Reliability analysis of highway network in China based on complex network theory

Hui-di Jia¹*, Qian Li², Guang-Jing Zhou¹ and Ran Zuo¹

¹ School of Traffic and Transportation, Beijing Jiaotong University, Beijing, Province, 100000, China
² Integrated Transportation Research Center, China Academy of Transportation Sciences, Beijing, 100000, China
*Corresponding author’s e-mail: huidijia@fox.mail

Abstract. In this paper, complex network theory is introduced to analyze the Chinese highway network. 72 cities were extracted from highway information in China and topological network was established based on L space. Then degree distribution, betweenness, clustering coefficient and other network characteristics of Chinese highway network were analyzed. Through computer simulation, the reliability of the Chinese highway network was carried out. The results show that the highway network in China is relatively complete, the connection between cities is convenient.

1. Introduction

In the past 10 years, network research has gradually attracted people's interest. And complex network theory is applied to many subjects, such as physics [1], biology [2] and economics [3]. Transportation network supports the movement of people, goods and information, so for us to know its structure and performance is critical. Wang uses a complex network approach to examine the network structure and nodal centrality of individual cities in the air transport network of China (ATNC) [4]. Dai and Liu capture the main topological and spatial changes of Asian air transport network (SAAN) [5]. Mohmand and Wang contributes a weighted complex network analysis of travel routes on the national highway network of Pakistan [6].

Zhang and Peng modelled and simulated the Taian city bus lines and concluded that the Taian city network had small-world character [7]. Wang and Li study the public transport network of Shenyang as a complex network. On the basis of a large number of statistical data, the characteristics of complex network in Nanjing are researched through Space L and Space P methods [8]. Xu and Zu studies the public transport networks of 330 cities in China based on the space P concept [9]. Li, Zhu and Zhou study the urban public transport network (UPTN) in Nanjing, and the empirical data indicates that it is a small-world network [10]. Gu, Li and Qin analyzes the spatial structure of Chinese railway network with indicators of degree distribution, the clustering coefficient, the average path length and centralization [11].

In recent years, China's expressways have been developing faster and faster. By the end of 2018, the total length of expressways in China had exceeded 140,000 kilometers, ranking first in the world. However, few studies have analyzed high-speed networks at the national level. To this end, this paper attempts to analyze the characteristics of the Chinese highway network based on complex network theory.
The innovation of this paper lies in the establishment of the highway network topology of China’s transportation hub cities and the analysis and evaluation of the highway network based on the complex network theory.

The rest of this paper is organized as follows. Section 2 discusses the data of this study and Section 3 discusses the network construction methodology. In Section 4, we will analyze the topological properties of Chinese highway network systematically. Finally, the conclusions are given in Section 5.

2. Methodology
The methodology of the study basically includes various steps, such as (1) data collection of Chinese highway network; (2) network construction of Chinese highway network; (3) analysis of Chinese highway network. And (3) is comprised of three major components: degree and betweenness; clustering coefficient and distance distribution; network reliability.

2.1. Degree and betweenness
Node degree refers to the number of edges associated with the node. It is also known as correlation degree. The larger the degree of a node, the more important it is. So the importance of the city can be measured by the degree between two nodes. The mathematical expression is as equation (1):

\[ k_{v_i} = \sum_{j} x_{ij} \]  

(1)

Among them, N is the number of total nodes, \( k_{v_i} \) represent the degree of node \( v_i \). If \( v_i \) is connected with \( v_j \), then \( x_{ij} = 1 \), or \( x_{ij} = 0 \).

By counting how many nodes have each degree, we form the degree distribution \( p_{deg}(k) \), defined by the fraction of nodes in the graph with degree \( k \).

The node betweenness is defined as the proportion of the total number of shortest paths in the network to the total number of shortest paths. The mathematical expression is as equation (2):

\[ B_{v_i} = \sum_{v_j \neq v_i \neq v_l} \frac{N_{v_j,v_l}(v_i)}{N_{v_j,v_l}} \]  

(2)

Where, \( v_j \) and \( v_i \) are two nodes in the network that are not adjacent. \( N_{v_j,v_l}(v_i) \) is the number of the shortest path between node \( v_j \) and \( v_l \) passing through the node \( v_i \). \( N_{v_j,v_l} \) is the total number of minimum paths between nodes \( v_j \) and \( v_l \).

The nodes with a large number of media are often the key nodes of the optimal transport route of cargo or the optimal route of passenger journey.

2.2. Clustering coefficient and distance distribution
The clustering coefficient and distance distribution are important parameters to describe the characteristics of a network.

The coefficient of agglomeration is used to describe the degree of clustering between the vertices in a graph. Specifically, the degree to which the adjacent points of a point are connected to each other. The clustering coefficient \( C \) of a network is defined as the average of the clustering coefficient \( C_i \) of all nodes in the network. \( C_i \) is defined as follow:

\[ C_{v_i} = \frac{2e_{v_i}}{s_{v_i}(s_{v_i}-1)} \]  

(3)

Among them, \( e_i \) is the number of edges that actually exist between the \( s_{v_i} \) neighbor nodes of node \( v_i \). The denominator is the largest possible number of edges. The higher the average aggregation coefficient, the closer the connection between nodes is.
The distance $d_{ij}$ between node $v_i$ and node $v_j$ is defined as the number of edges contained in the shortest path connecting these two nodes. The average path length $L$ of the network is defined as the average of the distance between any two nodes. $L$ is defined as follow:

$$ L = \frac{2}{N(N-1)} \sum_{1 \leq i < j \leq N} d_{ij} \quad (4) $$

Among them, $N$ is the number of total nodes. The distance between any two nodes reflects the minimum number of points or edges needed to get from one point to the other. The smaller the distance, the better the connectivity.

2.3. Network reliability

There are many ways to measure network reliability and damage resistance, Xing, Lu and Chen took the Shanghai Metro Network (SMN) as an example to investigate vulnerability of a weighted metro network in responding to random failures as well as malicious attacks [12]. Hui, Cai and Rui studied the random failure and selective attack of Lanzhou public transport complex networks [13]. Ruan and He proposes a CA-SVM based Monte Carlo approach to evaluate network reliability [14].

In this paper, the global efficiency of the network is used to measure the reliability of CHN. The global efficiency of the network is defined as follows:

$$ E = \frac{2}{N(N-1)} \sum_{1 \leq i < j \leq N} \frac{1}{d_{ij}} \quad (5) $$

3. Network construction of the highway network

In this study, according to the 13th five-year plan for the development of the modern comprehensive transportation system, we choose representative 72 cities to evaluate the highway network. The data is collected from Baidu map.

Generally, traffic network can be defined as P space, R space, L space and C space. This study is based on L space. In the CHN, a city is considered as a node, and an edge represents a connected route between two cities. Highways are bidirectional, so we consider CHN as an undirected network. The typical topology of CHN is shown in Figure 1.

![Figure 1. Topological structure of Chinese highway network.](image-url)

4. Characteristics of highway network in China

4.1. Degree and betweenness
Figure 2 shows the degree distribution of the Chinese highway network. According to the entire highway network, the average node degree is 12.2. There are 49% of nodes with degree value bigger than 12.2. Zhengzhou and Jiujiang have the highest degrees, which 23. The degree value of most nodes is concentrated between 10 and 15, and a few nodes have lower or higher degree value, that is, Chinese highway network is relatively complete.

![Figure 2. Degree distribution.](image)

In order to measure the importance of each station node in the Chinese highway network more comprehensively, the node betweenness is calculated in this paper.

Table 1 shows the top 15 stations in degree and medium betweenness. We find there is a certain correlation between the betweenness and degree, and the correlation coefficient of the two is 0.59. In other words, the betweenness of nodes with low degree value is also relatively lower, while the betweenness of nodes with high degree value is often higher.

| City      | Degree | City             | Degree |
|-----------|--------|------------------|--------|
| Zhengzhou | 23     | Wuhan            | 0.3145 |
| Jiujiang  | 23     | Zhengzhou        | 0.299  |
| Changsha  | 22     | Shijiangzhuang   | 0.2823 |
| Ganzhou   | 22     | Nanjing          | 0.2639 |
| Hefei     | 20     | Changsha         | 0.2636 |
| Chongqing | 19     | Nanchang         | 0.2274 |
| Xian      | 19     | Jinan            | 0.2153 |
| Jinan     | 19     | Guangzhou        | 0.184  |
| Guiyang   | 19     | Shenyang         | 0.1647 |
| Nanjing   | 18     | Beijing          | 0.161  |
| Nanchang  | 18     | Chongqing        | 0.1311 |
| Shangrao  | 18     | Lanzhou          | 0.116  |
| Tianjin   | 17     | Tianjin          | 0.1094 |
| Luoyang   | 17     | Fuzhou           | 0.1076 |
| Guilin    | 17     | Xian             | 0.0773 |

4.2. Clustering coefficient and distance distribution

The cumulative clustering coefficient frequency diagram of the Chinese highway network is shown in figure 3. More than 67% of the node clustering coefficient over 0.5, the average clustering coefficient is 0.55, showing high aggregation, this shows that the cities in Chinese highway network have establish close contact, and with the increasingly highway construction, the network average clustering coefficient will increase. This means that the relationship between node will be more and more closely.
By studying the node distance of the Chinese highway network, we found that the average path length is 2.48. This means that it takes an average of 2-3 cities to get from one city to another, which is a smaller average path length. Since the network has a high clustering coefficient and a short average path length, the Chinese highway network constructed in this paper is a small-world network.

Table 2 lists the top 20 nodes with the longest average shortest path length of all nodes. Due to their geographical location on the edge of China's territory, these nodes are less connected to other cities and are far away from other cities.

Table 2. The first 20 nodes with the longest average shortest path length.

| Station  | Length  | Station  | Length  |
|----------|---------|----------|---------|
| Sanya    | 4.507042| Shenyang | 2.760563|
| Haikou   | 3.535211| Xining   | 2.746479|
| Dalian   | 3.492958| Yantai   | 2.670563|
| Qiqihaer | 3.450704| Quanzhou | 2.647887|
| Jilin    | 3.422535| Ningbo   | 2.613803|
| Keshen   | 3.408451| Baotou   | 2.591549|
| Kuerle   | 3.140845| Shantou  | 2.591549|
| Yingkou  | 2.971831| Qinhuangdao | 2.591549|
| Haerbin  | 2.887324| Zhenjiang| 2.591549|
| Lasa     | 2.774648| Qujing   | 2.591549|
| Sanya    | 4.507042| Shenyang | 2.760563|
| Haikou   | 3.535211| Xining   | 2.746479|
| Dalian   | 3.492958| Yantai   | 2.670563|
| Qiqihaer | 3.450704| Quanzhou | 2.647887|
| Jilin    | 3.422535| Ningbo   | 2.633803|

4.3. Network reliability

Through computer simulation, random attacks and deliberate attacks were conducted on the nodes of the Chinese highway network, in which the deliberate attacks of the nodes were carried out according to the maximum degree of node attack and the maximum number of intermediate node attack.

According to the figure 4 and figure 5, we can find that when attacking nodes, deliberate attacks are more destructive than random attacks, but the difference is not obvious. One reason is that the Chinese highway network is multi-core, and even a deliberate attack would not immediately bring down the global network. On the other hand, the network has a large connection coverage, so the connection path between cities will not be completely destroyed because of one or two attacks.
5. Conclusion
In this paper, based on the research method of complex network, the network characteristics of the established Chinese highway network is analyzed. It is found that the Chinese highway network is a small-world network. The average clustering coefficient is 0.55 and the average path length is 2.48. The betweenness is positively correlated to the degree. In the face of attack, deliberate attack is more destructive than random attack, but the difference is not obvious. It indicates that Chinese highway network connection coverage is relatively large.
A sound transportation network can promote economic development and strengthen exchanges between regions. In this paper, the analysis of the structure characteristics and reliability of the highway network is helpful to the optimization of highway transportation. This also provides a basis for the further improvement and maintenance of the expressway network.

Acknowledgments
The research is supported by China Academy of Transportation Sciences (Project ID: 20196661184). The author wishes to thank the agency.

References
[1] E. Cervi, A.; Cammi, E. Zio. (2019) A new approach for nuclear reactor analysis based on complex network theory. Progress in Nuclear Energy, 112:96-106.
[2] K McGarry.; S McDonald. (2018) Complex network theory for the identification and assessment of candidate protein targets. Computers in Biology and Medicine, 97:113-123.
[3] G Cao, Q Zhang, Q Li. (2017) Causal relationship between the global foreign exchange market based on complex networks and entropy theory. Chaos, Solitons & Fractals, 99:36-44.
[4] Wang J, Mo H, Wang F, et al. (2011) Exploring the network structure and nodal centrality of China’s air transport network: A complex network approach. Journal of Transport Geography, 19(4):0-721.
[5] Dai L, Derudder B, Liu X. (2018) The evolving structure of the Southeast Asian air transport network through the lens of complex networks, 1979–2012. Journal of Transport Geography, 68:67-77.
[6] Mohmand Y T, Wang A. (2013) Weighted Complex Network Analysis of Pakistan Highways. Discrete Dynamics in Nature and Society, 2013:1-5.
[7] Zhang L, Peng L, Zhang Y, et al. Urban Bus Transport Network Optimization from Complex Network. In: 2012 Proceedings of International Conference on Modelling, Identification and Control, Wuhan, Hubei, China. pp. 1159-1164.
[8] Wang D, Li X. (2013) Analysis of the Public Transport Network of Shenyang as a Complex Network. Applied Mechanics and Materials, 433-435:1370-1373.
[9] Xu Q, Zu Z, Xu Z, et al. (2013) Space P-Based Empirical Research on Public Transport Complex Networks in 330 Cities of China. Journal of Transportation Systems Engineering and Information Technology, 13(1):193-198.
[10] Li, P., Zhu, Z T., Zhou, J., Ding, J Y., Wang, H W., & Wei, S S.. (2009). Performance analysis of public transport systems in nanjing based on network topology. Complex Sciences, First International Conference, Complex 2009, Shanghai, China. pp. 254-264.
[11] Gu X, Li D, Qin L. (2012) Spatial Structural Characteristics of Chinese Railway Passenger Network Based on Complex Network Theory. In: 2012 National Conference on Information Technology and Computer Science. Lanzhou, China. 750-753.
[12] Xing Y, Lu J, Chen S, et al. (2017) Vulnerability analysis of urban rail transit based on complex network theory: a case study of Shanghai Metro. Public Transport, 9: 501–525.
[13] Hui-Fang F, Cai-Hong L I, Rui W. (2016) Vulnerability Study for Public Transport Network of Valley City:Case of Lanzhou. Journal of Transportation Systems Engineering and Information Technology, 16(1):217-222.
[14] Ruan Y P, He Z. (2013) A CA-SVM Based Monte Carlo Approach for Evaluating Complex Network Reliability[J]. Springer-Verlag Berlin Heidelberg, 19: 609-617