View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Available online at www.sciencedirect.com
SciVerse ScienceDirect
ELSEVIER
Procedia Computer Science 9 (2012) 1338 – 1344
Procedia Computer Science

International Conference on Computational Science, ICCS 2012

Online Social Network Evolving Model Based on Damping Factor*

Guoyong Cai , Ruili Wang, and Baohua Qiang

Department of Computer Science and Engineering, Guilin University of Electronic Technology, China

Abstract

Recently the evolving model of network has become one of hot research topics. Since the BA model is presented, many kinds of evolving models have been proposed. However the generation mechanism of existing evolving model of online social network has certain limitation, i.e. researchers mostly focus on the factors that facilitating network growth but ignore the factors that delaying network growth. In order to resolve this limitation, a new online social network evolving model (DFEM) is proposed in this paper. In this model, the factors delaying network growth are defined as damping factor and which is divided into the decline of initial attraction, the loss of node heat, and irresistible natural factors. In addition, the damping factors, the attractive factors, and the degree of nodes are taken into consideration together in the preference link. The results of theoretical analysis and numerical simulation suggest that the degree distribution of network generated by DFEM model is more correspond with existing online social network.

Keywords: *Online Social Network; attractive factor; damping factor; BA model; CALW model.*

1. Introduction

With the application of the technology of Web 2.0 and the emergence of large online platforms such as QQ, Facebook[1] , and Blog, the online social network has attracted many researches' attention. Now the research on the structure of social network is improving, meanwhile, the evolving law of online social network has become a hot research field[2, 3].

As is known, the research of network evolving law is based on the establishment of evolving model. The essence of evolving model is network generative mechanism, specifically including the relation of every component and the whole characteristic generated by these relations. Research shows that the evolving model of most complex network has two characteristics, which is (1) growth, that is, network is growing and (2) preference link, that is, a new node tends to connect to the nodes which have larger degree. Recent years the research on evolving model of complex

* Corresponding author. Department of Computer Science and Engineering, Guilin University of Electronic Technology.
E-mail address: *wangruili1207.1@163.com*.

1877-0509 © 2012 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.procs.2012.04.147

network focuses on improving the characteristic of preference link. Based on BA model, many kinds of evolving model have been proposed [4, 7, 9-12].

Online social network is a complex network comprise of a lot of nodes and edges, in which nodes represent users and edges represent the links between users. There exists the following phenomenon when online social network evolving, Take Blog as an example, the blog of a celebrity may have a very high click rate because of his/her popularity at first, but as time growth the quantity and quality of the papers published by the bloger play the leading role, rather than the bloger's popularity when a user decides whether read the bloger' s paper or not. The problem of low cluster coefficient of evolving model is solved in literature [15] by introducing the concept of attraction and dividing it into initial attraction and evolutional attraction. However, the factors which delay the growth of network are neglected in this literature, so based on the existing evolving model of complex network, a new Damping factor-based evolving model of online social network (DFEM) is proposed in this paper. This evolving model studies the damping factors that delay the growth of network by introducing an absolutely new concept of damping factors and comprehensively considers how degree of nodes, attractive factors, and damping factors influence the generative mechanism of network in the way of preference link. The theory analysis and numerical simulation show that the network generated by DFEM model is more suitable for existing online social network.

2. Background

2.1. BA scale-free network evolving model

Barabási and Albert proposed the BA scale free model[5,6] in 1999, this model has growth character and preference link character. BA scale free network evolving model is described specifically as follows:

- Growth. First initialize the network in which the number of nodes and edges is m_0 and n_0 separately, and then introduce a new node at one timestep which will link to the m existing nodes ($m \le m_0$) of the initial network.
- Preference link. The probability that the new node links to the existing node i and the degree (k_i)of node i follows

$$\prod_i = \frac{k_i}{\sum_j k_j}$$

The existing researches show that the BA model has certain limitations[4,8]. On the one hand, the index of degree distribution of the scale-free network generated by BA model is 3, which is not suitable for the real network whose index of degree distribution is between 1 and 3. On the other hand, researchers only consider the influence of the degree of nodes on degree distribution in this model.

After t steps, a new network with N = $t + m_0$ nodes and mt edges will be generated by this algorithm. From Fig 1, we can clearly see the evolving process of BA network evolving model, where m_0=m=2.There are only two nodes in the network at first, the new node will link to the two existing node of the network according to the preference attachment mechanism.

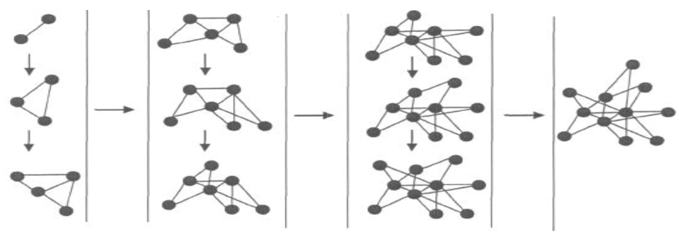

Fig.1. the evolving process of BA network evolving model

The research results show that the degree distribution of BA network evolving model follows power-law distribution, which is consistent with the theoretical analysis result, see Fig 2.

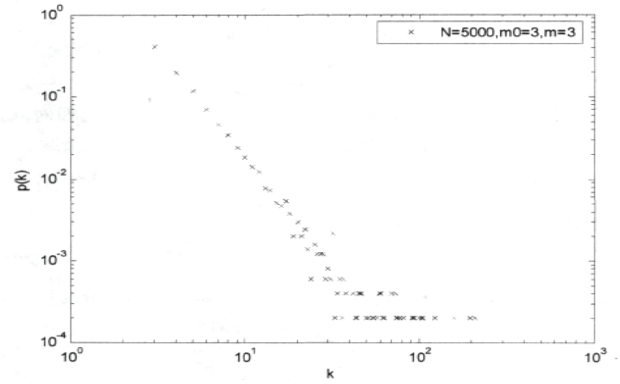

Fig.2. the degree distribution of BA network evolving model

2.2. CALW model

Wensheng Tian and Hongyong Yang proposed CALW model[15] which introduced the concept of attraction and divided it into initial attraction and evolutional attraction. The algorithm is described as follows:

- Initial condition. Initializing the network in which the number of nodes and edges is m_0 and e_0 separately.
- Growth. Introducing a new node A and m edges at every timestep.
- Preference link. Connecting the new node A with the m-1 old nodes in the local world according to preference probability $\prod_{local}(q_i)$.

$$\prod_{local}(q_i) = \prod'_{(i \in local\_world)} \frac{\beta_i}{\sum\limits_{j \in local\_world} \beta_j}$$

The experimental result shows that the lower cluster coefficient problem presented in BA model is well solved in this model, while researchers only consider the factors that facilitate network growth such as initial attraction and evolutional attraction but do not consider the factors that delaying network growth.

3. DFEM evolving model

In this paper the factors delaying network growth are defined as damping factor and divided it into the decline of initial attraction, the loss of node heat, and irresistible natural factors. For instance, for the node j going to entering into blog at t timestep, the major attraction of the node i entered into blog at $t-1$ timestep is initial attraction. However, as time increasing the initial attraction is increasingly decline and the decrease of initial attraction is represented by R_1 in this paper. Moreover, the numbers node i connecting with other nodes per timestep after it entering into blog is defined as node heat. The more times node i connecting with other nodes, the higher the node heat is, and vice versa. If a node whose connection number is smaller than n (here let $n = 7$) at timestep t, this node will be defined as zero-heat node, and representing the number of zero-heat nodes at t timestep in network by R_2. Actually, the remove of some edges is result from some irresistible natural factors, such as natural calamities and computer losing power suddenly, making all these natural calamities represented by R_3.

3.1. DFEM model algorithm

The evolving model proposed in this paper is defined as follows:

- Initialize network. Initializing the network to the entire coupling network with n_0 nodes and e_0 edges.
- Growth. Introducing a new node N at per timestep and connecting this node with the *m* existing nodes.
- Preference link. The probability the new node N connecting with the *m* existing nodes is

$$\prod_i = \frac{(k_i + A_i + D_i(t) - (R_{1i}(t) + R_{2i}(t) + R_{3i}(t))}{\sum_j \left(k_j + A_j + D_j(t) - (R_{1j}(t) + R_{2j}(t) + R_{3j}(t))\right)} \tag{1}$$

where A_i is the initial attraction of the existing node *i* attracting the new node N at t_i timestep, that is the degree of the new node *N*.

$D_i(t)$ is the evolving attraction of node *i* attracting the new node N from t_i to *t*, that is $D_i(t) = \frac{\varnothing(i)}{(t - t_i)}, (t - t_i) > 0$, $\varnothing(i)$ is the newly increase connection number of node i at t timestep.

Damping factor $R_{1i}(t)$ is the decline of the initial attraction from t_i to *t*, that is, $R_{1i}(t) = A_i - \frac{A_i}{(t - t_i)}$, $(t - t_i) > 0$, in which $\frac{A_i}{(t - t_i)}$ is the initial attraction of the existing node *i* at *t* timestep.

Damping factor $R_{2i}(t)$ is the number of zero-heat nodes at *t* timestep from t_i to *t*, that is, $R_{2i}(t) = \frac{\phi_i}{(t - t_i)}$, $(t - t_i) > 0$, in which ϕ_i is the number of zero-heat nodes at *t* timestep.

Damping factor $R_{3i}(t)$ is the decline of edges resulting from some irresistible natural factors. Because it is small probability events, in this paper let $R_{3i}(t) = 0$.

3.2. Theoretical analysis of degree distribution of DFEM model

The degree of node is the number of edges the node connecting to, and degree distribution is a probability distribution function of node degree, that is, the probability that the node having *k* edges. As a kind of scale-free complex network, the degree distribution of online social network follows by power-law. In order to describe conveniently, let θ represent the sum of initial attraction and evolving attraction in formula *(1)*, that is $\theta = A_i + D_i(t)$ and R represent the sum of damping factor $R_{1i}(t)$, $R_{2i}(t)$, and $R_{3i}(t)$, that is $R = R_{1i}(t) + R_{2i}(t) + R_{3i}(t)$.

Assumption that the new node *N* enters network at t_i timestep and the degree of node *i* connecting to N is $k_i(t_i) = m$, so node *N* connects to node *i* according to formula (1) and obtains the following formula

$$\frac{\partial k_i}{\partial t} = m \frac{k_i + \theta_i - R_i}{\sum_j \left(k_j + \theta_j - R_j\right)} \tag{2}$$

The above formula illustrates that as time increasing the change rate of k_i is not only related with k_i and θ but also related with damping factor *R*.

From $k_i(t_i)=m$ and formula *(2)*, the equation of $k_i(t)$ can be deducted

$$k_i(t)=(m+\theta_i-R_i)\left(\frac{2mt+\sum_j(\theta_j-R_j)}{2mt_i+\sum_j(\theta_j-R_j)}\right)-(\theta_i-R_i) \tag{3}$$

From formula *(3)*, the probability that $k_i(t)$ smaller than *k* is

$$\begin{aligned} p(k_i(t)<k) &= p\left(t<\frac{2mt_i+\sum_j(\theta_j-R_j)}{2m(m+\theta_i-R_i)^2}(k+\theta_i-R_i)^2-\frac{\sum_j(\theta_j-R_j)}{2m}\right) \\ &= p\left(t_i>\left(t+\frac{\sum_j(\theta_j-R_j)}{2m}\right)\left(\frac{m+\theta_i-R_i}{k+\theta_i-R_i}\right)^2-\frac{\sum_j(\theta_j-R_j)}{2m}\right) \\ &= 1-p\left(t_i\le\left(t+\frac{\sum_j(\theta_j-R_j)}{2m}\right)\left(\frac{m+\theta_i-R_i}{k+\theta_i-R_i}\right)^2-\frac{\sum_j(\theta_j-R_j)}{2m}\right) \end{aligned} \tag{4}$$

Assumption that the time internal every node entering network is the same, so time distribution is

$$p(t_i)=\frac{1}{m_0+t} \tag{5}$$

From formula *(4)* and formula *(5)*, the following formula is obtained

$$p(k_i(t)<k)=1-\frac{1}{m_0+t}\left(t+\frac{\sum_j(\theta_j-R_j)}{2m}\right)\left(\frac{m+\theta_i-R_i}{k+\theta_i-R_i}\right)^2-\frac{\sum_j(\theta_j-R_j)}{2m}$$

So the node degree distribution is

$$\begin{aligned} p(k) &= \frac{\partial(k_i<k)}{\partial k} \\ &= \frac{2}{m_0+t}\left(t+\frac{\sum_j(\theta_j-R_j)}{2m}\right)(m+\theta_i-R_i)^2(k+\theta_i-R_i)^{-3} \end{aligned} \tag{6}$$

From formula*(6)*, we can deduce that the node degree distribution of network $p(k)$ tends to be $2(m+\theta_i-R_i)^2(k+\theta_i-R_i)^{-3}$, when t tends to infinity. And when θ_i=0 and R_i=0 or $\theta_i=R_i$, $p(k)=2m^2k^{-3}$, this evolving model is BA model.

4. The simulation results and degree distribution of DFEM model

Fig 3 describes the degree distribution of DFEM model in which network size is 10000, m=m_0=5, and A_i=5. Fig 3 compares the degree distribution of DFEM model, BA model, and CALW model. Result shows that the degree distribution of DFEM model is similar with BA and CALW. From this result the degree distribution is not only related with node degree and the attraction itself but also related with damping factors. Moreover, because of damping factors, the degree distribution function of DFEM model is between that of BA model and CALW model.

In order to study how damping factor R_1 and R_2 influence the degree distribution of DFEM model, let R_1 and R_2 take different values. Fig 5 describes the degree distribution when R_1 is different but evolving attraction $D_i(t)$ and R_2 have no change. Fig 6 describes the degree distribution when R_2 is different but evolving attraction $D_i(t)$ and R_1 have no change. From Fig 5 and Fig 6, we can know that damping factor R_1 mainly influences the node whose degree is smaller and R2 mainly influences all of the network nodes especially the node whose degree is lager.

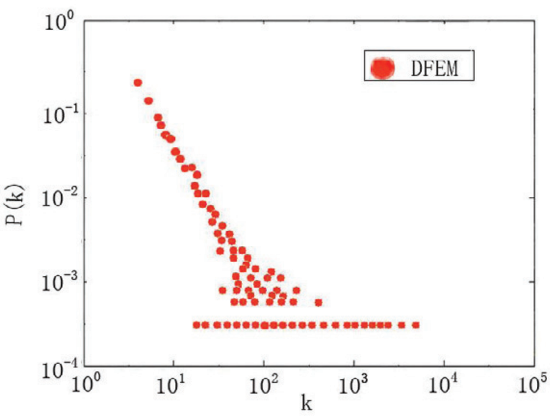

Fig.3. DFEM model

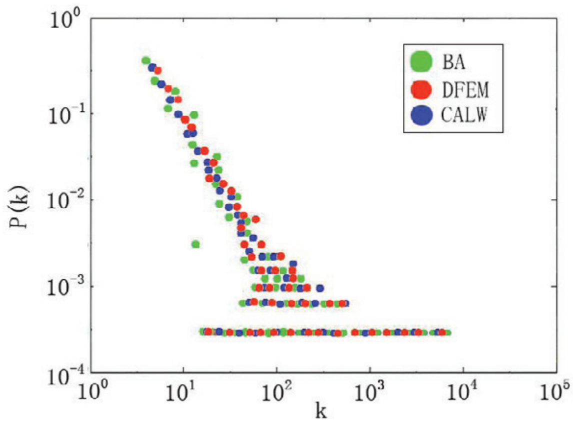

Fig.4. Three models' DD comparison

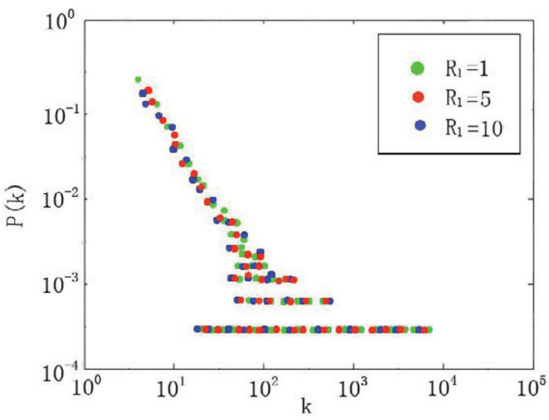

Fig.5. R_1 value is different

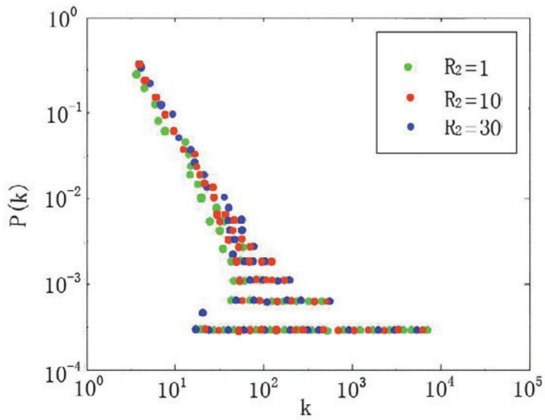

Fig.6. R_2 value is different

5. Conclusions

This DFEM evolving model is based on BA model, but when researching the factors that influent degree distribution, we not only sufficiently consider the node attraction but also take the real existing damping factors into consideration. In this paper we divide damping factors into the decline of initial attraction, the loss of node heat, and irresistible natural factors, and then study how the decline of initial attraction and the loss of node heat effect degree distribution of DFEM model. Using this method, the problem, only unilaterally researches how the node degree and attraction effect the degree distribution, appears in BA model and CALW model. The simulation results illustrate the exact of theoretical analysis which shows that the model proposed in this paper is more suitable for existing online social network.

Acknowledgements

This work is supported in part by Chinese National Science Foundation(Grant No. 61063039) and 2011 graduate student scientific research innovation projects in Guangxi(Project Numbers: 2011105950812M20) . We extend our gratitude to the anonymous reviewers for their insightful comments.

References

1. A. Mislove, M. Marcon, K. P. Gummadi,et al. Measurement and analysis of online social networks[C]. Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement. New York: ACK Press, 2007: 29-42.
2. H.H. Bo, W. Ke. Based on the complex network theory of online social network analysis[J]. Complex system and complexity science, 2008, 5(2): 1-14.
3. S. Golder, D. Wilkinson, B. Huberman. Rhythms of social interaction: Messaging within a massive online network[C]. 3rd International Conference on Communities and Technologies (CT2007), 2007: 41-66.
4. A. L. Barab6si, R. Albert, H. Jeong. Mean-field theory for scale-free random networks[J]. Physica A, 1999, 272: 173-187.
5. A. L. Barab6si, R. Albert. Emergence of scaling in random networks[J]. Science, 1999, 286: 509-512.
6. A. L. Barab6si, R. Albert, H. Jeong. Power-law distribution of the world wide web[J]. Science, 2000, 287: 2115a.
7. P. L. Krapicsky, S. Redner, F. Leycraz. Connectivity of growing rando networks[J]. Phys Rev Lett, 2000, 85: 4629-4632.
8. A. L. Barab6si , R. Albert . Statistical mechanics of complex networks[J]. Reviews of Modern Physics, 2002, 74(1): 47-98.
9. S. N. Dorogovstsev, J. F. F. Mendes. Scaling properties of scale-free evolving networks: Continuous approach[J]. Phys Rev E, 2001, 63: art. no. 056125 Part 2(1-19).
10. S. N. Dorogovstsev, J. F. F. Mendes. Evolution of reference networks with aging[J]. Phys Rev E, 2000, 62: 1842-1847.
11. G. Bianconi, A. L. Batabasi. Topology of evolving networks:local events and universality[J]. Phys Rev Lett, 2000, 85: 5234-5237.
12. W.N. Wang, J.Y. Lee, Z.K. Mi. Scale-free statistics of bulletin Board system server visiting[J]. Chinese Physics Letters, 2004, 21(2): 243-245.
13. S. N. Dorogovstsev, J. F. F. Mendes. Evolution of networks[M]. Oxford University Press, 2003.
14. J. Jost, M. P. Joy. Evolving networks with distance preferences[J]. Phys Rec E, 2002,66: 36126-36132.
15. S.W. Tian, H.Y. Yang. Based on clustering effect of complex network node attractive model[J]. Computer Engineering, 2010, 35(10): 58-60.