1,information spread through contact between different users [13],

we subdivide unknown nodes to i class, each class has a different value threshold θ_i, when unknown node receive an information from informed nodes and then to compare the information value with threshold, if $\sum_{w \in in(v)} b_{wv}^i > \theta_i$, the individuals in ith class will transmit the information to its neighbors with probability λ, that is means the individual change it unknown state to informed state with probability λ, if he not transmit the information, it will come into known state with probability $1-\lambda$ because he has known it. If $\sum_{w \in in(v)} b_{wv}^i < \theta_i$, the unknown node will change into refractory state. From the known individual's point of view, they become informed state with probability α because that they have confirmed the authenticity of the information or affected by informed nodes and other reasons. After informed individuals send information to their neighbors, they will become refractory state with probability β. The refractory nodes remain unchanged but when they interacted with informed nodes, they also have a probability γ change into informed state.

We consider the interactions in homogeneous networks and use $U_i(t)$, $K(t)$, $I(t)$ and $R(t)$ to denote the density of unknown nodes of ith class, known nodes, informed nodes and refractory nodes, respectively. Clearly, these variables obey the normalization condition:

$$U(t) = \sum_i U_i(t), U(t) + K(t) + I(t) + R(t) = 1$$

For ith class unknown nodes, the mean-field equation can be written as

$$\frac{dU_i(t)}{dt} = -kU_i(t)I(t),$$

so for all the unknown nodes and other nodes, we have the corresponding change rate equations as follows:

$$\frac{d\sum_i U_i(t)}{dt} = -kI(t)\sum_i U_i(t) \tag{1}$$

$$\frac{dK(t)}{dt} = k(1-\lambda)I(t)\sum_i(1-\theta_i)U_i(t) - k\alpha K(t)I(t) \tag{2}$$

$$\frac{dI(t)}{dt} = k\lambda I(t)\sum_i(1-\theta_i)U_i(t) + k\alpha K(t)I(t) - \beta I(t) + \gamma kI(t)R(t) \tag{3}$$

$$\frac{dR(t)}{dt} = kI(t)\sum_i \theta_i U_i(t) + \beta I(t) - \gamma kI(t)R(t) \tag{4}$$

Where k denotes the average degree of the network. In the initial period of information spreading, we assumed there is only one informed node to transmit information, and all the other nodes are unknown state, so we can get the initial condition of U_i KIR model,

$$U(0) = \sum_i U_i(0) = \frac{N-1}{N} \approx 1, I(0) = \frac{1}{N}, K(0) = R(0) = 0,$$

From equation (1) and (4), we can get

$$\frac{dR(t)}{dU(t)} = \frac{dR(t)}{d\sum_i U_i(t)} = \frac{kI(t)\sum_i \theta_i U_i(t) + \beta I(t) - \gamma kI(t)R(t)}{-kI(t)\sum_i U_i(t)} \tag{5}$$

Here, the density of unknown nodes in ith class is $U_i(t)$, so we assumed the fraction of $U_i(t)$ in $U(t)$ is p_i, so

$$p_i = \frac{U_i(t)}{U(t)}, \quad \sum_i p_i = 1 \tag{6}$$

From equation (5) and (6), we can get the simple equation as bellow:

$$\frac{dR(t)}{dU(t)} = \frac{kI(t)U(t)\sum_i \theta_i p_i + \beta I(t) - \gamma kI(t)R(t)}{-kI(t)U(t)} = -\sum_i \theta_i p_i - \frac{\beta}{kU(t)} + \frac{\gamma R(t)}{U(t)} \tag{7}$$

Assume $a = -\sum_i \theta_i p_i$, $b = -\frac{\beta}{k}$, so we have a simple formula as follow:

$$\frac{dR(t)}{dU(t)} = a + \frac{b}{U(t)} + \frac{\gamma R(t)}{U(t)}$$
(8)

Equation (8) is a non-homogeneous linear equation, we solve the correspond homogeneous linear equation of (8) as equation (9):

$$\frac{dR(t)}{dU(t)} - \frac{\gamma R(t)}{U(t)} = 0$$
(9)

So,

$$R(t) = CU(t)^{\gamma}$$
(10)

C is a constant, Transform C to m(U), so equation (10) convert into equation (11):

$$R(t) = mU(t)^{\gamma}$$
(11)

Compute the derivation of equation (11) on both sides respectively, and then there is an equation

$$\frac{dR(t)}{dU(t)} = m'U(t)^{\gamma} + m\gamma U(t)^{\gamma-1}$$
(12)

substitute equations (11) and(12) into equation (8) and get two equations:

$$m' = aU(t)^{-\gamma} + bU(t)^{-(\gamma+1)}$$
(13)

$$m = \frac{a}{1-\gamma}U(t)^{1-\gamma} - \frac{b}{\gamma}U(t)^{-\gamma} + C$$
(14)

At last, substitute equation (14) into equation (11), so the general solution of equation (8) is

$$R(t) = \frac{a}{1-\gamma}U(t) - \frac{b}{\gamma} + CU(t)^{\gamma}$$
(15)

When t=0, U(t)=1,R(t)=0, substitute into equation (15), then get

$$C = \frac{b}{\gamma} - \frac{a}{1-\gamma},$$

so equation (15) is changes into

$$R(t) = \frac{a}{1-\gamma}U(t) - \frac{b}{\gamma} + (\frac{b}{\gamma} - \frac{a}{1-\gamma})U(t)^{\gamma}$$
(16)

At $t = \infty$, diffusion process terminated, $R(\infty) = R, U(\infty) = 1-R, R(0) = 0, U(0) = 1,$ from the equation (16),we get the equation as below:

$$R = \frac{a}{1-\gamma}(1-R) - \frac{b}{\gamma} + (\frac{b}{\gamma} - \frac{a}{1-\gamma})(1-R)^{\gamma}$$

Because of follow conditions,

$$a = -\sum_i \theta_i p_i, \quad b = -\frac{\beta}{k},$$

We get the final equation about R:

$$R = -\frac{\sum_i \theta_i p_i}{1-\gamma}(1-R) + \frac{\beta}{k\gamma} + (\frac{\sum_i \theta_i p_i}{1-\gamma} - \frac{\beta}{k\gamma})(1-R)^{\gamma}$$

From the above equation, we can compute R by other parameters in the equation.

3. Numerical Analysis

We assume there have two classes unknown nodes in a social network and in the initial condition there is only one informed node in the network, so that U(0)=N-1/N,K(0)=0, I(0)=1/N,R(0)=0,and suppose that the average degree k=10,λ=0.7,α=0.6,β=0.7,γ=0.1, different effect of θ_i and p_i on R as Table 1.

Table 1. Different Effect of θ_i and p_i on R

	θ_1	p_1	θ_2	p_2	R

A	0.7	0.5	0.4	0.5	0.4439
B	0.7	0.5	0.5	0.5	0.1978
C	0.6	0.5	0.4	0.5	0.6199
D	0.7	0.6	0.4	0.4	0.3069
E	0.7	0.4	0.4	0.6	0.5566

The value of R represent how may nodes have known the diffusion information. From the table, we can see that between row A and row B, there have different threshold θ_2 of second unknown node class ,and the diffusion result tell us the bigger of θ_2,the lower of R, that means the bigger threshold lead to information diffusion process more difficult, and there have the similar result between row A and row C, meanwhile, the threshold θ_1 and θ_2 in row B are bigger than row C, the value of R is more small.

Meanwhile, there have the same threshold θ_1 and θ_2 in row A, D, E, and $\theta_1>\theta_2$,the percentage of unknown node in first class and second class have changed, As an example to compare row A and row D, we known that the more bigger percentage of one unknown node class which have more bigger threshold, the value of R is more small.

4. Conclusion

An extend information diffusion model is proposed in this paper based on considered the different information value of different social network users, we study that there are different information value threshold because of different social network users affected by different friends and their judgment. From the model, we found that information value threshold and the proportion of users in different class affects the range of information diffusion.

In the future, with the research development of information diffusion, there are many influenced factors will be find, but there are more complicated problems to need research, for example, there are many online social networks, one user may receive an information and transmit to another online social network, some users may quit from one network and some users entrance. Therefore, we must base on research more impacted factors to improve the diffusion model in order to better explain the diffusion process.

Acknowledgements

References

[1]   H. Ahmed, T. Daim and N. Basoglu, "Information technology diffusion in higher education", Technology in Society, vol. 29, no. 4, **(2007)**.
[2]   J. P. Hsiao, C. Jaw and T. Huan, "Information diffusion and new product consumption: A bass model application to tourism facility management", Journal of Business Research, vol. 62, no. 7, **(2009)**.
[3]   Y. Xue and R. Gençay, "Hierarchical information and the rate of information diffusion", Journal of Economic Dynamics and Control, vol. 36, no. 9, **(2012)**.
[4]   F. Xiong, "An information diffusion model based on retweeting mechanism for online social media", Physics Letters A, vol. 376, no. 30, **(2012)**.
[5]   H. Wang, F. Wang and K. Xu, "Modeling Information Diffusion in Online Social Networks with Partial Differential Equations", America, **(2013)**.
[6]   S. Livingstone and G. M., "Information-geometric Markov Chain Monte Carlo methods using Diffusions", Entropy, no. 16, **(2014)**.
[7]   L. Small and O. Mason, "Information diffusion on the iterated local transitivity model of online social networks", Discrete Applied Mathematics, vol. 161, no. 10, **(2013)**.
[8]   W. Goffman and V. A. Newill, "Generalization of epidemic theory", nature, no. 10, **(1964)**.
[9]   W. Goffman, "Mathematical approach to the spread of scientific ideas--the history of mast cell research", nature, Nature, no. 212, **(1966)**.
[10] H. F. Zhang and M. S. A. F., "Different Epidemic Models on Complex Networks",Commun. Theor. Phys, vol. 52, no. 1, **(2009)**.

[11] X. C. Fu, M. Small, D. M. Walker and H. F. Zhang, "Epidemic dynamics on scale-free networks with piecewise linear infectivity and immunization", Physical Review E, no. 77, **(2008)**.

[12] 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)**.

[13] L. J. Zhao, "SIR rumor spreading model in the new media age", Physics A: Statistical Mechanics and its Applications, vol. 39, no. 4, **(2013)**.

[14] Z. Li, "Predicting the scale of information diffusion in social network services", The Journal of China Universities of Posts and Telecommunications, no. 20, **(2013)**.

[15] H. Liu, "Affinity based information diffusion model in social networks", International Journal of Modern Physics C, vol. 25, no. 5, **(2014)**.

[16] Z. W. Tian, "Research on information dissemination of micro-blog based on complex network", Industry University of Harbin, Harbin, China, **(2012)**.

[17] L. Lu, D. Chen and T. Zhou, "Small world yields the most effective information spreading", New Journal of Physics, no. 13, **(2011)**.

[18] D. Kempe, J. Kleinberg and E. Tardos, "Maximizing the spread of influence through a social network", in Proceeding of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, **(2003)**.

[19] H. Che and Y. T. Wang, "Threshold-Based Heuristic Algorithm for Influence Maximization", Journal of Computer Research and Development, vol. 49, no. 10, **(2012)**.

Authors

Sun Qingsong, born in 1982, Anqing city, Anhui province, Current position, grades: PhD student of East china science and technology university and a lecturer at the Anqing normal college University studies: Master of Science, Anhui agriculture university, 2006; study for a doctor degree in East china science and technology university at present Scientific interest: information diffusion, complex network, knowledge management.