Vol 16 No 6, June 2007
1009-1963/2007/16(06)/1576-05

Chinese Physics

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

Normalized entropy of rank distribution: a novel measure of heterogeneity of complex networks*

Wu Jun(吴 俊)†, Tan Yue-Jin(谭跃进),
Deng Hong-Zhong(邓宏钟), and Zhu Da-Zhi(朱大智)

College of Information Systems and Management,
National University of Defense Technology, Changsha 410073, *China*

(Received 28 September 2006; revised manuscript received 14 November 2006)

Many unique properties of complex networks result from heterogeneity. The measure and analysis of heterogeneity are important and desirable to the research of the properties and functions of complex networks. In this paper, the rank distribution is proposed as a new statistic feature of complex networks. Based on the rank distribution, a novel measure of the heterogeneity called a normalized entropy of rank distribution (NERD) is proposed. The NERD accords with the normal meaning of heterogeneity within the context of complex networks compared with conventional measures. The heterogeneity of scale-free networks is studied using the NERD. It is shown that scale-free networks become more heterogeneous as the scaling exponent decreases and the NERD of scale-free networks is independent of the number of vertices, which indicates that the NERD is a suitable and effective measure of heterogeneity for networks with different sizes.

Keywords: complex networks, heterogeneity, rank distribution, scale-free networks
PACC: 0520, 0250

1. Introduction

Networks with complex topology describe a wide range of systems in nature and society.[1,2] In the past few years, the discovery of small-world[3] and scale-free properties[4] has stimulated a great deal of interest in studying the underlying organizing principles of various complex networks.[5−11]

There are major topological differences between random and scale-free networks. For random networks, vertices have approximately the same degree $k \approx \langle k \rangle$, the exponential decay of degree distribution guaranteeing the absence of vertices with significantly more edges than $\langle k \rangle$. In contrast, the power-law distribution implies that vertices with only a few edges are numerous, but a few vertices have a large number of edges. The presence of few highly connected vertices (i.e. 'hubs') is the most prominent feature of scale-free networks and indicates that scale-free networks are heterogeneous. The heterogeneity leads to many unique properties of scale-free networks. For example, Albert *et al*[12] have demonstrated that scale-free networks possess the robust-yet-fragile property, which means that they are robust against random failures of vertices but fragile to intentional attacks. Moreover, it is found that the homogeneous networks are more synchronizable than heterogeneous ones, even though the average network distance is larger.[13] Consequently, the measure and analysis of heterogeneity are important and desirable to the research of the behaviours and functions of complex networks.

Several measures of heterogeneity have been proposed. Nishikawa *et al*[13] quantitated the heterogeneity of complex networks using the standard deviation of degree σ. Solé and Alverde[14] proposed using the entropy of the remaining degree distribution $q(k)$ to measure the heterogeneity. Wang *et al*[15] measured the heterogeneity of complex networks using the entropy of the degree distribution $p(k)$. However, these conventional measures do not accord with the normal meaning of heterogeneity within the context of complex networks. For example, we are generally inclined to believe that a scale-free network is more heterogeneous than a random network, but it does not hold true for the measure above. In addition, a star network is generally considered to be very heterogeneous because of the presence of the only hub-vertex, but the star network is quite homogeneous for the conven-

*Project supported by the National Natural Science Foundation of China (Grant No 70501032).
†E-mail: wujunpla@163.com

tional measures.

In this paper, we first propose the rank distribution as a new statistic feature of complex networks. Then we propose a novel measure of the heterogeneity of complex networks called a normalized entropy of rank distribution (NERD) and compare it with conventional measures. Finally, we study the heterogeneity of scale-free networks using the NERD.

2. Rank distribution of complex network

A complex network can be represented by a graph G with N vertices and M edges. Assume that G is an undirected and simply connected graph. Let $d_{\rm i}$ be the degree of vertex $v_{\rm i}$. Obviously, $0 < d_{\rm i} < N$. As shown in Fig.1, we sort all vertices in descend order of degree. Note that different vertices may have the same degree, and then we can group all the vertices into $N-1$ groups according to their degree. That is, the degree of the vertices in the sth group is $N-s$. Let l_s be the number of vertices in the sth group, i.e. the frequency of vertices with degree $N-s$. Let s_i be the index of the group in which vertex v_i is located, and r_i be the global rank of v_i among all of the vertices in descending order of degree. Let $p(k)(0 < k < N)$ be the degree distribution, i.e. the probability at which a randomly chosen vertex has degree k.

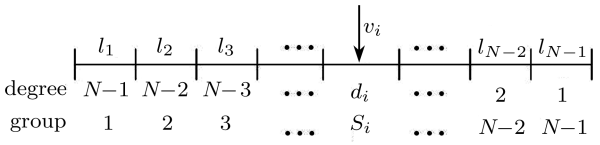


Fig.1. Sorting all vertices in descending order of degree and grouping all vertices into $N-1$ groups according to their degree.

We define the degree–rank function of G as the relationship function between degree and rank in the degree sequence $D(G)$ denoted by $d = f(r)(1 \le r \le N)$. The degree–rank function is non-stochastic, in the sense that there need be no assumption of an underlying probability distribution for the sequence.

Now we consider the vertex reached by following a randomly chosen edge. The probability distribution of the rank of the vertex is a new important statistic feature of complex networks. We define it as the rank distribution denoted by $Q(r)(1 \le r \le N)$. Such an edge arrives at a vertex with a probability proportional to the degree of the vertex. The correctly normalized distribution is

$$Q(r) = \frac{D_r}{\sum_s D_s} = \frac{f(r)}{\sum_s f(s)} = \frac{f(r)}{2M}. \tag{1}$$

As shown in Fig.2, we consider a graph with $N = 5$ vertices and $M = 6$ edges. The degree sequence is $D = (4, 3, 2, 2, 1)$. The degree distribution is $p(1) = 1/5$, $p(2) = 2/5$, $p(3) = 1/5$, and $p(4) = 1/5$. The degree–rank function is $f(1) = 4, f(2) = 3, f(3) = f(4) = 2$, and $f(5) = 1$. The rank distribution is $Q(1) = 1/3$, $Q(2) = 1/4$, $Q(3) = Q(4) = 1/6$, and $Q(5) = 1/12$.

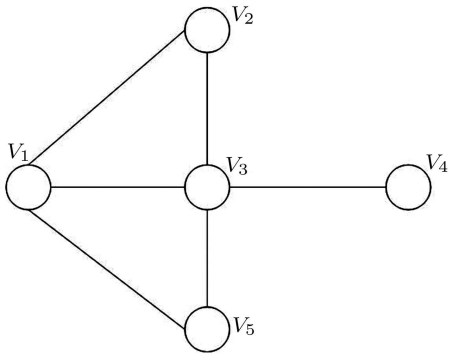


Fig.2. Illustration of a graph with five vertices and six edges.

3. Normalized entropy of rank distribution

To measure the heterogeneity of complex networks, we define an entropy of a rank distribution (ERD) as

$$\mathrm{ERD} = -\sum_{r=1}^{N} Q(r) \ln Q(r). \tag{2}$$

Obviously, the maximum value of ERD is $\mathrm{ERD}_{\max} = \log(N)$ obtained for $Q(r) = 1/N$, i.e. $d_1 = d_2 = \cdots = d_N$. Note that $d_i > 0$ and the values of d_i are all integers, so the minimum value of ERD is $\mathrm{ERD}_{\min} = \ln 4(N-1)/2$ and occurs when $D(G) = (N, 1, \cdots, 1)$. The maximum value of ERD corresponds to the most homogeneous network, i.e. a regular network, and the minimum value of ERD corresponds to the most heterogeneous network, i.e. a star network.

The normalized entropy of rank distribution (NERD) can be defined as

$$\mathrm{NERD} = \frac{\mathrm{ERD}_{\max} - \mathrm{ERD}}{\mathrm{ERD}_{\max} - \mathrm{ERD}_{\min}}. \tag{3}$$

For comparison, we present the definitions of entropy of the remaining degree distribution (ERDD) in Ref.[14]

$$\mathrm{ERDD} = -\sum_{k=0}^{N-2} q(k)\ln q(k), \tag{4}$$

where $q(k) = (k+1)p(k+1)/\sum_j jp(j)$, and entropy of degree distribution (EDD) in Ref.[15]

$$\mathrm{EDD} = -\sum_{k=1}^{N-1} p(k)\ln p(k). \tag{5}$$

For ERDD or EDD, the maximum value is $\log(N-1)$, which corresponds to the most heterogeneous network, and the minimum value is 0 obtained for $d_1 = d_2 = \cdots = d_N$, which corresponds to the most homogeneous network. To be consistent with the NERD, we define a normalized entropy of the remaining degree distribution (NERDD) and a normalized entropy of degree distribution (NEDD), respectively, as

$$\mathrm{NERDD} = \frac{\mathrm{ERDD} - \mathrm{ERDD}_{\min}}{\mathrm{ERDD}_{\max} - \mathrm{ERDD}_{\min}}, \tag{6}$$

and

$$\mathrm{NEDD} = \frac{\mathrm{EDD} - \mathrm{EDD}_{\min}}{\mathrm{EDD}_{\max} - \mathrm{EDD}_{\min}}. \tag{7}$$

Then a network becomes more heterogeneous as NERD or NERDD or NEDD increases.

Now we use NERD, NERDD, NEDD and standard deviation of degree σ to measure different complex networks including (a) a regular network with $N = 1000$; (b) a random network (ER model) with $N = 1000$ and connection probability $p = 0.3$; (c) a star network with $N = 1000$; (d) a scale-free network (Barabási–Albert model) with $N = 1000$ and $m_0 = m = 3$, the results are shown in Table 1.

Table 1. Results of comparison between the conventional measures and ours.

Type of network	NERD	NERDD	NEDD	σ
Regular network	0	0	0	0
Random network	0.0014	0.6493	0.6540	25.8196
Scale-free network	0.1118	0.4305	0.2960	6.4841
Star network	1	0.1004	0.0011	31.5595

With the NERDD or NEDD, the order of heterogeneity is: random network > scale-free network > star network > regular network. With the standard deviation of degree σ, the order of heterogeneity is: star network > random network > scale-free network > regular network. With the NERD, the order of heterogeneity is: star network > scale-free network > random network > regular network. We are generally inclined to believe that a scale-free network is more heterogeneous than a random network and a star network is very heterogeneous because of the presence of the only hub-vertex. So our measure accords with the normal meaning of heterogeneity within the context of complex networks compared with conventional measures.

4. Heterogeneity of scale-free networks

To obtain the ERD of scale-free networks, we first derive the relationship between the degree distribution $p(k)$ and the degree–rank function $f(r)$.

Since v_i is located in the s_ith group, then we obtain $\sum_{s=1}^{s_i} l_s \geq r_i$ and $\sum_{s=1}^{s_i-1} l_s \leq r_i$. Namely, s_i is the minimum T that satisfies $\sum_{s=1}^{T} l_s \geq r_i$, i.e.

$$s_i = T_{\min}\left\{\sum_{s=1}^{T} l_s \geq r_i\right\}. \tag{8}$$

Noting that $l_s = Np(k = N - s)$, we obtain

$$s_i = T_{\min}\left\{\sum_{s=1}^{T} p(k = N - s) \geq r_i/N\right\}. \tag{9}$$

Assuming that $p(k)$ is integrable, in a continuous approximation for $p(k)$, we obtain

$$s_i = T_{\min}\left\{\int_1^T p(k = N - s)\mathrm{d}s \geq r_i/N\right\}. \tag{10}$$

Owing to $p(k = N - s) \geq 0$, $\int_1^T p(k = N - s)\mathrm{d}s$ is an ascending function with respect to T, leading to

$$\int_1^{s_i} p(k = N - s)\mathrm{d}s = r_i/N. \tag{11}$$

If the degree distribution $p(k)$ is given, we can solve Eq.(11) for s_i, then we obtain the degree–rank distribution $d = f(r) = N - s$. Contrarily, if the

degree–rank function $d = f(r)$ is given, substituting $s = N - d$ into Eq.(11) we obtain

$$\int_d^{N-1} p(k)\mathrm{d}k = r/N. \tag{12}$$

Assuming that $d = f(r)$ is differentiable, we differentiate both sides of Eq.(12) with respect to r, then we obtain the degree distribution expressed as $p(d) = -1/(Nd')$, where d' is the derivative of d with respect to r.

For scale-free networks with power law degree distributions $p(k) = Ck^{-\lambda}$, where λ is the scaling exponent, according to the relationship above, we can obtain the degree–rank function of scale-free networks as follows:

$$d = f(r) = \left(\frac{\lambda-1}{NC}\right)^{\frac{1}{-\lambda+1}} \times \left[r + \frac{N(N-1)^{-\lambda+1}}{\lambda-1}\right]^{\frac{1}{-\lambda+1}}. \tag{13}$$

Note that the scaling exponent of most scale-free networks in the real world ranges between 2 and 3. We have $N(N-1)^{\lambda+1}/(\lambda-1) \to 0$ as $N \to \infty$. Then, Eq.(13) reduces to

$$d = f(r) \approx C_1 \cdot r^{-\alpha}, \tag{14}$$

where $C_1 = [(\lambda-1)/(NC)]^{-\alpha}$ and $\alpha = 1/(\lambda-1)$. We refer to α as a degree–rank exponent of scale-free networks.

Substituting Eq.(14) into Eq.(1), we obtain the rank distribution of scale-free networks as follows:

$$Q(r) = \frac{f(r)}{\sum_s f(s)} = \frac{C_1 r^{-\alpha}}{\sum_s C_1 s^{-\alpha}} = \frac{r^{-\alpha}}{\sum_s s^{-\alpha}}. \tag{15}$$

According to Eq.(15), the rank distribution of scale-free networks is only independent of the scaling exponent λ.

Substituting Eq.(15) into Eq.(2), we obtain the ERD of scale-free networks as follows:

$$\mathrm{ERD} = -\sum_{r=1}^{N} \left\{ \frac{r^{-\alpha}}{\sum_s s^{-\alpha}} \ln\left(\frac{r^{-\alpha}}{\sum_s s^{-\alpha}}\right)\right\}. \tag{16}$$

In a continuous approximation, we obtain the ERD of scale-free networks as a function of the degree–rank exponent α,

$$\mathrm{ERD} = \frac{\alpha \cdot N^{1-\alpha} \cdot \ln N^{1-\alpha}}{(1-\alpha)(N^{1-\alpha}-1)} + \ln\frac{N^{1-\alpha}-1}{1-\alpha} - \frac{\alpha}{1-\alpha}. \tag{17}$$

Substituting $\alpha = 1/(\lambda-1)$ into Eq.(17), we obtain the ERD of scale-free networks as a function of the scaling exponent λ,

$$\mathrm{ERD} = \frac{N^{\frac{\lambda-2}{\lambda-1}} \ln N^{\frac{\lambda-2}{\lambda-1}}}{(\lambda-2)(N^{\frac{\lambda-2}{\lambda-1}}-1)} + \ln\frac{(\lambda-1)(N^{\frac{\lambda-2}{\lambda-1}}-1)}{\lambda-2} - \frac{1}{\lambda-2}. \tag{18}$$

Substituting Eq.(18) into Eq.(3), we can obtain the NERD of scale-free networks. In Fig.3, we show the plots of NERD versus $\lambda \in (2,3)$ for different values of N. We can find that a scale-free network becomes more heterogeneous as λ decreases. Moreover, we can find that the plots of NERD for different values of N overlap with each other, which indicates that the NERD is independent of N and then the NEDS is a suitable measure of heterogeneity for different sizes of scale-free networks.

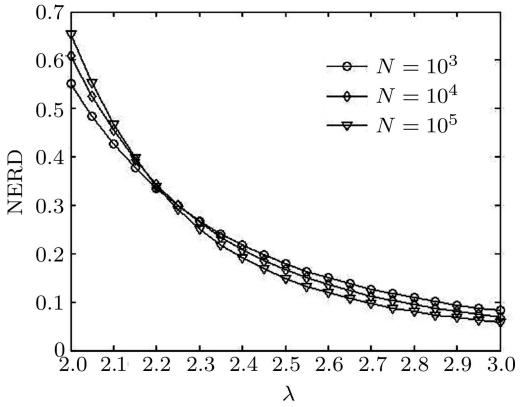


Fig.3. NERD versus $\lambda \in (2,3)$ for different values of N, showing that the NERD decreases as λ increases and NERD is independent of N.

5. Conclusions

In this paper, we have proposed a new statistic feature of complex networks called a rank distribution. Based on the rank distribution, we have proposed a normalized entropy of rank distribution (NERD) to measure the heterogeneity of complex networks. The maximum value of NERD corresponds to the most heterogeneous network, i.e. a star network, and the minimum value of NERD corresponds to the most homogeneous network, i.e. a regular network. We measure different networks by using the conventional measures and ours separately. By comparison, the NERD accords with the normal meaning of heterogeneity within the context of complex networks.

We have studied the heterogeneity of scale-free networks using the NERD and derived an analytical expression of NERD for scale-free networks. We have demonstrated that scale-free networks become more heterogeneous as scaling exponent decreases. We have also demonstrated that the NERD of scale-free networks is independent of the size of networks indicating that the NERD is a suitable and effective measure of heterogeneity.

References

[1] Albert R and Barabási A L 2002 *Rev. Mod. Phys.* **74** 47

[2] Strogatz S H 2001 *Nature* **410** 268

[3] Watts D J and Strogatz S H 1998 *Nature* **393** 440

[4] Barabási A L and Albert R 1999 *Science* **286** 509

[5] Huang Z X, Wang X R and Zhu H 2004 *Chin. Phys.* **13** 273

[6] Li Y, Liu Y, Shan X M, Ren Y, Jiao J and Qiu B 2005 *Chin. Phys.* **14** 2153

[7] Hu K and Tang Y 2006 *Chin. Phys.* **15** 2782

[8] He Y, Zhang P P, Xu T, Jiang Y M and He D R 2004 *Acta Phys. Sin.* **53** 1710 (in Chinese)

[9] Zhang P P, He Y, Zhou T, Su B B, Chang H, Zhou Y P, Wang B H and He D R 2006 *Acta Phys. Sin.* **55** 60 (in Chinese)

[10] Wang B H, Zhou T and He D R 2005 *China Basic Science* **7** 37 (in Chinese)

[11] Che H A and Gu J F 2004 *Systems Engineering—Theory & Practice* **24** 11 (in Chinese)

[12] Albert R, Jeong H and Barabási A L 2000 *Nature* **406** 378

[13] Nishikawa T, Motter A E, Lai Y C and Hoppensteadt F C 2003 *Phys. Rev. Lett.* **91** 014101

[14] Solé R V and Alverde S V 2004 *Lect. Notes. Phys.* **650** 189

[15] Wang B, Tang H W, Guo C H and Xiu Z L 2005 *Physica A* **363** 591