Advances in Computer Science Research, volume 71
4th International Conference on Machinery, Materials and Information Technology Applications (ICMMITA 2016)

Key Nodes Detection of Aviation Network Based on Complex Network Theory

Tu Congliang[1, a], Wu Minggong[2, b], Wen Xiangxi[3], Han Cheng[4]

[1]College of Air Traffic Control and Navigation, Air Force Engineering University,710051, China

[2]College of Air Traffic Control and Navigation, Air Force Engineering University,710051, China

[3]College of Air Traffic Control and Navigation, Air Force Engineering University,710051, China

[4]College of Air Traffic Control and Navigation, Air Force Engineering University,710051, China

[a]email: 534995138@qq.com, [b]email:597931972@qq.com

Keywords: Aviation Network; Complex Network; Closeness; Evaluation Matrix; Edge Weight

Abstract. In order to detect the key nodes in the aviation network, a node detection algorithm is proposed. In view of the fallibility of previous methods, this algorithm integrates the improved closeness sorting algorithm and the importance evaluation matrix. Firstly, the model of the aviation network is constructed. Based on the closeness sorting algorithm, a weight function is proposed, which is set to reflect the position information of the node. Then, considering the edge weights, the importance matrix is constructed, and the importance of each node is obtained. Experimental results show that the algorithm is capable of simulating the actual aviation network, using the advantages of the two methods to the most extent, at the same time, considering the route traffic.

Introduction

With the rapid development of air transport industry and the increase of air traffic, China's aviation network has gradually formed and continues to expand. As the lifeblood of the traffic network, the aviation network is closely linked with the development of politics, economy, science and technology. If the congestion and invalidation of the key nodes are caused, the whole network will be paralyzed seriously, which will heavily affect the progress of the society. In modern warfare, it is important to determine the key nodes in the air network accurately and quickly, and we can not only strike the enemy network system accurately, but also protect the key nodes of our aviation network if we do so. Therefore, it is of great theoretical and practical significance to study the key node identification of aviation network.

Aviation network has the characteristics and nature of complex network, so we reckon it as a typical complex network. At present, the theoretical research on the detection of critical nodes in complex network has become an important issue. In literature [1], a shortest path destroying network algorithm is proposed to detect the critical nodes. After removing a node from the starting point to the end point, the shortest path between two points is calculated, if the shortest path becomes longer, then the removed node is the key node. But this method has great flaws, when the network is no longer connected due to the missing node, it can’t distinguish between the intermediate nodes. The social network analysis methods [2] can be applied to complex network analysis in various fields. Such as literature [3], first, the node degree and the number of the median are compared, then the key nodes were identified based on game theory; In literature [4], the average path and the clustering coefficient weight are taken as the importance evaluation indexes; In literature [5], The shortest path increasing evaluation node algorithm was proposed. But these methods are often too rough and can’t evaluate the node comprehensively. Their application to key nodes detection of aviation network is still rare. The importance evaluation matrix reflects the role played by the nodes in the network transportation and the importance contribution of the neighboring nodes. The closeness sorting algorithm can reflect the difficulty of the nodes reaching to other nodes in the network, it considers the node degree, the position of the node, and overcomes the limitation and fallibility of the above method. The closeness sorting algorithm has high accuracy, but they only consider the case without edge, which is route traffic data in reality. Airline traffic is an important attribute index of air traffic, which reflects the role of airports and routes. However, the traditional methods of detecting complex network nodes are generally focused on the interrelationship between nodes and the topological structure of network, no consideration is given to the influence of edge weights on node importance in real air networks, and thus lack of accuracy.

In this paper, a new key node detection algorithm for aviation network is proposed by using the improved weighted closeness sorting algorithm and the importance degree evaluation matrix. This algorithm makes up for the shortcomings of the traditional method which neglects the edge weight. It can not only consider the position information of the node, but also consider the impact of the node on the whole network and the adjacent node, as well as route traffic. This method has important significance to the accurate detection of key nodes.

Basic theory

Complex network has not been precisely defined, it can be seen as a large number of real complex system topology abstraction. At present, there is a relatively strict definition of it, that complex network [6] is composed of a huge number of complex nodes and the intricate relationship between nodes. Here are some basic concepts.

The shortest path is from the node v to v' if there exists a path (when $v_1 = v, v_n = v'$) such that $sum\sum_{i=1}^{n-1} f(e_i, i+1)$ is minimized. The least number of routes needed to transfer from one airport to the designated airport is known as the shortest path of the aviation network, the number of edges e_i on the path is the shortest distance.

Node closeness refers to how close a node is to a local community. The airport node closeness reflects the relative location of the airport in a regional air network and the impact of the airport on the entire aviation network.

Key Node Detection Model

Air network model assumptions

In order to simplify the calculation process, the following assumptions are made to adjust to the actual aviation system in the process of modeling the aviation network. We take into account of the characteristics of the aviation data, as well as the definition of complex network (Unicom diagram, simple, undirected) :

(1) The node v_i in the topological model represent all the general-purpose airports with transport capacity in the whole country, the nodes set is $V = \{v_1, v_2, ...v_n\}$, and the number is $N = |V|$.

(2) Edge e_j represents the transport relationship between one airport and another. If there are routes between the two airports, then we consider it is connected by edge, or edgeless. There is at most one edge connected between two airports. The set $E = \{e_1, e_2, ...e_M\}$ is the set of edges whose number is $M = |E|$.

(3) Assuming that the topological characteristics of the route are the same, that is, the traffic flow is regarded as the edge weight in the network, with no consideration of the transmission efficiency and transmission line characteristics.

Weighted Closeness Calculation

The position of an air network node is closely related to its importance. The more the node is toward the center of the network, the more likely it is to become a key node, the more the node is biased toward the edge of the network, the less important it is. Closeness can measure the difficulty of arriving at any node from the specified node in the air network, that is, reflecting the position of the node and its distance from the rest nodes, so as to do the critical evaluation on the air network node. The traditional closeness can be calculated by (1), (2):

$$CL_i = \frac{1}{\sum_{j=1}^{N} d_{ij}} \tag{1}$$

N indicates the number of airports in the network, d_{ij} indicates the minimum number of airports i to the airport j. Then CL_i is normalized to:

$$P_i^c = \frac{N-1}{\sum_{j=1}^{N} d_{ij}} \tag{2}$$

It can be seen from the formula that the higher the value is, the higher the node's importance is, and the more critical the node is. However, the traditional closeness sorting algorithm can’t be adjusted according to the characteristics of different networks, but simply add the shortest distance between a node and the remaining nodes indiscriminately. The shortest distance between the nodes in our country is very close, which is no more than 3, so the closeness sorting algorithm is not very appropriate.

Considering the characteristics of China's aviation network, a weighted closeness sorting algorithm is proposed to highlight the influence of path distance between nodes on the detection results. Thus, here set a weight function for the shortest path, improved closeness calculation formula is as follows:

$$P_i^c = \frac{N-1}{\sum_{j=1}^{N} f\left(d_{ij}\right)} \tag{3}$$

In this paper, $f\left(d_{ij}\right)$ is a logarithmic function, when the change of d_{ij} is in a small range, the effect on the result is great; when d_{ij} increased to a certain extent, the effect on the result will become smaller and smaller. This is done in order to highlight the influence of adjacent air network nodes on their closeness and reduce the impact of distant nodes.

Critical Evaluation Matrix

In the process of detecting the critical nodes of the air network, besides considering the importance of the location of the airport nodes, the route flow of the airport nodes is also taken into account. However, the key evaluation matrix can combine both the position of the airport node and the flight flow, therefore reflects the importance of the airport in the actual aviation network relative to other airports, and the influence of the airport on the whole traffic flow. In this paper, the key evaluation matrix is proposed based on the importance evaluation matrix. The node degree is used to compose the importance degree relations among the nodes, and edge weights and closeness are used as the evaluation parameters. The establishment of the critical evaluation matrix is described below.

In the last section of the air network model, $N = |V|$ is the number of nodes, D_i is the degree of airport node v_i. In the model, a is the adjacency matrix for the aviation network. v_i gives the ratio D_i / k^2 of its own importance degree to every adjacent nodes. Thus form a matrix after calculating the importance degree output of all the nodes $v_1, v_2, ...v_n$ to their neighbors, and this matrix is denoted as H_{IC}:

$$H_{IC} = \begin{bmatrix} 1 & \delta_{12} D_2 / k^2 & \cdots & \delta_{1n} D_n / k^2 \\ \delta_{21} D_1 / k^2 & 1 & \cdots & \delta_{2n} D_n / k^2 \\ \vdots & \vdots & \cdots & \vdots \\ \delta_{n1} D_1 / k^2 & \delta_{n2} D_2 / k^2 & \cdots & 1 \end{bmatrix} \tag{4}$$

In the matrix, δ_{ij} is the contribution allocating parameter, if the two nodes are connected, the value of δ_{ij} is 1. Otherwise, the value is 0. The element on the diagonal is 1, which means the contribution of the node to the node itself is 1. In addition, k is the average degree of all nodes,

$$k = \frac{\sum_{i=1}^{|V|} D_i}{|V|} \quad (5)$$

In order to be able to reflect the role of the airport node in the course of the network transportation, combined with the flight flow, calculate the weight of each airport node, denoted as:

$$S_i = \sum_{j \in N_i} w_{ij} \quad (6)$$

N_i is the neighbor nodes set of node i, w_{ij} is the weight of the edge directly connected with the node i. The larger the weight, the closer the airport node is to the surrounding airport.

Considering the node's closeness value, we replace the node's importance degree contribution ratio of H_{IC} with importance degree contribution value, and the node importance degree evaluation matrix can be obtained:

$$H_E = \begin{bmatrix} P_1^c S_1 & P_2^c \delta_{12} S_2 D_2 / k^2 & \cdots & P_n^c \delta_{1n} S_n D_n / k^2 \\ P_1^c \delta_{21} S_1 D_1 / k^2 & P_2^c S_2 & \cdots & P_n^c \delta_{2n} S_n D_n / k^2 \\ \vdots & \vdots & \cdots & \vdots \\ P_1^c \delta_{n1} S_1 D_1 / k^2 & P_2^c \delta_{n2} S_2 D_2 / k^2 & \cdots & P_n^c S_n \end{bmatrix} \quad (7)$$

In the formula, H_{Eij} is the node j's importance degree contribution to the node i. It can be seen that the contribution value of an airport node to its adjacent airport is related to its closeness, node degree, and traffic passing through the route. The higher the closeness value, the node degree, the route traffic are, the greater the importance contribution the node to its adjacent nodes is.

Importance Calculation

The calculation of airport node importance is carried out from two aspects, one is the closeness, the other is the right side. As a method of probing the nodes of the aviation network, the closeness can determine the influence of the position of the airport on the degree of its importance. The logarithm function is used to weight the influence of the neighboring nodes, which is more pertinent to the actual aviation network. The setting of edge weights takes into account the traffic flow, which is closer to the actual reflection of the role of nodes than the previous methods. In order to combine the advantages of these two aspects, so that node detection algorithm is accurate and efficient, we define the importance of node i as C_i:

$$C_i = \sum_{j=1, j \neq i}^{n} P_j^c \delta_{ij} S_j D_j / k^2 \quad (8)$$

In the formula, C_i means the summation of all the airports' importance contribution to their adjacent airports, which is base of the sorting standard.

Detection Algorithm Steps

Here we give the flow chart of the algorithm to evaluate the node of its importance:

Input: Adjacency Matrix $A = (a_{ij})_{n \times n}$, Edge Weight w_i

Output: importance C_i of node i and sort.

As shown in the flow chart in Figure1, the algorithm combines the node's closeness, node degree and edge weight, which is equivalent to using the node and edge location information and network information at the same time. This method can evaluate the importance of the nodes in the air network and obtain more accurate evaluation results.

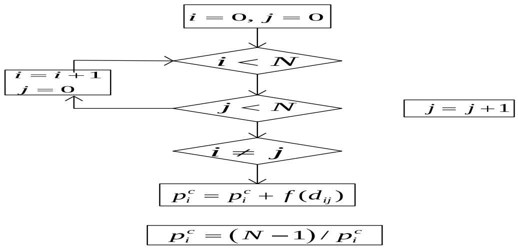

Figure 1. Key Node Detection Algorithm Flow Chart

Simulation

Algorithm Application

In the experiment, we take the weekly flight frequency of the 199 airports in china in May 2016 as statistics. $M=\left(a_{ij}\right)_{199\times 199}$ is the matrix the number of flights between airports, and we reckon the number of flights as the route traffic, that is, the edge weight. When the number of flights between airports is not equal to 0, the value is 1, otherwise 0, so we get an adjacency matrix $A=\left(a_{ij}\right)_{n\times n}$ of the Chinese aviation network.

Based on the adjacency matrix, pairs of airports with scheduled flights are connected to each other by a straight line, as shown in figure 2.

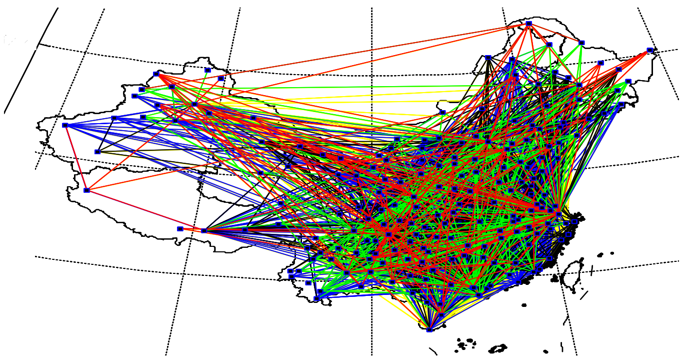

Figure 2. Aviation Network Of China(except for Hong Kong, Macao and Taiwan)

According to the algorithm model, input adjacency matrix $A=\left(a_{ij}\right)_{n\times n}$, and we get the shortest distance matrix $Dis=[d_{ij}]$ of the airport nodes using Freud's algorithm, obtained the importance C_i of each node and sorting in Table 1 according to formula (8).

According to the importance degree of each airport node obtained by the key node detection algorithm, the following conclusions are obtained:

(1) The three cities of Beijing, Shanghai and Guangzhou are expected to be the top three, which is in line with the reality, verifies the accuracy of the algorithm.

(2) Xi'an is the geographical center of China, followed by the fourth place. Although Xi'an is not the economic center as Beijing, Shanghai and Guangzhou, Xi'an is the hub of transportation between the east and the west, and plays an important role in the entire aviation network. This shows that the detection algorithm fully considers the global importance of nodes.

(3) Chengdu, Kunming, Chongqing, Shenzhen and other cities are in the forefront, these cities

are the center of the some regions, indicating that these nodes have a great impact on the local network.

(4) Although cities such as Haikou and Urumqi are not located in the geographical center of China, the results of detection algorithms show that they are still important nodes.

(5) Most of the nodes in the last few cities are located in the edge of the geometric margin, and the air traffic is not much, indicating that the algorithm can perceive their location information and edge information.

Table 1. Node Importance and Ranking

City name	Importance	City name	Importance	City name	Importance	City name	Importance	City name	Importance
Beijing	2.07556	quanzhou	0.423394	Weihai	0.224981	Wuhai	0.149873	Hechi	0.086889
Shanghai	2.025213	Lijiang	0.399548	Huaian	0.219972	Weifang	0.149084	Tengchong	0.085311
Guangzhou	1.340661	Jieyang	0.396312	Xilinhot	0.217598	Manchuria	0.148881	Chaoyang	0.084573
Xi’an	1.290995	Lhasa	0.355548	nyingchi	0.217054	Bayannur	0.148194	Panzhihua	0.083189
Chengdu	1.237558	Mianyang	0.351598	Wanzhou	0.210932	Rizhao	0.147295	Daocheng	0.078024
Kunming	1.119479	Baotou	0.346264	Dazhou	0.208261	Zhaotong	0.146982	Karamay	0.07768
Chongqing	1.101845	Hailer	0.343683	Liupanshui	0.205938	Taizhou	0.146952	Haixi	0.076649
Shenzhen	1.049477	Zunyi	0.336444	Jining	0.202741	Jiuzhaigou	0.1424	Yizhou	0.076215
Hangzhou	0.843087	Wuxi	0.335045	Fuyang	0.200536	Tongliao	0.140012	Yichun	0.075865
Xiamen	0.807869	Nantong	0.331213	Yanji	0.200297	Baise	0.139026	Jinchang	0.06957
Tianjin	0.784716	Yancheng	0.322171	Xingyi	0.198501	Chizhou	0.137698	Hetian	0.068331
Harbin	0.760046	Ordos	0.321483	Huizhou	0.196473	Jiamusi	0.136187	Turpan	0.068299
Dalian	0.743025	Zhanjiang	0.315984	Yichun	0.19566	Jinzhou	0.134429	Zhangye	0.066463
Shenyang	0.738116	Zhangjiajie	0.312876	Changed	0.190567	Hengyang	0.130951	Yongzhou	0.065978
Changsha	0.722533	Linyi	0.309181	Nanyang	0.188692	Changbaishan	0.128914	Changhai	0.060788
Qingdao	0.718613	Beihai	0.299021	Pu’er	0.187635	Anshun	0.127223	Lincang	0.060434
Haikou	0.70434	Yichang	0.295044	Zhangjiakou	0.186985	Qiqihar	0.125744	Aershan	0.059265
Guiyang	0.694962	Yangzhou	0.292802	Enshi	0.185844	Qingyang	0.12538	Changdu	0.057164
Urumqi	0.69285	Bijie	0.291269	Jingangshan	0.184959	Hami	0.123861	Tonghua	0.056581
Nanning	0.686065	Yulin	0.289168	Daqing	0.182365	Tangshan	0.123598	Anshan	0.053624
Lanzhou	0.662109	Tongren	0.279551	Mangshi	0.180641	Tianshui	0.123022	Bole	0.052996
Hohhot	0.645848	Changzhou	0.276704	Kasha	0.177444	Tacheng	0.120608	Shennongjia	0.052086
Fuzhou	0.64229	Lianyungang	0.276693	Dandong	0.175612	Zhijiang	0.120463	Yushu	0.051209
Sanya	0.639879	Yiwu	0.269481	Korla	0.172575	Qianjiang	0.117687	Liping	0.04375
Wuhan	0.639012	Xiangyang	0.267422	nanchong	0.1703	Mudanjiang	0.116314	Wuzhou	0.043703
Nanjing	0.632363	xishuangbannan	0.26556	Lvliang	0.169837	Qinhuangdao	0.114796	Aba	0.03644
Zhengzhou	0.624564	Yuncheng	0.259494	Aksu	0.167221	Handan	0.112005	Erenhot	0.035802
Yinchuan	0.570937	Dunhuang	0.251979	Wuyishan	0.164378	Yining	0.111061	Wenshan	0.033863
Xining	0.562041	Dali	0.251331	Jingdezhen	0.163561	Kuche	0.10935	Shigatse	0.030707
Guilin	0.535874	Luzhou	0.246895	Baoshan	0.161741	Mohe	0.106593	Kaili	0.030616
Shijiazhuang	0.530163	Xuzhou	0.246758	Jixi	0.161295	Fuyuan	0.105169	Golmud	0.024755
Ningbo	0.527426	Huangshan	0.240725	Wulanhaote	0.160609	Hanzhong	0.104559	Jiujiang	0.017298
Jinan	0.524535	Ganzhou	0.234739	meixian	0.160079	Guangyuan	0.103882	ali	0.014823
Changchun	0.511933	Jiayuguan	0.233945	Chifeng	0.158207	Liancheng	0.102371	Alxa	0.012783
Taiyuan	0.49395	Liuzhou	0.231604	Anqing	0.158195	Quzhou	0.099959	Ninlang	0.009487
Zhuhai	0.491233	Yibin	0.230322	Foshan	0.157107	Nalati	0.093691	Fuyun	0.004016
Wenzhou	0.462114	Datong	0.229129	Yan’an	0.156587	Guyuan	0.093239	Altay	0.004016
Hefei	0.459002	Changzhi	0.227682	Zhoushan	0.153112	Zhongwei	0.089875	Ankang	0.002481
Nanchang	0.456762	Xichang	0.226303	Shangri-la	0.151385	Heihe	0.089632	Alxa right	0.002481
Yantai	0.438553	Luoyang	0.226124	Dongying	0.150056	jiagedaqi	0.088213		

Conclusion

Key nodes of the airports in China's aviation network have been detected based on the improved closeness sorting algorithm and the importance degree evaluation matrix. Simulation results show that the algorithm proposed in this paper is well-targeted for the detection of aviation network nodes compared with the previous method. This method can reflect the importance of each airport node comprehensively, which is very important for the defense of key airport nodes. The next step of my study will be focused on key nodes detection of the civil and military airports network at home and abroad.

References

[1] H.W.Corley, H.Chang, Finding the most vital nodes into a flow network. Management Sei. 1974,(21):362-364.

[2] Pei Lei, Ma Fei-Cheng. The application of social network analysis(SNA) in information science. Library Tribune,2006,(06):40-45.

[3] Daniel Gomezb, Enrique Gonzalez Aiangtiena, etal. Centrality and Power in social networks:a game theoretic approach[J].MathematiealSoeialSciencesVolume4,ISSuel, 2003.

[4] Zhou Xuan, Zhang Feng-Ming, Li Ke-Wu. Finding Vital Node By Importance Evaluation Matrix In Complex Networks, Acta Phys. Sin.2012, (61):1-6

[5] He Nan,Li De-Yi,Gan Wen-Yan, etal. Mining vital nodes in complex networks. Computer Science,2007,34(12):1-5.

[6] Albert R, Barabasi A-L. Statistical mechanics of complex networks. Reviews of Modem Physics[J],2002,74(1):47-97.