Journal of Southeast University (English Edition) Vol. 25, No. 4, pp. 545 – 548 Dec. 2009 ISSN 1003—7985

Generation of scale-free knowledge network with local world mechanism

Shan Haiyan Wang Wenping

(School of Economics and Management, Southeast University, Nanjing 210096, China)

Abstract: In order to simulate the real growing process, a new type of knowledge network growth mechanism based on local world connectivity is constructed. By the mean-field method, theoretical prediction of the degree distribution of the knowledge network is given, which is verified by Matlab simulations. When the new added node's local world size is very small, the degree distribution of the knowledge network approximately has the property of scale-free. When the new added node's local world size is not very small, the degree distribution transforms from pure power-law to the power-law with an exponential tailing. And the scale-free index increases as the number of new added edges decreases and the tunable parameters increase. Finally, comparisons of some knowledge indices in knowledge networks generated by the local world mechanism and the global mechanism are given. In the long run, compared with the global mechanism, the local world mechanism leads the average knowledge levels to slower growth and brings homogenous phenomena.

Key words: knowledge network; network structure; scale-free; local world mechanism

Network structure is of vital importance in the topological characterization of complex systems in reality[1–2]. It has been observed recently that the distribution of several components in real growing networks, such as the Internet, metabolic network, and social network, has a power-law form[3–5].

Many generating mechanisms have been proposed to explain this phenomenon. Barabási and Albert[3] proposed an evolving network model in order to mimic the growing process of the real complex network. The Barabási-Albert model considers two fundamental mechanisms: growth and preferential attachment(PA), which capture the basic mechanism responsible for the power-law degree distribution. The effect of the PA is that nodes which already have many edges connecting to them will derive even more edges. This indicates the rich get richer scenario.

However the Barabási-Albert model has some limitations. For example, it only predicts a fixed exponent in the power-law degree distribution, and the preferential attachment works on the global network. There are many generalizations and extensions of the PA mechanism to overcome these limitations, usually with additional features such as adding links using nonlinear PA[6], rewiring[7], removal[8], fitness of nodes[9], link weights[10] and local-world linking[11–12].

Recently, the study of the knowledge network has been an attractive issue. Some academic literature has presented that many factors affect knowledge innovation such as the properties of the network structure[13–17].

In this paper, we model the formation of the knowledge network in which the new coming enterprise is not only based on the existing enterprises' connectivity but also on the correlation degree between the coming one and the existing ones in choosing cooperators to innovate together, which is much closer to reality. Moreover, we know the fact that in many real-world networks, the connection of nodes is usually limited due to various kinds of physical constraints, which may have a non-negligible impact on the characteristics of the network[18]. For instance, the global preferential attachment mechanism does not work for the world trade web, an enterprise usually pays its attention to a subset of correlated ones in the network but not on the whole[12]. When a node is added into the knowledge network, the newly added node will not be able or willing to obtain the global information about the network, and what it obtains usually is relevant to itself, i. e. local world mechanism.

Besides the local world mechanism, we slightly modify the original preferential attachment mechanism, i. e, adding the tunable parameters to each node's degree. Generally, the tunable parameter remains unvaried, and usually it is equal to zero in much of the literature[3–4, 8–9, 11–12, 17].

1 Generation Model

Let $S = S\{1, 2, \ldots, N\}$ denote a finite set of knowledge-based enterprises, the scale of which evolves over time. For any $i, j \in S$, we define the binary variable $\chi(i, j) = 1$ if a connection exists between i and j, and $\chi(i, j) = 0$, otherwise. If two enterprises are willing to innovate together, there exists a connection between the two. Each firm is characterized by L distinct types of knowledge. We represent this as a vector of length $L \geqslant 2$, which allows us to consider each firm as located at a point in knowledge space.

Suppose that at time t, node i and node j pool their knowledge together in order to innovate, where $j \in \Gamma_i$, $\Gamma_i = \{j \mid \chi(i, j) = 1, j \in S\}$. Then in the next timestep, each element of node i's knowledge level can be written as

$$v_{il}^{t+1} = \max\left\{v_{il}^t, \ \max_{j \in \Gamma_i}\left\{\frac{c(v_{il}^t)^{1+\alpha_i}(v_{jl}^t)^{\alpha_j}}{v_{il}^t + v_{jl}^t}\right\}\right\} \tag{1}$$

where α_i, α_j denote node i and node j's innovation ability in collaboration, and c is the coefficient of innovation.

Received 2009-03-06.

Biographies: Shan Haiyan(1981—), female, graduate; Wang Wenping(corresponding author), female, doctor, professor, wpwang@seu.edu.cn.

Foundation items: The National Natural Science Foundation of China (No. 70571013, 70973017), Program for New Century Excellent Talents in University (No. NCET-06-0471), Human Social Science Fund Project of Ministry of Education (No. 09YJA630020).

Citation: Shan Haiyan, Wang Wenping. Generation of scale-free knowledge network with local world mechanism[J]. Journal of Southeast University (English Edition), 2009, 25(4): 545 – 548.

The knowledge correlation degree between any two enterprises at time t is defined by

$$\tau_{ij}^{t} = \left(\sum_{l=1}^{L} (v_{il}^{t} - v_{jl}^{t})^2 \right)^{-1/2} \tag{2}$$

where $i, j \in S$; v_{il}^{t} denotes the l-th type knowledge level of node i at time t.

The following is the generation algorithm of the knowledge network:

1) Start with a small number m_0 of nodes and a small number e_0 of edges;

2) According to the knowledge correlation degree between the new coming node and any existing node, select M nodes from the existing network, referred to the local world of the new coming node;

3) At every timestep, we link the new node to m nodes in its local world determined in step 2), using a preferential attachment with probability $\Pi_{\text{local}}(k_i)$ defined by

$$\Pi_{\text{local}}(k_i) = \Pi'(i \in \text{local-world}) \frac{k_i + a}{\sum_{j \in \text{local-world}} (k_j + a)} \tag{3}$$

where $a(\geqslant 0)$ is a constant, and we call it the tunable parameter; local-world refers to all the nodes in which the new coming node has interest in period t.

After t timesteps the model leads to a network with $N = t + m_0$ nodes and $mt + e_0$ edges. The new coming node connects to m nodes, which are selected from its local world according to the knowledge correlation degree, but it does not choose from the whole system as in the Barabási-Albert scale-free model. If $M(t) = m_0 + t$, it means that the local world is the same as the whole network; i. e., the global world is the local world's special case. When $M(t) = m_0 + t$ and $a = 0$, the model simplifies to the original Barabási-Albert model.

2 Network Analysis

2.1 Theoretical prediction of degree distribution

Owing to the random generation of each existing node's initial knowledge level and new coming node's knowledge level, the probability that node i is selected into the new coming node's local world is approximately equal to $M/(m_0 + t)$. Hence Eq. (3) can be written as

$$\Pi_{\text{local}}(k_i) = \frac{M}{m_0 + t} \frac{k_i + a}{\sum_{j \in \text{local-world}} (k_j + a)} \tag{4}$$

We assume that k is continuous[4], and thus the probability $\Pi_{\text{local}}(k_i)$ can be interpreted as a continuous rate of change of k_i. Consequently, from Eq. (4) we can obtain

$$\frac{\partial k_i}{\partial t} = m\Pi_{\text{local}}(k_i) = \frac{mM}{m_0 + t} \frac{k_i + a}{\sum_{j \in \text{local-world}} (k_j + a)} \tag{5}$$

To simplify the following analysis, we assume that

$$\sum_{j \in \text{local-world}} (k_j + a) \approx (\langle k_i \rangle + a) M \tag{6}$$

where the average degree is

$$\langle k_i \rangle = \frac{2mt + 2e_0}{m_0 + t} \tag{7}$$

From formulae (5) to (7), we obtain

$$\frac{\partial(k_i + a)}{\partial t} = \frac{m(k_i + a)}{2mt + 2e_0 + a(m_0 + t)} \approx \frac{m(k_i + a)}{(2m + a)t}$$

The solution of this equation, with the initial condition that node i is added to the system at time t_i with degree $k_i(t_i) = m$, is

$$k_i(t) = (m + a)\left(\frac{t}{t_i}\right)^{m/(2m+a)} - a \tag{8}$$

Using Eq. (8), the probability that a node has a degree $k_i(t) < k$, $P(k_i(t) < k)$, can be written as

$$P(k_i(t) < k) = P\left(t_i > \left(\frac{m + a}{k + a}\right)^{(2m+a)/m} t\right) \tag{9}$$

Assuming that we add nodes at equal time intervals to the system, the probability density of t_i is

$$P(t_i) = \frac{1}{m_0 + t}$$

Then from Eq. (9), we obtain

$$P\left(t_i > \left(\frac{m + a}{k + a}\right)^{(2m+a)/m} t\right) = 1 - \frac{\left(\frac{m + a}{k + a}\right)^{(2m+a)/m} t}{m_0 + t}$$

The probability density of $P(k)$ can be obtained by

$$P(k) = \frac{\partial P(k_i(t) < k)}{\partial k} \sim \left(2 + \frac{a}{m}\right)(m + a)^{2 + a/m}(k + a)^{-(3 + a/m)}$$

i. e., $P(k) \sim A(k + a)^{-\gamma}$, where $A = \left(2 + \frac{a}{m}\right)(m + a)^{2 + a/m}$, $\gamma = 3 + \frac{a}{m}$.

So in the limit of large t, the degree distribution of the knowledge network follows a power-law.

2.2 Numerical simulations of degree distribution

Here we give the numerical simulations of the degree distribution $P(k)$. In all numerical simulations, α_i, L, c are equal to 1, 2, 3, respectively and the initializing network has three nodes no matter whether they are connected or not. At each timestep, a new node is added into the network.

Fig. 1 describes $P(k)$ in the case of the same number of new added edges ($m = 2$) and the same tunable parameter ($a = 1$) but different local world sizes (M_1, M_2), where $M_1(t) = \max\left\{M(t-1), \left[\frac{4(m_0 + t)}{5}\right]\right\}$, $M_2(t) = 3$, [ ] represents integral function.

As shown in Fig. 1, this network evolves into a scale-invariant state, and the probability that a vertex has k edges follows a power law. Although the local world size is not equal, the scaling exponent is nearly unchanged and approaches the theoretical coefficient in both two cases.

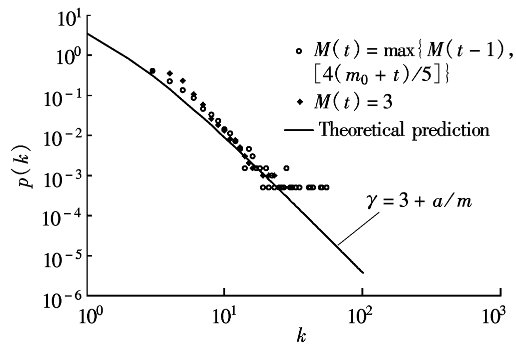

Fig. 1 Degree distribution of local world knowledge networks with different local world sizes M in logarithmic scales

We also find that the number of nodes having more edges is much more when the local world size is larger. This is because the nodes in the network can obtain more opportunities to connect with the new node when the new node's local world size is larger.

3 Parameter Analysis

3.1 Number of new added edges

Fig. 2 compares $P(k)$ in the case of the same local world size ($M=3$) and the same tunable parameter ($a=1$) but different numbers of new added edges (m_1, m_2), where $m_1=1$ and $m_2=2$. We can see that the simulation results almost approach the theoretical prediction and the scaling exponent is dependent on m. Furthermore, the scaling exponent decreases as m increases.

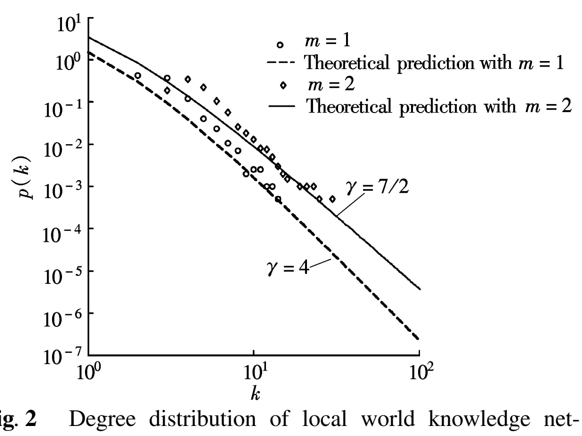

Fig. 2 Degree distribution of local world knowledge networks with different numbers of new added edges m in logarithmic scales

3.2 Tunable parameter

In Fig. 3, we compare $P(k)$ in the case of the same local world size ($M(t) = M_1(t)$) and the same number of new added edges($m=2$) but different tunable parameters a_1 and a_2, where $a_1=1$, $a_2=4$. We can see that the scaling exponent is dependent on a and increases as a increases.

In all, the degree distribution $P(k)$ shown in Figs. 1, 2 and 3 is as a function of the node degree k. From the figures, we can see that the majority of nodes approximately have the same number of connections and the number of

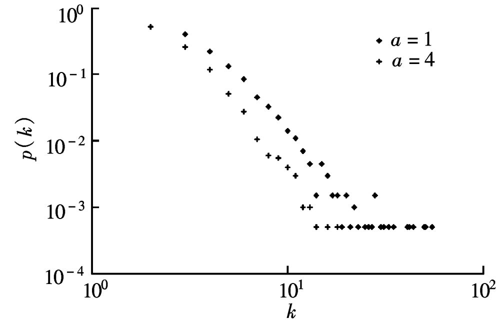

Fig. 3 Degree distribution of local world knowledge networks with different tunable parameters a in logarithmic scales

very highly connected nodes is very small. Sometimes the degree distribution $P(k)$ of this network changes no longer pure scale-free but is truncated by an exponential tail. Note that the theoretical predictions in Figs. 1 and 2 do not strictly have linear properties. The reason is that we suppose that $a=1$, not $a=0$.

4 Comparisons of Local World Mechanism and Global Mechanism

In this section, we compare the average knowledge levels and the variances in knowledge allocation in two connection mechanisms. One is local world and the other is global. We know that the global mechanism is a special case of the local world one, which brings some conveniences in simulation. First we give some definitions as follows.

Agent i's average knowledge level in period t is $\bar{v}_i^t = (\sum_l v_{il}^t)/L$. The average level of knowledge in the network in period t is $\bar{v}(t) = (\sum_{i\in S} \bar{v}_i^t)/N$ and the variance in knowledge allocation is

$$\sigma^2(t) = \frac{\sum_{i\in S}(\bar{v}_i^t - \bar{v}(t))^2}{N} = \frac{\sum_{i\in S}\bar{v}_i^{t2}}{N} - \bar{v}^2(t)$$

Fig. 4 and Fig. 5 show some index comparisons generated by the local world mechanism and the global mechanism. From Fig. 4, we can see that the average knowledge level generated by the local world mechanism is greater than the global mechanism at the very start, and then the trend changes. With time increasing, this gap expands. However, little change in the variances in knowledge allocation is generated by the local world mechanism, which is advantageous to network formation but leads the network to become homogenous.

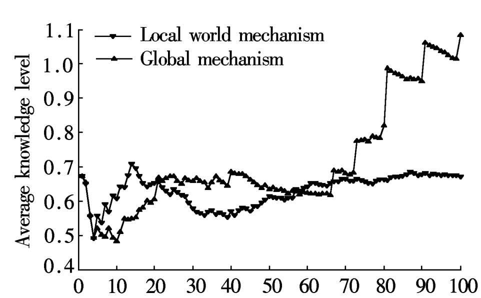

Fig. 4 Average knowledge levels in two mechanisms

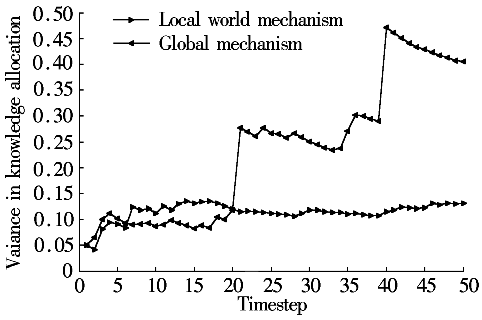

Fig. 5 Variances in knowledge allocation in two mechanisms

5 Conclusion

In this paper, we propose an evolving knowledge network model based on the Barabási-Albert mechanism and the local world mechanism to investigate the effects of the local world size, the tunable parameter and the number of new added edges on the knowledge network structure. It is found that the knowledge network generated by the mechanisms we give has the property of scale-free. When the number of new added edges decreases or the tunable parameter increases, the scaling exponent increases. We also find that the global mechanism increases average knowledge levels and variances of knowledge allocation in the knowledge network compared with the local world mechanism. High variances of knowledge allocation in the network can cause heterogeneity and lead the network to collapse because of inequity. In reality, a majority of networks have such properties, for example, enterprises involved in the world trade web.

References

[1] Albert R, Barabási A L. Statistical mechanics of complex networks[J]. *Reviews of Modern Physics*, 2002, **74**(1): 47 − 97.

[2] Dorogovtsev S N, Mendes J F. Evolution of networks[J]. *Advances in Physics*, 2002, **51**(4): 1079 − 1187.

[3] Barabási A L, Albert R. Emergence of scaling in random networks[J]. *Science*, 1999, **286**(15): 509 − 512.

[4] Barabási A L, Albert R, Jeong H. Mean-field theory for scale-free random networks[J]. *Physica A*, 1999, **272**(1/2): 173 − 187.

[5] Chakrabarti D, Faloutsos C. Graph mining: laws, generators and algorithms[J]. *ACM Computing Surveys*, 2006, **38**(1): 1 − 70.

[6] Mukherjee G, Manna S S. Quasistatic scale-free networks [J]. *Physical Review E*, 2003, **67**(1): 012101.

[7] Ree S. Generation of scale-free networks using a simple preferential-rewiring dynamics[J]. *Physica A*, 2007, **376**(1/2): 692 − 698.

[8] Deng K, Zhao H, Li D. Effect of node deleting on network structure[J]. *Physica A*, 2007, **379**(2): 714 − 726.

[9] Li M, Wu J, Wang D, et al. Evolving model of weighted networks inspired by scientific collaboration networks[J]. *Physica A*, 2007, **375**(1): 355 − 364.

[10] Bedogne C, Rodgers G J. Complex growing networks with intrinsic vertex fitness[J]. *Physical Review E*, 2006, **74**(4): 046115.

[11] Li X, Chen G. A local-world evolving network model[J]. *Physica A*, 2003, **328**(1/2), 274 − 286.

[12] Gu Y, Sun J. A local-world node deleting evolving network model[J]. *Physics Letters A*, 2008, **372**(25): 4564 − 4568.

[13] Watts D. Strogatz S. Collective dynamics of “small-world” networks[J]. *Nature*, 1998, **393**(6684): 440 − 442.

[14] Watts D. Networks, dynamics and the small-world phenomenon[J]. *American Journal of Sociology*, 1999, **105**(2): 493 − 527.

[15] Cowan R, Jonard N. Network structure and the diffusion of knowledge[J]. *Journal of Economic Dynamics and Control*, 2004, **28**(8): 1557 − 1575.

[16] Cowan R, Jonard N, Ozman M. Knowledge dynamics in a network industry[J]. *Technological Forecasting and Social Change*, 2004, **71**(5): 469 − 484.

[17] Tang F, Xi Y, Ma J. Estimating the effect of organizational structure on knowledge transfer: a neural network approach [J]. *Expert Systems with Applications*, 2006, **30**(4): 796 − 800.

[18] Yin C, Wang B, Wang W, et al. Efficient routing on scale-free networks based on local information[J]. *Physics Letters A*, 2006, **351**(3): 220 − 224.

局域世界机制下无标度知识网络的生成模型

单海燕　　王文平

(东南大学经济管理学院, 南京 210096)

摘要: 为了更真实地模拟现实知识网络的成长过程, 构造了一类基于局域连接机制下的知识网络生成模型. 利用统计物理学中的平均场方法, 给出了知识网络度分布的理论预测, 并运用 Matlab 仿真进行了验证: 当局域世界规模很小时, 网络度的分布函数近似服从无标度分布, 当局域世界规模不是很小时, 网络度的分布会从纯粹的无标度状态变化成尾状物服从指数分布的近似无标度状态, 且无标度指数随着可调参数增加而增加, 随着新增边数的增加而减少. 最后,比较了在局域连接机制和全局连接机制下生成的知识网络的一些知识指标. 从长期来看, 与全局机制相比, 局域机制会导致网络平均知识水平增长缓慢, 而且网络同质化现象严重.

关键词: 知识网络; 网络结构; 无标度; 局域世界机制

中图分类号: C93-03