

Step-by-Step Random Walk Network with Power-Law Clique-Degree Distribution *

YANG Han-Xin(杨涵新)^{1**}, WANG Bing-Hong(汪秉宏)^{1,3***}, LIU Jian-Guo(刘建国)^{1,2},
HAN Xiao-Pu(韩筱璞)¹, ZHOU Tao(周涛)^{1,2}

¹Department of Modern Physics and Nonlinear Science Center, University of Science and Technology of China, Hefei 230026

²Department of Physics, University of Fribourg, Chemin du Muse 3, CH-1700 Fribourg, Switzerland

³Institute of Complex Adaptive System, Shanghai Academy of System Science, Shanghai 201800

We propose a simple mechanism for generating scale-free networks with degree exponent $\gamma = 3$, where the new node is connected to the existing nodes by step-by-step random walk. It is found that the clique-degree distribution based on our model obeys a power-law form, which is in agreement with the recently empirical evidences. In addition, our model displays the small-world effect and the hierarchical structure.

PACS: 89.75.Fb, 89.75.Hc, 05.40.Fb

In the past few years, the study of complex networks has drawn many interests from natural and social scientists.^[1-7] Previously empirical evidences have shown that small-world effect^[8] and scale-free property^[9] widely exist in many natural and social networks, such as the World Wide Web (WWW),^[10,11] email communication networks,^[12,13] food webs,^[14] scientific and technological collaboration networks,^[15-19] sexual relations,^[20] opinion networks,^[21] and so on. Small-world effect reflects that networks have a small average distance and a high clustering coefficient. A typical small world network model was proposed by Watts and Strogatz.^[22] Scale-free property indicates that the degree distribution obeys a power law $P(k) \sim k^{-\gamma}$, where k is the degree and $P(k)$ is the probability density. A pioneering model which generates power-law degree distribution was presented by Barabási and Albert (BA).^[9]

Very recently, Xiao *et al.*^[23] have investigated the distributions of clique-degree, which are the extension of degree and can be used to measure the density of cliques in networks. The word clique in the graph theory equals the term complete subgraph.^[24] That is to say, the s order clique (s -clique for short) represents a fully connected subgraph with s nodes and $s(s-1)/2$ edges. The s -clique degree of a node i , denoted by k_i^s , is defined as the number of different s -cliques containing i . Obviously a two-clique is an edge and k_i^2 equals the degree k_i . An illustration of the clique-degree of node i is shown in Fig. 1. It has been found that many real-world networks exhibit power-law clique-degree distributions.^[23]

The BA model is constructed as follows: starting from m_0 nodes, every time step a new node with m edges is added and each edge of the new node is attached to an existing node i with the probability proportional to the degree of the i th node. This mechanism results in power-law degree distribution with exponent $\gamma = 3$. Note that the new node can obtain the global information of existing nodes' degrees in the BA model. While in many real-world networks such as social networks and the WWW, no one knows exact information of global structure. On the other hand, m edges of the new node are connected to the existing nodes independently in the BA model, and there is no relation among m edges. While in the social network, people are more likely to know a new friend through those friends they have already known. Analogously in the WWW, people can find out new web pages by using the hyperlinks of those web pages they have already known. Based on the above considerations, in this Letter we propose a growing network model which introduces the mechanism of the step-by-step random walk.

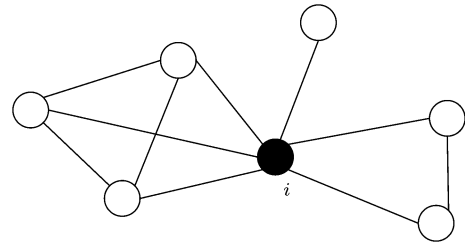


Fig. 1. Illustration of the clique degree of node i (black dot): $k_i^2 = 6$, $k_i^3 = 4$, $k_i^4 = 1$, $k_i^5 = 0$.

Supported by the National Basic Research Programme of China under Grant No 2006CB705500, the National Natural Science Foundation of China under Grant Nos 60744003, 10635040, 10532060 and 10472116, the Special Research Funds for Theoretical Physics Frontier Problems (NSFC Nos 10547004 and A0524701), the President Funding of Chinese Academy of Sciences, and the Specialized Research Fund for the Doctoral Programme of Higher Education of China.

**Email: hxyang@mail.ustc.edu.cn

***Email: bhwang@ustc.edu.cn

Our model can be constructed as follow: (1) Initial condition: the model starts with m_0 fully connected nodes. (2) Growth: a new node with m ($\leq m_0$) edges is added at every time step. (3) Step-by-step random walk: a random node is chosen as the starting point of the random walk, at the same time this node is marked. At each step of the walk, the walker moves to a randomly chosen neighbour (not yet marked) of the currently marked node, and the newly chosen node is also marked. After $m - 1$ ($m \geq 2$) steps, there are m marked nodes, and the new node is connected to the m marked nodes by undirected links.

We need to point out that this step-by-step random walk is self-avoiding, and the walker is not allowed to trace its steps backwards. Note that m edges of the new node are interdependent, each edge (except the first one) is connected to a randomly chosen neighbour of the prior neighbour. When $m = 2$, our model degenerates to the model proposed by Dorogovtsev *et al.*^[25] In all our following simulations, we set $m = m_0$.

We firstly investigate the degree distribution of our model. As shown in Fig. 2, the degree distributions exhibit a power-law behaviour, and the power-law exponents are about 3 for different m . It is interesting to note that if m edges of the new node are connected to randomly chosen nodes independently, the degree distribution will be exponential.^[9] While the mechanism of the step-by-step random walk can induce the emergence of scale free structure as the linear preferential attachment does.

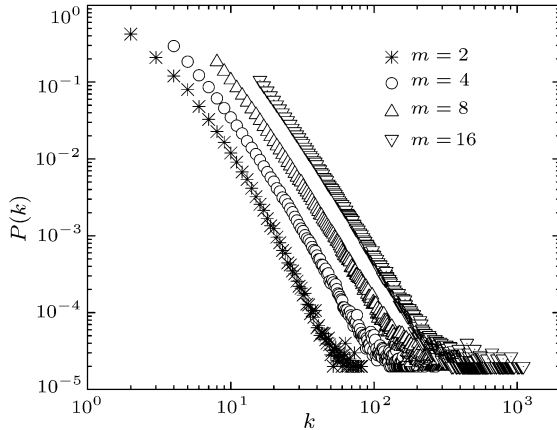


Fig. 2. Degree distributions of our model. Network size $N = 5 \times 10^4$, each data is gained by averaging over 20 independent runs. The power-law exponents γ of four curves are $\gamma_{m=2} = 3.04 \pm 0.05$, $\gamma_{m=4} = 2.98 \pm 0.04$, $\gamma_{m=8} = 2.93 \pm 0.03$, $\gamma_{m=16} = 2.91 \pm 0.02$.

From Fig. 3, one can observe that distributions of clique-degree on our model also exhibit a power-law form. Those higher-order ($s \geq 3$) power-law clique-degree distributions are extensively existent in many real-world networks,^[23] which could not be reproduced by the preferential attachment.^[9] Therefore, the evolving mechanism proposed in this study may

be helpful for further explorations on the clique-degree distribution.

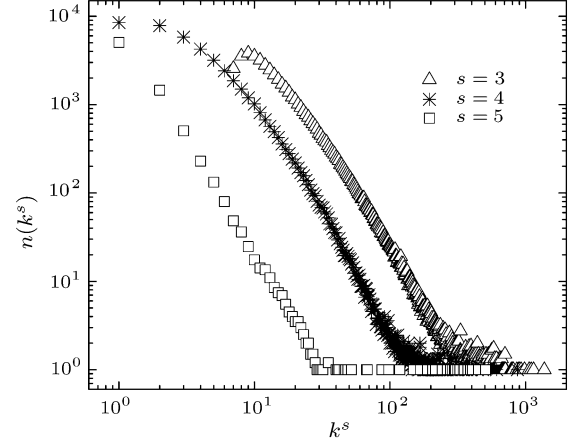


Fig. 3. Clique-degree distributions with the clique size $s = 3, 4, 5$, where k^s denotes the s -clique degree and $n(k^s)$ is the number of nodes with s -clique degree k^s . We set $N = 5 \times 10^4$ and $m = 8$. Each data is gained by averaging over 20 independent runs. The power-law exponents γ of three curves are $\gamma_{s=3} = 2.32 \pm 0.04$, $\gamma_{s=4} = 2.40 \pm 0.04$, $\gamma_{s=5} = 2.52 \pm 0.05$.

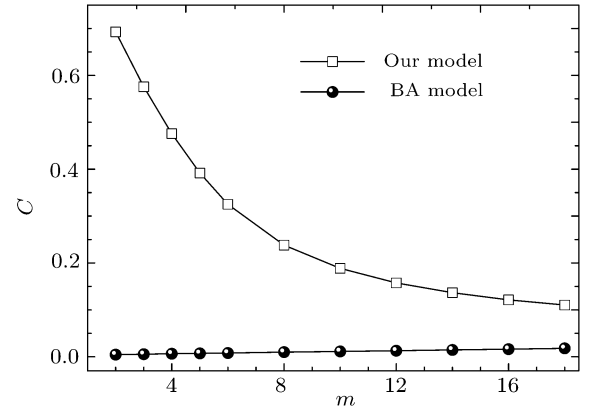


Fig. 4. Clustering coefficient as a function of m for our model and the BA model. The size of network is fixed as $N = 10^4$. Each data is gained by averaging over 20 independent runs.

We next check if the current model exhibits small-world effect. One of the important measures to judge small-world effect is the clustering coefficient. The clustering coefficient for node i is defined as

$$C(i) = \frac{2E(i)}{k_i(k_i - 1)}, \quad (1)$$

where $E(i)$ is the number of edges among node i 's neighbours, and k_i is the degree of node i . The clustering coefficient C of the whole network is defined as the average of $C(i)$ over all nodes, i.e.

$$C = \frac{1}{N} \sum_{i=1}^N C(i). \quad (2)$$

As shown in Fig. 4, the clustering coefficient of our model is much larger than that of the BA model.

Another important measure to judge the small-world effect is the average distance, which is defined as the mean distance over all pairs of nodes. From Fig. 5, one can see that the average distance approximately has a logarithmic dependence of network size: $L \sim \log N$. This means that the average distance increases very slowly with the increasing network size. From Figs. 4 and 5, it can be concluded that our model has the small-world effect.

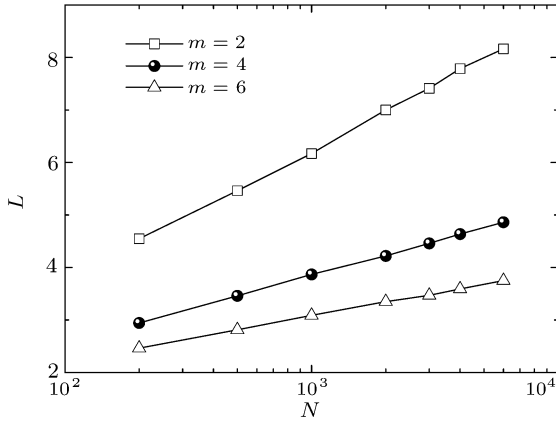


Fig. 5. Relationship between the average distance L and the network size N in a linear-log plot. Each data is gained by averaging over 20 independent runs.

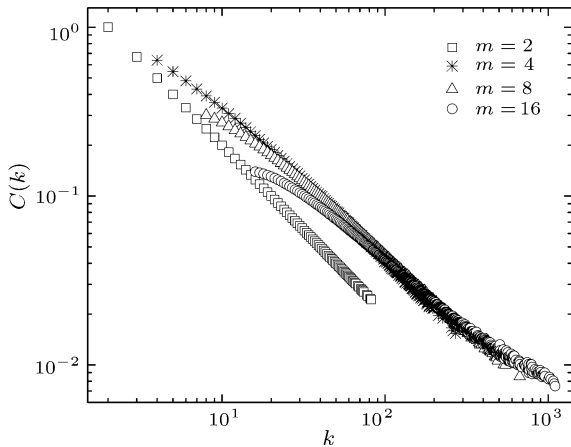


Fig. 6. Clustering coefficient as a function of the vertex degree for different m , network size $N = 5 \times 10^4$. Each data is gained by averaging over 20 independent runs. The power-law exponents γ of four curves are $\gamma_{m=2} = 1$, $\gamma_{m=4} = 0.92$, $\gamma_{m=8} = 0.87$, $\gamma_{m=16} = 0.75$.

In addition, many real-world networks including Internet, World Wide Web, the actor collaboration network, as well as some well-known models like the Holme-Kim model [26] and random Apollonian network [27] random Apollonian network are characterized by the existence of hierarchical structure, [28,29] which can usually be detected by the negatively power-law correlation between the clustering coefficient and the degree. It is known that the BA network

does not possess hierarchical structure and the clustering coefficient $C(i)$ of node i is independent of its degree k_i . [28] From Fig. 6, we find a power-law clustering-degree correlation in our model, i.e. $C(k) \sim k^{-\gamma}$, and the power-law exponent decreases with the increasing m .

In summary, we have proposed a growing network model based on the mechanism of the step-by-step random walk, which generates a power-law clique-degree distribution without utilizing any knowledge of node degree. The high clustering coefficient and small average distance show that our model has the small-world effect. In addition, our model possess hierarchical structure: $C(k) \sim k^{-\gamma}$, in agreement with the observations of many real networks. The evolving mechanism, highlighted in this study, is inspired from the real growing processes of social networks and WWW, and can reproduce many statistical properties shared by various kinds of real networks. Therefore, we believe this model will shed some light in the in-depth understanding of structural evolution of real networks.

References

- [1] Watts D J 1999 *Small Worlds* (Princeton: Princeton University Press)
- [2] Barabási A L 2002 *Linked: The New Science of Networks* (Cambridge: Perseus)
- [3] Newman M E J 2000 *J. Stat. Phys.* **101** 819
- [4] Hayes B 2000 *Am. Sci.* **88** 104
- [5] Strogatz S H 2001 *Nature* **410** 268
- [6] Albert R and Barabási A L 2002 *Rev. Mod. Phys.* **74** 47
- [7] Newman M E J 2003 *SIAM Rev.* **45** 167
- [8] Amaral L A N et al 2000 *Proc. Natl. Acad. Sci. U.S.A.* **97** 11149
- [9] Barabási A L and Albert R 1999 *Science* **286** 509
- [10] Albert R, Jeong H and Barabási A L 1999 *Nature* **401** 130
- [11] Barabási A L et al 2000 *Science* **287** 2115
- [12] Ebel H et al 2002 *Phys. Rev. E* **66** 035103(R)
- [13] Newman M E J et al 2002 *Phys. Rev. E* **66** 035101(R)
- [14] Camacho J et al 2002 *Phys. Rev. Lett.* **88** 228102
- [15] Newman M E J 2001 *Phys. Rev. E* **64** 016131
- [16] Newman M E J 2001 *Proc. Natl. Acad. Sci. U.S.A.* **98** 404
- [17] Jin E M et al 2001 *Phys. Rev. E* **64** 046132
- [18] Li M H et al 2007 *Physica A* **375** 355
- [19] Zhang P P et al 2006 *Physica A* **360** 599
- [20] Liljeros F et al 2001 *Nature* **411** 907
- [21] Zhou T, Ren J, Medo M and Zhang Y C 2007 *Phys. Rev. E* **76** 046115
- [22] Watts D J and Strogatz S H 1998 *Nature* **393** 440
- [23] Xiao W K, Ren J, Qi F, Song Z W, Zhu M X, Yang H F, Jin H Y, Wang B H and Zhou T 2007 *Phys. Rev. E* **76** 037102
- [24] Derényi I, Palla G and Vicsek T 2005 *Phys. Rev. Lett.* **94** 160202
- [25] Dorogovtsev S N, Mendes J F F and Samukhin A N 2001 *Phys. Rev. E* **63** 062101
- [26] Holme P and Kim B J 2002 *Phys. Rev. E* **65** 026107
- [27] Zhou T, Yan G and Wang B H 2005 *Phys. Rev. E* **71** 046141
- [28] Ravasz E and Barabási A L 2003 *Phys. Rev. E* **67** 026112
- [29] Trusina A, Maslov S, Minnhagen P and Sneppen K 2004 *Phys. Rev. Lett.* **92** 178702