Empirical Study on the Temporal-spatial Behavior of Tourist in Tourist Attractions Based on Geography Tagging Technique

Yongfeng Yang

College of Geography and Tourism, Chongqing Normal University, Chongqing 401331, China

Abstract

The data obtained by the former ways cannot fully and accurately record the temporal-spatial change and traveling route of tourists. The rapid development of geographic information system and Internet sharing platform provides more advanced technical support for further tracking and collecting tourist data. All kinds of digital cameras have become one of the most essential electronic products for most people to travel, so that they will keep a beautiful moment and frame the beautiful scenery along the way. During taking the photograph, the time of the photo and the detailed geographic coordinates are recorded at the same time. In order to make more accurate empirical analysis temporal-spatial behavior of tourist, taking the tourist attractions in Xi'an as the research area and based on the traffic condition, travel distance and terrain, scenic spots distribution, this paper obtains photo data of geography tagging in research range from Flickr and Panaramio through the relevant technical data combined with previous research. And based on guaranteeing the integrity of the tourist trajectories reflected by the photo information, the final data range is delimited by ArcGIS.

Keywords: Geography Tagging, Tourist Behavior, Temporal-Spatial, Model.

1. BACKGROUND OF RESEARCH

1.1 Literature review

Tourists are the main service object of the tourism activities and tourism industry. Only to understand the characteristics and causes of tourists' behavior, can we find out the methods of developing the tourism industry, strengthening the scenic transportation construction, enhancing the level of tourism services. Therefore the study on temporal-spatial behavior of tourist is always the core of tourism research (Yang and Jiang, 2014). Although the content of temporal-spatial distribution of tourists is relatively rich in tourism research, there is a lack of macroscopic, systematic and in-depth study on dynamic changes of temporal-spatial behavior of tourist, occurring mechanism with deep reason and cultural exchanges. Especially there is a lack of theoretical innovation and empirical study of tourist space-time dynamic model, decision of temporal-spatial behavior and behavior differentiation. Driven by the wave of Internet development, the emergence of 3G, the Internet concept and the corresponding technology system provide users with experience, diversity and interaction services. The new era of geographic information is coming. With the advent of Flickr and Panaramio, all Internet users can upload and share live photos on the same space service platform, making ordinary users become data providers (Miao and Yang, 2014). Geographically tagged photo information provided by network photo sharing platform can be realized data visualization through GIS to display tourists' spatial distribution and behavior track and provide decision-making reference for managers through data processing and analysis.

1.2 Purposes of research

In this paper, based on the geographically tagged photo data with a relatively high authenticity, the distribution rule of tourists in time, space and travel path can be reflected comprehensively and accurately. The research area will be locked in Xi'an tourism circle based on the scope defined by the previous administrative division and other scenic spots. The temporal-spatial behavior of tourists will be made comprehensive, refined and profound research by using the tourist route recorded through geography tagged photos of tourist (Zhu and Hu, 2016). Its regularities of distribution and trajectory characteristics in time and space are summed up in this paper. Finally, the detailed location and function of spatial network structure of tourism destination are measured and comprehensively evaluated.
2. RESEARCH METHOD AND TECHNICAL ROUTE

2.1 Main research methods

2.1.1 The method of space hotspot analysis

The method of space hotspot analysis is a spatial clustering method for identifying high value (hot spots) and low value (cold point) with statistical significance. First, the network of squares is generated by using ArcGIS. Then the layer of squares network is stacked with the dot element layer of the research area to get the geographic information in each unit (Zhang and Zhao, 2017). The data distribution of "hot spots" and "cold points" is explored by using the Hotspot Analysis (Getis-Ord Gi*) method. The local statistics of formula for Getis-Ord Gi* is:

\[ G_i^* = \frac{\sum_{j=1}^{n} W_{ij} X_j - \bar{X} \sum_{j=1}^{n} W_{ij}}{\sqrt{\sum_{j=1}^{n} W_{ij} - \left( \sum_{j=1}^{n} W_{ij} \right)^2}} \]  

Among them, the attribute value of the element J is \( X_j \), the spatial weight between the element i and j is \( W_{ij} \), and the n is the sum of the elements. Among which:

\[ \bar{X} = \frac{\sum_{j=1}^{n} X_j}{n} \]  

\[ S = \sqrt{\frac{\sum_{j=1}^{n} X_j^2}{n} - \left( \bar{X} \right)^2} \]  

The \( G_i^* \) statistics that return for each element of the data set is the score of \( z \). For the positive score of \( z \) with statistical signficance, the higher the \( z \) score is, the closer the clustering of the high efficiency (hot spot) is. For the significant negative score of \( z \) with statistical characteristic, the lower the \( z \) score is, the closer the clustering of the low value (cold point) is. ArcGIS uses \( z \) score, P value and confidence interval to create a new output feature class for each element in the input element feature class, and applies the default rendering to the Gi Bin field from hot to cold (Wu and Xiang, 2017). In this paper, the method of space hotspot analysis is used to identify the hot spots and cold spots in the spatial distribution of Xi'an tourism circle in different periods, so as to reflect the distribution characteristics of tourists in different periods.

2.1.2 The method of kernel density analysis

In order to get a better understanding of the concrete distribution of tourist attractions in Xi'an, one-dimensional kernel density estimation is introduced in this study. It is assumed that the spot geographic elements can be distributed in any space with different probability in this method. The distribution probability of spot geographical elements is different in aggregation and scattered areas. The former is larger, and the latter is smaller (Guo and Jia, 2017). In the process of computation, kernel function is used to create a smooth surface around every raster point with the circle as neighborhood. The point geographic elements falling into search area will be given different weight values. The closer to the grid center is, the larger the weight value is and the higher the kernel density value is and reduce to the periphery. When the density value at threshold range within a certain distance away the network center reduces to 0, the formula model is:

Assuming that the random sample \( X_1, X_2, ..., X_n \) is the distribution of the density function \( f(x) \), its calculation is estimated to be:

\[ \hat{f}(x) = \frac{1}{n \Delta x} \sum k \left( \frac{x-x_i}{\Delta x} \right) \]  

Among them, the \( f(x) \) is one of the accounting estimates for the total density \( f(x) \), \( K (\cdot) \) is the accounting function, \( h_n \) is the bandwidth, and \( n \) is the sample size within the range of the search area.

3. EMPIRICAL STUDY ON THE TEMPORAL-SPATIAL BEHAVIOR OF TOURIST IN TOURIST ATTRACTIONS BASED ON GEOGRAPHY TAGGING TECHNIQUE
3.1 Definition of tourism circle

Tourism circle is the spatial organization and product form of regional tourism cooperation. It first appeared in the early 90s. It was introduced into the tourism industry by the way of metropolitan area or city circle. It is a native word in China, which is used to express the phenomenon of inter-regional tourism cooperation. (Gao and Zhao, 2015) But in fact there is not a clear definition on it in the academic circles. In this paper, the common cognition among different opinions of scholars are summarized and the main features of the tourism circle are as following: (1) tourism circle is the spatial organization of interaction among various factors; (2) there is hierarchy in tourism circle; (3) tourism circle is an important form of regional tourism cooperation; (4) the tourism circle is constructed around one or more kernel with a big city, or tourism resources or tourism hub; (5) the tourism circle forms a close connection with the link of traffic. The figure 1 below is the definition of tourism circle range in Xi'an.

![Figure 1. Definition of Xi’an Tourism Circle](image)

3.2 The distribution characteristics of time behavior of tourists

3.2.1 The distribution characteristics of four seasons.

![Figure 2. Four Seasons Distribution of Xi'an Geographically](image)

According to the climatic characteristics of Xi'an, the winter of the Xi'an tourism circle is in December, January and February, the spring is 3-5 months, the summer is 6-8 months, and the autumn is 9-11 months. The seasonal percentage of geographically tagged photos is shown in Figure 2. The most geographically tagged photos of Xi'an tourism circle are shown in summer and more than 1/3 of total annual photos from summer, which indicates that the number of tourists in summer is the largest (Wu and Yang, 2015). Followed by autumn, accounting for 27% of the total annual number. The difference between spring and winter is not great, only three percentage points,
but the difference between two seasons of spring and winter and summer and autumn is obvious, which indicates that spring and winter are the off-season of Xi'an tourism circle.

According to the figure 3 blow, we can see that it is a multimodal seasonal form in the tourism circle of Xi'an. There are four peaks in the geographically tagged photos of Xi'an tourism circle. They are the January and December in winter, April and May in spring, July and August in summer, October in autumn. In the whole year, more visitors go to visit Xi'an tourist attractions from 4 to 10 months and the number of photos is all more than 1500. In October, the climate in Xi'an is pleasant. And due to the National Day holiday, many scenic spots will appear the phenomenon of tourist blowout and the number of tourists reached the peak of the year. (Ding and Li, 2015) From 7 to 8 months, although the climate is hot, many teachers and students companied with relatives and friends to participate in tourism activities due to summer holiday. Qinling Mountains is suitable for people to spend summer so that tourists reached a second peak. The climate of 4-6 months in Xi'an is pleasant. It is a good time for people to travel, and the number of photos in the three months is on average, which is the third peak.

Figure 3. Monthly Changes in Xi'an Geographically Marked Photos

3.2.2 Difference of seasonal distribution of tourists

In China, the seasonal intensity index and Gini coefficient are used to measure the degree of imbalance in the tourism season. (1) Based on seasonal intensity index: in order to illustrate the seasonal impact on tourist flow in Xi'an tourism circle, the seasonal intensity index R is introduced.

\[
R = \sqrt{\frac{\sum (Y_i - 8.33)^2}{12}}
\]

(5)

Among them, R represents the concentration index of passenger flow, and \(Y_i\) represents the percentage of passenger flow per month in the whole year. The larger the R is, the larger the seasonal difference of passenger flow is; the smaller the R value is, the more balance in distribution of the passenger flow throughout the year is. In this paper, due to the large number and randomness of photos, the percentage of photos taken every month accounting for the percentage of the total number of photos taken every year is used to reflect the percentage of passenger flow volume per month accounting for the whole year. (Huang and Shen, 2015) And then \(R = 3.52\) can be obtained, indicating that the tourist flow of tourist attractions in Xi'an is affected to some extent by the season. (2) Based on Gini coefficient: compared with Gini coefficient, the defect of seasonal intensity index is R's value range \([0, +\infty]\), which can only reflect the seasonal variation by comparison, and the measurement result is relative. Gini coefficient is the absolute value of the seasonal distribution of tourist and the value range is \([0, 11/12]\). The greater the Gini coefficient is, the more centralized the number of photos in time distribution is. The closer to zero the Gini coefficient is, the more average the time distribution of photographs quantity is, and it is very little affected by the extreme value in the view of the calculation of the standard deviation. (Yang and Li, 2015) Its calculation method is:

\[
G = \frac{2}{n}\sum_{i=1}^{n}(X_i - Y_i)
\]

(6)
Among them, $G$ represents the Gini coefficient and $n$ is the number of month, that is to say $n=12$. $i$ represents the number of photos among 12 months in ascending order and $X_i$ is the ratio of ranking number and total number of month, that is to say $X_i = i/n$, $Y_i$ represents the total frequency of misdeeds in Lorenz curve. After the formula is simplified, it can be obtained:

$$G = \frac{2}{n} \left[ \sum_{i=1}^{n} if_i - \frac{n+1}{2} \right]$$ (7)

Among them, the $f$ is the ratio of the number sum of geographically tagged photos taken in the month ranking $i$. The result is $G=0.23$, which indicates that the number of tourists in the tourist attractions of Xi'an shows a certain seasonal characteristic and centrality in the time distribution (Tao and Fu, 2016). Although the difference is not obvious, it should also be attached importance to them, which is similar to the results of the seasonal intensity index analysis.

3.3 The distribution characteristics of spatial behavior of tourists

3.1.1 Spatial distribution density of 3.1 tourists in 2012-2016

The hot tourist attractions of Xi'an tourism circle in 2012-2016 is mainly distributed in the Xi'an urban district. The number of photos accounts for 58.4%, mainly concentrated in the train station, Bell Tower and Drum Tower, Muslim's Quarter, Mosque, Walls, Beilin Museum, Xingqing Palace Park, Xingqinggong campus of Xi'an Jiao Tong University, Small Goose Pagoda, Da Ci'en Temple, Wild Goose Pagoda, Lotus palace of Tang Dynasty. The sub hot tourist area is mainly distributed in the Eastern Line scenic area and Huashan scenic area mainly based on the Huaqing Pond, The Tomb of Emperor Qinshihuang, Terracotta Warriors and Horses, Lishan forest park. The Eastern Line scenic spot accounts for 14.3%, and the Huashan scenic area is 13.3%. Cold spot tourist areas are mainly distributed in Zige natural scenic spots, Caotang temple, Lianzhu Lake scenic area, Cui Huashan National Geological Park, Jia Wu Tai scenic area, Nanwutai scenic areas of the south of Qinling Mountains tourist scenic; the sub cold tourism areas are mainly distributed in the Taiping National Forest Park and Zhongnanshan National Forest Park (Zheng and Wu, 2016). The general tourist areas are mainly distributed in the west tourism district of Famen Temple tourist area, Qianling Mausoleum, Zhao Mausoleum, Hammao Mausoleum, Han Yangling Mausoleum, Qinling Mountains tourist area of Taibai Mountain tourism area, Louguantai, Zhuque Forest Park, Shaohuashan Forest Park and the north line tourist area of Yaowangshan Mountain. The detailed distribution is shown in Figure 4.

3.2 The degree of spatial agglomeration of tourist attractions
There are three forms of dot elements in space, random distribution, uniform distribution and aggregated distribution. The nearest neighbor index can help us accurately estimate the distribution of the scenic spots. The nearest neighbor ratio is defined as:

\[ ANN = \frac{\bar{D}_O}{\bar{D}_E} \]  

Among them, \( \bar{D}_O \) refers to the average value of average observation distance and \( \bar{D}_E \) refers to the expected average distance. The calculation formula of \( \bar{D}_O \) is: \( \bar{D}_O = \frac{\sum_{i=1}^{n} d_i}{n} \), and the calculation formula of \( \bar{D}_E \) is: \( \bar{D}_E = \frac{0.5}{\sqrt{n/A}} \). Among them, \( A \) represents the area of spatial distribution of spot geographical elements, which refers to the overall area of the tourist attractions in this paper. The area is calculated by ArcGIS for about 297.58 square kilometers, and \( n \) represents the number of spot geographic elements in the study area (Ding and Wu, 2016). When \( ANN=1 \), and \( \bar{D}_O = \bar{D}_E \), geographical dot elements in the study area are randomly distributed; when \( ANN>1 \), that is \( \bar{D}_O > \bar{D}_E \), geographical dot elements in the study area are dispersedly distributed; when \( ANN<1 \), that is \( \bar{D}_O < \bar{D}_E \), geographical dot elements in the study area are distributed in aggregation state.

4. THE TRACK CHARACTERISTICS OF PATH OF TOURISM FLOW

4.1 Track model of single node path

In this paper, there are 900 tourists on a trip who just upload a photo or upload several photos with same geographical marker at the same position and leave the single node path, which accounts for 57.8% of total tourists; There are 7069 photos shoot by them accounting for 21.4% of the total number of photos; everyone takes 5.4 photos in average. Among all the counties, the number of photos with 2713 in Lianhu District is the most prominent accounting for 51% of the single node path. It shows that the considerable cultural relics, traditional blocks, special snacks and religious buildings in Lianhu District have great attraction for tourists. Beilin District is ranked on second and the number of photos is 739 accounting for 16.4% of the single node track. The well-preserved city wall of the Ming Dynasty, the Small Goose Pagoda, Wolong temple, Hua Pagoda of Baoqing Temple, and Temple of the Eight Immortals as well as prosperous social and commerce activities make Beilin District called "golden treasure" more popular. The Weiyang District is ranked at third and the number of photos is 849 accounting for 14.6% of the single node path. Its resources advantage is the ancient city ruins with Changan city of Han Dynasty as the representative. Figure 5 below is the spatial distribution of path tracks of single node tourists.

![Figure 5. Spatial Distribution of Path Trajectories of Single Node Tourists](image)

4.2 Evaluation of node effect
4.2.1 Degree centrality

Degree centrality is used to measure the number of connections between a node and other nodes, indicating the position and role of a node in the network structure (Yan and Tang, 2016). In the network structure of tourist flow, the degree centrality of introversion refers to the total number of tourists flowing from other nodes into a node and is used to measure the concentration degree of tourism node; and the degree centrality of extroversion refers to total number of tourists flowing one node into other nodes and is used to reflect the degree of divergence for tourism node. The expression is as following:

\[
C_{D\text{-in}}(n_i) = \sum_{j=1}^{n} n_{ji}
\]

(9)

\[
C_{D\text{-out}}(n_i) = \sum_{j=1}^{n} n_{ij}
\]

(10)

Among them, \(C_{D\text{-in}}(n_i)\) represents the inward centrality of the node \(n_i\) and \(n_{ji}\) represents the number of node \(n_j\) flowing to \(n_i\); \(C_{D\text{-out}}(n_i)\) represents the extroverted centrality of the node \(n_i\) and \(n_{ij}\) represents the number of node \(n_i\) flowing to \(n_j\).

Betweeness centrality

The betweeness centrality is used to measure the mediating capability of the node and to reflect the degree of control of a node to other nodes. The expression is as following:

\[
C_B(n_j) = \sum \frac{g_{jk}(n_i)}{g_{jk}}
\]

(11)

Among them, the \(C_B(n_j)\) represents the betweenness centrality of the node \(n_j\). \(g_{jk}\) is the shortest path number from the point J to the point K and \(g_{jk}(n_i)\) represents the shortest path number through the point \(n_i\) in the shortest path from point J to point K.

4.2.3 Closeness centrality

The closeness centrality is used to measure the closeness degree between one node and others and the node which is investigated is not controlled by other nodes. In the tourist flow network, the smaller the value is, the better the traffic accessibility of a node and other nodes is and the closer the contact degree is. Because the network matrix after binarization in the research is connectivity graph, that is there are some nodes could never get to another node in the network, small groups of Liquan County and Jingyang County as well as isolated node are deleted, and the closeness centrality of the main network is calculated. The expression is as following:

\[
C_{D\text{-in}}(n_i) = \sum_{j=1}^{n} d(n_j, n_i)
\]

(12)

\[
C_{D\text{-out}}(n_i) = \sum_{j=1}^{n} d(n_i, n_j)
\]

(13)

Among them, \(C_{D\text{-in}}(n_i)\) represents the inward closeness centrality of the node \(n_i\) and \(d(n_j, n_i)\) represents the shortest path from node \(n_j\) to \(n_i\); \(C_{D\text{-out}}(n_i)\) represents the extroverted closeness centrality of the node \(n_i\) and \(d(n_i, n_j)\) represents the shortest path from node \(n_i\) to \(n_j\).

5. CONCLUSION

According to the acquisition, process and analysis of the geographically tagged photos, the author thinks that the way of the geography tagged photo data can record the tourist’s track well. But in the future, if we want to do further research based on this way, we need to start with improving the source of data. Affected by the policy, location of tourism area, distribution of resources, economic development, transportation construction, infrastructure construction and many other aspects, the research on tourist’s path track and extension of tourism circle structure are very complex issues needing interdisciplinary research. The future research should dig into the formation of tourist path and the mechanism of the evolution of the tourism circle. Only by providing good photo sharing platform for tourists can we provide more comprehensive and representative data for the research.
ACKNOWLEDGEMENTS

Social science research project of Chinese Education Department, Tourism supply and demand contradiction and coordination mechanism under the background of Golden week in China, No. 13YJC790180. Tourism supply and demand contradiction and coordination mechanism under the background of Golden week in China, No. 14XYY006.

REFERENCES

Huang W.W., Shen F. (2015), An overview of the research on marginal tourist sites and Prospect, Human geography, 30 (04), 24-31.
Zhang J.R., Zhao Z.B. (2017). The expression of uncivilized behavior of tourists based on the "field theory", Resource science as an example of the portrayed the Badaling Great Wall, 39 (07), 1291-1302.
Zheng T.X., Wu R. (2016). Based on computer reasoning technology, the decision support system of theme park tourists' time and space diversion is studied - design, simulation and comparison, Tourism science, 30 (01), 60-77.