A LBS Supported Mining System for Self-Help Tourism—A Chinese Case Study

Qiu-hua HUANG

Department of City and Environment Science, City College of Dongguan University of Technology, 523106, China

Keywords: LBS, Self-Help tourism, Trajectory, Spatial-Temporal behavior.

Abstract. This paper analyses the mobile positioning data used in researching self-help tourism. Mobile positioning data which is stored in the memory files of mobile device with GPS positioning function are as follows: the location, date, total time and etc. From this paper we can find that mobile positioning data has meaningful applications for the development of tourism. This paper proposed models to calculate the experience value of tourist and the attraction index of trajectory. The results illustrate that location data has virtues: location data can be collected conveniently; the precision of which is higher than traditional statistics.

Introduction

Self-help tourism is of growing importance for the global economy, since positioning technology has risen quickly during recent decades. In operationalizing tourism flows, a majority of the research emphasizes the visitors’ behavior, as well as their impacts on the physical environment around them in recent decades[1]. The rapid development of the information and communication technologies (ICT) such as digital databases and geographical information science (GIS), which are advancing surveying methods in self-help tourism studies. The GIS have lots of applications in tourism studies[2-5].

There is a hot subject which connecting with mobile (cellular) phone positioning datasets and location-based services (LBS) [6]. Mobile positioning data becomes more and more important in the studies of spatial-temporal behavior. In the field of tourism scholarship very few experts studied the mapping and modeling of tourist movements[7-8]. There are some studies about the spatial effects of differences in attraction site visits [7,9] , but very few about the characteristics of travel behavior of self-help tourist.

The objective of this paper is to study the characteristics of the travel behavior of self-help tourist. In this study self-help tour were divided into seven kinds, e.g. self-driving tourism, mountaineering, mountain biking, hiking, high way biking, sightseeing and others, among them mountaineering and hiking accounted for the largest proportion. In this paper the trajectories were divided into four typologies based on their shapes, e.g. (T1) Echoing Trajectory; (T2) Branch Trajectory; (T3) Closed Trajectory; and (T4) Combination Trajectory, among them Echoing Trajectory is the most common type, which accounted for 52.6% of the total. At last, this paper proposed models to calculate the experience value of tourist and the attraction index of trajectory. Travel enthusiastic can choose the most popular track with the shortest time using trajectory’s attraction index. There are a huge amount of self-help tourism trajectories available on the web portals, which can be used in the studies of the characteristics of self-help tourists’ space-time behavior.

Data and Methods

Data

The database is downloaded from the web portal[14], which consisting of 445 trajectories of 162 self-help tourists from China between August 1, 2010 to July 31, 2011. The province of the self-help tourist is determined by the user account information or by his tourism trajectories.
Methods

The trajectories achieved from mobile positioning devices have some features, for example the low cost and the high precision than traditional methods. The tourists flows have been studied by some experts in some specific environments, for example island environments[7], urban locations[9], heritage attractions[10] and crowded environments[11]. Nevertheless, there are very few studies about the mapping and modeling of self-help tourist movements, and the spatial-temporal behavior of self-help tourist. The data acquisition method of mobile positioning have become mature, but there are still a lot of people who wouldn't wear a locator for privacy reasons. From the web portal[14] we can get some trajectories of self-help tourism, which were caught by mobile positioning.

Location data can be analyzed using many different methods of spatial-temporal behavior analysis[12-13]. Unfortunately, the self-help tourism trajectories only have a few features for each trajectory: date of start of trajectory, province city where trajectory start, self-help tourism typology what the trajectory belong to, the total length of the trajectory, trajectory’s total time, the self-help tourist ID of the trajectory which allows recognition of the same tourist of different trajectories, Latitude and longitude of the track regional center, the shape of the trajectory and the degree of difficulty of the trajectory. In this study, we use the following software: ArcGIS 10.0, Microsoft Excel, and Spss 12.0.

Results

In this study there are 445 trajectories, which were obtained by 162 self-help tourists (user IDs) from China during August 1, 2010 to July 31, 2011.

The Movement Patterns of Backpackers

From the data we can see that, every trajectory taken by a backpacker has linear path characteristic. Every trajectory is connected by a starting point, a number of mid-points and an end point. Some experts divided the linear itinerary into four patterns, they are as follows: (a) a circle tour; (b) a single destination and return trip; (c) multiple stop circle tour; and (d) a complex combination[15]. In this paper the trajectories were divided into four movement patterns based on their shapes in Fig.1. They are (T1) Echoing Trajectory; (T2) Branch Trajectory; (T3) Closed Trajectory; and (T4) Combination Trajectory.

Type T1: Echoing Trajectory. This pattern is unique for backpackers to travel from a start to a destination. About this pattern one or more attractions can be visited by backpackers before they arrive in the destination. This pattern is the most common for a short time travel, because there’s no limitation of the terrain and it’s relatively easy to perform. Some experts studied this pattern in the automobile tourists at regional Australian destinations[16].

Type T2: Branch Trajectory. This movement pattern sometimes is called repetitive point-to-point[15]. In this pattern backpackers travel to the same attractions twice during their travel. When limited by the terrain, you have to choose this pattern. In this pattern the start point and the end point may be the same one or not, and the travel time can be easily arranged. Backpackers don’t like to choose this movement pattern if condition permitted, because they don’t want to waste time to travel the same route.
Type T3: Closed Trajectory. This movement pattern starts and ends at the same or near point, and includes a set of attractions in a circular route. This may be the best efficient movement pattern considering the attractions visited and the transportation network established. It is the most widely used pattern often[15]. Although backpackers like to choose this movement pattern, it is limited by the terrain and the travel time is difficult to be estimated in this pattern.

Type T4: Combination Trajectory. This pattern could be a combination of one or several of the Echoing Trajectory, Branch Trajectory and Closed Trajectory. Personality style may influence this movement pattern, for example backpackers who eschew large environmental bubbles might have a higher tendency to wander. These trajectories are flexible, opportunistic, and process oriented. While on the surface the trajectories may seem chaotic, there may be an underlying logic to their movements[15].

In this study the number of each pattern is shown in table 1. These data illustrates that Echoing Trajectory is the most common type, which occupied 52.6% of the total; then is the Closed Trajectory, which occupied 21% of the total.

Table 1 also showed that T1 is the main trajectory type in mountaineering and hiking, which accounted for 63.1% and 61.2%; T1 and T3 are the main trajectory types in mountain biking, which accounted for 73.9%; T1 and T2 are the main trajectory types in high way biking, which accounted for 59.6%; T1 and T4 are the main trajectory types in sightseeing and self-driving tour, which accounted for 65.7% and 71.4%. All these data illustrated that backpackers don’t like to travel the same route twice in mountaineering and hiking, and it’s difficult to take the closed trajectory pattern limited by individual physical strength and terrain, so much of them are Echoing Trajectories
pattern. Backpackers like the Echoing Trajectory pattern in sightseeing and self-driving tour because of the relative larger freedom.

Table 1. The distribution of the four movement patterns in each backpacker taxonomy.

<table>
<thead>
<tr>
<th>Movement pattern</th>
<th>total</th>
<th>mountaineering %</th>
<th>hiking %</th>
<th>mountain biking %</th>
<th>high way biking %</th>
<th>sightseeing %</th>
<th>self-driving tour %</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>234</td>
<td>63.1</td>
<td>61.2</td>
<td>38.5</td>
<td>31.0</td>
<td>37.1</td>
<td>46.4</td>
</tr>
<tr>
<td>T2</td>
<td>50</td>
<td>8.8</td>
<td>5.8</td>
<td>9.2</td>
<td>28.6</td>
<td>17.1</td>
<td>14.3</td>
</tr>
<tr>
<td>T3</td>
<td>94</td>
<td>20.6</td>
<td>16.5</td>
<td>35.4</td>
<td>19.0</td>
<td>17.1</td>
<td>14.3</td>
</tr>
<tr>
<td>T4</td>
<td>67</td>
<td>7.5</td>
<td>16.5</td>
<td>16.9</td>
<td>21.4</td>
<td>28.6</td>
<td>25.0</td>
</tr>
</tbody>
</table>

Calculation of Trajectory Attraction Index

Calculate the Experience Value of Tourist. In this study there are 445 trajectories, which were obtained by 162 self-help tourists (user IDs) from China during August 1, 2010 to July 31, 2011, the number of per capita trajectories is 2.74. This paper only studies the tourist whose trajectory’s number is greater than or equal to 3, after counting find that 52 tourists meet the requirements. The following four independent variables will be used to calculate the experience value of tourists, e.g. the number of trajectories of tourist a \((NTr (a), 1≤a≤52)\), the number of provinces and cities of tourist a \((NPc (a), 1≤a≤52)\), the total travel time of tourist a \((Tt (a), 1≤a≤52)\), the number of travel types of tourist a \((NTy (a), 1≤a≤52)\). Among them the NTr (a), Tt (a) and NTy (a) are calculated directly from the data attribute table, yet NPc (a) is calculated by that the different provinces incremented by 1 and the different cities in the same province incremented by 0.5. The experience value of tourist a \(( Ev(a), 1≤a≤52)\) is calculated by the following formula:

\[
Ev(a) = \frac{w_1 \ln(NTr(a)) + w_2 \ln(Tt(a)) + w_3 \ln(NTy(a)) + w_4 \ln(NPc(a))}{n}
\]  

(1)

In this formula \(w_1, w_2, w_3\) and \(w_4\) are the weights of each variable, in this article they are taken as 0.5, 0.2, 0.2 and 0.1 according to their degree of influence on the dependent variable Ev(a), and \(w_1 + w_2 + w_3 + w_4 = 1\); \(n\) is the number of independent variables, which equals to 4.

In this study tourists are divided into 4 ranks, when Ev (a) is between \([0.21, 0.40]\) belonging to rank 1, between \([0.41, 0.60]\) belonging to rank 2, between \([0.61, 0.80]\) belonging to rank 3, and between \([0.81, 1.00]\) belonging to rank 4. The distribution of tourist experience value calculated using formula (1) is shown in Table 7. There are only 6% tourists belonging to the highest rank, and the most tourists belong to rank 2 accounted for 63%. These data indicates that there is still much more room for the development of self-help tourism.

Table 2. The distribution of the experience values of the tourists.

<table>
<thead>
<tr>
<th>Tourist rank</th>
<th>Experience value interval</th>
<th>The number of tourists</th>
<th>Percentage %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.21-0.40</td>
<td>6</td>
<td>0.12</td>
</tr>
<tr>
<td>2</td>
<td>0.41-0.60</td>
<td>33</td>
<td>0.63</td>
</tr>
<tr>
<td>3</td>
<td>0.61-0.80</td>
<td>10</td>
<td>0.19</td>
</tr>
<tr>
<td>4</td>
<td>0.81-1.00</td>
<td>3</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Calculate the Trajectory Attraction Index (TAI). As more and more people upload their travel trajectories to the site, there are many similar trajectories, which caused the trajectories redundant. For example a tourist wants to mountain in Mentougou District of Beijing, but he doesn’t know the travel itinerary, then he downloads 19 mountaineering trajectories in Mentougou District of Beijing from the website. Generally travel enthusiastic wants to choose the most popular track with the shortest possible time, rather than observe one by one. If people know the attraction index of each trajectory, people can choose the most popular trajectory based on it. The attraction index is higher, the trajectory is more popular.

In this paper, the TAI is calculated by the trajectory’s Similarity and the Experience Value of the
trajectory’s tourist. Let’s encode the 19 trajectories as 1 to 19, Sim (i, j) represents the similarity between trajectory i and trajectory j, Li represents the length of trajectory i, Lj represents the length of trajectory j, L (i, j) represents the length of the overlap between trajectory i and trajectory j. The formula is as follows:

\[
Sim(i, j) = \frac{L(i, j)}{L_i + L_j - L(i, j)} \quad (1 \leq i, j \leq 19, \ 0 \leq Sim (i, j) \leq 1)
\]

(2)

SSim (i) represents the similarity of trajectory i, SSim (i) is calculated by the following formula:

\[
SSim(i) = \sum_{j=1}^{n} \text{Sim}(i, j) \quad (1 \leq i, j \leq 19, \ n=19)
\]

(3)

TAI (i) represents the interest degree value of trajectory i, which is calculated by formula (4), \( a_i \) represents the tourist of trajectory i.

\[
TAI(i) = SSim(i) * EV(a_i) \quad (1 \leq i, j \leq 19)
\]

(4)

Conclusion
In this paper the trajectories were divided into four typologies based on their shapes, e.g. (T1) Echoing Trajectory; (T2) Branch Trajectory; (T3) Closed Trajectory; and (T4) Combination Trajectory, among them Echoing Trajectory is the most common type, which accounted for 52.6% of the total. At last, this paper proposed models to calculate the experience value of tourist and to calculate the trajectory attraction index. Travel enthusiastic can choose the most popular track with the shortest time using the attraction index.

Acknowledgements
This study received support from Dongguan social science and technology development project (2019507151565).

References


