International Journal of Modern Physics C
Vol. 19, No. 12 (2008) 1909–1918

DEGREE AND WEIGHTED PROPERTIES OF THE DIRECTED CHINA RAILWAY NETWORK

LONG GUO* and XU CAI

Complexity Science Center, Institute of Particle Physics
Huazhong (Central China) Normal University, Wuhan 430079, China
**longkuo0314@gmail.com*

Received 20 March 2008
Accepted 24 September 2008

As well known, many real complex systems are directed and weighted ones. For understanding the topology structure of the directed China Railway Network (CRN) further, we analyze the degree properties of the directed CRN and propose a new method to measure the weight of station (i.e., the utilized efficiency of station) in CRN according to how CRN works really. Rigorous analysis of the existing CRN data shows that the CRN is an assortative network with scale-free degree distribution in space L. On the other hand, the cumulative distribution of station's relative weight, the cumulative distribution of the shortest path length between stations, and the relationship between relative weight and in-(out-)degree have been studied. Our present work provides a new perspective to define the weight of transport complex network, which will be helpful for studying the dynamics of CRN.

Keywords: Assortative complex network; exponential property; utilized efficiency; transport system.

PACS Nos.: 89.75.Fb, 89.40.Bb, 05.90.+m.

1. Introduction

The last few years have witnessed an increasing interest in studying the statistical and dynamical properties of transport systems in the physics community. Several real-world transport networks, such as airport networks,[1,2] railway networks,[3–5] subway networks[6–8] and bus-station networks,[9–11] have been studied using various concepts of statistical physics of complex networks. Most previous studies have revealed that most transport networks appear to share well-known small-world properties.[12]

Generally, there are two basic different topological representations during studying the structural properties of transport system. One typical topological representation is space P, which is proposed by Sen *et al.*[3] The idea of space P indicates that two stations are considered to be connected by a link when there is at least one train, which stops at both the stations. In space P, the shortest path length between two arbitrary stations is the number of changing trains one has to take to get from one station to another. Hence, the space P can be called the space of changes.[13] The other typical topological representation is space L,[5,10] in which two stations are connected by an edge if they are two consecutive stop stations on a route. The shortest path length between two arbitrary stations is the minimal number of stops one need to make in space L. Then, space L has another name — space of stops.[13] Note that the spaces L and P are abstract spaces that are different from the Euclidean space.[14] In this paper, we pay most of our attention to the structure properties of China Railway Network (CRN) in space L, since the structure properties of CRN in space P had been studied in Ref. 4.

As well known, the formation of statistical and dynamical properties of CRN, which plays a crucial role in people's daily lives and in China economy, is influenced by the policy, economy, and culture. However, many previous works did not consider those factors in abstract spaces L and P, which lost some important information of transport systems when analyzing their dynamical properties. Here, we propose a new method to measure station weight, namely, the utilized efficiency of station, according to how CRN really works. Then, the cumulative distribution of station's relative weight, the cumulative distribution of the shortest relative weight between stations, and the relationships between relative weight and in-(out-)degree have been studied and the exponential properties have been identified.

The main goal of this paper is to study the degree and weighted properties of the directed CRN consisted of $N_S = 3467$ stations and $N_T = 2535$ trains in space L. we demonstrate that, on the one hand, CRN is an assortative network with scale-free degree distribution in space L. On the other hand, a few of stations, which play the same role as hubs in Internet, have higher utilized efficiency. The paper is organized as follows. Section 2 presents the results of degree distributions and degree–degree correlation of CRN in spaces L. In Sec. 3, we show the results of the cumulative distribution of station's relative weight, the relative weight-degree correlation, and the cumulative distribution of the shortest path length between stations in both space L and P. Conclusions are given in Sec. 4.

2. Degree Distribution and Degree–degree Correlation

We consider the directed CRN as a graph with N_S nodes (stations) and E edges. The connectivity is represented by the $N_S \times N_S$ adjacency matrix A whose element a_{ij} is equal to one when there is at least one train which consecutive stops at stations i and j and zero otherwise in space L.

In the directed CRN, the degree of station has two components: the number of outgoing trains $k_i^{\text{out}} = \sum_{j=1}^{N_S} a_{ij}$ leaving from the station i (called the out-degree of the station i) and the number of ingoing trains $k_i^{\text{in}} = \sum_{j=1}^{N_S} a_{ji}$ leaving for the station i (called the in-degree of the station i). The total degree is then defined as $k_i = k_i^{\text{out}} + k_i^{\text{in}}$. The most basic topological property of the directed CRN can be obtained in terms of the in-(out-)degree cumulative distribution in the both

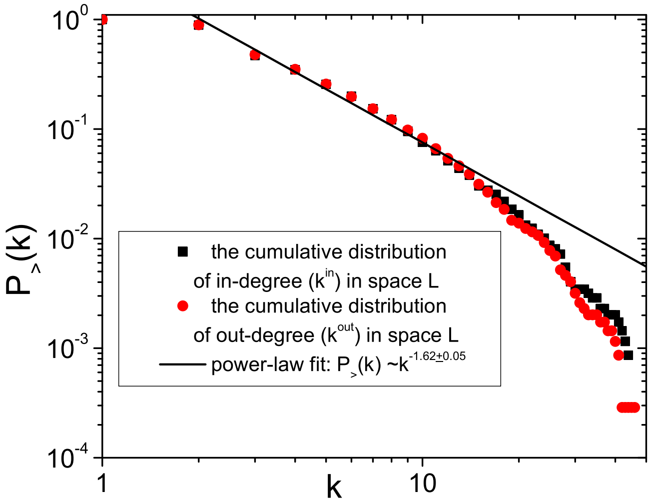


Fig. 1. Plots for the cumulative distribution of in-(out-)degree in space L in log–log representation.

abstract spaces. Figure 1 shows the plots for typical cumulative distribution of in-(out-)degree in space L. In Fig. 1 we can see that the cumulative distribution of in-(out-)degree follows power laws $P_>(k) \sim k^{-\gamma}$ with the power exponent $\gamma = 1.62$ in space L. Such power-law degree distribution can be found in other transportation networks, take three bus-station networks of China[11] and 22 public transport networks in Poland,[10] for example. Further, the points of in-degree distribution and out-degree distribution overlap each other in Fig. 1, which will be explained below. Interestingly, there exist some dots below the power-law line in Fig. 1, probably caused by some fluctuations due to the finite size effect.

Further light can be shed on the correlation properties of the CRN. To do this, the first choice is the degree–degree correlation and the in- and out-degree correlation, which maybe can explain the overlap phenomenon of the in-(out-) degree distribution.

As we know, the degree distribution is formally characterized by conditional probability $P(k'|k)$ that be defined as the probability that an edge from a node of degree k points to a node of degree k'. However, the direct evaluation of the condition probability gives extremely noisy results for most real networks because of their finite size N_S. In order to overcome this problem, we define the average degree of the neighbors of nodes with degree k, denoted as $k_{nn}(k)$.[15] Such a quantity can be expressed in terms of the conditional probability as

$$k_{nn}(k) = \sum_{k'} k' P(k'|k)\,. \tag{1}$$

where $P(k'|k)$ is the conditional probability, being defined as the probability that a link from a node of degree k points to a node with degree k'. $k_{nn}(k)$ increases

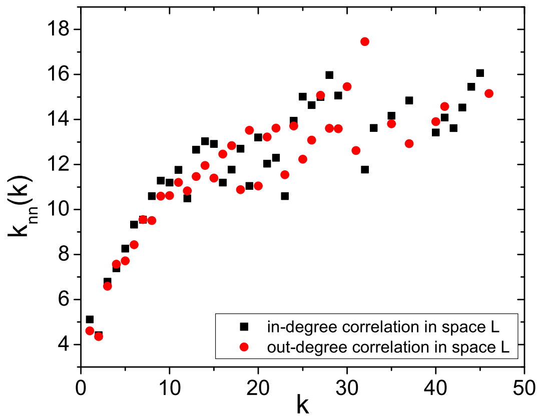


Fig. 2. Plots for the evolution of $k_{nn}(k)$ as a function of k in space L.

as a function of degree if nodes correlated by degree (assortative networks), and it decreases if they are anti-correlated (disassortative networks).[16]

In Fig. 2, we represent the evolution of both k_{nn}^{in} and k_{nn}^{out} as a function of the in-degree k^{in} and out-degree k^{out}, respectively, in space L. We can see that both of k_{nn}^{in} and k_{nn}^{out} increase as a function of in-degree k^{in} and out-degree k^{out}, respectively. Hence, the directed CRN is an assortative scale-free network in space L. That is to say, the stations tend to connect to their connectivity peers in assortative CRN, since there exist numerous smaller stations connected between two major ones in railway network. The assortative property of CRN in space L is different from that of the CRN[4] and the IRN[3] in space P because of the part geographical factor in space L.

We also analyze the in- and out-degree correlation following the method of degree–degree correlation mentioned above. We study the average out-degree of a node with in-degree k^{in}, denoted as $k^{\text{out}}(k^{\text{in}})$, which can be defined as follows:

$$k^{\text{out}}(k^{\text{in}}) = \sum_{k^{\text{out}}} k^{\text{out}} P(k^{\text{out}}|k^{\text{in}})\,, \tag{2}$$

where $P(k^{\text{out}}|k^{\text{in}})$ is the conditional probability that defined as the probability that a node with in-degree k^{in} has an out-degree k^{out}.

In Fig. 3, we present the evolution of k^{out} as a function of in-degree k^{in}. We find that there are most dots lie on the line $k^{\text{out}}(k^{\text{in}}) = k^{\text{in}}$. Namely, most stations have the same in-degree and out-degree, which can explain the overlap phenomenon of the cumulative distributions of in-degree k^{in} and out-degree k^{out} in Fig. 1. Both of the results reflect that the CRN is a balanced transportation system in space L, since there is one train leaving for the station, and then there is another train leaving from the same station in order to keep the traffic flow balance. The phenomenon of traffic flow balance, which has been found in China airport network,[1] is a basic phenomenon in transportation systems. Interestingly, there are some dots below

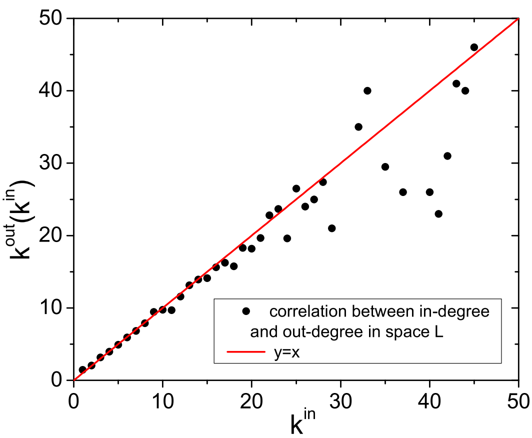


Fig. 3. The plots for the evolution of $k^{\text{out}}(k^{\text{in}})$ as a function of the in-degree k^{in}. The slope of the solid line in the picture is one, namely, there exists nodes that have the same number of out-degree and in-degree while the dots on the line.

the line $y = x$ in Fig. 3, namely, the number of outgoing trains leaving from one station is not always equal to the number of ingoing trains leaving for the same one, probably caused by the capacity of trains. We find that those stations with larger in-degree k^{in} and out-degree k^{out} play the same role as hubs in Internet.

3. Weighted Properties of the Directed CRN

Another important feature of CRN, along with many other real-world networks, is that each node (station) plays different role in the system. Some nodes have more weight than others and therefore play a greater role in contributing to the dynamics of the whole network. In fact, the trains running on CRN are divided into through trains, red balls, rapidoes, and way trains in China. The condition of stops of train is related to its kind, i.e., the through trains and red balls stop at the important stations — which always have been built in the cities that are the political, economic, and cultural centers of Provinces — and the way trains stop at any one. Therefore, it is necessary to consider the station weight and its distribution. According to how CRN works really, we define the weighted degree of station i as follows:

$$w_i = \sum_{j=1}^{N_T} \frac{1}{n_j(i)}\,, \tag{3}$$

where $N_T = 2535$ is the number of trains running through the CRN and $n_j(i)$ is the number of stops of train j that stops at station i. w_i represents the utilized efficiency of stations, which is related to the number of and the kind of trains. The larger is the weight w_i of station i, the more important role the station i plays in CRN.

For the sake of convenience, we define the relative weight rw_i of station i as follows:

$$rw_i = \frac{1/w_i}{\max\{1/w_1, 1/w_2, 1/w_3, \ldots, 1/w_{N_S}\}}\,, \tag{4}$$

where N_S is the number of stations, and the denominator is equal to the maximum value of the set $\{1/w_j\}(j = 1, 2, \ldots, N_S)$ in CRN. Hence, the relative weight $rw_i \in (0, 1]$, namely, we set the relative weight of the station with the lowest utilized efficiency as one. The smaller is the relative weight rw_i of station i, the more important role the station i plays in CRN.

Generally, the role and utilized efficiency of a node is proportion to its node degree in complex network. The nodes with higher node degree, which always have higher utilized efficiency (i.e., lower relative weight rw), are called hubs. Proven the conclusion, in Fig. 4 we represent the correlation between the in-degree k^{in} and the relative weight (rw) in both abstract space L and P. We find that there exist some stations with smaller in-(out-)degrees have smaller relative weights rw, namely, the higher utilized efficiency of those stations. Take 302 (the serial number of East Guangzhou station) and 315 (the serial number of Longchuan station) stations, for example, although they have almost the same in-(out-)degree ($k^{\text{in}}_{302} = 16, k^{\text{out}}_{302} = 21, k^{\text{in}}_{315} = 17, k^{\text{out}}_{315} = 18$), their relative weights are quite different in space L, i.e., $rw_{302} = 4.08 \times 10^{-4}$ and $rw_{315} = 5.6 \times 10^{-3}$. The same result is also found in space P, see the left picture of Fig. 4. Hence, the role and utilized efficiency of the East Guangzhou station is more important than that of the Longchuan station, since the Guangzhou is the political, economic, and cultural center of Guangdong Province and there are many through trains and red balls leaving for and from the station. On the other hand, only from the in-(out-)degree of the two stations, we cannot find any difference between them. What is more, we find that the stations with

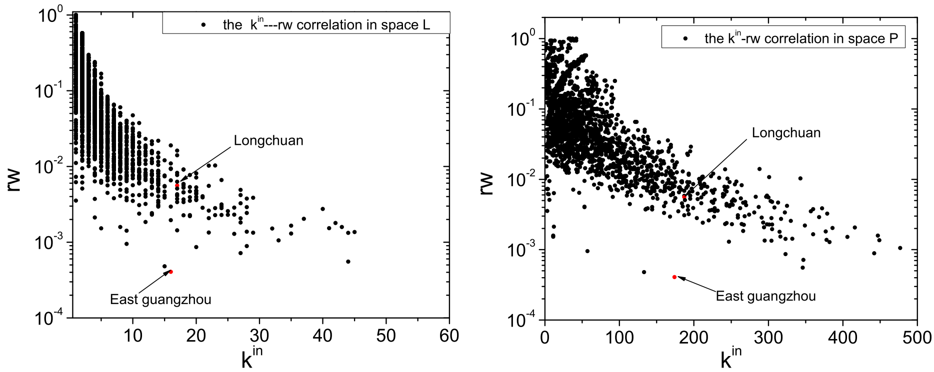


Fig. 4. Left: The plots for the correlation between the related weight (rw) and the in-degree (k^{in}) in space L. Right: The plots for the correlation between the related weight (rw) and the in-degree (k^{in}) in space P. Each dot represents the station with (k^{in}, rw) accordingly.

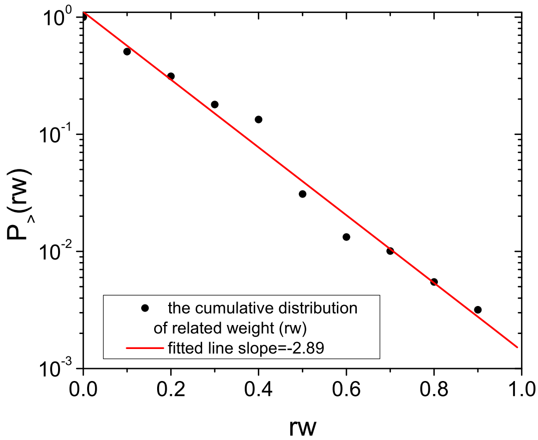


Fig. 5. Plots for the evolution of $P_>(rw)$ as a function of the relative weight rw in log-normal representation. The slope of the fitted line is -2.89.

higher utilized efficiency always have been built in cities, which are the political, economic, and cultural centers of local and global regions in China.

We consider the distribution of rw_i firstly. In order to reduce the statistical errors arising from the system finite size, we adopt the cumulative weight distribution. The cumulative distribution form, $P_>(rw)$, which is relative to the original distribution $P(rw)$ through the following formula:

$$P_>(rw) = \int_{rw}^{rw_{\max}} P(rw')drw' \tag{5}$$

where $rw_{\max} = 1$ is the maximal relative weight available in CRN. In Fig. 5, we represent the cumulative distribution $P_>(rw)$ as a function of the relative weight rw. We find that the cumulative distribution of the relative weight for CRN approximately fits to an exponential decaying distribution $P_>(rw) \sim \exp(-\beta rw)$ with $\beta = 2.89$, which is different from the distribution of degree in Fig. 1. Both the quantities of relative weight that we defined and degree play different roles in the structure and dynamics of CRN. Hence, we should consider the political, economical, and cultural factor (i.e., the role of relative weight) during studying the traveling dynamics of people by trains in abstract spaces L and P.

As well known, in abstract spaces L and P, we consider the structure topological properties of transport systems from the viewpoint of advantage of people's traveling. We hope to arrive at our destination expediently. Therefore, we study the shortest path length d, which describes the most advantage of a route between stations quantificationally, in the weighted CRN. In order to do this, we should define the relative weight of edge e_{ij} as follows:

$$e_{ij} = a_{ij} \times (rw_i + rw_j)\,, \tag{6}$$

where a_{ij} is the element of adjacency matrix A of CRN and rw_i is the relative weight of station i. Hence, the shortest path length d_{ij} is defined as the minimum summation of relative weights along the shortest route S from station i to station j and can be defined as follows:

$$d_{ij} = \min \sum_{(u,v)\in S} e_{uv} \tag{7}$$

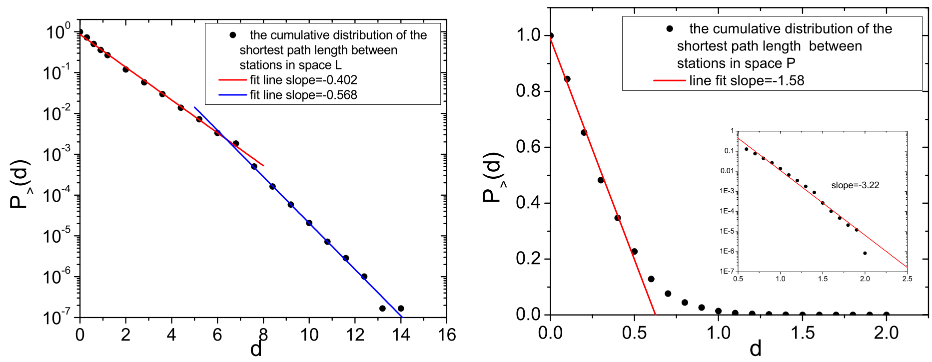
where the route S is the set of edges $\{a_{ik}, a_{kl}, \ldots, a_{uv}, \ldots, a_{wj}\}$. The route S provides the more advantage than other routes when you traveling from station i to station j. Therefore, it is necessary to study the property of the shortest path length of weight CRN before researching the dynamics of CRN in abstract space L and P. Here, we focus on the cumulative distribution $P_>(d)$ of the shortest path length in weighted CRN. The cumulative distribution $P_>(d)$ is defined as follows:

$$P_>(d) = \int_d^{d_{\max}} P(x')dx' \,, \tag{8}$$

where $P(x')$ is the probability that defined as the probability of the shortest path length x' between stations in weighted CRN, and $d_{\max}$ is the diameter of weighted CRN.

In Fig. 6, we represent the evolution of the cumulative distribution of the shortest path length $P_>(d)$ as a function of d in both spaces L and P. Interestingly, Fig. 6(a) shows that the cumulative distribution of the shortest path length approximately fits to a two-regime exponential decaying distribution with different exponents in space L. The turning value of the shortest path length $d_c^l \simeq 7$ in space L and the two-regime exponential decaying distribution can be well prescribed by the following expression:

$$P_>(d) \sim \begin{cases} \exp(-0.402d)\,, & \text{if } d < d_c^l\,, \\ \exp(-0.568d)\,, & \text{if } d > d_c^l\,. \end{cases} \tag{9}$$



(a) (b)

Fig. 6. Plots for the cumulative distribution of the shortest path length as a function of the shortest path length (a) in space L on log-normal representation and (b) in space P on normal representation. The inset in (b) shows the evolution of the cumulative distribution in space P when $d > d_c^p \simeq 0.6$, where d_c^p is the turning value of the shortest path length in space P.

On the other hand, Fig. 6(b) shows the evolution of the cumulative distribution $P_>(d)$ as the shortest path length d in space P. We find that the cumulative distribution $P_>(d)$ follows a linear function of the shortest path length d with the slope (-1.58) when $d < d_c^p \simeq 0.6$, where $d_c^p \simeq 0.6$ is the turning value of the shortest path length in space P. When $d > d_c^p \simeq 0.6$, the cumulative distribution $P_>(d)$ approximately fits to an exponential decaying distribution with the exponent (-3.22), see the inset in Fig. 6(b). It is amazing to find that $d_c^l \gg d_c^p$, probably caused by the geographic factor in space L. Furthermore, we also find that there are about 80% shortest path lengths that are lower than 0.5 in space P and about 50% lower than 0.5 in space L.

4. Conclusions

In this paper, we analyzed the degree and weighted properties of the directed CRN in abstract spaces. We find that the degree distribution is approximately given by power-law function and the degree–degree correlation $k_{nn}(k)$ increases as the function of degree, which shows that the topological structure of CRN is an assortative scale-free transport network. On the other hand, according to how CRN works really, we proposed a new method to measure the utilized efficiency (i.e., the weight and relative weight) of stations, which reflects the role of the political, economic, and cultural factors in the CRN working. The relation between in-degree and relative weight of station, the cumulative distribution of relative weight, and the shortest path length in weighted CRN have been studied, and exponential properties of the directed CRN have been identified. Our present work provides a new viewpoint to measure the weight (the utilized efficiency) of station according to how CRN works really and shows that it is necessary to consider the effects of weight during studying dynamics of CRN in abstract spaces.

Acknowledgments

L. Guo thanks Prof. Li for valuable suggestions and comments. This work was supported by the National Natural Science Foundation of China under Grant Nos. 70571027 and 10635020 and by the Ministry of Education in China under Grant No. 306022.

References

1. W. Li and X. Cai, *Phys. Rev. E* **69**, 046106 (2004).
2. R. Guimerà and L. A. N. Amaral, *Eur. Phys. J. B* **38**, 381 (2004).
3. P. Sen, S. Dasgupta, A. Chatterjee, P. A. Sreeram, G. Mukherjee and S. S. Manna, *Phys. Rev. E* **67**, 036106 (2003).
4. W. Li and X. Cai, *Physica A* **382**, 693 (2007).
5. K. A. Seaton and L. M. Hackett, *Physica A* **339**, 635 (2004).
6. M. Marchiori and V. Latora, *Physica A* **285**, 539 (2000).
7. V. Latora and M. Marchiori, *Phys. Rev. Lett.* **87**, 198701 (2001).

8. V. Latora and M. Marchiori, *Physica A* **314**, 109 (2002).
9. J. Sienkiewicz and J. A. Holyst, *Acta Phys. Polon. B* **36**, 1771 (2005).
10. J. Sienkiewicz and J. A. Holyst, *Phys. Rev. E* **72**, 046127 (2005).
11. X. Xu, J. Hu, F. Liu and L. Liu, *Physica A* **374**, 441 (2007).
12. D. J. Watts and S. H. Strogatz, *Nature* **393**, 440 (1998).
13. M. Kurant and P. Thiran, *Phys. Rev. E* **74**, 036114 (2006).
14. M. A. Cardillo, S. Scellato, V. Latora and S. Porta, *Phys. Rev. E* **73**, 066107 (2006).
15. R. Pastor-Satorras, A. Vazquez and A. Vespignani, *Phys. Rev. Lett.* **87**, 258701 (2002).
16. M. E. J. Newman, *Phys. Rev. Lett.* **89**, 208701 (2002).

The Fractal Dimensions of Complex Networks *

GUO Long(郭龙)**, CAI Xu(蔡勖)

Complexity Science Center and Institute of Particle Physics, Huazhong Normal University, Wuhan 430079

(Received 5 June 2009)

It is shown that many real complex networks share the distinctive features, such as the small-world effect and the heterogeneous property of connectivity of vertices, which are different from the random networks and the regular lattice. Although these feathers capture the important characteristic of complex networks, their applicability depends on the style of networks. To unravel the universal characteristic many complex networks have in common, we study the fractal dimensions of complex networks using the method introduced by Shanker. We find that the average 'density' $\langle\rho(r)\rangle$ of complex network follows a better power-law function as a function of distance r with the exponent d_f, which is defined as the fractal dimension, in some real complex networks. Furthermore, we study the relation between d_f and the shortcuts $N_{\rm add}$ in small-world network and the size N in regular lattices. Our present work provides a new perspective to understand the dependence of the fractal dimension d_f on the complex network structure.

PACS: 89.75.Da, 05.45.Df

Recently, complex networks have been studied extensively in interdisciplinary fields including mathematic, statistical physics, computer science, sociology, economics, biology, etc. Complex networks are ubiquitous in the real world, e.g., there are technological networks such as the power grid,[1] biological networks such as the protein interaction networks,[2] and social networks such as scientific collaboration networks,[3,4] and human communication networks,[5] to name a few.

It has been shown that many real complex networks share distinctive characteristic properties that differ in many ways from the random and regular networks. One such property is the "small-world effect",[1] which means that the average shortest path length between vertices in network is short, usually scaling logarithmically with the size N of network, while maintaining high clustering coefficient. A famous example is the so-called "six degrees of separation" in social networks.[6] Another is the scale-free property that many networks possess. The probability distribution of the number of links per node, $P(k)$ (also known as the degree distribution) satisfies a power-law $P(k) \sim k^{-\gamma}$ with the degree exponent γ in the range of $2 < \gamma < 3$.[7] Although these properties capture the important characteristic of complex networks, their applicability depends on the style of networks. With the aim of providing a deeper understanding of the underlying mechanism of these common properties and unravelling the universal characteristics that many complex networks possess, many researchers have studied the self-similarity property and the dimension of complex networks. Song *et al.* discussed the mechanism that generates fractality, i.e., the repulsion between hubs, using the concept of renormalization.[8] In order to unfold the self-similar properties of complex networks, Song *et al.* calculated the fractal dimension using a 'box-counting' method and a 'cluster-growing' method, and found that the box-counting method is a powerful tool for further investigations of network properties.[9] The degree exponent γ can be related to a more fundamental length-scale invariant property, characterized by the box dimension d_B and the renormalized index d_k, as a function of $\gamma = 1 + d_B/d_k$.[9] Kim *et al.*[10,11] studied the skeleton and fractal scaling in complex networks using a new box-covering algorithm that is a modified version of the original algorithm introduced by Song *et al.* What is more, Kim *et al.* discussed the difference of fractality and self-similarity in scale-free networks, which has been helpful for us to understand the complex networks better.[12] Zhou *et al.* proposed an alternative algorithm, based on the edge-covering box counting, to explore self-similarity of complex cellular networks.[13] Furthermore, Lee and Jung studied the statistical self-similar properties of complex networks adopting the clustering coefficient as the probability measure and found that the probability distribution of the clustering coefficient is best characterized by the multifractal.[14] On the other hand, several algorithms have been proposed to calculate the fractal dimension of complex network, such as the box-covering algorithm[15] and the ball-covering approach.[16] Shanker defined the dimension of complex network in terms of the scaling property of the volume, which can be extended from regular lattices to complex networks.[17,18] Nevertheless, understand-

*Supported by the National Natural Science Foundation of China under Grant No 10635020,the Program of Introducing Talents of Discipline to Universities under Grant No B08033, the National Key Basic Research Program of China under Grant No 2008CB317106, and the Key Project of the Ministry of Education of China (306022 and IRT0624).

**Email: longkuo0314@gmail.com;longkuo0314@hotmail.com

ing the self-similar properties of complex networks remains a challenge.

In order to unfold the universal scaling properties of complex networks, we study the fractal dimension of some real complex networks using the dimension measurement algorithm based on the scaling property of the volume in Refs. [17,18]. We find that there exists a universal scaling relation between the average density $\langle\rho(r)\rangle$ and the box linear size r with the exponent d_f. Furthermore, we study the fractal dimension d_f in small-world networks and in the regular lattices. We find that the dependence of the fractal dimension d_f on the average adding shortcuts $N_{\rm add} = Np$ in the NW small-world networks and the size N in the regular lattices.

Generally, we adopt the abstract space, which is different from one-dimensional linear space and two-dimensional flat space, to analyze the characteristics of complex networks, such as the structure of complex networks and the dynamics behavior of and on complex networks. In order to analyze the dimension property of complex networks, we define the distance d_{ij} between two vertices, say i and j, is the shortest path length from vertex i to vertex j. We set all the nodes as the seeds in turn and a cluster of nodes centrad at each seed within the box of the linear size r. Then, the average density $\langle\rho(r)\rangle$, defined as the ratio of the number of nodes in all the boxes with the size r and the complex network size N, is calculated as a function of r to obtain the following scaling:

$$\langle\rho(r)\rangle \simeq kr^{d_f}, \quad (1)$$

where d_f is defined as the fractal dimension of complex network and k is a geometric constant which depends on the complex network. The most important is that the definition of the fractal dimension reduces the fluctuation of the heterogeneous property of connectivity degree of vertices in complex networks, since all the nodes as the seeds in turn during covering complex network. The definition here is different from the box-covering algorithm, where the fractal dimension relation $N(l) \sim l^{-d_B}$ and $N(l)$ is the minimum number of boxes needed to tile a given network. However, to identify the minimum $N(l)$ value for any give l belongs to a family of NP-hard problems.[16]

Table 1. General characteristics of several real networks. For each network we have indicated the type (undirected network or directed network) of complex network, the number of nodes, the average degree $\langle k\rangle$, the average path length l, the clustering coefficient C and the degree distribution $P(k)$. Here empty shows that there is no obvious degree distribution since the size is too small. The various types of networks datasets are obtained from the Pajek datasets (http://vlado.fmf.uni-lj.si/pub/networks/data/).

Network	Type	Size	$\langle k\rangle$	l	C	$P(k)$
Power grid	undirected	4941	2.67	18.7	0.08	$e^{-0.59k}$
C.Elegans	directed	306	7.66	3.97	0.147	
Yeast	directed	2361	2.82	4.62	0.04	$k^{-2.11}$
CNCG	undirected	7343	1.62	3.92	0.103	$k^{-2.17}$
E-mail	directed	1133	9.62	3.606	0.166	$e^{-0.11k}$

We apply the definition of the fractal dimension above mentioned to some real complex networks, e.g., the chemical biology networks such as the protein-protein interaction network (PIN) in budding yeast,[19] the neural network of the nematode worm C.elegans,[1] the social networks such as the email network of University at Rovira i Virgili (URV)[5] and the collaboration network in computational geometry (CNCG), the technological network such as the electrical power grid of the western United States.[1] All those real complex networks are of scientific interest. The PIN in budding yeast plays a key role in predicting the function of uncharacteristic proteins based on the classification of known proteins within topological structures. The C.elegans is an important example of a completely mapped neural networks. The graph of the email network at URV and the graph of CNCG are the surrogates for social networks where the agents interact with others by the means of collaboration and information transition. The graph of the power grid is related to the efficiency and robustness of power networks.[1] Table 1 shows that those real complex networks are sparse ones with the small-world effect and the heterogeneous property of connectivity degree of vertices.

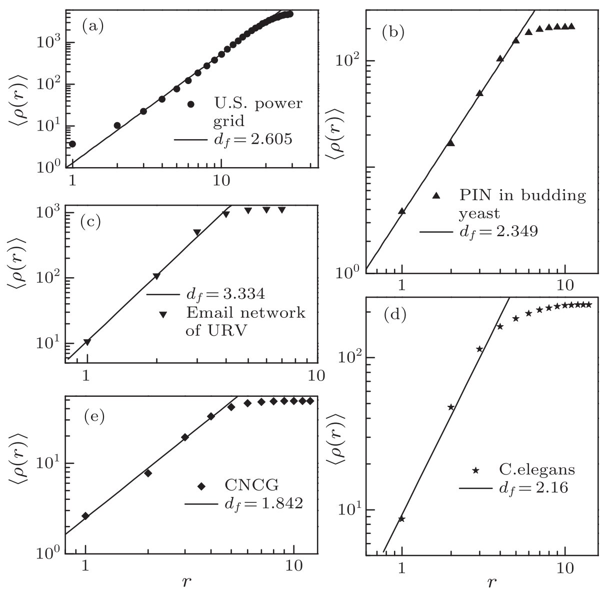


Fig. 1. (color online) The fractal dimensions in some real complex networks. (a) The U.S. power grid with $d_f = 2.286$. (b) The PIN in budding yeast with $d_f = 2.349$. (c) The email network of URV with $d_f = 3.334$. (d) The neural network of C. elegans with $d_f = 2.16$. (e) The CNCG with $d_f = 1.842$. The red solid lines represent the power-law fit for those real complex networks.

Figure 1 displays the evolution of $\langle\rho(r)\rangle$ as a function of r for various real complex networks. We find that $\langle \rho(r) \rangle$ evolves as a scaling function of r with the exponent d_f in all those complex networks. Interestingly, the scaling function is independent of the style of complex networks, which may show the universal scaling property in complex networks. However, the fractal dimensions d_f values are different in those real complex networks, such as the U.S. power grid with $d_f = 2.286$, The PIN in budding yeast with $d_f = 2.349$, the email network of URV with $d_f = 3.334$, the neural network of C. elegans with $d_f = 2.16$ and the CNCG with $d_f = 1.842$, see Fig. 1. The fractal dimension d_f maybe is related to the complex network structure, such as the shortcuts and the size N. Here, the d_f value is different from the d_B value obtained from the box-covering algorithm[12] and the d_{ball} value from the ball-covering approach,[16] because of the different physical quantities in those fractal definition. The average density $\langle \rho(r) \rangle$ of the vertices in the boxes with size r is an exact solution in our present work, and the minimum number $N(l)$ of boxes needed to tile a give network is an approximate solution in the box-covering algorithm. For example, in the C.elegans, $d_f = 2.16$ is smaller than $d_B = 3.5$ and $d_{\rm ball} = 3.7$,[16] respectively.

Further light can be shed on the dependence of the fractal dimension d_f on the complex network structure, such as the shortcuts in small-world network and the complex network size. In order to do this, we study the dimensions of the small-world network and the regular lattice with open boundary condition using the finite-size effect method.

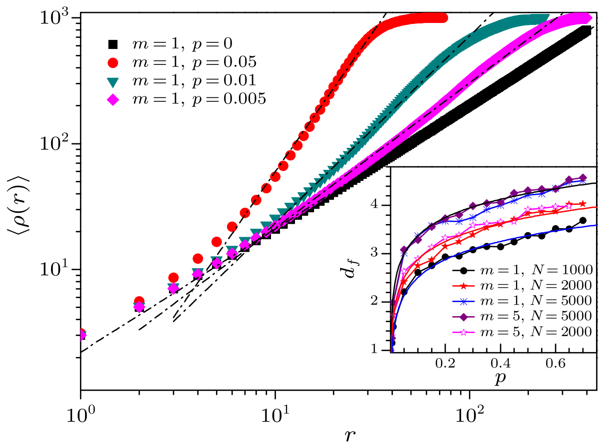


Fig. 2. (color online) The evolution of the density $\langle \rho(r) \rangle$ as a function of the boxes linear size r in NW small-world network with various shortcuts density p. The dot-dashed lines are the fit lines related to various p, respectively. The size of the network is $N = 1000$. Inset: the fractal dimension as a function of p in NW small-world network. The curves satisfy the function of $d_f = 1.25 \log(1 + Np)$ for $p > 0$, where $N_{add} = Np$ is the average number of shortcuts in the NW small-world network.

Here the small-world network is built as the algorithm of the Newman-Watts (NW) small-world network.[21] The NW small-world network is defined on a lattice consisting of N nodes arranged in a ring. Initially each node is connected to all of its neighbors up to some fixed range m to make the network with average coordination number $z = 2m$. Randomness is then introduced by taking each node in turn and, with probability p, adding an edge to a randomly chosen node, so that there are again (Np) shortcuts average. For convenience, we call m the first neighbor parameter (FNP) and p the shortcuts density. Tuning m and p, we can obtain a series of complex networks with different structural properties. This model is equivalent to the Watts-Strogatz model[1] for small p, whilst being better behaved when p becomes comparable to 1.[21]

In Fig. 2, we represent the evolution of the average density $\langle \rho(r) \rangle$ as a function of the box size r with the same FNP $m = 1$ and different shortcuts density p. We find that the relation between $\langle \rho(r) \rangle$ and r satisfies the scaling function as Eq. (1) with the fractal dimension d_f better. Furthermore, we find that $d_f = 0.998 \simeq 1$ for $p = 0$ and $d_f > 1$ for $p > 0$. Here d_f increases as the shortcut density p increases. Namely, the larger the shortcuts density p is, the larger the fractal dimension d_f of NW small-world network is. Hence, the dimension d_f can reflect the disorder degree of complex systems. On the other hand, we study the evolution of $\langle \rho(r) \rangle$ as a function of p using the finite-size effect, see the inset of Fig. 2. We find that the fractal dimension d_f, which is independent of the FNP m, increases as the size N and the shortcut density p of NW small-world network increases. We fit the evolution of d_f as a function of the shortcuts density p and the network size N for $p > 0$ using the nonlinear fitting method, and find that $d_f(N, p)$ satisfies the relation

$$d_f(N, p) = 1.25 \log(1 + Np), \tag{2}$$

where $N_{\rm add} = Np$ is the average number of shortcuts in the NW small-world network.

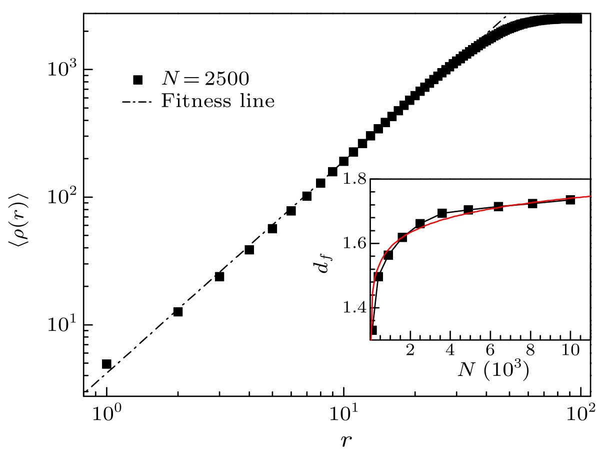


Fig. 3. (color online) The evolution of the density $\langle\rho(r)\rangle$ as a function of the boxes linear size r in the regular lattice with the size $N = 2500$. The dash dot line is the fitness line with the slope $\gamma = 1.649$. Inset: the fractal dimension as a function of the size N in regular network. The red curve satisfies the function of $d_f = 2 - exp(-N^{0.183}/4)$.

What is more, we also study the dimension of the regular lattice with open boundary condition using the above mentioned method. In Fig. 3, we represent the evolution of $\langle\rho(r)\rangle$ as a function of r with the size $N = 2500$. We find that the relation between $\langle\rho(r)\rangle$ and r satisfies the strictly scaling function as Eq. (1) with the exponent $d_f = 1.649$. Surprisingly, the dimension calculated by the above mentioned is not equal to the integer 2. We analyze the dependence of d_f on the size N of the regular lattice using the finite-size effect, since the regular lattice with finite size is embedded in the flat space. We find that the fractal dimension d_f increases as the size N increases, see the inset of Fig. 3. Interestingly, we also fit the function of $d_f(N)$ using the nonlinear fitting method, and find that $d_f(N)$ satisfies the relation

$$d_f(N) = 2 - \exp(-N^{0.183}/4), \qquad (3)$$

where 4 is the connectivity degree that most vertices are in the regular lattice. From Eq. (3), we find that $d_f \to 2$ for $N \to \infty$. Combining $d_f \simeq 1$ for $p = 0$ in the NW small-world network and $d_f \to 2$ for $N \to \infty$ in the regular lattice, we find that the definition of the fractal dimension here can be applied to the regular lattices. Hence, the finite size plays a crucial role in the complex network structure and the dynamics of and on complex networks.[22,23]

In summary, we have studied the fractal dimensions of complex networks using the method introduced by Shanker. We find that the evolution of the average density $\langle\rho(r)\rangle$ as a scaling function of the boxes linear size r in some real complex networks. The scaling property is independent of the style of complex networks and universal, since the calculation of the $\langle\rho(r)\rangle$ is averaged over all the vertices in complex networks in the definition of the fractal dimension. The average density reduces the fluctuation in complex networks. Furthermore, we study the dependence of d_f on the shortcuts (including the size N and the shortcuts density p) in small-world networks and the size N in the regular lattices. Our present work shows the important role of complex network structure in the fractal dimension d_f and provides a new perspective to understand the fractal dimension of complex networks.

References

[1] Watts D J and Strogatz S H *Nature* **393** 440 (1998)
[2] Jeong H, Mason S P, Barabási A L and Oltvai Z N 2001 *Nature* **411** 41
[3] Newman M E J 2001 *Phys. Rev.* E **64** 016131
[4] Newman M E J 2001 *Phys. Rev.* E **64** 016132
[5] Guimerà R, Danon L, Díaz-Guilera A, Giralt F and Arenas A 2003 *Phys. Rev.* E **68** 065103(R)
[6] Milgram S 1967 *Psychol. Today* **2** 60
[7] Albert R, Jeong H and Barabási A L 1999 *Nature* **401** 130
[8] Song C, Havlin S, and Makse H A 2006 *Nature Phys.* **2** 275
[9] Song C, Havlin S and Makse H A 2005 *Nature* **433** 392
[10] Goh K I, Salvi G, Kahng B and Kim D 2006 *Phys. Rev. Lett.* **96** 018701
[11] Kim J S, Goh K-I, Salvi G, Oh E, Kahng B and Kim D 2007 *Phys. Rev.* E **75** 016110
[12] Kim J S, Goh K I, Kahng B and Kim D 2007 *New J. Phys.* **9** 177
[13] Zhou W X, Jiang Z Q and Sornette D 2007 *Physica* A **375** 741
[14] Lee C Y and Jung S 2006 *Phys. Rev.* E **73** 066102
[15] Song C, Gallos L K, Havlin S and Makse H A 2007 *J. Stat. Mech.* **03** P03006
[16] Gao L, Hu Y and Di Z 2008 *Phys. Rev.* E **78** 046109
[17] Shanker O 2007 *Mod. Phys. Lett.* B **21** 321
[18] Shanker O 2007 *Mod. Phys. Lett.* B **21** 639
[19] Bu D et al 2003 *Nucl. Acids Res.* **31** 2443
[20] Gleiser P M and Danon L 2003 *Adv. Complex Syst.* **6** 565
[21] Newman M E J and Watts D J 1999 *Phys. Rev.* E **60** 7332
[22] Toral R and Tessone C J 2007 *Commun. Comput. Phys.* **2** 177
[23] Guo L and Cai X 2009 *Commun. Comput. Phys.* **6** 586