An Evolutionary Model of Reply Networks on Bulletin Board System

Min Wu, Hui Li*, Ke Zhang, Lijuan Qin
Educational Technology Department, Capital Normal University, Beijing, 100048, China
E-mail: wuminmosquito@sina.com

Abstract—**In order to study the relation between structure and behaviors of networks, an evolutionary model of reply networks on Bulletin Board System (BBS) is proposed in this paper. With the theories of scale-free model, we built the Barabási–Albert (BA) model on the directed network for the characteristics of reply network. Comparing the results of BA model with the virtual data, there exists a deep gap between them, and a new model will be proposed to narrow the gap under the deeply analysis of the reply network on BBS. The new model is closer to real network than BA model, but the gap is just narrower not disappear, so how to fill the gap is our next job.**

Key words-evolutionary model; reply network; scale-free; BA model

I. INTRODUCTION

With the deep researches on the complex networks, the relation between structure and behaviors of networks has been a hot issue [1]. It is well-known that the small-world model and scale-free model are the typical complex network models. Considered that this paper is based on the scale-free model, we will skip the small-world model. The origin of scale-free model was first addressed by Barabási and Albert in 1999 which was named after the two authors as BA model [2]. They took growth and preferential attachment shared by many real networks into consideration to built a model closer to reality.

The limits still exist because there is a deep gap between the results and reality. Many researchers have taken efforts to narrow the gap and a lot of new models have been put forward, such as the model based on nodes attraction [3], the model based on fitness [4]. Most of researches build undirected network based on the two regulations of BA model. But the structure of reply networks on BBS is directed representing the reply relationship. Building a directed network based on BA model is good for studying the evolutionary process of reply networks and search out its limits. A new model can be put forward to study the relations between network structure and network behavior for understanding the form and development of reply networks.

This paper is organized as followed: we build the directed network based on BA model and analyze the characteristics of the results in Section II, then a new model is proposed after the deeply analysis of the reply network on BBS, and finally a conclusion is drawn.

II. DIRECTED REPLY NETWORK BASED ON BA MODEL

A. Building the directed network

These two ingredients, growth and preferential attachment, are the spirit of the BA model, which led for the first time to a network with a power-law degree distribution. The algorithm of the BA model is the following [2]:

Growth: Starting with a small number (m_0) of nodes, at every time step, we add a new node with $m(\le m_0)$ edges that link the new node to *m* different nodes already present in the system.

Preference attachment: When choosing the nodes to which the new node connects, we assume that the probability $\prod$ that a new node will be connected to node i depends on the degree k_i of node i, such that

$$\prod(k_i) = \frac{k_i}{\sum_j k_j} . \quad (1)$$

After t time steps, the network will have $t + m_0$ nodes and edges. It is tested that the exponent of a power law is close to 3 while the clustering of coefficient and the characteristic path length close to those in random networks with equal number of nodes and links [2].

In the directed network, a node has input degree and out degree while the links between two nodes have two directions. So the ingredient, preference attachment, should be modified as following to take the characteristics of directed networks into consideration.

Preference attachment in directed network: When choosing the nodes to which the node connects, we assume that the probability $\prod$ that a new node will be connected to node i depends on the sum of input degree $k_{(i)output}$ and output degree $k_{(i)input}$ of node i, such that

$$\prod(k_i) = \frac{k_{(i)input} + k_{(i)output}}{\sum_j k_{(j)input} + k_{(j)output}} . \quad (2)$$

The direction between the new node and node i has two possibilities which can be determined by random selecting.

B. Result of the directed network based on BA model

Based on the modified BA model, we have got data for a month. In [5], we have built the reply network on BBS based on the data downloaded from a famous and hot BBS with

978-0-7695-4522-6/11 $26.00 © 2011 IEEE
DOI 10.1109/ICM.2011.92

network grab tool. We will compare the network produced by BA model with the data on board one from three sides, clustering of coefficient, characteristic path length and degree distribution, which are the prominent characteristics of networks.

Clustering of coefficient: A common property of complex networks is that clique's form represents circles of friends or acquaintances in which every member knows every other. This inherent tendency to clustering is quantified by the clustering coefficient. In a real social network of acquaintances, it reflects the average extend to which friend of given individuals' are also friends of each other. The detail of its algorithm can be got in [6].

Characteristic path length: The path length is the number of links in the shortest path between two nodes. The characteristic path length is the path length averaged over all pairs of nodes. The shorter the characteristic path length is, the closer the nodes' relationship is. The detail of its algorithm can be got in [7].

Degree distribution: The degree distribution $P(k)$ characterizes the spread in the number of links each node has, which gives the number of IDs who have exactly k links on the reply network. A scale-free distribution shows a straight line on a logarithmic scale histogram, and in which, the value of exponent λk corresponds to the slope of the straight line. The detail can be got in [8].

Fig. 1 shows the clustering of coefficient of the network produced by BA model and the real network. Fig. 2 shows the characteristic path length between the two networks. It is obviously observed that the clustering of coefficient of real network is larger than that of the network produced by BA model and the characteristic path length of real network is shorter than that of the other. That is to say, the network produced by BA model does not fit the characteristic of small-world of real network. Fig. 3 shows s the exponent of the power-law of the two networks. The exponent with the time of the network produced is close to 3 while the scale of the exponent of reply networks is from 1 to 2 [9]. Though the network produced with long tail, the value of the exponent does not belong to the reply networks.

As for the result of BA model, a new model should be proposed to be fitter.

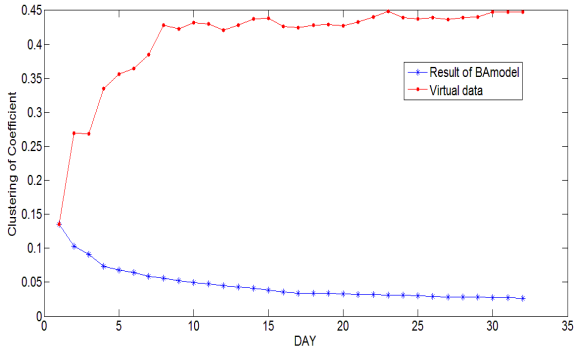

Fig. 1 Changes of clustering of coefficient with the time

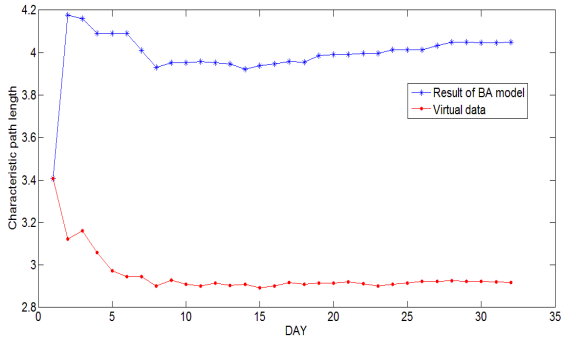

Fig. 2 Changes of characteristic path length with the time

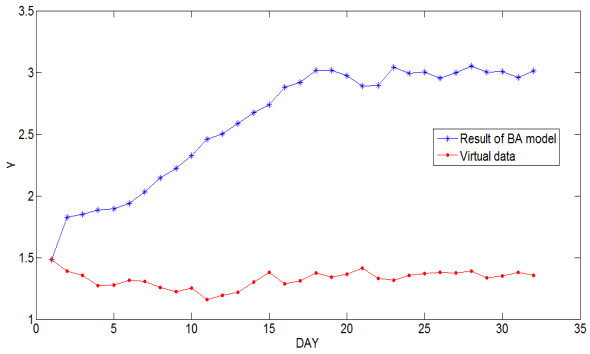

Fig. 3 Changes of the exponent with the time

III. A NEW MODEL PROPOSED

A. *Deeply analysis on the reply networks*

In [5], we have studied the additional node and links with the time. As we all known, a new node connected into the reply network have two methods, posting a reply article or posting an initial article. The additional of new nodes lead to a part of additional links. The other part of additional links contributes to the new links among the nodes already present in reply networks.

Fig. 4 shows the radio between sum of new nodes and new nodes post initial articles with the time. The amount of new nodes post initial articles is less than 50% of all the new nodes. So when a new node connects to node i, the direction of the link between them cannot be chose in random method.

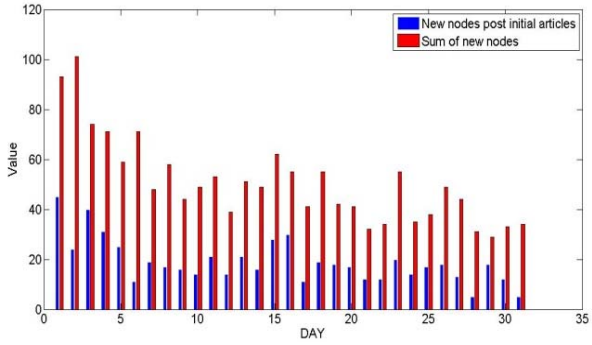

Fig. 4 Ratio between sum of new nodes and new nodes post initial articles with the time

Fig. 5 shows the ratio between all additional links and links contributing to new nodes posting initial articles with the time. It is obvious that only a few of the new links are caused by the new nodes. So the new links must be divided into two parts and most of them are caused among the nodes existed in the network.

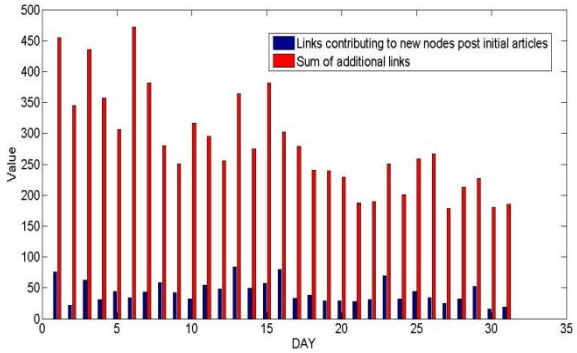

Fig. 5 Ratio between all additional links and links contributing to new nodes posting initial articles with the time

B. *Algorithm of the new model*

Based on BA model and the analysis on real reply network, the algorithm is the following:

This new model holds on two ingredients of BA model in the directed network, but in the growth ingredient, we will assume the sum of the new links. When a new node connects to node i, the direction of the link between them cannot be chose in random method. It depends on the real network.

Most of the new links are produced in the initial network, so we will add another ingredient, namely producing links in initial network [10]. When a new link adds into the network, two nodes should be chosen following the preference attachment ingredient.

C. *Result of the new model*

Fig. 6 shows the change of clustering of coefficient with the time between the network based on the new model and the real reply network, while Fig. 7 shows characteristic path length. Under the comparison, the structure of network fits the characteristic of small-world. In Fig. 8, we discover that the exponents belong to the scale of real reply network. In a word, the new network based on the new model is closed to real reply network than BA model.

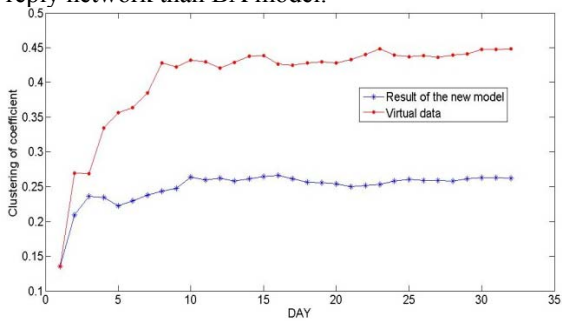

Fig. 6 Changes of clustering of coefficient with the time

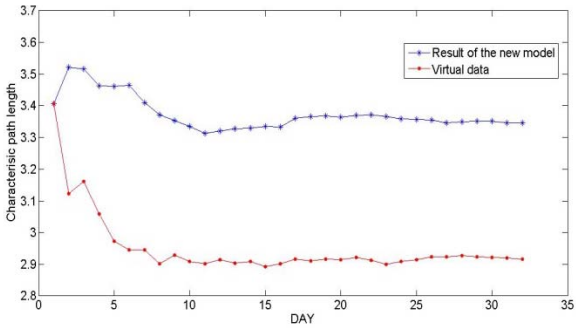

Fig. 7 Changes of characteristic path length with the time

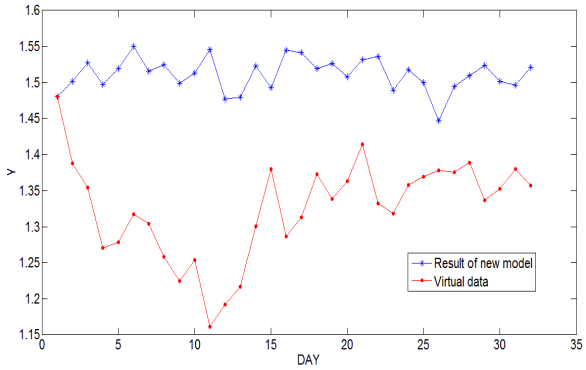

Fig. 8 Changes of the exponent with the time

IV. CONCLUSION

In this paper, we have transferred BA model from undirected graph to directed and proved that the network based on BA model does not fit the reply network without small-world characteristic. The exponent of degree distribution which is close to 3 is also far away from the real data.

Under the research in [5], we have analyzed the real reply network from the ratios. We found that the amount of new nodes post initial articles was less than 50% and most of the new links were caused among the nodes existed in the network.

With the jobs we have done, a new model was proposed. The network based on the new model was fitter than the BA model narrowed the gap between network we built and the real one. But the gap does not disappear and how to make it narrower is our next job.

ACKNOWLEDGMENT

This work is supported by National Social Science Foundation of P.R China (Grant No. 10CTQ012), Funding Project for Academic Human Resources Development in Institutions of Higher Learning under the Jurisdiction of Beijing Municipality (Grant No.PHR201108137)

REFERENCE

[1] L. Zhang and Y. Liu, "Research on the evolution process of virtual community networks", Chin. Phys. Soc, vol. 57, pp. 5419-5424, 2008. (in Chinese)

[2] R. Albert and A. L. Barabási, "Statistical mechanics of complex networks", Reviews of Modern Physics, vol. 74, pp. 47-97, 2002.

[3] S. H. Tao, C. Yang, H. N. Li and Y. Zhang, "Research on complex networks evolution model based on node attraction", Computer Engineering, vol. 35, pp. 111-113, 2009. (in Chinese)

[4] G. Bianconi and A. L. Barabási, "Bose-Einstein condensation in complex networks", Physical Review Letters, vol. 86, pp.5632-5635, 2001.

[5] M. Wu, H. Li, K. Zhang and L. J. Qin, "Studies on the evolutionary characteristics of reply network on BBS", ICNECS 2011, in press.

[6] Z. B. Kou and C. S. Zhang, "Reply Network on BBS", Phys. Rev. E, vol. 67, pp. 1-7, 2003.

[7] W. Meng, "Application of Complex Network Analysis in Scientific Co-authorship Network of Information Science", Beijing: ISTIC, 2008. (in Chinese)

[8] M. E. J Newman, "The Structure and Function of Complex Network", SIAM REVIEW, vol. 45, pp. 167-256, 2003.

[9] L. Wang and G. Z. Dai, "On Degree Distribution of Complex Network", Journal of Northwestern Polytechnical University, vol. 24, pp.405-409, 2006. (in Chinese)

[10] Y. Wu, K. Z. Xiao, H. T. Liu and H. Tang, "Evolution of BBS Virtual Community and its Simulation", Systems Engineering--Theory & Practice, vol. 30, pp.1183-1190, 2010. (in Chinese)