Analysis of China Aviation Network Structure Evolution Based on Ward System Clustering

Yikui Mo, Shulian Li*, Ying Zhao, and Lanfang Wang
College of Civil Engineering, Shenzhen University, Shenzhen City, Guangdong Province, 518060, China
*Corresponding author’s e-mail: Shulian2530@163.com

Abstract. The research uses the method of Ward system clustering. The hierarchical and changing trend of China aviation network structure is studied from the aspects of air transport operation, air transport facilities, urban support and city location. The research shows that China's aviation network is a structure with multiple hubs and axial radiation which is based on the framework of four central cities in Beijing, Shanghai, Guangzhou and Shenzhen. The level of aviation network is positively related to the network efficiency and scale. Therefore, the construction of the airport system in the future should attach importance to optimizing the network and grading system, and rationally guide the evolution of the aviation network structure.

1. Introduction
The airport is the hub of the air transport network. The hub nodes of aviation network can greatly improve the coverage of the entire air transport network through shunting and merging. The study of the aviation network structure can provide the basis for the layout and route planning of the National Airport, optimize the allocation of the resources of the air transport industry, and improve the overall utilization of the air transport network.

The study of hierarchy is relatively early in foreign countries. Simon[1] believes that the stability of hierarchical structure is the reason for the evolution of network to hierarchical state. Eppell and Zwart[2] points out that hierarchy is the key issue in road network planning. At present, qualitative analysis is often used in transport network research, lacking quantitative and objective analysis. At the same time, the establishment of the index is also judged by human subjectivity, which is not scientific. There are few studies on the hierarchy of transport network. Wang Lei[3] establishes the index system of aviation network hub based on multi-attribute decision-making, and classifies airports according to system clustering analysis. Based on the complex network, Xu Shunzhi[4] uses statistical theory analysis and optimization methods to study the evolution and complexity of the route network structure. Combined with the theory of complex network and data mining, Sun Mengdan[5] studies China's aviation network from four aspects of basic structure, cluster analysis, global destruction and destruction based on community structure.

Cluster analysis[6] is a multivariate statistical method to study classification problems. System clustering[7], also known as hierarchical clustering, is one of the most widely used clustering methods. The system clustering method does not depend on the input sequence and initial arrangement of elements. The result of analysis is not affected by the order of cluster process. The clustering results are relatively stable, and the generated clustering sets are relatively regular. The hierarchy of transport network is not predetermined. Considering the internal and external factors that affect the development
and operation of the airport, the research proposes an evaluation index which has an important impact on the comprehensive classification of airports and routes. Then, the research establishes the hierarchy evaluation index system of aviation network, and uses the Ward cluster analysis method to study the hierarchy structure of China's air transport network.

2. Research data and methods

2.1. Zoning index
Based on a large number of literature studies, this paper sets up a hierarchical evaluation system for China aviation network sample nodes by selecting four categories of one-level index of air transport operation, air transport facilities, urban support and city location. The specific influencing factors include the following: air transport operation contains "passenger throughput (10000 people) X1", "cargo throughput capacity (10000 tons) X2", "take-off and landing sorties (sorties) X3", "the terminal area (10000 square meters) X4" represents the level of aviation facilities. Urban support includes "travel population (10000 people) X5", "population (10000 people) X6", "population growth rate (%) X7", "urban non-agricultural population (10000 people) X8", "per capita GDP (yuan) X9", "GDP growth rate (%) X10". The influence of urban location includes "navigable country (one) X11" and "domestic location X12". In addition, we choose the four key indexes of "passenger transport volume (10000 people) Y1", "flight number (class) Y2", "cargo transport volume (ton) Y3" and "passenger seat rate (%) Y4" to establish the China aviation network sample route hierarchy evaluation system.

2.2. Research data
In this paper, the basic data of China aviation network samples are collected from the Statistics of Civil Aviation 2016, China's Statistical Yearbook 2016, 2016 National Economic and Social Development Statistics Bulletin, and Civil Aviation Statistical Bulletin (2005-2016).

2.3. Research methods
The aviation network classification problem is considered as a clustering problem. The hub airports and routes are taken as samples. Their characteristic mapping is considered as a feature vector. Each airport and route are classified as one category. Compute the distance between classes and merge the classes with the greatest similarity. According to the characteristic vector, the hub airports and routes are classified. The data are preprocessed by the method of extreme difference, and the 0-1 matrix of the evaluation index is obtained. There are too many hierarchical evaluation indexes for China's aviation network sample nodes. In this paper, the initial eigenvalue, variance contribution rate and cumulative variance contribution rate of correlation matrix are calculated by factor analysis. The research selects the commonly used Kaiser orthogonal rotation and regression method to calculate the factor score coefficient. Therefore, the key index is screened and the dimension reduction is realized.

3. Hierarchical analysis of sample nodes in China aviation network

3.1. Comprehensive treatment of system cluster analysis index
Twenty-nine airports in China are used as cluster units. SPSS is used for factor analysis of standardized data. The initial eigenvalue, variance contribution rate and cumulative variance contribution rate of the correlation matrix are calculated as shown in Table 1. According to the factor analysis variance contribution rate of the evaluation index, the first four eigenvalues have satisfied the condition that the cumulative contribution rate exceeds 85%, reaching 87.577%. Consider the principle that the eigenvalue is greater than 1, and determine the final number of principal factors. Therefore, the first three eigenvalues are extracted as the main factor. Finally, the score of the comprehensive factor of the sample node index of the aviation network is calculated as shown in Table 2.
Table 1. Aviation network node evaluation index factor analysis variance contribution rate.

| Components | Initial eigenvalue | Total variance of interpretation | Extracting square sum and loading | Rotated square sum loading |
|------------|--------------------|----------------------------------|----------------------------------|---------------------------|
|            | Total Variance     | %                                | Accumulate %                     | Total Variance            | %                                | Accumulate %                     | Total Variance            | %                                | Accumulate %                     |
| 1          | 6.796              | 56.631                           | 56.631                           | 6.796                     | 56.631                           | 56.631                           | 5.828                     | 48.564                           | 48.564                           |
| 2          | 1.689              | 14.075                           | 70.707                           | 1.689                     | 14.075                           | 70.707                           | 2.409                     | 20.076                           | 68.64                           |
| 3          | 1.164              | 9.7                              | 80.407                           | 1.164                     | 9.7                              | 80.407                           | 1.412                     | 11.767                           | 80.407                           |
| 4          | 0.868              | 7.171                            | 87.577                           | -                        | -                                | -                                | 1.036                     | 4.406                           | 4.406                           |
| 5          | 0.584              | 4.863                            | 92.44                            | -                        | -                                | -                                | 0.642                     | 2.954                           | 2.954                           |
| 6          | 0.406              | 3.386                            | 95.826                           | -                        | -                                | -                                | 0.343                     | 1.618                           | 1.618                           |
| 7          | 0.222              | 1.848                            | 97.674                           | -                        | -                                | -                                | 0.198                     | 0.936                           | 0.936                           |
| 8          | 0.1                | 0.835                            | 98.509                           | -                        | -                                | -                                | 0.086                     | 0.413                           | 0.413                           |
| 9          | 0.087              | 0.729                            | 99.238                           | -                        | -                                | -                                | 0.054                     | 0.259                           | 0.259                           |
| 10         | 0.066              | 0.552                            | 99.79                            | -                        | -                                | -                                | 0.049                     | 0.220                           | 0.220                           |
| 11         | 0.024              | 0.196                            | 99.986                           | -                        | -                                | -                                | 0.020                     | 0.091                           | 0.091                           |
| 12         | 0.002              | 0.014                            | 100                              | -                        | -                                | -                                | 0.002                     | 0.009                           | 0.009                           |

Table 2. Score of hierarchical evaluation index factor of China aviation network sample node.

| Airport name         | F1     | F2     | F3     | Airport name         | F1     | F2     | F3     |
|----------------------|--------|--------|--------|----------------------|--------|--------|--------|
| Beijing/Capital      | 3.2752 | -0.5089 | -0.6645 | Urumchi/Diwobao      | 0.0294 | -0.2740 | 0.4910 |
| Guangzhou/Baiyun     | 1.6712 | 0.4281 | 0.3204 | Harbin/Taiping       | -0.5493 | 0.0222 | -0.9363 |
| Shanghai/Pudong      | 2.8639 | -0.4936 | -1.0109 | Tianjin/Binhai       | -0.0684 | 0.3049 | -0.1816 |
| Shenzhen/Baoan       | 0.7100 | 0.8921 | 0.7826 | Guiyang/Longdongbao  | -0.5283 | 0.5392 | 1.1827 |
| Chengdu/Shuangliu    | 0.1174 | 1.5541 | -0.1327 | Fuzhou/Changle       | -0.2244 | -0.0167 | 0.2701 |
| Xi'an/Xiangan        | -0.2577 | 1.0586 | -0.1011 | Guilin/Liangjiang    | -0.6112 | -0.7005 | -0.3416 |
| Chongqing/Jiangebei  | 0.7999 | 0.2256 | 0.8366 | Nanchang/Changbei    | -0.6347 | 0.1416 | 0.2158 |
| Xiamen/Gaoqi         | -0.2438 | 0.2477 | -0.3166 | Hohhot/Baita         | -0.5923 | -0.2859 | -0.3650 |
| Wuhan/Tianhe         | -0.5808 | 1.9563 | 0.2395 | Yinchuan/Hedong      | -0.5132 | -0.8128 | -0.3426 |
| Changsha/Huanghua    | -0.4847 | 1.3542 | 0.6276 | Xishuangbanna/Gala   | -0.0703 | -1.9806 | 2.6726 |
| Nanjing/Lukou        | -0.2904 | 1.1776 | 0.0928 | Lhasa/Gonggar        | -0.4861 | -0.9351 | 0.8725 |
| Qingdao/Liuting      | -0.2766 | 0.7701 | -0.2933 | Kashgar              | -0.2963 | -1.6157 | 1.4715 |
| Sanya/Fenghuang      | -0.3530 | -0.9062 | -0.2218 | Xichang/Qingshan     | -0.8368 | -1.8011 | -2.1504 |
| Shenyang/Taoxian     | -0.7141 | 0.4940 | -2.2218 | Mudanjiang/Hailang   | -0.7464 | -1.2687 | -2.1385 |
| Zhengzhou/Xinzheng   | -0.1074 | 0.4336 | 0.4432 | -                    | -       | -       | -[

3.2. Hierarchical classification of sample nodes in system cluster analysis

The comprehensive factor score of sample nodes in Table 2 is used as a variable in cluster analysis. The name of each sample node (hub airport) is taken as an observation to conduct cluster analysis of Ward system. According to the analysis results of cluster analysis, the scale of airport operation and the economic development level of airport cities, then the twenty-nine airports are divided into four categories.

In Table 3, the first and second level hub airports are important aviation hub airports in China's aviation network planning and layout. It is located in a city with the most developed domestic economy, large population and large volume of passenger and freight traffic. It is the national large air transport hub. The third level hub airport is located in a densely populated large city. The level of economic development is inferior to that of the first or second level hub airport. It is a potential market for China's aviation network transportation development. The fourth level hub airport is mainly located in the regional economic centre. Its tourism resources are quite rich.

Table 3. Hierarchical classification of sample nodes in China's aviation network.

| Airport number | Airport name | Urban characteristics |
|----------------|--------------|-----------------------|
| 2              | Beijing / Capital, Shanghai / Pudong | Economic and cultural center, large population, well-developed economy, frequent |
Second level

| Routes                          | The provincial capital, the municipality directly under the central government, the deputy provincial city, large population density, large passenger flow momentum and high economic volume. |
|--------------------------------|----------------------------------------------------------------------------------------------------------------------------------|
| Guangzhou / Baiyun, Shenzhen / Baohan, Chengdu/Shuangliu, Xi'an / Xianyang, Chongqing / Jiangbei, Wuhan/Tianhe, Changsha/Huanghua, Nanjing/Lukou, Qingdao/Liuting |                                                                                        |

Third level

| Routes                          | The regional tourism population density is large, the economy is relatively developed, the tourism city is famous, and the volume of passenger and freight traffic is larger. |
|--------------------------------|----------------------------------------------------------------------------------------------------------------------------------|
| Xiamen/Gaoqi, Shenyang/Taoxian, Urumchi/Diwobao, Tianjin/Binhai, Fuzhou/Changle, Nanchang/Changbei, Yinchuan/Hedong, Mudanjiang/Hailang |                                                                                                                                  |
| Sanya/Fenghuang, Harbin/Taiping, Guiyang/Longdongbao, GuiLin/Liangjiang, Hohhot/Baita, Xichang/Qingshan, |                                                                                                                                  |

Fourth level

| Routes                          | Regional economic center and rich tourist resources. |
|--------------------------------|-----------------------------------------------------|
| Xishuangbanna/Gala, Lhasa/Gonggar, Kashgar |                                                                 |

4. Hierarchical analysis of airline network samples in China

Ward cluster analysis is carried out on sixty domestic routes. The airline network can be divided into three levels: the first level trunk line, the second level trunk line and the branch line.

Table 4. Hierarchical classification of China airline network sample routes.

| Classification | Routes |
|----------------|--------|
| First level trunk line | Beijing-Shanghai, Beijing-Shenzhen, Shanghai-Shenzhen, Guangzhou-Beijing, Guangzhou-Shanghai, Fuzhou-Shanghai, Chengdu-Xiamen, Fuzhou-Beijing, Nanjing-Beijing, Beijing-Xiamen, Guangzhou-Xiamen, Shanghai-Wuhan, Shenyang-Shenzhen, Nanjing-Shenyang, Nanchang-Shanghai, Beijing-Wuhan, Shanghai-Xian, Beijing-Qingdao, Beijing-Shenyang, Fuzhou-Changsha, Shanghai-Tianjin, Nanjing-Shenzhen, Beijing-Xian, Chongqing-Beijing, Beijing-Sanya, Guangzhou-Chengdu, Chengdu-Shanghai, Chongqing-Shenzhen, Chongqing-Shanghai, Guangzhou-Chongqing, Chengdu-Fuzhou Changsha-Sanya, Guangzhou-Tianjin, Hohhot-Shanghai, Yinchuan-Shanghai, Guangzhou-Zhengzhou, Guangzhou-Guiyang, Chengdu-Urumchi, Chongqing-Lhasa, Guilin-Shanghai, Xian-Xiamen, Hohhot-Beijing, Qingdao-Xian, Yinchuan-Beijing, Shenzhen-Tianjin, Guilin-Nanjing, Sanya-Shenzhen, Guilin-Xian, Chongqing-Urumchi, Guilin-Xiamen, Nanjing-Sanya, Shenyang-Quingdao, Yinchuan-Xian, Guilin-Yang-Sanya, Zhengzhou-Urumchi, Guangzhou-Nanchang, Nancheng-Shenzhen, Guiyang-Nanjing, Chengdu-Sanya, Zhengzhou-Chengdu |
| Second level trunk line |                                                                 |
| Branch line |                                                                 |

The first level trunk line connects Beijing, Shanghai, Guangzhou and Shenzhen. This is the four leading city in the national political and economic activities, which has the leading role and the ability of radiating and driving. It is the core starting and end point of China civil aviation route. Its route density and passenger and freight volume are huge, playing an important role in leading and radiating. The second level trunk line is mainly connected to the route between the first and second level aviation hub or the second level aviation hubs. The branch line has realized the route connection between the first and second level aviation hub and the third and fourth level aviation hub. China's aviation network has formed a development pattern with multiple hubs and axial radiation based on Beijing, Shanghai, Guangzhou and Shenzhen. It has a distinct hierarchy. The lines at all levels connect with each other and work together.

5. Comparison and analysis based on the development of hierarchical aviation network

Combined with the hierarchical classification results of the hub nodes of the aviation network, the take-off and landing sorties of the sample hub airports in 2005~2015 years were counted. Data show that in recent years, the weight curve of the first level aviation hub in China's aviation network is decreasing, and the weight curve of the third level aviation hub is increasing gradually. With the
The continuous development of China's aviation network, the second and third level hub airport has undertaken some of the freight volume of the previous first level hub airport in terms of air transport function in large and medium scale areas.

![Figure 1. 2005~2015 the weight change of aviation network hub at all levels.](image)

During the 2005-2015 year period, as the key index of the aviation network operation efficiency evaluation, the passenger transport volume, the cargo transport volume, the average passenger seat rate and the average flight carrier rate have increased. The annual growth rates of the four indicators were 12.42%, 8.03%, 1.54% and 1.13% respectively. By comparing and analyzing the operational indicators of China's aviation network in the past 2005-2015 years, the relationship between the aviation network hierarchy and network transport efficiency and scale is positively correlated. Through the interaction, interconnection and interweaving of different levels of aviation network, a stable hierarchical radial web structure is formed. Therefore, it can reduce the cost of distribution and transportation and improve the efficiency of aviation operation.

![Figure 2. 2005~2015 China civil aviation network flight passenger volume (cargo transport volume) trend.](image)

![Figure 3. 2005~2015 average passenger seat rate (carrier rate) trend of China civil aviation network.](image)
6. Conclusions
The research uses Ward cluster analysis method to study the hierarchy of China's aviation network. The twenty-nine airports in China are divided into four levels. And the sixty routes are divided into three levels: the first trunk line, the second trunk line and the branch line. This research comes to the following conclusions:
• China's aviation network has obvious hierarchy, that is, a structure with multiple hubs and axial radiation, which is based on the framework of four central cities in Beijing, Shanghai, Guangzhou and Shenzhen.
• With the intensification of the aviation network level, the four key indicators for evaluating the efficiency of the aviation network, such as the passenger transport volume, the cargo transport volume, the average passenger seat rate and the average flight carrier rate, have shown a growing trend. This growing trend shows that the hierarchy of aviation network is positively related to the operation efficiency and scale of network.

Acknowledgments
This research was supported by Project 51208308 supported by NSFC, Project 51208307 supported by NSFC, Science and Technology Projects of Transportation Department of Guangdong Province (No. 2012-02-055).

References
[1] Simon, H. (1962) The architecture of complexity. Proceedings of the American Philosophical Society, 1962:106-467.
[2] Eppell, V.A.T.A. (1997) Hierarchy the Next Step. In: 21st Australian Transport Research Forum. Australian. pp: 21-853.
[3] Wang, L. (2012) Research on Optimization Design of hub route network. Nanjing University of Aeronautics and Astronautics.
[4] Xu, S.Z. (2015) Structure evolving phase transition and optimization design of airline network. Nanjing University of Aeronautics and Astronautics.
[5] Sun, M.D. (2017) Research on Chinese aviation network based on complex network. Qingdao Technological University.
[6] Xie, L.H., Shang, T. (2012) SPSS System Analysis and Data Mining. Electronic Industry Press, Beijing.
[7] Jin, W., Chen H.P. (2007) K-means algorithm based on hierarchical clustering. Journal of Hohai University Changzhou, 01:7-10.
[8] Li, W.D. (2008) Application of multivariate statistical analysis. Peking University press, Beijing.