IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 1, 2019, accepted January 7, 2020, date of publication January 23, 2020, date of current version January 31, 2020.

*Digital Object Identifier 10.1109/ACCESS.2020.2969075*

# A Smart Path Recommendation Method for Metro Systems With Passenger Preferences

**WEI LI**[1,2], **QIN LUO**[1], **AND QING CAI**[3]
[1]College of Urban Transportation and Logistics, Shenzhen Technology University, Shenzhen 518060, China
[2]Key Laboratory of Optoelectronic Devices and Systems of Ministry of Education and Guangdong Province, College of Physics and Optoelectronic Engineering, Shenzhen University, Shenzhen 518060, China
[3]Department of Civil, Environment and Construction Engineering, University of Central Florida, Orlando, FL 32816, USA

Corresponding author: Qin Luo (465312799@qq.com)

This work was supported in part by the Research Projects of Natural Science Foundation of Guangdong Province under Grant 2018A030313119, and in part by the Ordinary University Innovation Project of Guangdong Province under Grant 2018KTSCX350.

**ABSTRACT** Passenger travel paths in metro networks have become more diversified with the development of network structures and the complexity of train schedules. Nowadays, passengers may have more than one alternative path in an OD (Origin-Destination) pair. In order to provide high-quality service to passengers, this paper proposes a smart path finding method to recommend fast and comfortable routes to passengers. By including the structure of the metro network as a two-dimensional plane and time as the third dimension, the space-time range of passenger activities is constructed. The accessible transfer stations for arrival are firstly computing by forward searching from O station, and the accessible transfer stations for departure are then identified by backward searching from D station. The intersection of the two feasible trajectories is regarded as the feasible path set of passengers. The path is recommended to passengers using generalized path impedance with the consideration of the situation that passengers have different perceptions of travel time, transfer penalty, and congestion tolerance. Finally, the proposed method is validated with the data of Shenzhen Metro under different traffic demand conditions, including the uncongested condition, the condition with congestion, the last train condition, and the condition with train delays, respectively. The time complexity of the proposed algorithm is also compared with that of some conventional algorithms. The results show that the proposed algorithm has lower algorithm complexity and is more suitable for the increasingly large-scale metro network.

**INDEX TERMS** Urban transport, metro, path finding, passenger preference, train schedule.

## I. INTRODUCTION

The metro system is a kind of transportation system with the advantages of large capacity, rapidness, punctuality and comfort in large cities. Meanwhile, the high-speed construction and development of the metro make the metro network more and more complex, diversifying the passengers' path choices in the network. In other words, there are more alternative paths for a passenger to choose within an OD pair in a large-scale network.

Meanwhile, passengers' travels in metro networks depend on the train movement. The train movement follows the train schedule, which has a strong periodic and programmatic. It should be noted that the train schedule is the basis of train operation in the metro system, which defines the arrival and departure time at the stations for each train. Therefore, the combination of network structure and train schedule leads the situation that the passengers' travel time, waiting time and passenger congestion are different over paths within the same OD pair, which makes the passenger path choice complex and diverse. Passengers might get confused with such many alternative paths when they travel by the metro. Hence, path finding and recommendation are needed to provide a high-quality service to passengers.

It is well known that path finding and recommendation are vitally important to the metro operation and management, such as passenger metro services, passenger flow distribution, metro tickets clearing, and network planning. For example, it can be used to calculate the congestion points of the network and formulate the corresponding passenger service. Besides,

The associate editor coordinating the review of this manuscript and approving it for publication was Zhigang Liu.

it can be applied to provide useful data to help enhance train schedules by analyzing the adaptability of train capacity. Moreover, it can provide high-quality passenger service for different kinds of passengers and recommend a fast and comfortable trip path guidance. Therefore, it is significantly necessary to study the path finding and recommendation algorithm.

In addition, the train delay also poses an urgent need for passenger path finding. Train delays are unavoidable due to the highly dense headways of metro trains and numerous involved systems. However, train delays will greatly affect passengers' metro travels. Passenger waiting time will increase with light delays, and even the potential safety of stations will be influenced by the continuous backlog of passengers flow with serious delays. Therefore, it is important to direct passengers to other paths according to the actual situation under the condition of train delays.

The objective of our study is to provide a path finding and recommendation method with the consideration of both time and space characteristics in passenger travels. The contributions of this paper are as follows.

1. A smart path finding algorithm is proposed to generate every possible path which passengers are likely to choose within a given OD pair. The time complexity of the proposed method is lower than some conventional methods, which implies that our method has great potential in practical transportation service systems in large scale networks with a shorter response time.

2. The path is recommended according to passenger preferences. Passengers may prefer different paths due to the perceptions of travel time, transfer penalty, and congestion tolerance over paths. The recommended method is more reasonable and efficient.

The remainder of this paper is organized as follows. The literature review is conducted in Section 2. Then a path finding algorithm is developed in Section 3. The complexity of the proposed algorithm is discussed in Section 4, followed by a path recommendation method for passengers in Section 5 and 6. Section 7 demonstrates a case analysis on the Shenzhen Metro network to validate the proposed method using different conditions. The last section draws conclusions and discusses some further research topics.

## II. LITERATURE REVIEW

Conventional methods for path finding in metro systems applied shortest path algorithm, deep-searching algorithm and K shortest path algorithm [1]–[4], in which K shortest algorithm was the most widely used. However, it should be noted that the algorithms for finding paths considered the spatial characteristics of the transit network, but failed to analyze its dynamic train schedule. To solve the problem, a transit assignment model and optimal path-finding algorithm for dynamic schedule-based transit networks were proposed in combination with time features, like the shortest travel path [5], [6] or K-shortest paths [7]–[10]. Specifically, Friedrich *et al.* [11] revised the model to trim paths using the branch and bound approach in order to improve the efficiency of the algorithm. According to the schedule-based model, Xu *et al.* [12] provided a two-stage labeling algorithm to search K shortest paths. Zhou and Xu [13] proposed an algorithm to obtain a possible path set by deleting infeasible paths from a valid passenger path set using the entry and exit time constraints. Guo and Jia [14] put forward a node label algorithm for finding the K shortest paths between two nodes. Temporal-arcs are put into the labels of nodes and arranged by ascending order. The algorithm is suitable for networks having parallel arcs with the same direction between two nodes.

The conventional way of K-shortest algorithm is to compute a sufficiently large number of overall shortest paths and delete the ones that do not satisfy the constraints. While considering time, the common practice of the above algorithms is to find all possible paths in the network first through K-shortest or other path algorithms. But the conventional methods might be time-consuming and inefficient when the number of paths is large or be inaccurate when the number of paths is small. Van Der Zijpp and Fiorenzo Catalano [15] proposed a new method to find the feasible shortest paths directly and can be applied in combination with a broad class of constraints. Li *et al.* [16] introduces space-time prism into path set generation to create a new path set generation algorithm considering both spatial and temporal characteristics of the metro system.

However, passengers prefer to different paths in the metro network. Some studies were conducted to recommend travel path for passengers. Wang *et al.* [17] pointed out three typical variants: the K-earliest arrival problem, the K-shortest travel time problem and the K-minimum number of transfers problem. Yang *et al.* [18] modified an excellent K-shortest paths algorithm proposed by Martins and Santos (MS) for the time-dependent model to solve the K-earliest arrival problem for the timetable information-based transportation system. Wang *et al.* [17] solved the problem by first modeling the timetable information with the time-expanded approach, then applying the MS algorithm. Jeon *et al.* [19] proposed an improved Schedule-based public transit routing algorithm that reflected transfer resistance and multi-path searching, which is applied at the time of transfer. Li *et al.* [20] analyzed passenger path choice based on smart card data and train schedule, and the result can be used to study passenger preference to different paths. Besides, Kim *et al.* [21] discussed the influence of congestion on the passenger route choice. The result showed that people reroute not only to avoid the delay from crowding but also to evade the large crowd.

Path finding and recommendation are especially important for last trains in metro networks. The last trains are the last chances for many passengers to travel by metro during late-night since metro trains will stop operating at midnight. Passengers need to obtain timely and effective path information to travel. Kang *et al.* [22]–[26] made great effects on last train transfers and last train schedule coordination. They studied the optimization of last train transfers to serve more

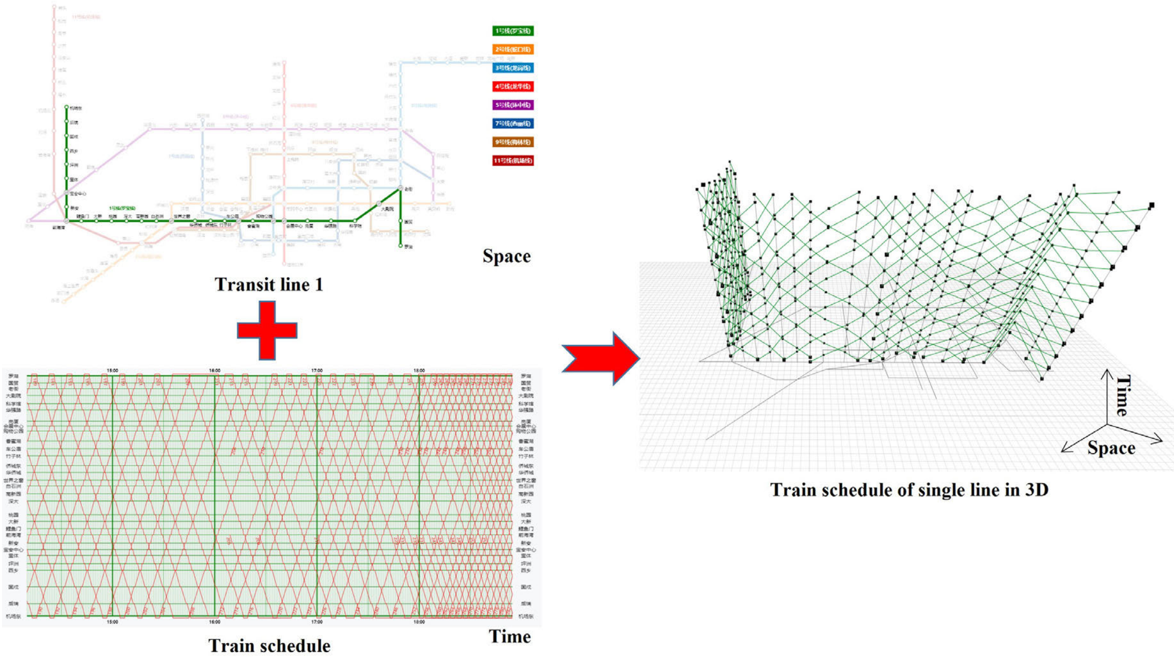



**FIGURE 1.** Train schedule of transit line 1 displayed in 3D.

passengers who can successfully transfer from one line to another. However, the problem of path finding under the last train condition is not investigated.

In conclusion, a new smart path finding algorithm and the corresponding recommendation method are needed considering the complexity of metro networks and the characteristics of train schedules. The proposed method should include the following three features:

1) The physical network structure of the metro system and the dynamic train schedule should be considered.

2) It can compute all the feasible paths fast and efficiently in a short time, especially for the path finding under the condition of train delays.

3) It is necessary to provide metro paths according to passenger preferences with the consideration of the diversification of passenger behavior.

## III. METHODOLOGY OF PATH FINDING

Different from other transport modes, the metro system is composed of a ''static'' network structure and a ''dynamic'' train schedule.

**Figure 1** shows the connection between the transit line and the train schedule. The train schedule is the diagram to display the relationship of train movements. By regarding the structure of the metro network as a two-dimensional plane and the time as the third dimension, we could obtain the train schedule displayed in 3D.

The network structure is the physical topology of the metro network, which provides the prerequisite for whether passengers can finish their trips successfully within OD pairs. Moreover, passengers traveling on the network also depend on train movement, which is constrained by the ''train schedule''. **Figure 2** shows the connection among network structure, train schedule, and passengers' trajectories. The space-time range of passenger activities can be constructed by the constraints of train schedules and the network structure. Thus passengers' trajectories in metro networks are three-dimensional space-time trajectories. All of the passengers' possible trajectories can be displayed in the 3D space-time range, as shown in **Figure 2**.

Hence, a smart path finding algorithm is proposed to search the possible passengers' trajectories. The purpose of the path finding algorithm is to determine the possible paths which passengers would choose to ride within a given OD pair. We forward search passengers' arrival trajectories from O station to obtain the accessible transfer stations for arrival. Then the accessible transfer stations for departure are found by backward searching from D station. Finally, we intersect the two feasible trajectories and the intersection is the feasible path set of passengers.

### A. BASIC ASSUMPTIONS

To facilitate the presentation of the essential ideas without the loss of generality, the following basic assumptions are made in this paper:

1. All trains run accurately according to the train schedule.
2. All Passengers interchange at transfer stations with the average transfer time.
3. Passengers will not alight the trains until they arrive at the destinations or the transfer stations.

### B. TRAIN SCHEDULE SEARCHING

In this subsection, some basic definitions are outlined about the train schedules of the metro. The train schedule is a

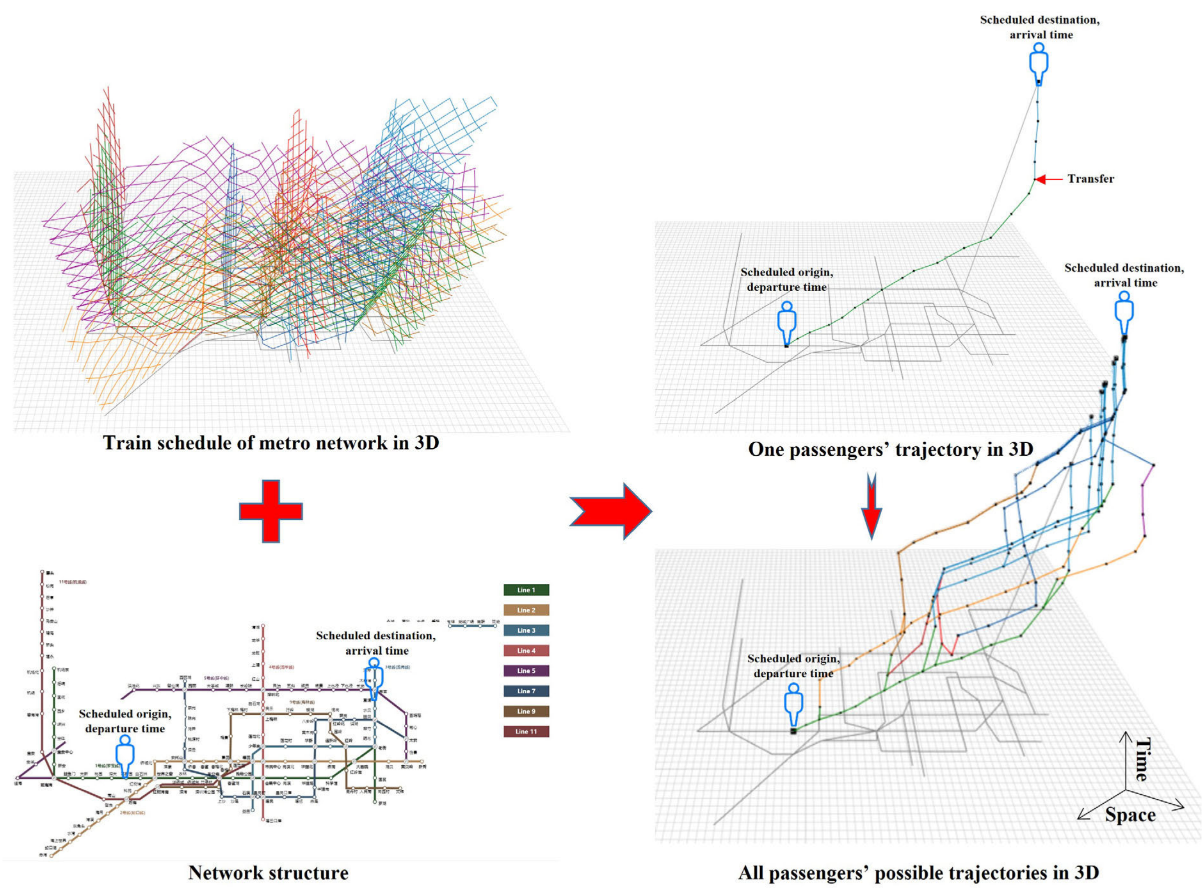



**FIGURE 2.** **Passengers' trajectories with the network structure and train schedule in metro systems.**

schedule that illustrates the relationship between space and time of train operation. The main messages it contains are the arrival and departure time of each station for all trains. Denote the set of metro lines as $L = \{1, 2, \ldots, l, \ldots, N\}$, and the set of stations in line $l$ as $S_l = \{1, 2, \ldots, i, \ldots, M\}$. Here, $S_{l,i}$ means station $i$ in line $l$, the arrival time $A^j_{l,i}$ and departure time $D^j_{l,i}$ of train $j$ at station $S_{l,i}$ can be described as $S_{l,i}\left(A^{(j)}_{l,i}, D^{(j)}_{l,i}\right)$. The train movement trajectory is defined as the collection $\left\{\forall i \in l | S_{l,i}\left(A^{(j)}_{l,i}, D^{(j)}_{l,i}\right)\right\}$. For each line, each station, each train, the train schedule is represented by $T = \left\{\forall j, l, i | S_{l,i}\left(A^{(j)}_{l,i}, D^{(j)}_{l,i}\right)\right\}$.

The definition of the train schedule provides the necessary foundation for generating the feasible arrival/departure transfer stations. Then, it is necessary to figure out which train to ride in the train schedule for passengers' trips. The method to forward search the train schedule describes as below.

Let the begin time be $t$, and a passenger is supposed to locate in the original station or transfer station. For each train in line $l$, the train $j$ that the passenger can board at station $S_{l,i}$ at time $t$ can be determined based on Equation 1. Thus, the passenger can choose train $j$ or next following trains to arrive at the destination or transfer station. The searching process is illustrated in **Figure 3**.

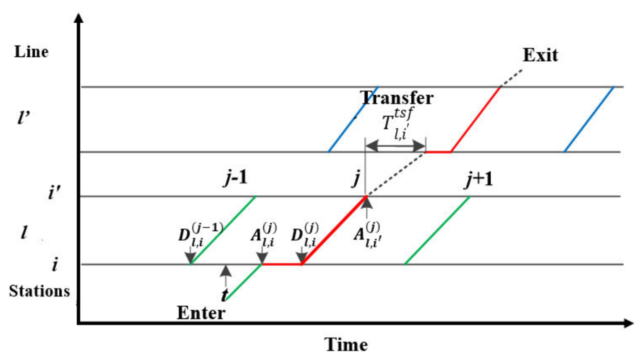



**FIGURE 3.** **Illustration of train schedule searching.**

$$D^{(j-1)}_{l,i} \leq t \leq D^{(j)}_{l,i} \tag{1}$$

Denote $S_{l,i'}$ as the passenger's destination or transfer station in line $l$. The arrival time $A^{(i)}_{l,i}$ of station $S_{l,i'}$ can be obtained from the schedule of train $j$. If a passenger arrives at the destination, he/she will end this trip and exit the station. Otherwise, the passenger will interchange to other lines to continue his/her trip. Note that transfer time should be added to the passenger's travel when station $S_{l,i'}$ is a transfer station in the path. Denote $T^{\text{tsf}}_{l,i'}$ as the transfer time of station $S_{l,i'}$. Then, the begin time is then updated as

$$t = A^{(j)}_{l,i'} + T^{\text{tsf}}_{l,i'} \tag{2}$$

In the same way to the backward searching train schedule, let the end time be $t'$. The train $j$ that the passenger alight at station $S_{l,i'}$ can be determined by Equation 3. Then end time can be updated as Equation 4.

$$A^{(j-1)}_{l,i} \leq t' \leq A^{(i)}_{l,i} \tag{3}$$

$$t' = D^{(j)}_{l,i} - T^{\text{tsf}}_{l,i} \tag{4}$$

### C. PATH FINDING ALGORITHM

As mentioned in the previous section, passenger travel is subject to both network structure and train schedule. The arrival/departure transfer stations can be generated through the method in **Section 3.B**. It is reasonable to generate an accessible path set under the constraints based on the space-time range, as shown in **Figure 4**. The path set generation algorithm is described below:

A passenger is assumed to travel from $O$ station to $D$ station with an expected departure time $T_{\text{start}}$ and a planned arrival time $T_{\text{end}}$.

**Step** 1: Start with forward searching the train schedule from the origin $O$ at time $T_{\text{start}}$ to obtain the trajectory by calling the function of train schedule searching. The accessible transfer stations for arrival before planned arrival time can be constituted from the trajectories, which is denoted by $R_o$ in **Figure 4a**.

**Step** 2: Backward search the train schedule from the destination $D$ at time $T_{\text{end}}$ to obtain his trajectories by calling

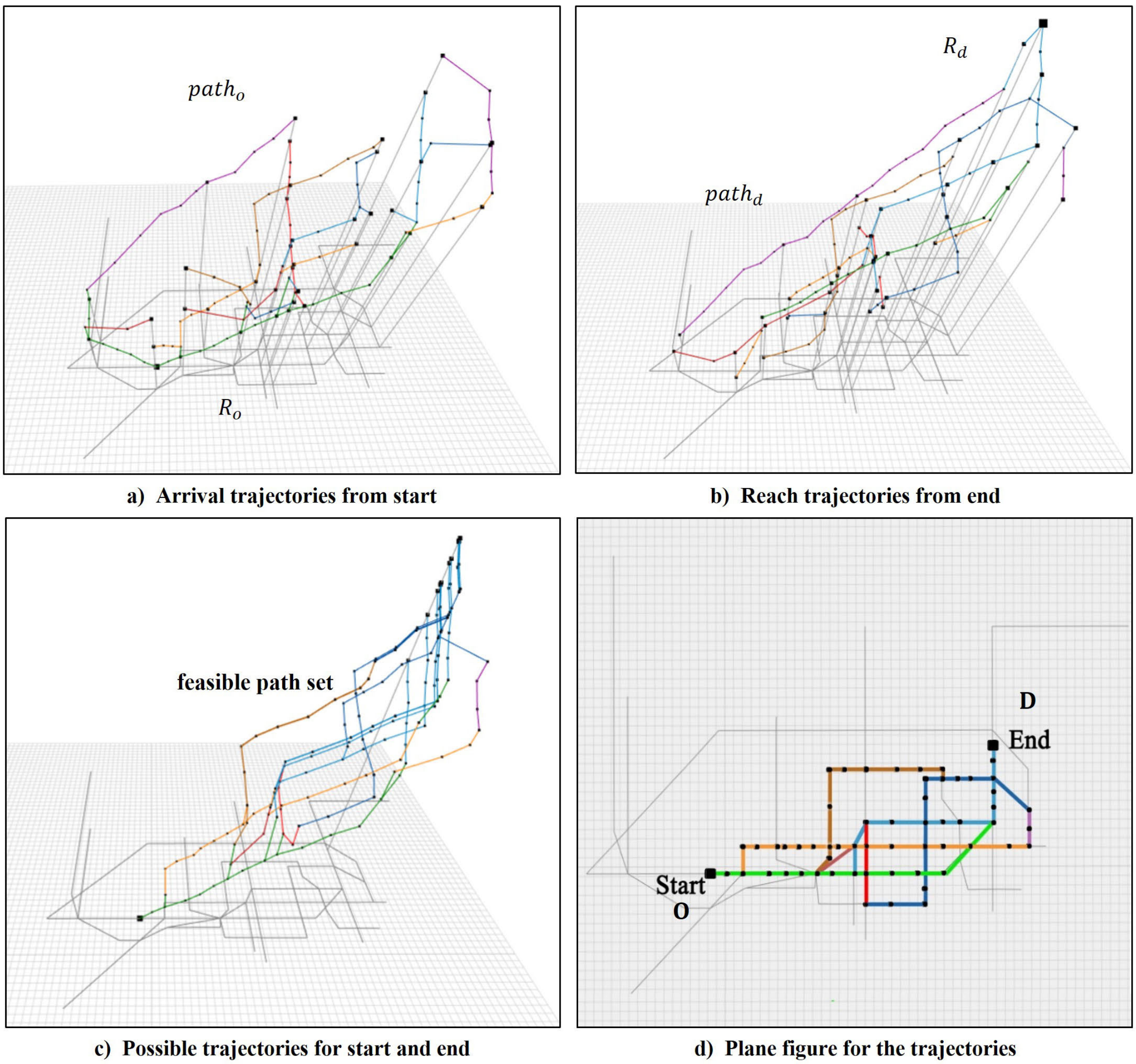



**FIGURE 4.** Illustration of path searching algorithm.

the function of reversed train schedule searching. Reversed train schedule searching is to search the schedule backward through time decreasing. While normal searching is to find accessible arrival transfer stations from origin $O$ at $T_{\text{start}}$, reversed searching is useful for finding feasible departure transfer stations from destination $D$ at $T_{\text{end}}$. Denote $R_d$ as the feasible departure transfer stations in **Figure 4b**.

**Step** 3: For each path in $R_o$, defined as path$_o$, calculate the earliest time at transfer stations the passenger can arrive from origin $O$. Similarly, for each path path$_d$ in $R_d$, calculate the latest departure time of transfer stations. Let $S_{l,i}$ be the transfer station exists both in path$_o$ and path$_d$. If its arrival time in path$_o$ is earlier than or equal to the departure time in path$_d$ minus an average transfer time, the path combining path$_o$ with path$_d$ is regarded as a feasible path, as shown in **Figure 4c**. The feasible path set is eventually generated after all paths in $R_o$ and $R_d$ are checked, as shown in **Figure 4d**.

**Step** 4: In order to obtain a practical path set, a threshold of trip time is set. The paths in the path set whose trip time exceeds the threshold are considered impossible and deleted from the set. Passengers are unlikely to choose them. The threshold can be got from a survey.

It should be noted that no additional modification is needed for the algorithm when to deal with train delays. The train movement in train schedules is set to prolong the section running time or to interrupt the train operation with the condition of train delays. Then, the path finding algorithm can be realized with train delays.

## IV. COMPLEXITY ANALYSIS OF THE ALGORITHM

In this section, we discuss the time complexity of the proposed algorithms. Denote $K$ as demanded path number, $m$ as the maximum number of links in a path, and $n$ as the number of transfer stations in the metro network. The value of $m$ is usually larger than that of $n$ in metro systems. Also, denote $P$ as the number of train trajectories that need to be checked. For the proposed algorithm, assume that the searching transfer station number is $n$ in a bad situation in the three steps. Step 1 is to search forward the network and the train schedule to obtain feasible arrival trajectories, and the complexity of step 1 is $O(n \log P)$. In the same way, the complexity of step 2 is $O(n \log P)$. In step 3, two feasible trajectories are intersected to form the feasible path set of passengers, and the complexity of step 3 is $O(Kn)$. Hence, the whole algorithm is $O(Kn + 2n \log P)$.

For the schedule-based K shortest algorithm [13], the first stage is to carry on the K-shortest algorithm. Then, the time complexity of the K-shortest algorithm is $O(Kmn \log n)$. The second stage is to check path validity according to the train schedule. Its time complexity is $O(Kn \log P)$. Therefore, the time complexity of schedule-based K shortest algorithm is $O(Kmn \log n + Kn \log P)$.

We compared the algorithm in this paper with two classical algorithms in past research. The algorithm in this paper is named as algorithm A. The schedule-based K shortest algorithm [13] is named as algorithm B. The algorithm proposed by Guo and Jia [14] is named as algorithm C. The comparison of time complexity is shown in **Table 1**. Compared to those time complexity, we can find that the algorithm efficiency of our algorithm is better than that of the conventional methods, especially for large scale networks. The result implies the great potential of the proposed algorithm in practical transportation service systems.

**TABLE 1. Comparison on time complexity of algorithms.**

| Algorithm | Time complexity |
|---|---|
| Algorithm A (proposed method ) | $O(Kn+2n\log P)$ |
| Algorithm B [13] | $O(Kmn\log n+Kn\log P)$ |
| Algorithm C [14] | $O(m\log P+2Km+2Kn\log Kn)$ |

## V. PATH CALCULATION WITH PASSENGER PREFERENCES

Because of the individual preference, passengers have different perceptions of travel time, transfer penalty and congestion tolerance, which results in preferences for different paths. In order to figure out the influence factors of passenger path preference, a survey was conducted by Shenzhen metro company (time: Oct 26 - Nov 8, 2017, location: Shenzhen Metro, sample size: 8,222, valid sample 6,205), which is shown in **Table 2**. The survey result in **Figure 5** shows that the main factors influencing passenger path choice are travel time (35.99%), congestion (25.92%), transfer coordination (including transfers and transfer time 24.87%) and stations (13.22%). It can be seen that the travel time have the biggest influence on passenger path choice over a whole day. Congestion has a little more influence on passengers in the flat period than that in the peak period.

**TABLE 2. Survey information about passenger path preference.**

| | | Sample Size | | |
|---|---|---|---|---|
| | Categories | Number | Sum | Total |
| Weekday | Morning peak | 1831 | 6136 | 8222 |
| | Flat period | 2463 | | |
| | Night peak | 1842 | | |
| Weekend | Peak period | 1253 | 2086 | |
| | Flat period | 833 | | |

Therefore, metro paths need to be planned for different situations when to recommend passengers' paths. For the factors affecting passenger path choice, the generalized path impedance is applied in the route recommendation. After the feasible path set is obtained in the previous section, the path is sorted by assigning the generalized path impedance, and the relevant path is recommended according to passenger preferences. Three kinds of paths are selected as the path recommendation to passengers.

1) Travel time cost: the travel time is used as the generalized path impedance, including the in-vehicle time, waiting time and transfer time, as shown in Equation 5. The path

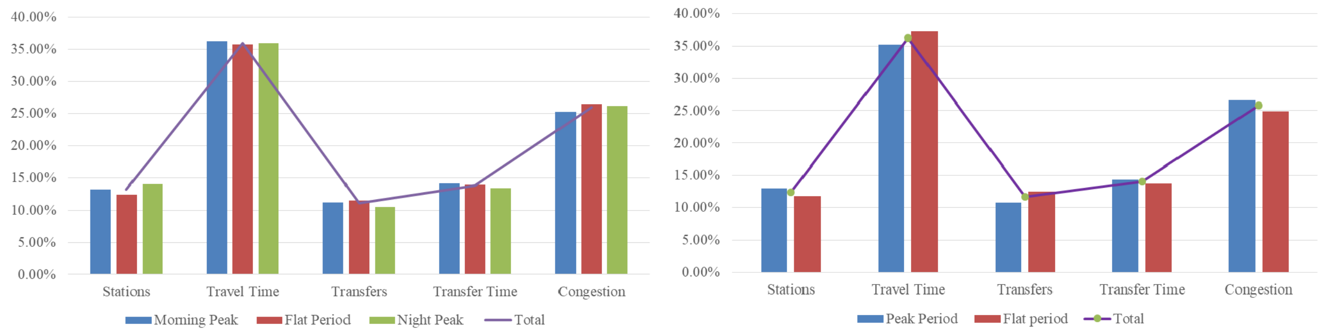



**FIGURE 5.** Passengers' preference to choose paths on weekday and weekend.

**TABLE 3.** Survey samples about passenger path choices.

| O | D | Choice 1 | | | | Choice 2 | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Transfer Time | Transfers | Travel Time | Sample Size | Transfer Time | Transfers | Travel Time | Sample Size |
| BX | LJ | 10.5 | 2 | 27.44 | 26 | 8 | 1 | 31.42 | 9 |
| CGM | HQB | 15.8 | 2 | 31.03 | 10 | 12.9 | 2 | 27.55 | 13 |
| FM | LJ | 7.9 | 1 | 26.58 | 1 | 4.2 | 1 | 19.41 | 14 |
| GX | HQB | 7.3 | 1 | 21.58 | 3 | 12.9 | 2 | 21.03 | 8 |
| GX | JT | 12.9 | 2 | 23.21 | 7 | 15.8 | 2 | 25.53 | 7 |
| GX | LJ | 12.9 | 2 | 19.84 | 10 | 15.8 | 2 | 22.15 | 2 |
| HQB | CGM | 13 | 2 | 27.65 | 8 | 15.5 | 2 | 30.73 | 7 |
| | | | | …… | | | | | |

with the shortest travel time is selected as the alternative path, expressed as $\underset{j}{\text{argmin}}\, T_j^{\text{cost}}, \forall j$.

$$T_j^{\text{cost}} = T_j^{\text{iv}} + T_j^{\text{wt}} + \sum_i^n T_{j,i}^{\text{tsf}} \tag{5}$$

where, $T_j^{\text{cost}}$ is the total travel time in path $j$ within the given OD pair; $T_j^{\text{iv}}$ is the in-vehicle time for passenger travel; $T_j^{\text{wt}}$ is the waiting time for passengers at the origin station; $T_{j,i}^{\text{tsf}}$ is the transfer time for passengers at the transfer station $i$; $n$ is the total number of transfer stations in path $j$; the symbol of argmin means to find $j$ to minimize the value $T_j^{\text{cost}}$ among all paths within the given OD pair.

2) Transfer coordination cost: As an indispensable part of metro travels, the transfer has a great impact on passengers' journeys. The influence on passenger path choices on transfer contains transfers and transfer time. Some passengers prefer less transfer or convenient (shorter transfer time) travel path. The transfer coordination cost is shown in Equation 6. The path with the minimum transfer coordination cost is selected as the alternative path, expressed as $\underset{j}{\text{argmin}}\, T_j^{\text{cost\_tsf}}, \forall j$.

$$T_j^{\text{cost\_tsf}} = \sum_i^n T_{j,i}^{\text{tsf}} * \alpha_i \tag{6}$$

where $T_j^{\text{cost\_tsf}}$ is the generalized transfer cost in path $j$ within the given OD pair; $\alpha_i$ is the transfer penalty coefficient, which increases geometrically with the increase of transfer times. The value can be obtained by the questionnaire survey.

**Table 3** displays the survey samples about passenger path choices. We calculate the sample size among different path choices during the same OD pairs. It is believed that the generalized travel cost between path choice 1 and 2 are regarded as the same by the passengers who choose them. The generalized travel cost can be expressed as

$$U(j) = T_j^{\text{cost}} - \sum_i^n T_{j,i}^{\text{tsf}} + \sum_i^n T_{j,i}^{\text{tsf}} * \alpha_i + \varepsilon \tag{7}$$

where $\varepsilon$ is a random error.

Therefore, the utility function is displayed as

$$S_j = e^{-U(j)} \tag{8}$$

$$P_j = \frac{S_j}{\sum S_j} \tag{9}$$

Here, $P_j$ is the selection probability for passengers to choose path $j$, which can be calculated by the information in Table 3. For example, $P_1 = \frac{26}{26+9} = 74.3\%$ in the first OD pair (BX-LJ).

Thus, the transfer penalty coefficient, $\alpha_i$, can be reflected by the survey according to the maximum likelihood estimation method.

$$\alpha_1 = 1.53;\ \alpha_2 = 1.79;\ \alpha_3 = 2.02 \tag{10}$$

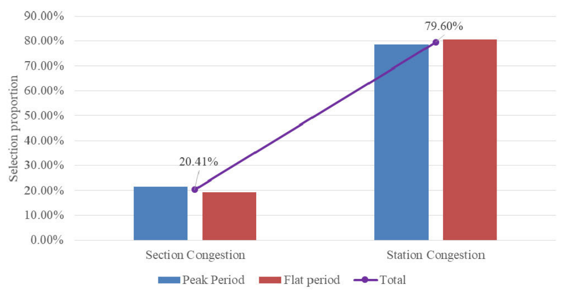



**FIGURE 6.** Passengers' preference to congestion between stations and sections.

where, the transfer penalty coefficient are 1.53 and 1.79 when passengers have one and two transfers, respectively, and it is 2.02 when passengers have three or more transfers.

3) Congestion cost: Some passengers are not sensitive to travel time; on the contrary, they are more concerned about comfort in their travels. They prefer the path with less crowded section/train passenger flow (section congestion), or the path with less crowded station passenger flow in the waiting or transfer process (station congestion). Here, the station congestion has a greater impact on passengers and will directly affect them. It will directly affect whether passengers can board the oncoming train while waiting at the origin stations or transfer stations, otherwise, they have to wait for the next available trains as the crowned passengers. The congestion cost is shown in Equation 11. The path with the minimum congestion cost is selected as the alternative path, expressed as $\underset{j}{\operatorname{argmin}} T_j^{\text{cost\_cgt}}$, $\forall j$.

$$T_j^{\text{cost\_cgt}} = \max_i^n \frac{F_{j,i}^{\text{sta}}}{C_{j,i}^{\text{sta}}} * \beta + \max_s^S \frac{F_{j,s}^{\text{sec}}}{C_{j,s}^{\text{sec}}} \tag{11}$$

where, $T_j^{\text{cost\_cgt}}$ is the generalized congestion cost in path $j$ within the given OD pair; $F_{j,i}^{\text{sta}}$ is the station passenger flow of transfer station $i$ in path $j$; $C_{j,i}^{\text{sta}}$ is the passenger capacity of transfer station $i$; $F_{j,s}^{\text{sec}}$ is the section passenger flow of section $s$ in path $j$; $C_{j,s}^{\text{sec}}$ is the passenger capacity of section $s$; $\beta$ is the magnification factor of congestion, which enlarges the impact of station congestion for passengers. The value is set as $\beta = \frac{79.60\%}{20.41\%} = 3.9$, which is obtained by the questionnaire survey, as shown in **Figure 6**.

## VI. PATH RECOMMENDATION

There are two methods to deal with the path recommendation when the proposed method is applied in phone and web applications.

1) Three paths with different preferences are presented to passengers after they select the origin, destination and scheduled time in the applications, as shown in Table 4. Passengers can choose the path that best meets their requirements.

2) The application records passengers' path choices once they have a trip. This information can form the unique travel preference of passengers, and it can be used in the following travels.

It is noted that the historical travel data (inbound and outbound swiping card data) acquired from ACC (AFC (Automatic Fare Collection) Clearing Center) data can be linked to the specific passengers with smart card ID. These data can be used to analyze passenger travel behavior with a long-range of date. However, there are several problems need to be solved, such as how to identify passengers travel path with just inbound and outbound swiping card data.

**TABLE 4.** A sample for path recommendation.

| Path No | Travel time cost | Transfer coordination cost | Congestion cost | Mark |
|---|---|---|---|---|
| 1 | 45 | 10.5 | 5.1 | Shortest travel time path |
| 2 | 48 | 6.2 | 5.5 | Less transfer penalty path |
| 3 | 51 | 9.3 | 3.6 | Less congestion path |

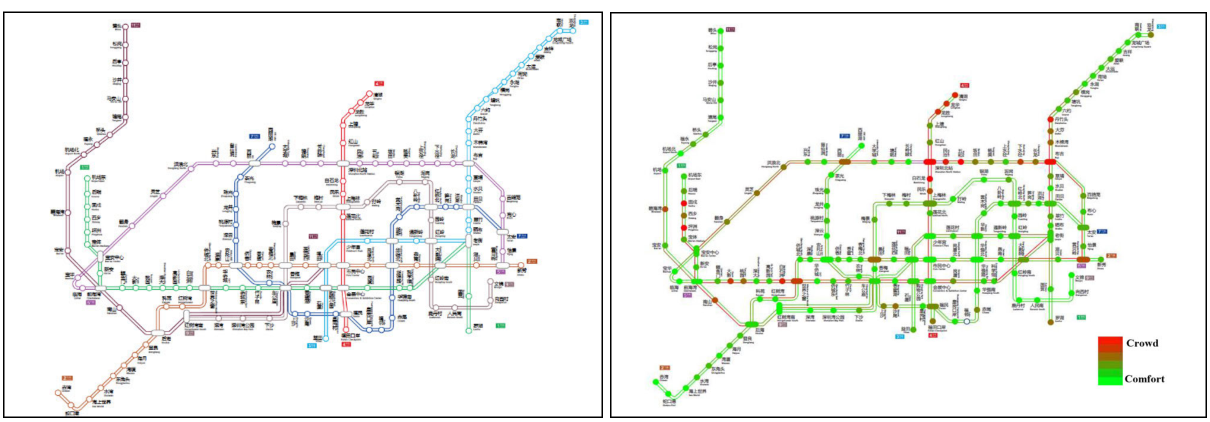



**FIGURE 7.** Network structure of Shenzhen metro.

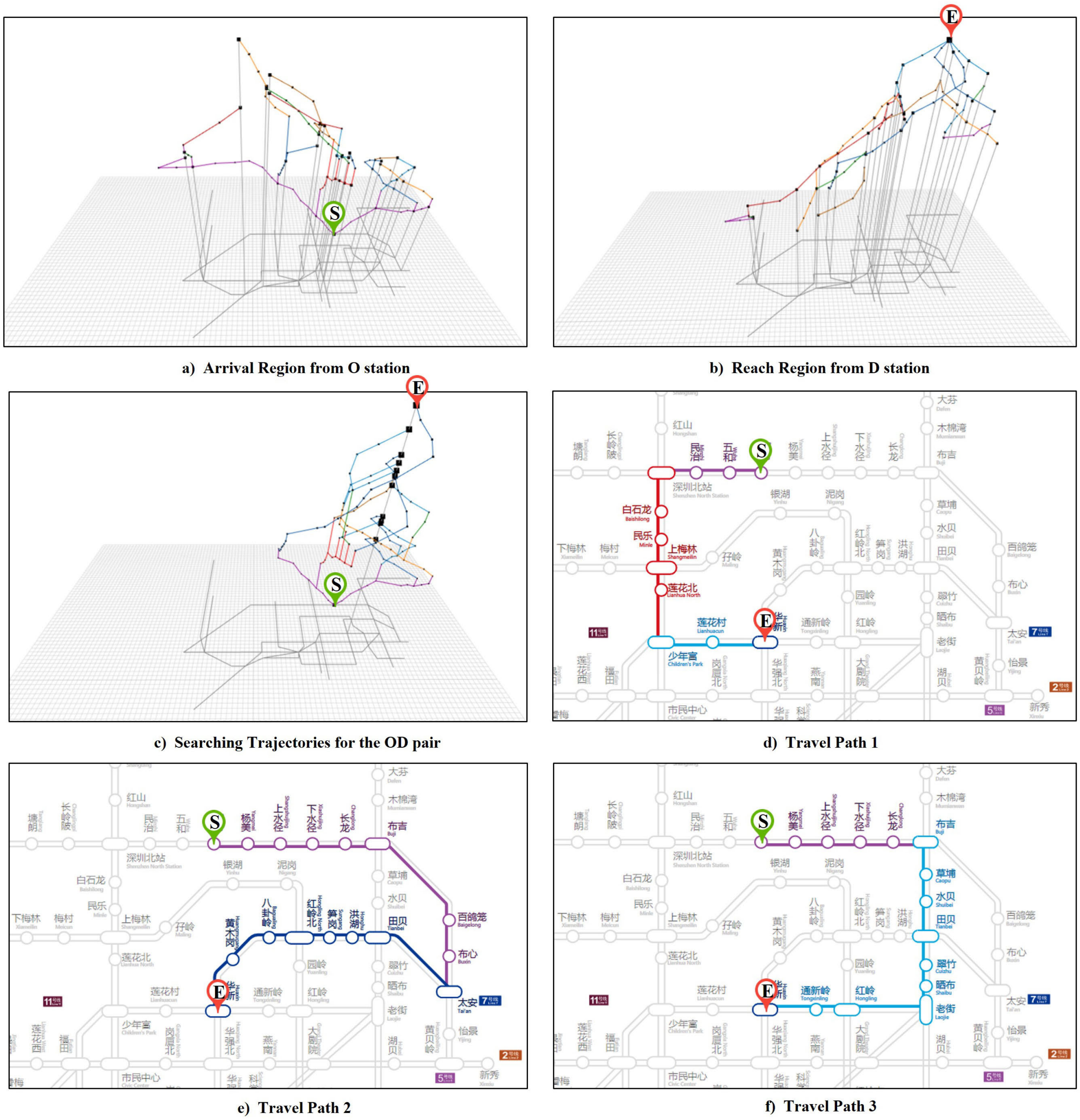

a) Arrival Region from O station b) Reach Region from D station c) Searching Trajectories for the OD pair d) Travel Path 1 e) Travel Path 2 f) Travel Path 3

**FIGURE 8.** Diagram of the searching paths for the OD pair.

## VII. CASE STUDY

To evaluate the proposed algorithm, a real-life metro network (Shenzhen metro) with a large number of lines and stations is presented as a case study application. Shenzhen metro network is composed of a total of 8 transit lines with 167 stations (including 28 transfer stations) by July 2019, as shown in the left part of **Figure 7**. In addition, the passenger flow is added to the stations and sections of the network to show the congestion, as shown in the right part of **Figure 7**. The source of passenger flow information used here is from the department of ACC (AFC (Automatic Fare Collection) Clearing Center) in Shenzhen Metro.

### A. NORMAL CONDITION

We first used the proposed algorithm to analyze the normal condition. Bantian and Huaxin stations are selected as O and D stations, where O station is marked with the green marker and D station is marked with the red marker. **Figure 8** shows the solving process and results of the algorithm. **Figure 8a-8c** shows the potential trajectories of passengers in the space-time range, which are generated by intersecting the trajectories forward searching from O and the trajectories backward searching from D. **Figure 8d-8f** shows the three most likely travel paths derived from the algorithm. These solved paths are consistent with the schedule-based K shortest algorithm [13], [14].

**Figure 8d**: Line 5 (purple) – Line 4 (red) – Line 3 (cyan), a total of 10 stations, 2 transfers, 33.5min of time cost, the shortest path;

**Figure 8e**: Line 5 (purple) – Line 7 (blue), a total of 15 stations, 1 transfer, 35.1min of time cost, the most transfer convenient path;

**Figure 8f**: Line 5 (purple) –Line 3 (cyan), a total of 14 stations, 1 transfer, 39.4min of time cost.

According to the proposed path recommendation method, it can be concluded that **Figure 8d** (travel path 1) holds the shortest travel time, and **Figure 8e**(travel path 2) is the most convenient transfer path. The transfer in **Figure 8f** holds longer transfer time for passengers even it has only one transfer.

### *B. NORMAL CONDITION WITH CONGESTION*

After adding the passenger flow of stations and sections, the path preference with congestion information for passengers is more obvious. **Figure 9a** (travel path 1, Line 5 to Line 4) and **Figure 9c** (travel path 3, Line 5 to Line 3) are crowded, especially at the transfer stations. As a contrast, **Figure 9b** (travel path 2, Line 5 to Line 7) is more comfortable with no congestion in its transfer station. Thus, **Figure 9b** (travel path 2) is recommended with congestion preference.

**Figure 9a**: Line 5 – Line 4 – Line 3, the stations and sections of Line 4 and Line 5 are crowded;

**Figure 9b**: Line 5 – Line 7, some sections of Line 5 are crowded, the impact on passengers is relatively small;

**Figure 9c**: Line 5 – Line 3, the transfer station from Line 5 to Line 3 is crowded, and it is not convenient for passengers to transfer.

In conclusion, when passengers prefer the shortest travel time, travel path 1 is recommended; when they prefer to the most convenient transfer, travel path 2 is recommended; and when they prefer to the least congestion, travel path 2 is recommended.

### *C. LAST TRAIN CONDITION*

In addition to the path finding under the normal condition, the proposed method can also be applied to the path finding under the last train condition. Because of the limit of the train schedule, metro trains will stop operation at midnight in Shenzhen Metro, leading there be a last boarding time for passengers to finish their trips. The last boarding time for each path may vary dramatically. It is often found that the common path with shorter travel time has stopped operation, while the path with longer travel time is still operated. This information is beneficial for passengers to travel by the metro at midnight.

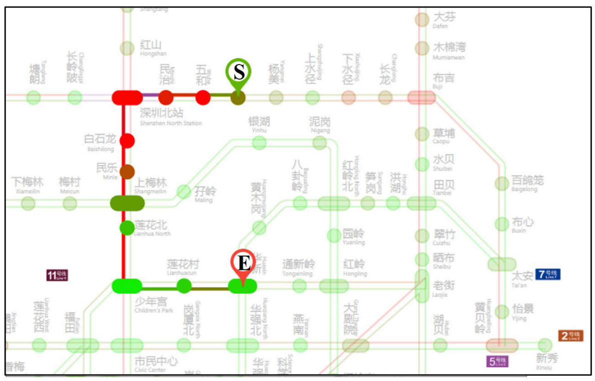
Longjin and Hongshuwan South stations are selected as O and D points. Two paths are calculated according to the proposed algorithm. Travel path 1 has a shorter travel time with the last boarding time of 23:09. On the contrary, travel path 2 has longer travel time and two transfers with the last boarding time of 23:29. There is a difference of 20 minutes between the two paths. In other words, passengers can travel 20 minutes later in path 2 than that in path 1.

**Figure 10a**: Line 7 (blue) – Line 11 (dark red), 1 transfer, 22.1minutes of time cost, 23:09 of the last boarding time;

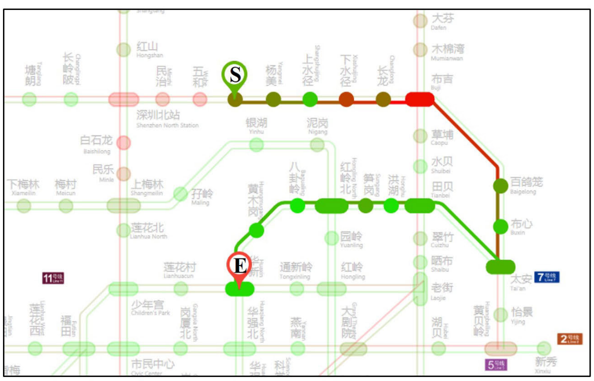
**Figure 10b**: Line 7 (blue) – Line 5 (purple) - Line 11 (dark red), 2 transfers, 39.3 minutes of time cost, 23:29 of the last boarding time.

**Generalized cost for each travel path**

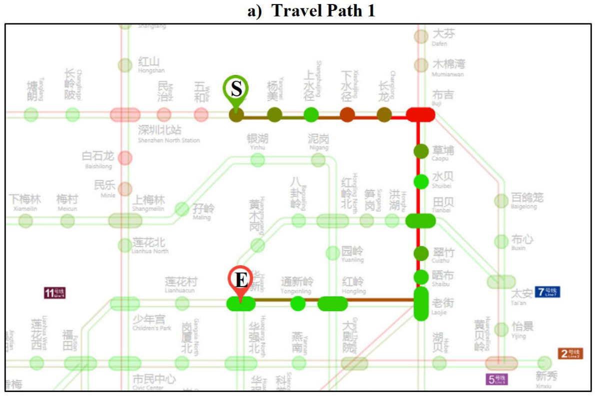
| BT-HX | Pass Lines | Stations | Travel time cost | Transfer coordination cost | Congestion cost |
|---|---|---|---|---|---|
| Path 1 | 5-4-3 | 10 | 33.50 | 12.97 | 5.18 |
| Path 2 | 5-7 | 15 | 35.10 | 2.30 | 1.66 |
| Path 3 | 5-3 | 14 | 39.40 | 5.20 | 4.76 |



a) Travel Path 1

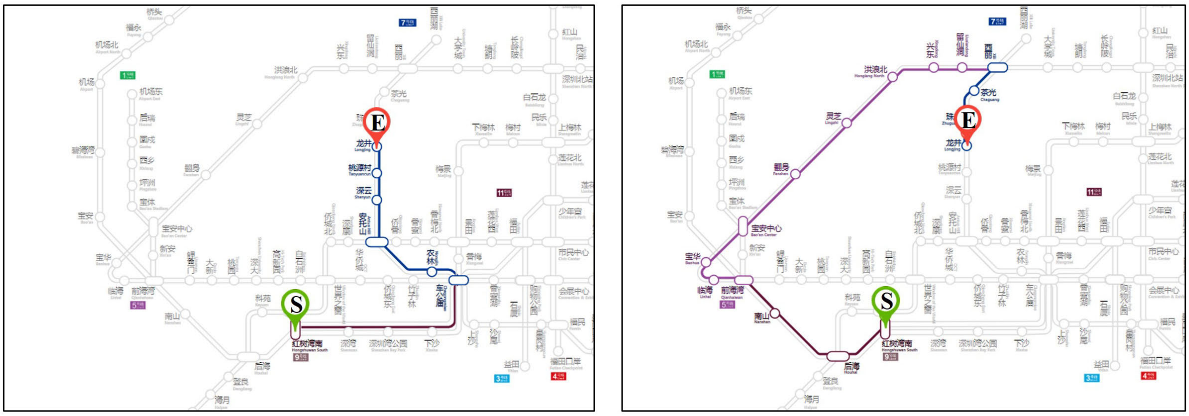


b) Travel Path 2



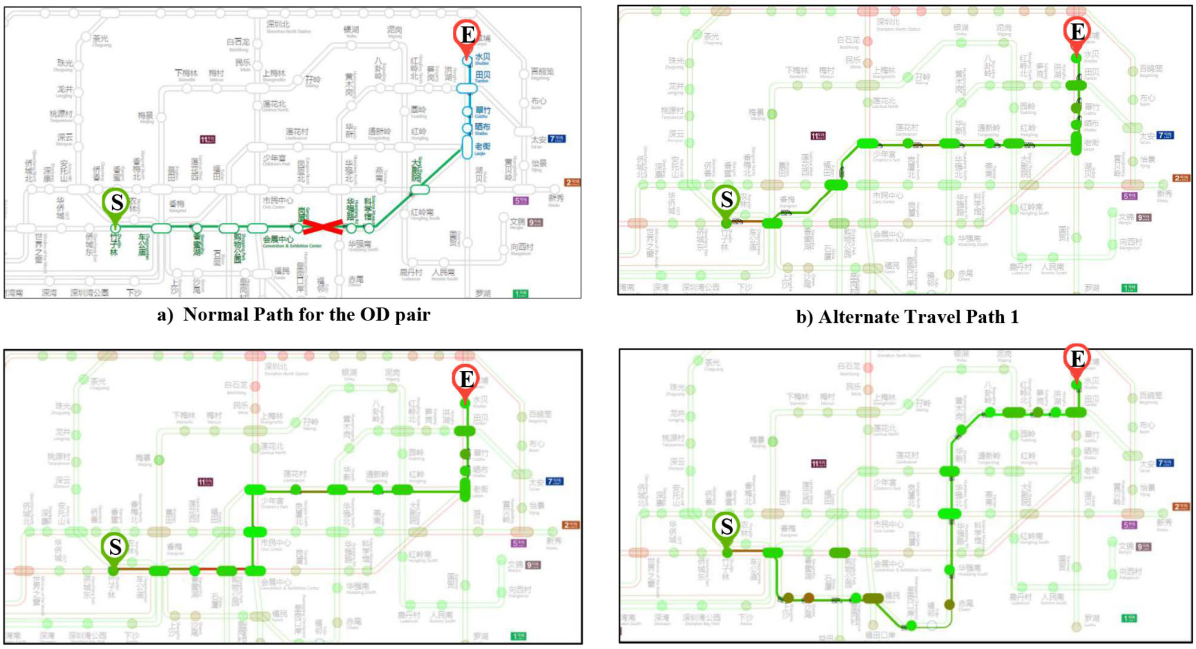
c) Travel Path 3

**FIGURE 9.** Diagram of the searching paths with congestion.



a) Travel Path 1 b) Travel Path 2

**FIGURE 10.** Diagram of the searching paths for the last trains.



a) Normal Path for the OD pair b) Alternate Travel Path 1

c) Alternate Travel Path 2 d) Alternate Travel Path 3

**FIGURE 11.** Diagram of the finding paths with train delays.

### D. ABNORMAL CONDITION WITH TRAIN DELAYS

The proposed path finding algorithm can also be applied to the abnormal condition with train delays. The alternative paths can be quickly determined by setting a longer section running time or setting an interruption in train operation in the train schedule.

Zuzilin and Shuibei stations were selected as O and D points. The section from Gangxia to Huaqiang Road is assumed to occur delay or interruption, as shown in the red fork in **Figure 11a**. **Figure 11** shows the normal path and alternate paths given by the proposed algorithm for the OD pair. **Figure 11a** is the path that passengers will choose under normal conditions. **Figure 11b-11d** are the alternative paths given by the proposed algorithm, which also accounts for the passenger flow of both sections and stations. The travel path given in **Figure 11c** is crowded in the middle of the route due to the delay of Line 1. Some stations in the path of **Figure 11d** is also crowded. Therefore, the better alternative path is the path given in **Figure 11b**.

## VIII. CONCLUSION

Focused on the increasingly complex metro system, this paper put forward a new path finding and recommendation method. The accessible transfer stations are forward searching from O station, and the feasible departure transfer stations are backward searching from D station. The intersection of the two feasible trajectories was regarded as the possible travel paths for passengers traveling within the OD pair.

The path recommendation was then explored according to passenger preferences, such as travel time, transfer penalty, and congestion tolerance. Compared with conventional algorithms, the proposed algorithm has the following advantages:

1) The generated accessible path set combines physical network structure with the dynamic train schedule;

2) No modification is needed for path guidance under the condition with train delays;

3) The time complexity of the proposed algorithm is low, and it can satisfy the requirement of real-time and fast computation;

4) Passengers' travel paths can be recommended according to their preference with travel time, transfer, and congestion.

Further studies could consider the congestion situation in the process of path search, and to take the passenger reservation (passengers cannot board the oncoming trains due to congestion, and they have to board the next following trains) into account for the path finding process. It is supposed to be applied in more metro networks to expand the application area of the algorithm.

## REFERENCES


[1] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numer. Math.*, vol. 1, no. 1, pp. 269–271, Dec. 1959.

[2] G. Y. Handler and I. Zang, "A dual algorithm for the constrained shortest path problem," *Networks*, vol. 10, no. 4, pp. 293–309, 1980.

[3] D. Eppstein, "Finding the k shortest paths," *SIAM J. Comput.*, vol. 28, no. 2, pp. 652–673, 1998.

[4] E. Hadjiconstantinou and N. Christofides, "An efficient implementation of an algorithm for finding K shortest simple paths," *Netw., Int. J.*, vol. 34, no. 2, pp. 88–101, 1999.

[5] Y. Sun and P. M. Schonfeld, "Schedule-based rail transit path-choice estimation using automatic fare collection data," *J. Transp. Eng.*, vol. 142, no. 1, Jan. 2016, Art. no. 04015037.

[6] Y. Zhu, H. N. Koutsopoulos, and N. H. Wilson, "A probabilistic Passenger-to-Train Assignment Model based on automated data," *Transp. Res. B, Methodol.*, vol. 104, pp. 522–542, Oct. 2017.

[7] C. Tong and S. Wong, "A stochastic transit assignment model using a dynamic schedule-based network," *Transp. Res. B, Methodol.*, vol. 33, no. 2, pp. 107–121, Apr. 1999.

[8] S. C. Wong and C. O. Tong, "Estimation of time-dependent origin–destination matrices for transit networks," *Transp. Res. B, Methodol.*, vol. 32, no. 1, pp. 35–48, 1998.

[9] R. Huang and Z.-R. Peng, "Schedule-based path-finding algorithms for transit trip-planning systems," *Transp. Res. Rec.*, vol. 1783, no. 1, pp. 142–148, Jan. 2002.

[10] X. Hu and Y.-C. Chiu, "A constrained time-dependent K shortest paths algorithm addressing overlap and travel time deviation," *Int. J. Transp. Sci. Technol.*, vol. 4, no. 4, pp. 371–394, 2015.

[11] M. Friedrich, I. Hofsaess, and S. Wekeck, "Timetable-based transit assignment using branch and bound techniques," *Transp. Res. Rec.*, vol. 1752, no. 1, pp. 100–107, Jan. 2001.

[12] W. Xu, S. He, R. Song, and S. S. Chaudhry, "Finding the K shortest paths in a schedule-based transit network," *Comput. Operations Res.*, vol. 39, no. 8, pp. 1812–1826, Aug. 2012.

[13] F. Zhou and R.-H. Xu, "Model of passenger flow assignment for urban rail transit based on entry and exit time constraints," *Transp. Res. Rec.*, vol. 2284, no. 1, pp. 57–61, Jan. 2012.

[14] J. Guo and L. Jia, "A new algorithm for finding the K shortest paths in a time-schedule network with constraints on arcs," *J. Algorithms Comput. Technol.*, vol. 11, no. 2, pp. 170–177, Jun. 2017.

[15] N. Van Der Zijpp and S. Fiorenzo Catalano, "Path enumeration by finding the constrained K-shortest paths," *Transp. Res. B, Methodol.*, vol. 39, no. 6, pp. 545–563, Jul. 2005.

[16] W. Li, J. Zhou, and F. Zhou, "Algorithm of valid path set generation for urban rail transit based on space-time prism," Presented at the 94th Annu. Meeting Transp. Res. Board, Washington, DC, USA, 2015.

[17] S. Wang, Y. Yang, X. Hu, J. Li, and B. Xu, "Solving the K-shortest paths problem in timetable-based public transportation systems," *J. Intell. Transp. Syst.*, vol. 20, no. 5, pp. 413–427, Sep. 2016.

[18] Y. Yang, S. Wang, X. Hu, J. Li, and B. Xu, "A modified k-shortest paths algorithm for solving the earliest arrival problem on the time-dependent model of transportation systems," in *Proc. Int. Multi Conf. Eng. Comput. Sci.*, vol. 2012, pp. 1562–1567.

[19] I. Jeon, H. Nam, and C. Jun, "A schedule-based public transit routing algorithm for finding K-shortest paths considering transfer penalties," *J. Korea Inst. Intell. Transp. Syst.*, vol. 17, no. 3, pp. 72–86, Jun. 2018.

[20] W. Li, Q. Luo, Q. Cai, and X. Zhang, "Using smart card data trimmed by train schedule to analyze metro passenger route choice with synchronous clustering," *J. Adv. Transp.*, vol. 2018, Jul. 2018, Art. no. 2710608.

[21] K. M. Kim, S.-P. Hong, S.-J. Ko, and D. Kim, "Does crowding affect the path choice of metro passengers?" *Transp. Res. A, Policy Pract.*, vol. 77, pp. 292–304, Jul. 2015.

[22] L. Kang, J. Wu, H. Sun, X. Zhu, and B. Wang, "A practical model for last train rescheduling with train delay in urban railway transit networks," *Omega*, vol. 50, pp. 29–42, Jan. 2015.

[23] L. Kang, X. Zhu, H. Sun, J. Wu, Z. Gao, and B. Hu, "Last train timetabling optimization and bus bridging service management in urban railway transit networks," *Omega*, vol. 84, pp. 31–44, Apr. 2019.

[24] L. Kang and Q. Meng, "Two-phase decomposition method for the last train departure time choice in subway networks," *Transp. Res. B, Methodol.*, vol. 104, pp. 568–582, Oct. 2017.

[25] L. Kang, X. Zhu, H. Sun, J. Puchinger, M. Ruthmair, and B. Hu, "Modeling the first train timetabling problem with minimal missed trains and synchronization time differences in subway networks," *Transp. Res. B, Methodol.*, vol. 93, pp. 17–36, Nov. 2016.

[26] L. Kang, J. Wu, H. Sun, X. Zhu, and Z. Gao, "A case study on the coordination of last trains for the Beijing subway network," *Transp. Res. B, Methodol.*, vol. 72, pp. 112–127, Feb. 2015.


**WEI LI** received the B.S. degree in transportation engineering and the Ph.D. degree in civil engineering from Tongji University, Shanghai, in 2011 and 2016, respectively.

From 2017 to 2019, he held a postdoctoral position with the Mechanical Engineering Department, Shenzhen University. He is currently an Assistant Professor with the College of Urban Transportation and Logistics, Shenzhen Technology University. His research interests include transportation engineering, especially for public transportation, travel behavior analysis, passenger flow analysis, train timetable optimization, urban rail transit operation, big data, and AI.

**QIN LUO** received the B.S. and M.S. degrees in transport engineering from the University of Tongji, Shanghai, China, in 2009, and the Ph.D. degree in civil engineering from Tongji University, Shanghai, in 2013.

From 2013 to 2014, he worked for South Korean subway. From 2014 to 2017, he was a Research Assistant with the College of Mechanical Engineering and Control, Shenzhen University. Since 2018, he has been an Associate Professor with Shenzhen Technology University. His research interests include travel behavior analysis, travel demand modeling, and urban rail transit operation.

**QING CAI** received the B.S. and M.S. degrees in civil engineering from Tongji University, Shanghai, China, in 2011 and 2014, respectively, and the Ph.D. degree in transportation engineering from the University of Central Florida, Orlando, FL, USA, in 2017.

Since 2017, he has been a Research Assistant Professor of transportation engineering with the University of Central Florida. His research intersects include transportation planning, big data, statistical modeling, deep learning, and intelligent transportation systems.

Dr. Cai serves as a member of the Transportation Research Board Visualization in Transportation Committee (ABJ95). He was a recipient of the Transportation Research Board 2018 Young Researcher Paper Award.