Front. Phys. China, 2007, 2(3): 364－367
DOI 10.1007/s11467-007-0039-y

RESEARCH ARTICLE

LIU Cui-mei, LI Jun-wei

Small- world and the growing properties of the Chinese railway network

Abstract By using the tools of statistical physics and recent investigations of the scaling properties of different complex networks, the structural and evolving properties of the Chinese railway network (CRN) is studied. It has been verified that the CRN has the same small-world properties of the Indian railway network (IRN). According to the class of small-world networks, we believe the CRN is a single scale. In addition, a novel result is obtained. Measurements on the CRN indicate that the rate at which nodes acquire links depends on the node's degree and follows a power law.

Keywords statistic mechanism, small-world network, complex networks, railway networks

PACS numbers 89.65.-s, 02.50.-r, 89.75.Hc

1 Introduction

Many systems take the form of networks—sets of nodes or vertices, connected by links or edges. Chemical-reaction networks [1], neuronal networks [2], food webs [3], and social networks [4] are commonly cited examples. These complex networks have attracted great attention and have been investigated because of their potential as models for the interaction networks of complex systems [5–10]. In recent years, two kinds of models have been constructed and studied, which are small-world networks [5] and scale-free networks [7, 8].

The small-world network (SWN) was proposed by Watts and Strogatz in 1998, which interpolates between regular and random networks. They argued that SWNs must have small diameters, which grow as $\ln N$ like random networks but should have large values of the clustering coefficients $C(N)$~1 like regular networks, where N denotes the number of nodes or vertices in networks. On the other hand, it was found that the degree distributions of scale-free networks follow some power laws: $P(k) \propto k^{-r}$, which emerge in the context of growing networks, where new sites connect preferentially to more highly linked sites. The world-wide web network [7, 8, 11] and internet network [12] were observed to have scale-free property. In general, scale-free networks show small-world properties while a small-world network is not necessarily scale-free.

Since SWNs and scale-free networks were proposed, authors have presented various models so that complex networks have been focused on in many scopes. Meanwhile, some scientists have investigated the structure and properties of real networks by using the tools of statistical physics, such as airport networks [13], navigation networks [15], bus networks [16] and other technique networks, etc.

A railway network is one of the important examples of transport systems. The complex topological structures of railway networks have attracted great attention. Benguigui studied the fractal nature of the structure of railway networks [17]. Small-world properties of the Boston subway [18] and Indian railway network (IRN) [19] have also been investigated.

In this paper, we first study the structure of the Chinese Railway network (CRN) in the light of recent investigations of complex networks. It has been verified that the CRN has the same small-world properties of the IRN. Secondly, we define the CRN as a growing network and study the evolving property of the CRN. The function of the rate of the CRN is obtained, which is distinguished from those of

LIU Cui-mei
Department of Physics and Information Engineering, Shangqiu Teachers College, Shangqiu 476000, China
E-mail: liu-cuimei@163.com

LI Jun-wei (✉)
School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
E-mail: 04114190@bjtu.edu.cn

Received January 20, 2007

scale-free networks.

2 Small-world properties of the CRN

Based on the study on the IRN, some definitions were followed. First, we select a representative graph G_N with the CRN of N stations. Here, the "vertices" of the graph represent the stations, whereas two arbitrary stations are considered to connect by a "link" when there is at least one train that stops at both the stations. If there is a link between two stations, these two stations are considered to be a unit distance irrespective of the geographical distance between them. As a result, the shortest distance l_{ij} between an arbitrary pair of stations s_i and s_j is the minimum number of different trains that one need to change to travel from s_i to s_j. For example, if $l_{ij} = 1$, it indicates that there is at least one train that will stop at both station s_i and s_j. Accordingly, if $l_{ij} = 2$, it shows that one must at least change once by some intermediate station. In accordance with this definition, if the trains $t_1, t_2, \cdots, t_n$ pass through some station s_i, all stations but s_i that n trains pass are a unit distance away from s_i and the number of the stations is considered as the degree of s_i.

Since the main motivation of all transport systems is to be fast and economical, the railways can achieve this. To do this, railways simultaneously run many trains, covering short as well as long routes. Thus, we don't need to change more than only a few trains to reach any arbitrary destination in the country. The CRN is a large space network that includes more than 6400 stations. Every day there are more than 10 000 trains in the network that ply its routes. To investigate the small-world properties of the CRN, we collect the data of CRN on a coarse-grained level following the recent "National Passenger Train Timetable", which was issued in March of 2004 [20]. The timetable contains L = 281 trains covering N = 664 stations, where only fast passenger trains are counted. An important adjacency matrix $T(N, N)$ among stations is constructed. Here, the ij-th element of this matrix is 1 if there is at least one train to stop both station s_i and station s_j, otherwise this element is zero.

The diameter of a graph is the maximal distance between any pair of its nodes in the graph theory. The graph G_N of the CRN is a connected graph so that we can obtain $N(N-1)/2$ shortest paths among the pairs of stations. Since we ignore space distances among the stations, the shortest path lengths are calculated using a burning algorithm similar to the IRN. We obtain that the diameter of the CRN is 4. It implies that one needs to change three trains to reach a station from any station in China. In Fig. 1 we plot the distribution Prob(l) of the shortest path lengths l. Here, there is a peak at l = 2 indicating that we can go to most stations in China by changing trains only once. The mean distance of a network is a measure of how good the connectivity of a network is. We calculate the approximate mean distance $D(N)$ of the CRN as 2.31.

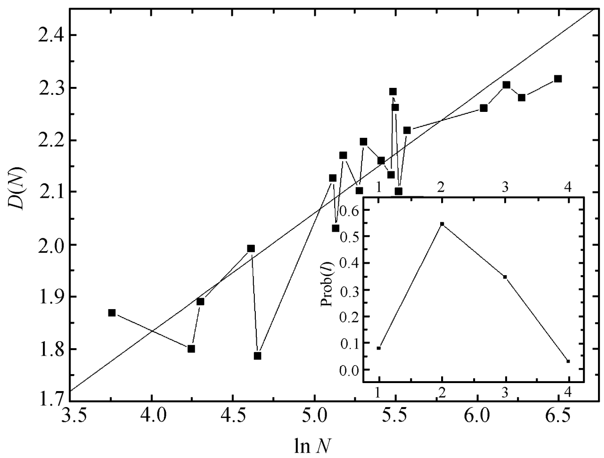

Fig. 1 The mean distance $D(N)$ of 20 different subsets of the CRN vs. ln N. For the result of linear fit, a function like $D(N) = a + b(N)$ is obtained where $a \approx 0.87$ and $b \approx 0.09$. The inset shows the distribution Prob(l) of the shortest path lengths l on the CRN.

In order to verify one of the important properties of small-world networks, how the mean distances $D(N)$ vary with N is investigated. The CRN is divided into 20 different subsets, which include six zones as well as their combinations. As a result we obtain 20 data points, reflecting the nature of variation of $D(N)$ with N. In Fig. 1, we plot this data on a semi-logarithmic scale. For small values of N there is some wild fluctuation, but for large values of N the linear behavior is quite apparent. The whole range fits with $D(N) = a + b \ln N$. Here, we obtain $a \approx 0.87$ and $b \approx 0.09$.

Clustering coefficient indicates the probability that two of these neighbors are connected. According to the clustering coefficient of small-world networks, the clustering coefficient C_i of railway networks is defined as follows. Suppose that a station s_i has k_i neighbors; then at most $k_i(k_i - 1)/2$ edges can exist between them. Let C_i denote the fraction of these allowable edges E_i that actually exist, where $C_i = \frac{2E_i}{k_i(k_i-1)}$. Define C as the average of C_i over all nodes. We obtain $C \approx 0.74$ for the CRN. Compared with a random graph with the same number N = 664 of nodes and average number of edges per node, the clustering coefficient of the CRN is far more than that of the random graph whose C is 0.079. It shows the other important property of small- world networks.

Recently, the clustering coefficients $C(N)$ of scale-free networks that decrease with the network size have been studied. In addition, the clustering coefficients $C(k)$ of the node with degree k have also been investigated. There is an important Barabási-Albert network that predicts $C(k) \propto k^0$ and $C(N) \propto N^{-0.75}$. The $C(k)$ in several networks like the actor, language or world-wide web shows a decrease (apparently a power law decay). However, in the internet network at the router level or power grid network of the Western US, $C(k)$ was found to be constant. In the IRN, $C(k)$ remains constant for small k and shows a logarithmic decay at large values of k. In the paper, for the $C(k)$ in the CRN we obtain the same result of the IRN in Fig. 2. We believe that the $C(N)$ in the CRN should be identical with that in the IRN.

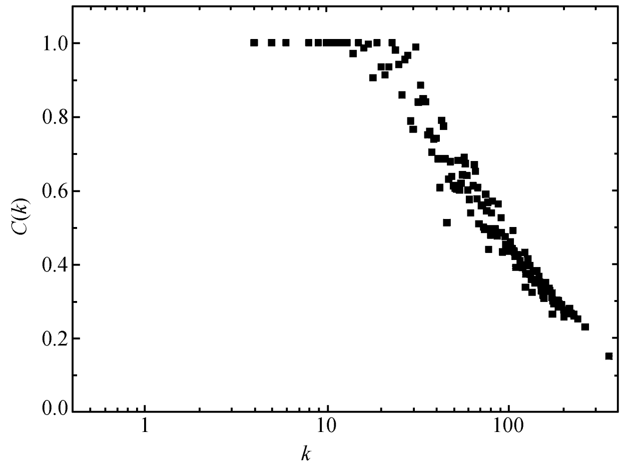

Fig. 2 The variation of the clustering coefficient $C(k)$ against the degree k for the CRN shows a logarithmic decay for large k.

To gain greater insight into the structure of the CRN, we calculate the degree distribution of the stations. Here, the degree distribution of the stations is donated by $P(k)$. We plot the cumulative distribution $F(k)=\int_k^{\infty} P(k)dk$ on a semi- log scale for the whole CRN in Fig. 3. In the plot, the distribution falls on a straight line, indicating an exponential decay of the cumulative distribution $F(k) \sim 10^{-\alpha k}$ with $\alpha = 0.02$. We obtain the same result of the cumulative distribution of the electric power grid of Southern California. According to classes of small-world networks [13], we believe that the CRN is a single scale.

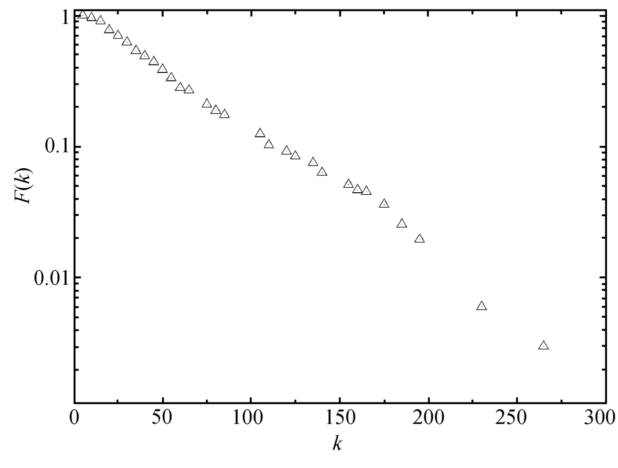

Fig. 3 The cumulative degree distribution $F(k)$ of the CRN with the degree k is plotted on the semi-logarithmic scale.

3 Evolving property of the CRN

When static networks were studied, the dynamics of networks themselves—how and why their topology change over time—also attracted great attention. Most evolving networks are based on two ingredients: growing and preferential attachment. The growth hypothesis suggests that networks continuously expand through the addition of new nodes and links between the nodes, while the preferential attachment states that the likelihood of receiving new edges increases with the node's degree. While most versions of such evolving network modes assume that the rate Πk with which a node with k links acquires new links is linear k and a monotonically increasing function of k, several authors have recently proposed that Πk could follow a power law. The science citation network, Internet, actor collaboration and science co-authorship network have been investigated. For first two networks the attachment rate depends linearly on the node degree, while for the last two the dependence follows a sublinear power law [21].

To measure Πk a numerical method was proposed to extract the functional form Πk directly from dynamical data on real evolving networks [22]. Consider the state of a network at a given time T_0, and record the number of "old" nodes present in the network and their degrees. Next measure the increase of the degree of the "old" nodes over a time interval ΔT, much shorter than the age of the network. Then, according to $\Pi k = \frac{k_i}{\sum_j k_j}$, plotting the relative increase $\Delta k_i/\Delta k$ in the function of the earlier degree k_i for every node gives the Πk function. Here, Δk is the number of edges added to the network in time ΔT. To reduce the fluctuation, the cumulative function is used. Here,

$$\kappa(k) = \sum_{k_i=0}^{k} \Pi k_i \tag{1}$$

Railways like roadways and airways are of crucial importance to the development of a country and indicators of its economic growth. To be fast and economical, new lines are continually constructed while locals are substituted fliers. Thus, we could regard the CRN as an evolving network with time.

Here, there are two questions to be raised. Firstly, is preferential attachment present in the CRN? Secondly, if Πk does indeed depend on k, what is its functional form? In order to answer the two questions, we follow the above numerical method to investigate the CRN. Considering the difficulty of collecting data, this database only contains the data from "National Passenger Train Timetable" of 1992, 1998, and 2004. Here, corresponding the numbers of stations for the above timetables are 487, 555, and 664, where we only count the stations that fliers pass. For the CRN, we choose ΔT = 6 years to consider that stations change less if the time interval is too short. In Fig. 4 the cumulative distributions $\kappa(k)$ for the CRN with k are plotted, where T_0 = 1992 and 1998 are chosen. We obtain $\kappa(k) = A + B \ln k$ and $\Pi k \propto 1/k$. Note that, apart from statistical fluctuations the function form of Πk is independent of T_0. The study indicates that the Πk in the CRN follow preferential attachment, which is a property of scale-free networks.

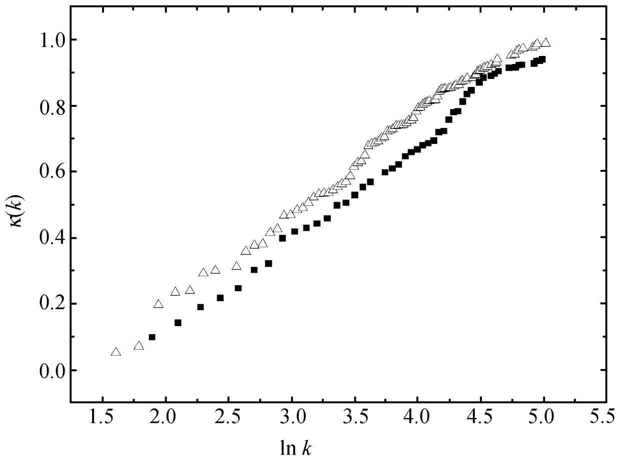

Fig. 4 The $\kappa(k)$-function determined numerically for the CRN. The top to the bottom correspond to measurements made at T_0 = 1992 and 1998.

4 Conclusion

The CRN has been verified as having the same small-world properties of the IRN. Here, the CRN has a small diameter but large clustering coefficient. Investigating the cumulative degree distribution of the CRN, we think that it should be a single-scale network. On the other hand, the evolving property of the CRN shows that the function Πk of the CRN follows a power law similar to scale-free networks. We believe that the CRN is scale-free. We are currently unable to study other real networks but it is believed that the obtained function Πk may have universality to a certain extent.

Acknowledgements This work was supported by the Key Project of National Natural Science Foundation of China (Grant No. 60572001).

References

1. Alon U., Surette M. G., Barkai N., and Leibler, Nature,1999, 379: 168
2. Pimm S. L., Lawton J. H, and Cohen J. E, Nature, 1991, 350: 669
3. Paine R. T, Nature, 1992, 355: 73
4. McCann K., Hastings A., and Huxel G. R., Nature, 1998, 395: 794
5. Watts D. J. and Strogatz D. H., Nature, 1998, 393: 440
6. Barthélémy M. and Amaral L. A. N., Phys. Rev. lett., 1999, 82: 3180
7. Albert R., Jeong H., and Barabási A. -L , Nature,1999, 401: 130
8. Barabási A. -L. and Albert R., Science, 1999, 286: 509
9. Newman M. E. J., Proc. Nat. Acad. Sci. (USA), 2001, 98: 404
10. Lago-Fernandez L. F., Hunerta R., Corbacho F., and Siguenza J. A., Phys. Rev. lett., 2002, 84: 2758
11. Lawrence S. and Giles C. L., Science, 1998, 280: 98
12. Faloutsos M., Faloutsos P., and Faloutsos C., Proc. ACM SIGCO MM, Comput. Commum. Rev., 1999, 29: 251
13. Amaral L. A. N., Scala A., Barthélemy M., and Stanley H. E., Proc. Natl. Acad. Sci., 2000, 97: 11149 (USA)
14. Guimerà R., Mossa S., Turtschi A., and Amaral L.A. N., 2003, arXiv: cond-mat/0312535
15. Kleinberg J. M., Nature, 2004, 406: 845
16. Wu J .J., Gao Z.Y., Sun H. T., and Huang H. J., Mod. Phys. Letts. B, 2004, 18: 1042
17. Benguigui L. and Daoud M., Geographical Analysis, 1991, 23: 362
18. Latora V. and Marchiori M., Physica A , 2002, 314: 109
19. Sen P., Dasgupta S., Chatterjee A., Sreeram P. A., Mukherjee G., and Manna S. S., Phys. Rev. E, 2003, 67: 036106
20. National Passenger Train Timetable, Chinese Raiway Press, 2004
21. Jeng H., Néda Z., and Barabási A. -L., Europhys Lett., 2003, 61(4): 567–572
22. Albert R. and Barabási A. -L., arXiv:cond-mat/0106096 (33–34)