# User Behavior Modeling Research Based on Group Level in Social Networks

Xie FENG[1, a], Wanli ZUO*[2,b]

[1,2] College of Computer Science and Technology, Jilin University, Jilin Province 132101,China

[a]xiefeng@yeah.net, [b]wanlizuo@yeah.net

**Keywords:** Complex Network; Social Network Analysis; Human Dynamic; Node Role; Topic Discovery

**Abstract.** Human behavior, due to its complexity, makes exploration of human behaviors very important and interesting. It is also because of the high complexity of human behavior, how to find and reveal the objective law has long attracted the research interest of scholars from sociology, psychology, economics, and other disciplines. With the rapid development of network technology, especially in recent years the online social network representative by personal online community, online dating network, social network has developed rapidly, the popularity of whose application directly lead to increase of the data amount, a large number of detailed user behavior data is recorded. Much data in online social network era gives us an unprecedented opportunity to study human behavior.

## Information dissemination process modeling under social networks forwarding mechanism

In fact, not only network activity in the community of users is in the continuous attenuation, other social network users in the same will exit the interaction due to lose enthusiasm. Whether the dissipative behavior of users on earth will affect the information transmission of social network? This chapter will study information dissemination process based on post forwarding, discusses the user activity attenuation and network topology's influence on the communication process.

### Social network information forwarding mechanism

In social network, the post will be released today by a user system will be pushed to the user's friends so these friends can easily see this post. When friends read the post, will consider whether need to forward this post. [1]If the post is very meaningful that some neighbors decided to forward, , post author neighbor will have a chance to see this post. At this time, the influence of the post has exceeded the original author's neighborhood, and is likely to be large. On the contrary, if the original author's neighbors have no interest, for the post will be quickly forgotten, its effects were small. Once a user forwards the post, the post will always exist in the user's page. Posts at the same time will be pushed to the user's neighbors, even the users no longer pay attention to this post, and the neighbors will still be able to read this article is forwarded posts. Unless under special circumstances, the posts are deleted by the website administrator, otherwise the user after forwarding the post will have been infection status will not change. At the same time, if the neighbors after reading the post did not forwarded immediately, their passion for the dissemination of information will keep falling until ignore this information.

 

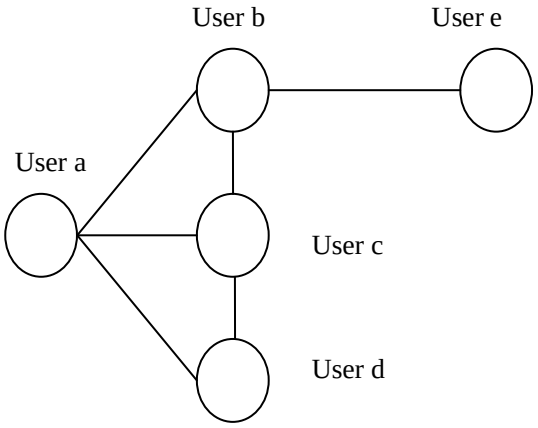


Figure1. Social network information transmission

**Social network information transmission models**

According to the social network information transmission mechanism, this model establishes N individuals in the network, and the interconnection of two individuals interacts. [2]Each time step, disseminator and other individual update rules is as follows:

(1) When the ignorant node comes across communication nodes, namely an infection is ignorant to infect at the probability of $\lambda$ and spread information, otherwise ignorant will enter a state of contact;

(2) Contact state node gradual loss of interest sharing and dissemination of information, to the probability $\delta$ to become immune, otherwise the rest of the contacts will be spread neighbor node to the probability $\lambda$ of infection.

Do average field analysis to the model. Because information dissemination process is closely related to the network topology, defined in the t time, degree of nodes for K ignorance, contact the density of nodes, infections, immune respectively $s(k,t)$ , $c(k,t)$ , $i(k,t)$ , $r(k,t)$ . [3]Assume probability of contact and ignorance node k infected during interval [t, t+1] are respectively $p_{si}(k,t), p_{ci}(k,t)$ . Ignorance node can only be transmitted infections, so average infection state neighbor degrees K node density as follows:

$$\sum_{k'} P(k' \mid k) i(k',t) \qquad (1\text{-}1)$$

If individuals have k neighbors, the individual has opportunity k times to become infected, the probability $p_{si}(k,t)$ is:

$$p_{si}(k,t) = 1 - \left(1 - \lambda \Delta t \sum_{k'} P(k' \mid k) i(k',t)\right)^{k} \qquad (1\text{-}2)$$

Under the limit of $\Delta t \to 0$ , obviously

$$p_{si}(k,t) = \lambda k \Delta t \sum_{k'} P(k' \mid k) i(k',t) \qquad (1\text{-}3)$$

It is noticed that there was no immediate immunized contact state node is likely to become infected, so the transition probability $p_{ci}(k,t)$ is as follows

$$p_{ci}(k,t) = \lambda k \Delta t (1-\delta) \sum_{k'} P(k' \mid k) i(k',t) \qquad (1\text{-}4)$$

**Research on user behavior model**

**Task priority queuing model**

In this model, in order to study of complex user behavior, model designer made some necessary simplification and assumption, including: 1) Time discretization; the continuous time discrete time step. 2) Unit processing time. Model assumes that the user at each discrete time step to deal with a

task. 3) Never stop working. [4]Users remove every step a task from the task queue and consumption. 4) Same rate of production and consumption. This model can be understood as essentially the problem of producers and consumers, producers keep production task, consumers constantly perform a task. Model assumes that the rates are the same. 5) The choice of the ways of consumption. In the model, the authors assume that the user agent missions there are three ways: first come, first service, selected at random, the highest priority task priority. Specific model is as follows:

1) Each user agent has the task queue of length L, each task assigned to the priority parameter $x_i$, $x_i \in [0,1]$, and $x_i$ obey uniform distribution.

2) At any step, producers produce a new task to the task queue, and record the task of time step. User agent from the task queue to probability P consumption of the current highest priority task, with probability 1 - P randomly select a task to consumption, consumption time step and record the current task.

$$P(\tau_w) = \begin{cases} 1 - \dfrac{1-p^2}{4p} \ln \dfrac{1+p}{1-p}, \tau_w = 1 \\ \dfrac{1-p^2}{4p(\tau_w - 1)} \left[ \left( \dfrac{1+p}{2} \right)^{\tau_w - 1} - \left( \dfrac{1-p}{2} \right)^{\tau_w + 1} \right], \tau_w > 1 \end{cases} \tag{2-1}$$

(2, 1) can be derived from the type, when $p \to 0$:

$$\lim_{p \to 0} P(\tau_w) = \left( \frac{1}{2} \right)^{-t_w} \tag{2-2}$$

That is, this theoretical derivation results and simulation results are of the same, obeying exponential decay Poisson process. The limiting cases illustrate random selection strategy the user agent represents. When $p \to 1$, from type (2-1) available:

$$-\lim_{p \to 0} P(\tau_w) = \begin{cases} 1 + o\left( \dfrac{1-p}{2} \ln(1-p) \right), \tau_w = 1 \\ o\left( \dfrac{1-p}{2} \right) \left( \dfrac{1}{(\tau_w - 1} \right), \tau_w > 1 \end{cases} \tag{2-3}$$

Among them:

$$\tau_0 = \left( \ln \frac{2}{1+p} \right)^{-1} \tag{2-4}$$

This situation almost every task of waiting time is 1, $\tau_w = 1$, almost all tasks after joining the list is consumed. The waiting time of those task that are not consumed obey the power law distribution, and rate of power attenuation index of 1. When $p < 1$ is fixed value, $\tau_w \to \infty$ from the type (2, 1) it can be obtained:

$$P(\tau_w) \sim \frac{1-p^2}{4} \cdot \frac{1}{\tau_w} \exp\left( -\frac{\tau_w}{\tau_0} \right) \tag{2-5}$$

**Adaptive interest drive model**

Analysis can be seen from the previous section that it is not under any condition, each one of us takes the principles of task first. [5]Take we are preparing a paper, for example, perhaps the paper is not the most important in our life, but because the deadline is coming soon, we may put more effort in preparing the paper. Specific rules of model are as follows:

1) Time discretization. The time step t, the probability of user agent in the certain events is r.

2) In the t + 1, r (t + l) = a (t) r (t), including:

$$a(t)=\begin{cases} a_0, \tau_i \le T_1 \\ a_0^{-1}, \tau_i \ge T_2 \\ a(t-1), T_1 < \tau_i \le T_2 \end{cases} \quad (2\text{-}6)$$


**Summary**

Find the user's behavior rules, and related research based on user behavior rules has been attracting people to explore unceasingly. In recent years, with the rapid development of web technology, mobile communication technology, it is becoming of feasibility to ensure the upgrading of the current Internet system. More importantly, from news and information share to users involve in web BBS, to the social network of manufacturing, modern Internet is more and more people-oriented. The current various online social networks take the user as the center, information generated by the user and used by the user. Compared with previous systems, these systems contain more abundant user behavior data. On the one hand, the data provide us with very good opportunities to explore user behavior, on the other hand presents new challenges to the existing methods.



**Acknowledgements**

This research was supported by the National Science foundation (Grant No:60973040), National Science Foundation for outstanding young scholars (Grant No: 60903098), Science and Technology development projects of Jilin Province (Grant No: 20070533), The ministry of education by the specialized research fund for the doctoral program of higher education (Grant No: 200801830021), interdisciplinary and innovation projects of higher education Supported by Basic scientific research business expenses of Jilin University (Grant No: 200810025), Key scientific and technological subject of jilin province(Grant No.20130206051GX).


**References**


[1] Porter M A,Mucha P J,Newman MELA network analysis of committees in the U.S. House of Representatives[C]. Proceedings of the National Academy of the United States of America (PNAS), 2005, 102(20):7057-7062.

[2] Schneider J J,Hirtreiter C. The impact of election results on the member numbers of the large parties in Bavaria and Germany [J]. International Journal of Modern Physics C, 2005, 16(8):1165-1215.

[3] JBattiston S,Catanzaro M. Statistical properties of corporate board and director networks[J]. The European Physical Journal B, 2004, 438:345-352.

[4] H. C. Dong, S. C. Hui,Y. L. He. Structural Analysis of Chat Messages for Topic Detection. Online Information Review, 2006, 30(5):496-516.

[5] J. P. Zeng, and S. Y. Zhang. Variable Space Hidden Markov Model for Topic Detection and Analysis [J], Knowledge-Based Systems, 2007, 20(7):607-613.