The Eighth International Symposium on Operations Research and Its Applications (ISORA'09)
Zhangjiajie, China, September 20–22, 2009

An Evolving Network Model With Local-World Structure*

Li Tan† Xiangxing Kong

School of Mathematics, Central South University, Changsha, Hunan, 410075, China.

Abstract In this paper, a new local-world evolving network model including triad formation mechanism is proposed and studied. Analytical expressions for degree distribution and clustering coefficient are derived using continuum theory. It is shown that the degree distribution transfers from exponential to power-law scaling, and the clustering coefficient is tunable with known parameters. Finally, the numerical simulations of degree distribution are given, and we can observe our analytical results are in good agreement with the simulations.

Keywords Local-world; Degree distribution; Clustering coefficient

1 Introduction

Networks exist in every aspect of our life and society. Many complex systems such as World Wide Web(WWW)[1], Internet[2], movie actor collaboration[3], human sexual contacts[4] and so on can be described as complex networks. Despite their diversities, such kind of networks have some common characteristics, like power-law degree distribution, high clustering coefficient and short average path length.

To model these systems and capture the properties of real networks, a lot of studies have been done in this field. At the beginning, the work was focused on regular graphs. Later, Erdös and Rényi[5-6] built the random graph theory, which was thought as the simplest realization of complex networks to explain real world phenomena. Being aware of the truth that many real complex networks are neither completely regular nor completely random, Watts and Strogatz[7] introduced the small-world model. It is found that this model displays high clustering and short average path length of real networks. The common feature of small-world model and random graph model is that the probability $P(k)$ that a node in the network connected to k other nodes is bounded, decaying exponentially for large k.

Although the small-world model captures some important traits of real networks, Barabási and Albert[1][3][8][9] found that the degree distribution of real networks are not Poisson but power-law. They had researches on WWW and revealed that the degree distribution of WWW is power-law[1]. In[8], they stated a growing network(BA model) with growth and preferential attachment. They used the mean-field method to compute

*Supported by the National Natural Science Foundation of China (Grant No. 10671212, 90820302)
†Email: tltanli@yahoo.com.cn

the degree distribution. It is shown that after many steps the proportion of nodes with degree k is power law, i.e. $P(k) \sim k^{-3}$. They also pointed out that most real networks are scale-free[9]. In order to make a better understanding of the phenomena in real networks, many scholars did researches on BA model, they also had extensions to BA model and got a lot of useful results. Bollobás[10] presented a modified BA model allowing for loops and multiple edges. Homle and Kim[11] extended the BA model to include triad formation(TF) step, which causes the clustering coefficient to be tunable.

Motivated by the fact that only part of network information is needed in networks like world trade web[12], scientific citation network, Li and Chen[13] put forward a local-world(LC) model and obtained a transition between power-law and exponential scaling to degree distribution. Following them, Wang and Zhang[14] studied an evolving model of scale-free network with adjustable clustering. Zhang and Rong[15] modified the LC model by adding an additional triad formation step.

Real networks like interpersonal relationship network, people are inclined to make friends with the similar background or hierarchy, thus give rise to local-world or communities. In a certain local world, new friends will come in to make friends with people already in the local-world, sometimes people introduce new friend to his old friend. Except for this, people already in network will open relationships among each other from time to time. Based on features of real networks, we propose a new extensive network model which take many possible cases together. Statistical characteristics like degree distribution and clustering coefficient are analyzed. Numerical simulations are also performed to inspect the validity of our theoretical results. It is found that our model unifies generic properties of real-life networks.

2 Model description

Start with a small number of m_0 nodes and e_0 edges. At each time step, select M nodes randomly from the present network as the local world. With probability p add a new node, and with probability $1-p$ add m edges between old nodes.

(a) When adding a new node, the new node v connects to m different nodes in the local-world of node v. For each of the m steps, we perform a LPA step first, then with probability q there is an additional TF step followed.

Local Preferential Attachment (LPA) step: The probability that an old node i of the local-world get a degree is given by:

$$\prod_{local}(k_i) = \prod(i \in local)\frac{k_i}{\sum\limits_{j \in local} k_j}, \tag{1}$$

where $\prod(i \in local) = M/N_t$, N_t is size of the network at time t.

Triad Formation(TF) step: In the previous LPA step if there is an edge connecting the new node v and an existing node i, then we add one more edge from node v to a randomly selected neighbor of node i with probability q. If all neighbors of node i have been connected to node v, this step is omitted.

(b) When adding edges between old nodes, firstly select a node from the local-world preferentially, then add an edge among its two different neighbors selected randomly. Repeat this step until m edges are added.

Note that $m \leq M \leq m_0 + pt$. After t time steps, the model develops to a network with $N_t = m_0 + pt$ nodes and expected $E_t = e_0 + mt(1+pq)$ edges. Thus the average degree of the network is

$$< k >= \frac{2e_0 + 2mt(1+pq)}{m_0 + pt} \approx \frac{2m(1+pq)}{p}. \quad (2)$$

3 Network analysis

We use continuum theory and rate equations to analyze two important statistical characteristics of the network. One is degree distribution, the other is clustering coefficient.

3.1 Degree distribution

A: Case of $M = m$

According to continuum theory, we assume that the degree k_i of node i is continuous, thus the degree change rate of node i is:

$$\frac{\partial k_i}{\partial t} = pm\frac{M}{N_t}\frac{1}{M} + pmq\frac{M}{N_t}\sum_{n\in\Omega_i}\frac{1}{M}\frac{1}{k_n} + 2(1-p)m\frac{M}{N_t}\sum_{n\in\Omega_i}\frac{1}{M}\frac{1}{k_n}. \quad (3)$$

Where Ω_i is the set of neighbors of node i, and k_n represents the degree of neighbor node n of node i. Simplify Eq.(3) and use the approximation of $< k >$ to replace k_n, so we have

$$\frac{\partial k_i}{\partial t} \approx \frac{m}{t} + \frac{2-2p+pq}{2(1+pq)}\frac{k_i}{t}. \quad (4)$$

Denote $\frac{2(1+pq)}{2-2p+pq} = A_{pq}$. Solve this equation with initial condition $k_i(t_i) = m(1+q)$, and the probability density $P_i(t_i) = \frac{1}{m_0+pt}$, thus the degree distribution $P(k)$ is given by:

$$P(k) = \frac{A_{pq}}{p}\left(\frac{m(pq^2+pq-2p+2q+4)}{2-2p+pq}\right)^{A_{pq}}(k+mA_{pq})^{-A_{pq}-1}. \quad (5)$$

This can be write in the form $P(k) \sim (k+\kappa)^{-\gamma}$, where $\kappa = mA_{pq} = \frac{2m(1+pq)}{2-2p+pq}$ and $\gamma = A_{pq}+1 = \frac{4+3pq-2p}{2-2p+pq}$. When $k \gg \kappa$, $P(k) \sim k^{-\gamma}$. When $k \ll \kappa$, $ln[P(k)] \sim -\gamma ln(k+\kappa) = -\gamma[ln(1+\frac{k}{\kappa}) + ln\kappa] \sim -\gamma[\frac{k}{\kappa} + ln\kappa]$. Thus

$$P(k) \sim \frac{1}{\kappa^\gamma}exp(-\frac{\gamma k}{\kappa}). \quad (6)$$

That is to say, in this case, the degree distribution is of exponential form.

B: Case of $m < M \leq m_0 + pt$

For node i with degree k_i, k_i changes as

$$\begin{aligned}\frac{\partial k_i}{\partial t} = {} & pm\frac{M}{m_0+pt}\frac{k_i}{\sum\limits_{j\in local}k_j} + pmq\frac{M}{m_0+pt}\sum_{n\in\Omega_i}\frac{k_n}{\sum\limits_{j\in local}k_j}\frac{1}{k_n} \\ & +2(1-p)m\frac{M}{m_0+pt}\sum_{n\in\Omega_i}\frac{k_n}{\sum\limits_{j\in local}k_j}\frac{1}{k_n}.\end{aligned} \quad (7)$$

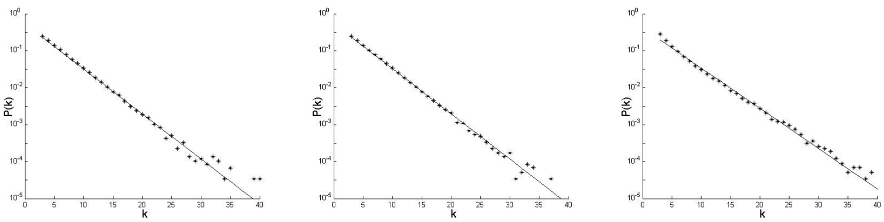
The first item in the sum is due to the degree increase of node i in a LPA step. The second item is a TF step. The last item is the edges adding between old nodes. Node i is one of the neighbors of preferentially selected node n in the local-world, then node i is randomly selected to add an edge to another neighbor of node n. Using (2) and the assumption that $\sum_{j \in local} k_j = M < k >$[13], we can get

$$\frac{\partial k_i}{\partial t} \approx \frac{pq - p + 2}{2(1+pq)} \frac{k_i}{t}. \tag{8}$$

With the initial condition, $k_i(t_i) = m(1+q)$, we have

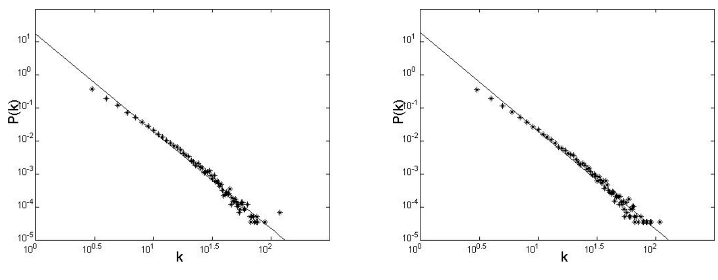
$$P(k) = \frac{2(1+pq)}{p^2 q - p^2 + 2p} [m(1+q)]^{\frac{2(1+pq)}{pq-p+2}} k^{-\frac{3pq-p+4}{pq-p+2}} \sim k^{-\frac{3pq-p+4}{pq-p+2}}. \tag{9}$$

This means that the degree distribution of $m < M \leq m_0 + pt$ follows a power law distribution, and the scaling exponent is $\gamma = \frac{3pq-p+4}{pq-p+2}$.



(a) p=0.98, q=0.2, m=3, M=3 (b) p=0.98, q=0.97, m=3, M=3 (c) p=0.9, q=0.97, m=3, M=4

Figure 1: Degree distribution $P(k)$ versus k in linear-log scale with $M \approx m$. The size of network is 60000. $*$ are the simulation results, while the solid lines are theoretical results predicted by Eq.(6).



(a) p=0.98, q=0.2, m=3, M=20 (b) p=0.98, q=0.97, m=3, M=20

Figure 2: Degree distribution $P(k)$ versus k in log-log scale with $M > m$. The size of network is 60000. $*$ are the simulation results, while the solid lines are theoretical results predicted by Eq.(9).

In order to verify our theoretical results, numerical simulations are performed. Figure 1 and Figure 2 show the analytical calculation and numerical results of case A and B for different M interval. In Figure 1, all the three are cases of $M \approx m$. Solid lines are theoretic results predicted by Eq.(6). Observe these three cases, it's easy to find that whatever the values of p and q, in case of $M \approx m$, the degree distribution of our model displays exponential scaling. Figure 2 is the situation of $M \gg m$. Solid lines are theoretic results predicted by Eq.(9), both of their slope are $\gamma = 3$. Figure 2 shows that under this condition, the degree distribution exhibits power-law form, and the degree exponent is 3.

3.2 Clustering coefficient

By definition, the clustering coefficient of node i is the ratio of the total number e_i of existing edges between all its k_i nearest neighbors and the number of $k_i(k_i-1)/2$ of all possible edges between them. That is $C_i = 2e_i/[k_i(k_i-1)]$. And the clustering coefficient of the whole network is the average of all individual C_i.

In our model there are 6 cases that will lead the change of e_i. Figure 3 shows the mechanisms of these cases.

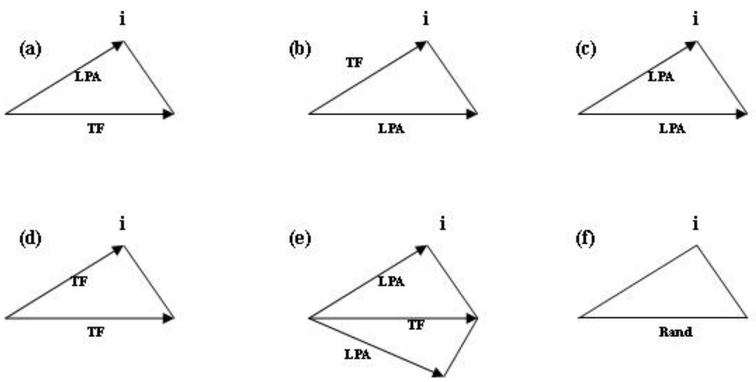


Figure 3: Possible cases of increasing e_i.

(a) node i is connected to the new node in a LPA step, followed by a TF step, one of i's neighbor is also connected to the new node.

(b) one of node i's neighbor is selected to the new node in the LPA step, and then in a TF step the new node connect to node i.

(c) in a LPA step node i is linked to the new node and in another LPA step, one of node i's neighbor is also linked to the new node.

(d) one of node i's neighbor is linked in a LPA step, followed by a TF step, node i is connected to the new node. Also in the same case, one of node i's another neighbor is linked to the new node by a TF step.

(e) node i is connected to the new node in a LPA step, and one of node i's neighbor is connected to the new node in a TF step after another LPA step.

(f) node i is preferentially selected to pick two of its neighbors to link randomly.

Sort this cases together, the rate equation for e_i is

$$\begin{aligned}\frac{\partial e_i}{\partial t} = \; & pm\frac{M}{m_0+pt}\frac{k_i}{\sum\limits_{j\in local}k_j}q+pm\frac{M}{m_0+pt}\sum_{n\in\Omega_i}\frac{k_n}{\sum\limits_{j\in local}k_j}\frac{1}{k_n}q \\ & +pm\frac{M}{m_0+pt}\frac{k_i}{\sum\limits_{j\in local}k_j}(m-1)\frac{M}{m_0+pt}\sum_{n\in\Omega_i}\frac{k_n}{\sum\limits_{j\in local}k_j} \\ & +pm\frac{M}{m_0+pt}\sum_{n^*\in\Omega_i}\frac{k_n^*}{\sum\limits_{j\in local}k_j}\frac{1}{k_n^*}q(m-1)\frac{M}{m_0+pt}\sum_{n'\in\Omega_n}\frac{k_{n'}}{\sum\limits_{j\in local}k_j}q \\ & +pm\frac{M}{m_0+pt}\frac{k_i}{\sum\limits_{j\in local}k_j}(m-1)\frac{M}{m_0+pt}\sum_{n'\in\Omega_n}\frac{k_{n'}}{\sum\limits_{j\in local}k_j}q+(1-p)m\frac{M}{m_0+pt}\frac{k_i}{\sum\limits_{j\in local}k_j},\end{aligned} \tag{10}$$

where node n and n^* are neighbors of node i, node n' is a neighbor of node n. Ω_i represent all neighbors of node i.

Solve this equation, we have

$$e_i = \frac{2pq-p+1}{1+pq}(k_i-e_0)+(1+q+q^2)\frac{(m-1)p^2}{16(1+pq)}k_i^2\frac{(lnN_t)^2}{N_t} \tag{11}$$

Thus the clustering coefficient for nodes with degree k is

$$C(k)=\frac{2e}{k(k-1)}\approx\frac{4pq-2p+2}{(1+pq)k}+(1+q+q^2)\frac{(m-1)p^2}{8(1+pq)}\frac{(lnN_t)^2}{N_t}. \tag{12}$$

From the above equation, we see that the clustering coefficient $C(k)$ can be tuned by parameters p and q. In particular, q is the main factor to have $C(k)$ changed.

In case of $p=1, q=0$, we get the clustering coefficient $C(k)$ of the LC model[13].

$$C(k)=\frac{m-1}{8}\frac{(lnN)^2}{N}. \tag{13}$$

where N is the network size.

4 Conclusions

In this paper, we proposed an extended local-world evolving network model. The degree distribution of this model can change from exponential to power-law with different parameters. Also, the clustering coefficient is tunable with p and q. It reflects some important properties of real networks. For the special cases, when $p=1, q=0$, and $m<M<m_0+t$, our model is the case of LC model proposed by Li and Chen[13]. When $p=1, q=0$, and $M=m$, it is a growing network with uniform attachment of case A in Barabási[8]. When $p=1, q=0$, and $M=m_0+t$, it turns to be the original BA model[8].

Also, it should be pointed out that, local-world networks with link-deleting and node-deleting and so on are also exist in real networks, so our future work should contain all these cases together to give a more realistic model.

References

[1] R. Albert, H. Jeong and A.L. Barabási, Diameter of the World-Wide Web, *Nature*. 401(1999), 130-131.

[2] M. Faloutsos, P. Faloutsos and C. Faloutsos, On power-law relationships of the internet topology, *Comput. Commun. Rev*. 29(1999), 251-262.

[3] A.L. Barabási and R. Albert, Emergence of scaling in random networks, *Science*. 286(1999), 509-512.

[4] F. Liljeros and C.R. Edling, The web of human sexual contacts, *Nature*. 411(2001), 907-908.

[5] P. Erdös and A. Rényi, On random graphs, *Publ. Math. (Debrecen)*. 6(1959), 290-297.

[6] P. Erdös and A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hung. Acad. Sci*. 5(1960), 17-61.

[7] D.J. Watts and S.H. Strogatz, Collective dynamics of 'small-world' networks, *Nature*. 393(1998), 440-442.

[8] A.L. Barabási, R. Albert and H. Jeong, Mean-field theory for scale-free random networks, *Physics A*. 272(1999), 173-187.

[9] R. Albert and A.L. Barabási, Statistical mechanics of complex networks, *Rev. Mod. Phys*. 74(2002), 47-97.

[10] B. Bollobás, O.M. Riordan, J. Spencer and G. Tusnády, The degree sequence of a scale-free random graph process, *Random Structures and Algorithms*. 18(2001), 279-290.

[11] P. Holme and B.J. Kim, Growing scale-free networks with tunable culstering, *Phys. Rev. E*. 65(2002), 026107

[12] M.A. Serrano and M. Boguna, Topology of the world trade web, *Phys. Rev. E*. 68(2003), 015101

[13] X. Li and G.R. Chen, A local-world evolving network model, *Physics A*. 328(2003), 274-286.

[14] B. Wang, Z.Z. Zhang and H.W. Tang, Evolving scale-free network model with tunable clustering, *Modern Physics B*. 19(2005), 3951-3959.

[15] Z.Z. Zhang, L.L. Rong and B. Wang, Local-world evolving network with tunable clustering, *Physics A*. 380(2007), 639-650.