Research Article

Schedule-Based Passenger Assignment for High-Speed Rail Networks considering the Ticket-Booking Process

Huanyin Su,¹ Feng Shi,¹ Guangming Xu,¹ Jin Qin,¹ and Xinghua Shan²

¹School of Traffic and Transportation Engineering, Central South University, Changsha 410075, China
²Institute of Computing Technologies, China Academy of Railway Sciences, Beijing 100081, China

Correspondence should be addressed to Feng Shi; shifeng@csu.edu.cn

Received 20 March 2016; Revised 10 May 2016; Accepted 11 May 2016

Academic Editor: Luca D’Acierno

Copyright © 2016 Huanyin Su et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a schedule-based passenger assignment method for high-speed rail networks considering the ticket-booking process. Passengers book tickets to reserve seats during the presale period in high-speed rail systems and passengers on trains are determined during the ticket-booking process. The ticket-booking process is modeled as a continuous and deterministic predecision process. A solution algorithm is designed using the discretization of the continuous process by partitioning the ticket-booking time and the optimal paths remain constant in any partition interval. Finally, an application to the Chinese high-speed rail network is presented. A comparison of the numerical results with the reality is conducted to validate the efficiency and precision of the method and algorithm. Based on the results, the operating efficiency of the current train schedule is evaluated and some specific improvement measures are proposed.

1. Introduction

In recent years, there is a rapid development of high-speed rail networks in the world, especially in China [1]. The passenger assignment is a key point for optimizing line planning problems and timetables in high-speed rail networks. The optimization effect is directly determined by the precision and efficiency of passenger assignment. Hence, designing an applicable assignment method is an important issue.

In recent studies, passenger assignment problems have been widely studied with frequency-based models and schedule-based models [2], which were mainly focused on urban transit systems. Urban transit systems are characterized by uncertainty, space-time priority (time priority means that passengers are first-come–first-served and space priority means that passengers waiting at the upstream stations of the lines have the priority to obtain seats over passengers waiting at the downstream stations), and crowding. Hence, the focuses of the studies are also on the above characters. For example, Sumalee et al. [3] proposed a dynamic transit assignment model that had an explicit seat-allocation process. Nuzzolo et al. [4] proposed a dynamic assignment model with a joint choice for the departure time and number of stops and a space-time path in which users decided to leave at given times, to access the network at given stops, and to board given trains to reach their destinations. Hamdouch et al. [5] proposed a schedule-based transit assignment model with travel strategies and loaded passengers on a first-come–first-served basis.

However, these features are not shared by high-speed rail systems. In high-speed rail systems, passengers can book tickets before the departure day, for example, two months early in China. Then, passengers can choose their optimal paths based on the available seats before departure, which is a predecision problem. Besides, the number of tickets sold is usually kept below each train’s seating capacity in many countries. Especially in China and Europe, a passenger should reserve a seat (In China, some standing tickets are sold but only during peak seasons, such as the Spring Festival) and this prevents overcrowding and overbooking on high-speed trains. Hence, in high-speed rail systems, a passenger’s choice of path depends entirely on the ticket-booking process, and then passengers on trains are determined in the process.

The ticket-booking process in the rail systems has not been considered by researchers in the passenger assignment methods so far. The methods having been designed are shown...
as follows. The system split method [6] assigning passengers to the different train types was used for line planning [7, 8]. Cascetta and Coppola [9] adopted a stochastic network loading model with no capacity constraints to estimate the passengers on individual trains. Kaspri and Raviv [10] assigned passengers on the paths with minimum time and the capacity of the trains was not binding. Fu et al. [1] assumed that passengers sought the most time-saving paths from all the available routes and were assigned very dispersely to the train service network. Passengers' response to crowding was studied in a train passenger assignment model [11]. Cascetta and Coppola [12] presented an assessment of schedule-based and frequency-based assignment models of high-speed rail services and predicted passengers on individual trains and station-to-station segments. In these studies, the ticket-booking process was not described and analyzed.

Hence, a new method for passenger assignment problems in high-speed rail systems should be built. Based on the characters in high-speed rail networks, the authors conclude that there are two key points in the ticket-booking process: the book-ticketing time and the path choice. In the literature, Carrier [13] analyzed the choice of airline itinerary and fare product using booking and seat availability data. Hetrakul and Cirillo [14] studied the ticket-booking time decisions of railway passengers based on the ticket fare variation over the sale horizon. Ticket fare discounts in the presale period were the mainly decision factor in the above two studies. However, the ticket fares of the high-speed rail system in China are fixed at present. Besides, the ticket-booking time is quite an uncertainty and complicated variable, which depends on many factors, such as when the decision of traveling is made, the available tickets, the travel distance, the destination, and many other personal factors. Hence, the distributions of passengers depending on the ticket-booking time are given as input data in this paper, which deserve deep study in a separate paper.

The purpose of this paper is to study the path choices of passengers at any given ticket-booking time and load passengers on the trains in high-speed rail systems. Hence, this paper proposes a schedule-based passenger assignment method for high-speed rail networks considering the ticket-booking process. At any ticket-booking time, the optimal paths for passengers are deterministic based on the available tickets. Then, the ticket-booking process is modeled as a continuous and deterministic predication process. A solution algorithm is designed using the discretization of the continuous process by partitioning the ticket-booking time and the expected departure time, which reduces the complexity of calculation.

(iii) In the application, passengers on any train segment and the paths with minimum cost at any ticket-booking time are calculated. A comparison of the results with the reality is conducted to validate the precision and efficiency of the method. Based on the results, the operating efficiency of the current train schedule is evaluated and some specific improvement measures are proposed.

2. Problem Definition

Schedule-based passenger assignment problems for high-speed rail networks can be described as follows:

(i) Inputs are as follows: high-speed rail network, high-speed train's schedule, daily travel demands between any two stations, and the demand distributions depending on the expected departure time and ticket-booking time.

(ii) Assumptions are as follows:

(a) All the unsold tickets are open to any passenger in a fair way and all passengers have the same opportunity to book tickets.

(b) All passengers have desired departure times based on their plans, which are called expected departure times, and the desired arrival time is not considered.

(c) All passengers book tickets in advance during the presale period. At any ticket-booking time, the number of passengers booking tickets between any two stations is given and fixed, that is, not considering ticket-canceling.

(d) Passengers reserve seats to choose paths with minimum cost when booking tickets. The number of tickets sold should not exceed the train seating capacity. The ticket class is not considered. The ticket fare variation is not considered. Every passenger on the train has a seat, so the crowding effect is not considered.

(e) For every station, there is a minimum time required for passengers to switch trains.

(iii) Output is as follows: passengers on every train segment.

A passenger's choice of path depends entirely on the ticket-booking process. Hence, passenger assignment in a high-speed rail network is determined by the ticket-booking process. Ticket-booking time is continuous and passengers always choose the paths with the current minimum cost based on the available train seating capacity at any ticket-booking time. The choices are deterministic and cannot be changed with the ticket-booking time continuing. Then, the ticket-booking process is modeled as a continuous and deterministic predication process.
3. Travel Demand and Network

3.1. Travel Demand and Distribution. A high-speed rail network is represented by \((V,E)\), where \(V\) is the set of station nodes and \(E\) is the set of arcs. The operating period for a given high-speed rail network is denoted by \([t_1,t_2]\) (trains running overnight are not considered). The set of O-D pairs is denoted as \(RS \subseteq V \times V\), and the daily travel demand of an O-D pair is denoted by \(q_{rs}\), \((r,s) \in RS\).

For any given O-D pair, there is a distribution depending on the expected departure time for the travel demand, for example, from Beijing to Shanghai, as shown in Figure 1. The intensity distribution function is denoted by \(f_{rs}(x), x \in [t_1,t_2], (r,s) \in RS\), satisfying the following:

\[
\int_{t_1}^{t_2} f_{rs}(x) \, dx = q_{rs}, \quad (r,s) \in RS.
\] (1)

According to the assumptions, all passengers book tickets in advance during the presale period \([\bar{t},0]\) (\(\bar{t}\) is negative, \(\bar{t}\) is the maximum presale time ahead of the departure day, and 0 means the departure day) and the number of passengers at any ticket-booking time is given. Hence, there is a distribution depending on the ticket-booking time for the travel demand, for example, from Beijing to Shanghai, as shown in Figure 2. The distributions were called booking curves by Tsai [15]. The intensity distribution function is denoted by \(g_{rs}(y), y \in [\bar{t},0], (r,s) \in RS\), satisfying the following:

\[
\int_{\bar{t}}^{0} g_{rs}(y) \, dy = q_{rs}, \quad (r,s) \in RS.
\] (2)

The authors assume that the above two variables are mutually independent; the combined probability density function of travel demand is expressed as follows:

\[
f_{rs}(x) g_{rs}(y) = \frac{q_{rs}}{q^*_{rs}}, \quad x \in [t_1,t_2], y \in [\bar{t},0].
\] (3)

The passenger intensity distribution function is expressed as follows:

\[
f_{rs}(x) g_{rs}(y) = \frac{q_{rs}}{\sigma_{rs}}, \quad x \in [t_1,t_2], y \in [\bar{t},0].
\] (4)

The distributions of travel demand depending on the expected departure time and ticket-booking time can be obtained utilizing the historical reservation data in the ticket-booking system. First, the historical actual boarding times of passengers can be directly obtained from the ticket-booking system; then, the intensity distributions depending on the actual boarding time can be calculated, which can be used to forecast the expected departure times of passengers. Hence, the intensity distributions depending on the expected departure time can be obtained using some forecasting methods, such as the exponential smoothing method and the moving average method. Similarly, the intensity distributions of travel demand depending on the ticket-booking time for all O-D pairs result from the passengers’ experiences of ticket-booking process and can be forecasted based on the historical reservation data.

3.2. Space-Time Network. Let \(\Omega = \{T\}\) be the set of trains running on the network \((V, E)\). For each train \(T \in \Omega, h(T)\) is the number of stops and \(v^T_j\) is the \(j\)th stop. At the \(j\)th stop of train \(T\), \(a^T_j\) is the arrival time and \(d^T_j\) is the departure time, which are also called the arrival node and the departure node, respectively. Let \(T(i,j)\) be the segment of train \(T\)’s route from stop \(v^T_i\) to stop \(v^T_j\), called the “train segment” here. The length of \(T(i,j)\) is represented by \(|T(i,j)|\). Specifically, \(T(i,i+1)\) is called a “segment” for short.

If a passenger of an O-D pair \((r,s)\) transfers in turn to train segments to reach the destination, then the path can be expressed as follows:

\[
P^\text{Tran}_{rs} = T_1(i_1,j_1) + T_2(i_2,j_2) + \cdots + T_n(i_n,j_n).
\] (5)

Here, “+” refers to the transfer link between the train segments.

A space-time network \((\mathcal{V}', \mathcal{A}')\) is designed based on the itineraries graph designed by Kaspi and Raviv [10], shown in Figure 3, where \(\mathcal{V}'\) is a set of space-time nodes and \(\mathcal{A}'\) is a set of space-time arcs. Any path \(P^\text{Tran}_{rs}\) matches a route on the network.

Let \(s_\infty\) be a virtual node for all passengers with destination station \(s\), called the end node. This produces the following:

\[
\mathcal{V}' = \left\{d^T_i \mid 1 \leq i < h(T), T \in \Omega\right\} \\
\cup \left\{a^T_i \mid 1 < i \leq h(T), T \in \Omega\right\} \cup \left\{s_\infty \mid s \in V\right\}.
\] (6)
Space-time arcs include five categories (for more details, see Kaspi and Raviv [10]):

(i) The set of arcs from node \( a^T_i \) to node \( a^T_{i+1} \) is denoted as \( \mathcal{A}_{\text{board}} \).

(ii) The set of arcs from node \( a^T_i \) to node \( a^T_{i+1} \) is denoted as \( \mathcal{A}_{\text{pass}} \).

(iii) The set of arcs from node \( a^T_i \) to node \( a^T_{i+1} \) is denoted as \( \mathcal{A}_{\text{tran}} \).

(iv) The set of arcs from node \( a^T_i \) to node \( a^T_{i+1} \) is denoted as \( \mathcal{A}_{\text{wait}} \).

(v) The set of arcs from node \( a^T_i \) to node \( s_{co} \) is denoted as \( \mathcal{A}_{\text{end}} \).

4. Choices of Paths with Minimum Cost

Travel cost on a high-speed rail network consists of the ticket fare, travel time, and transfer cost. The ticket fare rate per kilometer is based on train \( T \)'s route is denoted as \( r_p(T) \); the ticket fare of the path in (5) is calculated as \( \sum_{k=1}^{u} [r_p(T_k) \cdot |T_k(t_k, j_k)|] \).

There is a risk for passengers to switch trains at a station; then, a penalty cost is used and called the transfer cost, denoted as \( \rho \). The total transfer cost of the path in (5) is expressed as \( \sum_{k=1}^{u-1} \rho \), where \( u - 1 \) are transfer times.

The cost of the travel time is defined as the fatigue cost. Let the fatigue cost per unit time be \( w \); then, the fatigue cost of the path in (5) is calculated as \( w(a^T_{i+1} - a^T_i) \).

In conclusion, the cost of the path in (5) is expressed as follows:

\[
P^{\text{Tran}}_{rs} = \sum_{k=1}^{u} [r_p(T_k) \cdot |T_k(t_k, j_k)|] + \sum_{k=1}^{u-1} \rho + w(a^T_{i+1} - a^T_i).
\]

In the space-time network, the costs of arcs are expressed as follows:

\[
\begin{align*}
\Phi = \rho + w(d^T_i - a^T_i), & \quad (a^T_i, a^T_i) \in \mathcal{A}_{\text{pass}} \\
\rho + w(\widehat{a}^T_i - a^T_i), & \quad (\widehat{a}^T_i, \widehat{a}^T_i) \in \mathcal{A}_{\text{board}} \\
0, & \quad (a^T_i, s_{co}) \in \mathcal{A}_{\text{end}}.
\end{align*}
\]

For an O-D pair \( (r, s) \), a path is a route from node \( d^T_i \) to node \( s_{co} \) on the network \((V', \mathcal{A}')\) and the path with minimum cost is expressed as \( P^\text{Tran}_{rs}(d^T_i) \), which can be found through a shortest path algorithm on the network. Network \((V', \mathcal{A}')\) is acyclic, so a shortest path search over the space-time arcs from node \( s_{co} \) backwards toward all other nodes is used for any node \( s \in V \). The complexity of this method is linear regarding \( |\mathcal{A}| \) for each end node.

The actual boarding time of passengers may be earlier or later than the expected departure time and the time deviation would bring extra travel costs, which was also considered by Cascetta and Coppola [12]. For example, for a passenger of O-D pair \( (r, s) \) with an expected departure time \( x \) and actual boarding time \( d^T_i \), a travel plan is defined here as adding a start arc \((x, d^T_i)\) to path \( P^\text{Tran}_{rs}(d^T_i) \), denoted as \( P^\text{rs}(x) \) (note that a travel plan starts with an expected departure time and a path starts with an actual boarding time; the latter is part of the former). Let the “early” cost and “late” cost per unit time be \( \theta \); then, the cost of the start arc \((x, d^T_i)\) is expressed as \( \theta|x - d^T_i| \).

In conclusion, travel plan \( P^\text{rs}(x) \) is expressed as follows:

\[
P^\text{rs}(x) = (x, d^T_i) + P^\text{Tran}_{rs}(d^T_i).
\]

At any ticket-booking time \( y \in [\bar{t}, 0] \), passengers always choose the paths with the current minimum cost based on the available train seating capacity and such paths are effective. Hence, \( P^\text{Tran}_{rs}(d^T_i) \) is extended as \( P^\text{rs}_{rs}(d^T_i, y) \) and the travel plan with minimum cost is denoted as \( P^\text{rs}_{rs} \). The authors obtain the following:

\[
\bar{P}^\text{rs}_{rs}(x, y) = \min \left\{ \theta|x - d^T_i| + \|P^\text{Tran}_{rs}(d^T_i, y)\| \right\}
\]

\[
= r, \ T \in \Omega.
\]

5. Solution Algorithm

The calculation of (10) can determine the best paths for all passengers, and then passengers on every train can be calculated. Considering that the ticket-booking time and the expected departure time are continuous variables, the computational tasks are large. Hence, a discretization method by partitioning the ticket-booking period and the expected departure period into several discrete intervals is designed.
5.1. Discretization

(1) The Partitioning of the Ticket-Booking Period. As the ticket-booking time continues, some trains segments become fully occupied and the travel plans with minimum cost change. Hence, the authors take the time of any such change as a partition node to separate the ticket-booking period into several intervals, and, as a result, in any partition interval, the travel plan with the minimum cost remains constant for passengers of the same O-D pair and expected departure time.

The authors set vector $Y = (y_0, y_1, \ldots, y_{N'} )$ with $t = y_0 \leq y_1 \leq \ldots \leq y_{N'} = 0$, where $y_n (n = 1, 2, \ldots, N')$ is a partition node. The ticket-booking period is partitioned by vector $Y$. When the number of partition intervals is the minimum, such a partition $Y$ is denoted as the optimal partition. In any partition interval $[y_{n-1}, y_n)$ $(n = 1, 2, \ldots, N')$, the travel plan with the minimum cost remains constant for all passengers; hence, $\hat{P}_{Tran}(x, y)$ is denoted as $\hat{P}_{Tran}(x)$ and $\hat{P}_{rs}(d_{i,Tran}^T, y)$ is denoted as $\hat{P}_{rs}(d_{i,Tran}^T)$.

(2) The Partitioning of the Expected Departure Period. In the optimal partition interval $[y_n, y_{n+1}) (n = 0, 2, \ldots, N' - 1)$, the travel plan cost $\theta(x - d_{i,Tran}^T) + |\hat{P}_{Tran}(d_{i,Tran}^T)|$ is a piecewise linear function according to the expected departure time $x$. Hence, the expected departure period $[t_1, t_2]$ can be partitioned into $M_r = \{ |d_{i,Tran}^T| \mid v_i = r, T \in \Omega \}$ continuity intervals at most [16, 17], denoted as $[x_{Tran0}, x_{Tran1}]$, $[x_{Tran1}, x_{Tran2}]$, $\ldots$, $[x_{TranM_r-1}, x_{TranM_r}]$. The authors set the vector $X = (x_{Tran0}, x_{Tran1}, \ldots, x_{TranM_r})$. In any interval $[x_{Tran0}, x_{TranM_r}]$, the travel plans with minimum cost include the same path $\hat{P}_{Tran}(x_{Tran})$, where $x_{Tran}$ is called the common actual boarding time, and if $x_{Tran} \in [x_{Tran0}, x_{TranM_r}]$, then such a partition $X$ is denoted as the optimal partition. Obviously, the authors obtain the following:

$$\hat{P}_{Tran}(x_{Tran}) = \hat{P}_{Tran}(x_{Tran0}) \oplus \hat{P}_{Tran}(x_{TranM_r}) \oplus \hat{P}_{Tran}(x_{Tran})$$

In conclusion, if optimal partitions $(X, Y)$ are obtained, the travel plan with minimum cost is $\hat{P}_{Tran}(x) = (x, x_{Tran}) + \hat{P}_{Tran}(x_{Tran})$ for passengers in the partition area $[x_{Tran0}, x_{TranM_r}] \times [y_n, y_{n+1})$, $0 \leq m < M_r$, $0 \leq n < N'$, $(r, s) \in \Omega$.

5.2. Optimal Partitions. To obtain the optimal partitions $(X, Y)$, the authors start from the ticket-booking time $y_0 = \tilde{t}$. For any given optimal partition node $y_n$, first, a method is designed to calculate the optimal partition $X$ of the expected departure period; then, the next optimal partition node $y_{n+1}$ is calculated, and as a result some train segment is fully occupied when the passengers in the interval $[y_n, y_{n+1})$ are loaded or $y_{n+1} = 0$. The process is repeated until $y_{n+1} = 0$ or until a feasible assignment does not exist.

5.2.1. Optimal Partitioning of Expected Departure Period. A simple example is shown in Figure 4 to illustrate the partition process. There are seven departure nodes at station $r$ and the paths with the minimum cost are $\hat{P}_{Tran}(x)$, $i = 1, 2, \ldots, 7$. Passengers always choose the travel plan with the minimum cost, so only paths $\hat{P}_{Tran}(x_{1})$, $\hat{P}_{Tran}(x_{4})$, $\hat{P}_{Tran}(x_{5})$, and $\hat{P}_{Tran}(x_{6})$ are chosen by passengers, and the costs of travel plans which include the other three paths are too high, so no passengers choose them. Then, the service period $[t_1, t_2]$ is partitioned into four intervals by the three vertical dashed lines. The time sequence $[x_{Tran0}, x_{Tran1}, \ldots, x_{Tran4}]$ in Figure 4 is the optimal partition. For example, passengers with an expected departure time in the second interval choose path $\hat{P}_{Tran}(x_4)$.

The rooftops method [16, 17] was used based on only one line and all trains have the same origin and destination, which is not suitable for complicated networks in this paper. Hence, a modified rooftops method is designed as follows. Assuming that there exists at least one effective travel plan for the O-D pair $(r, s)$, the set of trains passing through station $r$ is denoted as $\Omega_r$. According to the definition of the optimal partition of
the expected departure period, if departure node $d_i^T$ is the common actual boarding time of passengers at station $r$ in an optimal interval, then it satisfies (11), which is equivalent to the following:

$$\left| \tilde{P}_{Trn}^{Trans} (d_i^T) \right| < \min \left\{ \theta \left| d_i^T - d_i^r \right| \right\}$$

$$+ \left| \tilde{P}_{Trn}^{Trans} (d_i^r) \right| \left| d_i^r \neq d_i^T, y_i^r = r, T^r \in \Omega \right\}. \quad (12)$$

Given $y_n$ and O-D pair $(r, s)$, the correctness of (12) for all departure nodes at station $r$ can be verified with complexity $O(M_r)$, shown in the Appendix. Then, the departure node satisfying (12) is the common actual boarding time; that is, $\tilde{x}_{Trn}, 0 \leq m < M_r$. $y_n+1 - \tilde{x}_{Trn} = \tilde{x}_{Trn}$, the number of passengers through node $\tilde{x}_{Trn}$, the probability density function is the same for any $(r, s) \in RS$.

Those passengers choose the same path $\tilde{P}_{Trn}^{Trans} (\tilde{x}_{Trn})$, so the interval $[y_n, y_{n+1}]$, the number of passengers passing through node $\tilde{x}_{Trn}$ is

$$F_{Trn} = \int y_n \int_{x_{Trn}}^{x_{Trn+1}} \frac{1}{q_{rs}} f_{rs} (x) g_{rs} (y) \, dx \, dy. \quad (15)$$

In the network $(\mathcal{V}', \mathcal{A})$, passengers boarding on segments all pass through arrival nodes; hence, the capacity constraint on segment $T(i-1, i)$ $(1 < i \leq h(T), T \in \Omega)$ actually acts on node $d_i^T$, denoted as $C_T(i)$, and capacity constraints are considered only on nodes in the set $\mathcal{V}' = \{ a_i^T | 1 < i \leq h(T), T \in \Omega \}$ in the following parts of the paper.

According to the definition of the optimal partition node $y_{n+1}$, the authors obtain the following:

$$y_{n+1} = \max \left\{ y'_{n+1} | C_T (i) - F_T (n, i) - F_T (i) \geq 0, a_i^T \right\} \in \mathcal{V}' \right\}. \quad (18)$$

Then, in the interval $[y_n, y_{n+1}]$, the number of passengers passing through node $a_i^T$ is as follows:

$$F_{Trn} (i) = \delta \left( \tilde{P}_{Trn}^{Trans} (\tilde{x}_{Trn}), a_i^T \right) F_{Trn}^{Trans}. \quad (19)$$

5.3. Algorithm Framework. Based on the above description, the general structure of the solution algorithm is developed as shown in Algorithm 1.

In Algorithm 1, $n$ denotes the outer iterations. For example, $n = 3$ means the third ticket-booking phase. In phase every, all the travel plans with minimum cost remain constant.

5.4. A Special Case. In this section, the authors assume that there exists a probability density function $g(y)$ that satisfies $g(y) \equiv g_{rs} (y)/q_{rs}$ for any $(r, s) \in RS$; that is, the probability density function is the same for any $(r, s) \in RS$, and $f_{rs} (x)$ is a step function with an hourly interval.

In Algorithm 1, the optimal partition node $y_{n+1}$ calculated by (18) has two situations. One is $y_{n+1} = 0$ and the other is $y_{n+1} < 0$. If $y_{n+1} = 0$, then Algorithm I terminates; if $y_{n+1} < 0$, then the following is satisfied:

$$\min \left\{ C_T (i) - F_T (i) - F_T (i) | a_i^T \in \mathcal{V}' \right\} = 0. \quad (20)$$

Let $a_i^{r'}$ be the solution of the above equation; then, the following is true:

$$C_{T'} (i') - F_{T'} (i') = F_{T'} (i'). \quad (21)$$

And because $g(y) \equiv g_{rs} (y)/q_{rs}$, then

$$F_{Trn} = \int_{y_n}^{y_{n+1}} \int_{x_{Trn}}^{x_{Trn+1}} f_{rs} (x) g (y) \, dx \, dy \quad (22)$$
Input high-speed rail network \((V, E)\), space-time network \((V, A)\), service period \([t_1, t_2]\), train set \(\Omega\), O-D demand \(q_{rs}\), \((r, s) \in RS\);

Output passengers on every train segment, that is, \(F_T(n, i), 1 < i \leq h(T), T \in \Omega\).

Begin

\(n \leftarrow 0, y_0 \leftarrow T;\)

for all \(i = 2, 3, \ldots, h(T) - 1, T \in \Omega\) do \(F_T(n, i) \leftarrow 0;\)

while \(y_n < 0\) do 

begin

Search for the path with the minimum cost on the network \((\mathcal{Y}, \mathcal{A})\), then obtain

\[
\{ \tilde{P}_{Tran}^{i}(d_T) \mid v_T = r, 1 \leq i \leq n(T), T \in \Omega, (r, s) \in RS \};
\]

If an effective path does not exist, then break;

Obtain the optimal partition \(X\) by the modified rooftops method;

Calculate the next optimal partition node \(y_{n+1}\) using (18);

Calculate \(F_T^n, (r, s) \in RS, 0 \leq m < M\) using (15);

Calculate \(F_T^{(i)}, 1 < i \leq h(T), T \in \Omega\) using (19);

Load passengers on network \((V, A)\):

for all \(a_T^i \in V^a\) do \(F_T(n + 1, i) \leftarrow F_T(n, i) + F_T^n(i);\)

\(n \leftarrow n + 1;\)

end

end

Algorithm 1

Hence,

\[
C_T(i) - F_T(i) = \int_{y_n}^{y_{n+1}} g(y) dy
\]

\[
= \sum_{(r,s) \in RS} \sum_{m=0}^{M_r-1} \delta \left( \tilde{P}_{Tran}^{i} \left( x_{rsmn} \right), a_T^{im} \right) \int_{x_{rsmn}}^{x_{rsmn+1}} f_{rs}(x) dx.
\]

(23)

Then,

\[
\int_{y_n}^{y_{n+1}} g(y) dy
\]

\[
= \sum_{(r,s) \in RS} \sum_{m=0}^{M_r-1} \delta \left( \tilde{P}_{Tran}^{i} \left( x_{rsmn} \right), a_T^{im} \right) \int_{x_{rsmn}}^{x_{rsmn+1}} f_{rs}(x) dx
\]

Let

\[
\Delta(n) = \min \left\{ \frac{C_T(i) - F_T(i)}{\sum_{(r,s) \in RS} \sum_{m=0}^{M_r-1} \delta \left( \tilde{P}_{Tran}^{i} \left( x_{rsmn} \right), a_T^{im} \right) \int_{x_{rsmn}}^{x_{rsmn+1}} f_{rs}(x) dx} \mid a_T^i \in \mathcal{Y}_a \right\}.
\]

(25)

Then, \(\Delta(n) = \int_{y_n}^{y_{n+1}} g(y) dy\). Because \(g(y)\) is a probability density function, \(0 < \Delta(n) < 1\). Equation (22) is rewritten as follows:

\[
F_{rsnm} = \Delta(n) \int_{x_{rsmn}}^{x_{rsmn+1}} f_{rs}(x) dx.
\]

(26)

\[
\sum_{m=0}^{M_r-1} F_{rsnm} = \Delta(n) \sum_{m=0}^{M_r-1} \int_{x_{rsmn}}^{x_{rsmn+1}} f_{rs}(x) dx = \begin{cases} 
\Delta(n) \int_{t_1}^{t_2} f_{rs}(x) dx = \Delta(n) q_{rs}, & \text{if } \exists v_T^r = r, T \in \Omega: \left| \tilde{P}_{Tran} \left( d_T^r \right) \right| < \infty \\
\Delta(n) \int_{t_1}^{t_2} f_{rs}(x) dx = 0, & \text{if } \not\exists v_T^r = r, T \in \Omega: \left| \tilde{P}_{Tran} \left( d_T^r \right) \right| < \infty.
\end{cases}
\]

(27)
Equation (27) shows that if an effective travel plan exists for the O-D pair $(r, s) \in RS$, the number of passengers booking tickets is $\Delta(n)\gamma_n$, during period $[y_n, y_{n+1})$. The number of passengers booking tickets for all O-D pairs is determined with the same proportion $\Delta(n)$ by (25) during the period $[y_n, y_{n+1})$. Calculating $\Delta(n)$ using (25) does not require any specific probability density function $g(y)$ or partition time $y_{n+1}$. Hence, if the authors obtain $\Delta(n)$, then the number of passengers in period $[y_n, y_{n+1})$ can be calculated using (26).

Based on the above analysis, at the nth iteration, the process of passenger assignment is determined as follows. First, for the O-D pair $(r, s) \in RS$ and $m : 0 \leq m < M_r$, according to the expectation of passengers, there are $\int_{x_{stem}}^{x_{stem+1}} f_r(x)dx$ passengers loaded on the path $P_{rs}^\text{Trans}(x_{stem})$ and the expected number of passengers in every segment is obtained. Second, proportion $\Delta(n)$ is calculated using (25), and the actual number of passengers loaded on segments is $\Delta(n)$ times the expected one. Finally, if $\Delta(n) = 1 - \sum_{u=0}^{n-1} \Delta(u)$, that is, $\sum_{u=0}^{n} \Delta(u) = 1$, then the algorithm terminates (i.e., equal to $y_n = 0$); otherwise, if $\Delta(n) < 1 - \sum_{u=0}^{n-1} \Delta(u)$, that is, $\sum_{u=0}^{n} \Delta(u) < 1$, then the algorithm continues to the next period. Hence, the authors obtain a simple method for the calculation of Algorithm 1.

Under the assumption, the authors analyze the computational complexity of Algorithm 1 as follows. The maximum iteration of the outer loop is $N = O(|\mathcal{A}|)$. In the body loop, at the nth iteration, there are four parts. These are, in order, as follows.

First, the complexity of searching all paths with minimum cost for an end node is $O(|\mathcal{A}|)$ and $|\mathcal{A}| < 5|\mathcal{A}_{\text{board}}|$; hence, the complexity of searching all paths with minimum cost for all end nodes is $O(|\mathcal{A}_{\text{board}}|)|V|).

Second, the authors solve the optimal partition X. This is a two-layer loop. The outer layer is for all O-D pairs, and the iterations of the inner layer are $O(M_r)$. Hence, the complexity is $O(|V|\sum_{v \in V} M_v) = O(|\mathcal{A}_{\text{board}}|)|V|).

Third, the authors calculate $\int_{x_{stem}}^{x_{stem+1}} f_r(x)dx$, $0 \leq m < M_r$, $(r, s) \in RS$. The number of integrals is $O(|V|\sum_{v \in V} M_v) = O(|\mathcal{A}_{\text{board}}|)|V|)$. The complexity of calculating each integral is $O(1)$ because $f_r(x)$ is an hourly step function. Hence, the complexity for calculating all integrals is $O(|\mathcal{A}_{\text{board}}|)|V|).

Fourth, the authors load passengers on the paths with minimum cost. All the shortest paths to the end node $s_{oo}$ generate a shortest path tree with a root node $s_{oo}$. After removing the nodes finishing the passenger loading process from the shortest path tree, the authors select a leaf node of the shortest path tree as the current node performing the passenger loading process. Following the shortest path, the authors load the passengers accumulated at the current node to the next neighboring node until the passengers at each node are loaded at the root node $s_{oo}$. Because every arc is processed at most once, the computational complexity for loading all passengers at one root node $s_{oo}$ is $O(|\mathcal{A}_{\text{board}}|)|V|).

According to the third and fourth steps, the complexities of calculating $\Delta(n)$, $F_{r,m}$, $0 \leq m < M_r$, $(r, s) \in RS$, and $F_{r,m}(i), 1 \leq i \leq h(T)$, $T \in \Omega$, are all $O(|\mathcal{A}_{\text{board}}|)|V|).

In conclusion, the computational complexity of Algorithm 1 under the assumption is $O(|\mathcal{A}_{\text{board}}|)|V|).

6. Application to a Large-Size Network

China's high-speed rail network was used. There are 444 stations and 966 segments (Figure 5). The train schedule on 1 July 2014 was used, and there were 2100 trains. The number of segments was $|\mathcal{A}_{\text{board}}| = 15,263$. In the space-time network $(\mathcal{V'}, \mathcal{A})$ based on the train schedule, there were $|\mathcal{V'}| = 30,970$ nodes and $|\mathcal{A}| = 72,027$ arcs. The total number of passengers was 1,876,255.

The parameters are set as follows: The units of cost are Yuan (¥, Chinese monetary unit). Ticket fare rate is determined as ¥0.45 per kilometer for all trains based on the current ticket prices of high-speed trains in China. For example, from Beijingnan to Shanghai-Hongqiao, the travel distance is 1318 km and the ticket prices are ¥553 for second class, ¥933 for first class, and ¥1748 for business class. And most of the tickets are second class, but the classes of tickets are not considered in this paper; hence, ¥0.45 per kilometer is the average data. The fatigue cost is ¥0.5 per minute, which is determined by the average hourly wage (¥30/h) in 2014 in China. Let $\theta = ¥0.4$ per minute, which is set to be lower than the fatigue cost, because, before boarding trains, passengers can do other things and the penalty for time deviation is thought to be lower than the fatigue cost in the travelling process. The minimum transfer time at a station is 40 min, and transfer cost at a station is ¥30, which is equivalent to the fatigue cost of an hour.

6.1. Indexes. To validate the results of passenger assignment, some indexes are used. The load factor refers to the number of on-board passenger kilometers divided by the number of seat kilometers of trains, denoted as LF. The larger the load factor of a train, the higher the operational efficiency of the train:

$$LF = \frac{\sum_{t=2}^{h(T)} F_T(i) [T(i-1, i)]}{\sum_{t=2}^{h(T)} C_T(i) [T(i-1, i)]},$$

(28)

where $F_T(i)$ denotes the actual passengers loaded on the train segment $T(i-1, i)$.

To evaluate how well a train schedule coincides with time-dependent travel demand, the average time deviation between the expected departure time and the actual boarding time is used as another index. For passengers on all O-D pairs, the average time deviation is expressed as follows:

$$\text{ATD} = \frac{\sum_{s=0}^{N-1} \left( \Delta(n) \sum_{(r,s) \in RS} \sum_{m=0}^{M_r-1} \int_{x_{stem}}^{x_{stem+1}} [\hat{x}_{stem} - x] f_r(x)dx \right) \sum_{(r,s) \in RS} \theta_t}{\sum_{(r,s) \in RS} \theta_t}$$

(29)

Because the passenger assignment is restricted to one day, the indexes analyzed above are also restricted to one day.
6.2. Numerical Results and Analysis. All codes were written in C#, and computations were conducted on a 3.20 GHz PC with 8 GB of RAM under the assumption in Section 5.4. The ticket-booking time in the passenger assignment was partitioned into 127 intervals, and the average running time was 3 min. For some ticket-booking time intervals, the optimal partitions of the expected departure time and paths with the minimum cost from city Beijing to city Shanghai are shown in Tables 1 and 2. From the 1st ticket-booking time partition interval to the 16th one, the optimal partition showed no
changes and there were 16 departure time partition intervals (Table 1). From the 112th ticket-booking time partition interval to the 116th one, the optimal partition showed no changes and there were 6 departure time partition intervals (Table 2).

The paths in Table 1 require no transfers in any of the departure time partition intervals, and the travel time is less than 6 h. However, in Table 2, there are fewer expected departure time partition intervals and some paths require two transfers or transfers to other stations in the same city. Also, travel time increases visibly. The path in the last expected departure time partition interval [11:48, 23:59] in Table 2 is direct. It involves taking train G151 with a travel time of 6 h 1 min, which is longer than the direct path shown in Table 1. Based on the analysis above, as ticket-booking time continues, the changes are as follows. There are fewer partition intervals for the expected departure time. A larger average time deviation, longer transfer times, and longer travel time occur.

6.3. Comparison. To validate the passenger assignment results, a comparison with the reality data was conducted. In this experiment, there are 177 trains on Jingguang Line and some stations on this line are numbered in Table 3. The load factor of every train and the passengers on every train segment in this experiment (E) were compared with the ones in the reality (R) and let \( \varepsilon \) be the absolute error of the load factor. A part of the results are shown in Tables 4–6.

There are 83% trains with the absolute error less than 0.3, 34% trains with the absolute error less than 0.05, and only 5% trains with the absolute error more than 0.5 (see Table 4). For more details, some trains are chosen to show the passengers on segment and load factors (see Tables 5 and 6). By the comparison, the precision of the passenger assignment method in this paper is acceptable and the results can be used to evaluate the current train schedule.

6.4. Evaluations and Improvement Measures. The current train schedule was evaluated and improvement measures were proposed as follows.

(1) Load Factor. As shown in Table 7, 52.96% of trains have load factors greater than 70%, indicating the high operating efficiency. On Jinghu Line, which is located in the developed east coast regions of China and has a dense population and high travel demand, 70.54% of the trains have load factors greater than 70%, making it one of the busiest lines in China. Jingha-Hada Line is the passenger corridor in northeastern China, and 61.9% of trains on this line have load factors above 50%. Jingguang Line is the main passage between south and north, and 53.25% of trains on this line have load factors above 50%. Although the transport capacity of these lines involves a surplus, to maintain a high service level, the lower load factors of some trains are also reasonable. But for some trains with very low load factors can be canceled and for some busy lines, more trains can be added.

(2) Average Time Deviation. As shown in Figure 6, the average time deviation of 90% of passengers was less than one hour and the average time deviation of 69.74% of passengers was less than half an hour. According to (29), the average time deviation for all passengers was 27.36 minutes. In this way, the train schedule matches time-dependent travel demand well. This evaluation index can be used to optimize the departure time distribution of trains to improve the level of service of the train timetable. For passengers of some O-D pairs with large time deviation, the departure times of trains should be adjusted or more trains should be added.

(3) Improvement Measures. 31 trains from Beijingnan to Shanghai-Hongqiao were selected (see Table 8). Through the absolute error \( \varepsilon \) of the load factor of every train, it is shown that the assignment results on those trains are close to the reality. Based on the results in Table 8, the authors propose some improvement measures as follows:

(i) Capacity: the load factors of trains departing before 9:00 are about 0.5 except for train G11, of which the load factors are not high, and all their capacities are above 1000. Hence, the capacities of these trains can be reduced to about 600. Similarly, for some trains with very high load factors but low capacities, the capacities can be increased, for example, train G19.

(ii) Stop pattern: for trains G11, G1, G13, G15, G3, G17, and G19, there are two common characters. First, the average distances of stops are relatively long; second,
### Table 1: Optimal partitions of the expected departure time from the first ticket-booking time partition interval to the 16th one.

<table>
<thead>
<tr>
<th>Expected departure time interval</th>
<th>Travel time</th>
<th>Boarding station</th>
<th>Boarding time</th>
<th>Train number</th>
<th>Arrival time</th>
<th>Arrival station</th>
</tr>
</thead>
<tbody>
<tr>
<td>[00:00, 07:16]</td>
<td>5 h 37 min</td>
<td>Beijingnan</td>
<td>7:05</td>
<td>G103</td>
<td>12:42</td>
<td>Shanghai-Hongqiao</td>
</tr>
<tr>
<td>[07:16, 08:18]</td>
<td>5 h 9 min</td>
<td>Beijingnan</td>
<td>8:00</td>
<td>G11</td>
<td>13:09</td>
<td>Shanghai-Hongqiao</td>
</tr>
<tr>
<td>[08:18, 09:35]</td>
<td>4 h 48 min</td>
<td>Beijingnan</td>
<td>9:00</td>
<td>G1</td>
<td>13:48</td>
<td>Shanghai-Hongqiao</td>
</tr>
<tr>
<td>[09:35, 10:30]</td>
<td>4 h 55 min</td>
<td>Beijingnan</td>
<td>10:00</td>
<td>G13</td>
<td>14:55</td>
<td>Shanghai-Hongqiao</td>
</tr>
<tr>
<td>[10:30, 12:00]</td>
<td>4 h 55 min</td>
<td>Beijingnan</td>
<td>11:00</td>
<td>G15</td>
<td>15:55</td>
<td>Shanghai-Hongqiao</td>
</tr>
<tr>
<td>[12:00, 12:26]</td>
<td>5 h 35 min</td>
<td>Beijingnan</td>
<td>12:13</td>
<td>G129</td>
<td>17:48</td>
<td>Shanghai-Hongqiao</td>
</tr>
<tr>
<td>[12:54, 14:36]</td>
<td>4 h 48 min</td>
<td>Beijingnan</td>
<td>14:00</td>
<td>G3</td>
<td>18:48</td>
<td>Shanghai-Hongqiao</td>
</tr>
<tr>
<td>[14:36, 15:41]</td>
<td>4 h 59 min</td>
<td>Beijingnan</td>
<td>15:00</td>
<td>G17</td>
<td>19:59</td>
<td>Shanghai-Hongqiao</td>
</tr>
<tr>
<td>[15:41, 16:41]</td>
<td>5 h 16 min</td>
<td>Beijingnan</td>
<td>16:00</td>
<td>G19</td>
<td>21:16</td>
<td>Shanghai-Hongqiao</td>
</tr>
<tr>
<td>[16:41, 17:09]</td>
<td>5 h 35 min</td>
<td>Beijingnan</td>
<td>17:00</td>
<td>G21</td>
<td>22:35</td>
<td>Shanghai-Hongqiao</td>
</tr>
<tr>
<td>[17:16, 17:34]</td>
<td>5 h 42 min</td>
<td>Beijingnan</td>
<td>17:22</td>
<td>G155</td>
<td>23:04</td>
<td>Shanghai-Hongqiao</td>
</tr>
<tr>
<td>[17:34, 17:46]</td>
<td>5 h 38 min</td>
<td>Beijingnan</td>
<td>17:45</td>
<td>G157</td>
<td>23:23</td>
<td>Shanghai-Hongqiao</td>
</tr>
<tr>
<td>[17:46, 23:59]</td>
<td>5 h 38 min</td>
<td>Beijingnan</td>
<td>17:50</td>
<td>G159</td>
<td>23:28</td>
<td>Shanghai-Hongqiao</td>
</tr>
</tbody>
</table>

### Table 2: Optimal partitions of the expected departure time from the 112th ticket-booking time partition interval to the 16th one.

<table>
<thead>
<tr>
<th>Expected departure time interval</th>
<th>Travel time</th>
<th>Boarding station</th>
<th>Boarding time</th>
<th>Train number</th>
<th>Arrival time</th>
<th>Arrival station</th>
</tr>
</thead>
<tbody>
<tr>
<td>[00:00, 07:44]</td>
<td>8 h 8 min</td>
<td>Beijingnan</td>
<td>7:05</td>
<td>G103</td>
<td>12:24</td>
<td>Beijingnan</td>
</tr>
<tr>
<td>[07:44, 08:13]</td>
<td>8 h 25 min</td>
<td>Beijingnan</td>
<td>8:06</td>
<td>G107</td>
<td>13:29</td>
<td>Nanjingnan</td>
</tr>
<tr>
<td>[08:13, 08:41]</td>
<td>7 h 56 min</td>
<td>Beijingnan</td>
<td>8:35</td>
<td>G109</td>
<td>13:50</td>
<td>Nanjingnan</td>
</tr>
<tr>
<td>[08:41, 09:52]</td>
<td>7 h 57 min</td>
<td>Beijingnan</td>
<td>9:05</td>
<td>G113</td>
<td>14:06</td>
<td>Nanjingnan</td>
</tr>
</tbody>
</table>

### Table 3: Station numbers on Jingguang Line.

<table>
<thead>
<tr>
<th>Station</th>
<th>Number</th>
<th>Beijngxi</th>
<th>Shijiazhuang</th>
<th>Zhengzhoodong</th>
<th>Wuhan</th>
<th>Xiaanningbei</th>
<th>Cibibe</th>
<th>Yueyangdong</th>
<th>Miluodong</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>S1</td>
<td>S2</td>
<td>S3</td>
<td>S4</td>
<td>S5</td>
<td>S6</td>
<td>S7</td>
<td>S8</td>
<td>S9</td>
</tr>
<tr>
<td>Station</td>
<td>Changshan</td>
<td>Zhuzhouxi</td>
<td>Hengshaxi</td>
<td>Hengyangdong</td>
<td>Leiyangxi</td>
<td>Chenzhouxi</td>
<td>Shaoguan</td>
<td>Yingdexi</td>
<td>S10</td>
</tr>
<tr>
<td>Number</td>
<td>S9</td>
<td>S10</td>
<td>S11</td>
<td>S12</td>
<td>S13</td>
<td>S14</td>
<td>S15</td>
<td>S16</td>
<td>S18</td>
</tr>
<tr>
<td>Station</td>
<td>Qingyuanshui</td>
<td>Guangzhoubei</td>
<td>Guangzhouan</td>
<td>Shenzhenbei</td>
<td>S17</td>
<td>S18</td>
<td>S19</td>
<td>S20</td>
<td>S17</td>
</tr>
</tbody>
</table>
the load factors are very high. This is mainly because longer average distances of stops mean fewer stops for the same O-D pair, so the total travel time is shorter. Then, passengers with long travel distance would like to choose such trains to save time (the ticket fares of these trains are the same). Hence, to improve the service level, if the passengers on some trains are mostly with long travel distances, the middle stops of such trains can be reduced properly. This can save the time of passengers and attract more passengers.

(iii) Departure time: the load factors of trains departing at [9:00, 11:00] and [14:00, 16:00] are very high, which are the peak hours in one day, and then more trains departing at the peak hours can be added if there are enough passengers. The load factors of trains departing before 8:00 are some low, and then the departure times of such trains can be pushed back properly.

7. Conclusion

Schedule-based passenger assignment for high-speed rail networks has the following features. Passengers book tickets and reserve seats during the presale period, train capacity is taken up in the booking process, and the crowding effect does not occur. These are the essential differences from passenger assignment in urban transit networks. For a high-speed rail network, passenger assignment is completed in the ticket-booking process and the process can be modeled as a continuous and deterministic predecision process. Since the crowding effect is not considered, the minimum cost of travel plans remains constant during some stages of the booking process. Based on this feature, the authors design a method to optimally partition the ticket-booking period. For any optimal partition interval, the modified rooftops method is used to generate an optimal partition of the expected departure time. Finally, a solution algorithm is designed, and a large-scale application to the Chinese high-speed rail network is shown to validate the efficiency and precision of the proposed method and the solution algorithm.

The limits of the proposed method are as follows:

(i) The assumption in Section 5.4 does not agree with the reality. But considering that the ticket-booking time is quite an uncertainty and complicated variable, high cost is needed to determine the distributions of passengers depending on ticket-booking time. Hence, the method based on this assumption can be seen as a simplified computational method and can be applied in reality with low cost. What is more, the
Table 6: Load factor of some trains on Jingguang Line.

<table>
<thead>
<tr>
<th>Origin</th>
<th>Destination</th>
<th>Train</th>
<th>LF (E)</th>
<th>LF (R)</th>
<th>(\varepsilon)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>S19</td>
<td>G65</td>
<td>0.89</td>
<td>0.89</td>
<td>0</td>
</tr>
<tr>
<td>S1</td>
<td>S19</td>
<td>G69</td>
<td>0.87</td>
<td>0.85</td>
<td>0.02</td>
</tr>
<tr>
<td>S1</td>
<td>S20</td>
<td>G71</td>
<td>0.84</td>
<td>0.84</td>
<td>0</td>
</tr>
<tr>
<td>S1</td>
<td>S3</td>
<td>G89</td>
<td>0.83</td>
<td>0.87</td>
<td>0.04</td>
</tr>
<tr>
<td>S1</td>
<td>S9</td>
<td>G501</td>
<td>0.65</td>
<td>0.61</td>
<td>0.04</td>
</tr>
<tr>
<td>S1</td>
<td>S9</td>
<td>G505</td>
<td>0.84</td>
<td>0.82</td>
<td>0.02</td>
</tr>
<tr>
<td>S1</td>
<td>S4</td>
<td>G513</td>
<td>0.86</td>
<td>0.82</td>
<td>0.04</td>
</tr>
<tr>
<td>S1</td>
<td>S3</td>
<td>G563</td>
<td>0.26</td>
<td>0.24</td>
<td>0.02</td>
</tr>
<tr>
<td>S1</td>
<td>S4</td>
<td>D2031</td>
<td>0.76</td>
<td>0.77</td>
<td>0.01</td>
</tr>
<tr>
<td>S1</td>
<td>S2</td>
<td>G6705</td>
<td>0.21</td>
<td>0.17</td>
<td>0.04</td>
</tr>
<tr>
<td>S1</td>
<td>S2</td>
<td>G6707</td>
<td>0.16</td>
<td>0.18</td>
<td>0.02</td>
</tr>
<tr>
<td>S1</td>
<td>S2</td>
<td>G6713</td>
<td>0.56</td>
<td>0.61</td>
<td>0.05</td>
</tr>
<tr>
<td>S3</td>
<td>S20</td>
<td>G73</td>
<td>0.97</td>
<td>0.94</td>
<td>0.03</td>
</tr>
<tr>
<td>S3</td>
<td>S20</td>
<td>G75</td>
<td>0.82</td>
<td>0.86</td>
<td>0.04</td>
</tr>
<tr>
<td>S4</td>
<td>S20</td>
<td>G77</td>
<td>0.88</td>
<td>0.9</td>
<td>0.02</td>
</tr>
<tr>
<td>S4</td>
<td>S20</td>
<td>G1001</td>
<td>0.76</td>
<td>0.73</td>
<td>0.03</td>
</tr>
<tr>
<td>S4</td>
<td>S20</td>
<td>G1005</td>
<td>0.83</td>
<td>0.82</td>
<td>0.01</td>
</tr>
<tr>
<td>S4</td>
<td>S20</td>
<td>G1007</td>
<td>0.88</td>
<td>0.93</td>
<td>0.05</td>
</tr>
<tr>
<td>S4</td>
<td>S19</td>
<td>G105</td>
<td>0.72</td>
<td>0.71</td>
<td>0.01</td>
</tr>
<tr>
<td>S4</td>
<td>S19</td>
<td>G107</td>
<td>0.81</td>
<td>0.78</td>
<td>0.03</td>
</tr>
<tr>
<td>S4</td>
<td>S20</td>
<td>G109</td>
<td>0.82</td>
<td>0.81</td>
<td>0.01</td>
</tr>
<tr>
<td>S4</td>
<td>S20</td>
<td>G111</td>
<td>0.82</td>
<td>0.84</td>
<td>0.02</td>
</tr>
<tr>
<td>S4</td>
<td>S19</td>
<td>G1125</td>
<td>0.85</td>
<td>0.87</td>
<td>0.02</td>
</tr>
<tr>
<td>S4</td>
<td>S19</td>
<td>G1131</td>
<td>0.9</td>
<td>0.89</td>
<td>0.01</td>
</tr>
<tr>
<td>S4</td>
<td>S19</td>
<td>G1133</td>
<td>0.8</td>
<td>0.76</td>
<td>0.04</td>
</tr>
<tr>
<td>S4</td>
<td>S19</td>
<td>G1135</td>
<td>0.96</td>
<td>0.97</td>
<td>0.01</td>
</tr>
<tr>
<td>S4</td>
<td>S19</td>
<td>G1139</td>
<td>0.85</td>
<td>0.87</td>
<td>0.02</td>
</tr>
<tr>
<td>S4</td>
<td>S19</td>
<td>G1157</td>
<td>0.35</td>
<td>0.33</td>
<td>0.02</td>
</tr>
<tr>
<td>S9</td>
<td>S20</td>
<td>G6001</td>
<td>0.98</td>
<td>0.94</td>
<td>0.04</td>
</tr>
<tr>
<td>S9</td>
<td>S20</td>
<td>G6011</td>
<td>0.66</td>
<td>0.65</td>
<td>0.01</td>
</tr>
<tr>
<td>S9</td>
<td>S20</td>
<td>G6017</td>
<td>0.86</td>
<td>0.9</td>
<td>0.04</td>
</tr>
<tr>
<td>S9</td>
<td>S20</td>
<td>G6019</td>
<td>0.93</td>
<td>0.89</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Table 7: The distributions of trains depending on the load factor on some lines.

<table>
<thead>
<tr>
<th>LF interval</th>
<th>(90%, 100%)</th>
<th>(70%, 90%)</th>
<th>(50%, 70%)</th>
<th>(30%, 50%)</th>
<th>(0, 30%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The whole rail network</td>
<td>20.52%</td>
<td>32.44%</td>
<td>19.52%</td>
<td>12.71%</td>
<td>14.81%</td>
</tr>
<tr>
<td>Jinghu Line</td>
<td>29.46%</td>
<td>41.08%</td>
<td>20.54%</td>
<td>4.46%</td>
<td>4.46%</td>
</tr>
<tr>
<td>Jingha-Hada Line</td>
<td>16.40%</td>
<td>24.87%</td>
<td>20.63%</td>
<td>17.99%</td>
<td>20.11%</td>
</tr>
<tr>
<td>Jingguang Line</td>
<td>6.49%</td>
<td>30.53%</td>
<td>16.23%</td>
<td>20.13%</td>
<td>26.62%</td>
</tr>
</tbody>
</table>

(ii) The ticket class is not considered in this paper. The choices for different seat classes involve the different levels of income for passengers, which are planned to be further studied in another paper.

The method developed in this paper can be applied with low cost for high-speed rail networks in reality and can also be expanded to other resource allocation problems with ticket-booking.

**Appendix**

Equation (12) is decomposed into (A.1)-(A.2). Then, if and only if (A.1)-(A.2) are both satisfied, \(d^T_i\) is the common actual boarding time in an optimal interval. In particular, the earliest departure time from station \(r\) is not necessary to satisfy (A.1)
For convenience, let \( \Omega_r = \{T_r(1), T_r(2), \ldots, T_r(M_r)\} \) and let the departure time of train \( T_r(k) \) from station \( r \) be \( d^k \) for simplicity, satisfying \( d^k < d^{k+1} \), \( k = 1, 2, \ldots, M_r - 1 \). For \( 1 < k \leq M_r \), (A.1) can be rewritten as follows:

\[
\left| \tilde{P}^{\text{Tran}}_{r,m} (d^k) \right| < \min \left\{ \theta^* (d^k - d^j) + \left| \tilde{P}^{\text{Tran}}_{r,m} (d^j) \right| \mid j < k \right\}.
\]

Let \( Z^- (k) = \min \{\theta^* (d^k - d^j) + |\tilde{P}^{\text{Tran}}_{r,m} (d^j)| \mid j < k\} \), which is the minimum cost of the travel plan if the actual boarding time is earlier than the expected departure time \( d^k \) for passengers. The authors obtain

\[
Z^- (k) = \min \left\{ \left| \tilde{P}^{\text{Tran}}_{r,m} (d^j) \right| \mid j < k \right\}.
\]

For \( k = 1 \), (A.1) is not required; let the following be true:

\[
Z^- (1) = \infty.
\]

All \( Z^- (k) \) (1 \( \leq k \leq M_r \)) can be calculated using (A.4)-(A.5) with complexity \( O(M_r) \), and the correctness of (A.3) is also verified at the same time.

Symmetrically, for \( d^k \), \( 1 \leq k \leq M_r \), (A.2) can be rewritten as follows:

\[
\left| \tilde{P}^{\text{Tran}}_{r,m} (d^k) \right| < \min \left\{ \theta^* (d^k - d^j) + \left| \tilde{P}^{\text{Tran}}_{r,m} (d^j) \right| \mid j > k \right\}.
\]

Let \( Z^+ (k) = \min \{\theta^* (d^k - d^j) + |\tilde{P}^{\text{Tran}}_{r,m} (d^j)| \mid j > k\} \), which is the minimum cost of the travel plan if the actual boarding time is later than the expected departure time \( d^k \) for passengers. Then, the following can be determined:

\[
Z^+ (k) = \min \left\{ \left| \tilde{P}^{\text{Tran}}_{r,m} (d^j) \right| \mid j > k \right\}.
\]

For \( k = M_r \), (A.2) is not required; let the following be true:

\[
Z^+ (M_r) = \infty.
\]
All $Z^+(k)$ ($1 \leq k \leq M_r$) can be calculated using (A.7)-(A.8) with complexity $O(M_r)$, and the correctness of (A.6) is also verified at the same time. The correctness of (12) is verified using (A.4)-(A.5) and (A.7)-(A.8) for all $d^k$ ($1 \leq k \leq M_r$) with complexity $O(M_r)$.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This research was supported by National Natural Science Foundation of China (no. U1334207) and Innovation Project of Central South University (no. 2015zzts046).

References


