Complex Network Flexibility Discussion

Han Jian   Feng Ye
Academy of Armored Forces Engineering
The 21st Mail Box, Dujiakan, Beijing 100072, China
Email: hjerpao@hotmail.com

Abstract

The complex network is adaptive network, which has three commonly used statistical attributes: degree distribution, clustering coefficient and average path length. In this article, looking back upon the development process of complex network, we introduce the current damage modes of complex network, random failure and intentional attack. The flexibility concept of network is stated and evaluation criterion is also established. Flexibility comparative experiments are carried out between BA scale-free network and ER network. The difference of these two networks in flexibility is supported by experiments, which comes to some related conclusions.

Keywords: complex network; flexibility; small-world network; scale-free network

Introduction

The famous physicist, Stephen Hawking believes that the 21st century is the era of complexity science. As the development of technologies, all kinds of complex system structures are emerging. The complexity of network is attributed to the complexity of interaction. These complex networks include transportation networks, telecommunication networks, Internet, World Wide Web, Film actor networks, scientific cooperation networks, and also biological system networks such as neural networks, genetic networks, metabolic network and the protein network etc.
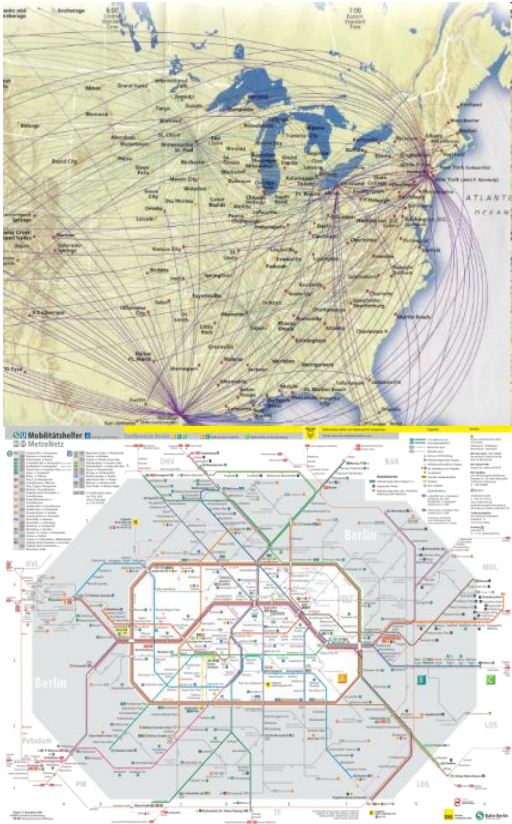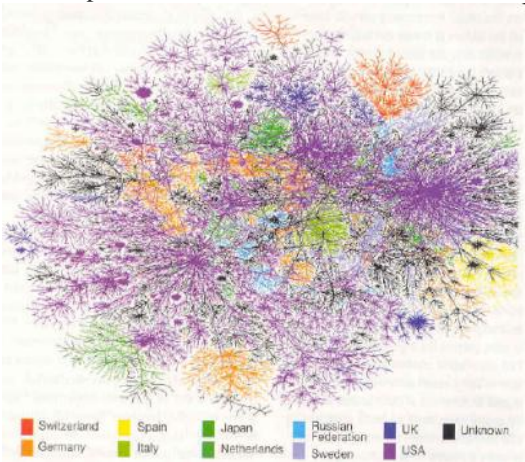

Fig.1 three kind of complex network instances

What is exact definition of complex networks? Complex network is not a graph composed of nodes and their relationships; however, it's the collection of graph, in other words, a random process of diagram. In this perspective, complex network is an extension of chart. In reality, because of high-speed development and changes occurring in the social environment all the time, permanent structure doesn't exist in this world. Every node and the relationship between each other in system keep changing constantly. These characteristics mentioned above constitute a complex network. That is, complex network is actually an effective tool studying the actual complex systems.

In recent years, the study of complex networks is at the stage of vigorous development. Accompany computing ability is enhanced day by day; research on big data of actual networks can be carried out. Scientists began to deal with actual complex network through high-performance computers. Regular static network, which is the core of complex network research, now is replaced by irregular structure and varied

network. More and more people turn to analyze the network system containing hundreds of billions of nodes instead of a small network. The network formed by adaptable individuals is focused.

According to the latest research, scientists in different fields using complex network theory explain many previously unexplained phenomena, and even subvert a lot of blameless conclusions. Complex network flexibility research can throw a significant impact upon network structure optimization, epidemiology etc. This article studies effects on flexibility of complex network from adjusting parameters of topology based on typical random network and BA scale-free network model. When the network encounters random failures and deliberate attacks, by analyzing its performance changes, corresponding robustness and fragility of complex network are discussed.

1. Statistical Attributes of Complex Network

Complex network in real world holds tens of thousands of nodes and countless edges. Therefore, it is very hard to grasp the law by traditional deconstruction. By means of statistics, analyzing the complex network standing in the global perspective is the only way. Therefore, describing and understanding statistical attributes which embody nature of large-scale nodes and their connection is the first step of the study of complex network[1]. In order to describe statistical attributes of network structure, many concepts and measurement methods are put forward. There are three common attributes: degree distribution, clustering coefficient and the average path length.

1.1 Degree Distribution

The degree k of network node i is defined as the total number of edges connecting i and the other nodes in this network. The meaning of degree is not the same in different network. In friend network, everyone is a node; degree means some person's friends number. Intuitively speaking, the greater the degree of some node is, more important the node is.

Network degree distribution can be represented by probability distribution function P (k), the implication of which is the probability of any node's edges number equaling k exactly. In the current network study, two kinds of distribution are more common: one is exponential degree distribution $P(k) \sim e^{-k}$, $P(k)$ decreases exponentially with k increasing; the other is power-law distribution $P(k) \sim k^{-\gamma}$, γ is called degree index, different γ leads to different network dynamics.

1.2 Clustering Coefficient

Clustering coefficient C is utilized in describing the gathering degree of nodes in the network. For example, in social network, your friends may also be my friend, or two of your friends may also be friends. Clustering coefficient C_i delivers the message that how close connection relationships are between the nodes directly linked with node i in the network. That is to say, C_i is the proportion of actual edge-number between node i and its adjacent nodes accounting for the largest possible number of edges. $C_i = 2e_i/k_i (k_i-1)$ $(0 \leq C_i \leq 1)$, k_i is the degree of node i, e_i represents the actual number of edges between node i and adjacent nodes. Network clustering coefficient C is the arithmetic mean of all nodes in the cluster coefficient, i.e., $C = \frac{1}{N}\sum_{i=1}^{N} C_i$, N is node number of the network. When C = 0, all nodes in the network are isolated points; when C = 1, any node in the network owns edges.

1.3 Average Path Length

The distance that can be represented by d_{ij} between two nodes is defined as the minimum number of edges gone through from one node to the other node. The average path length is the average distance between all pairs of nodes, which describes the degree of separation between the nodes in the network, i.e. how small the network is. The formula of average path length is $L = \frac{1}{N(N-1)}\sum_{i \neq j} d_{ij}$, d_{ij} is the shortest distance between node i and node j. It is found that the average path length of the network is amazing small, although the number of nodes in actual complex network is huge.

2 The Course of Complex Network Development

Looking back the history of science, the development of complex networks has gone through three stages:

1) More than two hundred years ago, Euler pioneered the disciplines of graph theory; scientists believed that relationships between actual system elements can be described by some rules, this is regular network.

2) About 40 years ago, Erdo and Renyi brought forward random graph theory. During the formation of large-scale network, there are no clear design principles. Connections between nodes are completely random, this is random network. Regular network is a symbol of order, while random network is a sign of chaos, and actual network is unlikely to be one of these two extremes.

3) Over the past decade, scientists have

discovered a large number of actual networks are neither regular network nor random network, but the network with statistical characteristics. At the end of the 20th century, two pioneering work opened a new era in the study of complex networks: first, in June 1998, Watts and Strogatz of Cornell University published "collective dynamics of small-world networks" in "nature" magazine. Second, in October 1999, Barabasi and Alber of Notre Dame University published "the emergence of scale in random network" in "science" magazine. These two tasks reveal small-world effect and scale-free effect of complex network; moreover, corresponding models are established to explain forming mechanism of these effects [2].

2.1 Regular Network

In a long time, people thought that the relationship between various elements of the actual system could be indicated by some of regular network disciplines. Usually the regular network is a ring network consisting of N nodes, where each node is connected only with its nearest k nodes. The degree of each node and clustering coefficient is same in the network. The node degree distribution function is P (k) =δ(k-K), node clustering coefficient C = 3 (k-2d)/4(k-d) (d is dimension of network). The higher clustering coefficient of regular network is, the longer average path length will become. The most three common regular networks are shown as following:
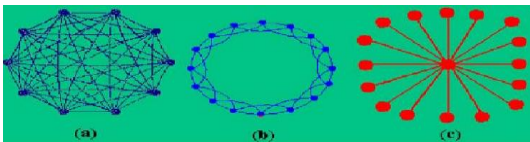

Fig.2 fully coupling network, neighbor coupling network and star coupling network

It can be seen from FIG, fully coupling network contains a ring with N nodes, and any node in the ring is connected with the remaining N-1 nodes. The minimum average path length of this kind of network is $L_{min} = 1$, the maximum clustering coefficient $C_{max} = 1$. Neighbor coupling network also contains an N-node ring, each node on the ring is connected to the K/2 neighbors (K is an even number) around it. The average path length of this network is L = N/2K, clustering coefficient C=$\frac{3(K-2)}{4(K-1)}$. There is a central point in star coupling network; each of the remaining N-1 point is only connected to this central point. The average path length of this network is L = 2-$\frac{2(N-1)}{N(N-1)}$, clustering coefficient C =$\frac{N-1}{N}$.

2.2 Random Network

Random network is put forward by the Hungarian mathematician Erdo and Renyi, it is so-called ER network. In graph theory, N nodes-graph can generate C_N^2 edges at the largest, picking up M edges from all of connections randomly, then random network is shaped.

As shown below, it is assumed that there are a lot of buttons (N >> 1) scattered on the ground, fastening each pair of buttons with the same probability p. In this way, ER random graph instance contains N nodes and PN (N-1) / 2 edges will be formed.
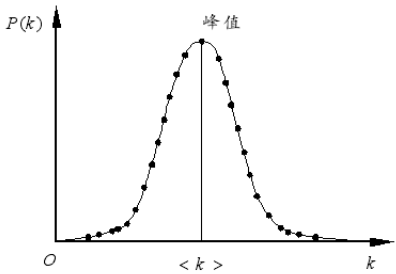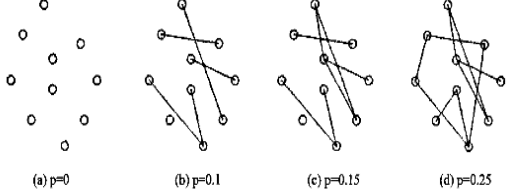

Fig.3 process of ER random graph formation

The network average path length L∝lnN/ln<k>; clustering coefficient C=p=<k>/N, (C<<1); network average degree <k>=p (N-1) ≈ pN; the degree of each node adheres to Poisson distribution approximately. Namely：

$$P (k) \approx \frac{<k>^k}{k!} e^{<k>}$$

It is a bell-shaped curve, as shown on the right. When network average degree k=<k>, peak value will be reached. In this graph, most of node degrees are concentrated in the vicinity of the peak; the proportion of node degree far away from <k> is very small. This phenomenon reveals that most of random network nodes have homogeneity.

Whether two nodes are connected with each other in ER network is no longer a definite thing. It depends on probability. However, in the past 40 years, scientists had been considered random network was the most suitable means to outline the real system. It should be noted that, ER network is classified as static network, because the node number is fixed in advance. Correspondingly, if the number of nodes keeps increasing or decreasing gradually, it will become dynamic and non-balanced, and also known as the evolving network.

2.3 Small World Network

Based on characteristics Integration of regular network and random network, Watts and Strogatz raised the formation mechanism of small-world network model (WS model). The steps of model construction algorithm are listed as follows:

(1) Taking N-nodes neighbor coupling network for instance, each node is respectively connected with adjacent K/2 nodes on the left and right sides. K is an even number.

(2) This N-nodes loop mentioned above, we remove some edges of each node randomly one by one, and reconnect them with probability p (except self-circulation and duplicated connections), these new edges are named as long-range connections.

Because long-range connections shorten average path length of the network greatly, clustering coefficient is improved higher. In view of these two properties (shorter average path length and higher clustering coefficient), we named this kind of network as small-world network (WS network).

The diagram below shows WS model formation process, left one is regular network; right one is random network; small world network is in the middle. Comparing with WS network, average path length of regular network is longer; random network owns more edges. The value of probability p determines the type of network. When p = 0, the network turns to be regular network; while p = 1, it becomes random network; in case of $0 < p < 1$, WS network will be constructed[3].
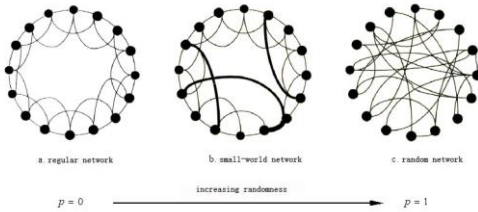

Fig.4 WS model formation process

The average path length, clustering coefficient and degree distribution of WS model are related to model construction probability p, so the small-world network average path length $L = L(p) = \frac{2N}{K} f(NKp/2)$, f (u) is the general scaling function.

$$f(u) = \begin{cases} constant, & u \ll 1 \\ \frac{\ln u}{u}, & u \gg 1 \end{cases}$$

Based on the mean-field method, f (u) can also be approximately expressed as follows:

$$f(u) \approx \frac{1}{2\sqrt{x^2 + 2x}} \arctan h \sqrt{\frac{x}{x+2}}$$

WS small-world network clustering coefficient:

$$C(p) = \frac{3(K-2)}{4(K-1)} (1-p)^3$$ the degree of

WS small-world model, the degree of each node is at least K, when $k \geq K$, the probability P (k) that degree of any node equals k can be calculated.

$$P(k) = \binom{N}{k-K} \left(\frac{Kp}{N}\right)^{k-K} \left(1 - \frac{Kp}{N}\right)^{N-k+K}$$

When k<K, P (k) = 0.

The scientists pointed out that a large number of instances are small-world networks, especially artificial networks including Internet, personal networks. A few years ago, American scientists carried out an experiment. They investigated all of old actor in Hollywood; assuming that each actor is a node. If two of actors cooperate with each other in the same movie, they drew a line between corresponding nodes; finally, actor cooperation network was formed. Statistical data shows that the average "distance" of any pair of actors is smaller than 6. Although the scale of complex network is huge, sometimes it owns even tens of thousands or more nodes, the average path distance between two nodes is much shorter than expectation. Taking friend relationship network for example, there is no small probability that any two persons of someone's friends know each other. In completely random network the probability should be very small. In comparison, the clustering effect of small-world network is more obvious. Empirical studies confirm that: clustering phenomena emerge in many actual networks. The fact has aroused widespread concern[4].

2.4 BA Scale-free Networks

It is found that many of the complex network connectivity distribution functions obey power-law form. Because of the type of network connection to the node is no obvious. There is no definite nodes' degree Eigen length, therefore, this kind of network is so-called scale-free network.

In 1999, Barabasi and Albert raised scale-free network model (also known as BA network model). Two important features of actual networks are taken into account in this model: growth and preferential selection.

The number of nodes in the entire formation process of small-world network is fixed. While the number of network nodes is constantly growing in BA network, that is, the real network evolution owes the continuous growth of nodes. This explanation coincides with actual situation of Internet or other social networks. More importantly, preferential selection tendency exists in the course of establishing connection between new nodes and old ones. For example, lots of new websites are generated every minute,

most of newly created site are clearly preferred to link to well-known websites, such as Sina, Sohu etc. In another example, new actors, of course, are more willing to cooperate with celebrities in the filmmaking.

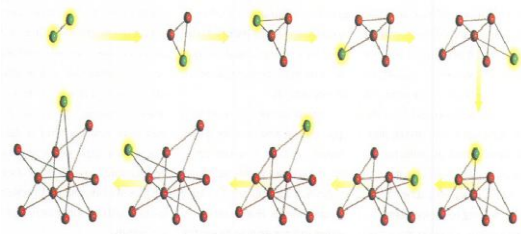BA scale-free network formation process is shown as below:

Fig. 5 formation of BA scale-free network

The steps of BA scale-free network formation are listed as follows[5]:

1) Initial condition: when t = 0, the number of nodes is m_0.

2) Growth: at each interval, a new node which should be connected with m (m $\leq$ m_0) nodes already exists is introduced into the network.

3) Preferential selection: the probability of selection is proportional to the number of connections (degree) of any node i.

$$P(k_i) = \frac{k_i}{\sum_j k_j}$$

In this formula, sum of denominator equals the number of all connections.

The average path length of BA scale-free network $L \propto \frac{\ln N}{\ln\ln N}$; clustering coefficient $C = \frac{m^2(m+1)^2}{4(m-1)}[\ln\left(\frac{m+1}{m}\right) - \frac{1}{m+1}]\frac{[\ln(t)]^2}{t}$; distribution function $P(k) = \frac{2m(m+1)}{k(k+1)(k+2)} \propto 2\, m^2 k^{-3}$. Distribution function of BA scale-free network can be described approximately by the power-law function, as the diagram shown at right.

In the past few years, scientists have found that scale-free is a common phenomenon in a lot of different systems, including Internet, communication networks, social networks and the body's cell metabolism net. The most important features of these networks are scale invariance; most of nodes only have a few attachments but few of the nodes which are called "hubs" (or central nodes) have a large number of connections with other nodes. These two mechanisms of Scale-free network contribute to the formation of hubs; when new nodes appear, they are more likely to attach themselves to hubs, as time goes by, these nodes will have much more connections. This is called "rich get richer" discipline[6].

3 Flexibility Comparison between Random Network and Scale-free Networks

Owing to further research, people gradually discover that actual complex networks are often faced with two attacks: random failure and deliberate attack. Random failure is the damage occurs randomly or with some probability, usually caused by the network itself. Deliberate attack is generally caused by the deliberate destruction; network nodes or edges are destroyed by a certain behavior on purpose. For example, in operation, enemies always select to attack our important military targets rather than random destruction.
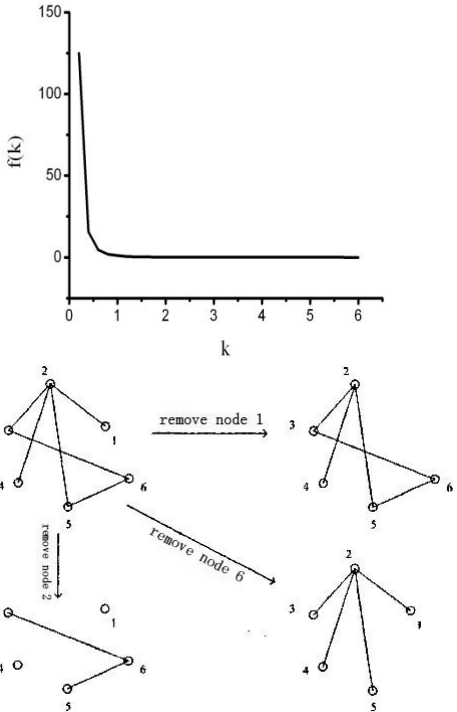

Fig. 6 different consequences of random failure and deliberate destruction

Complex network performances depend on nodes connectivity, we define recovery ability in network connectivity after destruction as network flexibility, and random failure and intentional attack are two ways to analyze network flexibility. The analysis of the former is referred to network robustness, the latter is vulnerability[7]. The focus of this chapter start

from these two analysis methods, robustness and fragility of the random network and scale-free network will be discussed in this chapter comparatively.

3.1 Establishment of Criterion

The connectivity of the network nodes (abbreviated as CN, $0 \leq CN \leq 1$) has a great impact on the network robustness and fragility. When CN = 0, all network nodes are isolated points. If CN = 1, the distance of any two nodes in the network are shortest. To a large extent, network performance is proportional to network connectivity. Better network connectivity, higher efficiency. We define network efficiency as E[8].

$$E = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}}$$

d_{ij} is the distance between node i and node j. It can be said that efficiency of the network reflects the degree of the network average connectivity. $CN \propto f(E)$. The larger CN increases, the higher the efficiency of the network E becomes, the robustness of the network will be better; on the contrary, it will be fragile.

3.2 Experimental Analysis

We made a comparison experiment. Firstly, random network and scale-free network are constructed as below, the left one is random network, the right one is scale-free network.
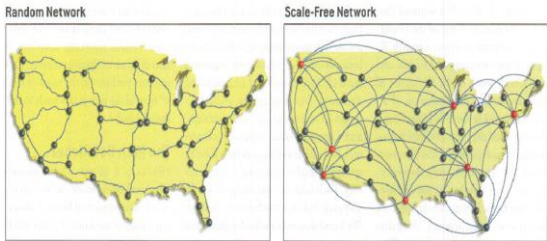

Fig.7 random network and scale-free network

(1) Random Failure Experiment

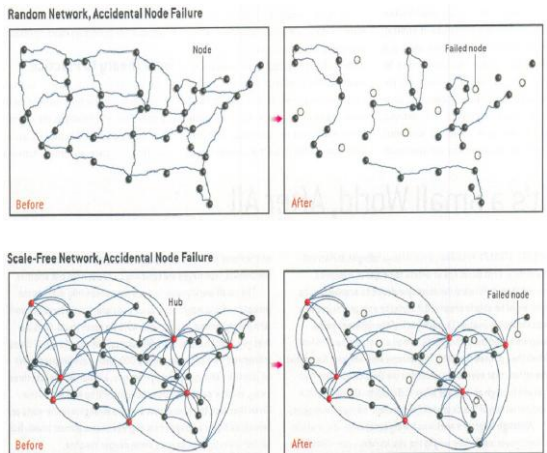After randomly delete some network nodes, two types of network connectivity are shown as below.

Fig.8 the first comparative experiment

(2) Deliberate Attack Experiment

After attacking the network hubs, scale-free network connectivity is shown as follows.
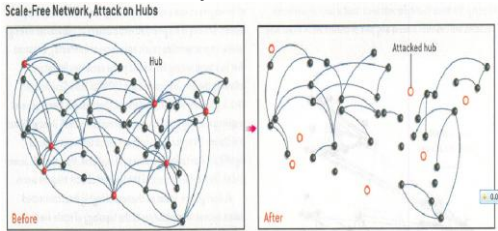

Fig.9 the second experiment

In accordance with these two experiment results, we drew function curves of efficiency E and delete node P.
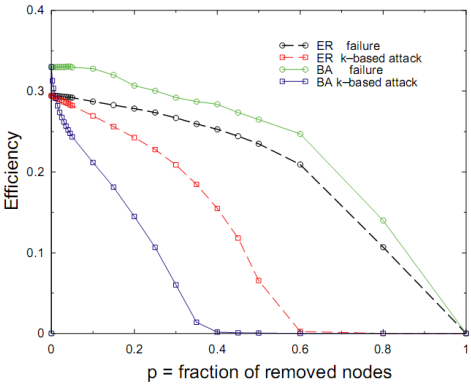

Fig.10 experimental result curves

Black curve and green curve represent respectively E-P function of random network and BA scale-free network in random failures experiment. Red curve and blue curve correspond to E-P function of random network and BA scale-free network separately in a deliberate attack experiment.

In case of removing the same node P randomly, it can be seen that E_{BA} is no less than E_{ran}; Scale-free network is more robust than random network. Under the condition of deliberate destruction, $E_{BA} < E_{ran}$, scale-free network is more vulnerable.

It has a strong stability or robustness in the face of random attacks, the reason is that network connections obey power law distribution. Scale-free network is a system predominated by a small number of hubs. Random failures are likely to damage unimportant nodes, whose number is much larger than the number of hubs, comparing with deliberate destruction. These unimportant nodes have only a small number of connections, even if the node is removed; there is no significant impact on network topology. For random network, node connection distribution follows bell curve distribution. According to this distribution, most of the nodes have around number of connections, hubs absolutely impossible exists in random network. Some

occasional random damages may lead to collapse of network.

Learning from the above experiments, we know that scale-free network is very fragile for deliberate attack. Removed a small portion of hub nodes, the entire network connectivity will decline severely; the robustness of the network is greatly reduced or even completely lost.

In summary, the following conclusions can be made: random failure throws no effect on efficiency of the BA network, on the contrary, in regard to deliberate attack; efficiency of BA network is declining much faster than random network. BA model both has strong robustness and frangibility relative to random network.

4 Summaries

In this article, gives the basic concept and three commonly used statistical attributes of complex network are clarified. Development process and commonly used models are introduced briefly. Taking network transmission efficiency E_i and frangibility V_i as criterion, the flexibility of ER network and BA scale-free network are researched respectively encountering with random failure and deliberate attack. At last, comparative experiments are carried out. Experimental data proofs that BA scale-free network have stronger robustness in random failure, but lower fragility in deliberate destruction comparing with ER network.

References

[l] Zheng Xiaojing, Xu Song Xu complex networks [J]. techno-economic, 2010, 29 (9): 11 ~ 16.

[2] Wu Jin-pei complex network of China [J]. Wuyi University (Natural Science), 2010, 24 (2): 1-12.

[3] He Da-Ren, Commercial Banks, WANG Bing-hong complex systems and complex networks [M]. Beijing: Higher Education Press, 2009.

[4] Liu Tao, Chen Zhong, Chen Xiao-rong complex network theory and its application research [J]. System Engineering, 2005, 23 (6): 1-6.

[5] Newman M E J. The structure and function of complex network s[J]. S IAM Review , 2003, 45: 167~ 256.

[6] Newman M E J. Assortative mixing in networks[J]. Phys. Rev. Lett, 2002, 89: 208701.

[7] Erdos P, R é nyiA. On random graphs[J]. Publication Mathematics, 1959, 6: 290~ 297.

[8] Barrat A ,Weigh t M. On the properties of small-world network models[J]. Eur. Phys. J. , 2000,B13: 547~ 560.