Combining Big Data Analysis to Study on the Spatial Structure of the Tourism Economy in the Three Northeast Provinces of China

Fuxiang Liu and Liqin Jiang(✉)

Harbin University of Commerce, Harbin, Heilongjiang, China
2924784347@qq.com

Abstract. Taking the digital economy under big data as the starting point, the article studies the network structure of tourism economy in the three eastern provinces, combining the gravitational model, social network analysis method and big data analysis method, by constructing the gravitational model to analyse and predict the spatial interaction ability between each city, after that, on the basis of obtaining the big data of 25 cities in the three northeastern provinces of China, the tourism economic linkage degree between each city in the three northeastern provinces is calculated, and the tourism economic linkage matrix, and then introduced the tourism economic linkage matrix into Ucinet software and used the method of social network analysis to analyse the network structure of the tourism economy in the three northeastern provinces of China. The results of the big data analysis found that the tourism economy of the three northeastern provinces of China is generally loose, and the top five cities with tourism economic linkages in the three northeastern provinces are Shenyang, Changchun, Jilin, Harbin and Dalian, and most of these cities are the capital cities of the three eastern provinces, according to which, it is recommended that the overall tourism economic development adopt a "core-led edge" development model.

Keywords: Big Data · The Modified Gravity Model · Social Network Analysis Method · Tourism Economy

1 Introduction

After decades of reform and opening up, China has accelerated the tertiary industry, especially tourism, and its status in the development of national economy is constantly improving. However, there are obvious differences in transportation, resources and tourism facilities in many regions, and the unbalanced development of tourism economy is more prominent. Tourism economic structure has been the hot issue of the current academic concern, as early as the 1960s, social network analysis was applied to tourism research abroad, while in China, it was not until the 1990s that social networks were used to analyse the spatial structure of tourism. The early research abroad provided a solid theoretical and practical basis for domestic scholars to conduct research in this area.

N. Radojević et al. (Eds.): ICAID 2022, AHIS 7, pp. 320–329, 2023.
https://doi.org/10.2991/978-94-6463-010-7_33

At present, academic research in this field in China mainly uses Ucinet software combined with social network analysis methods to analyse the structural characteristics of tourism economic networks between different cities, while there are also scholars who combine big data analysis and social network analysis. For example, Chen Chong combines social network analysis with current big data analytics and applies it to global networks [1]; Fang Yelin, Huang Zhenfang and Tu Wei took 16 cities in the Yangtze River Delta as the research object, and analysed the position of each node city in the whole by calculating the centrality, and structural holes [2]; Huang Changzhou, Li Ruyou made a relevant analysis of a whole tourism economic development in Jiangsu province based on social network analysis, and concluded that the tourism economic development varied greatly among cities in this province [3]; Wang Chunjuan and other scholars measured the linkages of the tourism economy in Zhejiang province, aiming to identify the factors that affect the development of the tourism economy and give suggestions [4]; Wang Jun, Xu Jinhai analysed the network structure of regional tourism in China from both overall and individual perspectives, and made relevant suggestions for our future tourism industry development [5]; Zhang Hong and Xia Ming measured the centrality and structural holes of the tourism economic structure network in Anhui Province from the individual network structure characteristics, and derived the spatial linkages of the tourism economic structure in Anhui Province [6]; In summary, the social network analysis method is widely used in the study of the spatial structure of the tourism economy, and there are fewer studies on the tourism economic linkages between individual cities in the three northeast provinces of China, therefore, this study is based on this.

2 Research Methods and Data Sources

2.1 Research Methodology

Firstly, based on the modified gravity model, the mutual gravitational force between two cities in the eastern provinces can be calculated, and every two cities constitute a network, so that the network structure of tourism space in the eastern provinces is formed, and the degree of tourism economic connection between two cities can be obtained. The following is the formula of the modified gravity model.

$$R_{ij} = \frac{\sqrt{P_i V_i} * \sqrt{P_j V_j}}{D_{ij}^2} \tag{1}$$

In which, R_{ij} represents the tourism economic linkage degree between cities i and j; P_i, P_j and V_i, V_j represent the total number of t tourist arrivals and total tourism economic income of cities i and j respectively, D_{ij} represents the shortest distance from city i to city j. The tourism economic linkage between each city and other cities is summed up to finally obtain the amount of tourism economic linkage of each city.

Secondly, in order to analyse the spatial structure characteristics of the tourism network of cities in the three northeast provinces of China, the social network analysis method was applied to analyse the network structure of the three northeast provinces of China with the help of Ucinet software, and the following four indicators were selected for the article.

(1) Network density, network density refers to the closeness of the connection between each node in an overall network, the network density of a network with N nodes and L actual connections is calculated as follows.

$$d = \frac{2L}{N(N-1)} \tag{2}$$

(2) Network centrality, network centrality mainly indicates what kind of position a point is in a network, network centrality degree includes degree centrality, closeness centrality and betweenness centrality. Degree centrality refers to the number of other points directly connected to a point; Closeness centrality depicts that the closer a point is to other points, the easier it is for the point to transmit information; and finally betweenness centrality means that a point has a high intermediate centrality if it is on a shortcut to many other pairs of points.

(3) Structural hole, a structural hole, also known as an intermediary, means that the two cannot be linked to each other and must be linked through a third person, who then occupies a structural hole in this network of relationships.

(4) Core-edge model, the core-edge model reflects a node's position in the overall tourism network, with cities in the core indicating a high position in the whole network and cities in the edge indicating a relatively low position in the whole network.

2.2 Data Source

Due to the impact of the epidemic since 2020, the article selects basic data related to tourism in the three northeast provinces of China in 2019. The study area includes the cities of Daqing, Mudanjiang, Qiqihar, Jixi, Harbin, Hegang, Ichun, Jiamusi, Heihe, Changchun, Tonghua, Jilin, Songyuan, White City, Shenyang, Liaoyang, Panjin, Yingkou, Fushun, Jinzhou, Huludao, Tieling, Dalian, Benxi and Fuxin. The total number of tourist arrivals and total revenue are from the 2019 National Economic and Social Development Statistical Bulletin of each city, and the distance between each city is based on the travel distance by car from Gaode Map.

3 Analysis of Results

Based on the survey research of big data, gravity model and social network analysis method, combined with the results of data calculated in Ucinet software, the article analyses the results of the tourism economy of the three eastern provinces of China in terms of its overall network structure and individual network structure.

3.1 The Overall Network Structure Analysis

According to the modified gravitational model formula, the tourism economic linkages between cities in the three northeast provinces of China can be obtained, the amount of tourism economic linkages between each city can be summed up to obtain the total economic linkages, as shown in Table 1, which can intuitively see the distribution of cities in the region with relatively strong tourism economic linkages in the three northeast

Table 1. The tourism economic income and the contact amount of the cities in the three northeast provinces of China

Cities	Total tourism economic income/billion yuan	Total number of tourist receptions/million people	Contact volume
Daqing	156	2703	261.707
Mudanjiang	205.4	2581.9	93.96
Qiqihar	182.7	3045.7	99.403
Jixi	83.1	1232.7	31.851
Harbin	1575.7	9544.2	1051.693
Hegang	64.2	668.8	30.002
Ichun	165.2	1782.5	57.232
Jiamusi	60.3	992	35.436
Heihe	114.7	1313.2	17.411
Changchun	2191.4	10156.5	2033.237
Tonghua	320.73	1853.1	206.149
Jilin	1184.4	7344.8	1545.52
Songyuan	180.07	1080.44	170.19
White City	95.46	530	35.509
Shenyang	853	9510	2161.74
Liaoyang	261.9	2993.5	975.854
Panjin	292.6	3597	540.925
Yingkou	305.5	3094.9	524.485
Fushun	377.7	4569	965.709
Jinzhou	253	3375.3	520.006
Huludao	246.4	2855.9	405.54
Tieling	191.5	2433.2	734.567
Dalian	1657	10268.3	527.636
Benxi	389.3	4941.4	599.734
Fuxin	117.7	1631.8	169.202

provinces of China, in which the cities with high tourism economic linkages are mostly located in the central region of the three northeast provinces of China. Conversely, cities such as Hegang, Jiamusi and Baicheng are located in the periphery of the eastern provinces and their own infrastructural development is not very good, combined with their remote location, making them generally less attractive to tourists.

According to the calculation results of the tourism economic linkage degree of each city in the three northeast provinces of China, the median value between the tourism economic linkage degree of each city in the three northeast provinces of China is 2.462,

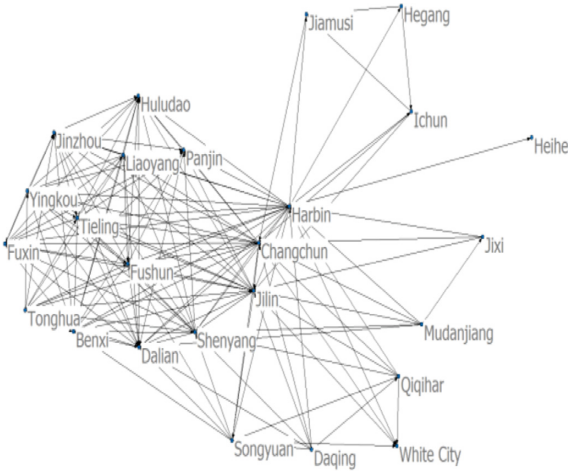

Fig. 1. Spatial structure of urban tourism networks in the three northeast provinces of China

taking this as the boundary and converting the attribute data into binary data, and the value greater than or equal to 2.462 is 1 and the value less than 2.462 is 0. Then, with the help of the Netdraw tool in Ucinet software, the network diagram of the tourism economic linkages of the three northeast provinces of China is drawn, as shown in Fig. 1. It shows that the network spatial structure of the tourism economic linkage structure of the three northeast provinces of China in 2019 is loose in Heilongjiang province and dense in Liaoning province and Jilin province, while the tourism economic network of the three northeast provinces of China is mainly centred on Shenyang, Changchun, Jilin, Harbin and Dalian, which is denser in the central part of the whole. The overall network density calculated by Ucinet software is 0.4867, whereas the theoretical maximum number of possible relationships for a network of 25 nodes is 600, the actual number is only 292, so the overall tourism economic development network in the three northeast provinces of China is not very dense.

3.2 The Individual Network Structure Analysis

By using ucinet software can calculate the centrality of each node and analyse its structural hole, from Table 2 we can intuitively see that whether from the degree centrality, closeness centrality and betweenness centrality, the top cities are Harbin, Changchun, Jilin, Shenyang and Dalian, indicating that these five cities have a more solid central position, and this indicates that these five cities have a strong central position and can play a good role in the development of the tourism economy and have a high degree of control over other cities. At the same time, it also reflects that these five cities play a strong intermediate role in realising the overall network connection.

In addition to this, as shown in Table 3, according to the data of the calculation results, it can be seen that the structural hole indicators of the five cities of Jixi, Hegang, Jiamusi, Heihe and Baicheng have one thing in common, that is, they are relatively small in terms of the effective size of the structural hole, and at the same time, the constraint

Table 2. The centrality of the tourism economic linkage network

Nodes	Degree	FreeClo	Betweenness
Daqing	0.29	0.59	0.06
Mudanjiang	0.25	0.57	0.18
Qiqihar	0.29	0.59	0.14
Jixi	0.17	0.55	0.00
Harbin	1.00	1.00	27.15
Hegang	0.13	0.53	0.00
Ichun	0.21	0.56	0.42
Jiamusi	0.17	0.55	0.12
Heihe	0.04	0.51	0.00
Changchun	0.92	0.92	12.00
Tonghua	0.50	0.67	0.00
Jilin	0.88	0.89	8.61
Songyuan	0.42	0.63	0.74
White City	0.21	0.56	0.00
Shenyang	0.75	0.80	3.07
Liaoyang	0.58	0.71	0.06
Panjin	0.58	0.71	0.06
Yingkou	0.58	0.71	0.06
Fushun	0.63	0.73	0.47
Jinzhou	0.58	0.71	0.06
Huludao	0.54	0.69	0.00
Tieling	0.58	0.71	0.06
Dalian	0.71	0.77	2.04
Benxi	0.63	0.73	0.47
Fuxin	0.54	0.69	0.00

of the structural hole in these cities is relatively high, which indicates that the cities of Jixi, Hegang, Jiamusi, Heihe and White City are located at the edge. In contrast, in the structural hole indicators of Harbin, Changchun, Jilin, Shenyang and Dalian cities, the constraint of the structural hole is lower and they have a higher effective size, indicating that these five cities are in a more central position, occupying more structural holes and being able to control interactions between other cities very well.

Table 3. Tourism Economic Linkages Network Structure Hole

Nodes	Structural Cave		
	EffSize	Efficiency	Constraint
Daqing	1.29	0.18	0.49
Mudanjiang	1.67	0.28	0.55
Qiqihar	1.57	0.22	0.49
Jixi	1.00	0.25	0.77
Harbin	13.83	0.58	0.15
Hegang	1.00	0.33	0.93
Ichun	2.20	0.44	0.61
Jiamusi	1.50	0.38	0.74
Heihe	1.00	1.00	1.00
Changchun	11.09	0.50	0.17
Tonghua	1.00	0.08	0.31
Jilin	9.76	0.46	0.18
Songyuan	3.00	0.30	0.36
White City	1.00	0.20	0.65
Shenyang	5.89	0.33	0.21
Liaoyang	1.29	0.09	0.27
Panjin	1.29	0.09	0.27
Yingkou	1.29	0.09	0.27
Fushun	2.33	0.16	0.25
Jinzhou	1.29	0.09	0.27
Huludao	1.00	0.08	0.28
Tieling	1.29	0.09	0.27
Dalian	4.76	0.28	0.22
Benxi	2.33	0.16	0.25
Fuxin	1.00	0.08	0.28

3.3 Spatial Classification of Tourist Sites

By using Core-Periphery in Ucinet, among the 25 cities in the eastern province, there are 5 cities are classified as core areas and 5 cities are classified as peripheral areas, as shown in the Table 3. Through the above analysis of the individual network spatial structure, it can be seen that the 5 cities of Harbin, Changchun, Jilin, Shenyang and Dalian, which have a good economic foundation themselves, have very satisfactory results for these city nodes, both in terms of centrality indicators and structural hole indicators, indicating that these cities can play a good role in connecting and linking the

Table 4. Core-edge model and role analysis

Analysis Method	Analysis results	
Core-edge model	Core Area	Harbin, Changchun, Jilin, Shenyang, Dalian
	Marginal Zone	Jixi, Hegang, Jiamusi, Heihe, Baicheng

entire tourism economic network structure in the eastern province, and therefore, these five cities are classified as core areas. Jixi, Hegang, Jiamusi, Heihe and Baicheng, which are located in the peripheral areas, are weakly developed and must rely on the core cities to drive their development. These cities also have less than ideal centrality and structural hole indicators, therefore these cities are classified as marginal areas (Table 4).

4 Conclusions

The article is mainly based on big data analysis, gravity model and social network analysis. The disadvantage of this method is that it is rather old-fashioned, but the main innovation lies in the novelty of the big data and the fact that few articles have chosen the eastern province as the research object.

The whole article first obtain the data of the tourism economy in the three northeast provinces of China, and uses the gravity model to calculate the tourism economic linkage degree between each city, and then adds up the tourism economic linkage degree of each city to get the total tourism economic linkage degrees of each city, so as to get the tourism economic linkage matrix, based on this, combined with the social network analysis method, the tourism economic linkage matrix is brought into the Ucinet software to carry out the overall tourism economic network structure of the eastern provinces.

Based on the big data survey, the analysis found that the tourism economic network structure of the three northeast provinces of China is relatively loose, with the degree of looseness in Heilongjiang Province as serious, indicating that the overall economic linkages of the three northeast provinces of China is not high, with the cities in Heilongjiang Province having the least tourism economic links with each other. Also comparing the table data of the above article, we can know that the cities with the best tourism economic development in the three northeast provinces of China are Shenyang, Changchun, Jilin, Harbin and Dalian, most of these cities are provincial capitals and have good economic development themselves, while the cities with poor tourism economic development like Hegang, Heihe and Jiamusi are mostly peripheral cities with poor tourism economic development, finally combining this feature with the tourism economic structure, the coordinated economic development of tourism in the three northeast provinces of China was proposed as a “core-led periphery” development model.

5 Recommendations

Firstly, the northern part of the three eastern provinces: Heilongjiang Province, which has the city of Harbin as its core area, plays a very important role in linking and connecting the overall one tourism economy of Heilongjiang Province, but the overall tourism economy of Heilongjiang Province is more loosely linked, such as the cities of Hegang and Heihe, which are relatively disadvantaged geographically due to their own more remote location. As a result, the development of Harbin plays an important leading role in the overall development of Heilongjiang Province, so the development of tourism in Heilongjiang Province relies heavily on the driving role of Harbin.

Once again, the central part of the three eastern provinces: Jilin Province, which is to have Changchun and Jilin as its core areas. Changchun City, as the capital of Jilin Province, occupies an absolute advantage in Jilin Province, both in economic terms and geographically, and with its well-developed transportation, it also provides great convenience for the development of tourism in other areas of the eastern provinces. Therefore, when developing the tourism economy, the central region should make use of the unique tourism resources and local advantages of the core cities to drive the common development of the surrounding cities.

Finally, the southern part of the three eastern provinces: Liaoning Province, which is the southernmost province of the three eastern provinces, should have the cities of Dalian and Shenyang as its core areas. Compared to Heilongjiang and Jilin Provinces, Liaoning Province has an advantage in terms of spatial distance as it is geographically located to the south, and tourists generally prefer Liaoning when visiting the three eastern provinces. Therefore, in the development of tourism in Liaoning Province, it is important to take Shenyang and Dalian as the centre and actively drive the joint development of tourism in other cities.

In summary, the main way to reduce the differences in the tourism economy of the three eastern provinces is to promote the development of the tourism economy in the core cities first, and then use the economic strength of the core cities and their own advantages to drive the development of the surrounding cities together. This will reduce the differences in the development of the tourism economy across the three eastern provinces and allow for the coordinated development of the tourism economy in the three eastern provinces.

References

1. Chen C (2021) The formation and evolution of global diplomatic access networks–a social network analysis based on big data. Foreign Aff Rev (J Foreign Aff Inst) 38(04):120–154+7–8
2. Fang Y, Huang Z, Tu W (2013) Spatial differences of tourism economy in Yangtze River Delta cities from the perspective of social networks. Trop Geogr 33(02):212–218. https://doi.org/10.13284/j.cnki.rddl.002327
3. Huang C, Li R (2014) Spatial analysis of regional tourism economic differences in Jiangsu Province. J Beijing Second Foreign Lang Inst 36(01):24–33
4. Wang C, Liu T, Wu Z, Ren H (2020) The measurement of tourism economic linkage and analysis of influencing factors in Zhejiang Province from the perspective of social network. J Hangzhou Normal Univ (Nat Sci Edn) 19(04):440–448

5. Wang J, Xu J, Xia J (2017) Research on the spatial association structure of China's regional tourism economy and its effects–based on social network analysis. J Tourism 32(07):15–26
6. Zhang H, Xia M (2011) Study on the spatial structure of tourism in Anhui Province–based on the perspective of tourism centrality and tourism economic linkage. Econ Geogr 31(12):2116–2121. https://doi.org/10.15957/j.cnki.jjdl.2011.12.030