Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Physica A 387 (2008) 3769–3780

PHYSICA A

www.elsevier.com/locate/physa

Unified index to quantifying heterogeneity of complex networks

Hai-Bo Hu, Xiao-Fan Wang*

Complex Networks and Control Lab, Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China

Received 13 September 2007; received in revised form 18 January 2008
Available online 9 February 2008

Abstract

Although recent studies have revealed that degree heterogeneity of a complex network has significant impact on the network performance and function, a unified definition of the heterogeneity of a network with any degree distribution is absent. In this paper, we define a heterogeneity index $0 \leq H < 1$ to quantify the degree heterogeneity of any given network. We analytically show the existence of an upper bound of $H = 0.5$ for exponential networks, thus explain why exponential networks are homogeneous. On the other hand, we also analytically show that the heterogeneity index of an infinite power law network is between 1 and 0.5 if and only if its degree exponent is between 2 and 2.5. We further show that for any power law network with a degree exponent greater than 2.5, there always exists an exponential network such that both networks have the same heterogeneity index. This may help to explain why 2.5 is a critical degree exponent for some dynamic behaviors on power law networks.

PACS: 89.75.-k; 89.75.Hc; 89.75.Fb

Keywords: Heterogeneity; Complex networks; Power law networks; Exponential networks

1. Introduction

In the last decade, researchers from different disciplines have made great and encouraging progresses in many fields of complex networks [1], including topological properties [1], network modeling [2], attack and error tolerance [3], dynamical processes taking place on networks [4,5] and network synchronization [6], of which the study of network structure possesses fundamental importance [1]. To a great extent, the structure of a network affects its function and behavior [7]. For instance, the topologies of social networks and the Internet affect the spread of information, gossip, infectious diseases or computer viruses [4,8], and the topology of the power grid influences the robustness and stability of power transmission [9].

Until now, two classes of networks have attracted intensive interests. The first class of networks is the exponential ones. The degree (or connectivity) distribution of an exponential network peaks at an average value and decays exponentially or faster. Such networks have also been widely called homogeneous networks, due to the fact that each node has about the same number of links. Typical exponential network models include the classical ER random graph investigated by Erdös and Rényi in 1960s [10] and the small-world models, such as the WS small-world

* Corresponding author. Tel.: +86 21 34204187.
E-mail address: xfwang@sjtu.edu.cn (X.-F. Wang).

0378-4371/$ - see front matter
doi:10.1016/j.physa.2008.01.113

model proposed by Watts and Strogatz in 1998 [11] and NW small-world model proposed by Newman and Watts in 1999 [12]. Another significant discovery in the field of complex networks is the observation that many real-world large-scale complex networks are scale-free, that is, their degree distributions have the power law form $P(k) \sim k^{-\gamma}$, where $P(k)$ is the fraction of nodes with k links and γ is the degree exponent [1]. Differing from an exponential network, a power law network with a degree exponent between 2 and 3 has been widely regarded as inhomogeneous (or heterogeneous): most nodes have very few links and yet a few nodes, called hubs, have many links. A widely investigated scale-free network model, BA model, was proposed by Barabási and Albert [13], which has a degree exponent 3. However, empirical studies have shown that the degree exponents of most real-world power law networks are between 2 and 2.5 [14,15]. Recent studies have also revealed that heterogeneity of scale-free networks has significant impact on network performance, such as robustness and attack tolerance [3], epidemic dynamics [4,16], evolutionary dynamics [17], traffic dynamics [18] and synchronizability [19].

However, there also exist many real-world networks whose degree distributions may not be well described by exponential or power law distributions. For instance, the degree distribution of a broad-scale network has a power law regime followed by a sharp cutoff like an exponential decay of the tail [20], examples of which include the movie–actor network [20], and the co-authorship networks in Medline, Physics E-print Archive [21], and Los Alamos Archive [22]. Therefore, *we need a unified index to quantify the heterogeneity of a network with any degree distribution, and to quantitatively compare the heterogeneity of networks with different types of degree distributions*.

As a measure of inequality of a distribution, the Gini coefficient has been widely used in economics, ecology, physics and sociology to describe inequality in wealth, size or other quantities [23–26]. We find that the Gini coefficient of the degree distribution of a network serves as a good heterogeneity index of the network, and therefore is called the heterogeneity index of the network. The heterogeneity index of a completely homogeneous network is 0; however, the heterogeneity index of a completely heterogeneous network will approach 1.

We analytically show the existence of an upper bound of heterogeneity index 0.5 for exponential networks, thus explain why exponential networks are homogeneous. On the other hand, we also analytically show that the heterogeneity index of an infinite power law network is greater than 0.5 only if its degree exponent $\gamma \in (1, 2.5)$ and approaches 1 if $\gamma \in (1, 2]$. We further prove that for any power law network with a degree exponent $\gamma > 2.5$, there always exists an exponential network such that both networks have the same heterogeneity index. This may help to explain why 2.5 is a critical degree exponent for some dynamic behaviors on power law networks [14,15]. Since BA model is a power law network with degree exponent 3, it may not be a good model for heterogeneous scale-free networks.

We compute the heterogeneity indices of several real-world large-scale scale-free networks, including the co-authorship network, the Internet, the WWW and the protein interaction network (PIN). We find that, except for the PIN, all other networks are heterogeneous with heterogeneity indices larger than 0.5. We also compute the heterogeneity indices of several real large-scale exponential networks, including the power grids of the western United States, China railway network and India railway network, and find that their heterogeneity indices are all less than 0.5 and therefore these networks are indeed homogeneous.

The paper is organized as follows. Section 2 introduces the definition of heterogeneity index of a network, and shows how it relates to the degree sequence or degree distribution of the network. Heterogeneity indices of several regular networks are presented. In Section 3, we present the analytical results for the heterogeneity of power law networks. Section 4 is devoted to the study of heterogeneity of exponential networks, including the ER random graph, the WS and NW small-world models, and a general exponential network model. In Section 5 we give our conclusions and point out directions for future research.

2. Heterogeneity index

2.1. Definition

Heterogeneity index of a network is defined according to Lorenz curve and Gini coefficient originally used in microeconomics to describe income inequality [23]. The Lorenz curve was developed by the American economist Max O. Lorenz in 1905 and is a graphical representation of income inequality, which shows, for the cumulative percentage $x\%$ of population (plotted on the x-axis) arranged from poorest to richest, their cumulative percentage $y\%$ of the total income (plotted on the y-axis). A perfectly equal income distribution in a society would be one in which every person has the same income. It can be depicted by the straight line $y = x$, the line of perfect equality.

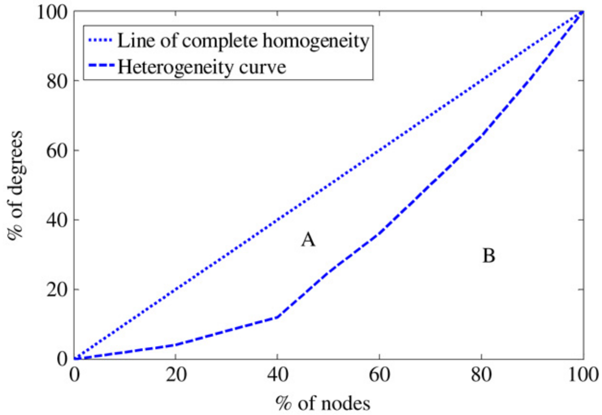

Fig. 1. Heterogeneity curve (dashed line) of a network. Dotted line is the line of complete homogeneity.

A perfectly inequal distribution, by contrast, would be one in which one person has all the income and everyone else has none. In that case, the curve would be at $y = 0$ for all $x < 100$, and $y = 100$ when $x = 100$, the line of perfect inequality. The Lorenz curve can be used to calculate the Gini coefficient, which was developed by the Italian statistician Corrado Gini in 1912 and is a ratio of the area between the line of perfect equality and the Lorenz curve, to the area between the line of perfect equality and the line of perfect inequality. Gini coefficient can quantify income inequality. With larger Gini coefficient comes greater income inequality and *vice versa*. Besides the application in microeconomics, Gini coefficient has also been widely used in ecology, physics and sociology to describe inequality in size or other quantities [24–26]. It is clear from the definition that the Gini coefficient can also be used to quantify the degree heterogeneity of networks.

We consider a connected network without self- and multiple-links of N nodes, which are labeled from 1 through N in increasing order of node degrees, i.e. $d_1 \le d_2 \le \cdots \le d_N$, where d_i is the degree of node i. In Fig. 1, we plot the cumulative percentage $y\%$ of the total degree of nodes, $\sum_{k=1}^{i} d_k / \sum_{k=1}^{N} d_k$ as a function of the cumulative percentage $x\%$ of the number of nodes, i/N. Such a curve is called heterogeneity curve. For a network with the same degree for each node, the heterogeneity curve is the diagonal line $y = x$, the line of complete homogeneity; the heterogeneity curve of a completely heterogeneous network which can only be obtained in the infinite network size limit, by contrast, would be the curve where $y = 0$ when $x = 0$, $y \to 0$ for all $0 < x < 100$ and $y = 100$ when $x = 100$. Heterogeneity index H of the network is defined as the ratio of the area S_A between the line of complete homogeneity and the heterogeneity curve, to the area $S_A + S_B$ below the line of complete homogeneity.

The income sequence in a population can be an arbitrary nonnegative real number one, whereas the degree sequence of a network cannot be any positive integer one. Some sets of positive integers may not be realizable as a network, and there exist some constraint conditions on realizability of a set of positive integers as degrees of the nodes of a network [27].

The H considers the difference of every two degree values in a degree sequence. Actually H is equal to one-half of the relative mean difference, i.e. the arithmetic average of the absolute values of the differences between all possible pairs of node degrees [28]:

$$H = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} \left| d_i - d_j \right|}{2N^2 \bar{d}}, \tag{1}$$

where $\bar{d}$ is the average degree. Heterogeneity index provides a measure of the average degree inequality in a network, and larger index implies higher level of heterogeneity. Apparently, $0 \le H < 1$. $H = 0$ for a completely homogeneous network and $H \to 1$ for a completely heterogeneous network. For any real-world networks, their H's are all less than 1, and H may approach one only for infinite networks. The H is superior to some other parameters in characterizing heterogeneity, such as variance or standard deviation, since for comparing heterogeneity they demand that two networks studied should have the same average degree.

Denote W_i $(1 \le i \le N)$ as the ratio of the ith node's degree to the total degree of all nodes, i.e., $W_i = d_i / \sum_{k=1}^{N} d_k$. S_B is the area of a polygon under the heterogeneity curve. Obviously S_B can be decomposed into the areas of one triangle and $N-1$ trapezia, which can be expressed as $\frac{1}{2}\left(\sum_{k=1}^{i} W_k - W_i + \sum_{k=1}^{i} W_k\right) \cdot 1/N (1 \le i \le N)$. Thus we can compute the heterogeneity index as follows:

$$\begin{aligned} H &= S_A/(S_A + S_B) = 1 - 2S_B \\ &= 1 - 2\left\{\frac{1}{2}\sum_{i=1}^{N}\left[\left(\sum_{k=1}^{i} W_k - W_i + \sum_{k=1}^{i} W_k\right) \cdot 1/N\right]\right\} \\ &= 1 - \sum_{i=1}^{N}\left[1/N \cdot \left(2\sum_{k=1}^{i} W_k - W_i\right)\right]. \end{aligned} \tag{2}$$

In case that the degree distribution $P(k)$ $(1 \le k \le M)$ of the network is known, where M is the maximum degree, S_B can be decomposed into the areas of one triangle and $M-1$ trapezia. Thus we have the following expression for S_B:

$$\begin{aligned} S_B &= \frac{1}{2} \cdot P(1) \cdot \frac{P(1) \cdot N}{T} + \sum_{j=2}^{M} \frac{1}{2} \cdot P(j) \cdot \left[\sum_{i=1}^{j-1} P(i) \cdot N \cdot i/T + \sum_{i=1}^{j} P(i) \cdot N \cdot i/T\right] \\ &= \frac{1}{2} \cdot P(1)^2 \cdot N/T + \sum_{j=2}^{M} \frac{1}{2} \cdot P(j) \cdot \left[2 \cdot \sum_{i=1}^{j-1} P(i) \cdot N \cdot i/T + P(j) \cdot N \cdot j/T\right], \end{aligned} \tag{3}$$

where the first term on the right-hand side is the area of the triangle contributed by nodes with degree one and the second term is the areas of $M-1$ trapezia contributed by nodes with degree $2 \le k \le M$, $P(i) \cdot N$ $(1 \le i \le M)$ is the number of nodes with degree i, and T is the total degree of the network:

$$T = \sum_{k=1}^{M} P(k) \cdot N \cdot k = N \cdot \langle k \rangle. \tag{4}$$

The heterogeneity index can then be computed as

$$\begin{aligned} H &= 1 - 2S_B \\ &= 1 - P(1)^2 \cdot N/T - \sum_{j=2}^{M} P(j) \cdot \left[2 \cdot \sum_{i=1}^{j-1} P(i) \cdot N \cdot i/T + P(j) \cdot N \cdot j/T\right] \\ &= 1 - P(1)^2/\langle k \rangle - \sum_{j=2}^{M} P(j) \cdot \left[2 \cdot \sum_{i=1}^{j-1} P(i) \cdot i/\langle k \rangle + P(j) \cdot j/\langle k \rangle\right]. \end{aligned} \tag{5}$$

2.2. *Heterogeneity of regular networks*

Fig. 2 illustrates the topological structures of six regular networks and their heterogeneity indices are summarized in Table 1.

A path network has 2 end nodes with degree 1 and $N-2$ nodes with degree 2 (see Fig. 2(a)). We find from Table 1 that $H \to 0$ as $N \to \infty$, which results from a large number of nodes with the same degree 2.

Nearest-neighbor coupled networks (see Fig. 2(b)) and globally coupled networks (see Fig. 2(c)), are completely homogeneous graphs with the same degree for each node, thus $H = 0$ for these networks.

In a large-scale star network (see Fig. 2(d)), large difference between degrees of the central node and the leaf nodes tends to increase the heterogeneity, but a large number of leaf nodes with the same degree one tend to decrease the heterogeneity. Such a tradeoff leads to a heterogeneity index of $1/2$ for an infinite star network.

A connected stars network (center-to-center, see Fig. 2 (e)) is a graph of a few star networks, connected from hub to hub by several edges which constitute a ring. It is interesting that the heterogeneity index of such a network is independent of G, i.e. hub number, and has an upper bound $1/2$, similar to that of star networks.

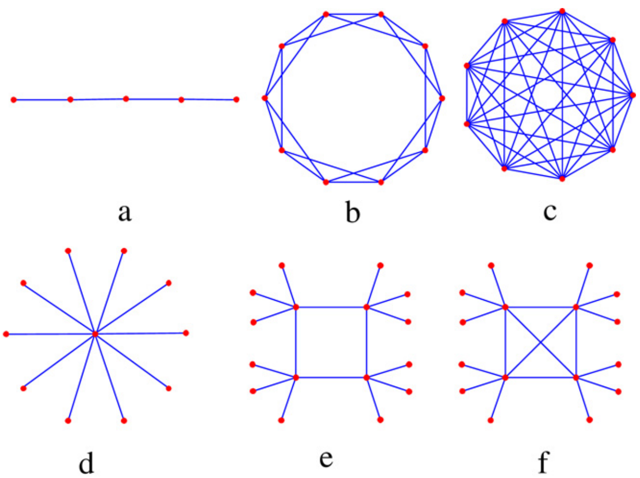

Fig. 2. Schematic illustration of six regular networks: (a) a path network with $N = 5$ nodes; (b) a K-nearest-neighbor coupled network with $K = 2$; (c) a globally coupled network with $N = 9$ nodes; (d) a star network with $N = 11$ nodes; (e) a connected stars network with $G = 4$ hubs and $L = 3$ leaf nodes to which every hub is connected; (f) a rich-hub stars network with $G = 4$ hubs and $L = 3$ leaf nodes to which every hub is connected.

Table 1
Heterogeneity indices of six regular networks shown in Fig. 2

Network	H	H_∞
Path network	$\frac{N-2}{N(N-1)}$	0
K-nearest-neighbor coupled network and globally coupled network	0	0
Star network	$\frac{N-2}{2N}$	1/2
Connected stars network	$\frac{L}{2(L+1)}\,(G \geq 3, L \geq 0)$	$\begin{cases} 1/2 & \text{as } L \to \infty \\ \dfrac{L}{2(L+1)} & \text{as } G \to \infty \end{cases}$
Rich-club stars network	$\frac{L^2+LG-2L}{(L+1)(2L+G-1)}\,(G \geq 2, L \geq 0)$	$\begin{cases} 1/2 & \text{as } L \to \infty \\ \dfrac{L}{L+1} & \text{as } G \to \infty \end{cases}$

H and H_∞ are heterogeneity indices of finite and infinite networks, respectively.

A rich-club stars network (see Fig. 2(f)) was proposed as a deterministic network with strong heterogeneity [29], which is a graph of a few star networks whose hubs are completely connected. We see from Table 1 that heterogeneity indices taking discrete values of this kind of network can vary in a wide range [0, 1).

3. Heterogeneity of power law networks

The degree distribution of a power law network has the form $P(k) \sim k^{-\gamma}$. It is well known that the smaller the value of γ, the more heterogeneous the network is and the more important the role of the hubs is in the network. If we sort the nodes of a network in decreasing degree sequence, we can define rank r of a node, to be the index of the node in the sequence. Heterogeneity index of power law networks can be analytically studied based on the relation between node rank and corresponding degree for such networks [29,30]:

$$D(r) = c \cdot r^{-\lambda}, \tag{6}$$

where $r = 1, 2, \ldots, N$ is node rank, $D(r)$ is degree value of node of rank r and $\lambda = 1/(\gamma - 1)$ is rank exponent. Substituting Eq. (6) into Eq. (2), we have the following expression of heterogeneity index for a power law network with N nodes:

$$H = 1 - \frac{1}{N}\sum_{i=1}^{N}\left[2 \cdot \sum_{k=1}^{i} \frac{(N+1-k)^{-\lambda}}{\sum\limits_{r=1}^{N} r^{-\lambda}} - \frac{(N+1-i)^{-\lambda}}{\sum\limits_{r=1}^{N} r^{-\lambda}}\right]. \tag{7}$$

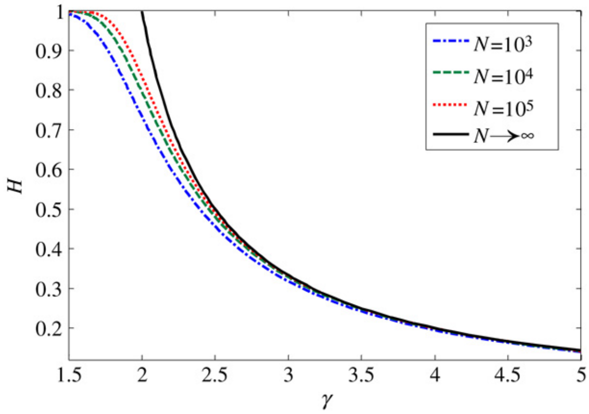

Fig. 3. Heterogeneity index H versus degree exponent γ for power law networks with different network sizes $N = 10^3$ (dash-dotted line), 10^4 (dashed line), 10^5 (dotted line) and $N \to \infty$ (solid line).

For an infinite power law network ($N \to \infty$), we have:

$$H_\infty = \begin{cases} \dfrac{\lambda}{2-\lambda} & \text{for } 0 \le \lambda < 1 \\ 1 & \text{for } \lambda \ge 1 \end{cases} = \begin{cases} \dfrac{1}{2\gamma-3} & \text{for } \gamma > 2 \\ 1 & \text{for } 1 < \gamma \le 2. \end{cases} \tag{8}$$

The H–γ relationship for finite and infinite power law networks is shown in Fig. 3 according to Eqs. (7) and (8). For any given network size N, H decreases with γ, which implies that small value of degree exponent indeed leads to high heterogeneity; On the other hand, for any given degree exponent γ, H increases with N, which indicates that large network size can lead to high heterogeneity. Particularly, as $N \to \infty$, $H \to 1$ if and only if $1 < \gamma \le 2$ and $H > 0.5$ if and only if $1 < \gamma < 2.5$. We can also see from Eq. (8) that $0.5 < H < 1$ if and only if $2 < \gamma < 2.5$. We will further show in the following section that for any power law network with a degree exponent $\gamma > 2.5$, there always exists an exponential network such that both networks have the same heterogeneity index. In fact, some recent researches have shown that $\gamma = 2.5$ is a critical point for different dynamic behaviors on power law networks [14,15].

Scale-free networks are thought to be highly heterogeneous and possess relatively large heterogeneity index, which can be demonstrated by several real networks (see Table 2), including the co-authorship network of the Los Alamos Condensed Matter archive, the Internet at the autonomous system level in 2004 and 2007, the WWW within nd.edu domain, and the protein interaction network (PIN) of the yeast *Saccharomyces cerevisiae*. Except for the PIN, all other networks have H_r larger than 0.5. The discrepancy between real heterogeneity index H_r and theoretical value H_t according to Eq. (7) mainly arises from two aspects. On the one hand, for real-world scale-free networks degree exponent γ is not an exact value and is only an estimated value obtained by various mathematical tools. On the other hand it is obvious that the degree distribution of real-world scale-free networks is not a strict power law. The non-power law form in the head or tail of a degree distribution can more or less lead to the H_r's deviation from H_t. The aforementioned two factors, together with the influence of network size on H result in the relatively large difference between H_r and H_∞. Especially, as shown in Table 2, the H_r for PIN is far different from H_t. Small scale of the network can induce the relatively large estimation error for γ; besides the PIN itself is not a strict power law network. Its degree distribution decays slower than exponential while faster than a power law for large degree [31].

4. Heterogeneity of exponential networks

In this section, as two special cases, we first numerically study the heterogeneity of the ER random graph and small-world models, and then we analytically study the heterogeneity of general exponential networks.

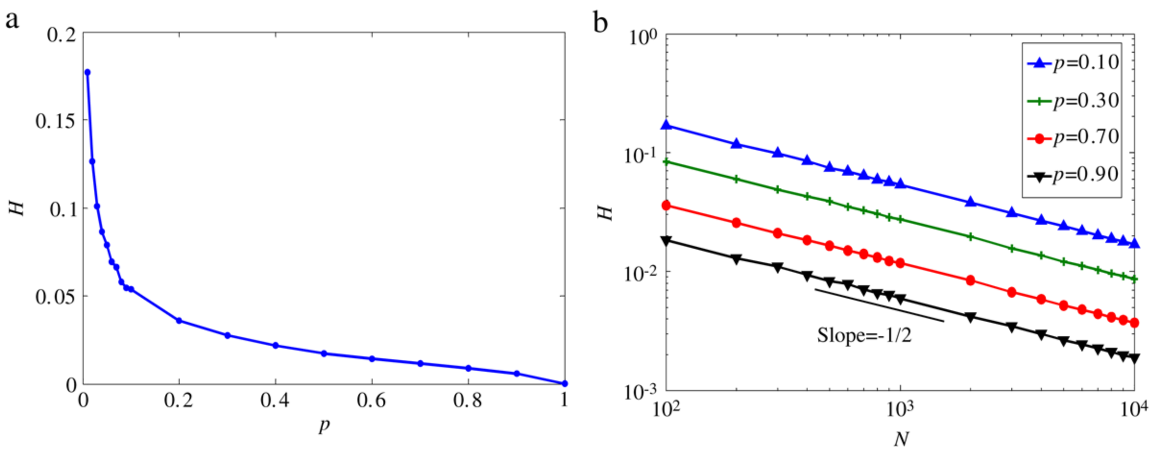

Fig. 4. Heterogeneity indices of ER random graphs. (a) Fixed node number $N = 1000$. (b) Fixed connection probability $p = 0.10$(▲), 0.30(+), 0.70(●), and 0.90 (▼); H and N satisfies power law relations with the same exponent $-1/2$. The plotted data are averaged over 20 independent realizations and all generated random networks are connected.

Table 2
Heterogeneity indices of five real scale-free networks

Network	N	$\langle k\rangle$	γ	H_r	H_t	e (%)	H_∞	Ref.
Co-authorship network	30 562	8.24	2.36	0.5146	0.5569	8.22	0.5814	[32]
Internet 2004	16 301	4.04	2.24	0.6221	0.6245	0.39	0.6757	[33]
Internet 2007	24 191	4.07	2.23	0.6309	0.6356	0.74	0.6849	[33]
WWW[a]	325 729	6.69	2.27	0.7215	0.6273	13.06	0.6494	[34]
PIN[b]	1 458	2.40	2.20	0.4662	0.6097	30.78	0.7143	[34]

$\langle k\rangle$ average degree, γ degree exponent, e relative error between real heterogeneity index H_r and theoretical value H_t according to Eq. (7), and H_∞ heterogeneity index of infinite network with the same γ according to Eq. (8)

[a] Directness of edges, multiple edges and self-loops of WWW within nd.edu domain have been ignored.

[b] The network consists of one big cluster and several isolated clusters. The parameters listed in Table 2 are based on the largest cluster containing 1458 out of 1870 proteins (∼78%).

4.1. ER random graph

Erdös and Rényi [10] brought forward an extremely simple model of a random network, i.e. taking some number N of nodes and connecting each pair randomly with probability p. Many properties of the ER random graph are exactly solvable in the limit of large graph size. Typically the mean degree satisfies $\langle k\rangle = p(N-1)$ and the model has a Poisson degree distribution:

$$P(k) = \frac{\langle k\rangle^k e^{-\langle k\rangle}}{k!}. \tag{9}$$

Heterogeneity indices of ER random graphs are shown in Fig. 4. For any given N, the heterogeneity index H is a decreasing function of the connection probability p. As p reaches 1, the ER random graph becomes a globally coupled network with $H = 0$ (Fig. 4(a)). In Fig. 4(b) which is plotted on log–log scale, for any given connection probability, H is a power law function of N with power law exponent $-1/2$: $H = a(p)\cdot N^{-1/2}$, here $a(p)$ is a decreasing function of p and $a(1) = 0$. This power law relationship is an intrinsically universal property for the ER random graphs which shows that heterogeneity indices decrease quite slowly with the increasing of node numbers.

4.2. Small-world models

In 1998, Watts and Strogatz proposed a simple model of social networks which interpolates between regular lattices and random graphs and displays both the clustering and small-world properties [11]. In this model, N nodes are placed

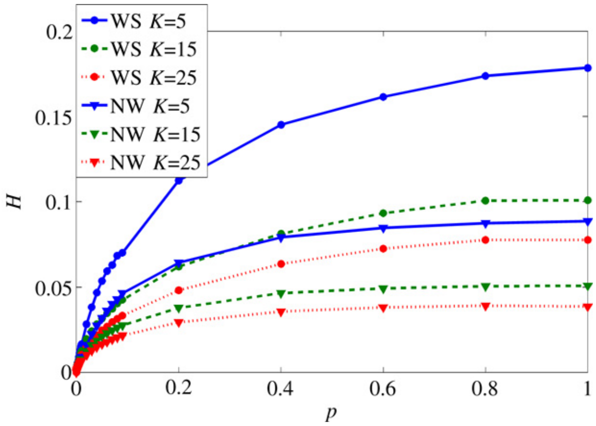

Fig. 5. Heterogeneity indices of WS and NW models with $N = 1000$ and $K = 5$ (solid line with • for WS and with ▼ for NW), 15 (dashed line with • for WS and with ▼ for NW), and 25 (dotted line with • for WS and with ▼ for NW). p is the rewiring probability for WS model and adding probability for NW model. All curves are averaged over 20 independent realizations and all generated small-world networks are connected.

on a regular one-dimensional lattice with nearest-neighbor connections out to some constant range K and periodic boundary conditions (the lattice is a ring and the degree of every node is $2K$). The small-world model is then created by taking a small fraction p of the edges randomly and "rewiring" them. The degree distribution expression of the model is [35]

$$P(j) = \sum_{n=0}^{\min(j-K,K)} \binom{K}{n} (1-p)^n p^{K-n} \frac{(pK)^{j-K-n}}{(j-K-n)!} \mathrm{e}^{-pK} \tag{10}$$

for $j \geq K$, and $P(j) = 0$ for $j < K$.

The shape of the degree distribution of the WS small-world model is similar to that of the ER random graph. It has a pronounced peak at $\langle k \rangle = 2K$ and decays swiftly for large or small degrees.

Another variant of the WS model was proposed by Newman and Watts [12]. In this NW model, no edges are rewired. Instead "shortcuts" joining randomly chosen node pairs are added to the lattice. The mean total number of shortcuts is NKp and the mean degree is $2NK(1 + p)$, here p is the adding probability. Degree distribution of the NW model is [36]

$$P(j) = \binom{N}{j-2K} \left[\frac{2Kp}{N}\right]^{j-2K} \left[1 - \frac{2Kp}{N}\right]^{N-j+2K} \tag{11}$$

for $j \geq 2K$, and $P(j) = 0$ for $j < 2K$.

The heterogeneity indices for WS and NW small-world networks are shown in Fig. 5. We find that H increases monotonically with p, implying the increase of randomicity in network structure can give rise to the enhancement of network heterogeneity. For the same K and p, H_{WS} is greater than H_{NW} (in fact, $H_{\mathrm{WS}} \approx 2H_{\mathrm{NW}}$), originating from the fact that in this case the proportion of random ingredient of network structure in WS model is larger than that in NW model.

4.3. A general exponential network model

Now we consider a general exponential network which has the degree distribution $P(k) = \beta \cdot c^{-k}$ ($c > 1, 1 \leq k \leq M$), where M represents the maximum degree. According to the normalization condition

$$\sum_{k=1}^{M} P(k) = 1, \tag{12}$$

we have $\beta = (c-1)/(1-c^{-M})$, therefore,

$$P(k) = \frac{c-1}{1-c^{-M}} \cdot c^{-k}. \tag{13}$$

According to Eq. (13), the average degree of an exponential network is

$$\langle k\rangle = \sum_{k=1}^{M} P(k)\cdot k = \frac{c^{M+1}-c-M(c-1)}{c^M(c-1)(1-c^{-M})}. \tag{14}$$

Its total degree is

$$T = N\cdot\langle k\rangle = \frac{N[c^{M+1}-c-M(c-1)]}{c^M(c-1)(1-c^{-M})}. \tag{15}$$

Let D_i be the total degree of nodes with degree i. According to the definition of heterogeneity curve, we have

$$S_B = \frac{1}{2}\cdot P(1)\cdot\frac{D_1}{T} + \sum_{j=2}^{M}\frac{1}{2}\cdot P(j)\left[\sum_{i=1}^{j-1} D_i/T + \sum_{i=1}^{j} D_i/T\right], \tag{16}$$

where

$$D_i = \frac{Ni(c-1)}{1-c^{-M}}\cdot c^{-i}.$$

Thus,

$$S_B = \left[\frac{N(c-1)^2}{2c^2(1-c^{-M})^2}\right]\Big/ T + \frac{A^2N}{2T}[2E+F], \tag{17}$$

where

$$A = \frac{c-1}{1-c^{-M}},\qquad E = \frac{(c^{-2}-c^{-M-1})}{c-1} + \frac{B-C}{c-1},\qquad F = \frac{2-c^{-2}-(Mc^2+c^2-M)\cdot c^{-2M}}{(c^2-1)^2},$$
$$B = \frac{2-c^{-2}-(c^2+Mc^4-Mc^2)\cdot c^{-2M}}{(c^2-1)^2},\quad \text{and}\quad C = \frac{2-c^{-1}-(c+Mc^2-Mc)\cdot c^{-M}}{c^M(c-1)^2}.$$

Thus,

$$H = 1-2S_B = \frac{c^{M+1}-c^{1-M}-Mc^2+M}{(1-c^{-M})(c^{M+1}-c-Mc+M)(c+1)}. \tag{18}$$

For infinite exponential networks, when $M\to\infty$, the following degree distribution holds:

$$P_\infty(k) = (c-1)\cdot c^{-k}(k\geq 1). \tag{19}$$

The mean degree is

$$\langle k\rangle_\infty = \sum_{k=1}^{\infty} P_\infty(k)\cdot k = \frac{c}{c-1}. \tag{20}$$

According to the definition of heterogeneity curve, in the limit of infinite network size, we have

$$S_{B_\infty} = \frac{c}{2(c+1)} \tag{21}$$

and the heterogeneity index is

$$H_\infty = 1-2S_{B_\infty} = \frac{1}{c+1}, \tag{22}$$

which can also be derived from Eq. (18) by taking $M\to\infty$.

Fig. 6(a) shows the H, whose value is given according to the colorbar, as functions of M and c. For any given c, H is an increasing function of M and it approaches the constant $1/(c+1)$ for sufficiently large M (see Fig. 6(b)).

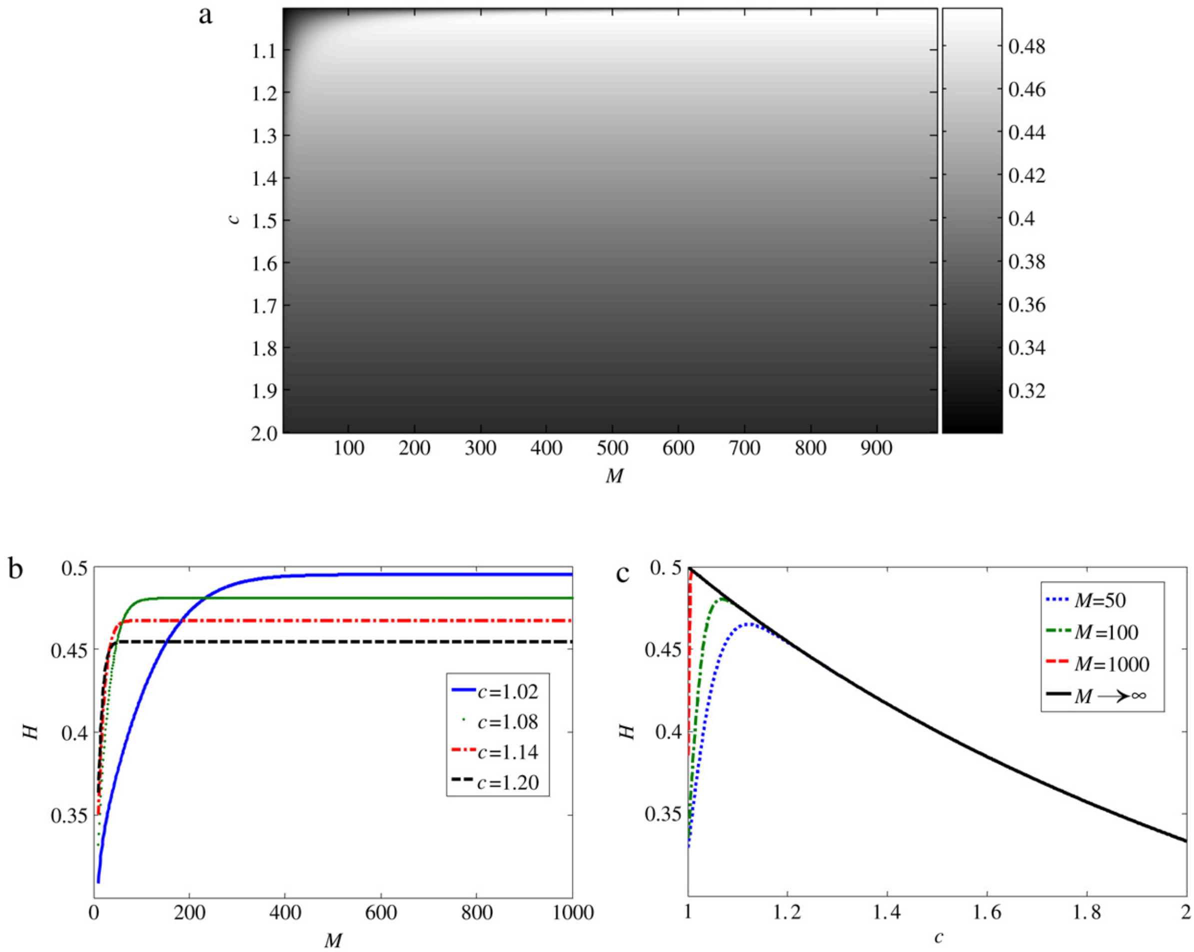

Fig. 6. Heterogeneity indices of exponential networks. (a) Image of H as functions of M from 10 to 1000 and c from 1.001 to 2. (b) H versus M for $c = 1.02$ (solid line), 1.08 (dotted line), 1.14 (dash-dotted line) and 1.20 (dashed line). (c) H versus c for $M = 50$ (dotted line), 100 (dash-dotted line), and 1000 (dashed line); also we picture the limit curve $H_\infty = 1/(c+1)$ (solid line) for comparison.

For any given M, H increases monotonously for small c and then decreases along the curve $1/(c+1)$ as c increases (see Fig. 6(c)). Since $c > 1$, we have $1/(c+1) < 1/2$, which implies that the heterogeneity index of an exponential network is always less than 0.5. This explains why exponential networks are called homogeneous networks.

In the last section, we have shown that the heterogeneity index of an infinite power law network is less than 0.5 if its degree exponent $\gamma > 2.5$. In fact, for any power law network with a degree exponent $\gamma > 2.5$, there always exists an exponential network such that both networks have the same heterogeneity index. On the one hand, for given M for exponential networks and N for power law networks, for every $\gamma > 2.5$ we always can find a c such that both networks have the same H. On the other hand, for infinite power law and exponential networks, according to Eqs. (8) and (22) we have $c = 2\gamma - 4$ for the same H.

To what degree can a real-world network be regarded as an exponential network? We compute the heterogeneity indices H_r of three real networks which have been regarded as exponential networks (see Table 3), including the power grids of the western United States (PGWUS), China railway network (CRN) and India railway network (IRN). For each network, we also compute the heterogeneity index H_t of the corresponding exponential network with the degree distribution Eq. (13).

It is worth noting that the relative error between H_t and H_r for PGWUS is relatively large. Strogatz regarded the network as an exponential one [7] and accordingly we have $H_t = 0.3668$, while Barabási and Albert regarded it as a

Table 3
Heterogeneity indices of three real networks. e is the relative error between H_t and H_r

Network	N	$\langle k \rangle$	c	M	H_t	H_r	e (%)	Ref.
PGWUS	4941	2.67	1.725	19	0.3668	0.3248	12.95	[37]
CRN	3717	57.67	1.014	564	0.4954	0.4679	5.86	[38]
IRN	587	66.79	1.009	379	0.4617	0.4547	1.55	[39]

power law network with degree exponent $\gamma = 4$ [13] and accordingly based on Eq. (7) we have $H_t = 0.1983$, a very small value. Actually PGWUS is an exponential network with a power law tail.

5. Conclusions

We have defined a unified heterogeneity index H to quantify the heterogeneity of a complex network with any degree distribution. We analytically study the heterogeneity of power law networks and find that the H of an infinite power law network is greater than 0.5 if and only if its degree exponent $\gamma \in (1, 2.5)$ and approaches 1 if and only if $\gamma \in (1, 2]$. We also analytically study the heterogeneity of general exponential networks and find the presence of upper bound 0.5 of H for such networks. For any power law network with a degree exponent $\gamma > 2.5$, there always exists an exponential network such that both networks have the same H. This may help to explain why 2.5 is a critical degree exponent for some dynamic behaviors on power law networks.

As an effective and unified measure, heterogeneity index can be applied to not only the characterization of heterogeneity of static networks but also the investigation of variation of heterogeneity for dynamic networks with dynamical processes taking place on them, and the influence of networks' heterogeneity on their dynamical behaviors.

Acknowledgements

We thank Parongama Sen for providing the India railway network data, Xiang Li for the helpful suggestions and Petter Holme for the valuable comments. We also thank the anonymous reviewers for their constructive remarks and suggestions which helped us to improve the quality of the manuscript to a great extent. This work was partly supported by the NSF of PRC for Creative Research Groups (60521002) and the NSF of PRC under Grant Nos. 70431002 and 60674045.

References

[1] R. Albert, A.-L. Barabási, Rev. Modern Phys. 74 (2002) 47;
S.N. Dorogovtsev, J.F.F. Mendes, Adv. Phys. 51 (2002) 1079;
S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Phys. Rep. 424 (2006) 175;
M.E.J. Newman, A.-L. Barabási, D.J. Watts, The Structure and Dynamics of Networks, Princeton University Press, Princeton, 2006.
[2] R. Albert, A.-L. Barabási, Phys. Rev. Lett. 85 (2000) 5234;
M.E.J. Newman, S.H. Strogatz, D.J. Watts, Phys. Rev. E 64 (2001) 026118.
[3] R. Albert, H. Jeong, A.-L. Barabási, Nature (London) 406 (2000) 378;
P. Crucitti, V. Latora, M. Marchiori, A. Rapisarda, Physica A 320 (2003) 622.
[4] R. Pastor-Satorras, A. Vespignani, Phys. Rev. E 63 (2001) 066117.
[5] S.N. Dorogovtsev, A.V. Goltsev, J.F.F. Mendes, Phys. Rev. E 66 (2002) 016104;
K.-I. Goh, D.-S. Lee, B. Kahng, D. Kim, Phys. Rev. Lett. 91 (2003) 148701;
P. Holme, Adv. Complex Syst. 6 (2003) 163;
E. Lieberman, C. Hauert, M.A. Nowak, Nature (London) 433 (2005) 312;
G. Szabó, G. Fáth, cond-mat/0607344.
[6] X.-F. Wang, G. Chen, Internat J. Bifur. Chaos 12 (2002) 187;
L. Kocarev, P. Amato, Chaos 15 (2005) 024101;
M. Barahona, L.M. Pecora, Phys. Rev. Lett. 89 (2002) 054101.
[7] S.H. Strogatz, Nature (London) 410 (2001) 268.
[8] D.H. Zanette, Phys. Rev. E 65 (2002) 041908;
Y. Moreno, M. Nekovee, A.F. Pacheco, Phys. Rev. E 69 (2004) 066130.
[9] R. Albert, I. Albert, G.L. Nakarado, Phys. Rev. E 69 (2004) 025103(R);
R. Kinney, P. Crucitti, R. Albert, V. Latora, Eur. Phys. J. B 46 (2005) 101.
[10] P. Erdös, A. Rényi, Publ. Math. Inst. Hung. Acad. Sci. 5 (1960) 17.

[11] D.J. Watts, S.H. Strogatz, Nature (London) 393 (1998) 440.
[12] M.E.J. Newman, D.J. Watts, Phys. Lett. A 263 (1999) 341.
[13] A.-L. Barabási, R. Albert, Science 286 (1999) 509.
[14] M. Aldana, P. Cluzel, Proc. Natl. Acad. Sci. USA 100 (2003) 8710.
[15] H.-J. Zhou, R. Lipowsky, Proc. Natl. Acad. Sci. USA 102 (2005) 10052.
[16] E. Volz, physics/0508160;
T. Zhou, Z.-Q. Fu, B.-H. Wang, Prog. Nat. Sci. 16 (2006) 452;
R. Pastor-Satorras, A. Vespignani, Epidemics and immunization in scale-free networks, in: S. Bornholdt, H.G. Schuster (Eds.), Handbook of Graphs and Networks: From the Genome to the Internet, Wiley-VCH, Berlin, 2003.
[17] T. Antal, S. Redner, V. Sood, qbio.PE/0601031.
[18] B.-H. Wang, T. Zhou, J. Korean Phys. Soc. 50 (2007) 134;
R. Guimerà, A. Díaz-Guilera, F. Vega-Redondo, A. Cabrales, A. Arenas, Phys. Rev. Lett. 89 (2002) 248701;
L. Zhao, Y.-C. Lai, K. Park, N. Ye, Phys. Rev. E 71 (2005) 026125;
G. Yan, T. Zhou, B. Hu, Z.-Q. Fu, B.-H. Wang, Phys. Rev. E 73 (2006) 046108;
B. Tadić, S. Thurner, G.J. Rodgers, Phys. Rev. E 69 (2004) 036102.
[19] A.E. Motter, C.-S. Zhou, J. Kurths, Phys. Rev. E 71 (2005) 016116;
M. Zhao, T. Zhou, B.-H. Wang, G. Yan, H.-J. Yang, W.-J. Bai, Physica A 371 (2006) 773;
M. Zhao, T. Zhou, B.-H. Wang, W.-X. Wang, Phys. Rev. E 72 (2005) 057102.
[20] L.A.N. Amaral, A. Scala, M. Barthélémy, H.E. Stanley, Proc. Natl. Acad. Sci. USA 97 (2000) 11149.
[21] M.E.J. Newman, Proc. Natl. Acad. Sci. USA 101 (2004) 5200.
[22] M.E.J. Newman, Phys. Rev. E 64 (2001) 016131.
[23] D.C. Colander, Microeconomics, McGraw-Hill, Boston, 2001.
[24] J. Weiner, O.T. Solbrig, Oecologia (Berlin) 61 (1984) 334;
C. Damgaard, J. Weiner, Ecology 81 (2000) 1139.
[25] P. Crucitti, V. Latora, S. Porta, Phys. Rev. E 73 (2006) 036125.
[26] M.J. Salganik, P.S. Dodds, D.J. Watts, Science 311 (2006) 854.
[27] S.L. Hakimi, SIAM J. Appl. Math. 10 (1962) 496;
V. Havel, Časopis Pěst. Mat. 80 (1955) 477.
[28] A.K. Sen, On Economic Inequality, Clarendon Press, Oxford, 1973.
[29] H.-B. Hu, L. Wang, Adv. Complex Syst. 8 (2005) 159.
[30] M. Faloutsos, P. Faloutsos, C. Faloutsos, Comput. Commun. Rev. 29 (1999) 251;
L.A. Adamic, B.A. Huberman, Glottometrics 3 (2002) 143.
[31] H. Jeong, S.P. Mason, A.-L. Barabási, Z.N. Oltvai, Nature (London) 411 (2001) 41.
[32] G. Palla, I. Derényi, I. Farkas, T. Vicsek, Nature (London) 435 (2005) 814.
[33] http://as-rank.caida.org/data/.
[34] http://www.nd.edu/˜networks/resources.htm.
[35] A. Barrat, M. Weigt, Eur. Phys. J. B 13 (2000) 547.
[36] M.E.J. Newman, SIAM Rev. 45 (2003) 167.
[37] http://cdg.columbia.edu/cdg/datasets.
[38] L. Wang, B. Zhu, H.-B. Hu, Systems Engng Suppl. 1 (2006) 140. (in Chinese).
[39] P. Sen, P. Dasgupta, A. Chatterjee, P.A. Sreeram, G. Mukherjee, S.S. Manna, Phys. Rev. E 67 (2003) 036106.