https://dx.doi.org/10.1016/j.tre.2020.101882

Train schedule optimization based on schedule-based stochastic passenger assignment

J. XIE [a], S.C. WONG* [b], S. ZHAN [c], S.M. LO [d], and Anthony Chen[e]

[a] *Department of Civil Engineering, The University of Hong Kong, China*
Email: jiemin@hku.hk
[b] *Department of Civil Engineering, The University of Hong Kong, China*
Email: hhecwsc@hku.hk
[c] *School of Transportation and Logistics, Southwest Jiaotong University, Chengdu, China*
Email: shuguangzhan@my.swjtu.edu.cn
[d] *Department of Architecture and Civil Engineering, City University of Hong Kong, China*
Email: bcsmli@cityu.edu.hk
[d] *Department of Civil and Environmental Engineering, Hong Kong Polytechnic University, China*
Email: anthony.chen@usu.edu

Declarations of interest: none.

*: Corresponding author.

Train schedule optimization based on schedule-based stochastic passenger assignment

Abstract

In this study, we propose a new schedule-based itinerary-choice model, the mixed itinerary-size weibit model, to address the independently and identically distributed assumptions that are typically used in random utility models and heterogeneity of passengers' perceptions. Specifically, the Weibull distributed random error term resolves the perception variance with respect to various itinerary lengths, an itinerary-size factor term is suggested to solve the itinerary overlapping problem, and random coefficients are used to model heterogeneity of passengers. We also apply the mixed itinerary-size weibit model to a train-scheduling model to generate a passenger-oriented schedule plan. We test the efficiency and applicability of the train-scheduling model in the south China high-speed railway network, and we find that it works well and can be applied to large real-world problems.

Keywords: Mixed itinerary-size weibit model, schedule-based, train scheduling

1. Introduction

The recent development of high-speed railways (HSRs) has attracted a great deal of attention. Railway companies wish to earn profits on their investments, and society has been promised that HSRs will improve the quality of travel. Achievement of these goals requires a good train schedule plan. The assessment of a schedule plan is based on passenger experience; hence, consideration of passenger-itinerary choices is important in scheduling. To improve the passenger service quality of train scheduling, this study introduces a new itinerary-choice model and integrates it with train scheduling.

1.1. Literature review

Our study involves two study areas, stochastic passenger assignment and passenger-oriented train scheduling, so that we divide the literature review into two parts to describe the recent developments of these two areas.

1.1.1. Stochastic passenger assignment

Traditionally, two kinds of models are used for transit systems. Frequency-based models are generally accepted in high-frequency transit systems, and they are sometimes used for low-frequency transit systems because they require cheaper and easier computation efforts than schedule-based models. Frequency-based models consider transit services in terms of lines, so they are not suitable to handle varying demand (e.g., morning peak). However, schedule-based models are more flexible and thus more suitable for dealing with low-frequency transit systems (Nuzzolo and Crisalli, 2009). They consider transit services in terms of runs (e.g., trains in railway systems) and consider the real departure/arrival time. Hence, schedule-based models are able to cope with varying demand. The HSR system is low-frequency, and its demand varies throughout a day. Taking HSR trips between Hong Kong and Shenzhen as an example, commuters go to work in the morning and come home in the evening, so there are two peak periods each day. Hence, schedule-based models are appropriate for this study.

Uncertainties have been highlighted in schedule-based passenger-assignment models. Passengers are assumed to be wise enough to make an itinerary choice to maximize their utility:

$$U(\boldsymbol{\beta},\boldsymbol{x}) = V(\boldsymbol{\beta},\boldsymbol{x}) + \varepsilon. \quad ()$$

The utility is estimated by the observed part $V(\boldsymbol{\beta},\boldsymbol{x})$, which considers the various costs $\boldsymbol{x}$ paid for this choice and the perceptions of passengers $\boldsymbol{\beta}$. Because of the passengers' unobserved behavioral patterns, part of the utility ε is random. This type of randomness is usually considered using the logit model or its variants (Domencich and McFadden, 1975; Hunt, 1990; Bekhor and Freund-Feinstein, 2006; Lin and Sibdari, 2009; Lurkin et al., 2017). Another type of randomness caused by heterogeneity in passengers' perceptions is represented by variations in $\boldsymbol{\beta}$. Tong and Wong (1999) suggested that $\boldsymbol{\beta}$ comprises independent random variables with probability density functions. They used the Monte Carlo algorithm to simulate random coefficients. To consider these two types of randomness, Adler et al. (2005), Warburg et al. (2006), and Parker (2017) recommended the mixed logit model, whereas Seelhorst and Liu (2015) suggested the latent class logit model.

However, the logit model has been criticized for its underlying assumption that ε should be independently and identically distributed (IID) (Sheffi, 1985), and nested logit models have been suggested to deal with the correlation. Hess and Polak (2006) and Drabas and Wu (2013) included nesting structures in their airline itinerary-choice models, and Freund-Feinstein and Bekhor (2017) considered both heterogeneity and correlation in their mixed cross-nested logit model. However, the characteristics of airline itineraries differ from those of train itineraries, so we cannot directly apply these findings to our study. For example, the overlapping problem, a common cause of correlation, is relatively simple in an airline network. If overlap exists between two airline itineraries, a single airline is used, and passengers take it from the origin to the destination. Hence, airline itineraries can be nested by airlines to address correlations. However, overlap in train itineraries may only occur in one part of a train journey. In the example shown in Figure 1, the three itineraries overlap because they all use IA7. Itinerary 1 and Itinerary 2 overlap at IA1, whereas Itinerary 1 and Itinerary 3 overlap at IA5 and IA6. IA5, IA6, and IA7 belong to the same train service. The selection of a suitable nesting structure for such cases is not easy.

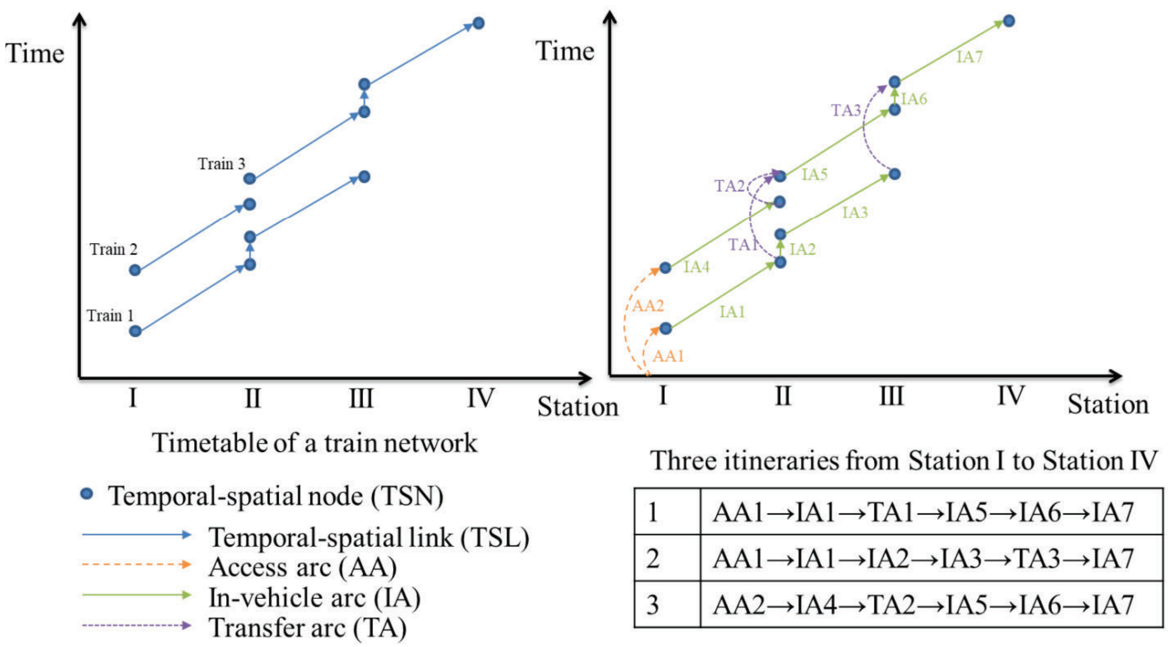


Figure 1. Example of itineraries.

There are other approaches to handling the overlapping problem in stochastic road assignments, such as the C-logit model (Cascetta et al., 1996; Cascetta et al., 1997; Zhou et al., 2012) and the path-size logit (PSL) model (Ben-Akiva and Bierlaire, 1999). These approaches add a correction term to utilities while maintaining the form of the logit model. Chen et al. (2012) examined the scaling effect and route overlapping problem in these approaches, and Lam and Xie (2002) and Tan et al. (2015) applied the PSL model to transit networks. The path-size factor in these models can alleviate the train overlapping problem. However, the identical distribution assumption is still needed in these models.

Under the identical distribution assumption, perceptions of trips of various lengths are the same, i.e., the absolute difference between trips decides the probability of their being chosen, rather than the relative difference. Castillo et al. (2008) proposed the weibit model to address the identically distributed assumption. The numerical results presented by Kitthamkesorn and Chen (2014) showed that the weibit model can account for the route-specific perception variance better than the logit model. Following this work, Kitthamkesorn and Chen (2013) developed the path-size weibit (PSW) model for a road network to relax the IID assumption. Numerical examples also illustrate that the PSW model works better in solving both the route-specific perception variance and the overlapping problem. Thus, the idea behind the PSW model is adopted for this study. The path-size factor for a road network can take several forms (Frejinger and Bierlaire, 2007). In this paper, we follow Kitthamkesorn and Chen (2013) and adopt the original form, which is derived from discrete choice theory for aggregate alternatives (Ben-Akiva and Lerman, 1985). The original form is defined as

$$\text{path-size factor} = \sum_{\text{all links in the path}} \left(\frac{\text{the length of the link}}{\text{total length of the path}} \times \frac{1}{\text{the number of paths using the link}}\right). \qquad ()$$

The ratio $\frac{\text{the length of the link}}{\text{total length of the path}}$ is an approximation of the path correlation, and the number of paths using the link indicates the influence of this link on the path correlation (Frejinger and Bierlaire, 2007).

However, the PSW model is built for a road network. The route choice in a transit network differs from that in a road network because the overlapping problem in a transit network is two-dimensional. We must therefore transfer it for itinerary case.

In summary, a suitable choice model for schedule-based passenger assignments in HSR networks should simultaneously consider the following aspects:

1) the characteristics of itineraries;
2) heterogeneity in passengers' perceptions;
3) the itinerary overlapping problem;
4) the identical distribution assumption.

To our knowledge, no model accounts for all of the above criteria. This study therefore proposes a mixed itinerary-size weibit (MISW) model to fill this gap. The utility function of the MISW model is formulated according to the characteristics of itineraries, and $\boldsymbol{\beta}$ comprises random coefficients to simulate heterogeneity in passengers' perceptions. The itinerary-size factor is added to the utility function to correct the effect of itinerary overlapping, and the weibit model is adopted to deal with the identical distribution assumption.

We apply the MISW model to the train-scheduling problem for its three main contributions:

1) It provides a way to assess train schedule plans according to passenger journey costs. For example, if a train schedule plan has a lower passenger journey cost, passengers pay less to complete their trips and are more likely to be satisfied with the train service.
2) It can contribute to the improvement of train schedule plans because it does not follow the assumption of homogeneity used by the deterministic passenger assignment. Thus, the MISW model can handle heterogeneity of passengers' perceptions and trace probabilities of different choices made by passengers. We can consider more aspects of passenger choice behaviors and a more precise passenger loading pattern will provide better guidance for train scheduling. For instance, priority of scheduling should be given to trains that carry more passengers, and the MISW model helps to identify these trains by simulating passenger loadings.
3) It enriches passenger-itinerary choices by finding several itineraries, rather than only one, to be improved. For example, if only one itinerary has been improved in the planning stage, when this itinerary cannot be used due to a system failure (such as a landslide or signal control incident), passengers may need to make several transfers with a long waiting time. Even worse, if there is no other possible itinerary, passengers can only wait for restored service or cancel their trips. However, if several possible itineraries are considered in the planning stage, passengers can use the other improved choices to continue their trip and enjoy the benefits of the improvement.

1.1.2. Passenger-oriented train-scheduling models

Several previous studies of train scheduling have been carried out. We refer interested readers to four recent surveys: Harrod (2012), Cacchiani and Toth (2012), Parbo et al. (2016), and Lusby et al. (2018). In this section, we focus on passenger-oriented train-scheduling models by discussing their development and solution methods. Based on this discussion, we identify research gaps in the literature, which require new studies, and justify our solution method.

1.1.2.1. Development of passenger-oriented train-scheduling models

Recently, train-scheduling models have begun to consider more aspects of passenger behavior. We denote such models "passenger-oriented" models, in light of the fact that the schedule is designed to minimize passengers' cost for their trips, so that passengers can enjoy a better train-travel experience. In this study, cost is considered in terms of journey time and fare, where the journey time comprises in-vehicle time (IVT), wait time, walking time for transfer, and travel time between stations and passengers' origin/destination.

Passenger-oriented scheduling can reduce passengers' cost. For instance, Talebian and Zou (2015) have focused on shortening passenger wait-time at stations, and smoothing transfer has also been addressed in the work of Bešinović et al. (2016), Kang and Meng (2017), and Kang et al. (2019b). However, the track allocation at stations for passenger transfers is rarely used in passenger-oriented train timetabling, although trains can be arranged to share the same platform to facilitate transfers between them. In addition, the track allocation at stations reflects the station-capacity problem (Kang et al., 2012; Wu et al., 2013; Dollevoet et al., 2017; Kang et al., 2019a), although it is completely ignored in macroscopic timetabling: for instance, Sparing and Goverde (2017) assumed the track allocation to be preassigned. In this study, the track allocation at stations is considered primarily to improve passenger transfers.

To reduce passengers' cost, a suitable passenger assignment is required to precisely simulate passenger behavior. However, previous passenger-oriented train-scheduling studies (Liebchen and Möhring, 2002; Vansteenwegen and Oudheusden, 2006; Wong et al., 2008; Niu and Zhou, 2013; Barrena et al., 2014; Canca et al., 2014; Niu et al., 2015; Kang et al., 2015) largely ignored passenger assignment because of its complexity. Later, with advancements in computer power, some studies used deterministic passenger-assignment models (Schmidt and Schöbel, 2015; Gattermann et al., 2016; Sels et al., 2016; Robenek et al., 2016; Borndörfer et al., 2017; Corman et al., 2017; Schöbel, 2017). The foundation of deterministic passenger-assignment models is the assumption that passengers make ideal decisions related to itinerary choices—that is, passengers are assumed to be homogenous and rational individuals who make the same perfect decisions, which is clearly unrealistic. The stochastic assignment of passengers can address these unrealistic assumptions by simulation of passenger-itinerary choices. Robenek et al. (2018) considered a stochastic passenger-assignment model using the logit model. As documented, the logit model has shortcomings caused by the IID assumption and fails to include heterogeneity in passengers' perceptions. Hence, a new itinerary-choice model is needed for passenger-oriented train scheduling to overcome the shortcomings of the present deterministic and stochastic passenger-assignment models.

As stated in Section 1.1.1, the MISW model proposed in this study can relax the IID assumption and consider heterogeneity in passengers' perceptions. Thus, this study applies the MISW model to the passenger-oriented train-scheduling problem to provide more accurate assessment and guidance for adjustment of train-schedule plans.

1.1.2.2. Solution methods of passenger-oriented train-scheduling models

Some previous passenger-oriented train-scheduling studies (Schmidt and Schöbel, 2015; Gattermann et al., 2016; Robenek et al., 2016) treated the entire problem as an integer linear problem (ILP) with the use of commercial software or by relaxing some constraints. However, the computational efficiency of this approach is questionable when dealing with large problems. To reduce the computation complexity for practical reasons, other researchers (Siebert, 2011; Sels et al., 2011; Sels et al., 2016; Borndörfer et al., 2017; Schöbel, 2017) divided the problem into two subproblems: train scheduling and passenger assignment. For example, Sels et al. (2011) suggested an iterative method with two steps: reflowing and retiming. Reflowing determines the number of passengers on trains, whereas retiming generates a timetable. Two steps are then connected via an iterative process. These researchers demonstrated the workability of the iterative method by applying it to their practical problems. In addition, the convergence of the iterative method has been analyzed from a theoretical perspective (Schöbel, 2017). We prefer the iterative method because of its advanced computational efficiency and adopt this approach to solve our rather complicated passenger-oriented train-scheduling model because it involves passenger-choice uncertainty, heterogeneity in passengers' perceptions, and correlated itineraries.

To further enhance the calculation efficiency of our method, the decomposition approach is included for train scheduling. Decomposition is a classical approach in which a large problem is decomposed into smaller subproblems that are then solved separately. It has been widely adopted in transportation research (Heydecker, 1996; Wong and Wong, 2002; Karabuk, 2009; Caimi et al., 2009; Talebian and Zou, 2015; Sinha et al., 2016; Zhou and Teng, 2016; Guo et al., 2016) and has also been applied to train scheduling. For instance, Ingolotti et al. (2006) introduced a meta-heuristic technique based on a decomposition approach to obtain a timetable for a railway corridor one train at a time. These studies have shown that the decomposition approach can be used to solve complex problems, with efficient computational tractability. Therefore, we adopt the decomposition approach for our iterative method.

In addition, for passenger assignment, we adopt the schedule-based route-search model developed by Tong and Richardson (1984) to search for possible itineraries. Tong and Richardson (1984), who performed the earliest studies on schedule-based transit assignments, suggested calculating wait/transfer times based on the scheduled times of transit vehicles. Hence, their model can be used to analyse the train system and passenger behaviour in a disaggregated manner, which is more accurate than using frequency-based models. This model has been widely used and updated in recent publications to make it more reliable for analysis of transit systems (Tong et al., 2001; Tong and Wong, 2004; Poon et al., 2004; Khani et al., 2014; Xie et al., 2017). Thus, we use most parts of this model and modify it to search for possible itineraries.

1.2. Objectives and contributions

This study proposes a new itinerary-choice model, the MISW model, that can handle the random part of utilities and the heterogeneity of passengers' perceptions. The MISW model relaxes the IID assumption, so it can address the overlap problem and itinerary-specific perception variance. Therefore, the MISW model can simulate a precise passenger loading pattern, which can be used to evaluate and improve the service quality of train scheduling. In this study, we apply the MISW model in the train scheduling which combines train timetabling, station-track allocation, a train cancellation plan, and stopping pattern design.

Section 2 describes the formulation of the MISW model. In Section 3, the capability of the MISW model is analyzed. Section 4 applies the MISW model to passenger-oriented train scheduling. In Section 5, two example networks are presented to test the application of the MISW model. A summary is given in Section 6.

2. Model formulation

In this section, we firstly provide a list of notations and basic assumptions. Then, we introduce the MISW model.

2.1. Notations and assumptions

2.1.1. Notation system

The used notations are as follows.

Table 1. Notations used in the model formulation.

Set	**Definition**
$\boldsymbol{A}_i$	Set of arcs itinerary i
$\boldsymbol{I}_r$	Set of all itineraries of OD pair r
$\boldsymbol{R}$	Set of all OD pairs
$\boldsymbol{y}_r$	Set of passenger characteristics of OD pair r with mean $\overline{\boldsymbol{y}}_r$ and variance $\boldsymbol{\sigma}_r^2$
$\boldsymbol{\theta}_{rk}$	Set of factors that reflects the influence of passenger characteristics of OD pair r on cost k
Element	**Definition**
a or $a^{'}$	Arc belonging to $\boldsymbol{A}_i$
i or i'	Itinerary belonging to $\boldsymbol{I}_r$
r	OD pair, $r \in \boldsymbol{R}$
Parameters	**Definition**
D_r	Passenger demand for OD pair r
T_r^{ar}	Latest time of arrival at destination for OD pair r
T_r^{de}	Preferred departure time of OD pair r in the origin zone
T^{wait}	Tolerance of passengers for waiting at the origin station
y_r^{IL}	Income level of OD pair r
y_r^{TP}	Trip purpose of OD pair r
$\overline{y}_r^{\mathrm{IL}}$	Mean of y_r^{IL}
$\overline{y}_r^{\mathrm{TP}}$	Mean of y_r^{TP}
β_r^k	Weights for cost k
ξ_r	Location parameter of the MISW model for OD pair r
γ_r	Shape parameter of the MISW model for OD pair r
$\theta_{rk}^{\mathrm{IL}}$	Coefficient of y_r^{IL} for cost k of OD pair r
$\theta_{rk}^{\mathrm{TP}}$	Coefficient of y_r^{TP} for cost k of OD pair r
θ_{rk}	Constant coefficient for cost k of OD pair r
σ_r^{IL}	Standard deviation of y_r^{IL}
σ_r^{TP}	Standard deviation of y_r^{TP}
Variables	**Definition**
b_{ai}^{O}	1, for arc a on itinerary i, else 0
$t_i^{\mathrm{wait-H}}$	Number of T^{wait} in the origin zone for itinerary i,

u_{ri}	Perceived travel cost of itinerary i for OD pair r
x_{ri}^{k}	Value of cost k of itinerary i for OD pair r, and cost k includes: k=1: journey time (IVT and access/egress time); k=2: wait time at stations; k=3: wait time in the origin zone; k=4: walking time; k=5: number of transfers; k=6: fare.
$\mathrm{Prob}_r(i)$	Probability of OD pair r selecting itinerary $i \in \boldsymbol{I}_r$
ω_i	Itinerary-size factor for itinerary i
ł_a	Time of arc a

2.1.2. Basic assumptions

According to the characteristics of our problem, we make the following assumptions:

1) Passengers begin or end their journeys in a zone, the unit of geography. The zone in which a journey begins is called the origin zone, whereas the zone in which a journey ends is called the destination zone. Passengers with the same origin and destination zones in a period are grouped as an origin-destination (OD) pair $r \in \boldsymbol{R}$, where $\boldsymbol{R}$ is the set of all OD pairs. The end of the period is assumed to be the preferred departure time of T_r^{de}. Limited by the passengers' activities at the destination, OD pair r has the requirement of the latest arrival time at the destination T_r^{ar}. The number of passengers of OD pair r is D_r. Because of the random components of utilities and heterogeneity in passengers' perceptions, the passenger-itinerary choice is not fixed on one itinerary; more than one itinerary could form a set of itineraries $\boldsymbol{I}_r$. Demand data are commonly obtained from historical ticket data, market research data, and demand forecast data.
2) Several stations may be placed near origin and destination zones. Passengers can select stations as their origin and destination station pair. This selection is considered in the itinerary choice.
3) Most itinerary-choice models, which are mainly used for air systems, include the following observed attributes to calculate the probability of choosing itinerary i:

$$\mathrm{Prob}_r(i) = f(\text{Flight time, Fare, Access/egress, Schedule time difference, Transfer, Alliance, Aircraft type}).()$$

Studies have illuminated the development of itinerary-choice models. However, the characteristics of airline itineraries differ from those of train itineraries, so we cannot directly apply them to our study. For example, trains in China are operated by the same company, so the alliance influence does not exist. In addition, the two main train types in China's HSR run at different speeds (250 km/h and 300 km/h) and charge different prices. Thus, the difference in vehicle type is reflected in the IVT and fare. This study considers five observed attributes in our itinerary-choice model, so that the probability that OD pair r chooses itinerary i is defined as

$$\mathrm{Prob}_r(i) = f(\text{Train time, Fare, Access/egress, Schedule time difference, Transfer}).\ ()$$

To deal with these five observed attributes, we assume that passengers bear six kinds of observed costs: journey time cost (IVT and access/egress time), wait cost at stations, wait cost at the point of origin, walking cost, transfer penalty cost, and fare cost. The IVT and access/egress time contribute to the journey time cost, and we assume that passengers hold the same perceptions about them. The difference between the ideal departure time and the actual departure time is the main concern for passengers, so the schedule time difference is reflected by the wait cost at the origin. The wait cost at the origin includes two parts: the wait cost in the origin zone and the wait cost at origin stations, because passengers tend to delay their arrival time at origin stations if the wait time at origin stations is too long. The wait costs in the origin zone are considered for the following reasons.

a) Passengers tend to enjoy their time in zones more than at stations because they can engage in useful activities in zones;
b) This division simulates departure time choices of passengers.

We assume that the wait time in the origin zone is a multiple of T^{wait} and that the wait time at the origin station is shorter than T^{wait}. T^{wait} is the longest time that passengers are willing to wait at an origin station. The wait time in the origin zone is set as a multiple of T^{wait} because of the passengers' departure tendency. If passengers wish to delay their departure, they are likely to wait for a multiple of a certain time interval, like 15 min or 30 min. This assumption can be easily relaxed if necessary. For itinerary i, the wait time in the origin zone is equal to $T^{\mathrm{wait}} \times t^{\mathrm{wait}-\mathrm{H}}_{i}$, where $t^{\mathrm{wait}-\mathrm{H}}_{i}$ is an integer variable. Hence, the departure time from the origin zone is $T^{\mathrm{de}}_{r} + T^{\mathrm{wait}} \times t^{\mathrm{wait}-\mathrm{H}}_{i}$.

Transfer is separated into three kinds of costs: wait cost at stations, walking cost, and an extra transfer penalty cost. The first two costs are for the extra journey time spent on transfers, whereas the last cost represents the inconvenience of transfers. Ticket prices, additive or nonadditive, must be transformed into a time unit so that all costs use the same unit (min).

The cost of itinerary i for OD pair r, u_{ri}, is the weighted sum of these costs, and it is weighted according to passengers' perceptions:

$$u_{ri} = \sum_{k} \beta^{k}_{r} x^{k}_{ri}, \qquad ()$$

where β^{k}_{r} is the weight coefficient of the respective cost x^{k}_{ri}.

4) Passengers may enjoy some parts of a journey more than others. For example, wait time at stations is usually considered to be annoying. β^{k}_{r} is determined by passenger characteristics, such as income and trip purpose. We assume that

$$\beta^{k}_{r} = g(\boldsymbol{\theta}_{rk}, \boldsymbol{y}_{r}), \qquad ()$$

where $\boldsymbol{y}_r$ is a set of passenger characteristics, and $\boldsymbol{\theta}_{rk}$ is a set of factors that reflects the influence of passenger characteristics of OD pair r on cost k. In this study, we assume that function () is a linear function and that the weight of the cost k is

$$\beta_r^k = \theta_{rk}^{\mathrm{IL}} y_r^{\mathrm{IL}} + \theta_{rk}^{\mathrm{TP}} y_r^{\mathrm{TP}} + \theta_{rk}, \qquad ()$$

where y_r^{IL} is the income level of OD pair r, y_r^{TP} is the trip purpose of OD pair r, and $\theta_{rk}^{\mathrm{IL}}$, $\theta_{rk}^{\mathrm{TP}}$, and θ_{rk} are coefficients.

Income level and trip purpose are usually the main attributes that influence passenger perceptions of trips. For example, passengers with high income or those who travel for business tend to weigh time as more important and fare cost as less critical. We therefore select these two passenger characteristics to reflect weight β_r^k. To show heterogeneity of passengers' perceptions, y_r^{IL} and y_r^{TP} are assumed to independently follow a normal distribution: $y_r^{\mathrm{IL}} \sim N(\overline{y}_r^{\mathrm{IL}}, (\sigma_r^{\mathrm{IL}})^2)$ and $y_r^{\mathrm{TP}} \sim N(\overline{y}_r^{\mathrm{TP}}, (\sigma_r^{\mathrm{TP}})^2)$. Hence, β_r^k also follows a normal distribution:

$$\beta_r^k \sim N(\theta_{rk}^{\mathrm{IL}} \overline{y}_r^{\mathrm{IL}} + \theta_{rk}^{\mathrm{TP}} \overline{y}_r^{\mathrm{TP}} + \theta_{rk}, (\theta_{rk}^{\mathrm{IL}} \sigma_r^{\mathrm{IL}})^2 + (\theta_{rk}^{\mathrm{TP}} \sigma_r^{\mathrm{TP}})^2). \qquad ()$$

All these variables can be calibrated with passenger data if they are available.

5) The train capacity is assumed to be sufficient for passengers, so passengers can always take their most-preferred itinerary. In this paper, we study a planning-stage problem and focus on typical daily operations rather than train operations during festivals (such as the Spring Festival in China). Hence, passenger demand is not very high. If there are many passengers on a train, we can consider arranging for a larger vehicle with more seats and carriages, such as a vehicle with 16 cars rather than 8. We can also regroup passenger demand to redistribute passengers. This approach is explained in detail in Section 5.1.1c.

2.2. Mixed itinerary-size weibit model

In this section, we develop the MISW model as the itinerary-choice model to assign passengers in schedule-based railway networks.

The probability of OD pair r selecting itinerary $i \in \boldsymbol{I}_r$, $\mathrm{Prob}_r(i)$, can be formulated as follows:

$$\mathrm{Prob}_r(i) = \int \mathrm{Prob}(i | \boldsymbol{y}_r = \tilde{\boldsymbol{y}}) \times \mathrm{Prob}(\boldsymbol{y}_r = \tilde{\boldsymbol{y}}) d\tilde{\boldsymbol{y}}, \qquad ()$$

where $\boldsymbol{y}_r = \tilde{\boldsymbol{y}}$ is an event that means that passenger perception $\boldsymbol{y}_r$ is reflected by $\tilde{\boldsymbol{y}}$; $\mathrm{Prob}(i | \boldsymbol{y}_r = \tilde{\boldsymbol{y}})$ is a conditional probability of choosing itinerary i when the event has already occurred, and Prob $(\boldsymbol{y}_r = \tilde{\boldsymbol{y}})$ is the probability of the event occurring. Multiplication of $\mathrm{Prob}(i | \boldsymbol{y}_r = \tilde{\boldsymbol{y}})$ and Prob $(\boldsymbol{y}_r = \tilde{\boldsymbol{y}})$ gives the probability that itinerary i is chosen and the event occurs. $\tilde{\boldsymbol{y}}$ is continual, and using integration to sum up all values of $\tilde{\boldsymbol{y}}$ can give the expected probability of choosing itinerary i.

It is easy to calculate $\mathrm{Prob}(\boldsymbol{y}_r = \tilde{\boldsymbol{y}})$ from the distribution of $\boldsymbol{y}_r$, but $\mathrm{Prob}(i | \boldsymbol{y}_r = \tilde{\boldsymbol{y}})$ is difficult to obtain because it is a joint probability related to a complex railway network. Here, we suggest the itinerary-size weibit (ISW) model, which follows the closed form of the PSW model (Kitthamkesorn and Chen, 2013), but the adjusted factor ω_i is changed to suit the itinerary case:

$$\mathrm{Prob}\left(i|\boldsymbol{y}_r=\tilde{\boldsymbol{y}}\right)=\frac{\omega_i\left(u_{ri}-\xi_r\right)^{-\gamma_r}}{\sum_{i^{'}\in\boldsymbol{I}_r}\omega_{i}^{'}\left(u_{ri^{'}}-\xi_r\right)^{-\gamma_r}}, \quad ()$$

where ξ_r is the location parameter ($\xi_r\in[0,u_{ri})$),

γ_r is the shape parameter ($\gamma_r\in(0,\infty)$), and

ω_i is defined as

$$\omega_i=\sum_{a\in\boldsymbol{A}_i}\frac{\text{ł}_a}{\sum_{a^{'}\in\boldsymbol{A}_i}\text{ł}_{a^{'}}}\frac{1}{\sum_{i^{'}\in\boldsymbol{I}_r}b_{ai^{'}}^{0}}, \quad ()$$

where ł_a is the time of arc a ($a\in\boldsymbol{A}_i$, $\boldsymbol{A}_i$ is the set of arcs of itinerary i) and b_{ai}^{0} is equal to 1 for arc a on itinerary i and 0 otherwise. Here, ω_i is the itinerary-size factor for itinerary i. The ratio $\frac{\text{ł}_a}{\sum_{a\in\boldsymbol{A}_i}\text{ł}_a}$ is an approximation of the itinerary correlation, and $\sum_{i^{'}\in\boldsymbol{I}_r}b_{ai^{'}}^{0}$ indicates the impact of arc a on the itinerary correlation.

We put formulation () back into formulation () and then have

$$\mathrm{Prob}_r(i)=\iint\frac{\omega_i\left(u_{ri}-\xi_r\right)^{-\gamma_r}}{\sum_{i^{'}\in\boldsymbol{I}_r}\omega_{i}^{'}\left(u_{ri^{'}}-\xi_r\right)^{-\gamma_r}}\frac{1}{\sqrt{2\pi(\sigma_r^{\mathrm{IL}})^2}}e^{-\frac{\left(y_r^{\mathrm{IL}}-\bar{y}_r^{\mathrm{IL}}\right)^2}{2\times(\sigma_r^{\mathrm{IL}})^2}}\frac{1}{\sqrt{2\pi(\sigma_{r1}^{\mathrm{TP}})^2}}e^{-\frac{\left(y_r^{\mathrm{TP}}-\bar{y}_r^{\mathrm{TP}}\right)^2}{2\times(\sigma_r^{\mathrm{TP}})^2}}dy_r^{\mathrm{IL}}dy_r^{\mathrm{TP}}. \quad ()$$

Here we include heterogeneous perception variance among passengers by considering the variance of β_r^k. Thus, the whole itinerary-choice model is actually a mixed itinerary-size weibit (MISW) model.

3. Model analysis

Three analyses are presented in this section to highlight the superiority of the MISW model from three aspects, handling the itinerary-specific perception variance, solving the itinerary overlapping problem, and considering the heterogeneous perceptions variance among passengers.

3.1. Analysis 1: The itinerary-specific perception variance

To illustrate how the MISW model addresses the itinerary-specific perception variance, a two-itinerary example is presented in Figure 2 with networks borrowed from Kitthamkesorn and Chen (2013) to model the railway system. To focus on the itinerary-specific perception variance, the variance of β_r^k is assumed to be zero in this analysis.

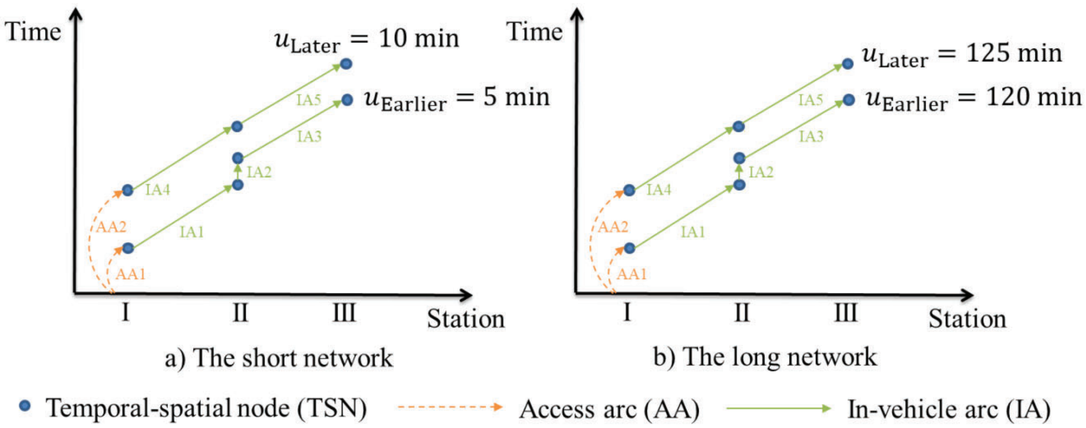


Figure 2. Two-itinerary networks.

The itineraries in the short network and the long network do not overlap, so $\omega_i = 1$. The left one is a short network, as the itinerary costs are 5 and 10 min. The right one is a long network with two larger itinerary costs (120 and 125 min). In both networks, the difference between the two itineraries is 5 min, but the cost of the later itinerary in the short network is twice as large as the earlier one, whereas it is around 4.17% larger in the long network. The likelihood of choosing the earlier itinerary should differ between these two networks; however, the logit model gives the same probability of 0.993. In contrast, the MISW model gives different probabilities (Figure 3). The probability of choosing the earlier itinerary in the short network is larger than that in the long network. The relative cost difference is used to accommodate the influence of the overall trip length (Kitthamkesorn and Chen, 2013). This example shows that the logit model treats these two cases in the same way, although trips in these two cases have very different relative lengths.

Furthermore, as the sensitivity test in Figure 3 shows, a larger ξ_r or γ_r results in a higher probability of choosing the earlier itinerary, meaning that passengers are more homogenous and more likely to choose the same itinerary which has the lower cost. This trend is more obvious in the short network.

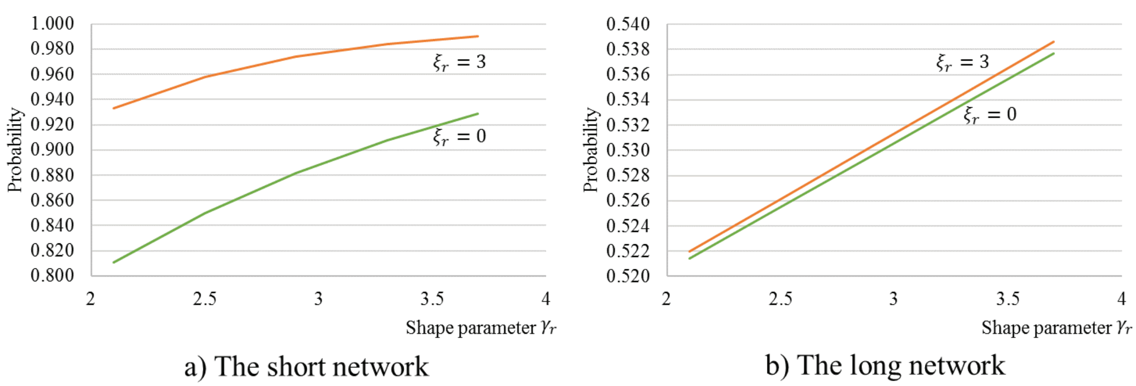


Figure 3. Probabilities of choosing the earlier itinerary in different networks.

In summary, the logit model fails to identify this route-specific perception variance because it considers only the absolute cost difference, whereas the MISW model identifies this variance because it also considers the relative cost difference.

3.2. Analysis 2: The itinerary overlapping problem

To demonstrate how the MISW model resolves the itinerary overlapping problem, a three-itinerary example is given. This example network is also borrowed from Kitthamkesorn and Chen (2013) to model the railway system shown in Figure 4. There are three itineraries from Station I to Station III, and the first two overlap in IA4. To simplify the example, only IVT is included in the itinerary cost. The itinerary costs for these three itineraries are the same: 100 min. Without loss of generality, we follow the assumption of Kitthamkesorn and Chen (2013) that all itineraries have the same variation of 0.3 and $\xi_r = 0$. According to the formula given by Kitthamkesorn and Chen (2013), $\gamma_r = 3.7$. Also, the variance of β_r^k is assumed to be zero for the focus of this analysis. Various values are used for the overlapping part to demonstrate how the itinerary choice probabilities change.

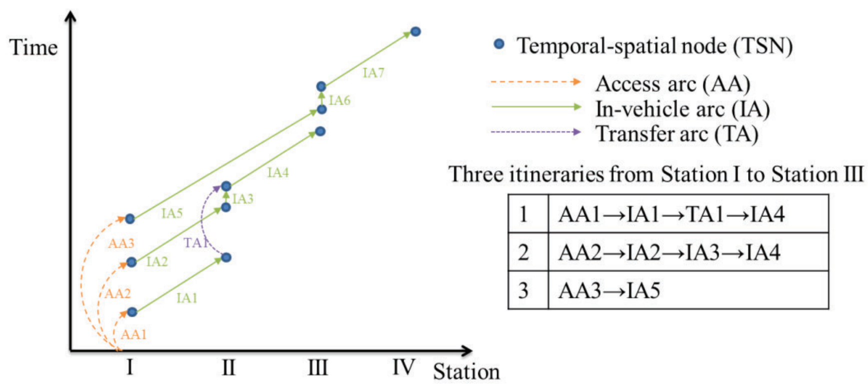


Figure 4. Three-itinerary network.

The test results are given in Figure 5.

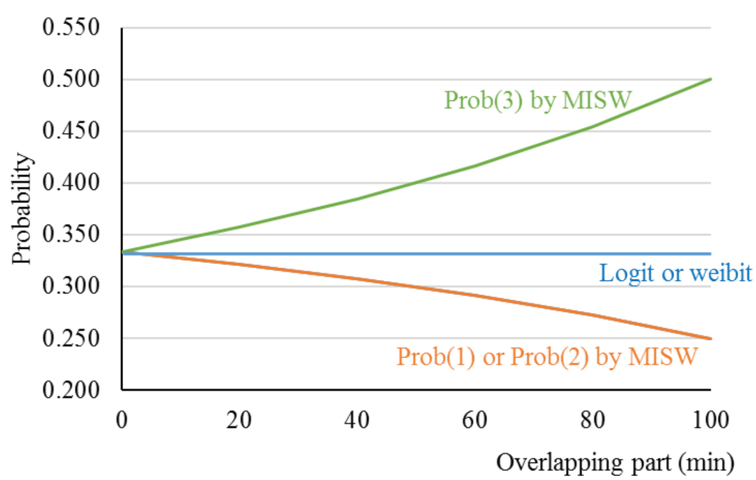
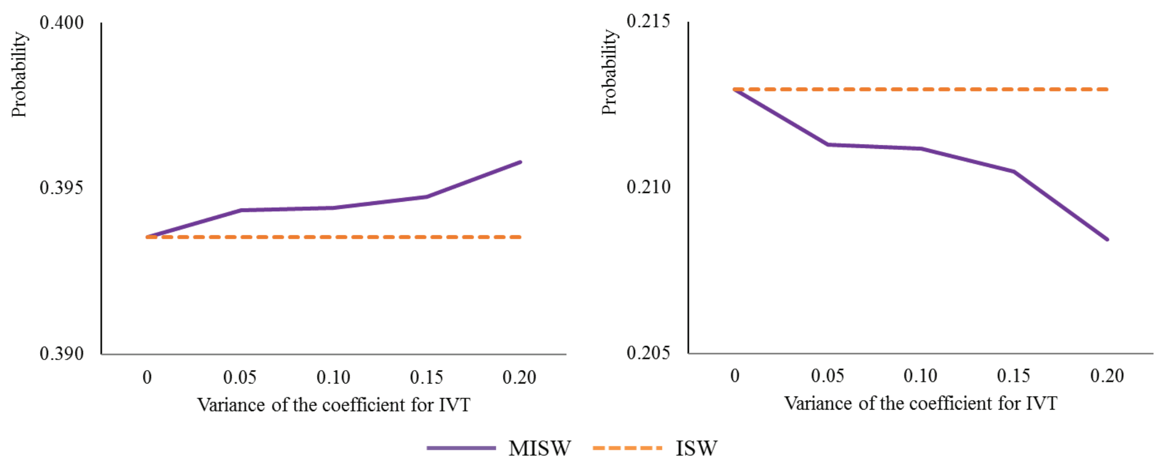


Figure 5. Probabilities of choosing itineraries.

The logit model provides the same probability for each itinerary when their itinerary costs are the same. The same result is given by the logit model regardless of how the overlapping part changes. The same result is also generated by the weibit model because both models assume that itineraries are independent; however, this may not be how passengers consider overlapping itineraries. In the extreme case, the overlapping part is infinitely close to the whole itinerary (i.e., 100 min). The first two itineraries would be regarded as one itinerary; that is, passengers would consider two itineraries with the same itinerary cost in the network. This extreme case indicates that this correlation will decrease the probability of choosing overlapping itineraries. Thus, the probability of choosing Itinerary 3 should be 0.500 rather than 0.333. In contrast, the MISW model can resolve this problem because of its itinerary-size factor. When the overlapping part is zero, the MISW model obtains the same result as the logit model and the weibit model. When the overlapping part increases, the MISW model shows that the probability of passengers choosing Itinerary 3 (Prob(3)) increases. The probability of choosing Itinerary 1 (Prob(1)) is the same as that of choosing Itinerary 2 (Prob(2)) because their IVTs and itinerary-size factors are the same. In the extreme case (i.e., when the overlapping part is 100 min), the MISW model determines that the probability of choosing Itinerary 3 is 0.500 and that the probability of choosing either Itinerary 1 or Itinerary 2 is 0.500 (Prob(1) + Prob(2) = 0.250 + 0.250 = 0.500). This example shows that adopting the itinerary-size factor to relax the independent assumption in the MISW model enables the model to consider the overlapping problem, whereas the logit and weibit model cannot do so because they assume that itineraries are independent.

3.3. Analysis 3: The heterogeneous perception variance among passengers

The three-itinerary example presented in Section 3.2 is used to show how heterogeneous perception variance among passengers affects choice probabilities. To simplify the example, only IVT and fare are included in the itinerary cost. The overlapping part of the first two itineraries is 60 min. We set the fares for these three itineraries as 50 yuan, 50 yuan, and 80 yuan. The coefficient for fares is 2 min/yuan. The mean of the coefficient for IVT is 1, and its variance ranges from 0 to 0.2. The probability of choosing Itinerary 1 is the same as that of choosing Itinerary 2 because their fares, IVTs, and itinerary-size factors are the same. The probabilities are shown in Figure 6.



a) Probabilities of choosing Itinerary 1 or probabilities of choosing Itinerary 2

b) Probabilities of choosing Itinerary 3

Figure 6. Probabilities of choosing itineraries with the variance of the coefficient for IVT.

When the variance of the coefficient for IVT is zero, the MISW model gives the same choice probabilities as the ISW model. However, when the variance changes, the MISW model gives different choice probabilities that reflect passengers' valuation of IVT, whereas the ISW model gives the same result regardless of how the variance of the coefficient for IVT changes. When the variance increases, passengers value IVT very differently: more passengers may be insensitive to IVT, and fares become a more important influence factor; hence, they are more likely to choose low-fare itineraries. As shown in Figure 6, the MISW model gives a lower choice probability for the high-fare itinerary (Itinerary 3) to reflect the increasing variance. This trend indicates that the MISW model can capture the heterogeneous perception variance among passengers by including random parameters, whereas the ISW model cannot do so because it assumes that the variance of parameters is zero.

4. The application in train scheduling

In this study, *a schedule plan details the number of scheduled trains, departure and arrival times, stopping pattern, and station-track allocation of trains in a railway network.* In this section, we integrate the MISW model with train scheduling. The integrated model is called as the passenger-oriented train-scheduling model. This study attempts to improve the train service by reducing passenger trip costs, and the objective of this train-scheduling model is to minimize the total itinerary cost C^{P} for passengers. This train-scheduling model can be described as follows:

$$\min C^{\mathrm{P}},$$

$$\text{s.t. } \mathbb{P} \text{ and } \mathbb{S},$$

where $\mathbb{P}$ is the set of passenger-itinerary choice constraints based on the MISW model to determine which itineraries are chosen and how likely the choices are to be made, and $\mathbb{S}$ is the set of train-scheduling constraints in consideration of the safety headway, running time, dwell time, station capacity, whether a train is scheduled, and which station should be skipped. $\mathbb{S}$ is built based on the pre-known line plan. *A line is a set of trains with the same passing stations, minimum running time, dwell time, and operating time-window.* A line plan of line l includes passing stations, maximum number of trains, the earliest departure time at origin terminals, the latest arrival time at destination terminals, minimum dwell time, and minimum running time between stations. The information about which station a train stops at or passes through is not given. Trains on the same line may have different stopping patterns. For example, trains n and n' belong to the same line. Train n may stop at station s, whereas train n' may pass station s without stopping because passengers on train n' do not board or alight there. Besides, it is possible that some trains will not be arranged in the final timetable. Our train-scheduling model determines how many trains are operated according to passenger needs. If train n in the given line plan is used by a negligible number of passengers, it will be canceled. Furthermore, the given line plans already consider the operating budget, so that a direct train for each OD pair may not be possible. In summary, we assume that the line plan is given.

However, we consider modifying the given line plan by cancelling trains or setting some trains to skip some stations based on the passengers' need, which is called line planning in the study.

To reduce the computational complexity, we transform the passenger-oriented train-scheduling model into a simultaneous optimization model that handles two optimization problems simultaneously, as follows (Equations () and ()):

$$\text{Passenger-oriented schedule-plan generation: } \boldsymbol{S} = \mathrm{Q}(\boldsymbol{I}); \qquad ()$$

$$\text{Schedule-based stochastic passenger assignment: } \boldsymbol{I} = \mathrm{G}(\boldsymbol{S}). \qquad ()$$

The first optimization problem (Equation ()) is used to generate the most convenient schedule plan $\boldsymbol{S}$ via model Q, based on the input information about passenger assignment $\boldsymbol{I}$. The second optimization problem (Equation ()) uses the model G to find the passenger assignments $\boldsymbol{I}$ that are most likely to be generated if the schedule plan $\boldsymbol{S}$ is used. These two problems can be regarded as a fixed-point problem, as follows:

$$\boldsymbol{S} = \mathrm{Q}(\mathrm{G}(\boldsymbol{S})). \qquad ()$$

The solution of the fixed-point problem may not be globally optimal because it may become trapped in a local optimum. However, it is an acceptable strategy to reduce the complexity of the whole problem and thus find one of the good, consistent, locally optimal solutions for practical use when it is very difficult to obtain the global optimum for the problem.

An iterative method is provided to solve the fixed-point problem, because iterations are a means by which a solution path can be used with a higher chance of identifying a reasonable, feasible solution that is likely to improve the performance of the railway system. The flow chart of this method is shown in Figure 7 and the output of each step is marked in green.

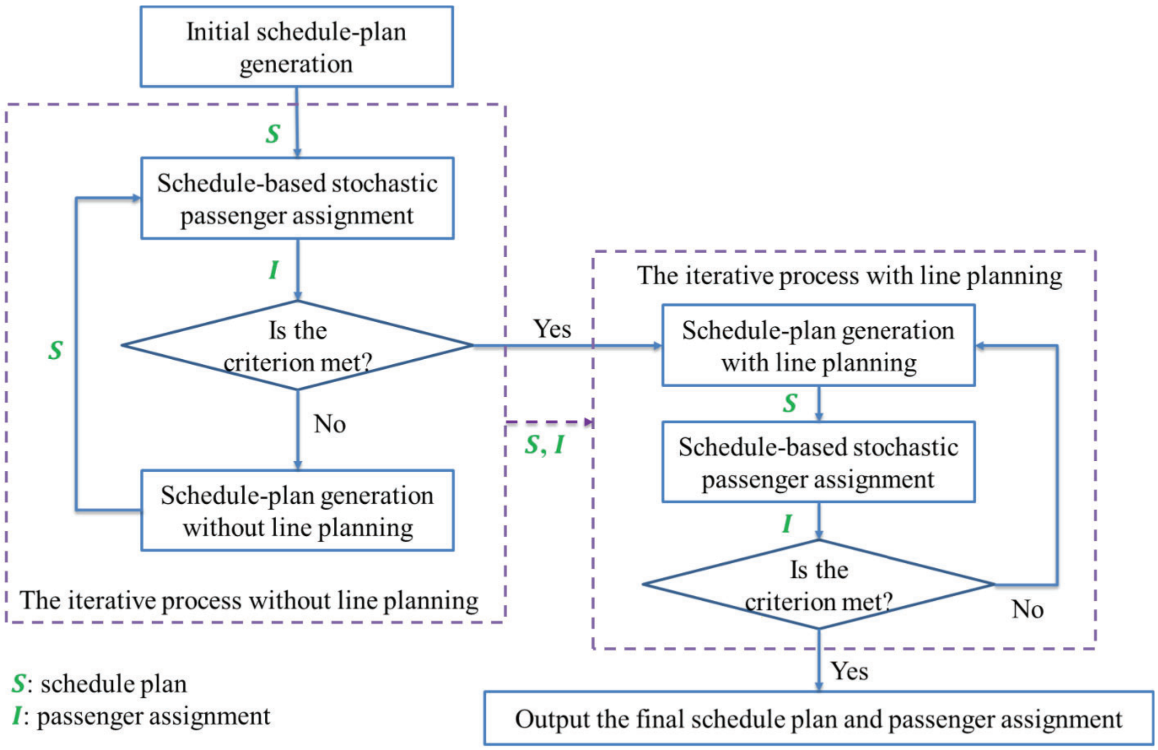


Figure 7. Flow chart of iterative method.

As the flow chart shows, the iterative method can capture the interactions between a schedule plan and passenger-itinerary choices. First, we run the initial schedule-plan generation (ISPG) to find the initial schedule plan (details of the ISPG are presented in Section 4.1). The line plan of the ISPG assumes that all trains are used and that each train stops at every station on its route. The ISPG only needs to confirm arrival and departure times and a station-track allocation.

The schedule plan is then adjusted via two iterative processes. In both iterative processes, the schedule-based stochastic passenger assignment (SSPA) indicates the direction of the adjustment of the schedule plan, whereas the SSPA assign passengers according to the schedule plan outputted by the schedule-plan generation (SPG) (the details of the SSPA and SPG will be introduced in Sections 4.2 and 4.3, respectively). The difference between these two iterative processes is whether the SPG considers line planning (i.e., canceling trains and setting some trains to skip some stations). In the iterative process without line planning, the line plan is fixed following the assumption of the ISPG, whereas the iterative process with line planning can change the line plan by canceling the unused trains and allowing trains to skip unused stations in the SPG. The improvement of the line plan can provide more time and space for scheduling and lower operating costs. For example, if a train is set to skip some stations, the running time of this train can be shortened so other trains may be arranged to use the station time slot to which it was originally allocated. Operating costs are also lowered by the fact that a train does not expend energy decelerating into and accelerating out of a station that it skips.

The criterion for stopping any one of these two iterative processes is the gap allowance between two successive calculated total passenger itinerary costs or a limited number of SPGs. After the criterion is met, the iterative process without line planning outputs a feasible schedule plan and passenger assignment to the iterative process with line planning for further adjustment, with the possibility of the line plan being adjusted. When the iterative process with line planning ends, the final schedule plan and passenger assignments are obtained.

The iterative method reduces the complexity of the original problem to improve the computation efficiency, but at the expense of optimality, as the iterative method tends to find a local optimum rather than a global optimum. The initial schedule plan, as the starting point, critically affects which local optimum is found, because the passenger assignment, i.e., the basis for guiding improvement, is generated based on the previous schedule plan. To moderate this weakness, we suggest generating multiple initial schedule plans to facilitate the finding of a good solution. A test in Section 5.1.2c is given to demonstrate the workability of this suggestion.

The following parts of this section detail the ISPG, SPG, and SSPA.

4.1. Initial schedule-plan generation

The aim of the ISPG is to find a feasible starting point for adjustment. The ISPG assumes that all trains are used and stop at every station on their routes, and that at least one feasible schedule plan exists. There is usually more than one feasible solution, and finding a good starting point helps to achieve a high-quality final result. In this study, we recommend two kinds of objective function F^{I}:

1) Minimize the total train running time

$$F^{\mathrm{I}} = \sum_{n \in \boldsymbol{T}} (t^{\mathrm{A}}_{ns^{\mathrm{t2}}_{n}} - t^{\mathrm{D}}_{ns^{\mathrm{t1}}_{n}}), \qquad ()$$

where t^{D}_{ns} is the departure time of train n at station s, t^{A}_{ns} is the arrival time of train n at station s, s^{t1}_{n} is the original terminal of train n, s^{t2}_{n} is the destination terminal of train n, and $\boldsymbol{T}$ is the set for all trains. This objective function has an obvious advantage in that it tries to reduce the IVT.

2) Minimize the schedule difference

$$F^{\mathrm{I}} = \beta^{+} \times \sum_{n \in \boldsymbol{T}} \sum_{r \in \boldsymbol{R}} \Delta t^{+}_{nr} \times D_r + \beta^{-} \times \sum_{n \in \boldsymbol{T}} \sum_{r \in \boldsymbol{R}} \Delta t^{-}_{nr} \times D_r. \qquad ()$$

The first part is the total positive schedule difference of all OD pairs, whereas the second part is the total negative schedule difference. Δt^{+}_{nr} is the schedule difference of OD pair r associated with train n when the scheduled time is later than the ideal time, β^{+} is the associated penalty, Δt^{-}_{nr} is the schedule difference when the scheduled time is earlier than the ideal time, and β^{-} is the associated penalty. Δt^{+}_{nr} and Δt^{-}_{nr} are obtained by the following formula:

$$t^{\mathrm{D}}_{ns^{\mathrm{t1}}_{n}} - t^{\mathrm{I}}_{nr} - \Delta t^{+}_{nr} + \Delta t^{-}_{nr} = 0, \qquad ()$$

$$\Delta t_{nr}^{+} \geq 0, \qquad ()$$

$$\Delta t_{nr}^{-} \geq 0\ , \qquad ()$$

where t_{nr}^{I} is the ideal departure time of train n at the original terminal s_n^{t1} for OD pair $r \in \boldsymbol{R}$. Constraint () connects the time differences (Δt_{nr}^{+} and Δt_{nr}^{-}) with the scheduled and ideal times of train n. Constraints () and () ensure that Δt_{nr}^{+} and Δt_{nr}^{-} are non-negative.

The strength of the second objective function is that it offers a standard for planning the initial schedule plan. However, it requires extra effort to obtain the ideal time. Here, we provide an approach for determination of the ideal time. We first build up a network (see Figure 8) in which nodes represent stations and links are for trains.

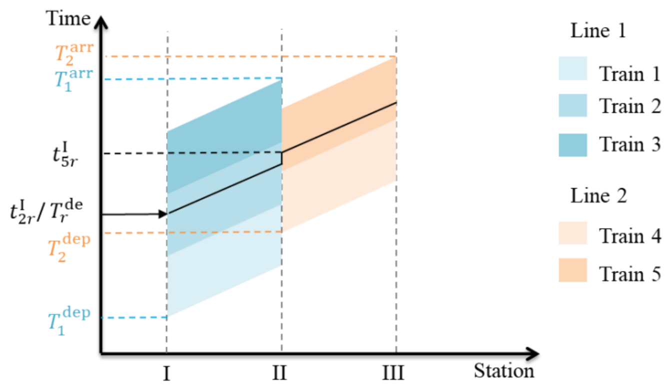


Figure 8. An example of finding the ideal departure time.

Each link has an operating time-window [T_l^{dep},T_l^{arr}] within which trains are allowed to run, where T_l^{dep} is the earliest time of departure from origin terminal for line l and T_l^{arr} is the latest time of arrival at destination terminal for line l. For example, the operating time-window of Line 1 is [T_1^{dep},T_1^{arr}], and it contains three trains. We assume that the three trains are evenly distributed. Hence, the operating time-window is divided into three parts indicated by various shades of blue. Each train is allowed to run in its respective operating time-window. Dijkstra's algorithm is then used to determine the shortest itinerary for each OD pair and deduce the ideal times. An example marked in black is shown in Figure 8. The preferred departure time T_r^{de} is within the operating time-window of Train 2, so passengers take Train 2 to leave Station I and arrive at Station II. We assume that the transfer time equals the maximum walking time for a transfer at stations. In this example, passengers can only transfer to Train 5 because their arrival time is outside the operating time-window of Train 4. Hence, the ideal time associated with Train 2 is $t_{2r}^{\mathrm{I}} = T_r^{\mathrm{de}}$, and that associated with Train 5 is t_{5r}^{I}.

In this study, we mainly use the second objective function for the ISPG because a better starting point can be found with some guidelines. The constraints include departure/arrival time constraints, headway constraints, station-track assignment constraints, dwell time constraints, and running time constraints. They are given in Appendix A.1.

Directly solving the ISPG for a practical network is time-consuming. Hence, we develop a heuristic based on the decomposition approach. The core idea of this approach is to create a schedule plan for the lines one-by-one. That is, after a feasible schedule plan is created for all lines, we improve it one line at a time: i.e., when the timetable of a line is being created, variables and constraints related to other lines are fixed, such that we improve the first line while the timetables of the other lines are fixed. Then we repeat the process for the second line, the third line, …, and the last line. The improvement that runs from the first line to the last line is called a completed adjustment. Several adjustments are made, and either a better schedule plan or an unimproved schedule plan is obtained after each adjustment. If the objective difference between two successive adjustments falls within the gap allowance, the heuristic stops, and the local optimum of the ISPG is found.

4.2. Schedule-plan generation

The SPG aims to reduce the total itinerary cost within the train-scheduling constraints. The SPG is formulated as follows:

$$\min F^{\mathrm{R}} = \sum_{r \in \boldsymbol{R}} \sum_{i \in \boldsymbol{I}_r} \mathrm{Prob}_r(i) \times D_r \times u_{ri},$$

$$\text{s.t. Eq. (), } \mathbb{S}, \text{ and } \mathbb{P},$$

where $\mathbb{S}$ includes the formulas stated in Appendix A.1 and the set of itinerary constraints $\mathbb{P}$ introduced in Appendix A.2 is built based on the previous SSPA. Passenger itinerary routing and passenger flow on them are given by the previous SSPA and acted as the input to the SPG. The itinerary cost u_{ri} is calculated by Eq. (), where the number of transfers, access/egress time, and fare are provided by the previous SSPA, whereas the IVT, wait time, and walking time are affected by the new schedule plan and determined by the respective equations included in $\mathbb{P}$.

We first order the lines in descending order based on the number of passengers assigned to each, and we then adopt a heuristic based on the decomposition approach (introduced in Section 4.1) to enhance computational efficiency. In the first iteration, a feasible schedule plan is generated by the ISPG, whereas in another iteration, a feasible schedule plan is given by the previous SPG. Therefore, the heuristic is used to improve the schedule plan to meet the objective. Similarly, the heuristic is used to perform several adjustments until the objective difference between two successive adjustments falls within the gap allowance, and each adjustment improves the schedule plan for each line, separately.

4.3. Schedule-based stochastic passenger assignment

In this section, we focus on the SSPA. First, we introduce the itinerary-search approach for the generation of itineraries for the SSPA. The approach consists of four stages, with Stages 2 and 3 representing techniques suggested by Tong and Richardson (1984) for downsizing a network to accelerate an itinerary search. The approach is described as follows.

1) Stage 1: Study network construction.
   A network is built to enumerate itineraries for OD pair r. In this network, nodes represent stations or zones, while links connecting nodes include train links between stations, walking links between platforms, and zone links between a zone and a station.
2) Stage 2: Forward pass of quickest itinerary.
   The quickest itineraries from an origin node to other nodes in the network are determined by Dijkstra's algorithm. The arrival time at a node via the quickest itinerary is the earliest arrival time from the origin node to this node. Passengers cannot board trains departing earlier than the earliest arrival time at this node; hence these trains can be ignored during the enumeration search and the earliest arrival time is set as the lower boundary of the feasible time-window at this node.
3) Stage 3: Backward pass of quickest itinerary.
   A reversed network is built, in which all links are reversed in direction and the timetables of trains are also reversed. The departure time of a train in the reversed network is equal to M minus the original departure time, where M is a sufficiently large number. As OD pair r is limited by the passengers' activities at the destination, it has the requirement of the latest arrival time at the destination T_r^{ar}. If OD pair r has no such requirement, T_r^{ar} is the arrival time of the last train at the destination station.
   T_r^{ar} is transformed to be the departure time from the destination in the reversed network, and the quickest itineraries from a destination node to other nodes are determined by Dijkstra's algorithm. The arrival time at a node via the quickest itinerary is transformed to be the normal time, which is set as the latest departure time from this node to the destination node. Passengers cannot board trains departing later than the latest departure time at this node; hence these trains can be ignored during the enumeration search and the latest departure time is set as the upper boundary of the feasible time window at this node.
4) Stage 4: Itinerary enumeration.
   If the earliest arrival time at a node is later than the latest departure time from a node, this node will not be used by passengers. Thus, such a node can be deleted from the network. After deleting all unused nodes, itinerary enumeration is performed within the feasible time-windows. If no feasible itinerary for OD pair r is found, a penalty ($\dot{u}_r$) is added to the total itinerary cost and the passengers of OD pair r will not be assigned.

After the itineraries are determined, the MISW model is used to assign passengers. The number of assigned passengers on itinerary i equals $\mathrm{Prob}_r(i) \times D_r$. This number acts as a weight coefficient in the objective function that minimizes the passenger cost, thus influencing the priority of scheduling trains during the SPG. If this number is too small, itinerary i is unlikely to be used by passengers, and therefore it is unnecessary to adjust timetables for such low demands. A criterion, the minimum passenger volume on an itinerary, is set to handle this. If the passenger volume on itinerary i is less than the minimum-flow criterion, the SPG and line planning will not consider itinerary i. As we do not set constraints to suit such low demands, itinerary i may no longer exist in the optimized timetable. In this case, the next SSPA will not assign any passengers to itinerary i.

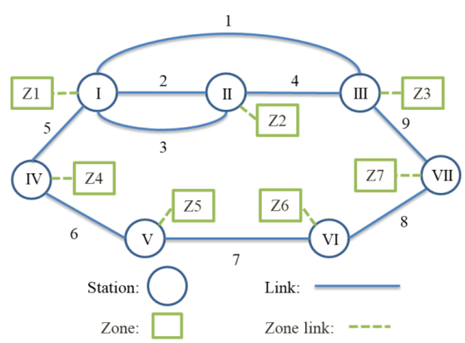
Alternatively, itinerary i may still be available but will become an itinerary with a longer travel time. If this is the case, then the next SSPA will assign even fewer passengers to this itinerary; consequently, after a few iterations, this itinerary will be naturally eliminated when the solution converges. If the travel time of this itinerary decreases while the timetable is optimized for other OD demands, there will be a few passengers who will take advantage of this timetable to travel to their destination. In the converged solution, there will be an equilibrium between the SSPA and SGP. The minimum-flow criterion value can be determined by operators, and in the examples of this study, the minimum-flow criterion value is one passenger.

5. Numerical examples

We test the train-scheduling model using the MISW model in a small network (Section 5.1) and a practical network (Section 5.2). The programming language used is C#, and the computation for the examples described below is performed on a personal computer (PC) with a 3.60 GHz i7-4790U central processing unit, eight cores, and 16 GB RAM. The PC runs Windows 7 Enterprise and has a 64-bit operating system. CPLEX Studio 12.7 is used to tackle scheduling problems in the examples.

5.1. Small example

We set up a loop network (Figure 9) to analyze the model and solution method.



Link	Minimum running time (min)
1	30
2	20
3	18
4	15
5	20
6	25
7	20
8	25
9	60

Figure 9. Abstracted-loop network for the small example.

The detailed setting of this example is listed in Appendix B. The iterative method ends after 24.78 s when it experiences one ISPG and five SPGs to reach the convergence. The total passenger itinerary cost of this example decreases from 15,148.91 min to 14,132.66 min. We provide a schedule plan of the final SPG, SPG (5), in Appendix C. The passenger assignments based on the schedule plans of the ISPG, SPG (1), and SPG (5) are presented in Table 2.

Table 2. Passenger assignments based on the schedule plans of the ISPG, SPG (1), and SPG (5).

	OD	Itinerary (Trains)	Volume (passengers)	Itinerary cost (min)	IVT (min)	Wait time in	Wait time at stations	Walking time	Fare (yuan)

						zones (min)	(min)	(min)	
ISPG	1	9	19.949	35.50	20	-	-	-	10
		4	0.045	190.45	20	180	10	-	9
		10	0.004	364.50	175	15	-	-	115
		2→10	0.003	368.05	45	30	110	5	31
	2	1	14.186	80.00	30	-	-	-	25
		9	4.705	108.00	37	30	-	-	16
		2	1.082	161.00	30	60	2	-	25
		10	0.027	434.50	158	45	-	-	109
	3	9	16.610	238.80	99	30	-	-	61
		10	25.426	249.00	96	45	-	-	60
		1→9	7.965	289.20	90	-	39	-	70
SPG (1)	1	9	19.967	33.50	18	-	-	-	10
		4	0.033	190.45	20	180	10	-	9
	2	1	8.481	80.00	30	-	-	-	25
		2	6.474	86.00	30	-	4	-	25
		9	2.746	106.00	35	30	-	-	16
		3→9	2.273	112.00	35	-	30	-	15
		10	0.025	385.00	158	-	6	-	109
	3	10	23.473	214.80	96	-	6	-	60
		9	9.643	236.80	97	30	-	-	61
		3→9	6.466	262.00	97	-	30	-	60
		1→9	5.243	285.60	90	-	37	-	70
		2→9	5.175	286.60	90	-	32	5	70
SPG (5)	1	9	20.000	33.50	18	-	-	-	10
	2	1	8.570	80.00	30	-	-	-	25
		2	6.678	86.00	30	-	4	-	25
		3→9	2.391	105.50	42	15	8	-	15
		9	2.362	106.00	35	30	-	-	16
	3	10	23.893	208.80	90	-	6	-	60
		9	8.835	236.80	97	30	-	-	61
		3→9	7.728	244.40	104	15	8	-	60
		1→9	4.803	285.60	90	-	37	-	70
		2→9	4.741	286.60	90	-	32	5	70

Section 5.1.1 investigates the capability of our model, while Section 5.1.2 discusses the workability of the proposed solution method.

5.1.1. Model analysis

a). Passenger service improvement

Although fares and IVT are a major part of the total itinerary cost, most IVTs cannot be reduced because of the maximum train speed limit, and fares are fixed. The ticket prices of trains given by the railway companies are preset, and the fare of an itinerary is calculated according to the ticket prices of the trains used. Because the previously generated itineraries still use these trains to reach these stations, the fares for them do not change. Hence, the change in the total itinerary cost is mainly caused by the reductions in wait time and transfer time. Compared with the total value, this kind of reduction does not seem significant even though wait time at origin stations and transfer time decrease substantially. For example, the wait time at the origin station of OD pair 2's itinerary which uses Train 2 and the transfer time of OD pair 3's itinerary (Train 3→Train 9) are reduced by 93.55% and 60.00%, respectively. In SPG (5), the line planning has been considered. Five trains are sufficient to support the passengers' needs, and two of these skip some stations because no itineraries used by more than one passenger require them. This schedule reduces operating costs and benefits passengers by shifting the train schedule to a better time slot to reduce travel time. For example, OD pair 3 using Train 10 have a 6-min reduction in IVT because Train 10 skips Stations IV, V, and VI.

Moreover, in SPG (5), two of the OD pairs have more than one itinerary. The availability of alternative itineraries provides more options to passengers. For example, the passengers in OD pair 2 have four possible itineraries. If Link 1 is blocked and Trains 1 and 2 are canceled, passengers can still use the other itineraries (that do not include Trains 1 and 2) to reach their destination. However, if only the itinerary with the lowest cost (i.e., directing use Train 1) is considered in train scheduling, passengers may suffer a long wait time rather than 23 or 30 min because the other itineraries are not considered when improving the train schedule plan.

b). Distributions of passenger flow on train

The MISW model calculates the probabilities of itinerary choices and assign passengers on trains. We select Train 9 as an example to show the distributions of passenger flow for the ISPG and SPG (5) (Figure 10). Because passengers with different OD pairs board and alight at different stations along the route of Train 9, the passenger flow varies during the operating period. The origin and destination terminal of a train is assumed to be fixed in this study; hence Train 9 runs on Links 5-8 even though it includes no passengers on these links. This indicates a resource waste that should be examined in future work. Also, due to the difference between the initial schedule plan and the final schedule plan, the itinerary cost changes and causes a change in the passenger flow from each OD pair. For instance, a decrease appears on Link 3, because more itineraries become feasible after some iterations, like OD pair 2's itinerary (Train 3→Train 9), and it is possible that some passengers will turn to select these newly-found itineraries.

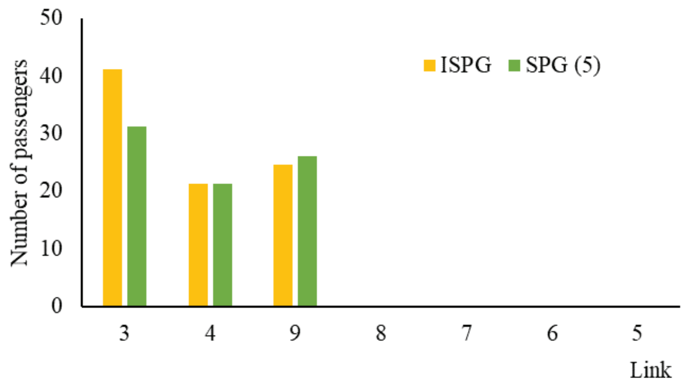


Figure 10. Distributions of passenger flow on Train 9 for the ISPG and SPG (5).

c). Train operation analysis

First, we analyze the track allocation. If a station has sufficient tracks, it can contain all trains that need to use the station at the same time. However, stations may not always have sufficient tracks. For example, a station with two tracks may need to serve three trains travelling in the same direction at the same time. This is an impossible task, thus one of these trains has to wait outside the station or be rescheduled to arrive at a later time, meaning that the serving time of these trains becomes longer. Similarly, Station I has two tracks. In the initial schedule plan, Trains 1 and 2 depart at 8:00 and 8:04, respectively at Station I and they stop at different tracks for a while (for instance, 2 min) to let passengers board before the train departure. Hence, Trains 1 and 2 stop at Station I during 7:58-8:00 and 8:02-8:04, respectively. However, if Station I has one track, Trains 1 and 2 need to share the same track and the train headway of using the same track should be not less than 4 min. Therefore, Train 2 has to delay its departure. Such cases illustrate the challenge of finding a suitable track allocation method, which is one reason that track allocation is included in our model.

Second, the assumption that the train capacity is sufficient may lead to a smaller number of scheduled trains. Although this example does not need to worry about the train capacity problem (we only have 90 passengers in total), it may be an issue for other cases. It can be handled by regrouping passengers. Initially, we group passengers with the same origin and destination within a 1-hour period to be an OD pair and assume that they depart at the end of this period. We then change 1 hour to 30 minutes, and the number of passengers in each OD pair is halved if we assume that demand is evenly distributed. Thus, we double the number of OD pairs, but the total number of passengers remains the same. We rerun the iterative method to find a schedule plan, and the number of scheduled trains increases from 5 to 6. We have shown a simple approach to reduce the impact of the capacity assumption. A more mature approach to handle it directly will be studied in the future.

5.1.2. Solution method analysis

a). Analysis of the SSPA

Firstly, the results of SSPAs shown in Table 2 support the discussion about the minimum-flow criterion in Section 4.3. There are four possible itineraries for OD pair 1 in the ISPG, but only the passenger flow of the first itinerary meets the minimum-flow criterion, i.e., one passenger. Hence, SPG (1) only considers the first itinerary. Train 10 will be shifted to an earlier time slot because of OD pair 3, therefore, the last two itineraries no longer exist in SPG (1). Train 4 will continue to run on the same time slot; thus, the second itinerary remains feasible. However, the cost of the first itinerary decreases due to the reduction of the IVT, such that more passengers are assigned to the first itinerary and fewer passengers are assigned to the second itinerary. Finally, the second itinerary is eliminated in SPG (5) because Train 4 is canceled, due to its having too few onboard passengers.

Secondly, we compare the SSPA and the deterministic passenger assignment which only considers optimal itineraries and does not load passengers to other possible itineraries. The final schedule plans of these two approaches are quite different. Train scheduling using deterministic passenger assignment ultimately results in a plan with three trains (Trains 3, 9, and 10) in operation, whereas the SSPA results in five trains. Hence, the number of possible itinerary choices is reduced if deterministic assignment is used. For instance, using deterministic assignment, OD pair 2 can only use Train 9 or Train 10 with a long travel time, because Trains 1 and 2 are cancelled in the train scheduling and Train 3 is set to arrive at Station II later than Train 9 for OD pair 1. If Link 3 is blocked, all passengers of OD pair 2 will be affected.

Thirdly, we conduct a test to analyze the influence of using different itinerary-choice models for the SSPA. The SSPA adopts different itinerary choice models, i.e., the MISW model and the mixed weibit (MW) model, where the latter does not consider the overlapping problem. The found itineraries, as shown in Table 2 and Table 3, use the same trains but run them based on different improved schedule plans (see Appendix C). As the analysis in Section 3.2 indicates, the model without the itinerary-size factor tends to heighten the chances of selecting overlapping itineraries.

Table 3. Passenger assignment generated by the MW model.

OD	Itinerary (Trains)	Volume (passengers)	Itinerary cost (min)	IVT (min)	Wait time in zones (min)	Wait time at stations (min)	Walking time (min)	Fare (yuan)
1	9	20.000	33.50	18	-	-	-	10
2	1	10.047	80.00	30	-	-	-	25
	3→9	3.623	105.50	42	15	8	-	15
	9	3.572	106.00	35	30	-	-	16
	2	2.758	113.50	44	15	-	-	25
3	10	19.338	198.00	90	-	-	-	60
	9	10.102	236.80	97	30	-	-	61
	3→9	8.835	244.40	104	15	8	-	60
	2→9	6.803	262.40	104	15	8	-	70
	1→9	4.922	285.60	90	-	37	-	70

The MW model therefore assigns more passengers to overlapping itineraries (i.e., those using Train 9). Thus, in this case, the SPG gives greater priority to those passengers and arranges Trains 1, 2, and 9 to share the same platform at Station III to smooth the transfer process for OD pair 3. In contrast, the MISW model results in a walking transfer in one itinerary of OD pair 3 (Train 2→Train 9) to reduce the travel time for OD pair 2. Furthermore, Trains 2 and 10 are adjusted accordingly to reduce the itinerary cost. For example, the running time of Train 2 increases so that those passengers of OD pair 3 who use Train 2 and then transfer to Train 9 can sit on the train rather than stand at the platform to wait, as long as the arrival time of Train 9 is not altered to meet other OD demands. However, the passengers of OD pair 2 who use Train 2 suffer from a longer journey time, compared with the case using the MISW model.

Hence, different itinerary choice models in the SSPA give different priorities with respect to adjusting schedule plans, because the SPG tries to balance the needs of different passengers according to the level of demand, and obtains a system optimum when faced by constraints on the supply side (such as the number of tracks in a station and the headway between trains). If the SSPA fails to reflect the actual situation, the priority will be inappropriately assigned, and the schedule improvement will deliver a reduced benefit. That is, the itineraries that should be improved will be sacrificed for those that are actually less demanding. As the analysis in Section 3 shows, the MISW model is more likely to provide superior priority assignment. Thus, we should adopt the MISW model in train scheduling.

b). Analysis of the SPG

To ensure that each solution is the local optimum, the ISPG and SPG are designed to have several adjustments and each adjustment improves the schedule plan line by line as stated in Section 4. The number next to each point in Figure 11 shows the number of adjustments in this ISPG or SPG.

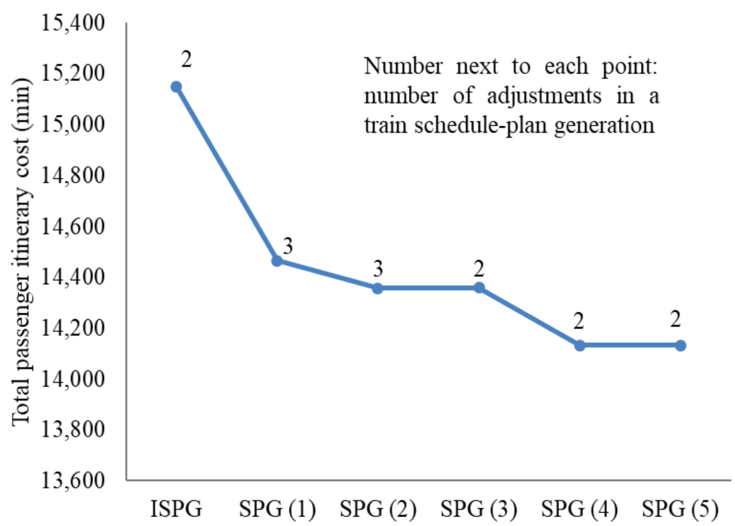


Figure 11. Change in the total passenger itinerary cost.

For instance, based on the initial schedule plan, the first stochastic passenger assignment is found. Then, SPG (1) runs to minimize the total passenger itinerary cost, and it involves three adjustments to obtain the local optimum. In the first adjustment, Line L1 carries the most passengers, and thus it is adjusted first. However, limited by the fixed timetables of the other lines (e.g., Line D2), Train 9 of Line L1 is not adjusted to the ideal situation. After the first adjustment, other trains (e.g., Train 3) are arranged to other time slots. Therefore, in the second adjustment the ideal arrangement for Train 9 can be made. As no further improvement is shown in the third adjustment, the iterative adjustment in SPG (1) ends. This case indicates the need for iterative adjustment.

In each adjustment, the decomposition approach is used; thus, ILPs for train scheduling of each line are small-scale. The average numbers of variables and constraints for an ILP are 140 and 267, respectively. These ILPs can be easily solved by CPLEX. The computation time is about 0.32 s for an ILP.

c). Analysis of solution quality

We can see from Figure 11 that the total itinerary cost continues to decrease. The gap of the total passenger itinerary costs between SPG (2) and SPG (3) is 0.01% falling within the gap allowance (0.25%). Hence, in SPG (4), the iterative method begins to cancel unused trains and set some trains to skip some stations. SPG (5) cannot further reduce total passenger itinerary cost. Therefore, the iterative method converges at this point and ends.

The convergence speed seems quick, and there are two possible reasons:

1) More than one adjustment exists in each SPG. In all, fourteen train-scheduling adjustments are required.
2) Line setting. For example, the operating time-window may be small, and there may not be many feasible solutions for searching.

Additionally, the iterative method tends to fall into a local optimum; we suggest the use of various starting points to avoid this. A test is performed to see how the initial schedule plan influences the solution quality. However, we note that the global optimum is difficult to find, because the passenger-oriented train-scheduling problem, which includes the itinerary search and the itinerary choice probability calculation, differs from the traditional scheduling problem, which can be formulated as linear. Thus, to analyze the difference between the global optimum and the solution found by our iterative method, the complexity must be reduced. In this analysis, we assume that there is no randomness and that the passengers select only the lowest-cost itinerary. In this example, the first scenario is the global optimum found by the CPLEX. We examine five initial schedule plans: three of these (Scenarios 2, 3, and 4) are generated by minimizing the schedule difference, whereas the other two (Scenarios 5 and 6) are generated by minimizing the train time. The comparison is shown in Table 4.

Table 4. Comparison of scenarios with different initial schedule plans.

Scenario	Final objective value (min)	Gap (%)
1	11,879	-
2	11,879	-
3	12,170	2.45
4	14,010	17.94
5	12,170	2.45
6	12,219	2.86

The gap in Table 4 is calculated by Equation (), as follows:

$$\text{Gap} = \frac{\text{Final objective value of the scenario - Global optimum}}{\text{Global optimum}} \times 100\%. \qquad ()$$

The second scenario finds the same objective as the global optimum, and most of the gaps are within 3%. These results show that various initial schedule plans lead to various final schedule plans. However, although all solutions are locally optimal, some are not satisfactory. To avoid this weakness of the solution approach and find a good-quality solution, if possible, we should investigate various initial schedule plans. Additionally, although minimizing the schedule difference in the ISPG is not a universally superior approach (as in Scenario 4), it is much more likely to find a more satisfactory solution (as in Scenarios 2 and 3) than minimizing the total train running time in the ISPG (i.e., Scenarios 5 and 6).

Besides, this test is further developed to analyze the workability of the iterative method when the problem size increases. The three OD pairs are retained, but the number of trains in each line increases by multiplying by a scale factor that varies from 1 to 5; hence, the total number of trains in the network increases from 10 to 50, and therefore the numbers of constraints and variables in the model also increase (shown in Table 5). We compare the objective values found by the iterative method with the global optimum found by CPLEX. The first three cases show no difference between the objectives, and the largest difference between objective values is 2.86%. The results indicate that the iterative method can find a good-quality solution.

Table 5. Comparison of the objectives as the scale factor increases.

Scale factor	Total number of trains in the network	Number of variables in the model	Number of constraints in the model	Objective found by CPLEX (min)	Objective found by the iterative method (min)	Gap (%)
1	10	1,572	3,354	11,879	11,879	0.00
2	20	5,259	13,191	11,870	11,870	0.00
3	30	11,066	29,568	11,870	11,870	0.00
4	40	18,993	52,485	11,870	12,210	2.86
5	50	29,040	81,942	11,870	12,210	2.86

We then compare the methods' computation times, as shown in Figure 12, where the x-axis represents the scale factor, and the y-axis represents the computation time required. As the scale factor increases, the computation time for CPLEX increases exponentially. The general trend of the computation time of the iterative method is also to increase. However, all computation times of the iterative method are within 19 s, whereas the longest computation time for CPLEX is approximately 193 s, more than 10 times that of the iterative method. The computation time required by CPLEX also increases much faster than that of the iterative method; hence, the latter has a computational advantage and can handle more complex problems than CPLEX. In summary, this analysis demonstrates that the iterative method can find a good-quality solution more rapidly than CPLEX.

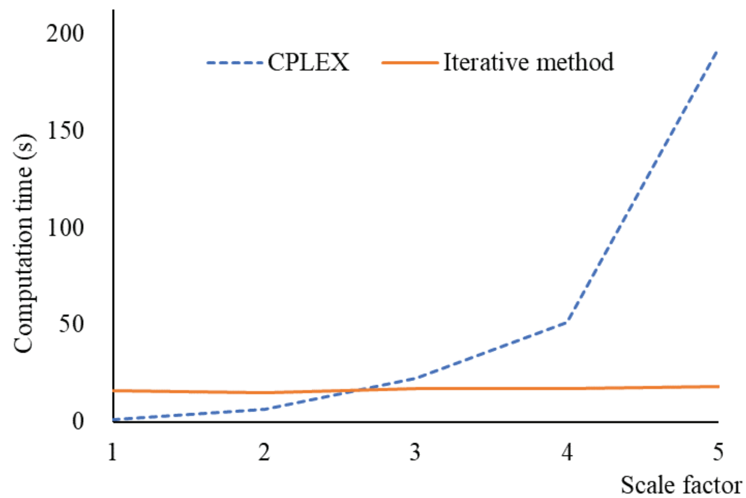


Figure 12. Comparison of the changes in computation time as the scale factor increases.

In addition, the initial schedule plan of the scenario when the scale factor is two contains one overtake event between two trains of Line L1. These two trains are named as Train A and Train B. Train A overtakes Train B at Station II. Figure 13 shows their times and spaces from Station I to Station III. As shown in Figure 13, a train stops at a station, i.e., it does not move on the x-axis but moves on the y-axis. Thus, vertical drawing lines represent train stops. At Station II, Train A stops for a long period, and during that period, Train B, a later coming train, overtakes it. This event shows that our model can handle train overtaking.

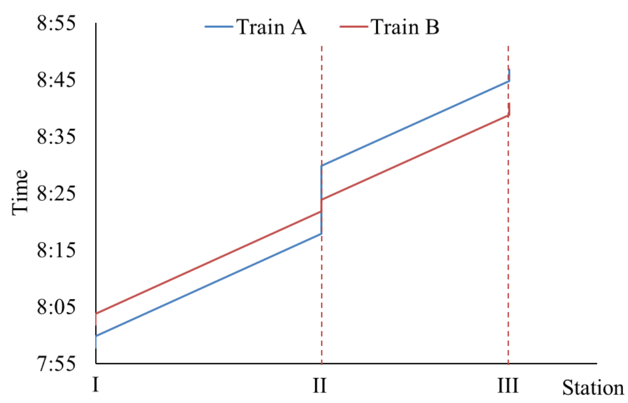


Figure 13. An overtaking event.

Moreover, we perform a test to investigate the two iterative processes in the suggested method, one without line planning and one with line planning. The problem can be resolved only by the iterative process with line planning, and the output is compared with the solutions from the two iterative processes. Predictably, running the iterative process with line planning requires a shorter computation time than running the two iterative processes (8.01 s compared with 24.78 s). However, the probability of finding new itineraries decreases because some trains are canceled and some trains are set to skip stations. For example, as shown in Table 2, a new itinerary for OD pair 2 (Train 3→Train 9) is found in SPG (1). However, if the iterative process with line planning is used after the ISPG, this itinerary cannot be found because Train 3, which carries zero passengers in the previous SSPA, will have been canceled by line planning. The decision of whether to consider the iterative process without line planning is a trade-off between the computation time and the probability of finding new itineraries, and this can be decided according to practical needs.

5.2. HSR network of southern China

We abstract a larger network, southern China's HSR network, as shown in Figure 14. The setting of this example is listed in Appendix D.

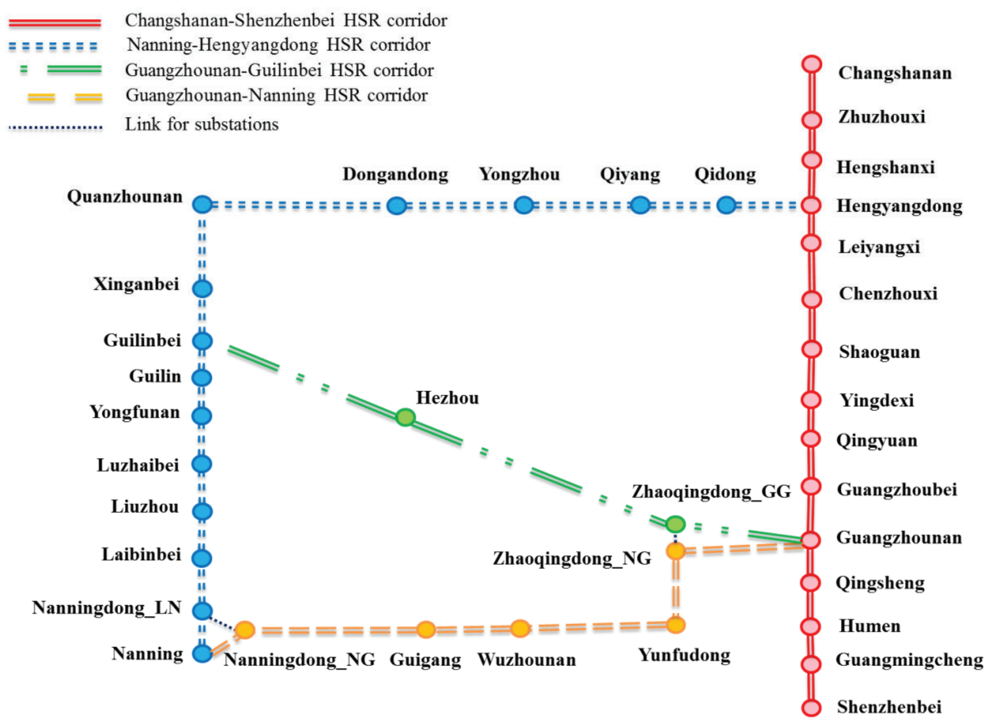


Figure 14. Abstract network for southern China's HSR.

5.2.1. Itinerary analysis

By running the iterative method stated in Section 4, we obtain a schedule plan for the whole network. The total passenger itinerary cost is calculated in the SSPA for each iteration, and it shows a generally decreasing trend (Figure 15).

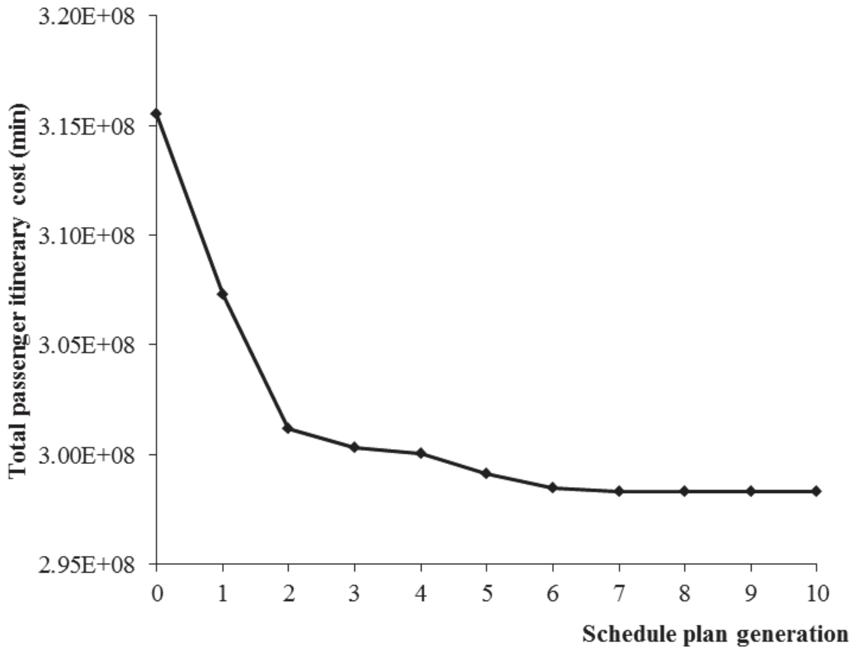


Figure 15. Total passenger itinerary cost for each train schedule-plan generation.

The iteration limit is set at 10, and no gap allowance is set for stopping the iterative process with line planning. The total passenger itinerary costs for the fifth and sixed schedule plans are approximately 299,130,440 min and 298,443,921 min, so the gap between them (0.23%) falls within the allowance of the absolute gap (0.25%). Thus, we begin to consider canceling and skipping in the seventh SPG. The computation time for the whole process (including the ISPG, 10 SPGs, and all SSPAs) is 45.0 hours.

The starting point ("0") is for the ISPG. The schedule difference is minimized in the ISPG; however, these differences cannot be zero because not all ideal departure times found in the approach for ideal-time determination can be met when generating the initial schedule plan. For example, two OD pairs want to use two different trains entering the same rail segment at the same time, but there is a safety headway between two successive trains so that one of these two OD pairs must wait. Or two OD pairs desire to use the same train, but they require different ideal departure times for this train. Hence, the initial schedule plan is unlikely to be perfect for all OD pairs, but the algorithm tries to find the schedule plan that best serves the majority of OD pairs. In the initial schedule plan, about 4.15% of the total OD pairs (377 OD pairs) are unable to find a feasible itinerary.

The solution converges at the eighth SPG after the program has run for 37.7 hours. 93.3% of the computation time is used for SSPAs. It takes around 1.5 second to find the itineraries and calculate the probabilities for each OD pair in a complex train network with 240 trains. The computation speed for each OD pair is reasonable, but there are nearly 10,000 OD pairs. Hence, a long computation time is needed. The computation for train scheduling is relatively quick (i.e., less than 7% of computation time) and the computation time for one SPG is around 17.6 min, because the decomposition approach is used. There are several adjustments in each SPG, and each adjustment is decomposed to some train-scheduling ILPs. The average numbers of variables and constraints for a decomposed train-scheduling ILP are 32,572 and 402,655, respectively. A such ILP handles one line while the other lines are fixed. Thus, the complexity of train scheduling reduces, and the number of feasible solutions becomes less. These ILPs can be handled by CPLEX, which averagely needs about 8.6 s for each of these ILPs. In addition, the Dijkstra's algorithm using in the ISPG for finding ideal time is efficient: it only takes about 3.5 s to find itineraries for 9076 OD pairs in a railway system with 240 trains. Because this is a planning stage problem, we believe that the computation time of 37.7 hours is acceptable for such a large network.

The total passenger itinerary cost is reduced by about 17,194,724 min after the eight SPGs, which illustrates the importance of the schedule plan adjustment based on the result of the stochastic passenger assignment. Most OD pairs (97.03%) have been served by the schedule plan obtained in the eighth SPG, in which 28.38% of the OD pairs that cannot find a feasible itinerary in the initial schedule plan find at least one feasible itinerary. This remains the case in later SPGs. For the 8806 served OD pairs, 29,678 feasible itineraries are found, so the average number of itineraries for an OD pair is 3.37. The total number of feasible itineraries and the average number of itineraries for an OD pair increase by 12.35% and 10.99%, respectively, compared with those from the initial schedule plan. Passengers have more options for their trips based on their needs.

5.2.2. Train operation analysis

Two trains are canceled from the schedule plan. All stations are served, and 26 stations are skipped by some of the passing trains. The total number of skips is 395. Some stations are more likely to be skipped:

1) Stations with low demand. For example, the Leiyangxi station, with 408 OD pairs, is skipped by 37 passing trains, whereas the neighboring Hengyangdong station, with 798 OD pairs, is skipped by only 10 passing trains.
2) Stations with low accessibility. For example, the Guangzhounan and Guangzhoubei stations are both located in Guangzhou, but the Guangzhounan station is more easily accessed by passengers than the Guangzhoubei station, which is closer to the city center. Furthermore, the Guangzhounan station can be accessed by two metro lines, whereas the Guangzhoubei station can only be accessed by one. Hence, the passenger travel time from the city center to the Guangzhounan station is shorter; thus, both stations serve the same zone but with different access times. More passengers therefore tend to select the Guangzhounan station as their terminal station. The Guangzhoubei station is skipped by 15 passing trains, whereas the Guangzhounan station is skipped by only 7 passing trains.

Canceling trains and allowing them to skip stations can reduce passenger journey time and train operating costs. These results indicate that our model also considers supply once demand is satisfied.

We calculate schedule plans for all trains in the network, and we choose the network's main HSR corridor, the Changshanan-Shenzhenbei HSR corridor, as an example to show its one-direction timetable in Figure 16. The high-speed trains in this HSR corridor travel from the Changshanan station to the Shenzhenbei station. This HSR corridor is part of the Beijing-Hong Kong HSR corridor, which is one of the busiest passenger channels in China. As Figure 16 shows, many high-speed trains pass through this HSR corridor, and their departure intervals are quite short. As the demand pattern in Table 14 of Appendix D shows, more OD demand appears before 14:00. Because the schedule plan is based on passenger-itinerary choices, more high-speed trains are scheduled before 15:00.

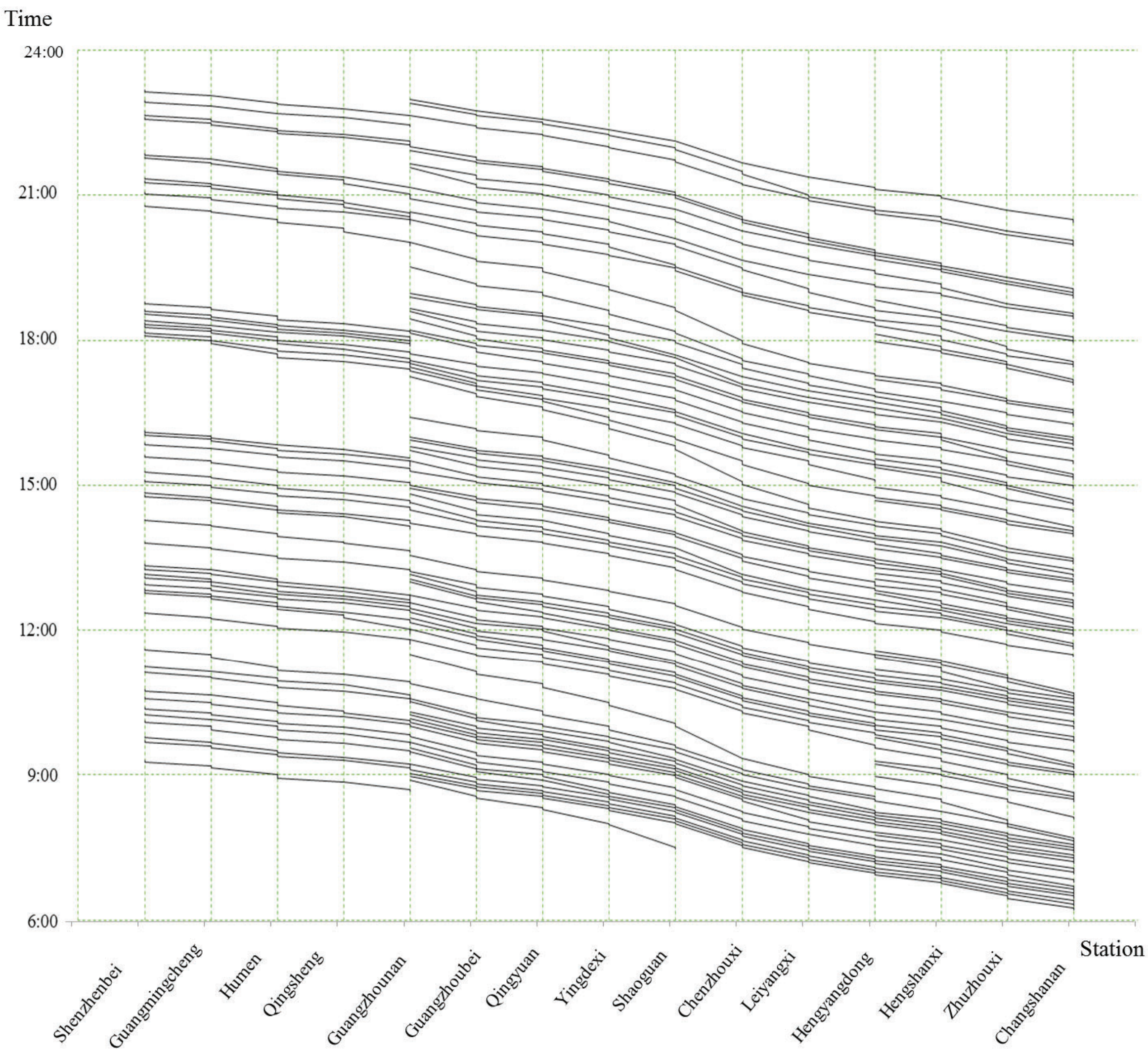


Figure 16. Timetable for trains in one direction of the Changshanan-Shenzhenbei HSR corridor.

6. Conclusions

This study develops a new itinerary-choice model, the MISW model, that addresses heterogeneity of passengers' perceptions and the IID assumption. Specifically, the Weibull distributed random error term resolves the perception variance with respect to various itinerary lengths, and an itinerary-size factor term is suggested to handle the itinerary overlapping problem. We analyze the MISW model in Section 3 with three numerical examples. The first two analyses indicate that the MISW model is superior to logit-type models because the MISW model differentiates itineraries according to their relative rather than absolute cost difference and adopts the itinerary-size factor to account for the correlation between overlapping itineraries, and the third analysis shows that it is essential to consider random parameters in the MISW model.

Furthermore, the MISW model is applied to a schedule-based passenger-oriented train-scheduling model that improves the quality of train services by adjusting the train timetable when the train cancellation plan, stopping pattern, and station-track allocation are also optimized. We describe the whole problem as a fixed-point problem and solve it using an iterative method. In addition, previous studies have rarely modeled timetable and station-track allocation simultaneously in the passenger-oriented scheduling problem. However, doing so is important because good station-track allocation can facilitate passenger transfers.

In the example in Section 5.1, we find that the passenger service quality of the planned schedule is improved by using the schedule-based stochastic passenger assignment based on the MISW model. In this example, the total itinerary cost decreases by about 1016 min due to better scheduling. For instance, skipping unused stations decreases IVT. In some cases, wait time and transfer time can be reduced by 93.55% and 60.00%, respectively. The solution quality of the iterative method is also tested. The comparison shows the ability of the iterative method to find a good-quality solution within a much shorter time than the commercial CPLEX software. In the extreme case of this test, the computation time of CPLEX is about ten times that of the iterative method.

We apply the train-scheduling model to southern China’s HSR network (Section 5.2) and the result also demonstrates that the model can help to reduce the passenger itinerary cost. The total passenger itinerary cost decreases by around 17,194,724 min. The consideration of several possible itineraries in the train-scheduling model enriches passenger choices. On average, each OD pair has three itinerary choices. In addition, the iterative method can help to obtain convergent suboptimal solutions for a regional train schedule plan in a reasonable time. 37.7 hours was required to find a convergent solution for such a problem with 36 stations, 240 high-speed trains, and 9076 OD pairs.

Although the MISW model is built and applied to the HSR system, it can also be used by non-HSR transit systems (such as bus systems and intercity metro networks). Because all public transit modes share some characteristics (e.g., operating on a timetable), only small modifications would be required to adapt this model to another type of transit system. For example, when the MISW model is used to plan a timetable for a bus system, the task of track selection for transfers would be removed.

This study makes a number of assumptions:

1) The capacity of trains is sufficient.
2) Passenger coefficients are related to trip purpose and income level.
3) There is a sufficient vehicle source for schedule plans.

However, these assumptions are not always the case in the real world. For example, all itineraries are available to passengers based on the assumption of train capacity and that even in a delay situation, passengers can easily shift to another unaffected itinerary. But it is possible that passengers may be unable to buy a ticket because of the limited train capacity, so that passengers may not benefit from the provision of several improved itineraries. Hence, a study should examine the effects of relaxing the above assumptions. Furthermore, the train-scheduling model used in this study focuses mainly on the passengers' benefit, which means that more trains may be scheduled to fulfill this purpose without consideration of the increased operating costs that it will generate. Our future study will investigate methods for balancing operators and passengers in the planning of railway systems to prevent the creation of service-cost imbalances.

Acknowledgments

The work described in this paper is supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. [T32-101/15-R]). The first author is also supported by two scholarships (a Postgraduate Scholarship from the University of Hong Kong and a University Postgraduate Fellowship from the Jessie and George Ho Charitable Foundation). The second author is supported by the Francis S.Y. Bong Endowed Professorship in Engineering. The third author is supported by the 2017 Hong Kong Scholar Program and the National Natural Science Foundation of China (NSFC) (No.71701174). The authors are grateful to Professor H.P. Lo for his constructive comments on the methodology.

References

ADLER, T., FALZARANO, C. S. & SPITZ, G. 2005. Modeling service trade-offs in air itinerary choices. *Transportation Research Record: Journal of the Transportation Research Board,* 1915, 20-26.

BARRENA, E., CANCA, D., COELHO, L. C. & LAPORTE, G. 2014. Exact formulations and algorithm for the train timetabling problem with dynamic demand. *Computers & Operations Research,* 44, 66-74.

BEKHOR, S. & FREUND-FEINSTEIN, U. 2006. Modeling passengers' preferences on a short-haul domestic airline with rank-ordered data. *Transportation Research Record: Journal of the Transportation Research Board,* 1951, 1-6.

BEN-AKIVA, M. & BIERLAIRE, M. 1999. Discrete choice methods and their applications to short term travel decisions. *In : Handbook of Transportation Science.* Boston, MA: Springer.

BEN-AKIVA, M. E. & LERMAN, S. R. 1985. *Discrete choice analysis: theory and application to travel demand,* Cambridge, Massachusetts: MIT Press.

BEŠINOVIĆ, N., GOVERDE, R. M. P., QUAGLIETTA, E. & ROBERTI, R. 2016. An integrated micro-macro approach to robust railway timetabling. *Transportation Research Part B: Methodological,* 87, 14-32.

BORNDÖRFER, R., HOPPMANN, H. & KARBSTEIN, M. 2017. Passenger routing for periodic timetable optimization. *Public Transport,* 9, 115-135.

CACCHIANI, V. & TOTH, P. 2012. Nominal and robust train timetabling problems. *European Journal of Operational Research,* 219, 727-737.

CAIMI, G., BURKOLTER, D., HERRMANN, T., CHUDAK, F. & LAUMANNS, M. 2009. Design of a railway scheduling model for dense services. *Networks and Spatial Economics,* 9, 25-46.

CANCA, D., ALGABA, E., BARRENA, E. & ZARZO, A. 2014. Railway rapid transit timetables with variable and elastic demand. *Procedia - Social and Behavioral Sciences,* 111, 538-548.

CASCETTA, E., NUZZOLO, A., RUSSO, F. & VITETTA, A. A modified logit route choice model overcoming path overlapping problems specification and some calibration results for interurban networks. *In:* LESORT, J. B., (ed.) *ISTTT Conference*, *1996*, *Lyon, France*. 697-711.

CASCETTA, E., RUSSO, F. & VITETTA, A. 1997. Stochastic user equilibrium assignment with explicit path enumeration: Comparison of models and algorithms. *IFAC Proceedings Volumes,* 30, 1031-1037.

CASTILLO, E., MENÉNDEZ, J. M., JIMÉNEZ, P. & RIVAS, A. 2008. Closed form expressions for choice probabilities in the Weibull case. *Transportation Research Part B: Methodological,* 42, 373-380.

CHEN, A., PRAVINVONGVUTH, S., XU, X., RYU, S. & CHOOTINAN, P. 2012. Examining the scaling effect and overlapping problem in logit-based stochastic user equilibrium models. *Transportation Research Part A: Policy and Practice,* 46, 1343-1358.

CORMAN, F., D'ARIANO, A., MARRA, A. D., PACCIARELLI, D. & SAMÀ, M. 2017. Integrating train scheduling and delay management in real-time railway traffic control. *Transportation Research Part E: Logistics and Transportation Review,* 105, 213-239.

DOLLEVOET, T., HUISMAN, D., KROON, L. G., VEELENTURF, L. P. & WAGENAAR, J. C. 2017. Application of an iterative framework for real-time railway rescheduling. *Computers & Operations Research,* 78, 203-217.

DOMENCICH, T. A. & MCFADDEN, D. 1975. *Urban travel demand: A behavioral analysis,* North Holland, Amsterdam.

DRABAS, T. & WU, C. L. 2013. Modelling air carrier choices with a segment specific cross nested logit model. *Journal of Air Transport Management,* 32, 8-16.

FREJINGER, E. & BIERLAIRE, M. 2007. Capturing correlation with subnetworks in route choice models. *Transportation Research Part B: Methodological,* 41, 363-378.

FREUND-FEINSTEIN, U. & BEKHOR, S. 2017. An airline itinerary choice model that includes the option to delay the decision. *Transportation Research Part A: Policy and Practice,* 96, 64-78.

GATTERMANN, P., GROßMANN, P., NACHTIGALL, K. & SCHÖBEL, A. 2016. Integrating passengers' routes in periodic timetabling: A SAT approach. *In:* HERBSTRITT, M., (ed.) *16th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*, *Dagstuhl, Germany*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik GmbH, Wadern/Saarbruecken, Germany.

GUO, X., WU, J., SUN, H., LIU, R. & GAO, Z. 2016. Timetable coordination of first trains in urban railway network: A case study of Beijing. *Applied Mathematical Modelling,* 40, 8048-8066.

HARROD, S. S. 2012. A tutorial on fundamental model structures for railway timetable optimization. *Surveys in Operations Research and Management Science,* 17, 85-96.

HESS, S. & POLAK, J. W. 2006. Airport, airline and access mode choice in the San Francisco Bay area. *Papers in Regional Science,* 85, 543-567.

HEYDECKER, B. G. 1996. A decomposition approach for signal optimisation in road networks. *Transportation Research Part B: Methodological,* 30, 99-114.

HUNT, J. D. 1990. A logit model of public transport route choice. *ITE Journal,* 60, 26-30.

INGOLOTTI, L., LOVA, A., BARBER, F., TORMOS, P., SALIDO, M. A. & ABRIL, M. 2006. New heuristics to solve the "CSOP" railway timetabling problem. *Advances in Applied Artificial Intelligence: 19th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, *2006*, *Annecy, France*. 400-409.

KANG, L., LU, Z. & MENG, Q. 2019a. Stochastic schedule-based optimization model for track allocations in large railway stations. *Journal of Transportation Engineering Part A: Systems,* 145, 1-9.

KANG, L. & MENG, Q. 2017. Two-phase decomposition method for the last train departure time choice in subway networks. *Transportation Research Part B: Methodological,* 104, 568-582.

KANG, L., WU, J. & SUN, H. 2012. Using simulated annealing in a bottleneck optimization model at railway stations. *Journal of Transportation Engineering,* 138, 1396-1402.

KANG, L., WU, J., SUN, H., ZHU, X. & WANG, B. 2015. A practical model for last train rescheduling with train delay in urban railway transit networks. *Omega,* 50, 29-42.

KANG, L., ZHU, X., SUN, H., WU, J., GAO, Z. & HU, B. 2019b. Last train timetabling optimization and bus bridging service management in urban railway transit networks. *Omega,* 84, 31-44.

KARABUK, S. 2009. A nested decomposition approach for solving the paratransit vehicle scheduling problem. *Transportation Research Part B: Methodological,* 43, 448-465.

KHANI, A., HICKMAN, M. & NOH, H. 2014. Trip-based path algorithms using the transit network hierarchy. *Networks and Spatial Economics,* 15, 635-653.

KITTHAMKESORN, S. & CHEN, A. 2013. A path-size weibit stochastic user equilibrium model. *Transportation Research Part B: Methodological,* 57, 378-397.

KITTHAMKESORN, S. & CHEN, A. 2014. Unconstrained weibit stochastic user equilibrium model with extensions. *Transportation Research Part B: Methodological,* 59, 1-21.

KROON, L. G. & PEETERS, L. W. P. 2003. A variable trip time model for cyclic railway timetabling. *Transportation Science,* 37, 198-212.

LAM, S. H. & XIE, F. 2002. Transit path-choice models that use revealed preference and stated preference data. *Transportation Research Record: Journal of the Transportation Research Board,* 1799, 58-65.

LIEBCHEN, C. & MÖHRING, R. H. 2002. A case study in periodic timetabling. *Electronic Notes in Theoretical Computer Science,* 66, 18-31.

LIN, K. Y. & SIBDARI, S. Y. 2009. Dynamic price competition with discrete customer choices. *European Journal of Operational Research,* 197, 969-980.

LURKIN, V., GARROW, L. A., HIGGINS, M. J., NEWMAN, J. P. & SCHYNS, M. 2017. Accounting for price endogeneity in airline itinerary choice models: An application to Continental U.S. markets. *Transportation Research Part A: Policy and Practice,* 100, 228-246.

LUSBY, R. M., LARSEN, J. & BULL, S. 2018. A survey on robustness in railway planning. *European Journal of Operational Research,* 266, 1-15.

NIU, H. & ZHOU, X. 2013. Optimizing urban rail timetable under time-dependent demand and oversaturated conditions. *Transportation Research Part C: Emerging Technologies,* 36, 212-230.

NIU, H., ZHOU, X. & GAO, R. 2015. Train scheduling for minimizing passenger waiting time with time-dependent demand and skip-stop patterns: Nonlinear integer programming models with linear constraints. *Transportation Research Part B: Methodological,* 76, 117-135.

NUZZOLO, A. & CRISALLI, U. 2009. The schedule-based modeling of transportation systems: Recent developments. *In : Schedule-Based Modeling of Transportation Networks: Theory and applications.* Boston, MA: Springer US.

PARBO, J., NIELSEN, O. A. & PRATO, C. G. 2016. Passenger perspectives in railway timetabling: A literature review. *Transport Reviews,* 36, 500-526.

PARKER, R. A. 2017. An agent-based simulation of air travel itinerary choice. *Procedia Computer Science,* 109, 905-910.

POON, M. H., WONG, S. C. & TONG, C. O. 2004. A dynamic schedule-based model for congested transit networks. *Transportation Research Part B: Methodological,* 38, 343.

ROBENEK, T., AZADEH, S. S., MAKNOON, Y., DE LAPPARENT, M. & BIERLAIRE, M. 2018. Train timetable design under elastic passenger demand. *Transportation Research Part B: Methodological,* 111, 19-38.

ROBENEK, T., MAKNOON, Y., AZADEH, S. S., CHEN, J. & BIERLAIRE, M. 2016. Passenger centric train timetabling problem. *Transportation Research Part B: Methodological,* 89, 107-126.

SCHMIDT, M. & SCHÖBEL, A. 2015. Timetabling with passenger routing. *OR Spectrum,* 37, 75-97.

SCHÖBEL, A. 2017. An eigenmodel for iterative line planning, timetabling and vehicle scheduling in public transportation. *Transportation Research Part C: Emerging Technologies,* 74, 348-365.

SEELHORST, M. & LIU, Y. 2015. Latent air travel preferences: Understanding the role of frequent flyer programs on itinerary choice. *Transportation Research Part A: Policy and Practice,* 80, 49-61.

SELS, P., DEWILDE, T., CATTRYSSE, D. & VANSTEENWEGEN, P. 2011. Deriving all passenger flows in a railway network from ticket sales data. *Proceedings of 4th International Seminar on Railway Operations Modelling and Analysis (IAROR): RailRome2011*, *Roma, Itialy*.

SELS, P., DEWILDE, T., CATTRYSSE, D. & VANSTEENWEGEN, P. 2016. Reducing the passenger travel time in practice by the automated construction of a robust railway timetable. *Transportation Research Part B: Methodological,* 84, 124-156.

SHEFFI, Y. 1985. *Urban transportation networks: Equilibrium analysis with mathematical programming methods,* Englewood Cliffs, N.J.: Prentice-Hall.

SIEBERT, M. 2011. *Integration of routing and timetabling in public transportation.* MSc thesis, Georg-August-Universität Göttingen.

SINHA, S. K., SALSINGIKAR, S. & SENGUPTA, S. 2016. An iterative bi-level hierarchical approach for train scheduling. *Journal of Rail Transport Planning & Management,* 6, 183-199.

SPARING, D. & GOVERDE, R. M. P. 2017. A cycle time optimization model for generating stable periodic railway timetables. *Transportation Research Part B: Methodological,* 98, 198-223.

TALEBIAN, A. & ZOU, B. 2015. Integrated modeling of high performance passenger and freight train planning on shared-use corridors in the US. *Transportation Research Part B: Methodological,* 82, 114-140.

TAN, R., ADNAN, M., LEE, D.-H. & BEN-AKIVA, M. E. 2015. New path size formulation in path size logit for route choice modeling in public transport networks. *Transportation Research Record: Journal of the Transportation Research Board,* 2538, 11-18.

TIAN, C. H., ZHANG, S. S., ZHANG, Y. S. & JIANG, X. L. 2015. Study on the train headway on automatic block sections of high speed railway. *Journal of the China Railway Society,* 37, 1-6.

TONG, C. O. & RICHARDSON, A. J. 1984. A computer model for finding the time-dependent minimum path in a transit system with fixed schedules. *Journal of Advanced Transportation,* 18, 145-161.

TONG, C. O. & WONG, S. C. 1999. A stochastic transit assignment model using a dynamic schedule-based network. *Transportation Research Part B: Methodological,* 33, 107-121.

TONG, C. O. & WONG, S. C. 2004. Minimum path algorithms for a schedule-based transit network with a general fare structure. *In:* WILSON, N. H. M. & NUZZOLO, A. (eds.) *In : Schedule-Based Dynamic Transit Modeling: Theory and Applications.* New York: Springer.

TONG, C. O., WONG, S. C., POON, M. H. & TAN, M. C. 2001. A schedule-based dynamic transit network model: Recent advances and prospective future research. *Journal of Advanced Transportation,* 35, 175-195.
VANSTEENWEGEN, P. & OUDHEUSDEN, D. V. 2006. Developing railway timetables which guarantee a better service. *European Journal of Operational Research,* 173, 337-350.
WARBURG, V., BHAT, C. & ADLER, T. 2006. Modeling demographic and unobserved heterogeneity in air passengers' sensitivity to service attributes in itinerary choice. *Transportation Research Record: Journal of the Transportation Research Board,* 1951, 7-16.
WONG, C. K. & WONG, S. C. 2002. Lane-based optimization of traffic equilibrium settings for area traffic control. *Journal of Advanced Transportation,* 36, 349-386.
WONG, R. C. W., YUEN, T. W. Y., KWOK WAH, F. & LEUNG, J. M. Y. 2008. Optimizing timetable synchronization for rail mass transit. *Transportation Science,* 42, 57-69.
WU, J., KANG, L., SUN, H. & JIA, X. 2013. Track allocation optimization in railway station: Mean-variance model and case study. *Journal of Transportation Engineering,* 139, 540-547.
XIE, J., WONG, S. C. & LO, S. M. 2017. Three extensions of Tong and Richardson's algorithm for finding the optimal path in schedule-based railway networks. *Journal of Advanced Transportation,* 2017, 1-16.
ZHAN, S., KROON, L. G., VEELENTURF, L. P. & WAGENAAR, J. C. 2015. Real-time high-speed train rescheduling in case of a complete blockage. *Transportation Research Part B: Methodological,* 78, 182-201.
ZHOU, W. & TENG, H. 2016. Simultaneous passenger train routing and timetabling using an efficient train-based Lagrangian relaxation decomposition. *Transportation Research Part B: Methodological,* 94, 409-439.
ZHOU, Z., CHEN, A. & BEKHOR, S. 2012. C-logit stochastic user equilibrium model: Formulations and solution algorithm. *Transportmetrica,* 8, 17-41.

Appendix A. Constraints

Appendix A.1. Train-scheduling constraints

The train-scheduling constraints consist of five kinds of constraint as follows.

1). Departure/arrival time constraints

A line consists of several trains. The operating time-window for trains on line l is $[T^{\mathrm{dep}}_{l},T^{\mathrm{arr}}_{l}]$, where T^{dep}_{l} is the earliest departure time at the origin terminal and T^{arr}_{l} is the latest arrival time at the destination terminal. The departure/arrival time ($t^{\mathrm{D}}_{ns}/t^{\mathrm{A}}_{ns}$) of train n on line l at station s should be within the following range:

$$T^{\mathrm{dep}}_{l} \leq t^{\mathrm{D}}_{ns} \leq T^{\mathrm{arr}}_{l}, \qquad ()$$

$$T^{\mathrm{dep}}_{l} \leq t^{\mathrm{A}}_{ns} \leq T^{\mathrm{arr}}_{l}. \qquad ()$$

2). Headway constraints

Our study assumes that double tracks are used between stations so that each track serves one running direction, and a train can only overtake another at a station. If trains n and $n^{'}$ arrive from or depart in the same direction, the following constraints should hold to maintain sufficient headway between them:

$$t^{\mathrm{A}}_{n^{'}s} + h \leq t^{\mathrm{A}}_{ns} + (1 - b_{n^{'}ns}) \times \mathrm{M}, \qquad ()$$

$$t^{\mathrm{D}}_{n^{'}s} + h \leq t^{\mathrm{D}}_{ns} + (1 - b'_{n^{'}ns}) \times \mathrm{M}, \qquad ()$$

$$b_{n^{'}ns} + b_{nn^{'}s} = 1\ , \qquad ()$$

$$b'_{n^{'}ns} + b'_{nn^{'}s} = 1, \qquad ()$$

$$b'_{n^{'}ns} - b_{n^{'}ns^{'}} = 0, \qquad ()$$

where h is the minimum headway for two successive trains,

M is a sufficiently large number,

$b_{n^{'}ns} = \begin{cases} 1, \text{ if train } n^{'} \text{ arrives before train } n \text{ at station s} \\ 0, \quad \text{else} \end{cases}$, and

$b'_{nn^{'}s} = \begin{cases} 1, \text{ if train } n^{'} \text{ departs before train } n \text{ at station } s \\ 0, \quad \text{else} \end{cases}$.

Constraints () to () guarantee that the safety headway between trains running in a segment or at a station is respected. In addition, station $s^{'}$ in constraint () is the downstream station of station s, and constraint () ensures that trains do not overtake others between stations. $b_{n^{'}ns}$ and $b'_{n^{'}ns}$ can be different; if they are, it means that overtaking occurs at station s.

3). Station-track assignment constraints.

Tracks at station s form a set of station tracks, $\boldsymbol{P}_s$. Figure 17 shows an example station.

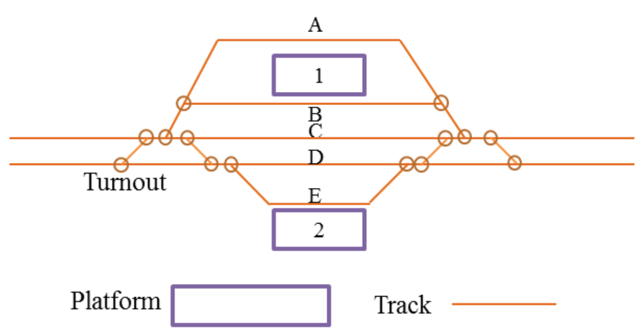


Figure 17. Track and platform layout of a station.

Some tracks are equipped with platforms for boarding and alighting (e.g., Track A in Figure 17), and others are not (e.g., Track C in Figure 17). Trains are guided via turnouts to arrive at different tracks. We generally divide the station tracks into two categories, $\boldsymbol{P}^1_s$ and $\boldsymbol{P}^2_s$, according to whether they are equipped with platforms. We note that the tracks in $\boldsymbol{P}^1_s$ can be used by any trains, but the tracks in $\boldsymbol{P}^2_s$

can only be used by non-stopping trains. The constraints for the station-track allocation are as follows:

$$t_{n's}^{\mathrm{D}} + h \le t_{ns}^{\mathrm{A}} + (3 - b_{np}^{\mathrm{Tr}} - b_{n'p}^{\mathrm{Tr}} - b_{n'ns}) \times \mathrm{M}, \qquad ()$$

$$\begin{cases} \sum_{p \in \boldsymbol{P}_s^2} b_{np}^{\mathrm{Tr}} + b_{ns}^{\mathrm{Tr}*} = 1, & \text{for non - stopping train } n \\ \sum_{p \in \boldsymbol{P}_s^1} b_{np}^{\mathrm{Tr}} = b_{ns}^{\mathrm{Tr}*}, & \text{for non - stopping train } n, \\ \sum_{p \in \boldsymbol{P}_s^1} b_{np}^{\mathrm{Tr}} = 1, & \text{for stopping train } n \end{cases} \qquad ()$$

where $b_{np}^{\mathrm{Tr}} = \begin{cases} 1, \text{ if train } n \text{ uses track } p \\ 0, \quad \text{else} \end{cases}$ and

$b_{ns}^{\mathrm{Tr}*} = \begin{cases} 1, \text{ if train } n \text{ skipping station } s \text{ uses a track belonging to } \boldsymbol{P}_s^1 \\ 0, \quad \text{otherwise} \end{cases}.$

Constraints () and () ensure that the safety headway between trains using the same station track is respected. Constraint () assigns station tracks to trains and states that a train can only take one station track. Stopping trains should take a track with a platform at a station so that passengers can get on or off. Non-stopping trains should first be planned to use tracks without platforms. If tracks without platforms are unavailable, the non-stopping trains are considered to use other tracks. $\sum_{n \in \boldsymbol{T}} \sum_{s \in \hat{\mathrm{S}}_l} b_{ns}^{\mathrm{Tr}*}$ ($\hat{\mathrm{S}}_l$ is the set of passing stations for train n on line l and $\boldsymbol{T}$ is the set for all trains) should be minimized to ensure this.

4). Dwell time constraints

The minimum dwell time for train n on line l at station s is $D_{sl}^{\min}$. Dwell time constraints are formulated as follows.

If train n stops, then

$$D_{sl}^{\min} \le t_{ns}^{\mathrm{D}} - t_{ns}^{\mathrm{A}}. \qquad ()$$

If train n skips the station, then

$$0 \le t_{ns}^{\mathrm{D}} - t_{ns}^{\mathrm{A}}. \qquad ()$$

Although train n skips station s, the dwell time may not be zero; train n may wait at station s to maintain a safe headway between successive trains.

5). Running time constraint

The minimum running time for train n on line l in a rail segment v is T_{lv}^{run}. This constraint is formulated as follows:

$$T_{lv}^{\mathrm{run}} \le t_{ns'}^{\mathrm{A}} - t_{ns}^{\mathrm{D}}, \qquad ()$$

where station s' is the downstream station of station s. We do not fix the running time, so that the train speed can vary under the maximum speed to meet passengers' need and the limitation of station capacity.

Appendix A.2. Itinerary constraints

The variables, parameters, and sets used in the itinerary constraints are provided in Table 6.

Table 6. Elements, sets, variables, and parameters used in the itinerary constraints.

Element	**Definition**
r	An OD pair
i	An itinerary
n or n'	A train
p or p'	A track
s or s'	A station
s_{in}^{A}	Station at which passengers on itinerary i alight from train n
s_{in}^{B}	Station at which passengers on itinerary i board train n
n_{i}^{O}	Train which passengers on itinerary i use to depart from the origin
n_{i}^{D}	Train which passengers on itinerary i use to arrive at the destination
n_{in}^{T}	Train to which passengers on itinerary i transfer from train n
z	A zone
z_{r}^{O}	Origin zone of OD pair r
Set	**Definition**
$\boldsymbol{R}$	Set of OD pairs
$\boldsymbol{P}_s$	Set of tracks at station s
$\boldsymbol{I}_r$	Set of itineraries used by OD pair r
Variable	**Definition**
x_{ri}^{1}	IVT of itinerary i
x_{ri}^{2}	Wait time at stations of itinerary i
x_{ri}^{3}	Wait time in the origin zone of itinerary i
x_{ri}^{4}	Walking time of itinerary i
$t_{inn's}^{\mathrm{walk}}$	For itinerary i, walking time for transfer from train n to train n' at station s
t_{ns}^{A}	Time of arrival at station s of train n
t_{ns}^{D}	Time of departure from station s of train n
$t_{i}^{\mathrm{wait-H}}$	For itinerary i number of T^{wait} in the origin zone
b_{np}^{Tr}	1 if train n stops at track p, else 0
$b_{nn'pp'ss'}^{\mathrm{TT}'}$	1 if train n uses track p at station s and train n' uses track p' at station s', else 0
$t_{pp'}^{\mathrm{wp}}$	Walking time between the platform serving track p at station s and the platform serving track p' at station s'; s and s' may be the same station or different stations linked by walkways
Parameter	**Definition**
T_{r}^{de}	Preferred departure time of OD pair r in the origin zone
T_{sz}^{zone}	Travel time between zone z and station s via other modes of transportation, such as buses or cars.
T^{wait}	Tolerance of passengers for waiting at the origin station

The itinerary constraints include:

$$x_{ri}^{1} = \sum_{n \in \boldsymbol{T}_i,\, s = s_{in}^{\mathrm{A}},\, s' = s_{in}^{\mathrm{B}}} (t_{ns}^{\mathrm{A}} - t_{ns'}^{\mathrm{D}});\ \forall i \in \boldsymbol{I}_r;\ \forall r \in \boldsymbol{R}; \quad ()$$

$$x_{ri}^{2} = \sum_{n \in (\boldsymbol{T}_i - n_i^{\mathrm{D}}),\, n' = n_{in}^{\mathrm{T}},\, s = s_{in}^{\mathrm{A}},\, s' = s_{in}^{\mathrm{B},}} (t_{n's'}^{\mathrm{D}} - t_{ns}^{\mathrm{A}} - t_{inn's}^{\mathrm{walk}}) + \sum_{n = n_i^{\mathrm{O}},\, s = s_{in}^{\mathrm{B}},\, z = z_r^{\mathrm{O}}} \left(t_{ns}^{\mathrm{D}} - T_r^{\mathrm{de}} - T^{\mathrm{wait}} \times t_i^{\mathrm{wait-H}} - T_{sz}^{\mathrm{zone}}\right);\ \forall i \in \boldsymbol{I}_r;\ \forall r \in \boldsymbol{R};$$

()

$$x_{ri}^{3} = T^{\mathrm{wait}} \times t_i^{\mathrm{wait-H}};\ \forall i \in \boldsymbol{I}_r;\ \forall r \in \boldsymbol{R}; \quad ()$$

$$x_{ri}^{4} = \sum_{n \in (\boldsymbol{T}_i - n_i^{\mathrm{D}}),\, n' = n_{in}^{\mathrm{T}},\, s = s_{in}^{\mathrm{A}},\, s' = s_{in}^{\mathrm{B},}} (t_{inn's}^{\mathrm{walk}});\ \forall i \in \boldsymbol{I}_r;\ \forall r \in \boldsymbol{R}; \quad ()$$

$$t_{ns}^{\mathrm{D}} - \left(T_r^{\mathrm{de}} + T^{\mathrm{wait}} \times t_i^{\mathrm{wait-H}} + T_{sz}^{\mathrm{zone}}\right) \geq 0;\ s = s_{in}^{\mathrm{B}};\ n = n_i^{\mathrm{O}};\ z = z_r^{\mathrm{O}};\ \forall i \in \boldsymbol{I}_r;\ \forall r \in \boldsymbol{R}; \quad ()$$

$$t_{n's'}^{\mathrm{D}} - t_{ns}^{\mathrm{A}} - t_{inn's}^{\mathrm{walk}} \geq 0;\ \forall n \in (\boldsymbol{T}_i - n_i^{\mathrm{D}});\ n' = n_{in}^{\mathrm{T}};\ s = s_{in}^{\mathrm{A}};\ s' = s_{in}^{\mathrm{B},};\ \forall i \in \boldsymbol{I}_r;\ \forall r \in \boldsymbol{R}; \quad ()$$

$$t_{inn's}^{\mathrm{walk}} = \sum_{p \in \boldsymbol{P}_s} \sum_{p' \in \boldsymbol{P}_{s'}} (t_{pp'}^{\mathrm{wp}} \times b_{nn'pp'ss'}^{\mathrm{TT}});\ \forall n \in (\boldsymbol{T}_i - n_i^{\mathrm{D}});\ n' = n_{in}^{\mathrm{T}};\ s = s_{in}^{\mathrm{A}};\ s' = s_{in}^{\mathrm{B},};\ \forall i \in \boldsymbol{I}_r;\ \forall r \in \boldsymbol{R};$$

()

$$\begin{cases} b_{nn'pp'ss'}^{\mathrm{TT}} \leq b_{np}^{\mathrm{Tr}} \\ b_{nn'pp'ss'}^{\mathrm{TT}} \leq b_{n'p'}^{\mathrm{Tr}} \\ b_{np}^{\mathrm{Tr}} + b_{n'p'}^{\mathrm{Tr}} \leq 1 + b_{nn'pp'ss'}^{\mathrm{TT}} \end{cases};\ p \in \boldsymbol{P}_s;\ p' \in \boldsymbol{P}_{s'};\ \forall n \in \left(\boldsymbol{T}_i - n_i^{\mathrm{D}}\right);\ n = n_{in}^{\mathrm{T}};\ s = s_{in}^{\mathrm{A}};\ s' = s_{in}^{\mathrm{B},};\ \forall i \in \boldsymbol{I}_r;\ \forall r \in \boldsymbol{R}. \quad ()$$

Eq. ()-() determines the IVT, wait time, and walking time for transfers of itinerary i. Because the objective is to minimize the total itinerary cost, Constraint () is set to ensure that the wait time of itinerary i for OD pair r at the origin station would be non-negative, i.e., t_{ns}^{D} is later than the arrival time of passengers using itinerary i ($T_r^{\mathrm{de}} + T_{sz_o}^{\mathrm{zone}}$). Constraints ()-() relating to station-track assignment constraints (formulas () and ()) ensure that the transfer time between trains n and n' is sufficient. The transfer can occur at the same station or at two substations connected by walking links.

Appendix B. Setting for the small example.

The setting for the small example is described here:

1) Network setting: Each station serves one zone. All travel times between the zones and stations are assumed to be 0. All stations have two tracks and two platforms, except Station II which having three tracks and three platforms. The walking time between platforms is 5 min except for those at Station II. The walking time between Platform P1 and Platform P3 at Station II is 10 min, whereas the others are 8 min.
2) Line setting: There are 6 lines in total and each line can have more than one train. The operating time of trains is [8:00, 12:00]. Trains on a line follow the respective line setting shown in Table 7. "D" in the line name in the first column indicates that this line is for

downstream trains, whereas “U” is for the lines with upstream trains. “L” is for the loop lines. The second column shows which trains belong to which lines. The last column indicates the order in which each line passes the links in the network. For lines that use only one link, the order of passing stations is given to show the running direction.

Table 7. Line information of the loop network.

Line name	Trains on the line	Links
D1	1 and 2	1 (Station I → Station III)
D2	3 and 4	2 (Station I → Station II)
U1	5 and 6	1 (Station III → Station I)
U2	7 and 8	2 (Station II → Station I)
L1	9	3→4→9→8→7→6→5
L2	10	5→6→7→8→9→4→3

Additionally, Table 8 provides fare information. Generally, the fare is set based on the journey time. A longer journey time asks for a higher fare. In addition, Links 1 and 3 are shortcuts. Operators set a higher ticket price for trains that use Links 1 and 3 because of the extra effort required for the construction and maintenance of shortcuts.

Table 8. Ticket prices for the loop network.

Line name	OD stations	Fare (yuan)	Line name	OD stations	Fare (yuan)	Line name	OD stations	Fare (yuan)	Line name	OD stations	Fare (yuan)
D1	I-III	25		III-VII	45	U1	III-I	25		VII-III	45
D2	I-II	9		III-VI	65	U2	II-I	9		VI-III	65
	I-II	10		III-V	77		II-I	10		V-III	77
	I-III	16		III-IV	97		III-I	16		IV-III	97
	I-VII	61		III-I	109		VII-I	61		I-III	109
	I-VI	81		VII-VI	20		VI-I	81		VI-VII	20
	I-V	93		VII-V	31		V-I	93		V-VII	31
	I-IV	113	L1	VII-IV	50		IV-I	113	L2	IV-VII	50
L1	I-I	125		VII-I	60	L2	I-I	125		I-VII	60
	II-III	6		VI-V	12		III-II	6		V-VI	12
	II-VII	51		VI-IV	31		VII-II	51		IV-VI	31
	II-VI	71		VI-I	42		VI-II	71		I-VI	42
	II-V	83		V-IV	20		V-II	83		IV-V	20
	II-IV	103		V-I	31		IV-II	103		I-V	31
	II-I	115		IV-I	12		I-II	115		I-IV	12

The minimum dwell time at all stations is 2 min. The minimum headway between two successive trains is 4 min. Let penalties for time difference between the ideal time and actual schedule time be: $\beta^+ = 2\beta^- = 0.02$.

3) Demand setting: The OD information is shown in Table 9-Table 11. Without loss of generality, 3 OD pairs are randomly selected. The longest wait time at the origin station is $T^{\text{wait}} = 15$ min. The penalty for not finding any itinerary is $\dot{u}_r = 4000$ min/passenger.

Without loss of generality, we follow the assumption of Kitthamkesorn and Chen (2013) and set $\xi_r = 0$ and $\gamma_r = 3.7$.

Table 9. Passenger characteristics.

Passenger characteristics		Mean	Variance
Trip purpose	Business trip	1	0.2
	Leisure trip	2	0.3
Income (yuan/month)	$\leq$ 5,000	1	0.2
	5,000-10,000	2	0.3
	$\geq$ 10,000	3	0.5

Table 10. Coefficients for β_{rk}.

k	$\theta_{rk}^{\mathrm{IL}}$	$\theta_{rk}^{\mathrm{TP}}$	θ_{rk}
1	0.00	0.00	1.00
2	0.35	-0.30	1.75
3	-0.25	-0.30	1.55
4	0.40	-0.20	1.80
5	1.00	-1.00	3.00
6	-0.25	0.20	1.85

Table 11. OD information.

OD	Origin zone	Destination zone	Preferred departure time	Income level	Trip purpose	Volume
1	Z1	Z2	8:30	2	1	20
2	Z1	Z3	8:00	1	2	20
3	Z1	Z7	8:00	1	1	50

Appendix C. Schedule plans for the small example.

Table 12. Schedule plan of SPG (5).

Train	Station	Track	Arrival time	Departure time	Train	Station	Track	Arrival time	Departure time
1	I	0	-	8:00	9	V	0	10:54	10:54
	III	0	8:30	-		IV	0	11:19	11:19
2	I	1	-	8:04		I	1	11:39	-
	III	1	8:34	-	10	I	0	-	8:06
3	I	1	-	8:15		IV	1	8:26	8:26
	II	2	8:42	-		V	1	8:51	8:51
9	I	1	-	8:30		VI	1	9:11	9:11
	II	2	8:48	8:50		VII	1	9:36	9:38
	III	0	9:05	9:07		III	1	10:38	10:38
	VII	0	10:07	10:09		II	1	10:53	10:53
	VI	0	10:34	10:34		I	0	11:11	-

Table 13. Schedule plan whose generation adopts the MW model.

Train	Station	Track	Arrival time	Departure time	Train	Station	Track	Arrival time	Departure time
1	I	1	-	8:00	9	V	0	10:54	10:54
	III	0	8:30	-		IV	0	11:19	11:19
2	I	0	-	8:15		I	0	11:39	-
	III	0	8:59	-	10	I	0	-	8:00
3	I	1	-	8:15		IV	1	8:20	8:20
	II	2	8:42	-		V	1	8:45	8:45
9	I	0	-	8:30		VI	1	9:05	9:05
	II	2	8:48	8:50		VII	1	9:30	9:32
	III	0	9:05	9:07		III	1	10:32	10:32

	VII	0	10:07	10:09		II	1	10:47	10:47
	VI	0	10:34	10:34		I	1	11:05	-

Appendix D. Setting for HSR network of southern China.

The setting for HSR network of southern China is described here.

1) We select 34 main stations from southern China's HSR network. Each station serves one zone. Because two stations, Nanningdong and Zhaoqingdong, serve two train corridors at the same time, we simulate them as two substations connected by walking links. The example network thus has 36 abstract stations.
2) We collect train data including passing stations for each train, running time, dwell time, and ticket prices from China Railway's official website.[1] In this example, 240 high-speed trains operate per day within the network. The running period is from 6:00 to 23:30. The network contains four main high-speed train corridors, as shown in Figure 14. According to their passing stations, the 240 high-speed trains are grouped as 22 lines for the scheduling problem.
3) According to previous studies, the minimum headway between two successive high-speed trains is 4 min (Kroon and Peeters, 2003; Zhan et al., 2015; Tian et al., 2015).
4) The following time-dependent OD matrix is generated for testing (Table 14). OD pairs are divided into three categories according to the minimum travel time (MTT), which is the time spent on the highest-speed train on the most direct route between the origin and the destination without transfers. As shown in Table 14, the MTT of a long-distance trip is longer than 6 hours, the MTT of a medium-distance trip is between 2 and 6 hours, and the MTT of a short-distance trip is shorter than 2 hours.

Table 14. OD patterns for three categories.

Category	MTT	Peak		Off-peak		Number of OD pairs
		Period	Number of passengers for an OD pair	Period	Number of passengers for an OD pair	
Long-distance trip	> 6 hours	[8:30,10:30]	40	[5:30,8:30] and [10:30,14:00]	20	2340
Medium-distance trip	2–6 hours		100	[5:30,8:30] and [10:30,17:00]	50	4848
Short-distance trip	< 2 hours		60	[5:30,8:30] and [10:30,21:00]	30	1888
Sum	-					9076

We assume that there is a peak period from 8:30 to 10:30. For off-peak periods, the different categories have different travel-time windows. For example, to take a long-distance trip, passengers must depart early enough to arrive at their destinations before midnight. Passengers with the same origin and destination within 30 min are regarded as an OD pair and are assumed to prefer to start their journey at the end of this interval. Because of journey

[1] China Railway's official website: http://www.12306.cn/mormhweb/ (Accessed August 23, 2018)

time limitations and the various levels of attractiveness of zones, we generate 9076 OD pairs for this example.

CRediT author statement

Jiemin Xie: Conceptualization, Methodology, Software, Writing - Original Draft

S.C. Wong: Conceptualization, Writing - Review & Editing, Supervision

Shuguang Zhan: Methodology, Writing - Review & Editing

S.M. Lo: Supervision

Anthony Chen: Methodology