Spatial Structural Characteristics of Chinese Railway Passenger Network Based on Complex Network Theory

Xuejing GU

School of Traffic and Transportation,
Beijing Jiaotong University, Beijing, China
07251035@bjtu.edu.cn

Dewei LI

State Key Laboratory of Rail Traffic Control
and Safety of China, Beijing Jiaotong
University, Beijing, China

Lu QIN

School of Traffic and Transportation, Beijing
Jiaotong University, Beijing, China

Abstract—**The railway network provides transport services for clients, it is very important for railway managers to analyzing the spatial structure of the network. This article analyzes the spatial structure of Chinese railway network with indicators of degree distribution, the clustering coefficient, the average path length and centralization. According to the definition of the small world network, the results illustrate that Chinese railway passenger network shows the characteristic of a small world network and is evolving to a scale-free network. What's more, the top level of the network has been formed.**

Keywords-raiway passenger transportation; spatial structure; complex network; China

I. Introduction

Railway is the artery of national economy and a popular traffic tool. It influences the development of national economy. The growth of railway passenger network promotes the development of Chinese traffic hub cities, urban concentration areas and the traffic economic belts. At present, Chinese railway passengers own about 4328 stations, which are responsible for about 34.7% of the passenger turnover. A wide range railway network, which is extended in all directions, has been formed, and is playing an irreplaceable role in the social economy development.

A variety of railway network models have been constructed to study the various properties of the system. Latora and Marchior (2002) [1] conducted a preliminary study on the Boston subway network (MBTA) of the characteristics. Through the study Sen (2002) [2] illustrated that the Indian Railway network (IRN) has the characteristics of a small world network. India's rail traffic network (IRN) was also studied by Sen P, Dasgupta P, Chatterjee A (2003) [3]. By constructing O-D network of the city bus, ziyou Gao and jianJun Wu etc [4] studied the scale-free characteristics and the degree distribution. Wei Zhao, hongsheng He and zhongcai Lin (2006) [5] thought that two different ways can be used to describe the Chinese railway system. They are geographic network and vehicle drift. Bo Zhou (2008) [6] pointed out that the key issue was to abstract individuals and their complex interactions into nodes and edges. Yue Zhao, wen Du, Shuang Chen (2009) [7] introduced the basic concepts of complex networks and gave a detailed analysis of the characteristics of urban transportation networks.

The complex network theory has gradually become a research trend in many areas, especially in the field of transportation. But literatures on overall characteristics and empirical researches of railway passenger network are still few. It is important to research major cities, hub locations, and the distribution and other issues. Because it can promote the economic advantages of railway technology, gain more market shares and ascertain a reasonable position in the future transport market, thus working out reasonable operation plans in the competitive market economy environment.

II. Model and Data

A Complex Network Models

A large number of systems in nature world can be described by complex networks whose characteristics are self-organizing, self-similarity, attractor, small world or scale-free. Most systems which are interconnected can be regarded as networks [8] under a certain abstract. From the perspective of complex networks there are three structure models of networks which are character networks, random networks and complex networks. Complex network models include the small world network and the scale-free network. The comparison of the indicators (the clustering coefficient and the average path length) of network models is shown in Table 1.

Table 1. The Comparison of the Two Indicators of Network Models [9]

network models	the clustering coefficient	average path length
regular networks	large	large
random networks	small	small
small world networks	large	small
scale-free networks	-	small
real networks	large	small

The "-" in the table means that the result relies on a specific network, the conclusion is not uniform.

B Construction and Description of Chinese Railway Passenger Network

Take the city as node, the relationship of railway passenger transportation between cities as edge [10], the passenger flow

of a city as point right, the number of railway passengers between cities as edge rights. Thus the railway passenger network can be abstracted as a complicated weighted directed network.

If there is a train bounding for city B from city A, then there is a train bounding for city A from city B too [9] [11]. That is, between them the passenger relationship is symmetrical. So the topology structure of Chinese railway passenger network can be got.

C   Data Source and Explanation

The data are from railway ticket database on February 15th, 2007 and selected mainly based on the following requirements:

- The single station must send more than 2000 passengers a day during the Spring Festival.
- The level of the station is higher than three-class.
- The station is located in the provincial capital or important economic city in China.

Special cases are also considered in south southwest and northwest of China such as Aladdin Su station and Korla station where the stations are few. Finally, 203 stations are selected and they undertake 52% of the passenger flow on that day.

Here are the instructions:

- The direct passenger flow is considered while the transfers are not included.
- Between the cities, take the notes which owe direct transportation relationships as the target. For a city with two or more stations, Shenyang for example, the data will be merged and 185 cities are got in this way at last.
- Mainly focus on the topology network, while the weighted network is used to research the passenger flow.

In view of the geographical distribution of the selected 203 stations, most of the stations which belong to 185 cities are in China's prefecture-level cities. From the scale point of view, nearly all of the passengers are form the first-class stations and above, some provincial capitals and tourist sites are also covered. Therefore, although the rate of the passenger flow of the selected stations in the total road is not great, the data can reflect the real laws of the passenger flow with good representation.

III. Network Indicators Analysis

In the Chinese railway passenger network, here are the indicators which are mainly discussed in this article: degree and degree distribution, clustering coefficient, average path length and centralization based on complex network theory.

A   Degree and Degree Distribution

The importance of the city can be measured by the degree between two nodes. According to the entire railway passenger network, the average nodal degree is 66.5. Therefore, this paper defines the nodes whose degree less than or equal to 66 as low-degree nodes. The others are high-degree nodes.
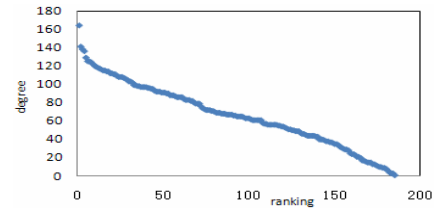

Figure 1. Distribution of nodal degree in Chinese railway passenger network

In the weighted network, Beijing sends the largest passengers, followed by Guangzhou, Shanghai, Chengdu, xi'an. Shanghai has the largest arriving passenger volume, followed by Beijing, Wuchang, Fuyang, Hengyang. The rank of the out-degrees and in-degrees are showed in figure 2 and 3.
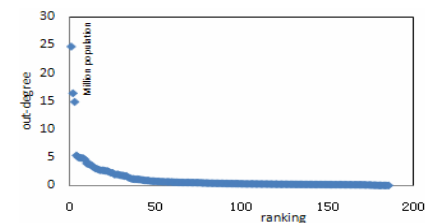

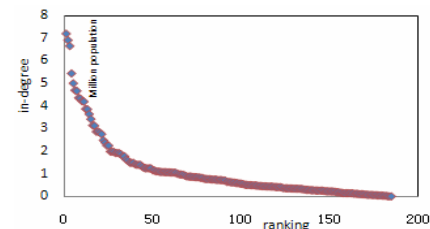Figure 2. Order of out-degrees in Chinese railway passenger network

Figure 3. Order of in-degrees in Chinese railway passenger network

The "output and input of railway passenger network" and the nodal degree reflect the central city of the network. Beijing's degree is the largest, and so is its total passenger flow. From that perspective, Beijing is in the central position of the network.

As can be seen from figure 4 and figure 5, although the degree distribution doesn't presented the typical power-law distribution, passenger out-degrees and in-degrees are various across 5 or 4 orders of magnitude respectively. It is an extreme "authoritarian" network [12]. The network has the similar properties with power-law distribution Therefore Chinese railway passenger network displays scale-free properties.
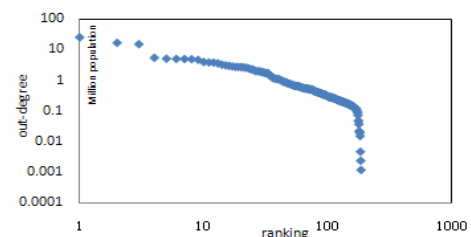

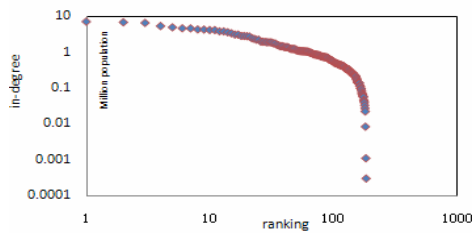Figure 4. Order of out-degrees in double logarithm coordinates

Figure 5. Order of in-degrees in double logarithm coordinates

According to the characteristics of the economical geography, most low-degree nodes are in the northeast, northwest, southwest and south coastal areas. A few cities whose nodal degrees are greater than 100 belong to the highly developed provinces. They are in the northern and eastern coastal economic zones, such as Hebei, Zhejiang, Jiangsu and other places. The four municipalities are all high degree cities. All capital cities except Hefei and Nanjing have the largest degree of their provinces, which is directly related to their economic development level.

According to the analysis of space, the nodal degree distribution shows a clear jump and discontinuity. Nodal degrees show a greater difference in east and west. High degree cities are concentrated on Beijing, Tianjin, and the Yangtze River Delta and Pearl River Delta region. Overall, the low degree nodes are distributed in all provinces.
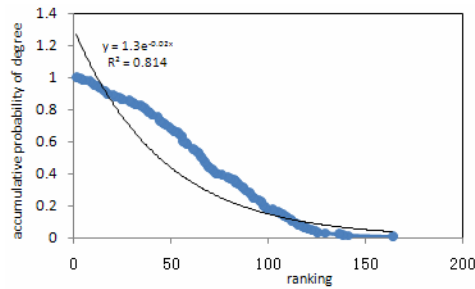

Figure 6. The accumulative degree distribution of the network accumulative probability of degree

According to the accumulative degree distribution in figure 6, the network, whose accumulative degree distribution function is $P(k) = 1.3e^{-0.02k}$, obeys exponential distribution and the confidence level ($R^2=0.814$) is high. In the network, a few nodes connect most of the cities, and only 5.4% of the nodal degrees are higher than 120. The majority of nodal degrees are between 20 and 100, which account for 70.8%. On the whole, the continuous condition of nodal degrees is good, and breakpoints seldom appear. About 20% of the cities undertake 60% of the passenger flow.

In conclusion, Chinese railway passengers mainly concentrate in Beijing, Guangzhou, Shanghai, Xi'an, Chengdu, Hangzhou, and Shenyang. They all are of great importance in railway network.

B   Clustering Coefficient

The clustering coefficient of Chinese railway passenger network is 0.689, which is larger than the same scale random network, whose clustering coefficient is 0.359. On the other hand, it is smaller than the same scale regular network, whose clustering coefficient is 0.995. This shows that Chinese railway passenger network nodes have a good probability to form short distance contacts with each other. According to the research, there are two types of clustering coefficient for low-degree nodes: very large or small. Most high-degree nodes show smaller clustering coefficients, Beijing and shanghai for example.

According to the analysis of space, the clustering coefficients also show a great difference. High-clustering coefficient nodes are distributed in all areas, but the clustering coefficients are various in the same region. In short, the clustering coefficients of Chinese railway passenger network are gathered compared with the degrees.

C   Reachability Analysis

The reachability of nodes in the network can be described by the average path length. For rail network, the shorter the average path length is, the easier passengers can reach their destinations. Table 2 summarizes the results of the shortest paths analysis.

Table 2. Shortest path statistics of the network [9]

Shortest path	No. of paths	No. of transit	Percentage of routes	Cumulative percentage
1	12318	0	0.3619	0.3619
2	20428	1	0.6001	0.9620
3	1292	2	0.0379	0.9999
4	2	3	0.0001	1.0000

Apart from other statistics, it also shows number of trains needed to be changed, which is an indicator of the convenience of travel in the network. Around 60% of paths are reachable by changing a maximum of 1 train. About 36% of the paths are connected by direct trains. The diameter of the network is 4, which means that one needs to change at most 3 trains to reach from any station to any other station.

The average path length of this network is 1.676. It is slightly larger than the same scale random network (lnN/ln<k> ≈1.244) and smaller than the regular network, whose average path length is 1.989. According to the definition [13] that the clustering coefficient of a small world network is larger than the same scale random network and the average path length is smaller than the same scale regular network, the railway passenger network is a typical small world network. By using low level of costs to meet the needs of passengers, the small world network has become the evolution trend.

According to analysis of spatial distribution, there are still large differences between reachabilities. In contrast to the west and northeast, accessibilities are higher in eastern and central regions. Accessibilities of Alashankou, Haikou and Nantong are the worst for the reasons below: the external contacts are less; the city, which has only one contact to the outside world, is a border node. Therefore, Beijing, Shanghai, Zhengzhou are centers of Chinese railway passenger network and Zhengzhou,

Xuzhou, Nanjing, Jinan, Shenyang, Chengdu and Xi'an have a leading position in accessibility.

D   Centrality in the Network

The centrality analysis will help understand the important nodes in network.

Most of the top cities are superposed by analyzing 5 centrality indicators which are Degree Centrality (53.52%), Closeness Centrality (59.21%), Eigenvector Centrality (11.39%), Betweenness Centrality (4.08%) and Flow Betweenness Centrality (1.182%). The superposed cities are Beijing, Shanghai, Zhengzhou, Xuzhou, Nanjing, Xi'an, Jinan, Chengdu, Shenyang and Shijiazhuang. It can deduce that the 10 cities play a very important role in the network and they are the central cities of Chinese railway passenger network.

IV. Conclusions and Outlook

A   The Railway Passenger Network Presents a Small World Network

For a common network, the most important indicators are the average path length and the clustering coefficient. Compared to other networks, Chinese railway passenger network owns a lower average path length and larger clustering coefficient. It is a small world network. While the accumulative degree distribution obeys exponential distribution, the network has been evolving to a combined model of scale-free and small world networks, and its spatial structure will be more complex.

B   Hierarchy Structure of the Network is Not Obvious

According to the centrality of the network, nodal degree, clustering coefficient and accessibility, the top structure of Chinese railway passenger network has been formed. They are Beijing, Zhengzhou, Shijiazhuang, Shanghai and so on, which mainly distribute in northern and eastern China where the economy develops well. But the next level is not obvious, the network structure is still in the process of development, and increasingly appears as the complex network.

C   Outlook

The article presents overall properties of railway passenger network in China. In the real network, however, the number of trains operating, vehicle levels and travel fares as well as transportation capacity are the factors which influence the accuracy. In order to fully grasp the structure of the network, more data should be collected and the directional and weight of the network should be taken into account.

Acknowledgment

References

[1]   V. Latora, M. Marchiori, "Is the Boston subway a small-world network?" Physic A. 314(1-4), pp. 109-113, 2002.
[2]   P. Sen, "Small-world properties of the Indian railway network.," Pllys.Rev.E.,67:36, pp, .106, 2003.
[3]   P. Sen, P. Dasgupta, A. Chatterjee et al Phys.Rev.E67 036106, 2003.
[4]   Ziyou GAO, Jianjun WU, Baohua MAO, Haijun HUANG, "Study on the complexity of traffic networks and related problems," Trans. Sys. Engin. Inf, vol. 5, 2005.
[5]   Wei ZHAO , Hongsheng HE, Zhongcai LIN, "The characteristics of Chinese rail passenger network," Phy. Jou, vol 55, 2006.
[6]   Bo ZHOU, "The forecast of rail freight volume based on complex network theory," Rail Freight.Jou, vol. 3, 2008.
[7]   Yue ZHAO, Wen DU, Shuang CHEN, "Application of complex network theory in urban transport," Urban. Trans. Vol. 7, 2009.
[8]   Tingting YE, "The connectivity reliability analysis of the national railway network based on complex network theory," D, 2009
[9]   Jiaoe WANG, Huihui MO, Fengjun JIN, "Spatial structural characteristics of Chinese aviation network based on complex network theory," Geo. Jou.vol.64,No.8, pp. 899-910, 2009.
[10] Feng GAO, Yaru DANG, "The distribution of nodal degree in aviation network of the world," Sci. Tech. Management, vol. 30, 2009
[11] Hongkun LIU, Tao ZHOU, "Empirical study of Chinese city airline network," Phi. Jou, vol. 56 ,2007.
[12] Haibo HU, Lin WANG, "A brief history of the power-law distribution," Phy, vol. 34, 2005.
[13] Lei SHEN, xin ZHANG, "The progress of small world networks and its study in transportation," pp.70-76, 2005.