Study on Tourist Flow Network Structure in Jiangsu Province

Study on Tourist Flow Network Structure in Jiangsu Province

Rong Pu[1,*], Shuying Lu[1], Yuansheng Pu[2], Zhuomin Tao[1], and Peixue Liu[3]

[1]School of Geographical Science, Nanjing Normal University, 210023 Nanjing, China

[2]Department of Business Administration, Jiangsu Province Nanjing Engineering Vocational College, 211135 Nanjing, China

[3]School of Geographic and Oceanographic Science, Nanjing University, 210093 Nanjing, China

[*] Corresponding author: pu_rong123@163.com

Abstract. In this paper, we used social network analysis (SNA) method to establish and analyse the network structure of the tourist flow of Jiangsu Province in China based on big data, and built a variable model to analyse its influencing factors. The results show that there is an obvious core-edge structure in the network of tourist flow. The centrality index in the southern region of Jiangsu Province is better than that in the northern region. The road distance is the main factor restricting the flow between regions, and the quantity of trains has a significant impact on the tourist flow. That is to say, a place with convenient transportation is the primary choice for people to travel. High quality of tourist facilities and online tourist products can attract tourists. They are also the driving factors for southern Jiangsu Province to play a great role in the tourist flow network.

1. Introduction

With the rapid development of tourism, inter-regional tourist flow are increasingly frequent, forming a complex network structure. Tourist flow and its network structure have become hotspots in domestic and international tourism research. Foreign scholars have studied the spatial structure of tourist flow for a long time. Representative theories are core-edge theory and tourist flow circle structure theory. Based on these theories, a large number of spatial structure models of tourist flow have been proposed, such as tourist flow model proposed by Campbell [1], tourist flow spatial stratification model proposed by Lundgren [2] and Dianne's multi-destination tourist flow spatial structure model, which is the preliminary ideal of tourist flow [3]. Scott found that the higher the level of economic development, the agglomeration effect could be more stronger by studying the network structure characteristics among the four tourism destination organizations in Australia [4]. Chua constructed the tourist flow network of Cilento in Italy based on Twitter geo-tagged photos [5]. In recent years, Chinese scholars have done a lot of research on the spatial structure and evolution mechanism of tourist flow. Zhang Jinhe first proposed the theory of spatial field effect of tourist flow and used factor analysis to study the spatial field effect of tourist flow [6]. Liu Fajian analyzed the characteristics of China's inbound tourist flow from a macroscopic perspective for the first time, providing a new method for exploring the spatial characteristics and mechanism of inbound tourist flow [7]. Feng Xuegang, based on the push-pull theory, studied the factors that affect the direction of the tourist flow [8]. Liu Hongying analysed the Network Structure Characteristics of Pan-Beibu Gulf of China from the perspective of tourism routes [9]; Yang Xingzhu studied the structural characteristics of urban tourist flow network

by using SNA method [10]; Ma Lijun discovered the structural characteristics of the domestic tourist flow network of urban residents [11]. Xu Min analysed the influencing factors of tourist flow in Yangtze River of China based on online reservation data [12].

It can be seen from the above that scholars at home and abroad have done a lot of work on the structural characteristics and influencing factors of tourist flow network, which has laid a good foundation for the development of tourism geography. However, most research is based on traditional statistics, and the relational data are rare. Moreover, with the rapid development of Internet technology, the combination of big data and tourism have become popular. The new method and data can supplement and enrich the study of traditional tourist flow network and provide a reference for the regional development and market positioning.

2. Data sources

The data in this study were obtained from Telecom data center. The sample includes the quantity of tourists from Jiangsu Province visitied 13 cities in the province in 2016. The road distance, the number of scenic spots and star hotels were come from *Jiangsu Statistical Yearbook*. The quantity of trains were obtained from www.12306.cn.

3. Methodology

3.1 Social network anlysis

Firstly, we used SNA method to build a tourist flow network. We took 13 cities in Jiangsu Province as the node, the city where the tourists were located as the starting points and the city where they traveled were located as the ending points. Establishing a relationship between them when the flow reached a certain scale [13]. The 13 cities are divided into a thirteenth order matrix M. The element m_{ij} in the matrix represents the tourist from the city i went to the city j during this period. In order to better reveal the evolution of network characteristics, 520 was selected as the cutoff value (520 is the median of matrix elements). Then the matrix was converted into a binary matrix M_b, when m_{ij}> 520, it was changed to 1, otherwise it was 0, this matrix is a 1-mode network that determines whether there is a tourist flow relationship between cities.

Secondly, we used NetDraw to draw the tourist flow relationship diagram (Fig.1). The diagram directly reflects the correlation between tourist flows in Jiangsu Province. Nanjing, Suzhou, Wuxi, Changzhou and other southern parts of Jiangsu Province have formed a closed pattern with tight network connection. There are multiple connections between Nanjing, Wuxi and Suzhou with frequent flow, while the northern part of Jiangsu Province has lots of isolated nodes and the network connection is weak.

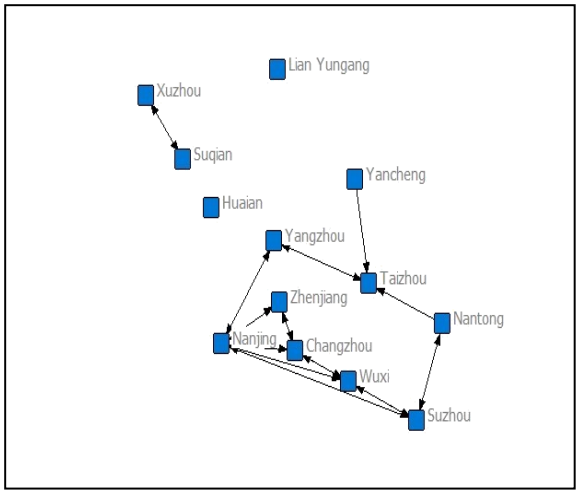

Fig. 1. The tourist flow diagram of Jiangsu Province.

Finally, we used Ucinet6.0 to calculate the structural indexes of tourism nodes. The structural characteristics of nodes are mainly analysed by degree centrality, betweenness centrality and core-edge analysis [14].

The degree of centrality includes the outdegree and the indegree, which respectively represent the outward and intward linkages between nodes. In the formal: C_{ADi} is the degree centrality of the point i; r_{ij} is the relationship between i and j; n is the number of tourism nodes.

$$C_{ADi} = \sum_{j=1}^{n} r_{ij} \qquad (1)$$

The betweenness centrality of the nodes measure the ability of the node to act as a transit stream intermediary. In the formal: C_{ABi} is the betweenness centrality of the point i; b_{jk} (i) is the probability that the point i on the shortcut between j and k; n is the number of tourism nodes.

$$C_{ABi} = \sum_{j}^{n} \sum_{k}^{n} b_{jk}(i) \qquad (2)$$

3.2 Influencing factors

In this paper, Quadratic Assignment Procedure (QAP) method was used to study the influencing factors of tourist flow in Jiangsu Province. The influence variables are summarized as the level of network marketing, road distance, the quantity of trains, the difference of tourism resources and the differences in tourist reception facilities between tourist sources and destinations. Based on this, the model was constructed as follows:

$$M=f\ (QFD, TRD, TFD, NML, RD) \qquad (3)$$

In the formula (3), M represents the tourist flow matrix; QFD is the quantity of trains, TRD is the difference of tourism resources; TFD is the difference of tourist reception facilities; NML is the level of network marketing; RD is the road distance. Among them, TRD was replaced by the number of scenic spots, TFD was replaced by the number of star hotels, and NML was replaced by the number of products on websites. In this formula, TRD and TFD were directly formed the difference matrix, the other variables are used to establish the relation matrix. Considering the different matrix quantization standards, the min-max standardization function was used to standardize the above matrix [15]. The formula is:

$$X_{std} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \qquad (4)$$

4. Results

4.1 Node structure index analysis

The table shows that Nanjing has the highest outdegree centrality, indicating that the city serves as the diffusion centers of the network. The centrality of outdegree is greater than that of indegree, which indicates the attribute of tourist source is stronger than that of tourist destination. Yangzhou, Zhenjiang, and Taizhou have attracted a large number of tourists due to the improvement of tourism infrastructure in recent years.

Table 1. The centrality of the tourist flow network.

Area	Degree centrality		Betweenness centrality	Role
	Outdegree	Indegree		
Nanjing	6.000	5.000	25.500	Core
Suzhou	5.000	4.000	13.000	Core
Wuxi	3.000	4.000	1.000	Core
Changzhou	4.000	5.000	3.500	Core

Nantong	3.000	2.000	2.000	Edge
Yangzhou	3.000	3.000	15.000	Edge
Zhenjiang	3.000	3.000	0.000	Edge
Taizhou	2.000	4.000	8.000	Edge
Yancheng	2.000	1.000	0.000	Edge
Huaian	1.000	1.000	0.000	Edge
Suqian	2.000	2.000	0.000	Edge
Lian Yungang	1.000	1.000	0.000	Edge
Xuzhou	2.000	1.000	0.000	Edge

Nanjing, Suzhou, and Yangzhou have much higher betweenness centrality values than other regions so they have much significant control over other regions in the whole network. In general, the centrality values of southern Jiangsu Province are better than the north, which have great attraction to most tourists in the province, and tourists in this region have stronger travel willingness.

In this paper, the core-edge analysis is implemented with the help of Network/ Core-Periphery in Ucinet. Core members of the network are Nanjing, Suzhou, Changzhou and Wuxi, while the other nine cities are all in edge. It shows that there is an obvious core-edge structure in the tourist flow network (Fig.2). The tourism development in Jiangsu Province is imbalanced, and the core area is distributed in the southern part of Jiangsu. It can be seen that the core-edge analysis is basically consistent with the results of the central analysis, so that the conclusion can be drawn more scientifically.

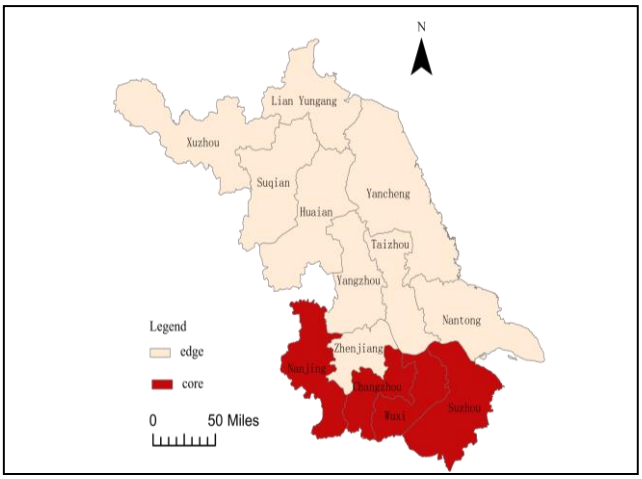

Fig 2. The Core-edge structure of tourist flow.

4.2 Results of influencing factors

Import the above data into Ucinet for calculation, and the correlation analysis results are as follows (Table 2). The significance level of QFT and NML are both under 1%. RD is significant at the level of 5%. These variables have a significant impact on the flow of tourism.

Among them, the coefficient of RD is negative, indicating that the farther the road distance is, the less the tourists are willing to drive to that place. Judging from the correlation coefficient, the coefficient of QFT is the largest, indicating that tourists have big demand on convenient transportations. Also, they are more concerned about the tourism area with better online marketing and abundant products when choosing destinations.

The regression analysis showed that three of the five variables reached a significance level of interpretation of the dependent variable. The coefficient of determination R^2 is 0.397, which is within the acceptable range. The explanatory factors of the three factors QFT, NML and RD for the independent variable are 38.1%. It indicates that the road distance is the main factor restricting the

flow between regions, and the quantity of trains has a significant impact on the tourist flow. That is to say, a place with convenient transportation is the primary choice for people to travel. Also, high quality of tourist service facility and items activities with local features and online promotion of travel products can attract tourists. They are also the driving factors for southern Jiangsu to play a great role in the tourist flow network.

5. Conclusion

In this paper, we analysed the characteristics of the network structure and tried to explore the factors affecting the network structure of tourist flows in Jiangsu Province. The main conclusions are as follows.

Table 2. Correlation and regression analysis of tourist flows.

Independent variable	Dependent variable			
	QAP correlation	Significance	QAP regression	Significance
QFT	0.552***	0.001	0.449***	0.000
TFD	0.201*	0.076	0.069	0.305
TRD	0.148	0.165	-0.031	0.410
RD	-0.301**	0.003	-0.303***	0.000
NML	0.384***	0.001	0.088*	0.060
R-square			0.397	
Adj s-sqr			0.381	

Note. ***p < 0.001; **p < 0.05; *p<0.1

The network in southern Jiangsu Province is connected closely, while the links between nodes in northern Jiangsu Province are relatively few. There is an obvious core-edge structure in the tourism network, and there is a big difference in tourism development between the north and the south. Considering the particularity of relational data, QAP analysis method was used to study the correlation and regression analysis between tourist flow matrix and factor matrix. The results show that tourists care much about the convenience of getting to the destination. High quality of tourist service facility and online promotion of travel products are also the driving factors in the southern part of Jiangsu Province to play a great role in the tourism flow network.

The northern part of Jiangsu Province has a great potential for development. It is more important to make reasonable plans to develop tourism in these areas. With the improvement of transportation in the province and the radiation of core regions, the tourist flow in northern Jiangsu Province will increase. They should learn from the development experience of southern areas, focus on tourism planning, and build a group of influential resort areas to attract tourists.

In this paper, there are some new ideas in the conclusion, but the related research still has limitations. The influence of tourists' behavior of different occupations, genders and family environments on tourist flow remains to be studied.

We are grateful to the National Natural Science Foundation of China for supporting our research (No. 41571139).

References

[1] C.K. Campbell, *An approach to Research in Recreational Geography* (British Columbia: Department of Geography, University of British Columbia, 1967)
[2] J.O.J. Lundgren, *The Development of Tourist Travel System: A Metro Politian Economic Hegemony Par Excellence (*Jahrbueh fur Fremdenverkegr, 1972)
[3] D. Dredge, Destination place planning and design. Annals of Tourist Research, **26**, 772-791 (1911)
[4] N. Scott, C. Cooper, R. Baggio, Destination Networks: Four Australian cases. Annals of Tourism Research, **35**, 169-188(2008)
[5] A. Chua, L. Servillo, E. Marcheggiani, et al,Mapping Cilento: Using geotagged social media data to characterize tourist flows in southern Italy. Tourism Management, **57**, 295-310(2006)

[6] J.H. Zhang, J. Zhang, N. Li, et al, Spatial field effect analysis of domestic tourist flow in China. Geographical Research, **24**, 293-303(2005)
[7] F.J. Liu, J. Zhang, D.D. Chen, Study on the Structural Characteristics and Motivation of China's Inbound Tourist flow Network. Acta Geographica Sinica, **65**, 1013-1024(2010)
[8] X.G. Feng, H.P. Huang, T.H. Wan, Study on the Temporal and Spatial Difference of Potential of Provincial Travel Market and Counter-seasonal Development Strategy in China. Tourism Science, **29**, 1-14(2015)
[9] X.Z. Yang, C.L. Gu, Q. Wang, Systematic Analysis of Driving Forces of Tourist flow. Geographical Research, **30**, 23-36(2011)
[10] H.Y. Liu, L.L. Wei, J. Zhang, Study on the Structural Characteristics of Regional Tourist flow Network Based on Tourism Routes. Human Geography, **27**, 131-136(2012)
[11] L.J. Ma, Y. Xiao, Structural Characteristics of Domestic Tourist flow Network of Typical Urban Residents. Economic Geography, *38*, 197-205+219(2018)
[12] M. Xu, Z.F. Huang, F.D. Cao, et al, Structural characteristics and influencing factors of tourist flow network based on online reservation data analysis-taking Yangtze River delta region as an example. Economic Geography, **38**, 193-202(2018)
[13] S. Hsinyu, Network characteristics of drive tourism destinations: An application of network analysis in tourism. Tourism Management, **27**, 1029-1039(2006)
[14] J. Liu, *Lecture Notes on Overall Network Analysis: Guidelines for Use of Ucinet Software* (Gezhi Press, 2009)
[15] Q.Y. Xiao, H.J. Liu, Re-examination of Provincial Labor Migration Motivation Theory Based on QAP Algorithm. China Population Science, **4**, 26-33+95(2007)