

Research Article

Discovering Public Transit Riders' Travel Pattern from GPS Data: A Case Study in Harbin

Shi An, Lei Wang, Haiqiang Yang, and Jian Wang

School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin, China

Correspondence should be addressed to Lei Wang; yutian@hit.edu.cn

Received 2 March 2017; Accepted 28 May 2017; Published 2 July 2017

Academic Editor: Manel del Valle

Copyright © 2017 Shi An 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.

This paper proposes a public transit riders' travel pattern measuring method based on divided cells and public transit vehicle's GPS data. The method consists of two parts: detecting urban origin and destination areas and measuring the public transit riders' travel pattern. Moreover, a series of indicators are proposed to reflect the public transit riders' travel pattern. A case study is carried out to evaluate the methods, which use the GPS data collected from taxis and buses in Harbin, China. The study is expected to provide a better understanding of public transit riders' travel patterns.

1. Introduction

The dramatic increase of urban vehicles leads to many serious problems including traffic congestions, road accidents, and air pollutions, which become general conundrums in many metropolitans all over the world, or even in smaller cities. Public transit is considered to be one of the most effective solutions for these general conundrums. Public transit includes various services that provide mobility to the general public, including buses, trains, ferries, shared taxi, and their variations [1, 2]. Public transit has many obvious advantages, such as less expenses, more effective mobility, and saving travel time, which makes more and more urban residents turn to public transit service. The city travelers and commuters proportion of Beijing who takes public transit system (including bus, taxi, and rail transit) has continuously increased in last decade. And this number came up to 54.2% in 2015, which represents more than 15.5 million trips per day [3]. This trend also occurred in other metropolitans and smaller cities all over the world.

Public transit riders usually exhibit a fixed travel pattern. That is to say, at a macrolevel, usually a fixed number of Origin-Destination pairs locate in same place of urban area, and the number of trips between these OD pairs stays steady every day [4–6]. The usage ratios of each kind of public transit stay steady every day. And the trips at morning or evening rush hours occupy a large proportion of the total trips every

day [3]. On the other hand, at a microlevel, a single traveler moves from resident area to work place in the morning and moves back in the evening for each day. If the travel pattern of public transit riders could be identified, the urban public transit manager can benefit from it. For example, analyzing those riders who prefer to choose public transit vehicles rather than private cars helps transit authorities to improve the strategies and even make new policy to attract new riders [7]. With a better understanding of the transfer behavior of public riders, transit agencies can adjust the bus route to make transfer easier, which can enhance the riders' satisfaction [8]. By calculating the shortest path lengths between all station pairs, the original-destination matrix, and trip lengths, transit agencies can develop fare change plans to manage demand or raise revenue [9, 10].

The Original-Destination (OD) matrix is a typical representation of residents' travel pattern, which reflects travel demand, trip generation, travel distribution, and so on. The traditional models of establishing OD matrix usually rely on travel behavior survey. In practice, household travel surveys are conducted in many countries [11]. However, the survey data contain many limitations and errors. For example, some metropolitans in Japan (Tokyo, Kyoto, Osaka, etc.) survey residents' travel behavior every 10 years. Since the cities are growing rapidly, the survey data will definitely be out of date [11]. Furthermore, the sampling rate is usually very low, which

brings sampling errors [12]. Meanwhile, many human factors may affect the accuracy of the OD matrix, such as, willfully filtering some trips, forgetfulness, and other related factors [13].

Compared with traditional survey data, GPS data and smart card data exhibit wider coverage, lower cost, and higher accuracy. With the rapid development of data-based technology, various intelligent transportation systems are widely applied in public transit system. These systems could collect residents' mobility data every day, including longitude, latitude, boarding time, and dropping off time [14]. In the last decade, various researches based on these data have been carried out, for example, mining urban recurrent congestion evolution patterns from GPS-equipped vehicle mobility data [14], comparing accessibility in urban slums using smart card and bus GPS data [15, 16], discovering functional zones using bus smart card data [17], and partitioning bus operating hours into time of day intervals based on bus GPS data [18], which makes the data based transportation research to be a hot spot of transportation field [19].

GPS data and smart card data are usually collected from different subsystems of one whole intelligent transportation system, or even from different systems. This is because bus, taxi, and rail transit usually belong to different public transit companies [3]. Therefore, most previous data based researches into transit traveler behaviors utilize smart data [20–22] or GPS data [23, 24], respectively. Accordingly, bus and rail transit riders' travel pattern can be generated [20–22] or travel behavior of taxi riders [23, 24], respectively. To the best of my knowledge, public transit riders' travel pattern researches are barely utilizing both smart card and GPS data. However, smart card data only provide riders' boarding and dropping off information and lack locating information [21, 22]. As a result, only approximate location can be acquired, which causes inaccuracy of origin and destination inference. Moreover, on the microlevel, GPS-equipped public transit vehicle riders' trips are far less than other public transit vehicles [3]. And the insufficient sampling number will definitely lead to inaccuracy of trip distribution. In this light, the smart data and GPS data should be integrated to discover public transit riders' travel pattern.

The aim of this paper is to propose an effective method to explore the public transit riders' travel pattern in an urban area. There are two subgoals identified: detecting urban origin and destination areas at a cell level and measuring public transit riders' travel pattern.

This paper is organized as follows. Section 2 discusses the definition of cells and locating points that will be applied to this research. Section 3 describes the proposed urban public transit riders' travel pattern measuring method. Section 4 applies the proposed public transit riders' travel pattern measuring methodology using taxi and bus GPS data and the urban road network of Harbin. Section 5 provides conclusions and recommendations for future research.

2. Definition of Cells and Locating Points

In this part, we are going to define some parameters of cells and locating points. Firstly, the urban area is divided into

$n \times m$ small cells with same size. $C(x, y)$ is one of these cells, where $x = 1, 2, \dots, n$ and $y = 1, 2, \dots, m$.

According to the taxi GPS dataset, there are four types of occupation status. In this paper, we are going to study the public transit riders' travel pattern, so we defined two types of locating points according to the boarding and dropping off status, which are described as follows:

Type 1. $P_{\text{taxi,boarding}}$ represents taxi vehicles' locating points whose occupation status value is 768 (i.e. the occupation status is boarding).

Type 2. $P_{\text{taxi,dropping off}}$ represents taxi vehicles' locating points whose occupation status value is 16640 (i.e. the occupation status is dropping off).

The locating points whose occupation status value is 256 (represents the taxi vehicles being vacant) are of no use to the public transit riders' travel pattern, so they are not utilized in our research.

$P_{\text{taxi}}(\text{ID}, \text{ts})$ represents a specific locating point of a taxi vehicle, where ID and ts are the Taxi ID and Timestamp from the taxi GPS datasets.

Passenger only needs to touch smart card once while boarding bus in many cities of China, such as Guangzhou, Xi'an, and Harbin, whereas, in Beijing, passenger needs to touch smart card once again while dropping off. It is easy to know both boarding locating points and dropping off locating points in Beijing. In Harbin, the dropping off locating points can be inferred based on boarding points and time period. In order to simplify the method, in this paper, we use the most periodic commuters' data from the bus datasets. That is to say, if a specific bus rider gets to work every morning and comes back home every evening, only this type of rider's locating points will be included in our research. This type of bus rider generates two trips each day. In this light, we defined two types of bus locating points as follows.

Type 1. $P_{\text{bus,boarding}}$ represents boarding passengers' locating points.

Type 2. $P_{\text{bus,dropping off}}$ represents dropping off passengers' locating points.

$P_{\text{bus}}(\text{ID}, \text{ts})$ represents a specific locating point of a bus vehicle, where ID and ts are the Card ID and Timestamp from the bus GPS datasets.

How to identify the boarding locating points and dropping off locating points from smart card data is described as follows.

Step 1. Extract the locating points with same card ID (i.e., $P_{\text{bus}}(\text{ID}_i, \text{ts}_1), P_{\text{bus}}(\text{ID}_i, \text{ts}_2), \dots, P_{\text{bus}}(\text{ID}_i, \text{ts}_n)$) from the whole smart dataset, where $P_{\text{bus},i}(\text{ID}_i, \text{ts})$ is the locating point whose card ID is ID_i and n is the number of this kind of locating points.

Step 2. Extract the locating points whose $n = 2$ and $\text{ts}_1 \in [07:00, 10:00]$ and $\text{ts}_2 \in [15:00, 18:00]$.

Step 3. Let $P_{\text{bus,boarding}} = P_{\text{bus}}(\text{ID}, \text{ts}_1)$ and $P_{\text{bus,dropping off}} = P_{\text{bus}}(\text{ID}, \text{ts}_2)$.

3. Methodology

The methodology for discovering public transit riders' travel pattern is described in this section. Two stepwise methods are proposed to achieve the main goal, including detecting Origin and Destination areas, measuring the commuter pattern between each OD pair.

3.1. Detecting Origin and Destination Areas. In this part, we are going to detect Origin and Destination areas from public transit riders' GPS data (i.e., taxi and bus GPS data). Because of the reasonable urban planning and constructions in recent decades, the urban area is usually divided into many different function zones. Therefore, all the origin points and destination points of passengers' trips will be clustered into several origin areas and destination areas. In this light, we use cluster algorithm to detect origin areas and destination areas in urban area. Moreover, we do not know how many clusters are in advance, and this type of cluster is usually not spherically-shaped. Therefore, we apply a customized Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm to solve this problem.

For a specific cell $C(x, y)$, we define a parameter $N(x, y)$ to present the locating point (i.e. P_{taxi} and P_{bus}) number in this cell during a specific period P . Taking a 3×3 cells, for example (as shown in Figure 1(a)), in one period all cells' parameters $N(x, y)$ can be easily calculated. For two specific cells $C(x_1, y_1)$ and $C(x_2, y_2)$, the distance between them $\text{dist}_{[C(x_1, y_1), C(x_2, y_2)]}$ is calculated as follows:

$$\text{dist}_{[C(x_1, y_1), C(x_2, y_2)]} = l \cdot \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}, \quad (1)$$

where l is the length of the cell. As shown in Figure 1(b), the distance between $C(i - 1, j)$ and $C(i - 1, j + 1)$ is l , and the distance between $C(i - 1, j - 1)$ and $C(i, j)$ is $l \cdot \sqrt{2}$.

Some relative parameters are defined as follows:

Object o represents the cell.

Core object co : the cell satisfies $N \geq \delta_N$, where δ_N is the threshold of parameter N .

ε -Neighborhood of a core object is the space within a radius ε centered at co .

Figure 2 illustrates the flow chart of the customized DBSCAN algorithm in this research. The Original DBSCAN algorithm is based on density. Comparatively speaking, the density reachable points in our customized algorithm are defined by $N \geq \delta_N$. And we define the *minPts* value as 1 in our paper.

3.2. Measuring Public Transit Riders' Travel Pattern. The public transit riders' travel pattern can be reflected by some indicators, like trip number between each OD pair, the proportion of different transit, path between each OD pair, travel time of the path, and so on. In order to show these indicators, we defined some preliminary terms as follows.

$\text{Tr}_{\text{mode}}(\text{CL}_i \rightarrow \text{CL}_j)$ is the trip number of a specific transit mode (i.e. taxi or bus) from i th cluster to j th cluster.

Based on clusters of origin and destination area, given a specific period P , the public transit OD matrix can be calculated.

$p_{\text{mode}}(\text{CL}_i \rightarrow \text{CL}_j)$ is the proportion of a specific transit mode (i.e. taxi or bus) from i th cluster to j th cluster, which is calculated as follows:

$$p_{\text{mode}}(\text{CL}_i \rightarrow \text{CL}_j) = \frac{\text{Tr}_{\text{mode}}(\text{CL}_i \rightarrow \text{CL}_j)}{\text{Tr}_{\text{MODE}}(\text{CL}_i \rightarrow \text{CL}_j)} \quad (2)$$

$$\times 100\%,$$

where $\text{Tr}_{\text{MODE}}(\text{CL}_i \rightarrow \text{CL}_j)$ is the trip number of all transit modes from i th cluster to j th cluster.

$\text{Tt}_{\text{mode}}(\text{CL}_i \rightarrow \text{CL}_j)$ is the travel time by the specific transit mode (i.e. taxi or bus) from i th cluster to j th cluster.

4. Case Study

We apply the proposed methods to Harbin city (China). First of all, the datasets used in this case study are described, and the stepwise methods are implemented one by one.

4.1. Datasets. The smart card data and GPS data are collected from all operating buses and taxis in Harbin. Harbin is the capital of Heilongjiang province in northeast of China. With 4.74 million populations and 7,086 km² areas in urban area [25], Harbin is a typical developing city in China. The public transit system of Harbin consists of taxi, bus, metro, and other rail transit. In addition, all taxis are equipped with GPS device and all buses are equipped with smart card system, which make collecting GPS data and smart card data possible.

Approximately 16,000 operating taxis equipped with GPS device are running around Harbin's urban area day and night. The location information is uploaded to the management system every 30 s during the day and 2 min at night. The data are accumulated to 2G in size and around 25 million rows each day. The taxi GPS data collected from 3rd Aug. to 7th Aug. 2015 is used, consisting of taxi ID, timestamp, latitude, longitude, and status, as shown in Table 1. There are 4 kinds of "status" in the table: 17152, 16640, 256, and 768, which represent occupation, dropping off, vacant, and boarding, respectively.

There are nearly 1500 buses traveling around urban area of Harbin, and these buses belong to 100 routes. The sampling frequency is 30 s. The bus IC records collected from 3rd Aug. to 7th Aug. in 2015 are used, consisting of Route ID, Bus ID, Card ID, Timestamp, Latitude, and Longitude, as shown in Table 2.

In our research, another dataset is digital map of Harbin, which consists of most urban areas of Harbin. The research area is nearly 100 square kilometers and just covers the range of 2nd Ring Road of Harbin, as shown in Figure 3. About 80 percent of GPS points locate in this research area. We divided this area into 250 square cells with same size. And each of the cell is 200 × 200 meters square, which is shown in Figure 3.

4.2. Detecting Origin and Destination Areas. In this paper, we only measure the public transit riders' travel pattern during

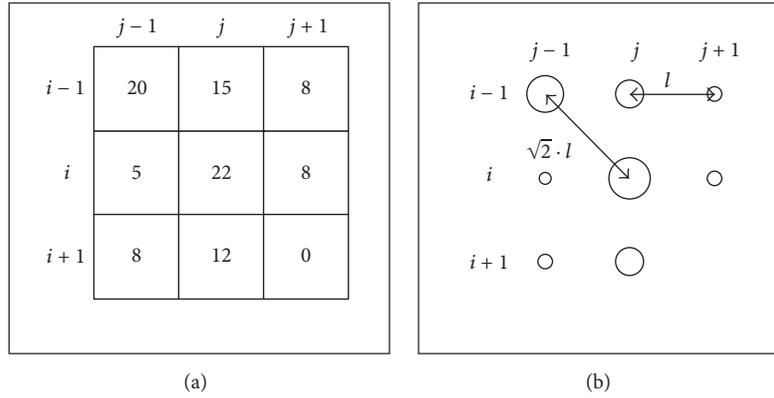
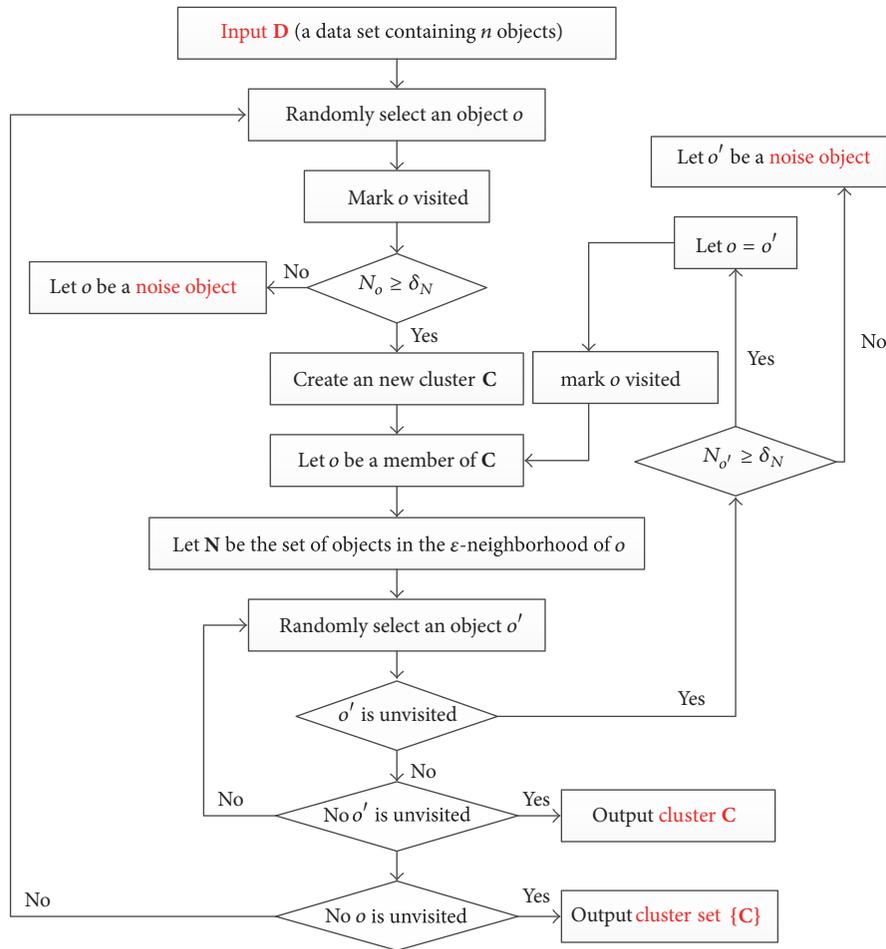
FIGURE 1: Example of cells with N value.

FIGURE 2: The flow chart of the customized DBSCAN algorithm.

working days. According to Section 3.1, we can calculate the $N(x, y)$ values of all cells. Taking one week's dropping off data (consisting of $P_{\text{taxi.dropping off}}$ and $P_{\text{bus.dropping off}}$) for example, that is, from 3rd Aug. to 7th Aug. in 2015, the relative data are illustrated in Table 3. And the sampling time is from 07:00 am to 19:00 pm.

And then we apply the customized DBSCAN algorithm to measure the origin and the destination clusters. We set ϵ value as 282.8 meters (i.e. $\sqrt{2} \cdot l$, where l is 200 meters). We apply different δ_N values to our experiment, in order to find out the optimal value. Table 4 illustrates the clustering results by computing with different δ_N values. According to the results,

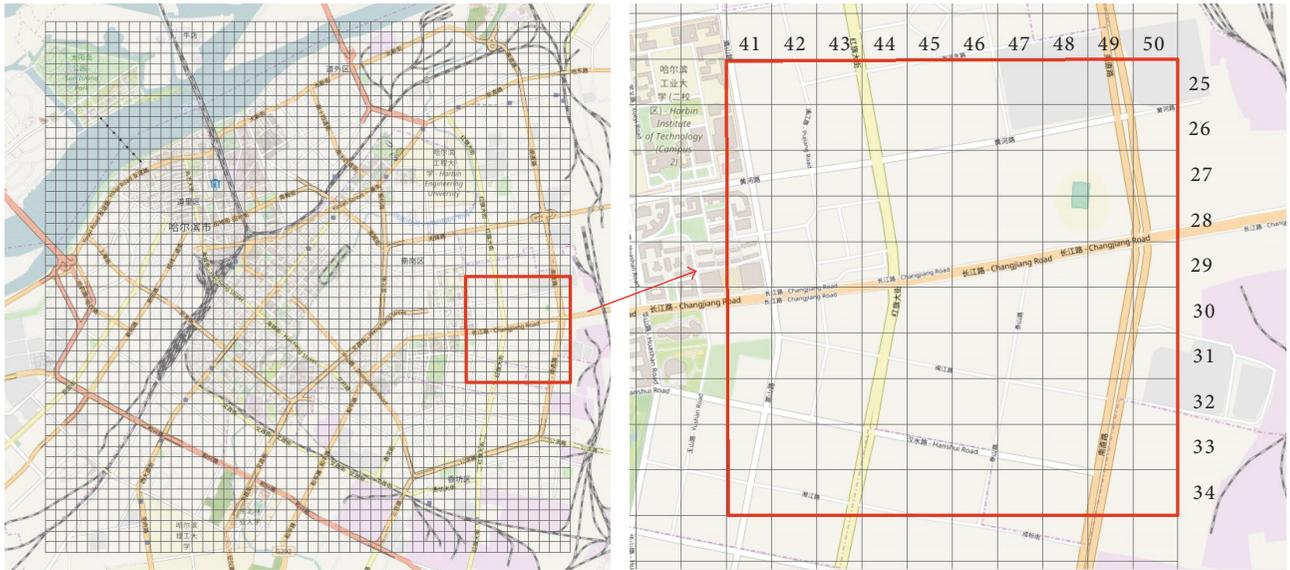
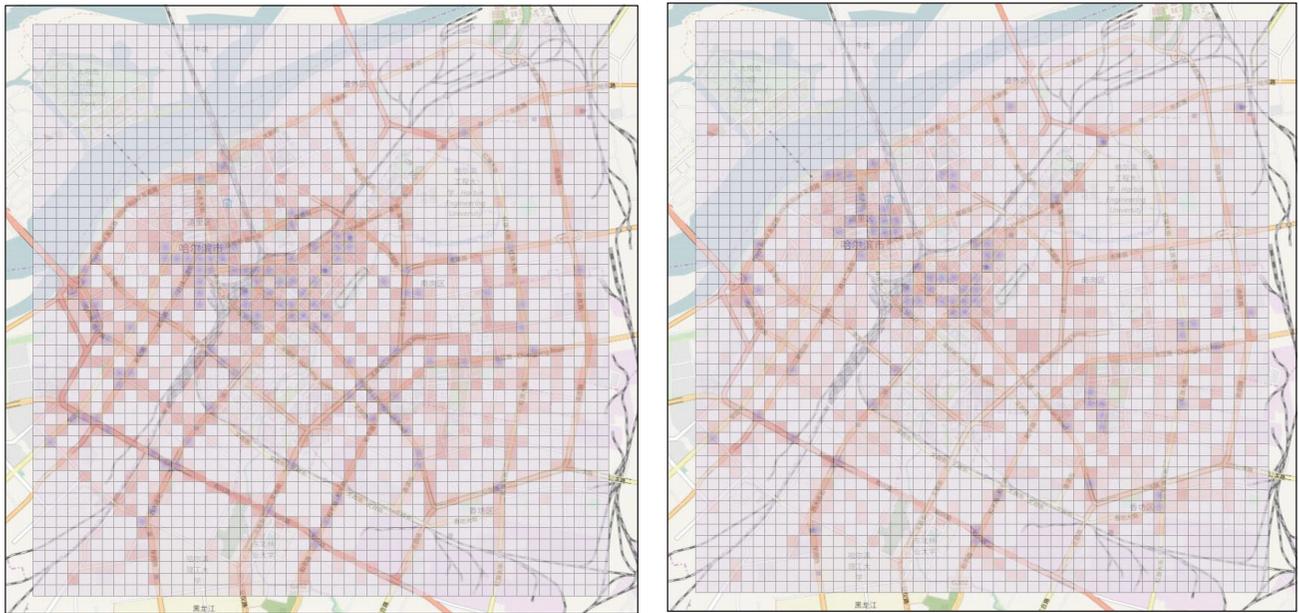


FIGURE 3: Illustration of divided cells. The base map is OpenStreetMap.



N
 0-500
 500-1000
 1001-1500

1501-2000
 2001-

(a) Origin clusters (boarding clusters)

N
 0-500
 500-1000
 1001-1500

1501-2000
 2001-

(b) Destination clusters (dropping off clusters)

FIGURE 4: Origin and destination clusters results.

when the value is larger than 1500, the cluster number is going to be stable. So we set the δ_N value as 2,000 each day and 10,000 for 5 days. In this light, we can measure the origin clusters (i.e. boarding clusters) and destination clusters (i.e. dropping off clusters) for the 5 days, as shown in Figure 4.

4.3. *Measuring Public Transit Riders' Travel Pattern.* For a specific OD pair, we can measure the travel pattern between them, that is, $Tr_{mode}(CL_i \rightarrow CL_j)$, $p_{mode}(CL_i \rightarrow CL_j)$, and $Tt_{mode}(CL_i \rightarrow CL_j)$, as mentioned in Section 3.2. Taking a 20×17 cells area, for example, as shown in Figure 5, there are

TABLE 1: Typical GPS records of a taxi.

Taxi ID	Timestamp	Latitude	Longitude	Status
0100320311	2015-08-21 17:08:21	45.726025	126.69883	17152
0100320311	2015-08-21 17:08:51	45.725937	126.70046	17152
0100320311	2015-08-21 17:09:21	45.72592	126.70041	16640
0100320311	2015-08-21 17:09:51	45.72424	126.70028	256
0100320311	2015-08-21 17:10:21	45.727768	126.678665	256
0100320311	2015-08-21 17:10:51	45.72415	126.70134	768
0100320311	2015-08-21 17:11:21	45.7243	126.70027	17152

Note. The columns from the left to right, respectively, represent the taxi identification; the taxi sampling time; the latitude and longitude of the position; and the occupation status.

TABLE 2: Typical Bus GPS records of a bus.

Route ID	Bus ID	Card ID	Timestamp	Latitude	Longitude
120	0100320311	313011090	2015-08-11 08:08:21	45.739658	126.643119
120	0100320311	313013116	2015-08-11 08:08:24	45.739658	126.643119
120	0100320311	313008101	2015-08-11 08:08:28	45.739658	126.643119
120	0100320311	313191231	2015-08-11 08:08:33	45.73966	126.643119

Note. The columns from the left to right, respectively, represent the bus route number; bus ID; smart card ID; the card sampling time; the latitude and longitude of the position.

TABLE 3: The dropping off data in sampling 5 days.

Date	Total	Average	Minimum	Maximum
20150803	1,121,198	448.5	0	26,371
20150804	1,112,459	445.0	0	24,508
20150805	1,112,058	444.8	0	25,100
20150806	1,111,987	444.8	0	24,766
20150807	1,110,522	444.2	0	24,958
Sum	1,113,645	445.5	0	25,140

Note. *Total* is the number of dropping off points in all cells, *average* is the average number of dropping off points in one cell, *minimum* is the minimum number of dropping off points in one cell, and *maximum* is the maximum number of dropping off points in one cell.

TABLE 4: Dropping off clustering results by different δ_N variously.

Date	Cluster number					
	$\delta_N = 500$	$\delta_N = 1000$	$\delta_N = 1500$	$\delta_N = 2000$	$\delta_N = 2500$	$\delta_N = 3000$
20150803	38	71	49	30	24	19
20150804	36	75	44	27	20	21
20150805	32	70	47	29	18	19
20150806	35	77	46	29	21	17
20150807	41	75	50	28	20	18
Average	36.4	73.6	47.2	28.6	20.6	18.8
Variance	11.3	8.8	5.7	1.3	4.8	2.2

one typical origin cluster and one typical destination cluster. And the three indicators' values in the survey 5 days can be calculated, which are shown in Table 5.

The Origin Cluster in this case study is the cell $C(48, 22)$, and the Destination Cluster consists of five cells (i.e. $C(35, 33)$, $C(35, 34)$, $C(36, 34)$, $C(36, 35)$, and $C(36, 36)$). The cell $C(48, 22)$ is located nearby Harbin Hongqi Resident District, Heping Resident District, Yuanda Central Park Resident District, and so on. And there is a bus station locating in this cell,

which consists of 12 bus routes (e.g. 14, 25, 31, and 44). And the Destination Cluster is located nearby Harbin International Golf Club, Harbin Wanda Plaza, several banks, and so on. This cluster area locates in an important commercial district in Harbin and definitely attracts many residents' travels. In this Cluster, there are several bus stations consisting of more than 20 bus routes (e.g. 17, 27, 34, and 71). From the Origin Cluster to Destination Cluster, there are at least three direct bus routes (i.e. 71, 82, and 209).

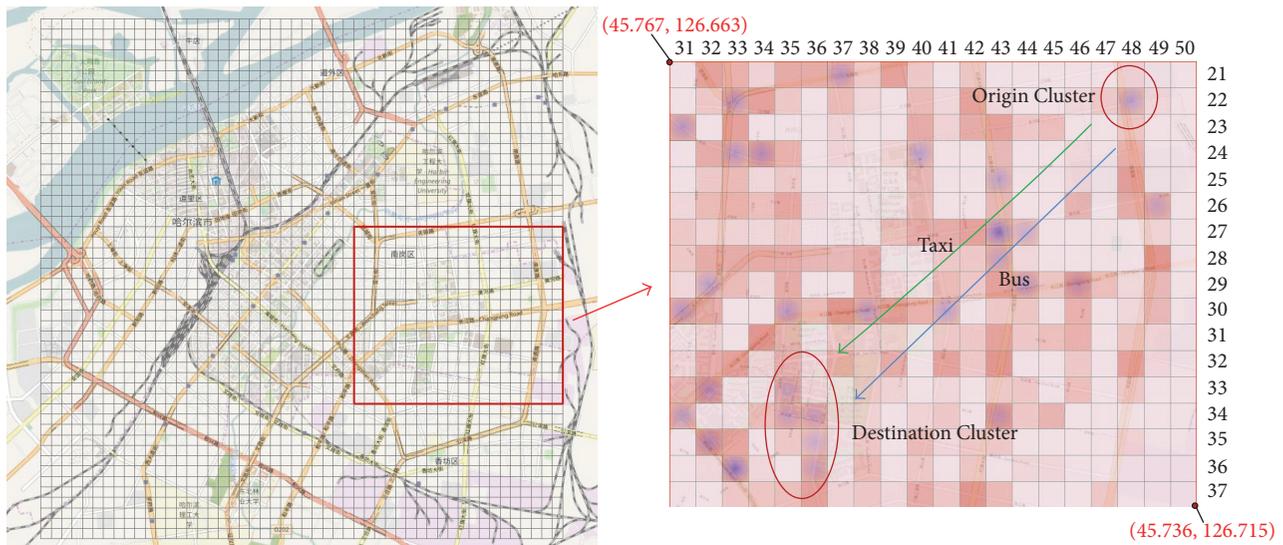
FIGURE 5: Illustration of a 20×17 cells area.

TABLE 5: Results of the three indicators.

	Taxi	Bus
Tr	367	509
p	41.9%	58.1%
Tt	16 min	40 min

5. Conclusion

This paper presented a cell-based urban public transit riders' travel pattern measurement method. The method used GPS-equipped public transit vehicle's locating data, which is more realistic and easy to collect. We proposed a customized DBSCAN algorithm to detect the origin and destination areas. We computed three indicators for each OD pair, which can reflect the relationship between the origin area and destination area. We carried out a numerical case study to evaluate our method, which uses taxi and bus GPS data in Harbin, China. The results can reflect OD pairs and relationship between each OD pair.

In the future, we will improve the customized DBSCAN algorithm by accuracy and efficiency. The travel pattern is a little simple in this paper, so we will further discover some other indicators in depth. And we plan to extend this research by utilizing more kinds of GPS data from different public transit vehicles.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (Project no. 51478151). This work was

performed at the Key Laboratory of Advanced Materials & Intelligent Control Technology on Transportation Safety, Ministry of Communications, China.

References

- [1] T. Litman, *Evaluating Public Transit Benefits and Costs*, Victoria Transport Policy Institute, British Columbia, Canada, 2015.
- [2] C. Ding, Y. Lin, and C. Liu, "Exploring the influence of built environment on tour-based commuter mode choice: a cross-classified multilevel modeling approach," *Transportation Research Part D: Transport and Environment*, vol. 32, pp. 230–238, 2014.
- [3] Beijing Transport Annual Report, Beijing Transportation Research Center, Beijing, China, 2015.
- [4] M. A. Munizaga and C. Palma, "Estimation of a disaggregate multimodal public transport origin-destination matrix from passive smartcard data from Santiago, Chile," *Transportation Research Part C: Emerging Technologies*, vol. 24, pp. 9–18, 2012.
- [5] C. Ding, S. Mishra, G. Lu, J. Yang, and C. Liu, "Influences of built environment characteristics and individual factors on commuting distance: a multilevel mixture hazard modeling approach," *Transportation Research Part D: Transport and Environment*, vol. 51, pp. 314–325, 2017.
- [6] C. Ding, D. Wang, C. Liu, Y. Zhang, and J. Yang, "Exploring the influence of built environment on travel mode choice considering the mediating effects of car ownership and travel distance," *Transportation Research Part A: Policy and Practice*, vol. 100, pp. 65–80, 2017.
- [7] A. Imaz, K. M. N. Habib, A. Shalaby, and A. O. Idris, "Investigating the factors affecting transit user loyalty," *Public Transport*, vol. 7, no. 1, pp. 39–60, 2014.
- [8] N. Nassir, M. Hickman, and Z.-L. Ma, "Activity detection and transfer identification for public transit fare card data," *Transportation*, vol. 42, no. 4, pp. 683–705, 2015.
- [9] Z.-J. Wang, X.-H. Li, and F. Chen, "Impact evaluation of a mass transit fare change on demand and revenue utilizing smart card data," *Transportation Research Part A: Policy and Practice*, vol. 77, pp. 213–224, 2015.

- [10] C. Ding, X. Wu, G. Yu, and Y. Wang, "A gradient boosting logit model to investigate driver's stop-or-run behavior at signalized intersections using high-resolution traffic data," *Transportation Research Part C: Emerging Technologies*, vol. 72, pp. 225–238, 2016.
- [11] Q. Ge and D. Fukuda, "Updating origin-destination matrices with aggregated data of GPS traces," *Transportation Research Part C: Emerging Technologies*, vol. 69, pp. 291–312, 2016.
- [12] A. Santos, N. McGuckin, H. Y. Nakamoto, D. Gray, and S. Liss, Summary of travel trends: 2009 national household travel survey (No. FHWA-PL-11-022), 2011.
- [13] K. M. Currans and K. J. Clifton, "Using household travel surveys to adjust ITE trip generation rates," *Journal of Transport and Land Use*, vol. 8, no. 1, pp. 85–119, 2015.
- [14] S. An, H. Yang, J. Wang, N. Cui, and J. Cui, "Mining urban recurrent congestion evolution patterns from GPS-equipped vehicle mobility data," *Information Sciences*, vol. 373, pp. 515–526, 2016.
- [15] R. O. Arbex, B. B. Alves, and M. A. Giannotti, "Comparing accessibility in urban slums using smart card and bus GPS data," in *Proceedings of the Transportation Research Board 95th Annual Meeting (No. 16-5614)*, 2016.
- [16] J. Hu, B. Park, and A. E. Parkany, "Transit signal priority with connected vehicle technology," *Transportation Research Record*, vol. 2418, pp. 20–29, 2014.
- [17] Y. Long and Z. Shen, "Discovering functional zones using bus smart card data and points of interest in Beijing," in *Geospatial Analysis to Support Urban Planning in Beijing*, vol. 116 of *Geo-Journal Library*, pp. 193–217, Springer International Publishing, Beijing, China, 2015.
- [18] Y. Bie, X. Gong, and Z. Liu, "Time of day intervals partition for bus schedule using GPS data," *Transportation Research Part C*, vol. 60, pp. 443–456, 2015.
- [19] J. Hu, M. D. Fontaine, B. B. Park, and J. Ma, "Field evaluations of an adaptive traffic signal—using private-sector probe data," *Journal of Transportation Engineering*, vol. 142, no. 1, Article ID 04015033, 2016.
- [20] X. Ma, Y. J. Wu, Y. Wang, F. Chen, and J. Liu, "Mining smart card data for transit riders travel patterns," *Transportation Research Part C: Emerging Technologies*, vol. 36, no. Part C, p. 12, 2013.
- [21] S. Tao, D. Rohde, and J. Corcoran, "Examining the spatial-temporal dynamics of bus passenger travel behaviour using smart card data and the flow-comap," *Journal of Transport Geography*, vol. 41, pp. 21–36, 2014.
- [22] L.-M. Kieu, A. Bhaskar, and E. Chung, "A modified density-based scanning algorithm with noise for spatial travel pattern analysis from smart card AFC data," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 193–207, 2015.
- [23] R. Paleti, P. Vovsha, D. Givon, and Y. Birotker, "Impact of individual daily travel pattern on value of time," *Transportation*, vol. 42, no. 6, pp. 1003–1017, 2015.
- [24] J. Cui, F. Liu, J. Hu, D. Janssens, G. Wets, and M. Cools, "Identifying mismatch between urban travel demand and transport network services using GPS data: a case study in the fast growing Chinese city of Harbin," *Neurocomputing*, vol. 181, pp. 4–18, 2016.
- [25] Harbin Statistics Yearbook Editorial Department, *Energy Statistics Yearbook*, China Statistics Press, Harbin, China, 2014.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

