Research On Evolution Model Based on Hierarchical Community Structure in Complex Networks

Prof. Yuhua Liu

Department of Computer Science,
Central China Normal University, Wuhan
430079,Hubei, China

HuaZhong Normal University

Author:

Yuhua Liu ,Jianzhi Jin ,Fang Hu, Zhenzhen Wang

Email:

yhliu@mail.ccnu.edu.cn

Homepage:

http://www.ccnulyh.com

Research On Evolution Model Based on Hierarchical Community Structure in Complex Networks

- Introduction
- Thoughts of Hierarchical Community Structure
- Construction of Evolution Model Based on Hierarchical Community Structure
- Construction of Dynamic Hierarchical Community Structure Evolution Model Based on Cellular Automata
- Conclusion
- Reference

Introduction

- Modelling of complex network experiences the stage of rule network, random network, small world network and the scale-free.
- For the modelling of complex networks it should have the following characteristics: global features and local characteristics.
- The global features has the power-law characteristics, isomerism and heterogeneity, dynamic growth and regression of nodes.
- the local characteristics has such as local priority connection and community structure.

Thoughts of Hierarchical Community Structure(1/6)

Through researching on metabolic network, researchers discover :

- The evolution process of this network is not based on the growth of nodes or edges, but through the iterative manner of modules.
- Then the network has obvious characteristics of hierarchical structure.

Thoughts of Hierarchical Community Structure(2/6)

In order to confirm that hierarchical structure in networks is exsiting,

- Barabasi summarizes the hierarchical structures in different networks, and collects a lot of data from different real networks .
- Finally Barabasi proposes the model of hierarchical structure in networks.

Thoughts of Hierarchical Community Structure(3/6)

- Traditional community structures cannot completely reflect the internal structure of complex networks
- In reality, in complex networks inherent hierarchical structure is existing.
- There is no confliction between hierarchy and community structure.
- In a sense the evolution of hierarchical structure generates the community structure.

Thoughts of Hierarchical Community Structure(4/6)

- It is found that a community in complex networks also contains some other communities.
- The community is made up of a series of smaller communities.
- There is a containment relationship between different hierarchical communities.
- We also call such associations having a hierarchical structure.

Thoughts of Hierarchical Community Structure(5/6)

- In general, information about the community structure is very less.
- It is impossible to know in advance how many communities can be divided into.
- The traditional graph partitioning methods often need to define the number of subgraph, and then proceed to the partitioning.
- In this case, we have to conjecture in advance the number of communities in the whole network. It is neither scientific nor reasonable.

Thoughts of Hierarchical Community Structure(6/6)

- There may be a large number of hierarchical structures in complex networks.
- So the partitioning methods of hierarchical community should be taken into account now.
- Our paper will explore the evolution model of hierarchical community structure.

Constructon of Evolution Model Based on Hierarchical Community Structure

Combined with the above characteristics and based on BA scale-free network model, a new complex network model has been established from the view of hierarchical community structure.

Specific Steps of Evolution Model Based on Hierarchical Community Structure(1/6)

The following is the specific steps of the model.

Step 1. First given an initial network which having m_0 backbone nodes and n_0 edges.

Step 2. In each time step the model random executes an operation as follows.

Specific Steps of Evolution Model Based on Hierarchical Community Structure(2/6)

(1) Generate a backbone node with probability *p*, and connect to the backbone node *i* which already exists in accordance with following probability Π_i .

$$\Pi_i = \frac{k_i}{\sum_{l \in \Omega} (k_l)} \qquad (1)$$

Where Ω is the local network composed by backbone nodes.

Specific Steps of Evolution Model Based on Hierarchical Community Structure(3/6)

(2)

- Generate an initial community network composed by m_1 common nodes and n_1 edges with probability q .
- The community network can be grown automatic dynamically, and has preferential attachment characteristics.
- Establish m_2 edges between the community network and the already exists backbone node i in accordance with the formula (1).

Specific Steps of Evolution Model Based on Hierarchical Community Structure(4/6)

(3) Establish m_2 edges in one community network and m_3 edges between random two community networks randomly with probability s, one end of the edges is randomly selected, and the other end is selected according to the formula (2).

$$\Pi_j = \frac{k_j}{\sum_l k_l} \qquad (2)$$

Where the denominator represents the sum of all common nodes' degree.

Specific Steps of Evolution Model Based on Hierarchical Community Structure(5/6)

(4) Remove m_4 edges in one community network and m_5 edges between random two community networks randomly with probability (1-s),one end of the edges is randomly selected, and the other end is selected according to the formula (3).

$$\Pi_j' = 1 - \frac{k_j}{\sum_l k_l} \qquad (3)$$

Where the denominator represents the sum of all common nodes' degree.

Specific Steps of Evolution Model Based on Hierarchical Community Structure(6/6)

Step 3. Repeat step 2 untill proper scale network is established.

The above parameters satisfy

$$0 \le p < q \le 1, 0.5 \le s \le 1, \; p+q=1 .$$

- The growth rate and the preferential attachment characteristics in (2) is randomly .
- So at time t the number of the nodes and edges in the community networks is different.

The Analysis of Theoretical Degree Distribution (1/7)

- Here we will use the mean field theory to obtain the degree distribution of the node *i* at time *t*.
- We assume that at time *t* the network has *N* nodes and *E* edges without considering the isolated nodes, then

$$N = m_0 + pt + qt(m_1 + t)$$

where in the backbone nodes are $N_B = m_0 + pt$, and

$$E = n_0 + pt + qt(n_1 + mt + m_1) + st(m_2 + m_3 + m_4 + m_5) - t(m_4 + m_5) .$$

The Analysis of Theoretical Degree Distribution (2/7)

a. Generate a backbone node with probability p , then

$$\left(\frac{\partial k_i}{\partial t}\right)_{(i)} = p(m_0 + pt) \cdot \frac{k_i}{\sum_{l \in \Omega}(k_l)}$$

The right of the equal sign on behalf of existing the selection rules of the already exists backbone node i.

b. Generate an initial community network with probability q , then

$$\left(\frac{\partial k_i}{\partial t}\right)_{(ii)} = qm_2(m_0 + pt) \cdot \frac{k_i}{\sum_{l \in \Omega}(k_l)}$$

The above equation describes a new community network associate with the network only through the backbone nodes.

The Analysis of Theoretical Degree Distribution(3/7)

c. Establish m_2 edges in one community network and m_3 edges between random two community networks randomly with probability s, then

$$\left(\frac{\partial k_i}{\partial t}\right)_{(iii)} = \frac{sm_2}{m_0 + pt} \cdot [\frac{1}{N_\Omega(t)} + (1 - \frac{1}{N_\Omega(t)}) \cdot \frac{k_j}{\sum_l k_l}] + sm_3[\frac{2}{(m_0 + pt)} - \frac{1}{(m_0 + pt)^2}] \cdot \frac{k_j}{\sum_l k_l}$$

The right of equal sign in the first paragraph describes the selection of the edge in one community network; the second paragraph describes the selection of the edge between random two community networks. One end of the edge is randomly selected; the other end is selected by the priority probability.

The Analysis of Theoretical Degree Distribution(4/7)

d. Remove m_4 edges in one community network and m_5 edges between random two community networks randomly with probability $(1-s)$, then

$$\left(\frac{\partial k_i}{\partial t}\right)_{(iv)} = \frac{(1-s)(m_4+m_5)}{m_0+pt}\left[\frac{1}{N_\Omega(t)} + \left(1-\frac{1}{N_\Omega(t)}\right)\cdot\frac{1}{N_\Omega(t)-1}\left(1-\frac{k_j}{\sum_l k_l}\right)\right]$$

$N_\Omega(t)$ represents the mean value of the ordinary nodes under the backbone node Ω at time t,

then $N_\Omega(t) = \dfrac{m_0+pt+qt(m_1+t)}{m_0+pt}$.

When t is large,

The Analysis of Theoretical Degree Distribution(5/7)

When t is large,

$$\frac{\partial k_i}{\partial t}=\frac{(m_0-n_0)(qm_2+p)k_i}{2pt}+\frac{sm_2}{m_0+pt}\cdot\frac{p}{p+qt}+\frac{sm_2}{m_0+pt}\cdot\frac{p}{p+qt}\cdot\frac{qt}{p}\cdot\frac{k_j}{\sum_l k_l}+\frac{2sm_3}{(m_0+pt)}\cdot\frac{k_j}{\sum_l k_l}$$

$$-\frac{sm_3}{(m_0+pt)^2}\cdot\frac{k_j}{\sum_l k_l}+\frac{(1-s)(m_4+m_5)}{m_0+pt}\cdot\frac{2p}{p+qt}-\frac{(1-s)(m_4+m_5)}{m_0+pt}\cdot\frac{p}{p+qt}\cdot\frac{k_j}{\sum_l k_l}$$

$$=[\frac{sm_2q}{p(p+qt)}-\frac{(1-s)(m_4+m_5)}{t(p+qt)}]\cdot\frac{pk_i}{2p+q(n_1+mt+m_1)+s(m_2+m_3+m_4+m_5)-(m_4+m_5)}$$

$$+\frac{(m_0-n_0)(qm_2+p)k_i}{2pt}+\frac{sm_2+2(1-s)(m_4+m_5)}{t(p+qt)}$$

Make $\frac{sm_2q}{p}=(1-s)(m_4+m_5)$, $a=\frac{[(m_0-n_0)(qm_2+p)-2sm_3]}{2p}$ and $b=\frac{sm_2+2(1-s)(m_4+m_5)}{q}$.

then $\frac{\partial k_i}{\partial t}=\frac{1}{t^2}\cdot b+\frac{k_i}{t}\cdot a$.

The Analysis of Theoretical Degree Distribution(6/7)

Solving the above equation, and when $a \neq -1$,

$$k_i(t) = \frac{t^{a+1}(b/t + m + am)}{x^{a+1}(a+1)} - \frac{bt^a}{(a+1)t^{a+1}}$$

Because $P(t_i) = \frac{1}{pt + qt^2}$, so

$$P[k_i(t) < k] = P[t_i > [\frac{k + b/t(a+1)}{m + b/t(a+1)}]^{-\frac{1}{a+1}} t] = 1 - \frac{1}{pt + qt^2} \cdot (\frac{k + b/t(a+1)}{m + b/t(a+1)})^{-\frac{1}{a+1}} t$$

The Analysis of Theoretical Degree Distribution(7/7)

Then we can get the degree distribution is

$$P(k)=\frac{\partial\{P[k_i(t)<k]\}}{\partial k}=\frac{1}{(a+1)(pt+qt^2)}\cdot[m+b/t(a+1)]^{\frac{1}{a+1}}\cdot[k+b/t(a+1)]^{-r}$$

where $r=\frac{1}{a+1}+1$.

The above indicates $P(k)\propto(k+b/t(a+1))^{-r}$, it shows that the network has the characteristics of power law, but it is different from BA model, the $b/t(a+1)$ illustrates the generation rule of the local network has very important influence for the global network.

Simulation and Estimation(1/9)

We compared 200 networks which were randomly generated in accordance with the following parameter values, where m_0=5, n_0=4, m_1=3, n_1=2, m_2=20, m_3=4, m_4=4, m_5=6, p=0.8,q=0.6,s=0.66.

Simulation and Results(2/9)

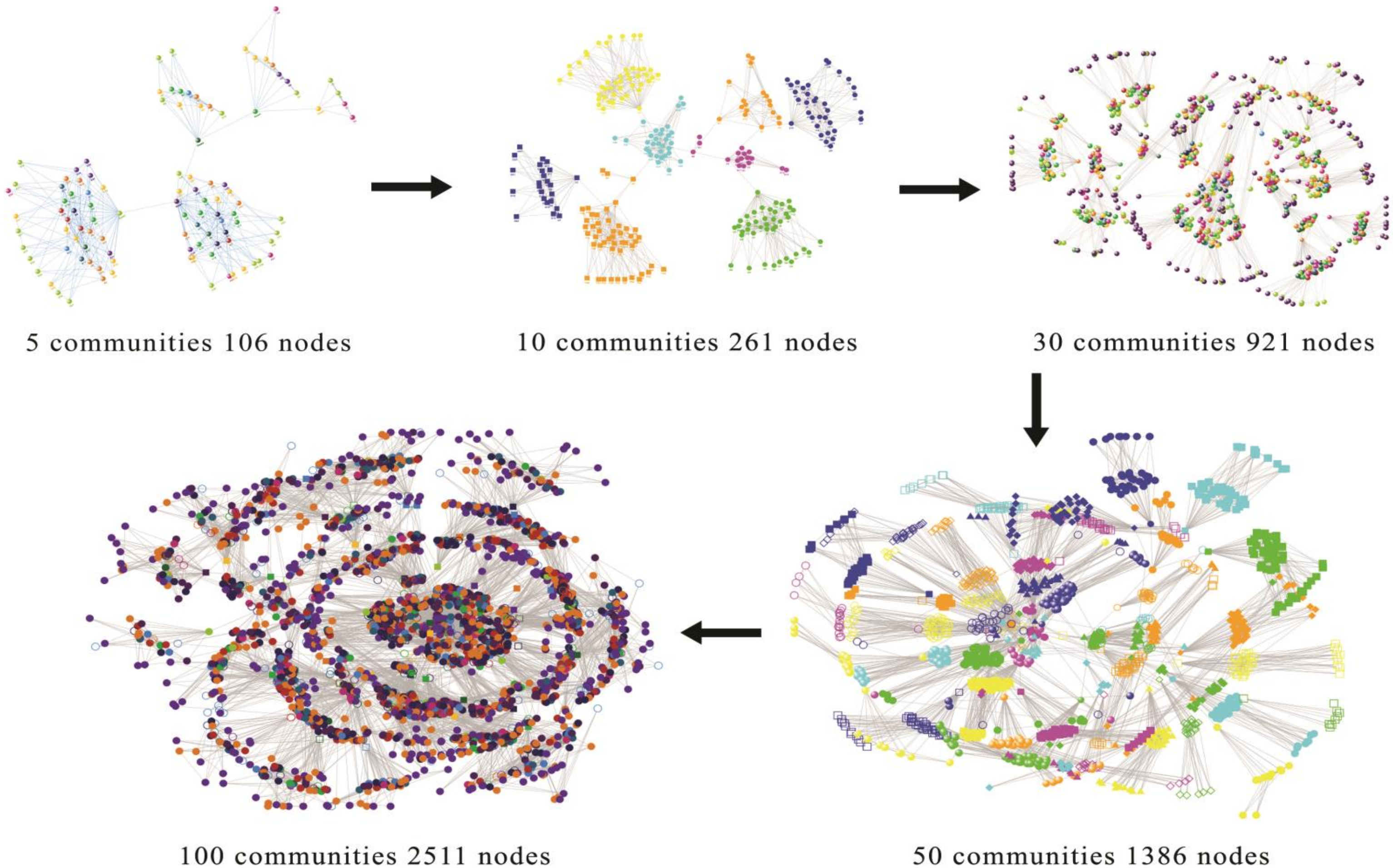


Figure. 1. The evolutionary process of the model.

Simulation and Results(3/9)

- Figure 1 is the evolution process of the model which community number increase gradually.
- with the increase of the community number, the network evolution became more complex.
- but the community structure is always visible easily.

Simulation and Results (4/9)

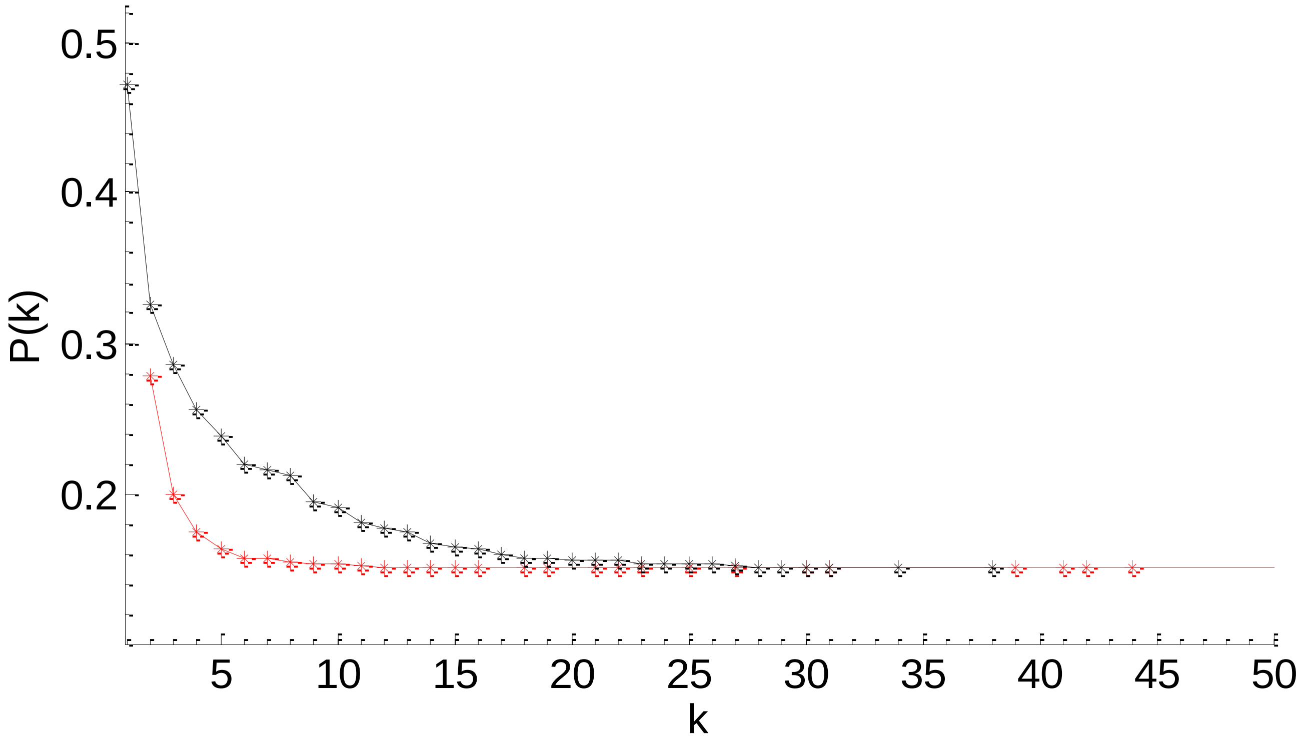
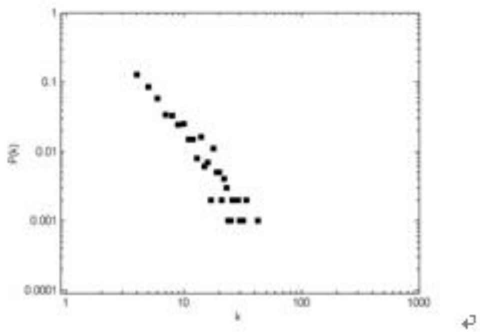


Figure. 2. The contrast of P (k) between the present model and the BA model, black curve is the present model, the red curve is the BA model.

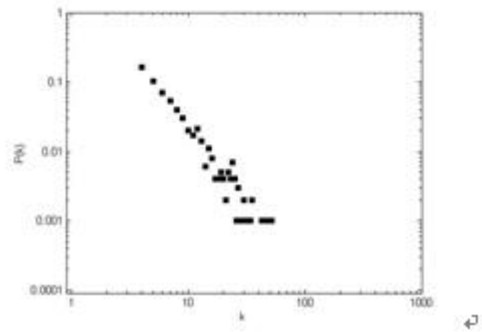
Simulation and Results (5/9)

- black curve is the present model, the red curve is the BA model.
- we can see that the degree distribution is follows the power law distribution, its distribution relative to the original BA model is more clear and smooth.
- the present model not only considers the node itself but also uses the local network as the basis of the network evaluation.
- which means that the generation rules of the local community has a very important impact on the global network.

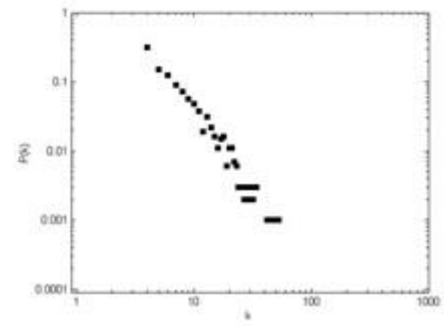
Simulation Experiments(6/9)



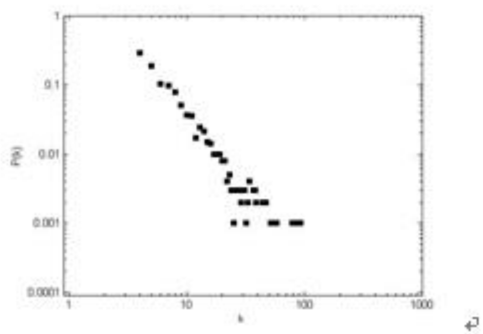
a.10 communities and 515 edges



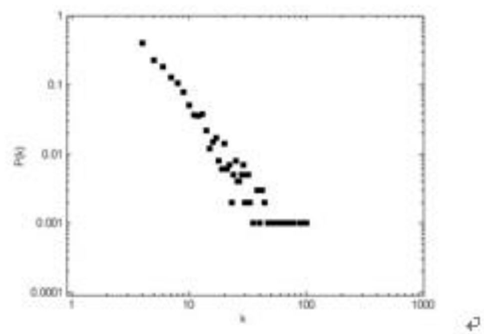
b.15 communities and 639 edges



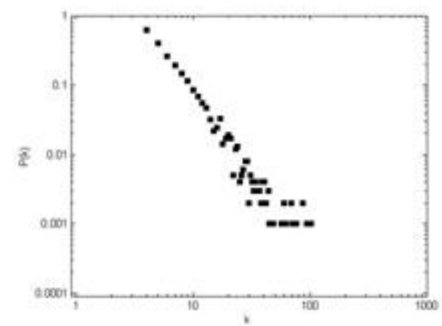
c.20 communities and 1130 edges



d.10 communities and 1082 edges



e.15 communities and 1489 edges



f.20 communities and 2332 edges

Figure. 3. The degree distribution of the generated network the different growth level of the network scale and community.

Simulation and Results (7/9)

- When the number of communities changed or the scale of the network changed, the degree distributions are all power-law.
- In the case of the same number of communities, the scale of the network in *d*, *e* and *f* is 2 times of the scale of the network in *a*, *b* and *c*.
- The larger the network scale is, the more obvious the power-law distribution trend is.

Simulation and Results(8/9)

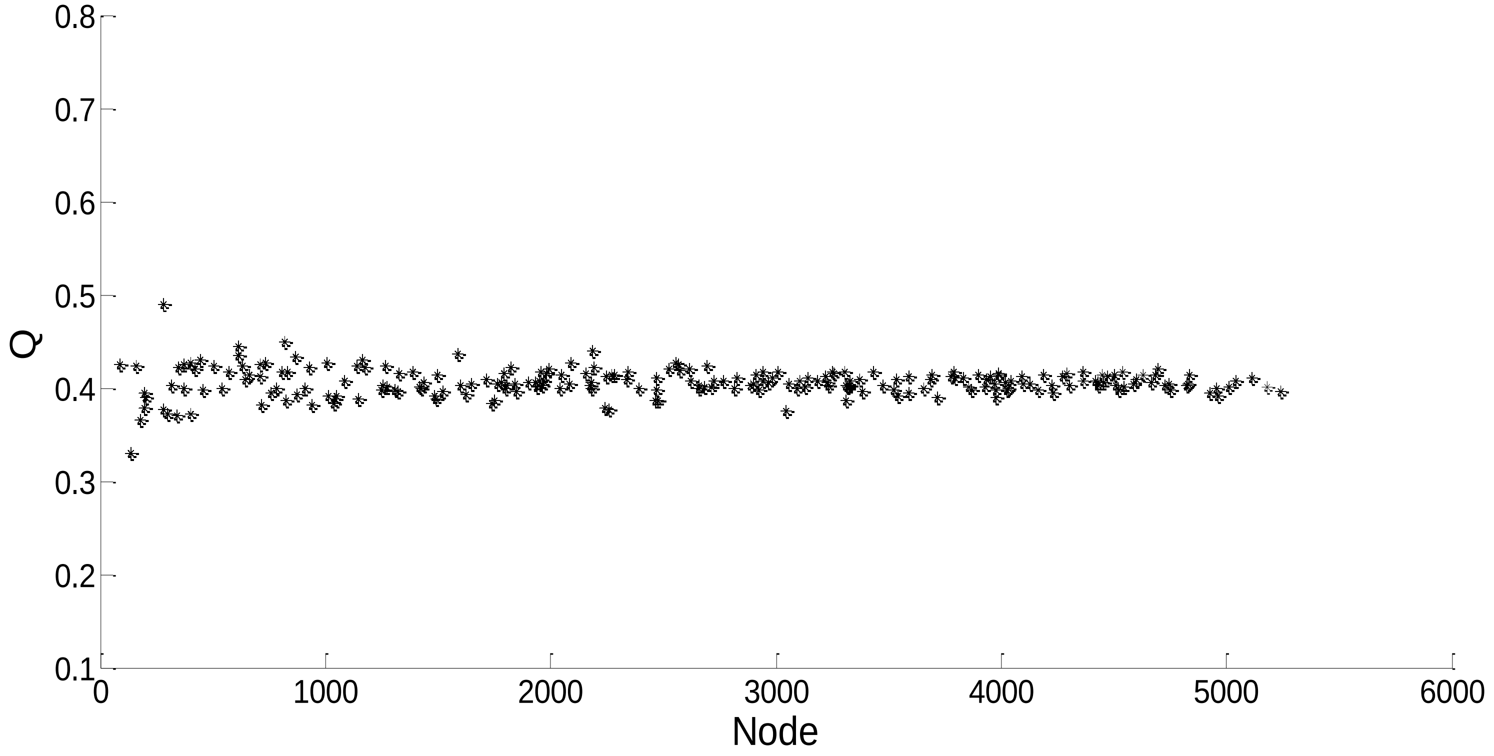


Figure. 4. The modularity scores of 200 networks randomly generated by the model.

Simulation and Results(9/9)

- Figure 4 is the modularity scores of 200 networks randomly generated by the model, where the nodes in the networks are in the range of [20, 5500].
- When its modularity score is between 0.3 and 0.6, we can consider the network has a significant community structure.
- Modularity scores stabilized at about 0.4, the community structures in these networks are obvious.

Construction of Dynamic Hierarchical Community Structure Evolution Model Based on Cellular Automata(1/4)

- A complex network model described above has been established from the view of hierarchical community structures.
- Furthermore, a dynamic evolution network model was established through evolution rules of cellular automata in our paper.

Construction of Dynamic Hierarchical Community Structure Evolution Model Based on Cellular Automata(2/4)

Theory of Cellular Automata

- A cellular automaton is a dynamic system that is discrete in time dimension change,
- namely, time is an integer with continuity straits and equal to the interval. If time interval *dt* = 1 and *t* = 0 at the initial time, then *t* = 1 for the succeeding moment.
- In the above conversion function, a cell's state at time *t+1* (directly) depends only on its states and its neighbors' state at time t.
- However, the states of the cell and its neighbors at time *t*-1 indirectly affect the state of the cell at time *t+1*.

Dynamic Hierarchical Community Structure Evolution Model Based on Cellular Automata(3/4)

Theory of Cellular Automata

- After the scale-free evolution, the adjacency matrix of the network is obtained. The adjacency matrix is modified as follows:
- the value of the edge whose one end is the backbone node is modified into 2. The cell is the data of the matrix, and the set of states includes {0, 1, 2} as indicated in the definition of cellular automata.

Dynamic Hierarchical Community Structure Evolution Model Based on Cellular Automata(4/4)

Theory of Cellular Automata

- The cell has three states: **0** means no edge exists;

1 corresponds to an existing edge whose ends are ordinary nodes;

2 corresponds to an existing edge, at least one end of which is the backbone node.

- A spatial network set in which cells are distributed is called cellular space. Cellular space is two dimensional, that is, the adjacency matrix of network nodes. Moore's selection method was employed in this study to select the cellular neighbors.

Specific Evolution Rules(1/3)

The specific evolution rules are as follows:

- **Rule 1** As the number of connections or attraction factors increases, ordinary nodes can become backbone nodes and would thus require rules:
- Ordinary nodes can transform into backbone nodes, but the probability is generally low. The condition is that the state of the node is 1, the state of three or more neighbors is 2, and the next state of the node is 2;

Specific Evolution Rules(2/3)

- **Rule 2** The no-edge nodes may also be connected at a specific moment and would thus require rules:
- a connection may exist between ordinary nodes. The condition is that the state of the node is 0, the state of three or more neighbors is 1, and the next state of the node is 1;
- **Rule 3** Ordinary nodes may lose their connection because they do not update for a long time or their activity decreases and may thus require rules:
- a connection may exist between common nodes. The condition is that the state of the node is 1, the state of five or more neighbors is 0, and the next state of the node is 0;

Specific Evolution Rules(3/3)

- **Rule 4** For the same reason, backbone nodes can be reduced or even lose their connection although the probability is low. Thus, rules are needed:
- Backbone nodes can transform into ordinary nodes, but the probability is relatively low. The condition is that the state of the node is 2, the state of five or more neighbors is 0, and the next state of the node is 1.
- **Rule 5** Based on the adjacency matrix border, if a node is a left boundary node, the right boundary nodes of its corresponding row and column will become its left neighbor nodes. This rule also applies to the top and bottom nodes.

Simulation and Results(1/9)

In order to verify that the generated network conforms to the key characteristics of social network, the evolution process, degree distribution, power-law behavior, average path length of this generated network are simulated and analyzed.

Simulation and Results(2/9)

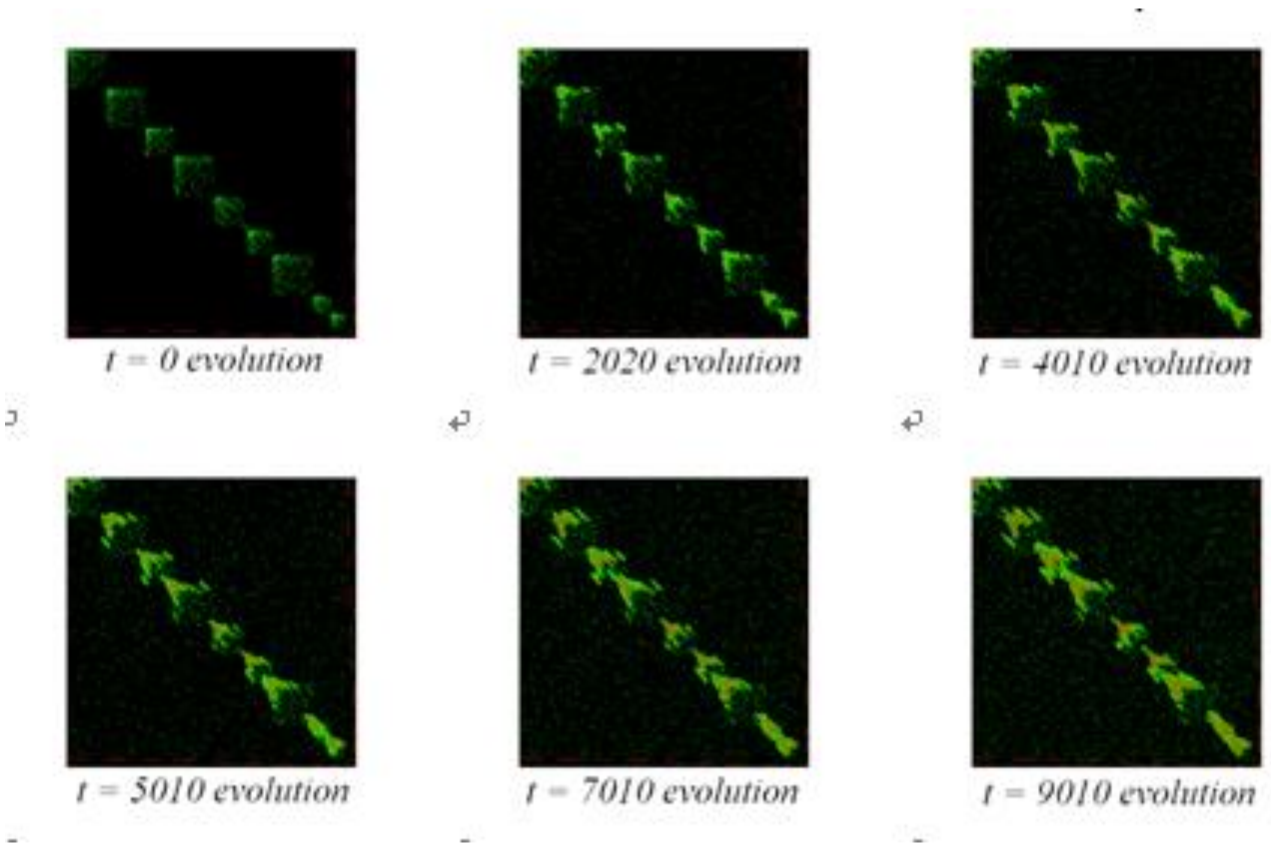


Figure. 5. Evolution of the network adjacency matrix consisting of 10 hierarchical communities.

Simulation and Results(3/9)

- Fig. 5 shows the adjacency matrix of a complex network consisting of 10 hierarchical communities under different evolutionary time *t*.
- The boundary of the communities blurs gradually, and small communities are derived aside from the large communities.
- Some randomly scattered connections emerge. However, these connections do not form a new community.

Simulation and Results(4/9)

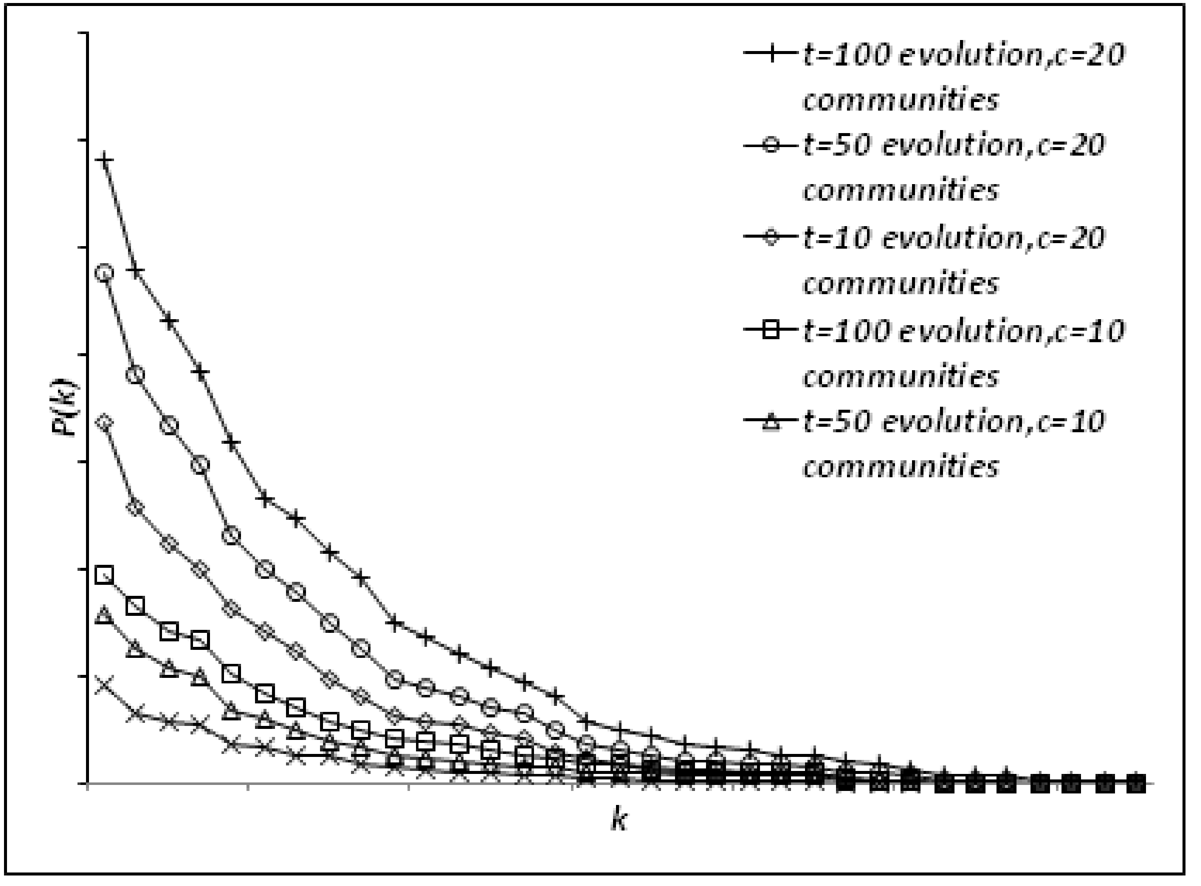


Figure. 6. Degree distribution of a complex network under different *c* and *t*.

Simulation and Results(5/9)

- The degree distribution of the above complex network under different *c* and *t*, where *t* represents the time number of evolution and *c* represents the number of communities in the network.
- According to the figure, the node degree distribution meets the power-law distribution.
- However, after a period of evolution, the curve trend slows and nodes with high connections gradually increase.

Simulation and Results(6/9)

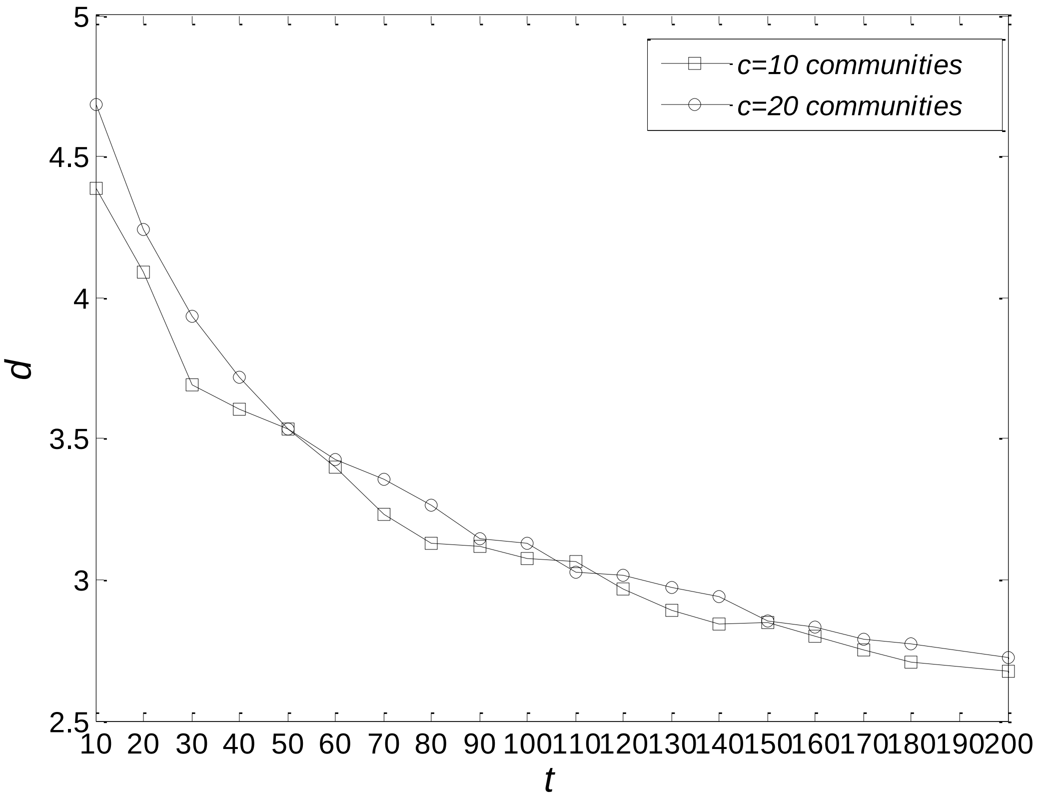


Figure. 7. Average path length variation trend of the network when *c* = 10 communities and 20 communities.

Simulation and Results(7/9)

- The average path length variation trend of the network when c = 10 and 20.
- The evolutionary time is long, and the average path length of the network is short. The network with more communities has a relatively larger average path length.
- After several tests, the average path length of the network stabilizes at approximately 1.9 when the number of evolutionary times is greater than 1000,
- It indicates that the network has a stable average path length at this time.

Simulation and Results(8/9)

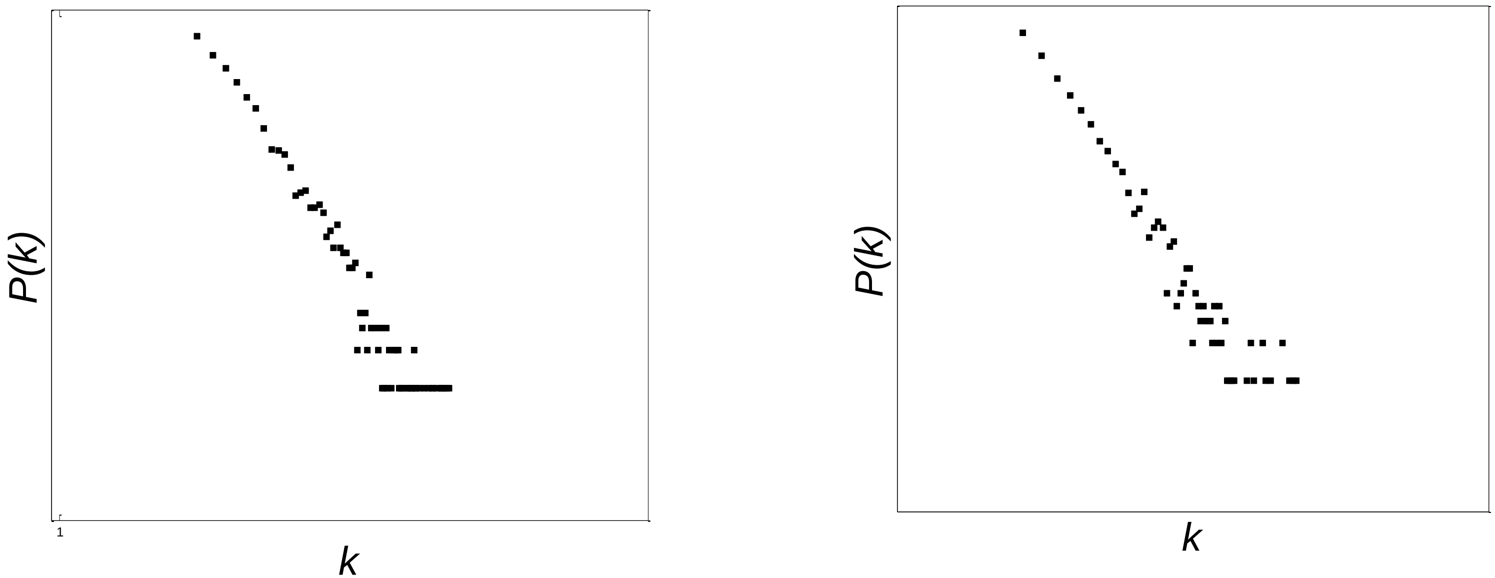


Figure. 8. Degree distribution comparison of our model on the left and Zheng's model on the right.

Simulation and Results(9/9)

The Degree distribution comparison of our model and Zheng's model, we can see that the distribution of the two models are independent of the network size and time asymptotic distribution.

Conclusion

- A evolution model of complex network is established from the view of hierarchical community structure.
- A dynamic evolution network model was also established through the establishment of a new evolution rule of cellular automata.
- It is proved that the first model has a power-law characteristic from the theoretical degree distribution and has a steady modularity score.
- It is proved that the second model has a power-law characteristic and has a clear, controlled community structure.
- These two models expand the BA model,are more consistent with the actual abstract and closer to the evolution of actual network

REFERENCE（1/4）

[1] Barabási, A-L., Albert, R., Emergence of scaling in random networks, Science, 1999, 286: 509-512.

[2] Albert, R, Barabási, A-L,. Statistical mechanics of complex networks, Rev. Mod. Phys., 2002, 74 (1): 47-51.

[3] Krapivsky, P. L., Redner, S., Leyvraz, F., Connectivity of growing random networks,Phys.Rev. Lett., 2000, 85: 4629-4632.

[4] Dorogovtsev, S. N., Mendes, J. F. F., Samukhin, A. N., Structure of growing networks with preferential linking, Phys. Rev. Lett., 2000, 85(21): 4633-4636.

[5] Dorogovtsev, S. N., Mendes, J. F. F., Effect of the accelerating growth of communications networks on their structure, Phys. Rev. E, 2001, 63(2): 025101(R).

[6] Bianconi, G., Barabási, A.-L., Competition and multiscaling in evolving networks, Europhysics Letters, 2001, 54(4): 436-442.

REFERENCE（2/4）

[7] Liu, Z. H., Lai, Y. C., Ye, N., et al., Connectivity distribution and attack tolerance of general networks with both preferential and random attachments, Physics Letters A, 2002, 303(5-6): 337-344.

[8] Li, X., Chen, G. , A local-world evolving network model, Physica A, 2003, 328: 274-286.

[8] Chen, G., Fan, Z. P., Li, X., Modeling the complex Internet topology, Complex Dynamics in Communication Networks, G. Vattay, L. Kocarev(eds). Berlin: Springer-Verlag, 2005, 213-235.

[10] Dorogovtsev, S. N., Mendes, J. F. F., Samukhin, A. N., Structure of growing networks with preferential linking, Phys. Rev. Lett., 2000, 85(21): 4633-4636.

REFERENCE（3/4）

[11] Jianzhi, Jin., Yuhua, Liu., Laurence, T. Yang., Naixue, Xiong., Fang, Hu., An Efficient Detecting Communities Algorithm with Self-Adapted Fuzzy C-Means Clustering in Complex Networks, 2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications, 2012, 1988-1993.

[12] Karsten, S., Nitesh, V. C., Identifying and evaluating community structure in complex networks, Pattern Recognition Letters, 2010, 31(5): 413-421.

[13] Ball, B., Karrer, B., Newman, M. E. J., Efficient and principled method for detecting communities in networks, Physical Review E, 2011, 84(3): 036103.

[14] Dorogovtsev. S N, Mendes. J F F, Povolotsky. A M, et al. "Organization of complex networks without multiple connections," Physical review letters, 95.19, pp. 195701, 2005.

REFERENCE（4/4）

[15] Peköz. E A, Röllin. A, Ross. N, “Degree asymptotics with rates for preferential attachment random graphs,” The Annals of Applied Probability, 23.3, pp. 1188-1218, 2013.

[16] Sen. Q, Guan-Zhong. D, “A new local-world evolving network model,” Chinese Physics B, 18.2, pp. 383, 2009.

[17] Meirong. Zheng, Yuhua. Liu, Kaihua. Xu, Wenwen. Dai, “Research on Coupling Coefficient of the Scale-free Network Model,” computer science in chinese, 39(5), 2012.

[18] Kayama. Y, “Complex networks derived from cellular automata,” arXiv preprint arXiv:1009.4509, 2010.

Thanks !

Date: 19/7/2014