

Research Article

Mining Travel Time of Airport Ferry Network Based on Historical Trajectory Data

Cong Ding ¹, Jun Bi ², Dongfan Xie,^{1,2} Xiaomei Zhao,^{1,2} and Yi Liu¹

¹School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China

²Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing, China

Correspondence should be addressed to Jun Bi; jbi@bjtu.edu.cn

Received 5 August 2021; Accepted 8 October 2021; Published 31 October 2021

Academic Editor: Ming Xu

Copyright © 2021 Cong Ding 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.

An airport ferry vehicle is a ground service vehicle used to transfer passengers between the far apron and the terminal. The travel time of ferry tasks in the airport ferry network is an important decision-making basis for ferry vehicle scheduling. This paper presents a graph-based method to mine the travel time between nodes in the airport ferry network. Firstly, combined with map and trajectory information, the method takes the terminal boarding gates, parking lots, and remote stands as road network nodes to build a complete airport ferry road network. Then, this paper uses big data processing technology to identify the travel time between regional connection nodes by data fusion through the temporal and spatial relationship between flight schedule and ferry vehicle GPS travel trajectory. Finally, the Floyd shortest path algorithm in graph theory is used to obtain the shortest path and travel time of all OD points. The experimental results show that all the ferry times calculated by the method proposed in this paper can better reflect the actual driving situation. This method saves the manpower, material resources, and time cost of on-site investigation and lays a foundation for the scheduling of ferry vehicles.

1. Introduction

When the aircraft is parked at a far apron, passengers need ferry vehicles to transfer between the aircraft and the terminal. After the airport operation center obtains the flight schedule and gate assignment [1, 2] results in a period of time, it will consider the travel time consumed by the ferry task in the road network and reasonably arrange the ferry vehicles for each flight parked in a remote stand. Only a reasonable ferry scheduling plan can ensure the normal implementation of the flight schedule. At present, the scheduling of ferry vehicles in many airports mainly depends on experienced staff. The efficiency of the scheduling results is difficult to be ensured, which is easy to lead to flight delays. Therefore, how to establish the optimal scheduling model of ferry vehicles by scientific means and improve the service level of airport ferry vehicles has attracted the attention of some scholars.

The premise of ferry vehicle scheduling modeling is to know the travel time information between OD points in the

ferry network. However, due to the closed environment in the airport flight area, nonstaff cannot enter the interior for field measurement. With the rapid development of smart airports, in order to realize the digitization and intelligence of airport management and meet the needs of airport safety management, some airports have established ground service vehicle GPS real-time monitoring and management systems to realize vehicle positioning, real-time tracking, track playback, cross-border alarm, and other functions. The historical GPS track data of ferry vehicles recorded by the system can reflect the time information and spatial information of vehicles in the process of performing flight support tasks. Therefore, using the trajectory data to mine and extract the travel time of the airport ferry network is a new idea.

This paper proposed a novel graph-based method to mine the travel time in the airport ferry network. The research work is mainly divided into three aspects. Firstly, combined with the map and trajectory information, this

paper takes the terminal boarding gates, parking lots, and remote stands as the road network nodes to build a complete airport ferry road network. Then, the nodes are divided into different regions, and the travel time between regional connecting nodes is identified by data fusion through the temporal and spatial relationship between flight schedule and ferry vehicle GPS travel trajectory. Finally, the Floyd shortest path algorithm in graph theory is used to obtain the shortest path and travel time of all OD points. The research results of this paper can save the manpower, material resources, and time cost of on-site investigation and provide a decision-making basis for the optimal scheduling model of ferry vehicles.

The rest of this paper is organized as follows. In Section 2, the literature on the optimal scheduling of airport ground support vehicles and GPS trajectory mining is reviewed. Section 3 takes Kunming Changshui International Airport in China as the research object and introduces the method of constructing the airport ferry network. In Section 4, the calculation method of OD point travel time in the road network is designed. The example verification and results are given in Section 5. Finally, conclusions and future research directions are proposed in Section 6.

2. Literature Review

With the rapid development of civil aviation, the disadvantages of airport management and operation relying only on manual experience are becoming more and more prominent. In order to ensure the normal operation of flights, how to improve the service level of airport ground support services through scientific means has attracted more and more attention of scholars. From the optimal operation of special vehicles providing various ground services to the overall optimization of joint scheduling of multiple vehicles, many achievements have been made. This paper summarizes the research literature of different ground service vehicles as a reference. Although there are differences in service content with ferry vehicles, there are many similarities in the optimization objectives and constraints in their scheduling.

Norin et al. [3] constructed an airport deicing vehicle optimal scheduling model aiming at minimizing the delay time and the minimum number of deicing vehicles and designed the corresponding solution algorithm based on a greedy randomized adaptive search algorithm. Du et al. [4] studied the trailer scheduling problem in flight transit service, described the problem as VRPTW, and constructed an integer programming model with minimizing the trailer operation cost as the optimization objective and the vehicle operation restriction as the constraint. Starting from the flexibility of vehicle service, Wang et al. [5] proposed a scheduling algorithm based on a greedy strategy to deal with the dynamic scheduling problem of airport refueling vehicles. With the diversification of airport ground service and considering the interaction between various vehicle services, many studies begin to pay attention to the joint scheduling of airport ground service vehicles. Padron et al. [6] constructed a double objective optimization model to minimize waiting time and turnaround time for the joint scheduling of

airport service equipment. Xu and Shao [7] evaluated the fluctuation of service equipment operation time by analyzing the airport historical data and proposed an optimization model of ground service support equipment with uncertain operation time. Fei and Shu'an [8] studied the optimization problem of airport service vehicle scheduling in peak hours and constructed a joint scheduling model of ground service vehicles to minimize the purchase cost and operation cost of service equipment. In terms of ferry vehicles, Zhao et al. [9] established an integer programming model with the goal of minimizing the number of ferry vehicles required in peak hours or a certain period and constructed a ferry vehicle sharing network, which transformed the model into the problem of maximum network flow. Han et al. [10] also proposed a ferry capacity network model, in which the directed edge indicates that the two associated nodes may be continuously served by the same ferry. The model aims to minimize the number of ferry vehicles required and is solved by the method of the graph. However, their research only roughly estimated the travel time between the terminal and the far apron area in the process of building the ferry network. For large airports, the terminal covers a wide area and is far away from different boarding gates. The unified estimation method will cause great errors in practical application.

With the development of positioning technology, many vehicles are equipped with GPS receivers. During the moving process, the vehicle continuously collects real-time information, including position, motion parameters, and positioning time, and transmits it to the data center. This type of data is called floating car data [11, 12]. The emergence of these data makes it possible to mine rich knowledge from GPS trajectory data by using big data analysis technology [13]. Palma et al. and Bhattacharya et al. [14, 15] proposed the method of analyzing motion characteristics (velocity, azimuth, acceleration, etc.) to mine important places related to people and objects. Zheng and Xie [16] proposed a trajectory mining algorithm to analyze the user's GPS trajectory data, so as to recommend personalized tourist attractions. In addition, Li et al. [17] analyzed the typical characteristics of the data of floating vehicles in the parking lot and used the DBSCAN algorithm to extract the parking lot location. Wang et al. [18] proposed the identification of key nodes and sections of urban road networks based on GPS trajectory data. In terms of path extraction, Schoredl et al. and Li et al. [19, 20] proposed two methods to extract a high-precision roadmap from GPS trajectory, which are applicable to high sampling frequency and low sampling frequency of positioning data respectively. Unlike them, Cao and Krumm [21] proposed a new gravity model to transform the original GPS track into a road network that can guide path selection. Tang et al. [22] proposed a lane-level road network information mining method based on lane number and turning rules. In addition to the region of interest and path recognition, scholars have also done a lot of research on spatiotemporal pattern extraction based on trajectory data, for example, the method of automatically extracting passenger train operation information from historical track data [23]. Dong et al. [24] proposed a study on the temporal and spatial change of

traffic accessibility under public health emergencies based on GPS trajectory. Lei et al. and Zhao et al. [25, 26] proposed a spatiotemporal analysis model to capture the motion mode of the object. Different from those, Lu et al. [27] proposed a visual analysis method to study the behavior of vehicles on a certain route. Some scholars have also studied periodic pattern recognition, such as the probabilistic periodic detection method of moving objects [28]. In general, most of the GPS data mining and analysis research by scholars is based on the road traffic network, and few studies use the GPS track of airport ferry vehicles to extract the road network structure and travel time of the flight area. Therefore, the research content of this paper is innovative.

3. Construction of Airport Ferry Network

This paper takes the real data of Kunming Changshui International Airport in China as the research object. The airport terminal covers an area of 548300 square meters, with 78 boarding gates, 104 remote stands, 2 entry ports, and 3 parking lots. Figure 1 shows the spatial layout of the airport flight area. The boarding gates and entry ports are distributed around the terminal, while the parking lots and remote stands are distributed at the far apron.

When a flight is assigned to an apron, the ferry vehicle needs to arrive at the corresponding stand in advance for the arriving flight. After the plane arrives, the ferry vehicle will drive along the planned route to deliver the passengers to the entrance port of the terminal. For departure flights, according to the flight schedule, the ferry vehicle needs to arrive at the boarding gate before ticket check-in. After the gate is opened, passengers will be sent to the designated parking stand. There is also a necessary transfer time between two consecutive services of the ferry vehicle, that is, the travel time required for the ferry vehicle from the end of the last service to the start of the next service. According to the arrival and departure attributes of two adjacent service flights, the transfer between two services can be divided into four categories, as shown in Table 1. After a ferry task, when the buffer time is sufficient, the ferry vehicle can go to the parking lot first and then go to the starting place of the next service when the next task is approaching. When the buffer time is insufficient, in order to avoid flight delays, the ferry vehicle can go directly to the starting point of the next service. The security efficiency of ferry vehicles is mainly affected by the transfer time. Flights with a remote stand at the airport usually change the boarding gate flexibly according to the situation. For large airports, the terminal covers a wide area and different boarding gates are far away from each other. Therefore, calculating the transfer path and time between all OD points in the whole airport is the prerequisite for the optimal scheduling of ferry vehicles.

Regarding the boarding gates and remote stands as road network nodes, directly calculating the travel time of any two points in the road network according to the coordinate information and the travel track of ferry vehicles will cause a large workload. In order to simplify the airport road network structure, firstly, the airport terminal boarding gate and apron are divided into several relatively independent areas

according to the adjacent relationship, and the spatial transfer of ferry vehicles between service points is regarded as the transfer between regions. According to the number of remote stands, apron 3 is divided into 1 to 4 parts, apron 5 is divided into 1 to 4 parts, and apron 7 is divided into 1 to 2 parts. The boarding gate area of the terminal is divided into 1 to 4 parts for the convenience of the road network structure. Figure 2 shows the historical track of ferry vehicles. Combined with the actual road network distribution of the airport, the connection relationship between regions can be obtained, as shown in Figure 3. The arrangement structure of service points in one region is relatively simple, including series and parallel. As shown in Figure 4, there are two rows of remote stands in area 5-2, which are distributed in parallel. Area 3-4 has only a single row of stands, which are distributed in series. Therefore, combined with the road network structure between regions, a complete road network structure of airport service points can be obtained, as shown in Figure 5. After obtaining the spatial connection relationship between all service points from the road network diagram of service points, any service point can communicate with other service points through the road network, and when combined with the distance information on the edge, the travel time between any two points can be calculated. Therefore, the next section will explore the travel time of each edge of the road network.

4. Calculation Method of Travel Time

According to the ferry road network established in the previous section, this section uses the historical GPS track data of ferry vehicles to obtain the weight of each edge. For the convenience of description, the following definitions are made in this paper: all nodes in the ferry network are represented by $V = [D_1, D_2, \dots, D_i, \dots, D_n]$, where D_i represents different regions. The nodes in one region are divided into internal nodes and connection nodes, $D_i = [P_{i1}, P_{i2}, \dots, P_{im}, Q_{i1}, Q_{i2}, \dots, Q_{ik}]$, the internal node P_{im} is only connected with the nodes in the region, and the connection node Q_{ik} can be connected with the nodes in other regions. The GPS tracks of all ferry vehicles are represented by $R = [G_1, G_2, \dots, G_i, \dots, G_n]$, where G_i represents the track set of ferry vehicle i , including the longitude and latitude coordinates $[\text{lng}_i, \text{lat}_i, t]$ of the vehicle at each time. The flight set is represented as $F = [f_1, f_2, \dots, f_i, \dots, f_n]$, where the arrival and departure attributes, time, and OD point of the ferry task of flight f_i are known.

4.1. Travel Time between Terminal and Far Apron. During the flight support task, the ferry vehicle will travel between the boarding gate area and the far stand area. Therefore, the path and time of the connecting edge between the gate area and the far stand area can be mined according to the temporal and spatial correlation between the flight schedule and the GPS track of the ferry vehicle. For arrival flights, the service flow of ferry vehicles is shown in Figure 6. t_i^{ETA} indicates the estimated arrival time of the flight. The ferry service starts from the time of arriving at the remote stand in advance, and

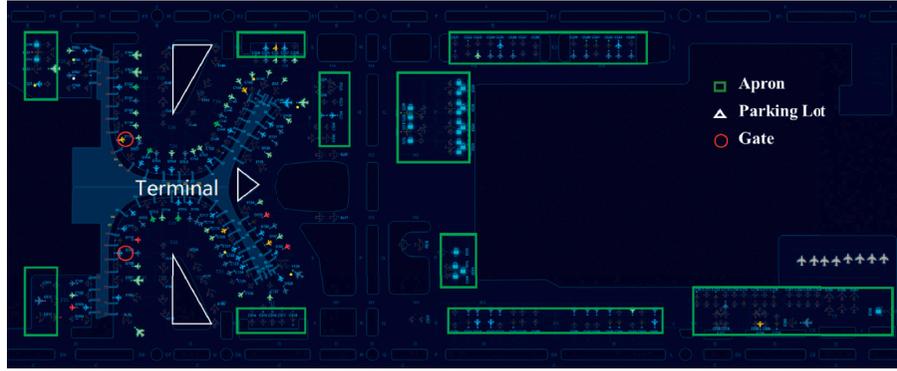


FIGURE 1: Spatial layout of the airport flight area.

TABLE 1: OD point conversion of continuous flights.

Flight 1	Flight 2	Origin	Destination
Arrive flight	Arrive flight	Arrive gate	Remote stand
Arrive flight	Departure flight	Arrive gate	Departure gate
Departure flight	Departure flight	Remote stand	Remote stand
Departure flight	Arrive flight	Remote stand	Departure gate

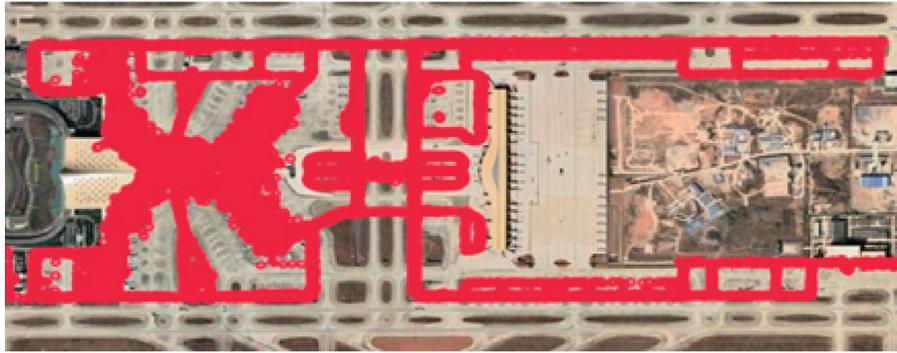


FIGURE 2: Historical GPS track of ferry vehicles.

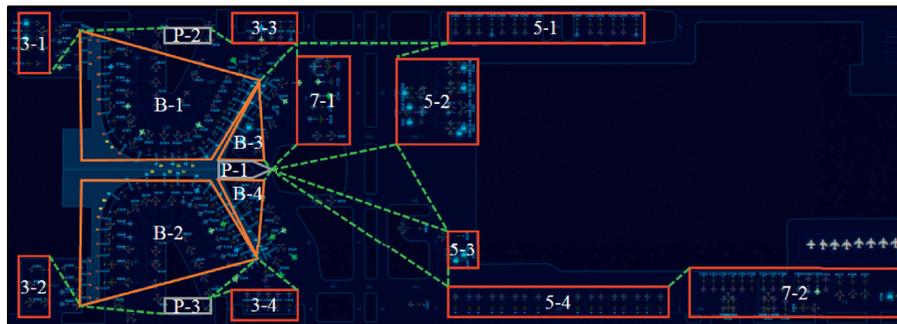


FIGURE 3: Road network connection between the regions.

the service start time TS_i^A can be calculated according to equation (1). After the flight arrives at the remote stand, it takes time T_{wait} to wait for passengers to board. The ferry service ends after the passengers get off at the port, and the service end time TE_i^A can be calculated according to equation (2).

$$TS_i^A = t_i^{\text{ETA}} - T_{\text{adv}}^A, \quad (1)$$

$$TE_i^A = t_i^{\text{ETA}} + 2 * T_{\text{wait}} + T_i^{\text{ferry}}. \quad (2)$$

Similarly, for departure flights, according to the flight schedule, the ferry vehicle arrives at the boarding gate before

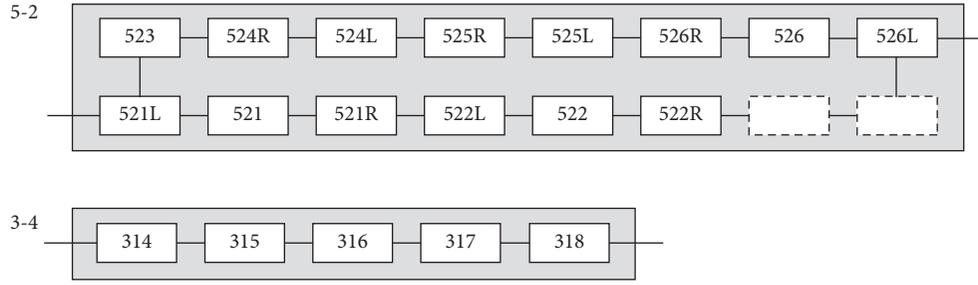


FIGURE 4: Nodes connection relationship in a region.

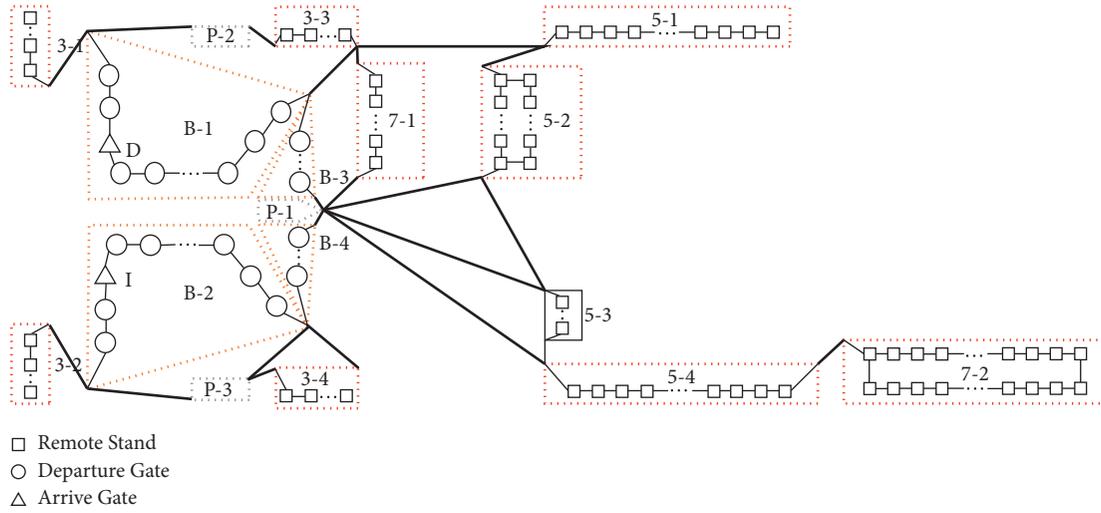


FIGURE 5: A complete structure diagram of airport ferry network.

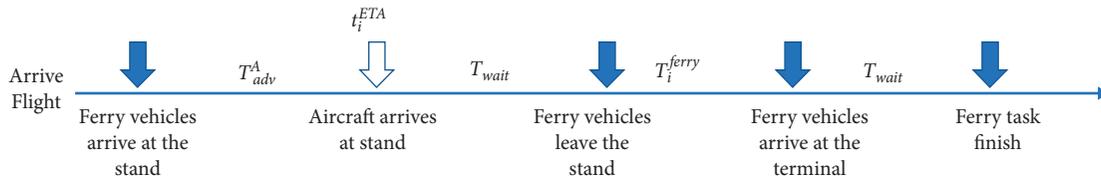


FIGURE 6: Ferry vehicle service process of arrival flight.

ticket check-in. After passengers get off at the parking stand, the service of the ferry vehicle ends. Figure 7 shows the ferry vehicle service process of departure flights. Assuming that the opening time of the boarding gate is 40 minutes (according to the service standards of China’s civil transport airport) before the departure time t_i^{ETD} of the flight, the start time of the ferry vehicle for the departure flight is calculated by equation (3) and the end time of the service is calculated by equation (4).

$$TS_i^D = t_i^{ETD} - 40 \text{ min} - T_{adv}^D, \quad (3)$$

$$TE_i^D = t_i^{ETD} - 40 \text{ min} + 2 * T_{wait} + T_i^{ferry}. \quad (4)$$

According to the ferry service process, the algorithm steps of mining the travel time T_i^{ferry} between the boarding

gate area and the apron by integrating GPS track data and flight data are as follows.

Step 1. According to the road network structure, find out the connecting nodes between the boarding gate area and the far apron area to form OD pairs $[Q_{ik}, Q_{jl}]$. Search the flight set F and find the flight i whose starting and ending point of the ferry task is $[Q_{ik}, Q_{jl}]$.

Step 2. If flight i found is a departure flight, obtain the departure time t_i^{ETD} from the flight information, take time t and distance m as the search range parameters, form a set $Ferry_{start}$ of all ferry vehicles that appear at the boarding gate Q_{ik} in the period $[t_i^{ETD} - 40 \text{ min} - t, t_i^{ETD} - 40 \text{ min} + t]$, and record the time T_{start} when each ferry car leaves the search

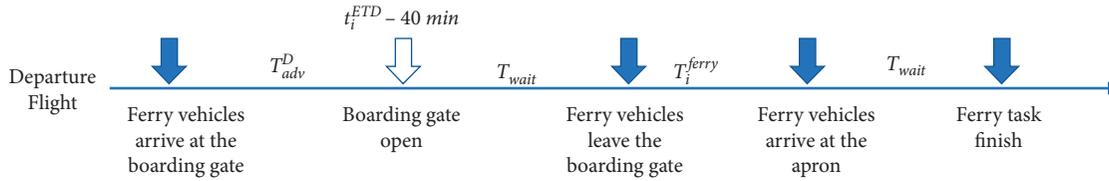


FIGURE 7: Ferry vehicle service process of departure flight.

range. Similarly, find out all the ferry vehicles that appear at the remote stand Q_{ji} in the period $[t_i^{ETD} - 40 \text{ min} + t, t_i^{ETD}]$, form a set $\text{Ferry}_{\text{end}}$, and record the time T_{end} when each ferry car enters the search range. If flight i found is an arrival flight, obtain the arrival time t_i^{ETA} from the flight information, and change the starting point to the remote stand Q_{ji} , the time range is $[t_i^{ETA} - t, t_i^{ETA} + t]$, the end point is the port Q_{ik} , and the time range is $[t_i^{ETA}, t_i^{ETA} + t_{\text{max}}]$ where t_{max} represents the maximum ferry time range.

Step 3. Take the ferry car G_i existing in both $\text{Ferry}_{\text{start}}$ and $\text{Ferry}_{\text{end}}$ as the ferry car i serving the flight, and calculate the travel time $T_{\text{end}}^i - T_{\text{start}}^i$ of the ferry car as the travel time from the boarding gate to the stand.

When judging whether the ferry car enters the search range, take the boarding gate or stand as the center of the circle, first calculate the longitude and latitude distance between the ferry car and the center of the circle, and then compare this distance with the reference distance m . If the calculated distance is less than the reference distance, judge whether the ferry car enters the search range.

4.2. Travel Time between Different Far Aprons. The starting and ending point of flight service is from the boarding gate to the stand, so it is impossible to obtain the time when OD points are both stands through flight information. There are two situations for the travel time of the ferry car between two target stands. The first situation is that the ferry car directly travels from the starting stand to the ending stand. At this time, the travel time represents the distance between the two stands. Another situation is that the ferry car drives from the starting stand to other places and then to the target stand. At this time, the travel time is obviously greater than the distance between the two stands. Therefore, without the mining time range, we can only intercept the GPS track of the ferry vehicle through the coordinate values of the connecting nodes in two far aprons, then count a large number of travel time values, and select the smaller value as the travel time between the aprons. In this way, the distance from the parking lot to other service points can also be obtained.

4.3. Travel Time between Any Two Nodes. After mining the travel time information of edges in the road network through the above method, the nodes, edges, and the weight of edges of the road network have been obtained. The travel time between any two nodes can be calculated by using the shortest path algorithm. Calculating the shortest path is a classical problem in graph theory. At present, there has been

a very mature research. The commonly used shortest path algorithms include Floyd [29], Dijkstra [30], and SPFA [31]. It is necessary to select the appropriate algorithm according to different use scenarios. From the characteristics of the algorithm, the Floyd algorithm is suitable for finding the shortest path from multiple sources, but the high time complexity makes the algorithm not suitable for road network maps with many nodes; the Dijkstra algorithm is the basis of all basic shortest path algorithms. It is the most stable algorithm, but it is suitable for finding the shortest path of a single source. It needs repeated operation to obtain the shortest path of all points; the SPFA algorithm is a queue optimization of the Bellman Ford algorithm. It is the shortest path algorithm based on BFS. It solves the problem of changing weight to negative value that the Dijkstra algorithm cannot solve. At the same time, the implementation of the algorithm is the most troublesome. The road network constructed in this paper has a total of 187 service points. Considering the scale of the road network, the positive right of way, and the need to obtain the path information between all points, the Floyd algorithm is selected to solve the shortest path.

Floyd algorithm is an algorithm that uses the idea of dynamic programming to find the shortest path between multiple source points in a given weighted graph. The algorithm has the characteristics of clear structure and fast implementation. Its core idea is as follows: for each pair of shortest vertices u and v , if there is a vertex w that makes the path from u to w and then to v shorter than the known path, update it. Considering that the nodes have been divided into different regions, this paper first uses the Floyd algorithm to calculate the shortest circuit between different regions and then calculates the shortest circuit between the nodes in the region and the connecting nodes. The algorithm flow of finding the shortest path of OD points of ferry task is as follows (Algorithm 1):

After the shortest path between regions is obtained by the above algorithm, combined with the distance from the internal node to the connecting node in the region, the shortest travel time between any two points can be obtained. The edge inside the area mainly represents the distance between two gates, and the fixed value C_{gate} can be adopted. If n stands are separated between the internal node and the connecting node, the distance is $n * C_{\text{gate}}$.

5. Experimental Results and Analysis

In order to verify the proposed travel time mining method of the ferry network, this paper uses the actual data of Changshui International Airport to verify the effectiveness

(i) Input: region ferry network $G = (Q, E)$, Q is the set of connection nodes representing each region, and E is the travel time between regions obtained by data mining as the weight of edges.
(1) Initialization matrix M , m_{ij} represents the travel time between vertices i and j . If i and j are not directly connected, the weight is set to ∞ . If i and j come from the same region, the weight is set to 0.
(2) Starting from the first node q_1 , calculate that $m_{ij}^{(1)} = \min\{m_{ij}^{(0)}, m_{i1}^{(0)} + m_{1j}^{(0)}\}$ is the shortest path from i to j that only allows the middle to pass through q_1 , and update the matrix M .
(3) Similarly, calculate that $m_{ij}^{(k)} = \min\{m_{ij}^{(k-1)}, m_{ik}^{(k-1)} + m_{kj}^{(k-1)}\}$ is the shortest path from i to j that only allows the middle to pass through q_1, q_2, \dots, q_k , and update the matrix M .
(4) When $k = n$, m_{ij} in matrix M is the shortest path from node i to j
(ii) Output: matrix M .

ALGORITHM 1: Calculating the shortest path between ferry OD points by the Floyd algorithm.

TABLE 2: Flight information (partial departure flight).

Flight number	Stand	Boarding gate	Check-in time	Close time	Departure time
TV6031	153	27	2019/1/1 0:01	2019/1/1 0:13	2019/1/1 0:37
3U8837	155	17	2019/1/1 0:03	2019/1/1 0:24	2019/1/1 0:51
3U8107	326	4	2019/1/1 0:12	2019/1/1 0:36	2019/1/1 1:03
KY8335	521L	55	2019/1/1 0:15	2019/1/1 0:49	2019/1/1 0:59
CA5785	521L	55	2019/1/1 0:15	2019/1/1 0:49	2019/1/1 0:59
ZH5335	521L	55	2019/1/1 0:15	2019/1/1 0:49	2019/1/1 0:59
MU2569	109	70	2019/1/1 0:18	2019/1/1 0:32	2019/1/1 0:50
ZH3181	145	59	2019/1/1 0:23	2019/1/1 0:45	2019/1/1 0:56
KY8281	145	59	2019/1/1 0:23	2019/1/1 0:45	2019/1/1 0:56
CZ3416	141	60	2019/1/1 0:23	2019/1/1 0:34	2019/1/1 0:50

TABLE 3: GPS trajectory information of the ferry car (partial).

Car ID	GPS time	Longitude	Latitude
1	2019/1/1 12:01:03	102.932983	25.111266
1	2019/1/1 12:02:01	102.936	25.108516
1	2019/1/1 12:03:01	102.938266	25.10955
1	2019/1/1 12:04:01	102.939033	25.109433
1	2019/1/1 12:05:31	102.939066	25.109433
1	2019/1/1 12:06:29	102.939066	25.109416
1	2019/1/1 12:06:59	102.939066	25.109416
1	2019/1/1 12:08:59	102.939083	25.1094
1	2019/1/1 12:09:29	102.939066	25.1094
1	2019/1/1 12:10:59	102.93905	25.1094

and practicability of the method. Flight data from January to June 2019, historical GPS track data of ferry vehicles, and coordinate information of each parking stand and boarding gate of the airport are collected. The data samples are shown in Tables 2 to 4.

Since the airport does not record the daily scheduling historical data of ferry vehicles, we do not know which vehicle performs each task. Therefore, we need to identify them according to the data fusion method proposed in Section 4 to obtain the specific ferry vehicles corresponding to the task and the time nodes of their entry and departure starting and ending points. According to the method in Section 4, we first excavate the distance between the boarding gate area and the far apron area. Ferry vehicles will not stop too far away from the aircraft and stop at the task

site for too long, so we set the time scaling parameter $t = 10$ min and the space distance scaling parameter $m = 20$ m. Take boarding gate area B-1 and far apron 5-2 as examples, according to the identification results, the corresponding starting waiting time, travel time, and terminal waiting time can be calculated, and the GPS track of the vehicle can be checked back according to the identification time of the starting and ending points to obtain the support path, as shown in Figure 8. The yellow circle indicates the search range, and the time difference of these seven GPS data can be calculated as the travel time of connection nodes gate B48 and stand 521L. According to this, the results of each calculation are collected, and the mean value of all results is taken as the transfer time between the two. For the distance between different far apron areas, we count a lot of travel

TABLE 4: Latitude and longitude information of service point (partial).

Number	Boarding gates		Number	Remote stands	
	Longitude	Latitude		Longitude	Latitude
<i>b1</i>	102.9274531	25.1011307	311	102.9290611	25.0955317
<i>b2</i>	102.9265117	25.1012863	312	102.9295471	25.0951637
<i>b3</i>	102.9271301	25.1010857	313	102.9300271	25.0946687
<i>b4</i>	102.9268037	25.1012533	314	102.9359601	25.1012577
<i>b5</i>	102.9257167	25.1015113	315	102.9362391	25.1016377
<i>b6</i>	102.9252457	25.1018223	316	102.9364991	25.1019157
<i>b7</i>	102.9248317	25.1021203	317	102.9367597	25.1021916
<i>b8</i>	102.9243962	25.1024135	318	102.9370517	25.1025726
<i>b9</i>	102.9236602	25.1022575	321	102.9222547	25.1004473
<i>b10</i>	102.9238932	25.1027735	322	102.9216707	25.1006893

TABLE 5: Distances between areas in the road network.

Region 1		Region 2		Travel time (min)
Number	Connection node	Number	Connection node	
<i>B-1</i>	<i>b9</i>	3-1	321	1.6
<i>B-1</i>	<i>b56</i>	3-3	328	1.7
<i>B-2</i>	<i>b72</i>	3-2	311	1.5
<i>B-2</i>	<i>b40</i>	3-4	318	2.3
<i>B-3</i>	<i>b48</i>	<i>P-1</i>	<i>p5 p6</i>	0.5
<i>B-4</i>	<i>b32</i>	<i>P-1</i>	<i>p5 p6</i>	0.5
<i>P-1</i>	<i>p5 p6</i>	7-1	330	3.1
<i>P-1</i>	<i>p5 p6</i>	5-2	521L	6.5
<i>P-1</i>	<i>p5 p6</i>	5-3	520	8.5
<i>P-1</i>	<i>p5 p6</i>	5-3	516	9.5
<i>P-2</i>	<i>p1</i>	<i>B-1</i>	<i>b13</i>	2.2
<i>P-2</i>	<i>p1</i>	3-3	324	1.4
<i>P-3</i>	<i>p3</i>	<i>B-2</i>	<i>b78</i>	2.1
<i>P-3</i>	<i>p3</i>	<i>B-2</i>	<i>b40</i>	2.8
<i>P-3</i>	<i>p3</i>	3-4	314	1.3
3-3	328	7-1	722a	1.8
3-3	328	5-1	531	3.6
5-2	526L	5-1	531	1.0
5-3	519R	5-4	516	1.4
5-4	501	7-2	710	1.3



FIGURE 8: Example of ferry car path calculation.

time between two connection nodes. Figure 9 shows the distance distribution characteristics from gate 48 to *P-1* parking lot connection point *p5|p6* and from stand 526L to stand 531. The corresponding distance is represented by

selecting a smaller value, which can be taken as 0.5 and 1.0 minutes. Through the above two methods, the transfer time between regional connection nodes in the outlet network can be calculated, as shown in Table 5.

With the travel time between connection nodes, the shortest travel time between any regions can be calculated according to the Floyd shortest path algorithm. The distance between aircraft stands in the area is $C_{gate} = 40$ m according to the actual situation of Changshui airport. When the maximum allowable speed of the ferry vehicle in the flight area is 25 km/h, the travel time between the inner edges of the area can be calculated as about 0.1 minutes. Therefore, the distance between any two nodes in the airport ferry network is equal to the distance between regions plus the distance between these two points and the connecting nodes in the region. The calculation results are shown in Table 6.

In order to check the accuracy of the distance information mined in this paper, we directly calculate the distance between the internal nodes of the two randomly selected areas as the real value through the above distance calculation method and

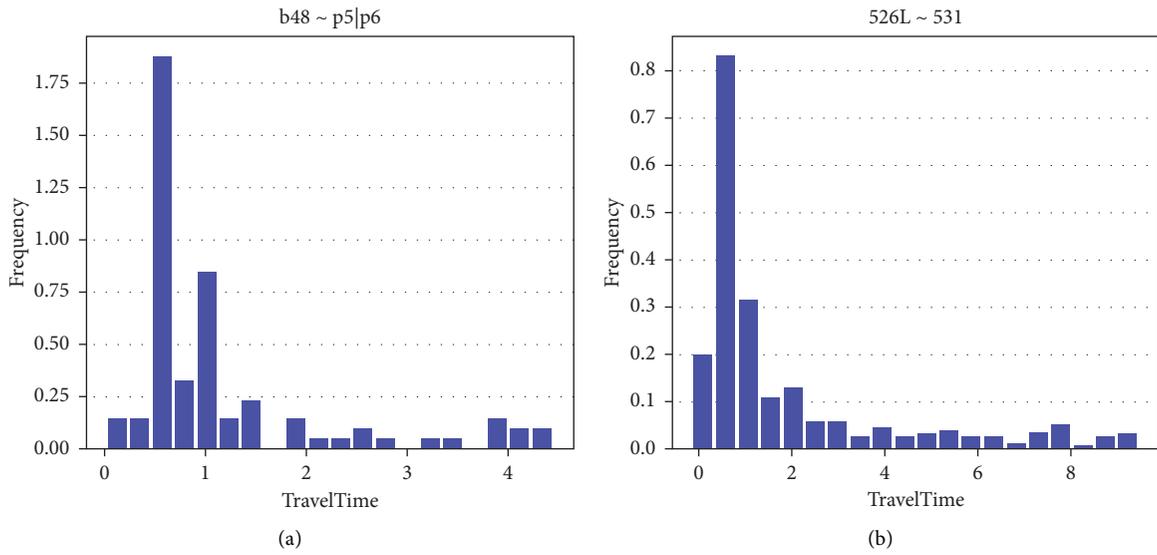


FIGURE 9: Histogram of OD distance distribution from b48 to p5|p6 (a) and 526L to 531 (b).

TABLE 6: OD distances calculated by the Floyd algorithm (partial).

	328	329	330	521	521L	521R	522	522L	522R	523
328	0.0	2.3	2.4	4.8	4.9	4.7	4.5	4.6	4.4	5.0
329	2.3	0.0	0.1	7.1	7.2	7.0	6.8	6.9	6.7	7.3
330	2.4	0.1	0.0	7.2	7.3	7.1	6.9	7.0	6.8	7.4
521	4.8	7.1	7.2	0.0	0.1	0.1	0.3	0.2	0.4	0.2
521L	4.9	7.2	7.3	0.1	0.0	0.2	0.4	0.3	0.5	0.1
521R	4.7	7.0	7.1	0.1	0.2	0.0	0.2	0.1	0.3	0.3
522	4.5	6.8	6.9	0.3	0.4	0.2	0.0	0.1	0.1	0.5
522L	4.6	6.9	7.0	0.2	0.3	0.1	0.1	0.0	0.2	0.4
522R	4.4	6.7	6.8	0.4	0.5	0.3	0.1	0.2	0.0	0.6
523	5.0	7.3	7.4	0.2	0.1	0.3	0.5	0.4	0.6	0.0

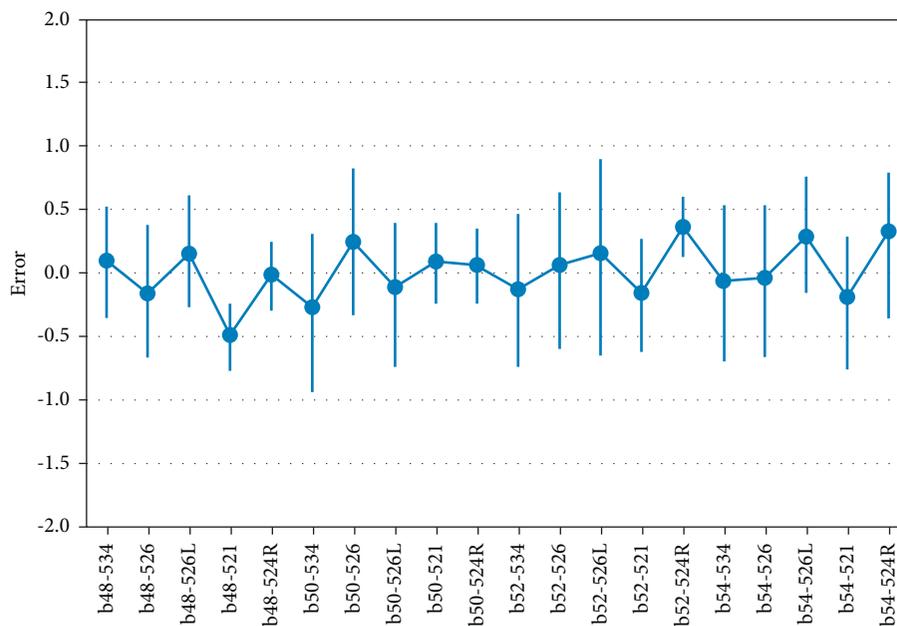


FIGURE 10: Error distribution of OD distance.

TABLE 7: Error index of OD distance.

Gate	MAE				RSME			
	b48	b50	b52	b54	b48	b50	b52	b54
521	0.6611	0.8489	0.7278	0.8766	1.0447	1.1118	1.1822	1.2811
524R	0.5761	0.707	0.7507	0.3767	0.7919	0.8516	0.9228	0.8409
526	0.7936	0.6212	0.8163	0.9594	0.8462	0.8449	1.1006	0.8715
526L	0.7636	0.7184	1.0703	0.8291	0.9939	0.8487	0.9494	1.1001
534	1.0132	1.0117	1.0679	1.2102	1.1882	1.3363	1.2691	1.3307

compare it with the corresponding distance in the distance matrix obtained by Floyd algorithm. In order to better measure this difference, this paper introduces absolute error (Error), mean absolute error (MAE), and root mean square error (RMSE) as evaluation indexes to test the accuracy of the distance matrix, as shown in equations (5) to (7).

$$\text{Error} = Y - \hat{Y}, \quad (5)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|, \quad (6)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}. \quad (7)$$

The error is shown in Figure 10, and the MAE and RMSE calculation results are shown in Table 7. The results show that the mean value of the distance error between the distance calculated by the road network and the GPS track is kept at about zero, and the deviation is kept at about 1 minute. Considering that there are many random factors in the actual driving of the ferry car, the current results can better reflect the actual driving situation and meet the needs of data accuracy for modeling.

6. Conclusion

This paper studies the ferry vehicles in the airport flight area. Based on the analysis and mining of ferry vehicle trajectory data, the ferry road network map is established, and a new method is proposed to extract the travel time between OD points in the ferry network. The experimental results show that the shape of the ferry road network constructed in this paper and the calculation results of travel time can reflect the real operation to a great extent. The purpose of this research paper is exactly to solve the problem of ferry vehicles' dynamic scheduling. With the travel time between OD points of the ferry network obtained in this paper, combined with the real-time GPS data of ferry vehicles, we can study how to predict the task state of vehicles and schedule vehicles dynamically and reasonably in the future.

Data Availability

The flight data and ferry vehicle GPS track data used in this paper are provided by Kunming Changshui Airport, China. The coordinates of each boarding gate and remote stand of the airport are obtained by Google satellite map.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (71621001) and an open fund of Key Laboratory of Airports Cluster Intelligent Operation under Grant no. KLAGIO20180801.

References

- [1] J. Bi, Z. Wu, L. Wang, D. Xie, and X. Zhao, "A tabu search-based algorithm for airport gate assignment: a case study in Kunming, China," *Journal of Advanced Transportation*, vol. 2020, pp. 1–13, 2020.
- [2] B. Liang, Y. Li, J. Bi, C. Ding, and X. Zhao, "An improved adaptive parallel genetic algorithm for the airport gate assignment problem," *Journal of Advanced Transportation*, vol. 2020, no. 4, pp. 1–17, 2020.
- [3] A. Norin, D. Yuan, T. A. Granberg, and P. Värbrand, "Scheduling de-icing vehicles within airport logistics: a heuristic algorithm and performance evaluation," *Journal of the Operational Research Society*, vol. 63, no. 8, pp. 1116–1125, 2012.
- [4] J. Y. Du, J. O. Brunner, and R. Kolisch, "Planning towing processes at airports more efficiently," *Transportation Research Part E: Logistics and Transportation Review*, vol. 70, pp. 293–304, 2014.
- [5] Z. Wang, Y. Li, X. Hei, and H. Meng, "Research on airport refueling vehicle scheduling problem based on greedy algorithm," in *Proceedings of the Intelligent Computing Theories and Application*, pp. 717–728, Wuhan, China, 2018.
- [6] S. Padrón, D. Guimarans, J. J. Ramos, and S. Fitouri-Trabelsi, "A bi-objective approach for scheduling ground-handling vehicles in airports," *Computers & Operations Research*, vol. 71, pp. 34–53, 2016.
- [7] C. Xu and Q. Shao, "Optimization of airport ground service support equipment scheduling under uncertain operation time," *Science Technology and Engineering*, vol. 018, no. 3, pp. 372–378, 2018.
- [8] T. Fei and L. Shu'An, "Research of aircraft ground service scheduling problems," in *Proceedings of the Control & Decision Conference*, May 2016.
- [9] P. Zhao, X. Han, and D. Wan, "Evaluation of the airport ferry vehicle scheduling based on network maximum flow model," *Omega*, vol. 3, pp. 35–47, 2019.
- [10] X. Han, P. Zhao, Q. Meng, S. Yin, and D. Wan, "Optimal scheduling of airport ferry vehicles based on capacity network," *Annals of Operations Research*, vol. 295, no. 1, pp. 163–182, 2020.

- [11] M. Rahmani and H. N. Koutsopoulos, "Path inference from sparse floating car data for urban networks," *Transportation Research Part C: Emerging Technologies*, vol. 30, pp. 41–54, 2013.
- [12] B. Y. Chen, H. Yuan, Q. Li, W. H. K. Lam, S.-L. Shaw, and K. Yan, "Map-matching algorithm for large-scale low-frequency floating car data," *International Journal of Geographical*, vol. 28, no. 1, pp. 22–38, 2014.
- [13] G. Draijer, N. Kalfs, and J. Perdok, "Global positioning system as data collection method for travel research," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1719, no. 1, pp. 147–153, 2000.
- [14] A. T. Palma, V. Bogorny, Kuijpers, and L. O. Alvares, "A clustering-based approach for discovering interesting places in trajectories," in *Proceedings of the 23rd Annual ACM-Symposium on Applied Computing*, pp. 863–868, Fortaleza Ceara, Brazil, 2008.
- [15] T. Bhattacharya, L. Kulik, and J. Bailey, "Extracting significant places from mobile user GPS trajectories: a bearing change based approach," in *Proceedings of the ACMSIGSPATIALGIS'12*, ACM, Redondo Beach, Calif, USA, 2012.
- [16] Y. Zheng and X. Xie, "Learning travel recommendations from user-generated GPS traces," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 1, 2011.
- [17] J. Li, Q. M. Qin, L. You, and C. Xie, "Parking lot extraction method based on floating car data," *Journal of Wuhan University (Information Science Edition)*, vol. 38, no. 5, pp. 599–603, 2013.
- [18] J. Wang, C. Dong, C. Shao, S. Huang, and S. Wang, "Identifying the key nodes and sections of urban roadway network based on GPS trajectory data," *Journal of Advanced Transportation*, vol. 2021, no. 1, pp. 1–14, 2021.
- [19] S. Schroedl, K. Wagstaff, S. Rogers, P. Langley, and C. Wilson, "Mining GPS traces for map refinement," *Data Mining and Knowledge Discovery*, vol. 9, no. 1, pp. 59–87, 2004.
- [20] J. Li, Q. Qin, J. Han, L.-A. Tang, and K. H. Lei, "Mining trajectory data and geotagged data in social media for road map inference," *Transactions in GIS*, vol. 19, no. 1, pp. 1–18, 2015.
- [21] L. L. Cao and J. Krumm, "From GPS traces to a routable road map," in *Proceedings of the Workshop on Advances in Geographic Information Systems*, pp. 3–12, New York, NY, USA, 2009.
- [22] L. Tang, X. Yang, Z. Kan, and Q. Li, "Lane-level road information mining from vehicle GPS trajectories based on naïve bayesian classification," *ISPRS International Journal of Geo-Information*, vol. 4, no. 4, pp. 2660–2680, 2015.
- [23] J. Li, Q. Li, Y. Zhu, Y. Ma, Y. Xu, and C. Xie, "An automatic extraction method of coach operation information from historical trajectory data," *Journal of Advanced Transportation*, vol. 2019, Article ID 3634942, 15 pages, 2019.
- [24] L. Dong, Y. Lv, H. Sun, D. Zhi, and T. Chen, "GPS trajectory-based spatio-temporal variations of traffic accessibility under public health emergency consideration," *Journal of Advanced Transportation*, vol. 2021, no. 7, 22 pages, Article ID 8854451, 2021.
- [25] P.-R. Lei, T.-J. Shen, W.-C. Peng, and I.-J. Su, "Exploring spatial-temporal trajectory model for location prediction," in *Proceedings of the 12th IEEE International Conference on Mobile Data Management (MDM '11)*, pp. 58–67, Lulea, Sweden, June 2011.
- [26] P. Zhao, X. Liu, W. Shi, T. Jia, W. Li, and M. Chen, "An empirical study on the intra-urban goods movement patterns using logistics big data," *International Journal of Geographical Information Science*, vol. 34, no. 10, pp. 1–28, 2018.
- [27] M. Lu, Z. C. Wang, and X. R. Yuan, "TrajRank: exploring travel behaviour on a route by trajectory ranking," in *Proceedings of the IEEE Pacific Visualization Symposium*, pp. 14–17, China, 2015.
- [28] J. Li, J. Wang, J. Zhang, Q. Qin, T. Jindal, and J. Han, "A probabilistic approach to detect mixed periodic patterns from moving object data," *GeoInformatica*, vol. 20, no. 4, pp. 715–739, 2016.
- [29] R. W. Floyd, "Algorithm 97: shortest path," *Communications of the ACM*, vol. 5, no. 6, p. 345, 1962.
- [30] E. W. Dijkstra, "A note on two problems in connexion with graphs[J]," *Numerische Mathematik*, vol. 1, no. 1, pp. 269–271, 1958.
- [31] F. D. Duan, "A faster algorithm for shortest-Path-SPFA," *Journal of Southwest Jiaotong University*, vol. 29, no. 2, pp. 207–211, 1994.