Community Structure in Small-world and Scale-free Networks

Haifeng Du
Institute for Population and Development Studies
School of Public Policy and Administration
Xi’an Jiaotong University
Xi’an, Shaanxi Province, 710049
China
and
Morrison Institute for Population and Resource Studies
Stanford University
Stanford, CA 94305
USA

Marcus W. Feldman*
Morrison Institute for Population and Resource Studies
Stanford University
Stanford, CA 94305
USA
Email: marc@charles.stanford.edu
Tel: (650) 723-7518
Fax: (650) 725-8244

Shuzhuo Li
Institute for Population and Development Studies
School of Public Policy and Administration
Xi’an Jiaotong University
Xi’an, Shaanxi Province, 710049
China

Xiaoyi Jin
Institute for Population and Development Studies
School of Public Policy and Administration
Xi’an Jiaotong University
Xi’an, Shaanxi Province, 710049
China
and
Morrison Institute for Population and Resource Studies
Stanford University
Stanford, CA 94305
USA

Zhongshan Yue
School of Management
Xi’an Jiaotong University
Xi’an, Shaanxi Province, 710049
China

August 15, 2006

* Corresponding authors.
Community Structure in Small-world and Scale-free Network

Abstract: The community structure of small-world networks and scale-free networks are explored in the framework of the modularity using computing simulations. We find: first, community structure in small-world networks is clear, but in scale-free networks it is not; second, community structure is affected by the density of the network, namely, the sparser the network connections, the clearer is the community structure; third, modularity in small-world networks is a dynamic variable which changes as a function of the probability of rewiring connections but not monotonically.

Key words: rural-urban migration, social network, small-world phenomena, scale-free network
LEGENDS FOR FIGURES 1-4

Figure 1. $L(p)/L(0)$ represented by $\rightarrow$, $C(p)/C(0)$ by $\Box$ and $Q(p)/Q(0)$ represented by $\uparrow$. These ratios are plotted for the values of rewiring probability $p$ shown.

Figure 2. The network with 60 nodes, namely $n=60$, represented by $\circ$; $n=80$ by $\checkmark$; $n=100$ by $\ast$; $n=120$ by $\times$; $n=140$ by $\ast$; and $n=160$ by $\Delta$. Modularity is shown as a function of rewiring probability $p$ for different values of $k$.

Figure 3. $L(m)/L(2)$ represented by $\rightarrow$, $C(m)/C(2)$ by $\Box$, $Q(m)/Q(2)$ represented by $\uparrow$, $L_r(m)/L(2)$ represented by $\cdots\cdots$, $C_r(m)/C(2)$ by $\cdots\cdots$ and $Q_r(m)/Q(2)$ represented by $\cdots\cdots$. These ratios are plotted for the values of $n$ shown.

Figure 4. $m=4$ represented by $\Box$, $m=6$ by $\Delta$, $m=8$ by $\checkmark$, $m=10$ by $\rightarrow$, $m=12$ by $\leftarrow$ and $m=14$ by $\uparrow$. Modularity is shown as a function of network size $n$ for different values of $m$.

Figure 5 legend
1. Introduction

Networks consist of a set of nodes connected by ties indicating interaction. Network analysis “makes it possible to bridge the ‘micro-macro gap’, the theoretical gulf between microsociology, which examines the interaction of individuals, and macrosociology, which studies the interaction of groups or institutions.” (Mustafa and Jeff, 1994). Mathematicians and physicists have found a number of distinctive structural properties in “real-world” networks, and study of complex networks is active in many fields (Wasserman and Faust 1994; Scott, 2002; Dorogovtsev and Mendes, 2001; 2002; Albert and Barabási, 2002; Watts et al., 2002). The most famous characteristics of the complex networks (Xiao and Chen, 2003; Zhou et al., 2005) are “small-world phenomena” (Watts and Strogatz, 1998; Dorogovtsev and Mendes, 2002) and “scale-free properties” (Barabási et al., 1999; Barabási, 2002).

The small world phenomenon produces effective diffusion of information in a society in that every node can be reached through a short chain of acquaintances (Kleinberg, 2000). Compared with regular and random networks, a small-world network consists of many local links and fewer long range “shortcuts”. In spite of the lack of an accurate definition, a network is usually regarded as having the small-world property if the average path length is small and the average clustering coefficient is big (Watts and Strogatz, 1998). On the other hand, the scale-free property reflects that the network’s degree distribution obeys a power law (Barabási et al., 1999; Albert and Barabási, 2002); that is, a few nodes with high degree act as hubs while the degree of most nodes is low. The main difference between a Poisson (or normal) and power-law distribution is in their tails. Unlike Poisson and normal distributions, whose tails decay exponentially, power-law degree distributions decay slower, probably as a result of the existence of hubs.
It has been suggested that the scale-free network is one type of small-world network (Amaral et al., 2000).

Further studies have revealed that many actual networks are highly inhomogeneous (Strogatz, 2001); they do not consist of an undifferentiated mass of nodes but have some subgroup structure (Flake et al., 2002; Girvan and Newman, 2002; Holme et al., 2003; Guimerá and Amaral, 2005; Mercedes and Jennifer, 2006). A network can often be divided into subgroups, within which the network connections are dense, but between which they are more sparse (Newman, 2004a, 2004b). This seems to be a common feature of many networks, especially social networks (Luo, 2005). Each of these subgroups is called a “community”, and the complete structure of these subgroups within a network is referred to as “community structure” (Newman, 2004a), which has been identified as one of the main characteristics of complex networks (Wang and Dai, 2005).

In order to evaluate community structure, several measures which are based on density have been proposed (Wasserman and Faust, 1994). Recently, Newman and Girvan (2003) proposed a new measure they call “modularity”, which is a numerical index of how good a particular partition is, and overcomes limitations of previous measures. In contrast to most other measures, it not only directly measures the quality of a particular clustering of nodes in a network, but also can be used to automatically select the optimal number of clusters (Scott and Smyth, 2005).

Most recent studies of small-world phenomena and scale-free properties focus on revealing the corresponding characteristics in “real-world” networks or exploring the relevant complex network models (Barabási and Albert, 1999; Davidsen et al., 2002; Konstantin and Vi´ctor, 2002; Latora1 and Marchiori, 2003). Comparisons between
network structures can be facilitated by studying community structure (Newman, 2002; Radicchi et al., 2004; Wu and Huberman, 2004; Newman, 2003a, 2003b; Newman and Girvan, 2003, 2004; Newman, 2004a, 2004b; Aaron et al., 2004; Scott and Smyth, 2005). Although community structure has attracted some attention (Wasserman and Faust 1994; Scott, 2002; Newman, 2004a; Wang and Dai, 2005), there are few studies that focus on comparing community structure of small-world networks and scale-free networks.

The main purposes of this paper are to detect the community structure in Watts and Strogatz’s small-world network and Barabási’s scale-free network, and to compare these two standard complex networks from the point of view of the community structure. In section 2, our analysis strategy will be briefly introduced. Results of simulation experiments will be shown in section 3. We conclude with some discussion of our results in section 4.

2. Analytic Strategy

Although more complex models have been developed, Watts and Strogatz’s small-world network and Barabási’s scale-free network, which will be denoted as W-S and B-A in the remainder of this note, are still the most common models for complex networks. These two types of networks are the focus of this paper (The W-S and B-A models are described in Appendices A and B). The main structural parameters of W-S include the size $n$, the number of nearest neighbors $k$ and the rewiring probability $p$. For the B-A model, starting size $m_0$, $m$ ($\leq m_0$) edges for a new additional node and total time steps $t$ are the main parameters, and the B-A network size is $n=m_0+t$.

Despite the many parameters that could be involved in social network analysis, such as degree, centrality, centralization transitivity, reciprocity, etc. (Wasserman and Faust,
complex network analysis generally focuses on the clustering coefficient $C$, the average path length $L$ and the degree distribution. The clustering coefficient $C$ and average path length $L$ are the most important properties in assessing small-world phenomena (Watts and Strogatz, 1998) and the degree distribution is used in describing the scale-free property. It is difficult to judge a power-law distribution statistically and the conventional strategy is to draw a simple histogram of the degree distribution and plot it on a log scale to see if it produces a straight line (Newman, 2005). The degree distribution will not be discussed in the following. Instead we focus on the modularity $Q$ together with $C$ and $L$ (These three statistics are described in Appendix C).

To explore community structure in small-world networks, we simulate many W-S networks with different $p$, $k$ and $n$. $L$, $C$ and $Q$ are calculated for each network. Then we examine that how the average values of $L$, $C$ and $Q$ change as $p$, $k$ and $n$ change separately. In this way, we can assess the main effects of these parameters on a small-world network’s community structure. Our strategy for studying the community structure of scale-free networks is similar to that for the small-world networks, but $m$ and $n$ are the two major parameters considered for B-A. Since $m_0<<n$ and $m_0$ has little effect on the B-A structure, we do not consider the role of $m_0$ in community structure.

$L$ and $C$ cannot be calculated strictly according to the theoretical formula because the real networks are usually sparse and have a few isolated nodes. We only calculate $C$ for nodes with more than 1 neighbor (namely $n_v > 1$, where $n_v$ is the number of elements in the set of nodes directly connected to node $v$) and $L$ for nodes possessing short path length (namely $\ell(i,j) \leq n-1$, where $\ell(i,j)$ is the minimum distance between nodes $i$ and $j$). Our results are in agreement with Ucinet software. $Q$ is computed according to the
algorithm of Aaron et al (2004). All analyses were carried out using Matlab.

3. Result

3.1 Community structure in the small-world network

For each rewiring probability $p$, we construct 100 networks according to W-S model and denote the average clustering coefficient $C$ as $C(p)$, the average path length $L$ as $L(p)$, and the modularity $Q$ as $Q(p)$. Hence, $C(0)$, $L(0)$ and $Q(0)$ are the parameters for the regular lattice network and $C(1)$ , $L(1)$, and $Q(1)$ are those for the random network.

When $n=100$ and $k=12$, figure 1 shows $L(p)$, $C(p)$ and $Q(p)$ for the family of randomly rewired graphs described in Appendix A. For convenience, the three functions are normalized by the values $L(0)$, $C(0)$ and $Q(0)$, respectively.

**Figure 1**

From figure 1, we see that $L$ and $C$ decrease monotonically as the rewiring connection probability $p$ increases. This is one of the small-world phenomena, and is intermediate between properties of regular lattices ($p=0$) and random networks ($p=1$). The change in modularity $Q$ is very different from that of $L$ and $C$. For a fixed value of $k$ and network size $n$ varying from 60 to 160 in increments of 20, Figure 2 shows the modularity $Q$ for the family of W-S model.

**Figure 2**

Analysis of figure 2 shows:

(1) For different $k$ and $n$, there is clearly a peak of modularity $Q$ as a function of the rewiring probability $p$. Generally, When $0.01 \leq p \leq 0.1$, the $Q$ value for W-S is greater than that of the corresponding regular lattice network ($p=0$) and the random network ($p=1$). The modularity $Q$ increases with network size $n$, but is not strongly affected by $k$,
the number of nearest neighbors.

(2) The density of the network is \( \frac{nk}{\binom{n}{2}} = \frac{2k}{n-1} \). For \( k \) fixed, the network becomes sparser as \( n \) is increased. Hence, the sparser the connections in W-S, the greater the value of \( Q \) and the clearer the community structure.

### 3.2. Community structure in the scale-free network

In figure 3, for different network sizes, the average path length, average clustering coefficient and modularity are compared between B-A and the corresponding random network, which has the same size and ties. For each network with the new additional number of edges \( m \), we denote its average clustering coefficient \( C \) as \( C(m) \), average path length \( L \) as \( L(m) \) and modularity \( Q \) as \( Q(m) \), with, \( C_r(m) \), \( L_r(m) \) and \( Q_r(m) \) for the corresponding random network. Actually, we can conceive of the B-A model with \( m=1 \), but the clustering coefficient would then be almost 0. We therefore recalculated the statistics for the B-A model for \( m=2 \). For convenient plotting, the three parameters are normalized by the values \( L(2) \), \( C(2) \) and \( Q(2) \) separately. Figure 3 shows how the standardized values of \( C \), \( L \) and \( Q \) vary as functions of \( m \). In figure 3 the network size \( n \) is fixed at 100, 150, 200, 250, 300, 350, 400, 450, 500 and 550, and \( m \) is simulated from 2 to 14 in increments of 1. The figure 4 shows the modularity \( Q \) for the family of networks created by the B-A model. Each data point in figure 3 and 4 is an average over 20 random realizations obeying the B-A model or the random network model.

### Figure 3

### Figure 4

From figures 3 and 4, we find:
(1) Compared with the corresponding random network, B-A has a similar average path length $L$ and a greater average clustering coefficient $C$, and most of the $C$ values of B-A are about double those of the random network. Hence, B-A networks display small-world phenomena in a “weak” sense.

(2) The modularity $Q$ for B-A is similar to that of the random networks, and is less than 0.3 (shown in figure 4), indicating that there is no clear community structure in B-A networks.

(3) $m$ is the key parameter for B-A. Among the B-A family with different sizes, $C$ increases while $L$ and $Q$ decrease with increasing $m$.

(4) Modularity $Q$ is also affected by the density of the network. For each curve in figure 4, the B-A network becomes sparser with $n$ increasing, while the modularity $Q$ increases.

4. Discussion

Our main findings are the following:

(1). Community structure is the basic characteristic of the networks considered here. Compared with the random network, the regular lattice network and W-S have clear community structures, while B-A does not. The rewiring probability $p$ of W-S has a remarkably strong effect on community structure. Our simulations reveal that the community structure in W-S is clearer than that for the corresponding regular lattices network when $0.01 \lesssim p \lesssim 0.1$. Cowan et al (2004) found that the long-run average level of knowledge diffusing through a W-S network is also a non-monotonic function of the rewiring probability $p$, with its absolute peak corresponding to values of $p$ around 0.06, which are clearly in the small-world region (between $p = 0.01$ and $p = 0.1$ approximately;
Morone and Taylor, 2004). Our findings are in agreement with these. The criterion for small-world phenomena suggested by Watts and Strogatz (1998) is qualitative. Using community structure, however, we obtain a more quantitative criterion, namely, a network with small-world phenomena should have a large average clustering coefficient, small average path length and clear community structure.

The community structure of the B-A network is clearly different from that of the W-S network. Table 1 summarizes the main differences in the properties of the different types of networks.

Table 1

(2). The density of the network affects the community structure. Our simulations suggest that a sparse network may have clear community structure. The structure of networks can vary a great deal and there are many different types of community structure for a network, so that detecting community structure is difficult. Hence, theoretical analysis for the relationship between the density and community structure is impossible. However, the situation for regular networks is comparatively simple. For a regular lattice network (as described in Appendix A) with $t$ communities, the modularity $Q$ is

$$Q = \frac{nk - tk'(k' + 1)}{nk} - \sum_{i=1}^{t} (m_i)^2, \quad (1)$$

where $\sum_{i=1}^{t} m_i = 1$, $k' = \frac{k}{2}$, and when $m_1 = m_2 = \cdots m_t = \frac{1}{t}$,

$$Q = 1 - \left( \frac{k + 2}{4n} t + \frac{1}{t} \right). \quad (2)$$

If $\frac{k}{n}$ is small, the network is sparse, and $Q$ becomes larger. Obviously, $t$ is the other
major contribution to $Q$. Furthermore, if we take $t = \left\lfloor \frac{4n}{k+2} \right\rfloor$ or $t = \left\lceil \frac{4n}{k+2} \right\rceil$ according to which gives the greater $Q$, we have

$$Q_{\text{max}} \approx 1 - \frac{\sqrt{\frac{k+2}{n}}}{n},$$

(3)

where $\lfloor x \rfloor$ rounds $x$ to the nearest larger integer and $\lceil x \rceil$ rounds it towards the nearest lower integer.

If $k \leq 0.49n - 2$, then $Q_{\text{max}} \geq 0.3$, and the regular lattice network has clear community structure. Obviously, this condition is very easily satisfied for regular lattice networks and W-S is constructed from the regular lattice network by rewiring a few connections. Hence, if the rewiring probability $p$ is small, the regular structure of the lattices is retained in W-S. With a few random long range “shortcuts”, the community structure may be clearer than for the regular lattice network. If $p$ is large, W-S looks more like a random network and the regular structure of the lattices is not seen in W-S with the result that the community structure is less noticeable. For the B-A network, the connections becomes denser with increasing $m$ but the connection between each pair of nodes is random because of preferential attachment. Consequently, there is no clear community structure in B-A network.

(3). It may be not accurate to claim that the scale-free network is one of the classes of small-world networks. Compared to the corresponding random network, B-A has a larger average clustering coefficient and similar average path length. Hence, the scale-free network might be regarded as one class of small-world networks (Amaral et al., 2000). In our simulations the average clustering coefficient of the B-A network is no more than 3 times that of the random network so it does not satisfy the criterion
“$C >> C_{\text{rand}}$” suggested by Watts and Strogatz (1998), where $C_{\text{rand}}$ is the average clustering coefficient of the random network. From the point view of community structure, the B-A network is closer to the random network than the small-world network. Together with their difference in degree distributions, the preceding remarks lead us to believe that small-world network and scale-free network should be regarded as two classes of complex networks that should be evaluated separately.

(4). Community structure vs. small-world phenomena, and community structure vs. scale-free properties. Many local links and few long range “shortcuts” are structural characteristics in networks with small-world phenomena. The “local links” may be considered as connections among nodes in the same community, and “long range ‘shortcuts’” may be connections among different communities. On the other hand, a network with clear community structure may exhibit small-world phenomena. For example, we can form a network as follows:

$s$ sub-networks with $\{n_1,n_2,\ldots,n_s\}$ nodes, respectively, are generated. For network $i$, the connection probability is $p_i$ for each pair of nodes, where $i = 1,2,\ldots,s$, and $p_c$ for a connection between nodes in any two sub-networks. These sub-networks can be combined to form a new larger network.

If $p_i >> p_c, \ i = 1,2,\ldots,s$, then the whole network has an obvious community structure (Du, et al., 2006) and we denote such a network as R-R. For the following three cases, we calculate $C, L$ and $Q$ for R-R networks and the corresponding random network and list them separately in Table 2.

Case 1: $n_1 = n_2 = \cdots = n_{15} = 10$; $p_1 = p_2 = \cdots = p_{15} = 0.5; p_c = 0.01$

Case 2: $n_1 = n_2 = \cdots = n_{15} = 10$; $p_1 = p_2 = \cdots = p_7 = 0.6$;
\[ p_8 = p_9 = \cdots = p_{15} = 0.3; \quad p_c = 0.01 \]

Case 3: \( n_1 = n_2 = n_3 = 10; \quad n_4 = n_5 = \cdots = n_{10} = 6; \quad n_{11} = n_{12} = \cdots = n_{15} = 15; \)

\[ p_1 = p_2 = p_3 = 0.4; \quad p_4 = p_5 = \cdots = p_{10} = 0.5; \quad p_{11} = p_{12} = \cdots = p_{15} = 0.3; \]

\[ p_c = 0.01 \]

**Table 2**

From table 2, R-R networks not only have clear community structure, but also exhibit small-world phenomena, although the R-R network, which obeys the definition of community structure, namely, within subgroups the network connections are dense, but between them they are sparse, is very different from W-S. This illustrates the relationship between small-world phenomena and community structure.

The relationship between community structure and scale-free properties is difficult to describe. On the one hand, the scale-free property suggests that the degree distribution of a network should obey a power law, and a few nodes named “hubs” control the whole network. Thus, community structure may not be obvious because most of the nodes should be connected to the “hubs” simultaneously. On the other hand, community structure provides a bridge between the single node and the whole network, because there are few connections between these sub-networks.

Real-world social networks such as those of the rural-urban migrants’ social networks obtained from the Shenzhen survey (Du, et al., 2006; Li, et al. 2006) are sparse and complex. These real social networks exhibit clear community structure, small-world phenomena and scale-free properties simultaneously. None of the current network models, such as the random network, W-S, or B-A model, fits these social networks very well.
new social network structure may be necessary.
Appendix A

Watts and Strogatz Small-World Network Model

Figure 5 shows how to construct the network: we start with a ring of \( n \) nodes, each connected to its \( k \) nearest neighbors by directed edges, making a regular network as in figure 5 (a). With probability \( p \), we reconnect each edge to a destination node over the entire ring, with duplicate edges forbidden; the small-world networks shown in figure 5 (b) will be obtained for \( p=0.1 \). For \( p=1 \) the network will be a random network, as in figure 5 (c) shown. According to Watts and Strogatz’s small-world network model, we require \( n \gg k \gg \ln n \gg 1 \), where \( k \gg \ln n \) guarantees that a random graph will be connected. Here, \( n=20 \) and \( k=4 \) in figure 5(a).
Appendix B

Barabási Scale-Free Network Model

A network with a power-law degree distribution can be constructed using the Barabási and Albert model (Barabási et al., 1999) in two steps:

(a) **Growth**: Starting with a small number \( m_0 \) of nodes, at every time step we add a new node with \( m \leq m_0 \) edges (i.e., a new individual, who will be connected to \( m \) of the nodes already present in the system).

(b) ** Preferential attachment**: When choosing the nodes to which the new node connects, we assume that the probability \( \Pi \) that a new node will be connected to node \( i \) depends on the connectivity \( k_i \) of that node:

\[
\Pi(k_i) = k_i \sum_j k_j .
\]

The time dependence of the connectivity of a given node can be calculated analytically using a mean-field approach. The degree distribution of the network obeys a power-law, namely,

\[
P(k) = \frac{2m^2}{m_0 + t} \frac{1}{k^3}.
\]

Growth and preferential attachment are essential ingredients of the BA model (Barabási, et al., 1999). A network with a power law degree distribution is scale-free (Barabási, 2002).
Appendix C

Parameters

A social network may be defined as $G = (V, E)$, where $V$ is the set of nodes or individuals in the network and $E$ is the set of edges, namely the relationships between each pair of nodes.

The clustering coefficient, $C$, represents the degree of local clustering, namely, the probability that a pair of nodes connected to a common node are also connected to each other. Suppose there are $n$ nodes. For a node $v$, $V_v$ is the set of nodes directly connected to $v$, and $n_v$ is the number of elements in $V_v$. The clustering coefficient for node $v$ may be defined as:

$$c_v = \frac{\sum_{l \in E, x \in V_v} \delta_l x \delta_l y}{\frac{n_v(n_v - 1)}{2}},$$

(6)

where the edge $l$ contains node $x$, then $\delta_l x = 1$. Otherwise $\delta_l x = 0$. The average clustering coefficient of the network is therefore

$$C = \frac{1}{n} \sum_{v=1}^{n} c_v.$$

(7)

The greater the value of $C$, the more local clustering occurs in the network.

The average path length, $L$, measures the efficiency of communication or passage time between nodes, namely, the average number of links in the shortest path between a pair of nodes in the network. If $l(i, j)$ is the minimum distance between nodes $i$ and $j$, $L$ can be defined as:
The smaller the value of the average path length $L$, the more efficient the communication through the network.

For a partition of a network into $t$ communities, we define a $t \times t$ matrix $E$ whose entry $e_{pq}$ is the fraction of edges in the original network that connect vertices in community $p$ to those in community $q$.

$$e_{pq} = \begin{cases} \frac{1}{2} \sum_{i \in V_p, j \in V_q} a'_{ij} & p \neq q \\ \frac{1}{2} \sum_{i \in V_p, j \in V_p} a'_{ij} & p = q \end{cases}$$

Here $V$ represents the set of nodes of the whole network and $V_p$ is the set of nodes of the community $p$. $a'_{ij}$ and $a_{ij}$ are entries of the adjacent matrix, generally,

$$a_{ij} = \begin{cases} 1 & \text{if node } i \text{ connects with } j \\ 0 & \text{if node } i \text{ unconnects with } j \end{cases}$$

The modularity is defined as (Girvan and Newman, 2002):

$$Q = \sum_{p=1}^{t} \left[ e_{pp} - \left( \sum_{q=1}^{t} e_{pq} \right)^2 \right].$$

The greater the value of $Q$, the stronger is the community structure of the network (Newman, 2004a, 2004b). In practice, $Q$ values for networks with strong community structure typically fall in the range from about 0.3 to 0.7 (Newman and Girvan, 2004).
Acknowledgements. This work is jointly supported by National Social Science Foundation of China Grant (05ARK001 and 05CRK002), Program for New Century Excellent Talents in University (NCET-04-0931), Morrison Institute for Population and Resource Studies at Stanford University and Santa Fe Institute International Program.

References

Aaron C., M. E. J. Newman, M. Cristopher. 2004. “Finding community structure in very large networks”. Phys. Rev. E. 70, 066111.

Albert R. and Barabási A.L., 2002. “Statistical mechanics of complex networks”. Reviews of Modern Physics. 74: 48-97.

Amaral L. A. N., A. Scala, M. Barthélémy, and H. E. Stanley. 2000. “Classes of small-world networks”. Proc. Natl. Acad. Sci. USA, 97(21): 11149-11152

Barabási A.L. and R. Albert. 1999. “Emergence of scaling in random networks”. Science. 286: 509-512

Barabási A.L., R. Albert, H. Jeong. 1999. “Mean-field theory for scale-free random networks”. Physic A. 272: 173-189

Cowan R., N. Jonard, M. Ozman. 2004. “Network structure and the diffusion of knowledge”. Journal of Economic Dynamics and Control. 28: 1557-1575.

Davidsen J., E. Holger, B. Stefan. 2002. “Emergence of a small world from local interactions: Modeling acquaintance networks”. Phys. Rev. Lett. 88, 128701

Dorogovtsev S.N. and J.F. Mendes. 2001. “Scaling properties of scale-free evolving networks: Continuous approach”. Phys. Rev. E. 63, 056125

Dorogovtsev S.N. and J.F. Mendes. 2002. “Evolution of networks”. Adv. Phys. 51:
1079-1187.

Du H.F., M. W. Feldman, S. Li, et al. 2006. An Algorithm for Detecting Community Structure of Social Networks Based on Prior Knowledge and Modularity (Manuscript)

Flake G. W., S. R. Lawrence, C. L. Giles, and F. M. Coetzee. 2002. “Self-organization and identification of Web communities”. *IEEE Computer*. 35: 66-71.

Girvan M. and M. E. J. Newman. 2002. “Community structure in social and biological networks”. *Proc. Natl. Acad. Sci. USA* 99, 7821-7826.

Guimerà R. and L. A. N. Amaral. 2005. “Functional cartography of complex metabolic networks”. *Nature*. 433: 895-900.

Holme P., M. Huss, H. Jeong. 2003. “Subnetwork hierarchies of biochemical pathways”. *Bioinformatics*. 19: 532-538.

Institute for Population and Development Studies at Xi’an Jiaotong University. 2005. Report on the survey of rural-urban migrants in Shenzhen, China. (In Chinese)

Kleinberg J. 2000. “Navigation in a small world.” *Nature*. 406: 845.

Konstantin K. and M. E. Víctor. 2002. “Growing scale-free networks with small-world behavior”. *Physical Review E*. 65 (5): 057102

Latora1 V. and M. Marchiori. 2003. “Economic small-world behavior in weighted networks”. *The European Physical Journal B*. 32, 249-263

Li S., Y. Xu, X. Jin, et al. 2006. Complexity of rural-urban migrants’ social networks in China: A quantitative study in Shenzhen (Manuscript)

Luo J. 2005. *Social Network Analysis*. Social Science Academic Press. (In Chinese).

Mercedes P. and A. D. Jennifer. 2006. *Ecological Networks: Linking Structure to*
Morone P. and R. Taylor. 2004. “Knowledge diffusion dynamics and network properties of face-to-face interactions”. *Journal of Evolutionary Economics*. 14: 327-351.

Mustafa E. and G.Jeff. 1994. “Network analysis, culture, and the problem of agency”. *The American Journal of Sociology* 99(6):1411-1454.

Newman M. E. J. and M. Girvan. 2003. “Mixing patterns and community structure in networks”, Pastor-Satorras, J. Rubi, and A. Diaz-Guilera (eds.), *Statistical Mechanics of Complex Networks*, R. Springer, Berlin.

Newman M. E. J. and M. Girvan. 2004. “Finding and evaluating community structure in networks”. *Phys. Rev. E*. 69, 026113.

Newman M E J. 2002. “Assortative mixing in networks”. *Phys. Rev. Lett*. 89, 208701

Newman M E J. 2003 a. “Mixing patterns in networks”. *Phys. Rev. E*. 67, 026126

Newman M E J. 2003 b. “The structure and function of complex networks”. *SIAM Review*. 45(2):167-256.

Newman M. E. J. 2004a. “Detecting community structure in networks”. *Eur. Phys. J. B*. 38, 321-330.

Newman M. E. J. 2004b. “Fast algorithm for detecting community structure in networks”. *Phys. Rev. E*. 69, 066133.

Newman M E J. 2005. “Power laws, Pareto distributions and Zipf's law”. *Contemporary Physics*. 46: 323-351.

Newman M E J. and P. Juyong. 2003. “Why social networks are different from other types of networks”. *Phys. Rev. E*. 68, 036122.

Radicchi F., C. Castellano, F. Cecconi, V. Loreto, D. Parisi. 2004. “Defining and
identifying communities in networks”. Proc. Natl. Acad. Sci. USA 101: 2658-2663.

Scott J. 2002. Social Network Analysis: A Handbook. Sage, London, 2nd edition.

Scott W. and P. Smyth. 2005. “A spectral clustering approach to finding communities in graphs”. SIAM International Conference on Data Mining, Newport Beach, CA, 2005.

Strogatz S. H. 2001.” Exploring complex networks”. Nature. 410: 268-276

Wang L. and G. Dai . 2005. “Community finding in complex networks: Theory and applications”. Science & Technology Review. 23(8):62-66. (In Chinese)

Wasserman S. and K. Faust. 1994. Social Network Analysis: Methods and Applications. New York and. Cambridge, ENG: Cambridge University Press.

Watts D.J. and S. H. Strogatz. 1998. “Collective dynamics of 'small-world' networks”.

Nature. 393(4): 440-442.

Watts D. J., P. S. Dodds, M. E. J. Newman. 2002. “Identity and search in social networks”. Science. 296: 1302-1305

Wu F. and B. A. Huberman. 2004. “Finding communities in linear time: A physics approach”. The European Physics Journal B. 38(10):331-338.

Xiao F. W. and G. R. Chen. 2003. “Complex Networks: Small-World, Scale-Free and Beyond”. IEEE circuits and systems magazine. 3(1): 6-20.

Zhou T., Bai W., Wang B., et al. 2005. “A brief review of complex network studies.”

Physics. 34(1): 31-36. (In Chinese).
Fig. 1. Three statistics as a function of rewiring probability $p$
Fig. 2. Modularity as a function of rewiring probability $p$

(e) $k=14$
Fig. 3. Three statistics as functions of $m$ for different values of $n$
Fig. 4. Modularity as a function of $n$ for different values of $m$
Increasing randomness

(a) Regular lattice network    (b) Small-world network      (c) Random network

Fig. 5. Construction process of network models
Table 1. Comparison of different networks

| Networks             | Parameters | Average path length | Clustering coefficient | Modularity | Degree distribution |
|----------------------|-----------|---------------------|------------------------|------------|--------------------|
| Regular Network      | Big       | Big                 | Big                    | Big        | \                 |
| Random Network       | Small     | Small               | Small                  | Small      | Poisson distribution |
| Small World Network  | Small     | Big                 | Big                    | Big        | undetermined       |
| Scale Free Network   | Small     | Big                 | Small                  |            | Power law distribution |
Table 2. R-R networks compared to the corresponding random network*

| Networks          | Parameters | Degree | Average path length | Clustering coefficient | Modularity |
|-------------------|------------|--------|---------------------|------------------------|------------|
| Case 1            | R-R network| 4.6267 | 2.7523              | 0.4015                 | 0.6150     |
|                   | Random network| 4.6267 | 2.4815              | 0.0659                 | 0.2700     |
| Case 2            | R-R network| 4.3067 | 2.8194              | 0.3675                 | 0.6280     |
|                   | Random network| 4.3067 | 2.5502              | 0.0635                 | 0.2896     |
| Case 3            | R-R network| 4.1905 | 2.8288              | 0.2907                 | 0.5696     |
|                   | Random network| 4.1905 | 2.5681              | 0.0550                 | 0.2841     |

* By chance, the degree of R-R network is not the same for each simulation. In order to compare with the corresponding random network, the results in table 2 are obtained from one simulation for each case. Similar results are obtained acquired for the other simulations.