

# Research on the Spatial Structure Characteristics of Tourist Flow Network in Guangxi

Yanhua Liu<sup>1,2</sup>, Guohua Ye<sup>2,\*</sup>

<sup>1</sup> Key Laboratory of Environment Change and Resources Use in Beibu Gulf, Ministry of Education / Guangxi Key Laboratory of Earth Surface Processes and Intelligent Simulation, Nanning Normal University, Nanning 530001, China

<sup>2</sup> School of Natural Resources and Surveying, Nanning Normal University, Nanning 530001, China

**Abstract.** Based on the online travel notes released by tourists, using social network analysis, this article analyses the spatial structure characteristics of tourist flow network under the county-level administrative units in Guangxi. The results show that: (1) The density of tourist flow network in Guangxi is sparse, and the tourist flow is dominated by landscape sightseeing tours and coastal holiday tours. (2) Nanning, Guilin, Beihai and Yangshuo are the core nodes of tourism system in Guangxi, which not only have great tourism attraction, but also have strong resource control capabilities. (3) Guangxi tourism presents a significant clustering phenomenon. The tourist flow is extremely uneven. The popularity of scenic spots has an important impact on the node flows. The level of the regional transportation network determines the central position of the node in the network to a large extent.

## 1 Introduction

Tourist flow, as a tourism geography concept with spatial attribute, shows the migration phenomenon of tourists in space locations [1]. For a long time, scholars have given a lot of attentions and researches to the tourist flow. According to the general view of researches made by scholars on the tourist flow, they have studied the temporal-spatial evolution law [2-3], network structure characteristics [4], flow pattern [5], cause and driving mechanism [6], simulation and prediction [7-8], influential effect [9-10] and so on. These researches have revealed the travel time characteristics, spatial flow track, tourism market scale and the influence of tourist flow on the economy, ecology and culture of destination from different perspectives. With the continuous improvement of tourism infrastructure and information management, mass tourism will become more and more popular, tourist flow phenomenon and tourism network system will also become more complex, which will push scholars to conduct more in-depth and detailed research on tourist flow. Exploring the characteristics of the spatial structure of the tourist flow network will help improve destination tourism management, optimize tourism planning, and expand the tourism market.

In the internet age, the internet has become an important part of people's life, study and work. From the perspective of tourism, a variety of new online media has become a general tool for tourists to understand destination information, formulate tourism strategies, publish travel notes and evaluate scenic spots. This expression with the UGC (User Generated Content) model as the core is the tourist's self-issued. Compared

with the survey questionnaire data, the online travel notes published using the UGC model are non-disturbing materials for tourists to objectively reproduce the travel experience and express their emotions freely [11]. They have the characteristics of both pictures and texts, spatiality, timeliness, authenticity and objectivity, and can restore the time and space movement trajectory of tourists, which provide a new perspective for scholars to study the movement trajectory of tourists in actual geographic space [12]. Therefore, UGC has become a very popular data acquisition method for tourist flow research, and it has also become an inevitable trend to use online UGC to study tourist flow phenomena.

Guangxi is rich in tourism resources and the tourism market is developing strongly. However, there are very few studies involving tourist flow in Guangxi. Existing studies are mainly based on statistical yearbook data to analyse the changes of tourist flow indicators, research on the coordinated development of tourist flow and regional economies, and the network characteristics of folk cultural tourism. With the continuous popularity of mass tourism, the expansion of tourism market in Guangxi is also facing huge challenges and fierce competition. It is particularly necessary to carry out researches on the spatial structure of tourist flow network in Guangxi. This article is based on the UGC data of online travel notes, from the spatial scale of county-level, combined with social network analysis methods, to analyse the characters of tourist flow network structure in Guangxi. It is hoped to further enrich the practice of tourist flow research, and provide reference for optimizing tourist flow network system in Guangxi.

\* Corresponding author: 140017@nnnu.edu.cn

## 2 Materials and Methods

### 2.1 Data collection

Using web crawler technology, 1664 valid online travel notes are collected, which were published by tourists in [www.qunar.com](http://www.qunar.com) from 2014 to 2018. Considering that some attractions are distributed linearly and span multiple districts within the city, this study considers all municipal districts of each city as a research unit. So Nanning, Guilin, Beihai and other cities mentioned in the article refer to municipal districts. At the same time, considering the special geographical location of Weizhou Island in Beihai, it is regarded as an independent research unit. 1664 travel notes contain 64 nodes, and a 64×64 multi-value directed matrix can be constructed. The idea is as follows: the scenic spot a and b belong to the county A, and the scenic spot c belongs to the county B, if there is a flow from a or b to c, the flow from A to B can be marked as 1, otherwise it will be marked as 0. If a tourist transits or stays for a short time in county D, but does not visit any scenic spots in D, the route connecting with D will be deemed invalid. The multiple accumulation between two nodes is a path flows, so as to convert travel notes into a flow's matrix. In order to eliminate the contingency that may appear in the sample data and better reflect the overall structural characteristics of the network, this article selects 2 as the path flow threshold, that is, excludes all paths with a flow of 1. Then the matrix totally consists of 36 nodes, 135 paths with 2784 tourist flows.

### 2.2 Research methods

Social network analysis method is a kind of research method to study the relationship of actors and its influence on network [13]. Based on the social network analysis method, the article selects network density and network centrality indicators to reflect the structure characteristics of the tourist flow network.

(1) Network density. Network density is used to express how close the nodes are connected to each other in the network, expressed as:

$$D = E/n(n - 1) \quad (1)$$

Where,  $D$  is the network density, the larger the  $D$ , the denser the network, otherwise, the sparser the network.  $E$  is the actual number of connections between nodes in the directed network.  $n$  is the number of nodes in the network, representing the network scale.

(2) Network centrality. The centrality of the network is characterized by the centrality of the point, which can effectively reflect the connection ability of a node and other nodes in the network. In a directed network, the centrality of a point can be refined by in-centrality, out- centrality, and point-centrality, which are expressed as:

$$C_{in}(i) = \sum_{j=1}^n c_{ji} \quad (2)$$

$$C_{out}(i) = \sum_{j=1}^n c_{ij} \quad (3)$$

$$C(i) = C_{in}(i) + C_{out}(i) \quad (4)$$

Where,  $n$  is the number of nodes in the network.  $C_{in}(i)$  is the in-centrality of node  $i$ ,  $C_{out}(i)$  is the out-centrality of node  $i$ ,  $c_{ji}$  indicates the connection from node  $j$  to node  $i$ ,  $c_{ij}$  indicates the connection from node  $i$  to node  $j$ ,  $c_{ji}$  and  $c_{ij}$  are 0 or 1,0 means that the two nodes are not connected, and 1 means that the two nodes are connected in a certain direction.

## 3 Result Analysis

### 3.1 Network characteristics

In the network of 36 nodes based on sample data, there may be up to 1260 network connections in theory, but only 135 connections actually appear, and the network density is only 0.107, indicating that the overall structure of tourist flow network in Guangxi is loose and the network density is sparse. Tourists mainly gather in Guilin, Yangshuo, Beihai. The network structure presents the characteristics of "large loose and small gathering", and tourism development in Guangxi presents extremely unbalanced spatial differences.

### 3.2 Network node characteristics

#### 3.2.1 Node flows

The node flows indexes show (Fig.1) that the top four regions in terms of outflow are Guilin, Yangshuo, Beihai, and Weizhou Island, with outflows accounting for about 69% of the total outflows. The top four regions in terms of inflow are Yangshuo, Beihai, Weizhou Island, and Guilin, with inflows accounting for about 64% of the total inflows. The top four regions for flows are Yangshuo, Guilin, Beihai, and Weizhou Island, with flows accounting for about 67% of the total flows. The results show that Guilin, Yangshuo, Beihai, and Weizhou Island are very popular tourist destinations in Guangxi, and these places have a high tourism reputation. Although Nanning has a 5A scenic spot, it is less well-known.

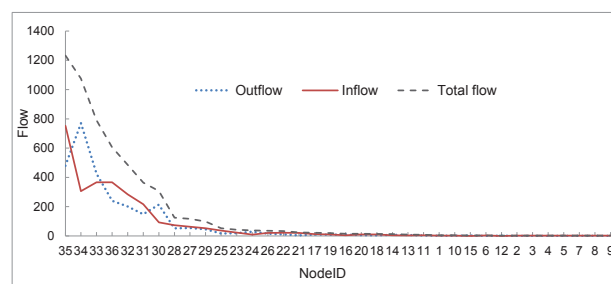
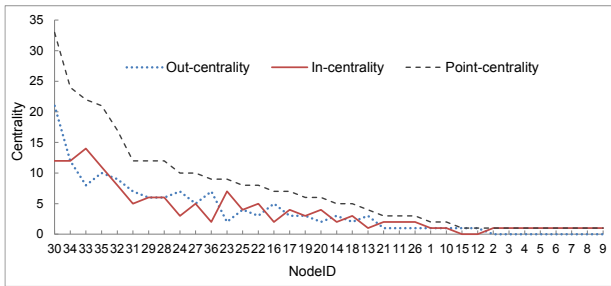


Fig. 1. Flows Statistics of Each Node.

#### 3.2.2 Node centrality

The node centrality indexes show (Fig.2) that the top four regions with out-centrality are Nanning, Guilin,

Yangshuo and Longsheng. The top four regions with in-centrality are Beihai, Nanning, Guilin, and Yangshuo. The top four regions with point-centrality are Nanning, Guilin, Beihai, and Yangshuo. The results show that Nanning, as the capital of Guangxi, has the most convenient transportation and the strongest connectivity in the network, which determines its position with the highest centrality. Weizhou Island has a large number of tourists, but its transportation is very inconvenient, which greatly reduces the node centrality. Therefore, the node centrality is closely related to the traffic network status.



**Fig. 2.** Centrality Statistics of Each Node.

### 3.3 Network path characteristics

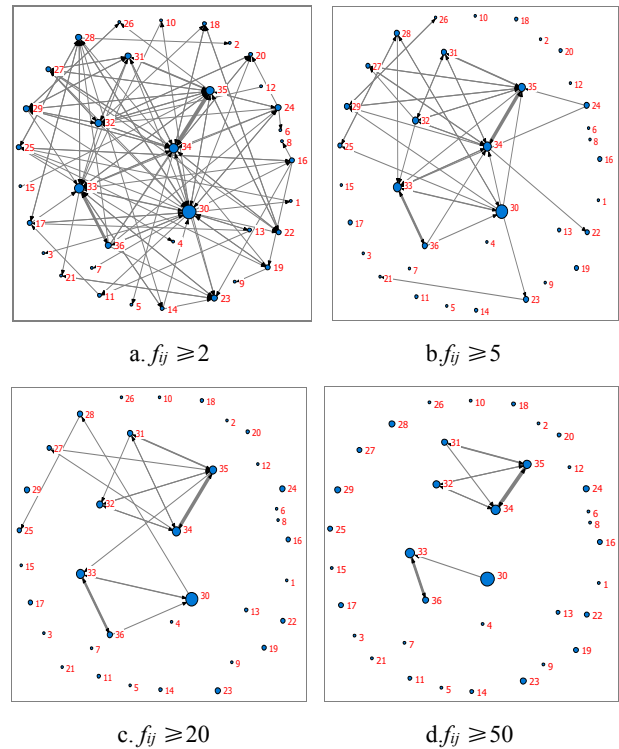
According to the path flows in this study, the number of network nodes (NN), the number of paths (NP), the network density (ND), the total flows (TF), the average flows (AF) and the proportion of the flows (PF) are calculated according to four different flows thresholds (Table 1), and the network structure diagram are drawn under the corresponding threshold (Fig.3).

**Table 1.** The statistical characteristics of the flow network

Threshold	NN	NP	ND	TF	AF	PF(%)
2	36	135	0.107	2784	20.62	95.3
5	18	59	0.193	2588	43.86	88.6
20	10	22	0.244	2252	102.36	77.1
50	7	12	0.286	1936	161.33	66.3

In the original network, there are 64 nodes and 271 paths, the network density is only 0.067. When the threshold is 2, there are only 36 nodes and 135 paths in the network (Fig.3a), but the flows still account for 95.3%, which shows that selecting 2 as the threshold for network analysis is appropriate and does not affect the overall characteristics of the network. With the increase of the threshold, the number of nodes and paths continue to decrease, the network density and the average flows of paths continue to increase, and the network flows show a slight decrease trend (Fig.3b-c). When the threshold is 50, there are only 7 nodes and 12 paths left in the network (Fig.3d), but the network flows still account for 66.3%. At this time, the tourist flow network structure is divided into two parts, one part is a sub-network composed of four nodes in Guilin, Yangshuo, Longsheng, and Lipu. The flow path is embodied in the

two-way flow between Guilin, Yangshuo and Longsheng three nodes, as well as the two-way flow between Yangshuo and Lipu, and the one-way flow from Lipu to Guilin. The other part is a sub-network composed of three nodes in Nanning, Beihai, and Weizhou Island. The flow path is reflected in the two-way flow between Beihai and Weizhou Island and the one-way flow from Nanning to Beihai. It can be seen that the tourist flow in Guangxi shows obvious regional agglomeration and imbalance. More than half of the tourists are concentrated in the above two sub-networks, and the connections between other nodes are very sparse, and the phenomenon of spatial polarization is significant.



**Fig.3.** Networks with different threshold (a-d).

## 4 Discussion

The flows and centrality of the nodes indicate that the flows and centrality of Guilin, Yangshuo, and Beihai are relatively high, and they maintain good consistency. Nanning, the capital of Guangxi, has the highest point-centrality, but has a small tourist flows. It shows that the flows of nodes are not completely consistent with the centrality, that is, the node with the most tourist flows is not completely the node with the highest centrality. Through the Baidu index search, it is found that if the existing 5A scenic spots in Guangxi are searched with the corresponding "city+scenic spot", the average Baidu indexes from 2014 to 2018 are all lower than 300. However, if the Baidu index is searched with "Guilin landscape", the average Baidu index from 2014 to 2018 reaches 2,331. In addition, the Baidu index search are conducted with "Beihai Silver Beach", "Beihai Weizhou Island", and "Yangshuo West Street", the average Baidu index are 1109, 583, and 961 respectively, which are much higher than that of the 5A scenic spot. It can be

seen that the relationship between node flows and the level of scenic spots in the region is not obvious, and is closely related to the popularity of scenic spots. It is well known that “Guilin landscape tops those elsewhere” and “Yangshuo landscape tops that of Guilin”, Beihai Silver Beach enjoys the reputation of "Chinese No. 1 Beach". Weizhou Island is the largest and youngest volcanic island in China, many tourists regard these areas as one of their destinations in Guangxi. Therefore, it is necessary to vigorously strengthen the superimposed publicity of the city and the scenic spot, and each administrative region will focus on promoting a scenic spot or tourism highlight, and create a "city+scenic spot" or "city+tourism highlight" business card to increase the reputation of the city and the scenic spot to attract more tourists.

## 5 Conclusions

(1) The density of tourist flow network in Guangxi is very low, and the spatial difference of network structure is obvious. Tourists traveling to Guangxi have obvious agglomeration characteristics, mainly concentrated in two clusters, a cluster composed of four nodes of Guilin, Yangshuo, Longsheng, and Lipu, and another cluster composed of three nodes of Beihai, Weizhou Island, and Nanning. Tourists mainly prefer landscape sightseeing tours and coastal vacation tours. Other types of tourism, such as ethnic minority tourism, China-Vietnam border customs tourism, cultural relics tourism, forest park tourism, leisure and health tourism, and red tourism are relatively small and fail to show a large tourism pattern and diverse tourism trends.

(2) In the network system, Nanning, Guilin, Beihai, and Yangshuo are the core nodes, which have strong tourism attraction and resource control capabilities. The flows of a node have little correlation with the grade of the scenic spots it contains, and are highly related to the popularity of the scenic spots. The higher the popularity of the scenic spot, the more flows in its area. The centrality of a node is related to the popularity of its scenic spots and regional traffic conditions. The level of transportation network is the key factor of node centrality. At present, the role of the core nodes mainly lies in their own tourism attraction and resource control. The relationship with the surrounding nodes is not close enough, and it fails to reflect the radiation ability of core nodes as regional tourism centres. Therefore, in the future, it is necessary to take advantage of the tourism resources of the core nodes to rationally plan the tourism routes in the surrounding areas, to give play to the tourism radiation and diffusion capabilities of the core nodes, and to further increase the density of the tourism network within a certain range around the core nodes.

## Acknowledgments

This article is financially supported by the Guangxi Natural Science Foundation (2020GXNSFAA159065), the Seventh Batch of Distinguished Experts in Guangxi,

and the Opening Fund of Key Laboratory of Environment Change and Resources Use in Beibu Gulf, Ministry of Education, Guangxi Key Laboratory of Earth Surface Processes and Intelligent Simulation, Nanning Normal University (GTEU-KLOP-K1701). Special thanks are given to all authors who provided references for this research. The authors also thank Nanning normal university for its no-cost literature resources.

## Appendix

**Table A1.** Comparison table of NodeID and Node name

ID	Name	ID	Name	ID	Name
1	Tiandong	13	Chongzuo	25	Jingxi
2	Fusui	14	Qinzhou	26	Hezhou
3	Wuming	15	Wuzhou	27	Lingchuan
4	Napo	16	Baise	28	Daxin
5	Quanzhou	17	Pingxiang	29	Zhaoping
6	Rongshui	18	Ziyuan	30	Nanning
7	Shanglin	19	Bama	31	Lipu
8	Leye	20	Sanjiang	32	Longsheng
9	Fengshan	21	Dongxing	33	Beihai
10	Yulin	22	Xing'an	34	Guilin
11	Ningming	23	Fangchenggang	35	Yangshuo
12	Gongcheng	24	Liuzhou	36	Weizhou Island

## References

1. Y.Z. Yang, C.L. Gu, Q. Wang, *Geographical Research* **30**, 23 (2011)
2. Y.H. Shao, S.S. Huang, Y.Y. Wang, et al, *Tour Manag Perspect* **36**,100752(2020)
3. T.H. Wan, H.P. Huang, *World Regional Studies* **23**, 128 (2014)
4. G. Xiao, W.M. Li, C. Zhang, *Areal Research and Development* **39**, 88 (2020)
5. L.N. Zhong, S. Sun, R. Law, et al, *Sustainability* **12**, 9125 (2020)
6. F.G.Santeramo,M.Morelli, *Curr Issues Tour* **19**,1077 (2016)
7. A.P.Zhang,Y.H.Liu, L.S. Zhong, et al, *Resources and Environment in the Yangtze basin* **24**,364(2015)
8. K.Q. Li, C.Y. Chang, W.X. Lu,et al, *ISPRS Int J Geo-Inf* **9**,676(2020)
9. M.S. Gallego, J. Francisco, L. Rodríguez,et al, *Econ Model* **52**,1026(2016)
10. A.Kolokotsa,M.Leotsinidis,L.Kalavrouziotis,et al, *J Appl Microbiol* **130**,516(2020)
11. J.Wang, J. Hu, Y.Y. Jia,et al, *Economic Geography* **36**,176(2016)
12. Y.Y. Zhang, J.Y. Li, M. Yang, *Human Geography* **129**,111(2016)
13. L. Adamic, A.Eytan, *Soc Networks* **27**,187(2005)