Research Article

Taxi Driver’s Operation Behavior and Passengers’ Demand Analysis Based on GPS Data

Xiaowei Hu, Shi An, and Jian Wang

School of Transportation Science and Engineering, Harbin Institute of Technology, No. 73, Huanghe River Road, Nangang District, Harbin, Heilongjiang Province 150090, China

Correspondence should be addressed to Xiaowei Hu; xiaowei_hu@hit.edu.cn

Received 8 August 2017; Revised 1 November 2017; Accepted 27 November 2017; Published 3 January 2018

Academic Editor: Anastasios Kouvelas

Copyright © 2018 Xiaowei Hu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The existing research outputs paid less attention to the relationship between land use and passenger demand, while the taxi drivers’ searching behavior for different lengths of observation period has not been explored. This paper is based on taxi GPS trajectories data from Shenzhen to explore taxi driver’s operation behavior and passengers’ demand. The taxi GPS trajectories data covers 204 hours in Shenzhen, China, which includes the taxi license number, time, longitude, latitude, speed, and whether passengers are in the taxi vehicle, to track the passenger’s pick-up and drop-off information. This paper focuses on these important topics: exploring the taxi driver operation behavior by the measurements of activity space and the connection between different activity spaces for different time duration; mainly focusing on eight traffic analysis zones (TAZs) of Shenzhen and exploring the customer’s real-time origin and destination demands on a spatial-temporal distribution on weekdays and weekends; taxi station optimization based on the passenger demand and expected customer waiting time distribution. This research can be helpful for taxi drivers to search for a new passenger and passengers to more easily find a taxi’s location.

1. Introduction

Urban land use and built environment have been considered to affect residents’ travel demand with three dimensions: design, density, and diversity [1]. Traffic engineers and urban planners have been paying more attention to explore the correlation between land use and transportation, including the land use influence on travel demand, the transport network impacts on the urban spatial development, and the integration of land use and transport system [2–6].

As an important mode, taxis play a key role in the urban passenger transportation market and provide a convenient and comfortable service for the passengers. In the taxi service study field, researchers usually adopt virtual customer origin-destination demand patterns to analyze the model [7–9], which is connected with the area land use situation, but cannot completely reflect the temporal and spatial characteristics of passenger demands. With the rapid development of Information and Communication Technologies (ICT), this provides more accurate access time and location information for the study of human mobility. Taxi vehicles equipped with Global Position System (GPS) can be served as city-wide probes, which can also provide the traffic condition, time, taxi speed, and location information, as well as whether there are passengers in the taxi. Based on the taxi GPS traces data, we can obtain the customers real-time origin and destination demand, which can assist researchers in validating the taxi service model [10–12].

As the GPS data on taxis are only recorded over a special period of time’s passenger trajectory, it is difficult to analyze the human mobility over a longer period, but taxi GPS traces can be adopted to analyze urban transport and land use situations [13]. For the past few years, researchers have achieved more progress in this field, such as the researches of Li et al. (2011) [14], Zheng et al. (2011) [15], and Yue et al. (2012) [16].

Recently researchers have combined taxi GPS data with mathematical models (Lévy flights model or Zipf distribution law) to analyze the passenger’s visiting frequency at one area [17], trip length distribution [18], and drivers’ behavior [11, 19]. However, the existing researchers paid less attention to the taxi drivers’ behavior for different lengths of observation
Table 1: An example trace of taxi GPS data, using the same vehicle as the example.

<table>
<thead>
<tr>
<th>Example</th>
<th>Data type</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>Character</td>
<td>Taxi vehicle ID</td>
</tr>
<tr>
<td>Date</td>
<td>Date</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Time</td>
<td></td>
</tr>
<tr>
<td>Longitude</td>
<td>Number</td>
<td></td>
</tr>
<tr>
<td>Latitude</td>
<td>Number</td>
<td></td>
</tr>
<tr>
<td>Status</td>
<td>Boolean</td>
<td>0: no passenger; 1: have passenger</td>
</tr>
<tr>
<td>Velocity</td>
<td>Number</td>
<td>Instantaneous velocity: kilometers per hour (km/h)</td>
</tr>
<tr>
<td>Angle</td>
<td></td>
<td>Driving direction: 0, east; 1, southeast; 2, south; 3, southwest; 4, west; 5, northwest; 6, north; 7, northeast</td>
</tr>
</tbody>
</table>

period; meanwhile, the relationship between land use and passenger demand has not been explored.

So this paper focuses on the time series distribution dynamic characteristic of passenger’s temporal variation in certain land use types and taxi driver’s searching behavior connection between different activity spaces for different lengths of observation period. This paper focused on the following topics.

(1) Exploring the taxi driver operation behavior by the measurements of activity space and the connection between different activity spaces for different time duration

(2) Mainly focusing on eight TAZs of Shenzhen and exploring the customer’s real-time origin and destination demand on spatial-temporal distribution on weekdays and weekends

(3) Taxi station optimization based on the passenger demand and expected customer waiting time distribution.

The structure of this paper is as follows. Section 2 reviews the urban land use and travel demand correlation, as well as taxi driver’s searching behavior. In Section 3, we present the taxi GPS traces data source and analysis measurements in detail. Section 4 presents the results and discussions. Finally, we conclude this paper in Section 5.

2. Literature Review

Researchers usually use virtual customer origin-destination demand patterns to analyze the taxi service model, which can refer to Arnott (1996) [7], Yang and Wong (1998) [8], Wong et al. (2001) [20], Bian et al., (2007) [21], and Luo and Shi (2009) [9]. With the development of GPS hardware and communication technology, now we can collect taxi GPS traces data over longer periods than previous typical survey [16] and it also can provide more information in detail, such as trip length, travel time, and speed by time of day, which can assist researchers to validate the taxi service model. At present, some researchers also work on this field [22, 23]; Zhang and He (2011) [22] focused more on the spatial distribution of taxi services in one day, while Hu et al. (2011) [23] mainly analyzed the one-day taxi temporal distribution of customers’ pick-up and drop-off times in Guangzhou, China.

Based on taxi GPS trace data, researchers can analyze urban transport and land use status at the macro level, which can cover the shortage of the traditional questionnaire survey [14–16, 24]. Yue et al. (2012) [16] calibrated the parameters of the spatial interaction models based on the taxi GPS traces data of the central business district in Wuhan. Liu et al. (2012) [25] explored the temporal patterns of urban-scale trip in Shanghai and found that urban land use and structure can be expressed by the taxi trip patterns.

Giraudo and Peruch (1988) [26] had divided the taxi operation into two phases, “the transport phase” and “the approach phase,” which also can be used to represent the taxi with passenger and without passenger operation, respectively. The taxi driver’s searching passenger behavior happens in “the approach phase.” When the driver has dropped off the prior passenger, then he/she drives around the area or region searching for the next passenger for a short time.

For the taxi driver’s individual characteristics (driving experience, road network familiarity, etc.) and randomness of the passenger’s arriving, the driver’s searching for the next passenger can be seen as a random variable. Luo (2009) [27] had expressed taxi driver’s searching for the next passenger as a double exponential (Gumbel) distribution.

Liu et al. (2010) [11] described the taxi driver’s operation patterns and difference between top drivers and ordinary drivers’ behavior in Shenzhen and discussed taxi drivers’ behavior based on the taxi daily GPS traces data; they analyzed the drivers’ spatial selection behavior, operation behavior, and route choice behavior. But in the research of Liu et al. (2010) [11], they did not mention drivers’ searching space behavior pattern.

This paper attempts to bridge these gaps between theoretical research and practical development, based on the taxi GPS trajectories data of Shenzhen to explore urban land use and taxi driver’s operation behavior.

3. Data Source and Taxi Operation Activity

3.1. Data Source. In this research, we use the taxi GPS traces data of Shenzhen, China, which contains 3198 taxi records over nine consecutive days, from 18 April, 2011 (Monday), to the noon 26 April, 2011 (Tuesday), with a total of 204 hours. Table 1 shows the typical format of taxi trajectory data, including taxi location (longitude, latitude), speed, direction (angle), and passenger pick-up and drop-off information (status), with associated time information. The data collection time interval is generally around 5 to 15 seconds. Delays or missing data may occur depending on the
GPS signal, and additional records are collected when taxi load status changes.

Here we use R statistics software (R 2.13.2, from https://www.r-project.org/) and ArcGIS 9.3 (https://www.esri.com/zh-cn/arcgis/products/arcgis-pro/overview) software to handle taxi GPS data processing and statistics work.

The data analysis level and scope can be divided into three aspects, including the mean center analysis of the pick-up and drop-off events each day, passengers’ spatial-temporal distribution of eight TAZs (traffic analysis zones) in the 204 continuous hours, and the taxi driver’s searching behavior exploring from different level.

3.2. Taxi Operation Activity Space Measurements. The taxi operation activity measurements mainly are based on the basic parameters distributions, such as mean center, standard deviation of the basic parameters distributions, such as mean center, standard deviation, and kernel density. Based on existing researches, we divided these measurements into two categories, the spatial distribution category and the extended second moments of activity locations measurement category.

3.2.1. The Spatial Distribution Category. The mean center (MC) is the average location of taxi operation activity service events in the space (including the pick-up and drop-off event), which can be calculated by the following [28]:

\[
(x_{mc}, y_{mc}) = \left[ \frac{\sum_{j=1}^{n} x_j}{n}, \frac{\sum_{j=1}^{n} y_j}{n} \right],
\]

where \(x_{mc}, y_{mc}\) are the coordinates of the mean center, which can determine the space location of the MC; \(x_j, y_j\) are the coordinates of taxi service event \(j\) in two dimensions; \(n\) is the total number of taxi service events.

Standard deviational ellipse (SDE) [28, 29] can determine the directional factors of the spatial distribution and find the main direction of the taxi service event in space. The calculation of SDE can be expressed by \(\theta\) (the SDE \(y\)-axis in the clockwise rotation angle); \(e_1\), the major axis of SDE; and \(e_2\), the minor axis of SDE. The detailed formulas are as follows:

\[
\theta = \arctan \left( \frac{\sum (x_j - \bar{x}_{mc})^2 - \sum (y_j - \bar{y}_{mc})^2}{2 \sum (x_j - \bar{x}_{mc})(y_j - \bar{y}_{mc})} \right) + \sqrt{\left[ \frac{\sum (x_j - \bar{x}_{mc})^2 - \sum (y_j - \bar{y}_{mc})^2}{2 \sum (x_j - \bar{x}_{mc})^2 - \sum (y_j - \bar{y}_{mc})^2} + 4 \sum (x_j - \bar{x}_{mc})(y_j - \bar{y}_{mc}) \right]^2}
\]

\[
e_1 = 2S_x = \frac{\sum [(x_j - \bar{x}_{mc}) \cos \theta - (y_j - \bar{y}_{mc}) \sin \theta]^2}{n - 2}
\]

\[
e_2 = 2S_y = \frac{\sum [(x_j - \bar{x}_{mc}) \sin \theta - (y_j - \bar{y}_{mc}) \cos \theta]^2}{n - 2}
\]

\[A = \pi S_x S_y,
\]

where \(A\) is area of SDE, \(S_x\) is the semimajor axis of SDE, \(S_y\) is the semiminor axis of SDE, and \(\bar{x}_{mc}, \bar{y}_{mc}\) and \(x_j, y_j\) are consistent with formula (1).

3.2.2. The Extended Second Moments of Activity Locations Measurement Category. Each taxi driver’s daily activity space area mean center may have the relationship with the centroid of the whole taxi drivers’ activity space [19], similar to the Susilo & Kitamura (2005) [30] analysis of the worker’s daily activity locations relationship. We can analyze taxi driver’s day-to-day variation on activity space and statistically analyze the second moments of activity locations.

Figure 1 shows an illustration of the drop-off (pick-up) locations mean center of each taxi driver and all taxi drivers, which can analyze each taxi driver’s day-to-day variation in the pick-up and drop-off activity space. Based on our statistics, the distance of the two MCs is mainly concentrated between 200 m and 400 m, which may reflect the taxi driver’s searching behavior around a certain MC.

4. Analysis Results

4.1. Taxi Driver’s Operation Behavior Analysis. In this section, we first explored the taxi driver operation behavior by the measurements of activity space and the connection between different activity spaces for different time duration. Here the MC and the locations mean center of each taxi driver and all taxi drivers have been used in the analysis. Figure 2 presents the spatial distribution of all taxi drivers’ drop-off activity space mean center, which is analyzed by each day.

From Figure 2, we can find that taxi driver’s drop-off activity space mean centers are mainly distributed around 22.562 to 22.576 (latitude) and 114.035–114.070 (longitude). And comparing the weekdays (from Monday to Friday) and weekends, there are two area distributions, which is from 1 a.m. to 6 p.m. and from 7 p.m. to 12 p.m., respectively. The red circle in Figure 2 shows the distribution from 7 p.m. to 12 p.m.

Figure 3 presents the spatial distribution of all taxi drivers’ pick-up activity space mean center, which is analyzed by each day.
Table 2: The detailed information of the eight TAZs.

<table>
<thead>
<tr>
<th>Number</th>
<th>TAZ name</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Area/km²</th>
<th>Land use and urban function</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAZ1</td>
<td>Futian CBD</td>
<td>114.055–114.065</td>
<td>22.530–22.540</td>
<td>2.0</td>
<td>CBD, Commercial center</td>
</tr>
<tr>
<td>TAZ2</td>
<td>Conventional financial center</td>
<td>114.105–114.123</td>
<td>22.534–22.548</td>
<td>0.9</td>
<td>Conventional financial center</td>
</tr>
<tr>
<td>TAZ3</td>
<td>Luohu Port</td>
<td>114.109–114.117</td>
<td>22.530–22.538</td>
<td>0.6</td>
<td>Port connecting with Hong Kong</td>
</tr>
<tr>
<td>TAZ4</td>
<td>Huanggang Port</td>
<td>114.068–114.080</td>
<td>22.518–22.529</td>
<td>1.0</td>
<td>Port connecting with Hong Kong</td>
</tr>
<tr>
<td>TAZ5</td>
<td>Huanggang Village</td>
<td>114.058–114.069</td>
<td>22.517–22.528</td>
<td>1.4</td>
<td>Communities are densely distributed</td>
</tr>
<tr>
<td>TAZ6</td>
<td>Huangbei Street</td>
<td>114.130–114.143</td>
<td>22.545–22.557</td>
<td>1.5</td>
<td>Residential area</td>
</tr>
<tr>
<td>TAZ7</td>
<td>Nanhu Street</td>
<td>114.113–114.131</td>
<td>22.539–22.545</td>
<td>1.3</td>
<td>Residential area</td>
</tr>
<tr>
<td>TAZ8</td>
<td>Dongmen Street</td>
<td>114.120–114.130</td>
<td>22.544–22.554</td>
<td>1.0</td>
<td>Commercial leisure and entertainment area</td>
</tr>
</tbody>
</table>

From Figure 3, we can also find that taxi driver's pick-up activity space mean centers are mainly distributed around 22.560 to 22.574 (latitude) and 114.035–114.070 (longitude). And comparing the weekdays (from Monday to Friday) and weekends, there are two area distributions, which is from 1 a.m. to 6 p.m. and from 7 p.m. to 12 p.m., respectively. The red circle in Figure 3 shows the distribution from 7 p.m. to 12 p.m.

From Figures 2 and 3, the scope of the taxi drivers' pick-up activity space mean center is more concentrated than the drop-off activity space mean center. Meanwhile, on the weekdays, the taxi drivers' pick-up activity space mean center and drop-off activity space mean center are more concentrated than on weekends, which can reflect the passengers' daily life change, and they may have a more comfortable weekday.

4.2. Eight TAZs' Passenger Demands' Spatial-Temporal Distribution Analysis. In this section, we present the analysis results between passenger's origin and destination demand on spatial-temporal distribution from 18 April, 2011 (Monday), to the noon 26 April, 2011 (Tuesday). And we mainly focus on eight TAZs (see in Table 2) of Shenzhen; Figure 4 presents the eight TAZs' passenger pick-up (in blue line) and drop-off (in red line) statistical chart.

Based on Figure 4, there is a relative equilibrium of the total passenger pick-up and drop-off situation in the eight TAZs, but, for different land use types, there exist different peak hours of the pick-up and drop-off demand.

For the commercial and CBD areas (TAZ1, TAZ2), the passenger pick-up and drop-off service frequency is higher than in residential areas (TAZ6 and TAZ7). Meanwhile, the peak hour for the eight different land use areas is quite different from each other (see in Table 3).

By reason of different land use types, the peak hours of the eight TAZs are different from each other, while the passenger's pick-up and drop-off events are not synchronized. In Shenzhen, the peak hour of taxi passenger's is almost at the midnight, such as in TAZ2, TAZ7, and TAZ8, which is similar to the research of Hu et al. (2014).

The trend of how pick-up and drop-off changes with time is almost the same from Monday to Friday for each TAZ. At weekends, the peak hour is a bit different with in weekdays, especially in TAZ1, TAZ5, and TAZ6.

Then the taxi vehicle's service frequency for each TAZ was analyzed, which is shown in Table 4. From this table it can be seen that, in each TAZ, the taxi vehicle's supply is different to each other and each taxi vehicle's service time in TAZ is quite different. In Table 4, we can find that some taxi drivers are cruising around some areas, especially for the taxi drivers who provide more than 130 pick-up service in 204 hours (see in Table 4).

Based on this phenomenon, we divide the taxi drivers into different categories, some drivers only provide random service in the whole city, but some drivers can provide a relatively fixed service just around a specific area, such as the CBD, and residential area. Then the distributions of taxi drivers' pick-up service time in the eight TAZs were analyzed (as shown in Figure 5).

In TAZ1, TAZ5, and TAZ7, more than 60% of taxi driver's pick-up service times are less than 5 times, while, in TAZ3, TAZ4, TAZ6, and TAZ8, more than 85% of taxi driver's pick-up service times are less than 20 times, so 20 times can be taken as the boundary for the two different categories of taxi driver's service pattern. From Figure 5, we can also find that, in TAZ2, the average service time of each taxi driver is 46.47 times, and the 85% of taxi driver's pick-up service times is 70 times, so in TAZ2 the 70 times can serve as the boundary for the two different categories of taxi driver's service pattern.

4.3. Taxi Station Optimization. From the analysis, we can find that the biggest passenger demand is in TAZ2, which is along the Shenzhen south road and international trade
Figure 2: Spatial distribution of taxi driver’s drop-off activity space mean center (from Monday to next Tuesday).

Table 3: Peak hour statistics of different TAZs.

<table>
<thead>
<tr>
<th>Number</th>
<th>Activity</th>
<th>April 18 Monday</th>
<th>April 19 Tuesday</th>
<th>April 20 Wednesday</th>
<th>April 21 Thursday</th>
<th>April 22 Friday</th>
<th>April 23 Saturday</th>
<th>April 24 Sunday</th>
<th>April 25 Monday</th>
<th>April 26 Tuesday</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAZ1</td>
<td>Pick</td>
<td>12</td>
<td>14</td>
<td>9</td>
<td>15</td>
<td>10</td>
<td>23</td>
<td>22</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Drop</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>12</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>TAZ2</td>
<td>Pick</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>21</td>
<td>23</td>
<td>22</td>
<td>21</td>
<td>23</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Drop</td>
<td>22</td>
<td>22</td>
<td>16</td>
<td>21</td>
<td>23</td>
<td>22</td>
<td>21</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>TAZ3</td>
<td>Pick</td>
<td>23</td>
<td>23</td>
<td>16</td>
<td>21</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Drop</td>
<td>7</td>
<td>7</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>9</td>
<td>21</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>TAZ4</td>
<td>Pick</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>24</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Drop</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>23</td>
<td>1</td>
<td>23</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>TAZ5</td>
<td>Pick</td>
<td>10</td>
<td>20</td>
<td>1</td>
<td>22</td>
<td>1</td>
<td>20</td>
<td>23</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Drop</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>24</td>
<td>23</td>
<td>23</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>TAZ6</td>
<td>Pick</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>22</td>
<td>10</td>
<td>20</td>
<td>23</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Drop</td>
<td>22</td>
<td>23</td>
<td>23</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>21</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>TAZ7</td>
<td>Pick</td>
<td>22</td>
<td>22</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Drop</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>23</td>
<td>22</td>
<td>23</td>
<td>22</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>TAZ8</td>
<td>Pick</td>
<td>22</td>
<td>22</td>
<td>23</td>
<td>23</td>
<td>22</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Drop</td>
<td>22</td>
<td>22</td>
<td>23</td>
<td>23</td>
<td>22</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>1</td>
</tr>
</tbody>
</table>

* April 26, 2011, only has 12 hours of data recorded.
Figure 3: Spatial distribution of taxi driver's pick-up activity space mean center (from Monday to next Tuesday).

Table 4: Summary of each taxi vehicle's service frequency of each TAZ.

<table>
<thead>
<tr>
<th>TAZ</th>
<th>TAZ2</th>
<th>TAZ3</th>
<th>TAZ4</th>
<th>TAZ5</th>
<th>TAZ6</th>
<th>TAZ7</th>
<th>TAZ8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1st quarter</td>
<td>4</td>
<td>28</td>
<td>8</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Median</td>
<td>7.392</td>
<td>46.47</td>
<td>14.92</td>
<td>4.163</td>
<td>15.35</td>
<td>5.737</td>
<td>18.02</td>
</tr>
<tr>
<td>Mean (average value)</td>
<td>76</td>
<td>84</td>
<td>20</td>
<td>130</td>
<td>85</td>
<td>130</td>
<td>109</td>
</tr>
<tr>
<td>3rd quarter</td>
<td>9</td>
<td>62</td>
<td>20</td>
<td>5</td>
<td>20</td>
<td>7</td>
<td>23</td>
</tr>
<tr>
<td>Maximum</td>
<td>12,661</td>
<td>44,414</td>
<td>30,811</td>
<td>38,044</td>
<td>12,661</td>
<td>44,414</td>
<td>30,811</td>
</tr>
<tr>
<td>Number of taxi vehicles</td>
<td>2,468</td>
<td>2,498</td>
<td>2,467</td>
<td>1,985</td>
<td>2,478</td>
<td>2,207</td>
<td>2,465</td>
</tr>
<tr>
<td>Total pick-up service times</td>
<td>18,244</td>
<td>116,072</td>
<td>36,812</td>
<td>8,264</td>
<td>38,044</td>
<td>12,661</td>
<td>44,414</td>
</tr>
</tbody>
</table>

From Figure 4, we can find the two peak hours of passengers’ pick-up service in TAZ2 is 2 p.m. to 3 p.m. and 9 p.m. to 10 p.m., which is connected with the land use and geographic location. So the taxi station optimization is based on the passenger demand and expected customer waiting time distribution, while we do not consider the setting form of the taxi station in this paper.

For the study field of taxi station’s service area, Daganzo (1978) [24] proposed the flexible transit design model (FTDM), and in 2012 he had optimized it into a transit optimization approach [31]. Based on existing research of Nourbakhsh and Ouyang (2012) [32] and Sathaye (2014) [33], here a taxi station optimization model is presented to determine the service radius $R$.

According to the research of Nourbakhsh and Ouyang (2012) [32], each passenger’s expected walk distance is shown in the following formula in km:

$$E = \frac{2D}{3}$$

where $D$ is the length of the side of one square; then each passenger’s expected walk time in hours is
The taxi pick-up and drop-off service frequency of TAZ1

The taxi pick-up and drop-off service frequency of TAZ2

The taxi pick-up and drop-off service frequency of TAZ3

The taxi pick-up and drop-off service frequency of TAZ4

The taxi pick-up and drop-off service frequency of TAZ5

The taxi pick-up and drop-off service frequency of TAZ6

The taxi pick-up and drop-off service frequency of TAZ7

The taxi pick-up and drop-off service frequency of TAZ8

Figure 4: Eight TAZs passenger pick-up (in blue line) and drop-off (in red line) service statistics.

\[ T = \frac{E}{v} = \frac{2D}{3v}, \quad (4) \]

where \( v \) is the average operation speed (km/h). Therefore, a taxi station's service radius \( R \) can be expressed by the following formula:

\[ R = \min \left( \frac{D}{2}, \frac{0.95D}{\sqrt{\pi Y}} \right), \quad (5) \]

where \( R \) is the service radius of the taxi station (km) and \( Y \) is the number of taxi stations.
Figure 5: Continued.
Figure 5: The taxi vehicles’ pick-up service frequency statistics of each TAZ in Shenzhen.

Table 5: Calculation table of taxi station’s service radius (unit: km).

<table>
<thead>
<tr>
<th>D (kilometer)</th>
<th>Y = 1</th>
<th>Y = 2</th>
<th>Y = 3</th>
<th>Y = 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>0.1</td>
<td>0.075799</td>
<td>0.061890</td>
<td>0.053598</td>
</tr>
<tr>
<td>0.4</td>
<td>0.2</td>
<td>0.151598</td>
<td>0.123779</td>
<td>0.107196</td>
</tr>
<tr>
<td>0.6</td>
<td>0.3</td>
<td>0.227397</td>
<td>0.185669</td>
<td>0.160794</td>
</tr>
<tr>
<td>0.8</td>
<td>0.4</td>
<td>0.293196</td>
<td>0.247559</td>
<td>0.214392</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
<td>0.378995</td>
<td>0.309448</td>
<td>0.267990</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.757990</td>
<td>0.618897</td>
<td>0.535980</td>
</tr>
</tbody>
</table>

For the given $D$ and $Y$, we can calculate the taxi station’s service radius; the results are shown in Table 5. Referring to the study by Zhang et al. (2015) [34], which is based on taxi GPS data and analysis, they recommend the taxi station’s service distance to be 300 m; this result can be matched with some results in Table 5 (the bold result).

5. Conclusions and Recommendations

This paper is based on taxi vehicle’s GPS data to analyze the time series distribution dynamic characteristics of passengers’ temporal variation in certain land use types and taxi driver’s searching behavior in connection with different activity spaces for different lengths of observation period. And adopting GPS data had identified the passengers’ demand hot area and proposed a taxi station optimization model, which can be served as reference to taxi station location decision.

By researching and proposing appropriate measures and statistics to properly measure and analyze activity spaces while recognizing their geographical dependence, this study can make some contributions to methodologies in measuring and analyzing behavioral dynamics. By understanding the dynamics in the activity spaces of taxi drivers over time, this study directly contributes to the field of travel behavior dynamics. The value that the study will potentially render in policy guidance also cannot be underestimated, understanding these dynamics at the driver level. The analysis will also provide a new method to optimize urban transport management, to investigate land-using planning, and to evaluate road network traffic conditions.

However, this paper from taxi vehicle’s GPS data can reflect driver’s behavior more accurately and, regarding the passenger level, it needs to combine the passenger’s characters survey and the booking data from an Internet-booking application with the taxi vehicle’s GPS data [35, 36] and to analyze passenger’s trip and the relationship with land use. In addition, taxi service and the public passenger transport system are strongly complementarity in big cities. In the future, we will take into account the main public transit facilities on taxi demand analysis.

On the other hand, based on the expected customer waiting time distribution, the approximate distribution of the customer waiting time formula can be validated and modified in different kinds of land use type, which will be more useful in describing the principal characteristics in the taxi market and affect customers’ decision and taxi driver’s cruising time [37, 38].
Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (71603063), the Natural Science Foundation of Heilongjiang (no. E2016032), China Postdoctoral Science Foundation funded Project (no. 2013M540299), and the Fundamental Research Funds for the Central Universities (no. HIT.NSRIF:2015075). The authors would like to express their sincere gratitude to Professor Xuesong Zhou (Arizona State University) for his valuable suggestions.

References


