Vol 16 No 12, December 2007
1009-1963/2007/16(12)/3571-10

Chinese Physics

© 2007 Chin. Phys. Soc.
and IOP Publishing Ltd

The effects of degree correlations on network topologies and robustness*

Zhao Jing(赵 静)[a)b)d)], Tao Lin(陶 林)[b)], Yu Hong(俞 鸿)[b)],
Luo Jian-Hua(骆建华)[a)], Cao Zhi-Wei(曹志伟)[b)†], and Li Yi-Xue(李亦学)[a)b)c)‡]

[a)] *School of Life Sciences & Technology, Shanghai Jiaotong University, Shanghai* 200240, *China*
[b)] *Shanghai Center for Bioinformation and Technology, Shanghai* 200235, *China*
[c)] *Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai* 200031, *China*
[d)] *Department of Mathematics, Logistical Engineering University, Chongqing* 400016, *China*

(Received 20 March 2007; revised manuscript received 20 April 2007)

Complex networks have been applied to model numerous interactive nonlinear systems in the real world. Knowledge about network topology is crucial to an understanding of the function, performance and evolution of complex systems. In the last few years, many network metrics and models have been proposed to investigate the network topology, dynamics and evolution. Since these network metrics and models are derived from a wide range of studies, a systematic study is required to investigate the correlations among them. The present paper explores the effect of degree correlation on the other network metrics through studying an ensemble of graphs where the degree sequence (set of degrees) is fixed. We show that to some extent, the characteristic path length, clustering coefficient, modular extent and robustness of networks are directly influenced by the degree correlation.

Keywords: network dynamics, random graphs, complex networks, degree correlation
PACC: 0250, 0500, 0520, 8000

1. Introduction

Complex networks have been playing an important role in understanding many natural systems.[1−6] The empirical analysis of many real-world networks from various realms has revealed some common features shared by the individual systems, such as power-law degree distribution,[7,8] modular or community organization,[9−14] small-world effect,[15] hierarchical organization,[16] as well as some apparent differences between networks, such as the positive degree correlation in social networks vs the negative degree correlation in technological and biological networks.[17−22] At the same time, several network metrics have been proposed to describe network topology quantitatively, in which the most important and commonly used ones are the degree distribution,[7], the characteristic path length,[15] clustering coefficient,[15] modularity,[10,23] and assortative coefficient.[18] On the other hand, a considerable number of network models have been developed to simulate networks in the real world,[7,24−26] among which the most widely used models are the Barabasi and Albert's (BA) preferential attachment model for the evolution of networks,[7] and the Ravasz and Barabasi (RB) model for the hierarchically modular organization pattern of networks.[24] Since the network metrics and models are derived from a wide range of studies, a systematic study is required to investigate the possible correlations among the network metrics, as well as the applicabilities of network models.

In this paper, we study how the network topologies and performances are influenced by the degree correlation through investigating a statistical ensemble generated from a certain 'seed' network, which includes the maximally random networks under the constraints of a given degree sequence and fixed assortativity.[27] Two model networks and two real bi-

*Project supported by the Research Foundation from Ministry of Science and Technology, China (Grant Nos 2006AA02Z317, 2004CB720103, 2003CB715901 and 2006AA02312), the National High Technology Research and Development Program of China (Grant No 2006AA020805), the National Natural Science Foundation of China (Grant Nos 30500107, 30670953 and 30670574), the International Cooperation Project of Science and Technology Commission of Shanghai Municipality, China (Grant No 06RS07109), and Grant from Science and Technology Commission of Shanghai Municipality, China (Grant Nos 04DZ19850, 06PJ14072 and 04DZ14005).
†Corresponding author. E-mail: zwcao@scbit.org
‡Corresponding author. E-mail: yxli@sibs.ac.cn

ological networks are applied as seed networks.

2. Seed networks

In the present study, the following four networks of which the degree distributions obey the power-law are studied as seed networks:

(1) Seed network A: the hierarchically modular network constructed by Ravasz and Barabasi (RB model)[24] in the 3rd iteration. At every step, the central node of the central module is connected to the external nodes of the four peripheral modules (see Fig.1 in Ref.[24]).

(2) Seed network B: a model network constructed by Barabasi and Albert's (BA) preferential attachment model with parameter values $m = m_0 = 3$.[7]

(3) Seed network C: the biggest connected cluster of the E.coli metabolic network,[9] in which nodes correspond to substrates and edges to reactions. In order to reflect biologically relevant transformations of substrates, the top 10 most common substrates are removed.[9,28]

(4) Seed network D: the biggest connected cluster of the protein interaction network CCSB-HI1, a proteome-scale map of human binary protein–protein interactions tested from a high-throughput yeast two-hybrid experimental system.[29]

In Table 1, we summarize the basic graph metrics of these networks. We quantitate the topological features of the networks by their characteristic path length (L), clustering coefficient (C) and modularity (M), and the degree correlation by assortative coefficient (r).

Table 1. Basic graph metrics of the seed networks.

Seed network	A	B	C	D
Number of nodes	625	1000	716	1307
Number of links	1976	2991	1425	2483
Degree distribution	$P(k) \sim k^{-2.42}$	$P(k) \sim k^{-3}$	$P(k) \sim k^{-2.78}$	$P(k) \sim k^{-2.51}$
Characteristic path length (L)	3.36	3.59	4.54	4.36
Clustering coefficient (C)	0.5821	0.0246	0.0928	0.0327
Modularity (M)	0.688	0.4046	0.6637	0.587
Assortative coefficient (r)	-0.1516	-0.0527	-0.084	-0.2228

3. Network ensemble based on the spectrum of degree-correlation

It has been found that networks with the same power-law degree distribution exhibit significant difference in their topological features.[30−32] Here we construct a network ensemble with a fixed degree sequence and varying degree-correlation.

3.1. Topological metrics of degree correlation

The degree correlation of a network is the pattern of correlations between degrees of neighbouring nodes.[17−21] It could be simply measured by the assortative coefficient,[18] defined as the Pearson's correlation coefficient of degrees on either side of an edge, of which theoretical range is $[-1, 1]$. High values of assortative coefficient mean the tendency for high-degree nodes to be attached to other high-degree ones, and low values represent the preference of linkages between high-degree nodes and low-degree ones.

There are some other ways of quantitating degree correlations.[20,21,30] Li *et al*[30] defined the s-metric as the sum of products of degrees of adjacent nodes as follows:

$$s = \sum_{(i,j)\in E} d_i d_j, \tag{1}$$

where d_i is the degree of node i, and E is the edge set of graphs. They also proved that s-metric was linearly correlated with assortativity coefficient.[30] Therefore, for a given degree sequence, the biggest and smallest s-values, denoted by $S_{\max}$ and $S_{\min}$ respectively, correspond to the extremes of assortativity. The simple connected graphs with the extreme s-values are called $S_{\max}$ graph and $S_{\min}$ graph respectively.

3.2. Generating extreme networks of degree correlation

To construct the ensemble of networks, we first generate the extreme graphs of degree correlation from the seed network. They are networks that have the highest and lowest values of assortative coefficient while the degree sequence of the seed network is conserved.

We apply the algorithm proposed by Li *et al* to constructing the extreme graphs.[30] This algorithm uses an ordered sequence of all potential links as its input, and then checks the sequence from top to bottom to select one edge at a time so as to construct a simple, connected graph of given degree sequence D. Reference [30] constructs the $S_{\max}$ graph for a given seed network, in which all potential links (i, j) for all node pairs $i, j (i < j,\ d_i \geq d_j;\ i, j = 1, 2, \ldots, N)$ are ordered according to their weight $d_i d_j$ in a descending manner.

Here, we first order the potential links for the $S_{\max}$ graph and $S_{\min}$ graph according to the rearrangement inequality:[30]

$$a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \geq a'_1 b_1 + a'_2 b_2 + \cdots + a'_n b_n \geq a_n b_1 + a_n b_2 + \cdots + a_n b_n,$$

where $a_1 \geq a_2 \geq \cdots \geq a_n$, $b_1 \geq b_2 \geq \cdots \geq b_n$, and $(a'_1, a'_2, \ldots, a'_n)$ is any permutation of $(a_1, a_2, \ldots, a_n)$. Then we apply the ordered sequences as inputs of the algorithm in Ref.[30] to constructing the $S_{\max}$ graph and $S_{\min}$ graph, separately.

Obviously, there are a lot of permutations of potential links that satisfy the rearrange inequality, especially when the degree sequence includes many nodes of the same degrees. We order the potential links as follows:

1. For the degree sequence of the seed network, order the nodes according to descending values of node degree, i.e. $d_1 \geq d_2 \geq \cdots \geq d_N$. More specifically speaking, the nodes are labelled as $(1, 2, \ldots, k_1, k_1+1, \ldots, k_2, \ldots, k_{m-1}+1, \ldots, k_m)$ such that

$$\begin{aligned} d_1 &= d_2 = \cdots = d_{k_1} > d_{k_1+1} \\ &= \cdots = d_{k_2} > \cdots > d_{k_{m-1}+1} = \cdots = d_{k_m}, \\ k_m &= N. \end{aligned}$$

2. For each node i $(i = 1, 2, \ldots, N-1)$, order its potential links for the $S_{\max}(S_{\min})$ graph descendingly (ascendingly) according to the products of degrees of the nodes on either side of an edge.

(a) For the $S_{\max}$ graph, order all potential links of node i as follows: $L_i = \{(i, i+1), (i, i+2), \ldots, (i, N)\}$.

(b) For the $S_{\min}$ graph, assuming $k_{p-1} \leq i < k_p$, order all potential links of node i as follows:

$$L'_i = \begin{cases} \{(i, k_{m-1}+1), \ldots, (i, N), (i, k_{m-2}+1), \ldots, (i, k_{m-1}), \ldots, (i, i+1), \ldots, (i, k_p), & \text{if } p < m, \\ \{(i, i+1), (i, i+2), \ldots, (i, N)\} & \text{if } p = m. \end{cases}$$

3. (a) For case 2(a), order all potential links of the $S_{\max}$ graph as $L_1, L_2, \ldots, L_{N-1}$. (b) For case 2(b), order all potential links of the $S_{\min}$ graph as $L'_1, L'_2, \ldots, L'_{N-1}$.

Figure 1 shows the resulting extreme graphs for a small seed network as an example. As can be seen, the extreme networks generated by this algorithm exhibit a chain-like topology. In this $S_{\max}$ graph, nodes with similar degrees are linked to form quasi-cliques, while in the $S_{\min}$ graph, high degree nodes are matched with low degree nodes to form bipartite quasi-cliques, and the remaining unmatched nodes are linked with each other in a clique-like modular manner.

3.3. Constructing network ensemble from the $S_{\max}$ and $S_{\min}$ graphs

Starting from the $S_{\max}$ and $S_{\min}$ graphs respectively, we carry out a series of edges rewiring in two steps as follows:

Step 1 In one direction, we rewire the links of $S_{\max}$ and $S_{\min}$ graphs respectively, during which two links between two randomly selected pairs of nodes are rewired only if this step generates no multiple edges while keeps the graph connected.[21]

The process is repeated pM times in which only performed exchanges are counted, where M is the total number of links in the graph, and p is a positive number representing the rewiring fraction of links. Switching p in the area [0, 3], we generate a set of randomized networks for the $S_{\max}$ graph and the $S_{\min}$ graph, respectively. For $p = 0$, the original $S_{\max}$ and $S_{\min}$ graphs are unchanged.

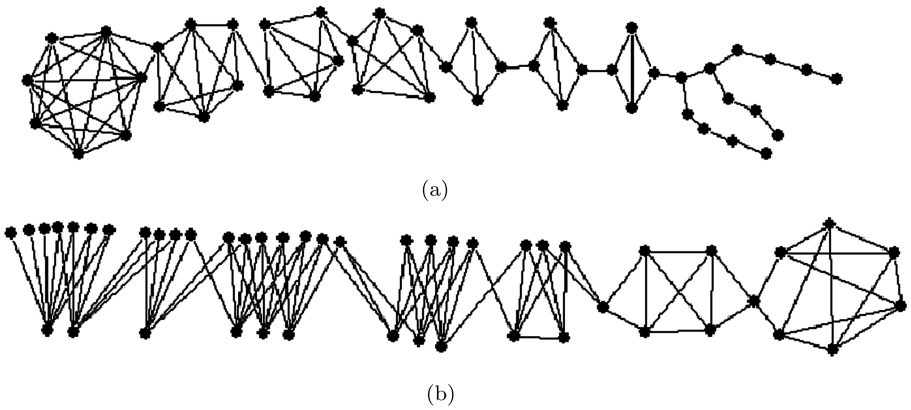
Step 2 In another direction, we rewire each network obtained in step 1 in the way that its degree sequence and assortative coefficient are kept. That is to say, a pair of randomly selected edges (v_1, v_2) and (v_3, v_4) is rewired to (v_1, v_3), (v_2, v_4) only if they satisfy the following conditions:

(1) Edges (v_1, v_3), (v_2, v_4) do not exist in the original graph,

(2) $d(v_1) \times d(v_2) + d(v_3) \times d(v_4) = d(v_1) \times d(v_3) + d(v_2) \times d(v_4)$,

(3) The new graph is still connected.

This process is repeated sufficiently many times. Thus each network generated from step 2 is the maximally random network under the constraints of a given degree sequence as well as fixed assortativity.[27] All of these resulting networks constitute a network ensemble in this study, and we represent the ensemble of seed network G as $\mathcal{G}_{\mathrm{D}}^{\mathrm{A}}(\mathrm{G})$.



(a)

(b)

Fig.1. The $S_{\max}$ graph (a) and $S_{\min}$ graph (b) for a small seed network.

4. Numerical results

In this section, we will study the possible relation between degree correlation and the other structural metrics numerically by using the resulting network ensemble $\mathcal{G}_{\mathrm{D}}^{\mathrm{A}}(\mathrm{G})$ of our four seed networks.

4.1. Assortative coefficient

To probe the spectrum of degree correlation in each ensemble, we plot the assortative coefficient r of the networks generated in step 1 as a function of the rewiring fraction p in Fig.2. It can be seen that the change tendencies of assortative coefficient r for the randomization of the $S_{\max}$ graph and the $S_{\min}$ graph are opposite. As p increases from 0 to 1, assortative coefficients change accordingly between the maximum and minimum values, and then they are convergent to a constant when $p \geq 1$. This tendency suggests that after performing more than M times ($p \geq 1$), the random rewiring process reaches a good mixing, thus generating the maximally random networks under only one constraint of a given degree sequence.[27] We represent this maximally random network ensemble of seed network G as $\mathcal{G}_{\mathrm{D}}(\mathrm{G})$, and the limit value of the assortative coefficient as r_{c}. Hence the latter corresponds to the average r-value of the $\mathcal{G}_{\mathrm{D}}(\mathrm{G})$ ensemble. It is worth noting that both the ensemble $\mathcal{G}_{\mathrm{D}}(\mathrm{G})$ and $\mathcal{G}_{\mathrm{D}}^{\mathrm{A}}(\mathrm{G})$ represent the statistical sets of randomized networks with a given degree sequence, yet the randomized extents of networks in these ensembles are different. In the process of randomization, networks in $\mathcal{G}_{\mathrm{D}}^{\mathrm{A}}(\mathrm{G})$ have one constraint of assortativity more than those in $\mathcal{G}_{\mathrm{D}}(\mathrm{G})$.

Figure 2 indicates that the rewiring process in step 1 allows us to tune the network between the maximum and minimum levels of degree correlation, and the resulting network set spans the whole range of assortativity based on the same degree sequence. Therefore, the ensemble $\mathcal{G}_{\mathrm{D}}^{\mathrm{A}}(\mathrm{G})$, which is constructed by a randomization process that keeps the degree sequence and assortativity as described in step 2, can be used to probe the transformation of network topologies depending on the maximum and minimum levels of its degree correlation.

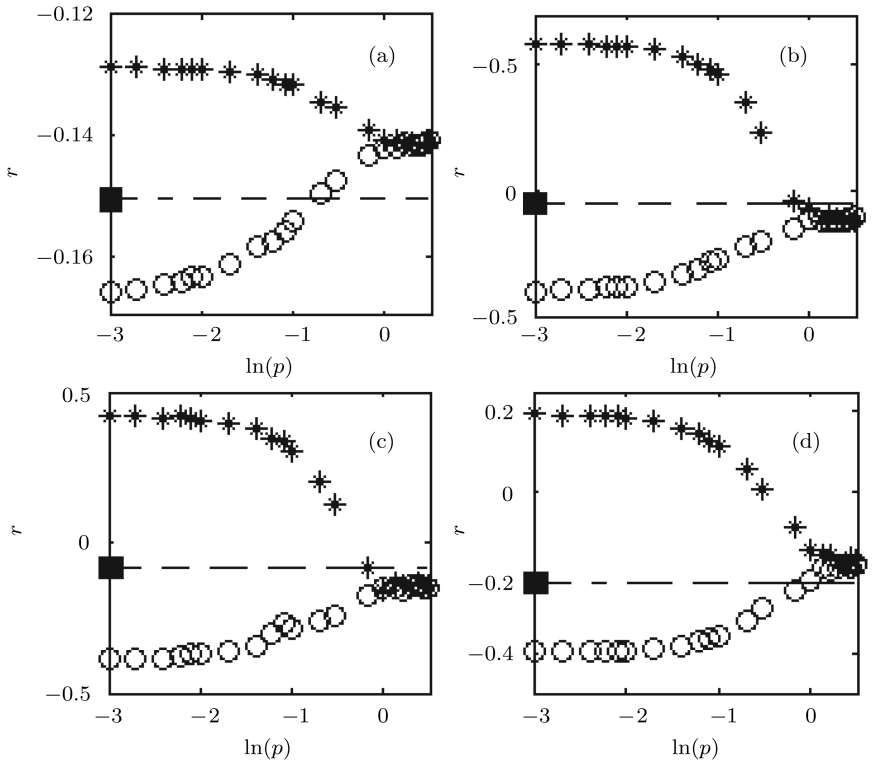


Fig.2. Assortative coefficient (r) as function of the randomization fraction (p). A logarithmic horizontal scale is used to depict the mini increasing of rewiring fraction. The solid square corresponds to the (p, r)-value of the seed network. The star and open circle indicate the values of the randomized networks from the $S_{\max}$ graph and the $S_{\min}$ graph, respectively.

Figure 2 also illustrates that the practical ranges of assortative coefficient are far less than the theoretical limit $[-1, 1]$, and not all the maximally random networks under the constraint of a given degree sequence ($p \geq 1$) have an assortative coefficient near 0. Except maximally random networks of seed network B, those for the other seed networks are disassortative mixing with negative assortative coefficient. This observation can be explained by the structural cut-off constraints, i.e. simple random graphs of which the degree distribution follows a power law with an exponent γ must satisfy the condition of $\Delta \leq \sqrt{N}$ (Δ and N are the highest degree and total number of nodes respectively) in order to have no degree-correlations.[33] This cut-off constraint is equivalent to $\gamma \geq 3$,[33,34] which is only satisfied by network B. Figure 2 also displays that even the assortative coefficient of the $S_{\max}$ graph for seed network A is negative, suggesting that any connected simple graph having the same degree sequence as that for network A is disassortative mixing, no matter how its nodes are connected. This can be caused by the highly heterogeneous distribution of node degrees in network A, in which $\Delta > \frac{1}{2}N$ but the total number of high-degree nodes is much less than $\frac{1}{2}N$. Thus the high-degree nodes have to link with much more low-degree nodes than high-degree ones, resulting in the disassortative mixing feature.

4.2. Characteristic path length

Characteristic path length is defined as the average length of the shortest paths between any pair of nodes in the network.[15] It can measure the average communication speed between the nodes of the network.

As can be seen from Fig.1, the extreme networks are linked in a chain-like way. It is true that all of the resulting extreme networks of the four seed networks, except the $S_{\max}$ graph for seed network D, have a chain-like structure with a very long path. That is to say, the characteristic path lengths of these networks are much larger than those in logarithmical scale of the network size. The reason for the exceptional case (the $S_{\max}$ graph for seed network D) is that network D has a lot of degree-one nodes that need to link with high-degree nodes for keeping the connectivity of the network. Although the characteristic path lengths of almost all scale-free networks in the real world scale log-

arithmically with the system size (called small-world networks), our result suggests that scale-free networks could also be large-world networks. Actually, network models with a power-law degree distribution but large characteristic path length have been proposed in Refs.[35–37].

Figure 3 displays the relationship between the characteristic path length (L) and assortative coefficient (r) in the network ensemble $\mathcal{G}_{\rm D}^{\rm A}({\rm G})$. As can be seen, the characteristic path length of the seed networks is close to the minimum value, which is also the average L-value in the $\mathcal{G}_{\rm D}({\rm G})$ ensemble. The curves in Fig.3 (except that in (a)) all have a similar changing tendency. The characteristic path length approaches a minimum when r is near to $r_{\rm c}$, and it decreases before $r = r_{\rm c}$ and increases after $r = r_{\rm c}$. Moreover, the closer to the extreme values the r-value is, the longer the corresponding characteristic path length will be. Such a tendency can be explained by the linkage patterns of extreme graphs. In networks with a higher assortativity, an increase in path length could be a compensation for the strict constraint of links between homogeneous nodes. On the other hand, the higher disassortativity, the stronger its bipartivity,[39] thus the longer its path length. This is because bipartite graphs usually have a longer path length than unipartite ones of the same sizes. In Fig.3 there exists only the exceptional case in which the disassortative networks in the ensemble of seed network A do not exhibit obviously a longer path length, which can also be the result of the highly heterogeneous distribution of node degrees in network A, as shown in Section 4.1. In highly disassortative graphs of seed network A, a few nodes with extremely large degrees are connected with a huge number of low-degree nodes to form a core-like bipartite sub-graph, thus effectively shortening the average distance between nodes.

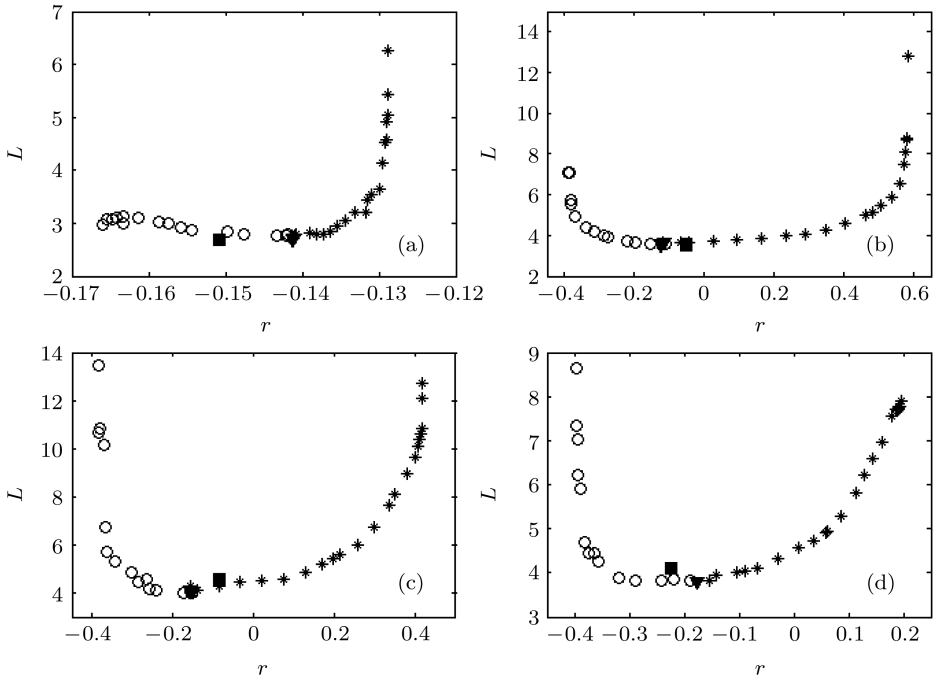


Fig.3. Relationship between characteristic path length (L) and assortative coefficient (r). The solid triangle corresponds to the average (r, L)-value of the $\mathcal{G}_{\rm D}({\rm G})$ ensemble, and the other symbols correspond to those in Fig.2. The data shown in the subfigures are averaged over 10 random realizations of the rewiring process.

4.3. Clustering coefficient

It has been reported that many categories of real-world networks are enriched with triangles. In friendship networks, a high density of triangles can be caused by the fact that friends of the same people are usually friends of each other. Clustering coefficient is a network metric to quantitate the density of triangles in a network. Here we apply the definition in Ref.[15], in which the clustering coefficient of node v is defined as the probability with which two neighbouring nodes of v are also connected with each other

$$CC(v) = \frac{2\,|N(v)|}{d(v)(d(v)-1)}, \tag{2}$$

where $|N(v)|$ denotes the number of links between neighbours of node v, $d(v)$ is the degree of node v. The clustering coefficient of a network is defined as the average of $CC(v)$ over all v.

Figure 4 shows that the clustering coefficient is a monotonically ascending function of assortative coefficient. This result can be explained intuitively from Fig.1: networks with higher assortativities include more quasi-cliques, in which there are high densities of triangles; whereas those with lower assortativities contain more bipartite sub-graphs, in which there are no triangles. The correlation between these two variables also agrees with the earlier observation.[38] In addition, Fig.4 displays that the seed network is significantly more clustered than the $\mathcal{G}_{\mathrm{D}}(\mathrm{G})$ average.

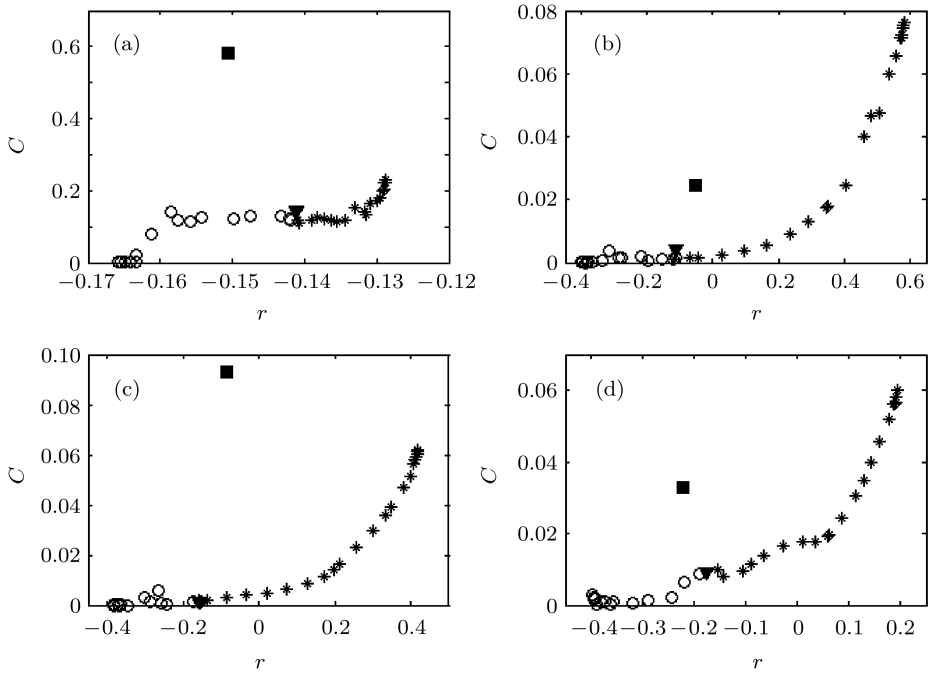


Fig.4. Relationship between clustering coefficient (C) and assortative coefficient (r). The symbols correspond to those in Fig.3. The data shown in the figure are averaged over 10 random realizations of the rewiring process.

4.4. Modularity

Many real-world networks, including social networks, technological networks, and biological networks, are found to divide naturally into communities or modules,[9,11], i.e. densely connected sub-networks within which there is a high density of edges, and between which there is a lower density of edges.[10] The modularity (M) metric was defined initially to quantitatively compare different partitions of a network[10]

$$M = \sum_{i=1}^{r} \left[e_{ii} - \left(\sum_{j} e_{ij} \right)^2 \right], \qquad (3)$$

where r is the number of clusters, e_{ij} is the fraction of edges that lie between vertices of cluster i and j. Guimera *et al*[23] then defined the modularity metric of a network as the largest modularity metric in all possible partitions of the network, and developed a simulated annealing algorithm to compute this metric.[39,40] Thus the modularity metrics of networks can be applied to measuring the modular extent of different networks.

Figure 5 depicts the modularity as a function of the assortativity, illustrating that networks with the assortative coefficient near to r_{c} exhibit the lowest extent of modularity, and the closer to the maximum or minimum value the assortative coefficient r of a network is, the bigger its modularity M will be. Newman and Park[19] have suggested that the formation of communities (or modules) can produce highly assortative mixing. Our results indicate that the significantly high extent of assortative mixing or disassor-

tative mixing can result in a high level of modularity. The phenomenon in which the networks of extreme assortativity exhibit a high extent of modularity can be observed directly from the extreme networks shown in Fig.1. Actually, networks that show both rich and variable community structure as well as disassortative mixing have also been found in some network models of mechanical assemblies.[41]

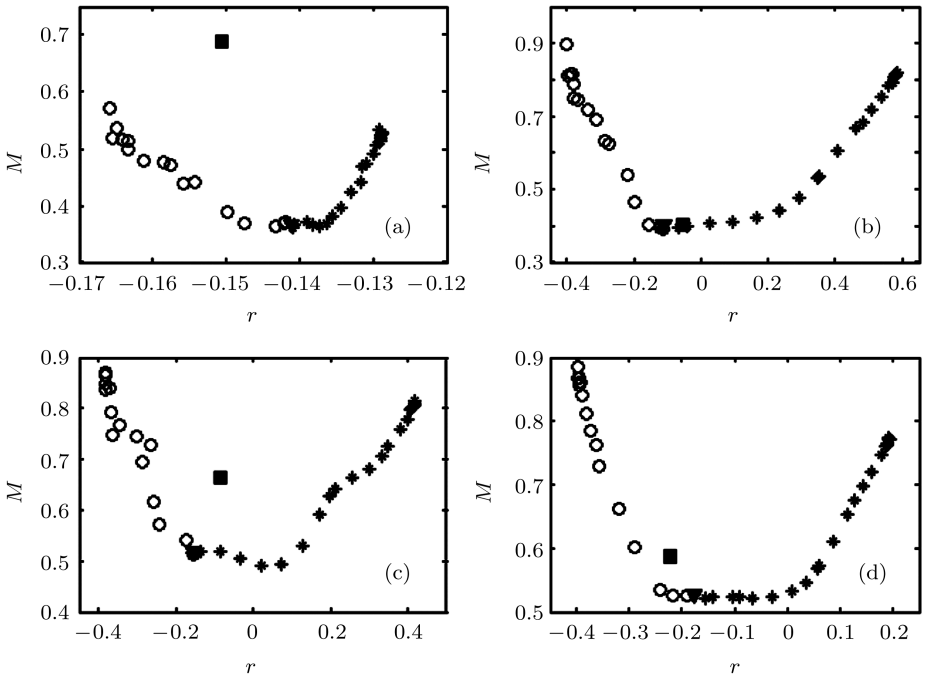


Fig.5. Relationship between modularity (M) and assortative coefficient (r). The symbols correspond to those in Fig.3. The data shown in the figures are averaged over 10 random realizations of the rewiring process.

Figure 5 also shows that the real biological networks, seed networks C and D, have a significantly higher extent of modularity than networks with the same assortativities in the ensemble $\mathcal{G}_{\mathrm{D}}^{\mathrm{A}}(\mathrm{G})$, suggesting that the modularity might be an intrinsic feature of biological networks developed during a long evolutionary process. However, seed network B, a realization of BA model, does not exhibit a significantly high level of modularity, indicating that the BA model cannot simulate the modular feature of real networks so well as the RB model can.

4.5. Network resilience

Most dynamical processes taking place in the topology defined by the network, such as information communication and disease spreading,[42−44] rely heavily on the connectivity of networks. When some nodes are removed, the network might be fragmentized and the interactions between nodes are thus interrupted. Along this line, two ways of damaging networks, i.e. intentional attacks on target hubs (attacks) and random removal of nodes (failures), and the network performances due to these actions have been usually studied.[45−48]

In this study, we measure the network performance under attacks or failures by the f-robustness metric, which is defined as the fraction of vertices expected to be removed so that the ratio between the size of the largest connected component for the resulting network and that of the original network is equal to $f \in (0, 1)$.[31] We set $f = 1/2$, and represent the 1/2-robustness under attacks or failures as R_{a} or R_{f}, respectively. In Fig.6, we display R_{a} and R_{f} respectively as a function of assortativity. A comparison of the subfigures in the first row with the corresponding ones in the second row shows that, for the same network, the value of R_{f} is significantly bigger than that of R_{a}. This observation agrees with the earlier finding that scale-free networks are resilient to random failures but relatively vulnerable to intentional attacks.[45]

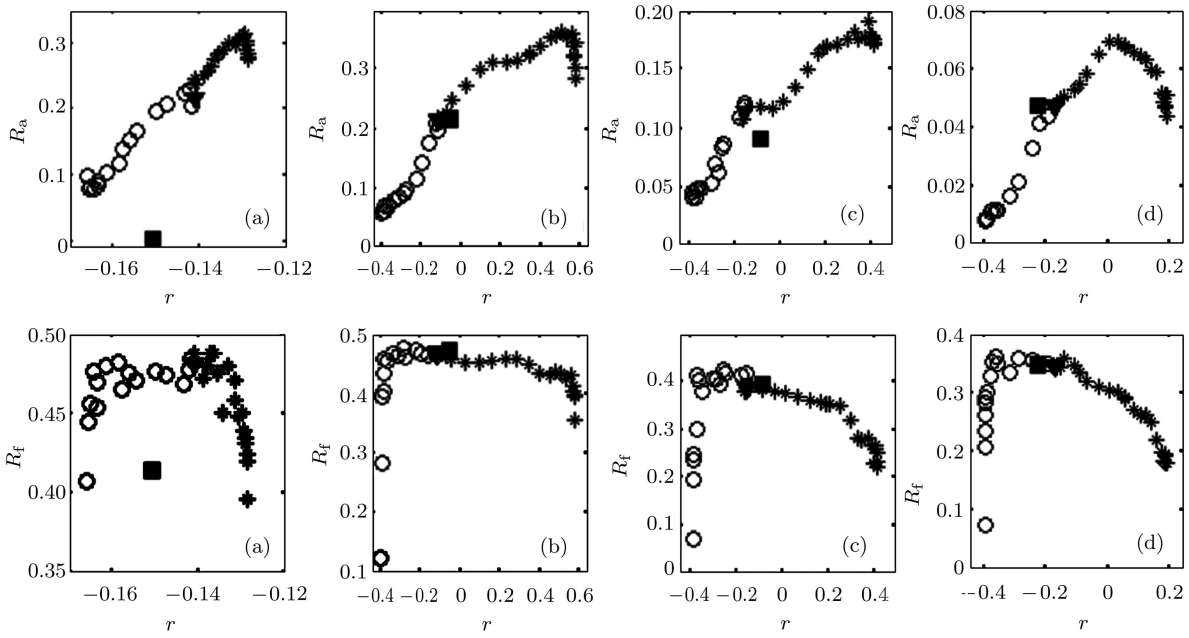


Fig.6. The effect of degree correlation on network robustness. The symbols correspond to those of Fig.3. Subfigures in the first and second rows depict the robustness under attacks and failures as a function of assortativity, respectively. The data shown in the subfigures are averaged over 10 random realizations of the rewiring process.

The subfigures in the first row of Fig.6 illustrate that the robustness under attacks is roughly an ascending function of assortativity. The only exception is the area of the assortative coefficient near the maximum value, in which R_a decreases with assortativity. This transition may stem from the chain-like topologies of these networks. Although R_a does not monotonically ascending in the area near the maximum assortativity, Fig.6 indicates that, compared with dissortative networks, assortative networks are robust to removal of their highest degree nodes, supporting the argument in Refs.[18,49]. The subfigures in the second row of Fig.6 show that networks with assortative coefficients near r_c exhibit the highest extent of robustness under failures, and the robustness decreases as the assortative coefficients approach to the extreme value.

As can be seen, the two real biological networks, i.e. seed networks C and D, exhibit almost the best performance under failures in their own ensembles, so does seed network B, a network which is constructed by the BA model. These three networks also show nearly the same extent of robustness under attacks as the networks with the same assortativity in the ensemble $\mathcal{G}_{\mathrm{D}}^{\mathrm{A}}(\mathrm{G})$. Thus the higher extent of robustness of real biological networks might be a result of evolution. However, although seed network A has been used as a model to characterize real-world networks such as metabolic networks and protein interaction networks,[24] its performances under attacks and failures are significantly worse than those of the real biological networks, suggesting that the RB model cannot simulate the robustness of real networks so appropriately as the BA model can.

5. Conclusion

In this study, through exploring the maximally random ensemble of networks under the constraints of a given degree sequence and fixed assortativity, we have conducted a system-level survey about how the degree correlation of networks is related to their topologies and performances. As illustrated, this ensemble can give us more information than before. By investigating this ensemble, we find that networks with the same power-law degree sequence may have very different topological features and performances under attacks and failures, and these features are partially determined by the degree correlation. In addition, our study indicates that both BA and RB models can simulate the small-world features of real biological networks, i.e. short path length and high clustering. However, the BA model is good at simulating the robustness but weak at mimicking the modular feature of real networks, whereas the RB model is just opposite, suggesting that these models still have limitations

in depicting the main features of real networks. Thus more network models are expected to be developed. The approach used in this study, which compares real networks with their random counterparts at different extents, places the network under study in a context of comparable statistical weight. This approach may help us have a clearer insight into the topological properties of the network, thus could be promising in the study of complex networks.

Acknowledgement

We thank Luís A. Nunes Amaral and Roger Guimerà for kindly providing us with the software Modul-w of computing the network modularity metric. Our gratitude must also be extended to John Doyle, Petter Holme and Lun Li for useful discussion and constructive comments.

References

[1] Newman M E J 2003 *SIAM Rev.* **45** 167
[2] Barabasi A L and Oltvai Z N 2004 *Nat. Rev. Gen.* **5** 101
[3] Boccaletti S, Latora V, Moreno Y, Chavez M and Hwang D U 2006 *Phys. Rep.* **424** 175
[4] Costa L D F, Rodrigues F A, Travieso G and Villas Boas P 2007 *Adv. Phys.* **56** 167
[5] Zhao J, Yu H, Luo J, Cao Z and Li Y 2006 *Chin. Sci. Bull.* **51** 1529
[6] Zhao J, Tao L, Yu H, Luo J H, Cao Z W and Li Y 2007 *Chin. Sci. Bull.* **52** 1036
[7] Barabasi AL and Albert R 1999 *Science* **286** 509
[8] Tanaka R, Yi T M and Doyle J 2005 *FEBS Lett.* **579** 5140
[9] Ravasz E, Somera A L, Mongru D A, Oltvai Z N and Barabasi A L 2002 *Science* **297** 1551
[10] Newman MEJ and Girvan M 2004 *Phys. Rev.* E **69** 026113
[11] Girvan M and Newman M E J 2002*Proc. Natl Acad Sci.* **99** 7821
[12] Holme P, Huss M and Jeong H 2003 *Bioinformatics* **19** 532
[13] Zhao J, Yu H, Luo J, Cao Z and Li Y 2006 *BMC Bioinformatics* **7** 386
[14] Zhao J, Ding G H, Tao L, Yu H, Yu Z H, Luo J H, Cao Z W and Li Y X 2007 *BMC Bioinformatics* **8** 311
[15] Watts D J and Strogatz S H 1998 *Nature* **393** 440
[16] Trusina A, Maslov S, Minnhagen P and Sneppen K 2004 *Phys. Rev. Lett.* **92** 178702
[17] Vazquez A, Pastor-Satorras R and Vespignani A 2002 *Phys. Rev.* E **65** 066130
[18] Newman MEJ 2002 *Phys. Rev. Lett.* **89** 208701
[19] Newman MEJ and Park J 2003 *Phys. Rev.* E **68** 036122
[20] Pastor-Satorras R, Vazquez A and Vespignani A 2001 *Phys. Rev. Lett.* **87** 258701
[21] Maslov S and Sneppen K 2002 *Science* **296** 910
[22] Huang Z X, Wang X R and Zhu H 2004 *Chin. Phys.* **13** 273
[23] Guimera R, Sales-Pardo M and Amaral LAN 2004 *Physical Review* E **70** 025101
[24] Ravasz E and Barabasi A L 2003 *Phys. Rev.* E **67** 026112
[25] Xu X P and Liu F 2007 *Chin. Phys.* **16** 282
[26] Chen F, Chen Z Q and Yuan Z Z 2007 *Chin. Phys.* **16** 287
[27] Dorogovtsev SN, Mendes JFF and Samukhin AN 2002 arXiv:cond-mat/0206131
[28] Huss M and Holme P 2007 *IET Syst. Biol.* **1** 280
[29] Rual J F, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, Berriz GF, Gibbons FD, Dreze M, Ayivi-Guedehoussou N *et al* 2005 *Nature* **437** 1173
[30] Li L, Alderson D, Doyle JC and Willinger W 2006 *Internet Math.* **2** 431
[31] Holme P and Zhao J 2007 *Phys. Rev.* E **75** 046111
[32] Burda Z, Erdmann J, Petersson B and Wattenberg M 2003 *Phys. Rev.* E **67** 026105
[33] Catanzaro M, Boguná M and Pastor-Satorras R 2005 *Phys. Rev.* E **71** 027103
[34] Boguna M, Pastor-Satorras R and Vespignani A 2004 *Eur. Phys. J.* B **38** 205
[35] Vazquez A, Boguna M, Moreno Y, Pastor-Santorras R and Vespignani A 2003 *Phys. Rev.* E **67** 046111
[36] Holme P 2007 *Physica A: Stat. Theor. Phys.* **377** 315
[37] Klemm K and Eguıluz VM 2002 *Phys. Rev.* E 036123
[38] Soffer SN and Váquez A 2005 *Phys. Rev.* E **71** 057101
[39] Guimera R and Amaral L A N 2005 *J. Stat. Mech.: Theory and Experiment* p02001
[40] Guimera R and Amaral L A N 2005 *Nature* **433** 895
[41] Whitney D E August 2005 *Working Paper Series.* (Massachusetts Institute of Technology, Engineering Systems Division)
[42] Li Y, Liu Y, Shan X M, Ren Y, Jiao J and Qiu B 2005 *Chin. Phys.* **14** 2153
[43] Hu K and Yi T 2006 *Chin. Phys.* **15** 2782
[44] Liu Z and Hu B 2005 *Europhys. Lett.* **72** 315
[45] Albert R, Jeong H and Barabasi A L 2000 *Nature* **406** 378
[46] Holme P, Kim BJ, Yoon CN and Han SK 2002 *Phys. Rev.* E **65** 056109
[47] Liu Z, Lai Y C and Ye N 2002 *Phys. Rev.* E **66** 036112
[48] Liu Z, Lai Y C, Ye N and Dasgupta P 2002 *Phys. Lett.* A **303** 337
[49] Váquez A and Moreno Y 2003 *Phys. Rev.* E **67** 015101