Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 8579354, 17 pages
https://doi.org/10.1155/2022/8579354

*Research Article*

# First-Train Timetable Synchronization in Metro Networks under Origin-Destination Demand Conditions

**Hetian Chai, Xiaopeng Tian, and Huimin Niu**

*School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China*

Correspondence should be addressed to Xiaopeng Tian; xptian@mail.lzjtu.cn

Received 13 November 2021; Revised 7 April 2022; Accepted 10 May 2022; Published 3 June 2022

Academic Editor: Yuan Gao



This paper focuses on how to synchronize network-wide timetables of first trains in an urban metro system, in which the train-connection-based route can be exactly determined for each first-train-attached origin-destination (OD) demand pair. With the help of even headway scheduling on each line, the problem is actually to adjust the departure times of first trains and connecting trains from their origin stations and the departure interval on each line. Subjected to train operation and connection constraints, a biobjective nonlinear integer programming model is formulated to minimize the total travel time of OD-dependent passenger demands and the deviation between the known and expected schedules. Then, the Nondominated Sorted Genetic Algorithm-II (NSGA-II) is adopted to solve the proposed model, and an improved technique is elaborated to reduce the alternative route choices. Finally, numerical experiments are conducted to demonstrate the effectiveness and availability of the proposed model and methods.

## 1. Introduction

As one of the network-wide train timetabling problems, the first train timetabling problem seeks to determine the departure times of the first trains in metro networks, to improve the connection efficiency of these first trains at transfer stations. Compared with the non-first train timetable synchronization problem, the first train timetabling problem is characterized by a longer train connection time between the first trains, which typically exceeds the maximum headway on lines. Such excessive passenger transfer waiting time can seriously decrease the travel quality of the passengers taking the first trains. It is necessary to design a high-level first-train timetable for these passengers. Furthermore, a well-designed first train synchronization timetable, which can provide high-quality transfer services to attract more passengers, is a matter of significant importance for operators.

The first train timetabling problem is known to be very complicated, so the existing literature has proposed some simplistic measures to address this problem. First, transfer direction-based passenger demand, which is actually the number of passengers from the feeder line to the connecting line at transfer stations, is considered to facilitate modeling and solving [1]. It is undeniable that such a passenger demand pattern is capable of representing passenger transfer behaviors when passenger travel routes are unique. However, in large-scale metro networks, passengers typically have several alternative transfer routes from the origin stations to designation stations, which means that passengers can flexibly select the transfer direction. That is, the transfer-direction-based demand may be a rough estimate of passenger transfer because it lacks a description of passenger choice behaviors. Second, the simplistic measures are from the identification of train connections. Based on the given train timetable, a typical approach is to optimize only the train connections between the first trains, ignoring the connections from the first trains to non-first connecting trains [2]. Although a standard mixed-integer linear programming model can be established by the above simplification, the corresponding first train synchronization efficiency is bound to be affected by the simplified connections and the given train timetable.

Therefore, this paper addresses passenger demand and train connection identification to improve the first train transfer efficiency. Specifically, we consider transfer passenger origin-destination (OD) demand that takes the first trains in the metro network and allows them to select flexible travel routes. Furthermore, all the connections from the first trains to the corresponding connecting trains are identified, and the variable headway on each line is considered. Thus, a biobjective optimization model is established to minimize the passenger travel times and the deviation times from the actual schedules. Meanwhile, the Nondominated Sorted Genetic Algorithm-II (NSGA-II) is designed to generate Pareto solutions for the problem under consideration.

*1.1. Literature Review.* The network-wide train timetabling problem has attracted extensive attention from the public transportation field, and the objective functions of the problem are diverse. First, much research has focused on single-objective functions for train timetable synchronization. Fleurent et al. [3] established a network flow model to reduce the deviation between the actual departure interval and the ideal interval. Cevallos and Zhao [4] presented a system-wide approach to minimize the total transfer times by shifting existing timetables. Considering heterogeneous transfer walking times, Chen et al. [5] proposed a mixed-integer quadratic programming model to maximize the number of transfer passengers for last train services on metro networks. A mixed-integer programming model was established by Chen et al. [6] to improve passenger accessibility. As for the first train timetabling problem, Guo et al. [2] proposed a first train timetable optimization model based on the importance of lines and transfer stations, which can minimize the total connection time at transfer stations coupled with the degree of importance of the station. Li et al. [7] proposed a reference-based piecewise function to maximize the satisfaction of passengers' transfer waiting times.

Second, multiple objective functions are often used to optimize the train transfer optimization problem. Chung Min Kwan and Chang [8] proposed a biobjective programming model to minimize the total passenger dissatisfaction index and total deviation index and employed the Nondominated Sorted Genetic Algorithm-II to generate Pareto solutions. Tian and Niu [9] considered a biobjective function with the maximum number of train connections and the minimum passenger waiting times for the network-wide train timetable problem. From the perspective of passengers and operators, Guo et al. [10] developed a multiobjective optimization model to determine optimal train service schedules with the maximum transfer events and the minimum transfer synchronization time. Kang et al. [11] simultaneously minimized passenger transfer waiting times, total train running times, and train energy consumption for the last train station-skipping problem. Furthermore, Nie et al. [12] established the last-shift schedule model to minimize the transfer waiting times and maximize the network connectivity. Obviously, it is more comprehensive to use multiple evaluation criteria to measure the complicated network-wide train timetabling problems. Note that, however, the studies associated with first train synchronization mainly focus on a single objective function with train connection efficiency.

In the design of train timetables, passenger demand plays an important role, such as Niu and Zhou [13], Barrena et al. [14], Niu et al. [15], Tian and Niu [16], and Shang et al. [17]. Due to the inherent complexity of train synchronization, transfer-direction-based passenger demand is usually considered to optimize network-wide train timetables. Given the number of passengers between each transfer direction, Wong et al. [18] presented a mixed integer programming optimization model to minimize the interchange waiting times of all passengers for this schedule synchronization problem with nonperiodic timetables. Similar to transfer-direction-based demand, Wang et al. [19] adopted arrival-rate-based passenger demand to establish a nonlinear integer programming model for the train transfer problem. Shafahi and Khani [20] formulated a mixed integer programming model to minimize the waiting time of station-based passenger demand for the transit network scheduling problem. For the first train synchronization problem, Kang et al. [1] established an optimization model by using the transfer-direction-based demand and analyzed the connectional regulation of first trains at transfer stations on each line. Indeed, such passenger demand patterns can effectively facilitate modeling and computation, but they cannot capture the travel choice behavior of passengers.

More recently, the passenger OD demand has also been applied to the network-wide train timetabling problem. For two interconnected high-speed rail lines, Niu et al. [21] proposed a nonlinear programming model based on the time-dependent OD demand to minimize passenger waiting times at stations and crowding disutility in trains. Li et al. [22] maximized the total number of passengers reaching their destinations in a network by using the OD demand. Zhou et al. [23] paid close attention to the entire travel process, including multiple transfer behaviors of passengers and loaded static OD passenger demand on the traveling path. For the last-train timetable problem, Yang et al. [24] expected to find a feasible train space-time path with accessibility for the OD passenger demand. Furthermore, Yin et al. [25] utilized the time-dependent OD demand pattern to construct a mixed-integer nonlinear programming model for the train timetabling problem with the transfer. Although the OD passenger demand has been considered for some train transfer problems, the first train timetable problem still lacks the application of the OD passenger demand.

*1.2. Focus of This Study.* To highlight the contributions of our proposed approach, Table 1 provides a systematic comparison with the most relevant literature. Specifically, we collaboratively optimize the train timetabling and the passenger routing by using the OD passenger demand instead of the transfer-direction-based passenger demand. Then, a biobjective function is proposed to more comprehensive formulate this problem. Furthermore, the NSGA-II algorithm is designed to solve the proposed biobjective optimization model. Note that, similar to the existing literature,

Table 1: Comparison between this paper and the most relevant literature.

| Publication | Problem type | Train capacity | Passenger demand | Objective function | Passenger travel route | Algorithm |
|---|---|---|---|---|---|---|
| Wong et al. [18] | All trains | No | Transfer-direction-based demand | Minimize waiting time | No | CPLEX-based heuristic |
| Shafahi and Khani [20] | All trains | No | Station-based demand | Minimize waiting time | No | GA |
| Zhou et al. [23] | Last train | No | Static OD demand | Maximize OD accessibility | Yes | CPLEX |
| Chen et al. [5] | Last train | No | Station-based demand | Maximize the number of transfer passengers | No | CPLEX |
| Chen et al. [6] | Last train | No | Static OD demand | Maximize percentage of accessible OD pairs | Yes | GA |
| Yang et al. [24] | Last train | No | Time-dependent OD demand | Maximize space-time accessibility | Yes | LR |
| Guo et al. [2] | First train | No | Transfer-direction-based demand | Minimize connection time | No | CPLEX |
| Kang et al. [1] | First train | No | Transfer-direction-based demand | Minimize waiting time | No | Heuristic |
| This paper | First train | No | Static OD demand | Minimize travel time<br>Minimize schedule deviation | Yes | NSGA-II |

this paper ignores the train capacity constraints to simplify this problem.

Compared with the existing research, the main contributions of this paper are summarized as follows:

(1) We consider the transfer OD passenger demand in the first train timetabling problem. Based on this OD demand, transfer passenger travel behaviors can be easily formulated during the modeling process, which is beneficial for designing passenger-oriented train timetables.

(2) To trade off transfer passenger travel efficiency and nontransfer passenger service quality, a biobjective nonlinear programming model is then established to minimize the transfer waiting times and the deviations between the actual schedule and the expected schedule on each line. In addition, a train connection-based route is introduced to calculate the total travel time of transfer passengers.

(3) An NSGA-II with customized acceleration strategies is proposed to effectively solve large-scale instances. These strategies are able to generate promising passenger travel routes and accelerate computational times by reducing the repetitive calculation of transfer times for different OD demands.

The remainder of this paper is organized as follows: the problem is described in detail in Section 2. In Section 3, a mixed integer programming model is established to describe the interest in the problem. Then, in Section 4, the method of determining the effective route alternative set and the OD demand classification method are introduced, and the implementation principle of the NSGA-II algorithm is also introduced. Section 5 conducts several numerical examples with different scales, and the authors conclude this study and provide further research directions in the last section.

## 2. Problem Statement and Passenger Travel Times

*2.1. Problem Statement.* This paper focuses on a directed urban metro network, where the set of lines is denoted by $L$, $l \in L$, the set of stations on line $l$ is indicated as $S_l$, $s \in S_l$, and the set of trains scheduled on line $l$ is represented by $I_l$, $i \in I_l$. Considering that the same-speed trains are scheduled in most metro systems, we assume that both the running times on each segment and dwelling times at each station for all trains are given.

In the metro network, let $K$ be the set of all transfer OD demands that need to take the first trains at their origin stations. For OD demand $k \in K$, let $k = (o_k, d_k, q_k)$, where $o_k$, $d_k$, and $q_k$ represent the origin station, the destination station, and the number of passengers from $o_k$ to $d_k$, respectively. The purpose of this paper is to generate high-quality network-wide train timetables for serving the considered transfer OD demand.

To facilitate modeling, we do not consider passenger waiting times at the origin stations and instead pay more attention to the passenger travel times. Such consideration is the same as in the existing literature [2], mainly for the following reasons. Specifically, in practice, the obtained first-train timetable is usually publicly announced, especially in China, which means that the passengers taking the first trains can select a reasonable time to arrive at the origin stations according to the published first-train timetables, thereby avoiding long waiting times. Additionally, the passengers taking the first trains may face a special transfer time exceeding the maximum headway of all lines due to the uncoordinated first-train timetables [1]. Thus, these passengers are typically more concerned about their travel times than waiting times at the origin stations. Furthermore, we specify in advance an actual timetable that can provide high-level service quality for the nontransfer passengers. By minimizing the deviation time between the optimized timetable and the actual timetable, we can indirectly guarantee the service quality of the nontransfer passengers.

In practice, the majority of the OD demands have multiple physical routes from the origin stations to the destination stations in metro networks. These physical routes are actually represented by a sequence of stations on different lines. In this paper, let $R_k$ denote all routes of demand $k$ ($r \in R_k$), and let $\overline{S}_r$ be all transfer stations in route $r$. At transfer station $s \in \overline{S}_r$, the corresponding feeder and connecting lines are indicated by $l_{r,s}^{\text{feeder}}$ and $l_{r,s}^{\text{connect}}$, respectively. As shown in Figure 1, the OD demand from stations 2 to 9 has two routes, i.e., $R_k = \{r_1, r_2\}$, and the sets of stations of the two routes are $\overline{S}_{r_1} = \{3, 7\}$ and $\overline{S}_{r_2} = \{4, 8\}$, respectively. At transfer station 3 in route $r_1$, we can obtain $l_{r_1,3}^{\text{feeder}} = 1\text{D}$ and $l_{r_1,3}^{\text{connect}} = 3\text{U}$.

To model and solve the first-train timetable synchronization problem as described previously, we integrate nonlinear modeling techniques, multiobjective optimization methods, and intelligent evolutionary algorithms. Specifically, the passenger travel times are used to measure the level-of-service of train timetables. However, the calculation of passenger travel times involves inherent nonlinear relationships caused by train connections between different lines. The nonlinear modeling technique is thus employed to represent the complicated calculation process of passenger travel times. Then, we systematically consider four kinds of constraints and two objective functions to build a biobjective first-train timetable synchronization model. Based on the multiobjective optimization method, an intelligent evolutionary algorithm (i.e., NSAG-II) is finally designed to solve the proposed biobjective optimization model. An illustrative description of the research process is shown in Figure 2.

*2.2. Passenger Travel Times.* Passenger travel times are often used as an important evaluator for passenger-oriented train timetabling problems. Generally, the travel times are composed of train running times, transfer waiting times, and transfer times. The train running times can be easily obtained according to the train operating process. The passenger transfer waiting times need to identify the effective connections between different trains at transfer stations, which is a more complicated task. Furthermore, the key to calculating travel times is determining the connection times between different trains at transfer stations.

For simplicity, we adopt two common assumptions: (1) the walking times of all passengers at transfer stations are fixed [18] and (2) the train capacity is not considered so that passengers can board the first arriving train [26]. That is, a successful train connection must satisfy the requirements that the difference between the departure time of the connecting train and the arrival time of the feeder train is minimal and not less than the transfer time. As shown in Figure 3, the connecting train of train 1 on line 1D is only train 2 on line 3U at transfer station 3.

From the two above-given assumptions, we can obtain that a feeder train corresponds to the unique connecting train at the transfer station. Meanwhile, it has been mentioned that the transfer OD demand only takes the first train at the origin station. Thus, the passengers along a given physical route definitely have unique train trajectories when the network-wide train timetable is known. In this paper, the train trajectory is called a train-connection-based route. In terms of a single passenger, his or her travel time along the given route is actually the duration of the corresponding train-connection-based route, which is equal to the difference between the arrival time of the train serving him or her at the destination station and the departure time of the first train from the origin station. For clarity, let $TA_{l,i}^{s}$ and $TD_{l,i}^{s}$ denote the arrival and departure times of train $i$ at station $s$ on line $l$, respectively.

Taking route $r_1$ in Figure 1 as an illustrative example, the given train timetables on the lines involved in route $r_1$ are shown in Figure 2. Clearly, for the OD demand $k$ from stations 2 to 9, Figure 2 shows that their travel route along $r_1$ is unique. Specifically, they take the first train at station 2 on line 1D, then transfer to train 2 on line 3U at station 3, and finally transfer to train 4 on line 2D at station 7. Furthermore, the travel time is the arrival time of train 4 at station 9 on line 2D minus the departure time of train 1 at station 2 on line 1D, i.e., $TA_{2\text{D},4}^{9} - TD_{1\text{D},1}^{2}$.

As mentioned previouly, many passengers need to transfer multiple times from their origin stations to the destination stations. To model the passenger travel process, it is necessary to record the trains involved in the passenger travel process. To this end, we introduce the symbol $\alpha(r, l, s)$ to indicate the corresponding connecting train on line $l$ at transfer station $s$ in route $r$. In fact, the identification of connecting trains at transfer stations is a recursive iterative process. For example, at transfer station 3 in Figure 2, the feeder train must be the first train on line 1D; thus, we easily obtain that train 2 on line 3U is its connecting train because this train satisfies $\min\{TD_{3\text{U},2}^{3} - TA_{1\text{D},1}^{3} \mid TD_{3\text{U},2}^{3} - TA_{1\text{D},1}^{3} \geq t_{1\text{D},3\text{U}}^{3}, 2 \in I_{3\text{U}}\}$. Furthermore, at the next transfer station 7, train 2 on line 3U can be used as the feeder train, so its connecting train (i.e., train 4 on line 2D) can be recursively obtained by the same method. By analogy, we can identify the associated trains on the route. Thus, the symbol $\alpha(r, l, s)$ can be represented as follows:

$$\alpha(r, l, s) = \underset{i \in I_l}{\operatorname{argmin}} \left\{ TD_{l,i}^{s} - TA_{l',\alpha(r,l',s^-)}^{s} \mid TD_{l,i}^{s} - TA_{l',\alpha(r,l',s^-)}^{s} \geq t_{l',l}^{s}, l' = l_{r,s}^{\text{feeder}} \right\} l = l_{r,s}^{\text{connect}}, s \in \overline{S}_r, r \in R_k, \quad (1)$$

where $t_{l',l}^{s}$ is the fixed transfer time from lines $l'$ to $l$ at transfer station $s$ and $\alpha(r, l', s^-)$ is the connecting train at the previous transfer station (denoted by $s^-$) of station $s$ on line $l'$ in route $r$. Furthermore, the passenger travel time of OD

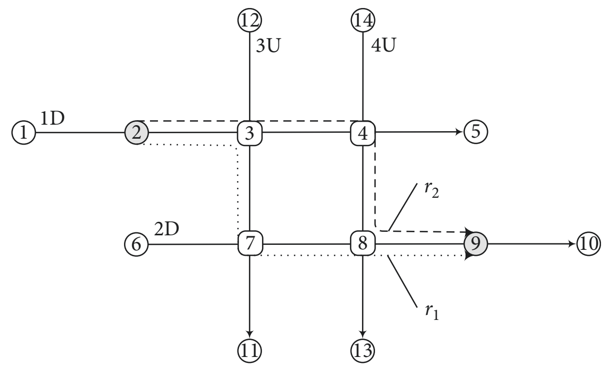


FIGURE 1: Illustration of the metro network and passenger transfer route.

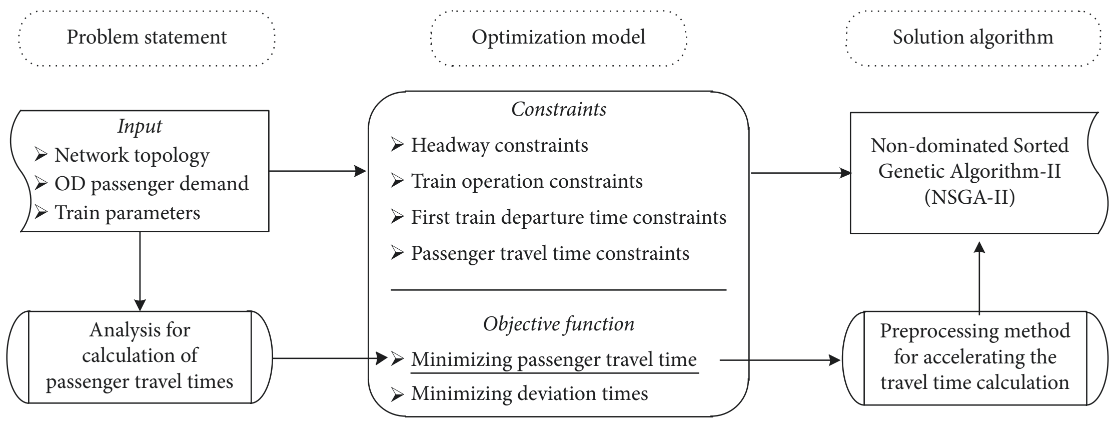


FIGURE 2: Illustration of the research process outline.

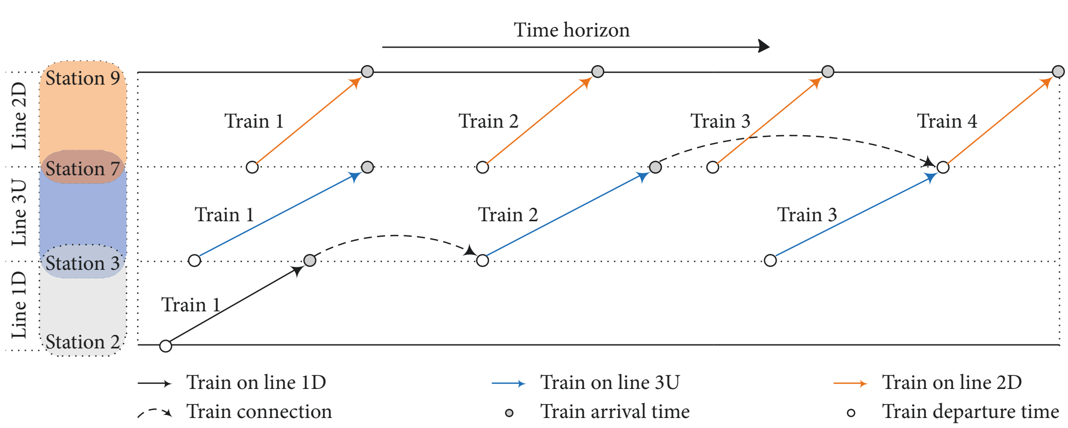


FIGURE 3: Train-connection-based route.

demand $k$ along route $r$, denoted by $T_k^r$, can be easily formulated as follows:

$$\begin{aligned} T_k^r &= q_k \cdot \left( TA_{l,\alpha(r,l,r^+)}^{d_k} - TD_{l',1}^{o_k} \right) \\ l &= l_{r,r^+}^{\text{connect}}, \\ l' &= l_{r,r^-}^{\text{feeder}}, r \in R_k, k \in K. \end{aligned} \tag{2}$$

where $r^-$ and $r^+$ are the first and last transfer stations in route $r$, respectively.

From the above-given analysis, it can be seen that the calculation of passenger travel time is a very complicated process, which involves the passenger in-train travel times and the transfer waiting times. The passenger in-train travel times depend utterly on the train running time on each segment. In the metro network, though the train running time is fixed, the passenger in-train travel times of different travel routes are distinguishable. Also, the passenger transfer waiting times can be determined by the corresponding train connections between the feeder trains and the connecting trains. To exactly calculate the passenger travel time, the passenger route selection and the train connection identification must be combined. Note that, such a combination greatly increases the complexity of this problem so that we need to explore a very large solution space for finding the optimal solution.

## 3. First-Train Timetable Synchronization Model

*3.1. Notations.* In this study, we introduce the following two types of decision variables:

$TD_{l,1}^{s(l)}$: departure time of the first train on line $l$ at its starting station $s(l)$

$H_l$: headway of two adjacent trains on the line $l$

For the convenience of readers, the notations used in the formulation of our problem are listed in Table 2.

*3.2. Constraints*

*3.2.1. Headway Constraints.*

$$H_l = TD_{l,i+1}^{s(l)} - TD_{l,i}^{s(l)} i, i+1 \in I_l, l \in L, \quad (3)$$

$$h_l^{\min} \leq H_l \leq h_l^{\max} l \in L. \quad (4)$$

Constraint (3) can ensure that the time interval of two adjacent trains is equal to headway $H_l$. That is, using the decision variables $TD_{l,1}^{s(l)}$ and $H_l$, we can easily obtain the departure times of nonfirst trains at the corresponding starting stations. Constraint (4) requires that the decision variable $H_l$ satisfy the minimum and maximum restrictions [27].

*3.2.2. Train Operation Constraints.*

$$TA_{l,i}^{s} = TD_{l,i}^{s-1} + t_{l,s}^{\text{run}}, \quad s \in S_l \backslash \{s(l)\}, i \in I_l, l \in L, \quad (5)$$

$$TD_{l,i}^{s} = TA_{l,i}^{s} + t_{l,s}^{\text{dwell}}, \quad s \in S_l \backslash \{s(l), e(l)\}, i \in I_l, l \in L. \quad (6)$$

These constraints can determine the arrival and departure times of all trains at each station according to the given running times and dwell times.

*3.2.3. First Train Departure Time Constraints.*

$$t_l^{\text{early}} \leq TD_{l,1}^{s(l)} \leq t_l^{\text{late}} l \in L. \quad (7)$$

In a metro network, the departure time of the first train at the starting station on each line is required to satisfy the earliest and latest departure times. In this paper, the earliest departure times are equal to the prespecified starting operation times, which can be comprehensively determined by many factors, such as passenger demand, maintenance constraints, and station management.

*3.2.4. Passenger Travel Time Constraints.*

$$T_k = \min_{r \in R_k} \{T_k^r\}. \quad (8)$$

Here, $T_k$ denotes the shortest travel time of demand $k$. As mentioned previously, OD demand $k$ in the metro network usually corresponds to several travel routes $(R_k)$. Thus, the travel time of demand $k$ is equal to the minimum travel time in $R_k$, namely, $\min_{r \in R_k}\{T_k^r\}$. Note that, the travel time of each route $(T_k^r)$ can be determined by (2), which is associated with the decision variables.

*3.3. Objective Functions.* In this paper, the optimization objectives of the first train timetable synchronization problem are considered from the following two aspects.

First, the minimum total transfer passenger travel time is regarded as one optimization objective. As mentioned previously, any OD demand $k$ corresponds to several travel routes. In general, passengers expect to select the shortest route for their journey. We use $T_k$ to denote the shortest travel time of demand $k$, i.e., $T_k = \min_{p \in P_k}\{T_k^p\}$. Therefore, the minimum total transfer passenger travel time, denoted by $Z_1$, can be represented as follows:

$$\min Z_1 = \sum_{k \in K} T_k. \quad (9)$$

Second, the minimum deviation between the actual schedule and the optimized schedule, denoted by $Z_2$, is used as the other objective function. Similar to Kwan et al. [8], this paper introduces the actual schedule to guarantee the service level for nontransfer passengers. Let $\overline{H}_l$ indicate the actual headway of trains on line $l$, and $\overline{D}_l$ indicate the actual departure time of the first train at the origin station on line $l$. Furthermore, it is required that the optimized schedule is as close as possible to the actual schedule. The corresponding objective function can be written as follows:

$$\min Z_2 = \sum_{l \in L} \left( \left| TD_{l,1}^{s(l)} - \overline{D}_l \right| + \left| H_l - \overline{H}_l \right| \right). \quad (10)$$

*3.4. Model Analysis.* Furthermore, the above first-train timetable synchronization model (FTSM) can be summarized as follows:

Model FTSM. Objective function (9) and (10) Constraints (1)–(8)

The above-given model FTSM is a typical biobjective nonlinear optimization model. Between the two objective functions (9) and (10), there exists a conflicting relationship that the deviation times may not be optimal if the passenger travel times are minimized and vice versa. Such a conflicting relationship easily causes severe solving difficulty to simultaneously minimize these two objectives. Additionally, the objective function (10) with the absolute value and the constraints (1), (2), and (8) are nonlinear, which further increases the complexity of the model FTSM.

Table 2: Sets, indices, parameters, and variables.

| Sets | Definition |
|---|---|
| $L$ | Set of lines in the urban rail transit network |
| $S_l$ | Set of stations on line $l$ |
| $I_l$ | Set of trains scheduled on line $l$ |
| $K$ | Set of OD demand in the urban rail transit network |
| $R_k$ | Set of transfer routes of OD demand $k$ |
| $\overline{S}_r$ | Set of transfer stations in route $r$ |
| *Indices* | |
| $l, l'$ | Index of lines, $l, l' \in L$ |
| $i$ | Index of trains, $i \in I_l, l \in L$ |
| $s$ | Index of stations, $s \in S_l, l \in L$ |
| $k$ | Index of OD demand, $k \in K$ |
| $r$ | Index of transfer route of OD demand $k$, $r \in R_k, k \in K$ |
| *Parameters* | |
| $s(l)$ | Starting station on line $l$ |
| $e(l)$ | Terminal station on line $l$ |
| $o_k$ | Origin station of OD demand $k$ |
| $d_k$ | Destination station of OD demand $k$ |
| $q_k$ | Number of passengers for OD demand $k$ |
| $r^-$ | The first transfer station in route $r$ of OD demand $k$ |
| $r^+$ | The last transfer station in route $r$ of OD demand $k$ |
| $l_{r,s}^{\text{feeder}}$ | Feeder line at station $s$ in route r |
| $l_{r,s}^{\text{connect}}$ | Connecting line at station $s$ in route $r$ |
| $t_l^{\text{early}}$ | Earliest departure time of the first train from the starting station online $l$ |
| $t_l^{\text{late}}$ | Latest departure time of the first train from the starting station online $l$ |
| $h_l^{\max}$ | Maximum departure interval of two adjacent trains on line $l$ |
| $h_l^{\min}$ | Minimum departure interval of two adjacent trains on line $l$ |
| $t_{l,s}^{\text{run}}$ | Train running time from stations $s-1$ to $s$ on line $l$ |
| $t_{l,s}^{\text{dwell}}$ | Train dwell time at station $s$ on line $l$ |
| $t_{l',l}^{s}$ | Transfer time from lines $l'$ to $l$ at transfer station $s$ |
| *Variables* | |
| $TA_{l,i}^{s}$ | Arrival time of train $i$ at station $s$ on line $l$ |
| $TD_{l,i}^{s}$ | Departure time of train $i$ at station $s$ on line $l$ |
| $\alpha(r,l,s)$ | Connecting train on line $l$ at transfer station $s$ in route $r$ |
| $T_k^r$ | Travel time of route $r$ of OD demand $k$ |
| $T_k$ | Minimum travel time of OD demand $k$ |

For the multiobjective optimization problem, nondominated sorting is a necessary step in the solution process. To represent the complexity of nondominated sorting in our proposed model, let $ND_l$ denote the set of departure times of the first train at the starting station on line $l$, $ND_l = \{t_l^{\text{early}}, t_l^{\text{early}}+1, \ldots, t_l^{\text{late}}\}$, and $NH_l$ indicate the set of accessible headways on line $l$, $NH_l = \{h_l^{\min}, h_l^{\min}+1, \ldots, h_l^{\max}\}$. Furthermore, the number of feasible solutions, denoted by $N$, for the proposed model can be obtained, namely, $N = \prod_{l=1}^{|L|}(|ND_l| \cdot |NH_l|)$. Similar to the discussion of Wu et al. [28], the computational complexity of nondominated sorting in our problem can be expressed as $O(2N^2)$. For the rail network with 4 lines shown in Figure 1, when assuming $|ND_l| = 5$ and $|NH_l| = 5$ for all lines, the number of feasible solutions is $N = 390{,}625$, and then the computational complexity of nondominated sorting is 305,175,781,250. Obviously, the computational complexity of nondominated sorting greatly increases the difficulty of solving this problem.

Based on the above analysis, the considered problem is very complex. Actually, the train synchronization optimization problem is known to be NP-hard [18]. It is difficult to find the optimal solutions in large-scale and complex instances. Thus, using multiobjective evolutionary algorithms (e.g., NSGA-II) is a good choice to overcome the difficulties mentioned previously.

## 4. Solution Algorithm

*4.1. Algorithm Framework.* For the above biobjective optimization model, we employ the NSGA-II to generate its Pareto solutions. During the solution process, the calculation of objective function $Z_1$ (i.e., minimizing the total travel time) is more complicated. Specifically, we need to search all the routes of each OD demand and calculate the corresponding travel times. However, in large-scale urban rail networks, there are thousands of OD demands, and each OD demand has a large number of routes. Obviously, the calculation of the travel times in the objective function is a very time-consuming process. Therefore, a preprocessing method (see Section 4.2) is embedded in the NSGA-II to speed up the travel time calculation.

Furthermore, the overall algorithm framework is shown in Figure 4. First, we need to input the associated information, such as the network topology, OD demand, actual schedule, and algorithm parameters, in the initialization stage. Second, we implement the preprocessing method and randomly generate several feasible solutions to activate the NSGA-II algorithm. Then, the key steps and basic genetic operations of the NSGA-II are executed. Finally, a new population is generated, and the corresponding objective value of each chromosome can be rapidly calculated by using the preprocessing method.

*4.2. Preprocessing Method for Accelerating the Travel Time Calculation.* As mentioned previously, each OD demand corresponds to numerous physical routes in large-scale metro networks. In practice, many of these routes may be unfavorable for passenger travel because of their longer travel times [29]. To save computational time, we select some promising routes instead of all routes to calculate the objective function. Specifically, according to the departure time windows of the first trains at the starting stations and the actual schedule of each line, we estimate the minimum and maximum travel times of each physical route using the enumerate method. Then, given any two routes $r$ and $r'$, we will take route $r$ as the promising route if the maximum travel of route $r$ is less than the minimum travel time of route $r'$.

We continue to take the small-scale metro network shown in Figure 1 as an example. The running time on each segment is shown in Figure 5. The transfer times at all transfer stations are set as 3 min, the dwelling times of all trains at each station are 1 min, and the actual headway of each line is 7 min. Furthermore, the departure time windows of lines 1D, 2D, 3U, and 4U are [5:50, 6:10], [5:55, 6:15], [5:55, 6:15], and [6:00, 6:20], respectively. For the OD demand from stations 1 to 10 in this network, there are two routes, i.e., $r_1 = (1 \longrightarrow 2 \longrightarrow 3 \longrightarrow 7 \longrightarrow 8 \longrightarrow 9 \longrightarrow 10)$ and $r_2 = (1 \longrightarrow 2 \longrightarrow 3 \longrightarrow 4 \longrightarrow 8 \longrightarrow 9 \longrightarrow 10)$. Using the enumeration method, we estimate the minimum and maximum travel times of the two routes, as shown in Table 3. Obviously, since the maximum travel time of route $r_1$ is less than that of route $r_2$, we select route $r_1$ as the promising route.

In addition, we mentioned previously that the key to the travel time calculation is to determine the train connection time at transfer stations. In practice, the routes of many OD demands have common transfer directions. That is, the travel times of these routes include the same train connection times. The OD demands (1, 9), (2, 9), (1, 10), and (2, 10) shown in Figure 4 can complete their journeys along route $r_1$ (i.e., $1 \longrightarrow 2 \longrightarrow \mathbf{3} \longrightarrow \mathbf{7} \longrightarrow 8 \longrightarrow 9 \longrightarrow 10$). Obviously, these OD demands have the same train connections at transfer stations 3 and 7. Thus, it is necessary to avoid repeatedly calculating train connection times to accelerate the computational time. To this end, we first determine the train connection times of each first train at transfer stations from the given network-wide timetables and then save the values of these train connection times to a data pool. Subsequently, the travel time calculation of each OD demand directly uses the associated connection times from the data pool.

*4.3. NSGA-II.* In order to clearly describe the NSGA-II algorithm, let *pop_size* denotes the population size, *Gen_max* indicates the maximum number of iterations, and *l_num* denotes the number of lines in the network. An integer coding chromosome with a length of $2 * l\_num$ is employed in the genetic operation. The first *l_num* genes express the departure times of the first trains on each line, and the last *l_num* genes indicate the headways on each line.

Figure 6 uses a chromosome with 8 genes to denote the corresponding solution to the problem illustrated in Figure 1. Specifically, the first four genes represent the departure times of the first trains on four lines, and the remaining genes denote the headways on all lines. Furthermore, if we assume that the starting timestamp of the study period is 6:00, then the departure times of the first trains on lines lD, 2D, 3U, and 4U are 6:10, 6:15, 6:21, and 6:08, respectively; and the corresponding headways are 8 min, 7 min, 6 min, and 4 min on the four lines, respectively.

The crossover and mutation operations in this paper are similar to Li et al. [30]. In the crossover operation, two crossover points are randomly generated in the range of [1, *l_num*] and [*l_num* + 1, 2 * *l_num*], respectively. In addition, the crossover operation is performed with the given crossover probability. In the mutation process, two feasible values are randomly generated for the departure times and headways, respectively. Likewise, if the randomly generated floating-point number is less than the mutation probability, the mutation operation is conducted.

To calculate the crowding distance of different objectives, the following equation is established:

$$D_w^i = \left(f_w^{i+1} - f_w^{i-1}\right)/\left(f_w^{\max} - f_w^{\min}\right). \tag{11}$$

Here, $f_w^{\max}$ and $f_w^{\min}$ represent the maximum and minimum values of the wth objective function at this level, respectively, and $f_w^{i+1}$ and $f_w^{i-1}$ represent the wth objective function values corresponding to the $i+1$th and $i-1$th individuals at this level, respectively.

In the process of generating a new parent population, individuals with higher levels are selected from the combined population to enter the new parent population. If the total number of individuals in the same level exceeds the population size of the new parent population, the individuals with a large crowding distance in the level are selected to enter until the number of individuals reaches the population size. The detailed procedure is represented in Algorithm 1.

## 5. Numerical Experiments

In this section, different-scale experiments regarding the Xi'an metro network are conducted to verify the correctness and effectiveness of the above model and algorithm. All experiments are implemented by C++ on a PC with an i5-4590s 3.00 GHz CPU and 4 GB RAM.

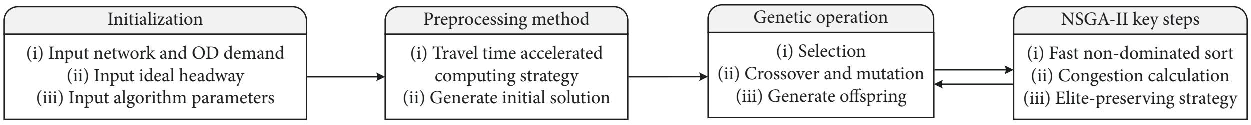


Figure 4: Algorithm framework.

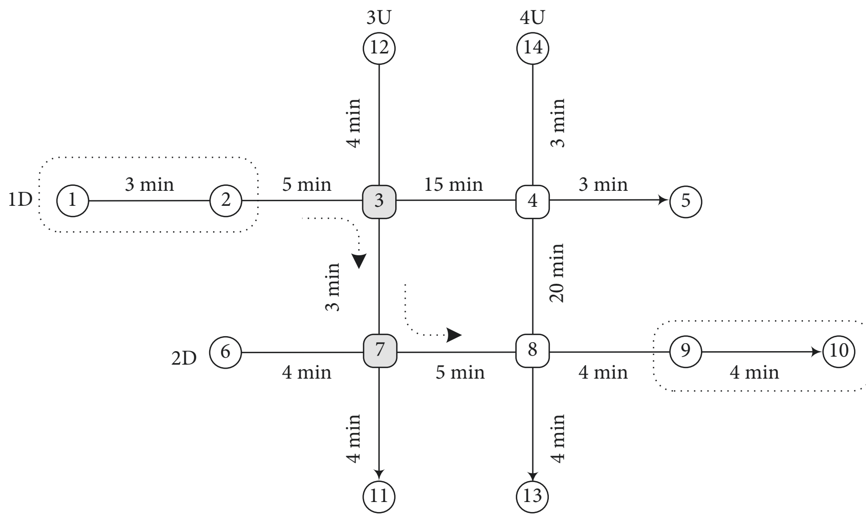


Figure 5: Illustration of the running time on each segment.

Table 3: Feasible route of OD demand from stations 1 to 10.

| Departure times and travel time | | Route $r_1$ | | Route $r_2$ | |
|---|---|---|---|---|---|
| | | Minimum case | Maximum case | Minimum case | Maximum case |
| Departure time of the first train at the starting station on each line | 1D | 5:50 | 5:50 | 5:50 | 5:50 |
| | 2D | 5:56 | 5:59 | 5:55 | 6:00 |
| | 3U | 5:57 | 6:15 | — | — |
| | 4U | — | — | 6:00 | 6:06 |
| Travel time | | 33 | **57** | **60** | 72 |

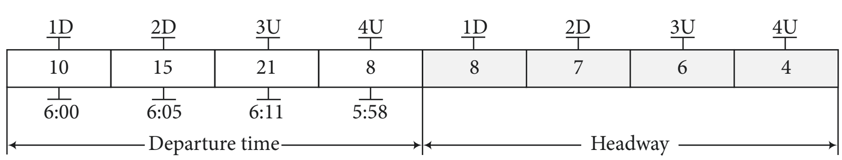


Figure 6: Chromosome structure for four lines.

*5.1. Large-Sized Numerical Experiment.* We take the Xi'an urban rail transit network, which includes 16 lines (8 bidirectional lines) and 153 stations, as the large-sized numerical experiment. In this experiment, each bidirectional line is used as two unidirectional lines. For example, line 1 represents the up-train and down-train directions as lines 1U and 1D, respectively. The lines and stations are renumbered as shown in Figure 7. The running times on each segment and the OD demands are listed at https://github.com/xpTianLZJT/TTP/blob/main/JAT2021/FurtherExperiments.zip. The dwell times of all trains at each station are 1 min, the passenger transfer walking times at each transfer station are 3 min, and the actual headway of each line is 7 min. In addition, according to the actual operation requirements of the Xi'an metro, we set the minimum headway and the maximum headway to 2 minutes and 10 minutes, respectively. Table 4 shows the earliest and latest departure times of the first trains on all lines.

We set the size of the population to 200, the crossover probability to 0.9, the mutation probability to 0.2, and the maximum number of iterations to 500. To reduce the influence of randomness in the NSGA-II algorithm, we repeatedly perform the NSGA-II algorithm 50 times for the large-scale experiment. Furthermore, we select the most typical 5 groups of Pareto solutions from multiple

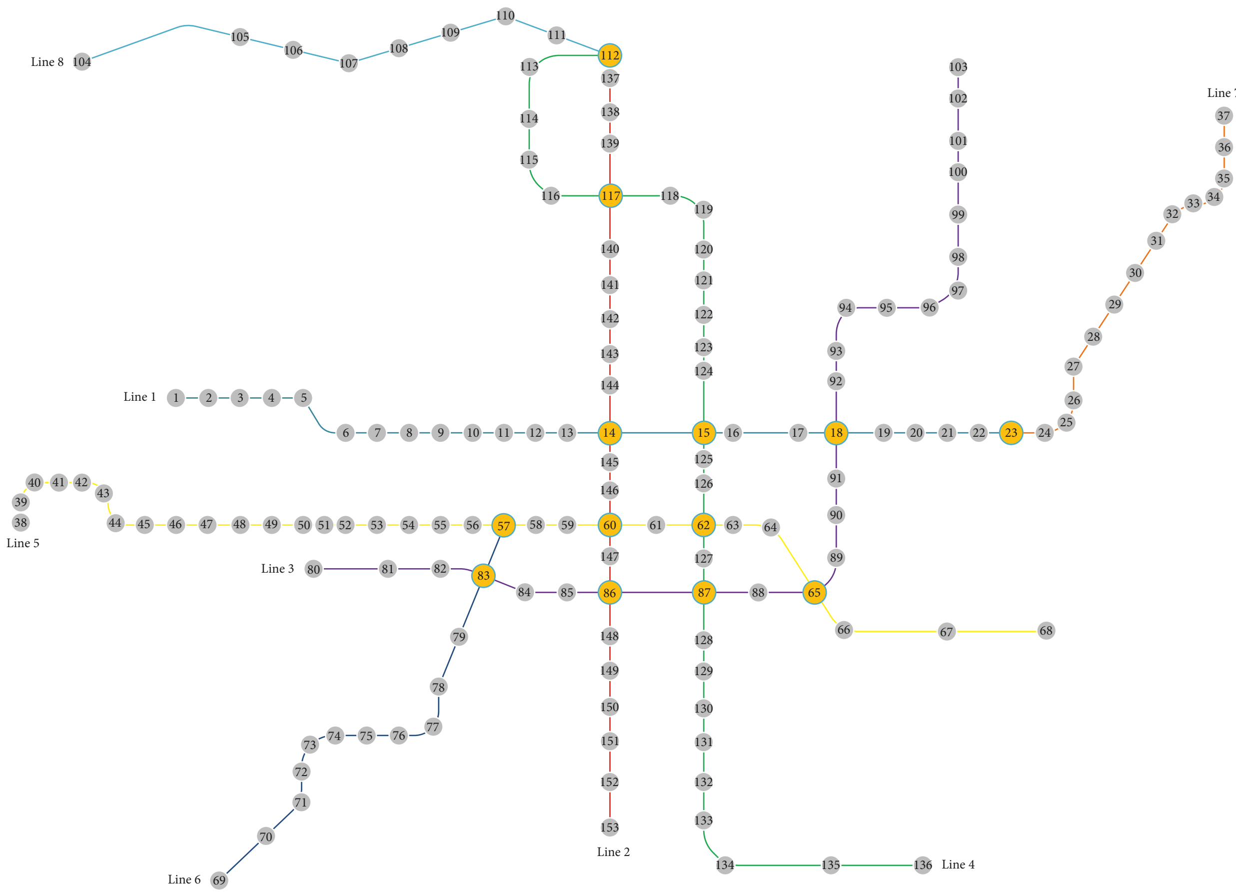


FIGURE 7: Map of the Xi'an metro network.

calculations, which are shown in Figure 8. For convenience, each solution is denoted by the corresponding coordinate of the objective value, i.e., (*value of passenger travel time*, *value of schedule deviation time*). For example, the first solution with a travel time of 5648027 min and a deviation time of 36 min is represented by (5648027, 36).

To compare the difference between the optimized timetable and the actual timetable, Table 5 lists the departure times of the first trains at the starting stations in the actual timetable, and the actual headways of all lines are equal to 7 min. The total travel time of transfer passengers is 5729497 min.

From the solutions in Figure 8, we can find that the passenger travel time of the solution (5236686, 99) is the smallest, which is 8.6% less than that of the actual schedule. The corresponding headways, the departure times of the first trains, and the deviations from the actual timetable are shown in Table 6, where the deviation is represented by the optimized headway (departure time) minus the actual headway (departure time). For example, the deviation headway of line 1D is −2 (i.e., −2 = 5−7). In addition, for the solution (5648027, 36) with the minimum deviation, the total passenger travel time is also the largest, which is reduced by 1.4% compared with the actual schedule. In practice, operators can select a suitable solution according to practical requirements.

Table 7 lists some train connection relationships from the possible feeder lines to the connecting lines 1U and 1D in solution (5236686, 99). At transfer station 14, (1, 2D)⟶(1, 1D) represents the first train on line 2D is connected by the first train on line 1D. Similarly, the passengers on the first train of line 3D can transfer to the third train on line 1U at transfer station 18. Note that there is a unique transfer direction at station 23. Among the optimized connection in Table 7, the "first train-first train" connection occurs six times, accounting for 46.2%. From the above-given results, we can further verify that not all passengers taking the first trains transfer to the first train on the connecting line. Also, compared with the actual train timetable, the connection efficiency and transfer waiting time of the optimized train timetable are improved.

### 5.2. Further Experiments

*5.2.1. Comparison of Different-Sized Experiment Instances.* By using different-sized networks, we further compare the efficiency of the proposed algorithm with the NSGA-II without the preprocessing method. These different-sized networks are constructed based on the above-mentioned Xi'an metro network, which is shown in Table 8. The detailed information of these instances is available at https://github.com/xpTianLZJT/TTP/blob/main/JAT2021/FurtherExperiments.zip. Among the

Step 1: (Initialization)
Step 1.1: Input the number and index of the lines in the network and the number and index of the intermediate stations and transfer stations on each line;
Step 1.2: Input OD demand set $K$ and promising routes;
Step 1.3: Input the actual schedule of each line in the network and other parameters;
Step 1.4: Input algorithm parameters, such as the population size *pop_size*, maximum number of iterations *Gen_max*, crossover probability, and mutation probability;
Step 2: (Initial population)
**For** $i = 1$ to *pop_size* **do**
Generate *l_num* departure times of first trains and headways of each individual randomly;
**Endfor**//*i*
Generate the parent population;
$Gen = 1$;
Step 3: (Sorting)
Step 3.1:
**For** $j = 1$ to *pop_size* **do**
Calculate the nondominated sorting rank of each individual;
**Endfor**//*j*
Step 3.2:
Calculate the crowded distance of each individual in the same rank;
Step 4: (Genetic operation)
According to the binary tournament selection rules, *pop_size*/2 individuals are selected from the parent population and the *pop_size*/2 individuals are genetically operated according to the randomly generated crossover probability and mutation probability until other *pop_size*/2 individuals are generated;
Step 5: (Offspring population)
These *pop_size* individuals generated in Step 4 compose the offspring population;
Step 6: (Combined population)
The parent population generated in Step 7 (if $Gen = 1$, generated in Step 2) and offspring population generated in Step 5 compose the combined population ($2 * pop\_size$);
Step 7: (New population)
Calculate the nondominated sorting rank of each individual in the combined population and calculate the crowded distance of each individual in the same rank;
Select *pop_size* new individuals from the combined population to compose the new parent population;
Step 8: (Solution)
**If** $Gen < Gen\_max$, **then**
Go to step 4;
Else
Output final solutions, such as the Pareto front, synchronized timetable, and train connection sequence.
Endif

Algorithm 1: NSGA-II.

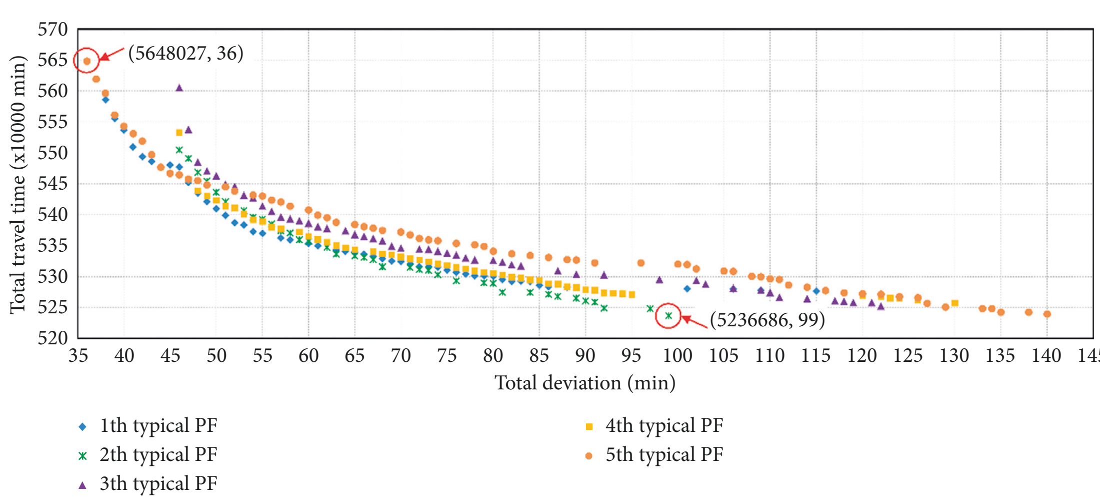


Figure 8: Pareto solutions of the large-scale experiment.

Table 4: The earliest and latest departure times of the first trains.

| Line | Earliest departure time | Latest departure time |
|---|---|---|
| 1D | 5:50 | 6:10 |
| 1U | 5:50 | 6:10 |
| 2D | 5:50 | 6:10 |
| 2U | 5:50 | 6:10 |
| 3D | 5:50 | 6:10 |
| 3U | 5:50 | 6:10 |
| 4D | 5:50 | 6:10 |
| 4U | 5:50 | 6:10 |
| 5D | 5:50 | 6:10 |
| 5U | 5:50 | 6:10 |
| 6D | 5:50 | 6:10 |
| 6U | 6:05 | 6:25 |
| 7D | 6:20 | 6:40 |
| 7U | 6:20 | 6:40 |
| 8D | 5:50 | 6:10 |
| 8U | 5:50 | 6:10 |

Table 5: Departure times of the first trains in the actual timetable.

| Line | Departure time |
|---|---|
| 1D | 6:00 |
| 1U | 6:00 |
| 2D | 6:00 |
| 2U | 6:00 |
| 3D | 6:00 |
| 3U | 6:00 |
| 4D | 6:00 |
| 4U | 6:00 |
| 5D | 6:00 |
| 5U | 6:00 |
| 6D | 6:00 |
| 6U | 6:15 |
| 7D | 6:30 |
| 7U | 6:30 |
| 8D | 6:00 |
| 8U | 6:00 |

Table 6: Optimized headways and departure times of the first trains for solution (5236686, 99).

| Line | Optimized headway | Deviation from the actual headway (min) | The optimized departure time of the first train | Deviation from the actual departure time of the first train (min) |
|---|---|---|---|---|
| 1D | 5 | −2 | 5:53 | −7 |
| 1U | 6 | −1 | 6:04 | +4 |
| 2D | 3 | −4 | 6:02 | +2 |
| 2U | 6 | −1 | 6:02 | +2 |
| 3D | 6 | −1 | 6:00 | 0 |
| 3U | 4 | −3 | 5:51 | −9 |
| 4D | 2 | −5 | 5:56 | −4 |
| 4U | 2 | −5 | 5:53 | −7 |
| 5D | 6 | −1 | 5:51 | −9 |
| 5U | 3 | −4 | 6:03 | +3 |
| 6D | 10 | +3 | 5:58 | −2 |
| 6U | 2 | −5 | 6:20 | +5 |
| 7D | 5 | −2 | 6:29 | −1 |
| 7U | 8 | +1 | 6:29 | −1 |
| 8D | 7 | 0 | 5:58 | −2 |
| 8U | 8 | +1 | 6:02 | +2 |

Table 7: Connection relationships for the connecting lines 1U and 1D in solution (5236686, 99).

| Transfer station | Optimized connection relationship | Transfer waiting time (min) | Actual connection relationship | Transfer waiting time (min) |
|---|---|---|---|---|
| 14 | (1, 2D) ⟶ (1, 1D) | 0 | (1, 2D) ⟶ (1, 1D) | 10 |
| | (1, 2D) ⟶ (1, 1U) | 0 | (1, 2D) ⟶ (2, 1U) | 3 |
| | (1, 2U) ⟶ (2, 1D) | 0 | (1, 2U) ⟶ (1, 1D) | 5 |
| | (1, 2U) ⟶ (2, 1U) | 1 | (1, 2U) ⟶ (3, 1U) | 5 |
| 15 | (1, 4D) ⟶ (1, 1D) | 0 | (1, 4D) ⟶ (1, 1D) | 3 |
| | (1, 4D) ⟶ (2, 1U) | 0 | (1, 4D) ⟶ (3, 1U) | 0 |
| | (1, 4U) ⟶ (1, 1D) | 1 | (1, 4U) ⟶ (1, 1D) | 2 |
| | (1, 4U) ⟶ (2, 1U) | 1 | (1, 4U) ⟶ (4, 1U) | 6 |
| 18 | (1, 3D) ⟶ (1, 1D) | 1 | (1, 3D) ⟶ (1, 1D) | 9 |
| | (1, 3D) ⟶ (4, 1U) | 0 | (1, 3D) ⟶ (5, 1U) | 3 |
| | (1, 3U) ⟶ (1, 1D) | 13 | (1, 3U) ⟶ (1, 1D) | 12 |
| | (1, 3U) ⟶ (2, 1U) | 0 | (1, 3U) ⟶ (5, 1U) | 6 |
| 23 | (1, 7U) ⟶ (12, 1U) | 0 | (1, 7U) ⟶ (12, 1U) | 6 |

algorithm parameters, we set the population size to 50, the number of iterations to 50, the crossover probability to 0.9, and the mutation probability to 0.2.

In general, the physical lines in metro networks are decreasing, and the corresponding transfer passenger demand is also decreasing. The opposite also holds. Table 9 lists the calculation time of solving the model with or without the preprocessing method. The scale of the problem always affects the calculation speed, and the preprocessing method can effectively reduce the computation time.

To validate the effectiveness of the proposed approach, we use statistical analysis to assess the results of different-scale experiments obtained by the NSGA-II with and without the preprocessing method. Taking the computation times in Table 9 as two related samples, we then perform the Wilcoxon test on the SPSS software to verify the two methods. The verification results are shown in Table 10. Specifically, the value of normalized Wilcoxon statistics (denoted by $Z$) is −2.023, and the asymptotic 2-sided significance (denoted by $P$) is 0.043. Due to $P = 0.043 < 0.05$, the computation times with and without preprocessing method are completely different, and the NSGA-II with the preprocessing method is conducive to improving the solution efficiency.

To further demonstrate the advantage of our proposed approach, we compare the difference between transfer OD passenger demand and transfer-direction-based demand. First, we convert the OD demand to the transfer-direction-based demand. Similar to Wong et al. [18], we assume that the path choices of passengers are fixed, and passengers choose a path with as few interchanges as possible. This assumption allows us to obtain transfer-direction-based demand by assigning OD demand. Second, we replace minimizing passenger travel time with minimizing passenger transfer waiting time. The revised approach is capable of solving instances with the transfer-direction-based demand. Then, we use the two approaches (i.e., the proposed approach for OD demand and the revised approach for transfer-direction-based demand) to solve five groups of instances with different-sized networks. For each instance, we select solutions with the same deviation from the Pareto solutions obtained by the two approaches, respectively. The comparison results of the two cases are shown in Table 11.

It is well known that a transfer passenger may need to transfer multiple times to reach his/her destination in a complex metro network, which can be found in Table 11. Specifically, the passenger waiting time for the first, second, and third transfers can be obtained in our proposed approach, but the approach for the transfer-direction-based demand only can retrieve the waiting time for the first transfer.

In other words, the approach for transfer-direction-based demand cannot measure the multiple transfers of passengers. As shown in Figure 9, in the solution obtained by our proposed approach in the instance 1, OD (12, 79) passenger demand is to transfer from train 1 on line 1D to train 2 online 2D at station 14, from train 2 on line 2D to train 2 on line 3U at station 60, and from train 2 on line 3U to train 3 on line 6U. When using the approach for transfer-direction-based demand, however, the transfer processes of OD (12, 79) demand at stations 60 and 57 cannot be considered.

*5.2.2. Analysis of the Actual Headway $\overline{H}_l$ and the Maximum Headway $h_l^{max}$.* To analyze the influence of the actual headways ($\overline{H}_l$) and maximum headways ($h_l^{\max}$) on the optimized solutions, we consider five instances by setting different $\overline{H}_l$ and $h_l^{\max}$ on the basis of the above lager-sized example regarding the Xi'an metro network. Furthermore, we set the population size to 50, the number of iterations to 50, the crossover probability to 0.9, and the mutation probability to 0.2. The obtained results are shown in Table 12, where we calculate the average travel time, deviation time, and average headway for all the obtained Pareto solutions. Obviously, both the average travel time and average headway increase with both the maximum and actual headways.

In addition, we select three solutions (i.e., (5974021, 75), (5669964, 108), and (5492363, 147)) from Pareto solutions of

Table 8: Instance information.

| Instance ID | Network | Number of lines | Number of stations | Number of transfer stations |
|---|---|---|---|---|
| 1 | Origin network | 16 | 153 | 13 |
| 2 | Except 6U and 6D | 14 | 142 | 11 |
| 3 | Except 6U, 6D, 8U, and 8D | 12 | 134 | 10 |
| 4 | Except 6U, 6D, 8U, 8D, 7U, and 7D | 10 | 120 | 9 |
| 5 | Except 6U, 6D, 8U, 8D, 7U, 7D, 3U, and 3D | 8 | 98 | 5 |

Table 9: Comparison of the computational efficiency under different network sizes.

| Instance ID | Number of OD pairs | Computational times of the proposed algorithm without the preprocessing method (min) | Computational times of the proposed algorithm with the preprocessing method (min) |
|---|---|---|---|
| 1 | 19559 | 2.98 | 0.212 |
| 2 | 16481 | 2.20 | 0.153 |
| 3 | 14353 | 1.63 | 0.138 |
| 4 | 11021 | 0.992 | 0.078 |
| 5 | 6886 | 0.536 | 0.034 |

Table 10: Wilcoxon test results.

| | With the preprocessing method – without the preprocessing method |
|---|---|
| $Z$ | −2.023 |
| $P$ | 0.043 |

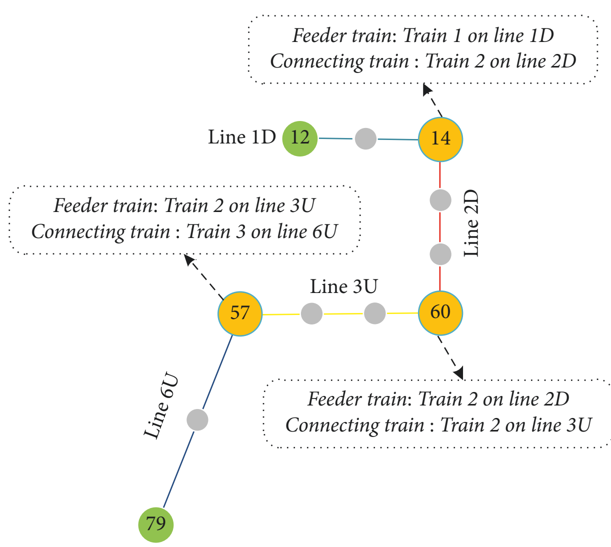


Figure 9: Transfer process for OD (12, 79) passenger demand.

Table 11: Comparison of the two approaches with different demand patterns.

| | Our proposed approach | | | | Approach for transfer-direction-based demand | | | |
|---|---|---|---|---|---|---|---|---|
| Instance ID | Passenger waiting time for the first transfer | Passenger waiting time for the second transfer | Passenger waiting time for the third transfer | Computational times (min) | Passenger waiting time for the first transfer | Passenger waiting time for the second transfer | Passenger waiting time for the third transfer | Computational times (min) |
| 1 | 367538 | 122574 | 74768 | 0.212 | 371544 | — | — | 0.052 |
| 2 | 313964 | 83143 | 44813 | 0.153 | 338915 | — | — | 0.033 |
| 3 | 275089 | 34666 | 15227 | 0.138 | 289690 | — | — | 0.030 |
| 4 | 211925 | 27188 | 0 | 0.078 | 211315 | — | — | 0.026 |
| 5 | 100121 | 23618 | 0 | 0.034 | 102418 | — | — | 0.015 |

Table 12: Analysis of different maximum headways and actual headways.

| Instance: Maximum headway ($h_l^{max}$) | Instance: Actual headway ($\overline{H}_l$) | Average travel time (min) | Average deviation time (min) | Average headway (min) | Computational times (min) |
|---|---|---|---|---|---|
| 12 | 7 | 5457096 | 90 | 6 | 0.225 |
| 14 | 9 | 5525758 | 81 | 10 | 0.217 |
| 16 | 11 | 5566354 | 113 | 10 | 0.223 |
| 18 | 13 | 5694798 | 108 | 14 | 0.284 |
| 20 | 15 | 5747981 | 110 | 14 | 0.325 |

Table 13: Analysis of optimized headway and connection of the first train.

| Solution | Optimized headway | | | | Connection of the first train | | |
|---|---|---|---|---|---|---|---|
| | Number of optimized headways (>13 min) | Number of optimized headways (=13 min) | Number of optimized headways (<13 min) | Average headway (min) | Number of connection times for first trains (>18 min) | Number of connection times for first trains (≤18 min) | Average connection time (min) |
| (5974021, 75) | 6 | 5 | 5 | 14 | 19 | 69 | 38 |
| (5669964, 108) | 2 | 3 | 11 | 10 | 18 | 70 | 36 |
| (5492363, 147) | 3 | 1 | 12 | 8 | 19 | 69 | 37 |

the instance with $h_l^{max} = 18$ and $\overline{H}_l = 13$. We count the number of optimized headways and the number of connections of the first trains, which are shown in Table 13. It can be found that the average headway of the solution (5974021, 75) is greater than the actual headway, but the average headways of the other two solutions are less than the actual headway. That is, the optimized Pareto solutions provide different train schedules such that operators can choose the preferred solutions according to their preferences. Also, although we have optimized the connection times of the first trains, the connection times of many first trains are still longer than the headway. Actually, this result is also the most distinctive feature of the first train timetabling problem.

*5.2.3. Analysis of Different Departure Time Ranges of the First Trains at the Starting Stations.* To analyze the influence of different departure time ranges of the first trains at the starting stations on the optimized solution, we further solve the 6 groups of instances with the different lengths of departure time ranges. Specifically, the Xi'an metro network is also used as input, and the length of the departure time ranges of the first trains are set to 10, 12, 14, 16, 18, and 20 minutes, respectively. The other parameters in the algorithm are the same as the setting of the above section, and Table 14 lists the optimized results. One of the most notable features is that as the length of the departure time range is increased, the average deviation time increases because the longer departure time range inevitably causes a larger solution space.

*5.2.4. Analysis of Different Decision Variables.* To analyze the influence of different decision variables, we propose the following three approaches: (i) optimizing departure times of first trains with fixed headways, denoted as FH-ODT for short, (ii) optimizing headways with fixed departure times of first trains, denoted as FDT-OH, and (iii) optimizing headways and departure times of first trains (i.e., our proposed approach). Based on the aforementioned large-sized instance in Section 5.1, we construct five instances by setting different headways and departure times of first trains. For simplicity, we assume that the headways of all lines are the same and the departure times of first trains are uniformly increased or decreased on the basis of Table 5. Thus, each instance is identified by the headways and the increment of the departure times of the first trains. For example, instance (5, −5) is that the headways of all lines are 5 minutes, and the departure times of the first trains are the departure times in Table 5 minus 5 minutes. Then, we employ the three approaches to solve each instance. The computational results are shown in Table 15. Compared with approaches FH-ODT and FDT-OH, our proposed approach can significantly improve passenger travel time while maintaining small changes in deviation times.

Table 14: Analysis of the different departure time ranges of the first trains at starting stations.

| The length of the departure time range for each instance (min) | Number of Pareto solutions | Average travel time (min) | Average deviation time (min) |
|---|---|---|---|
| 10 | 24 | 1369 | 647 |
| 12 | 18 | 1296 | 662 |
| 14 | 22 | 1387 | 814 |
| 16 | 22 | 1678 | 888 |
| 18 | 24 | 1758 | 897 |
| 20 | 27 | 2121 | 1133 |

Table 15: Influence of decision variables on results.

| Instance ID | Approach FH-ODT | | Approach FDT-OH | | Our proposed approach | | |
|---|---|---|---|---|---|---|---|
| | Average travel time | Average deviation time of departure times of first trains | Average travel time | Average deviation time of headway | Average travel time | Average deviation time of departure times of first trains | Average deviation time of headway |
| (5, −5) | 5338422 | 44 | 5477621 | 28 | 5328060 | 62 | 28 |
| (6, −3) | 5463306 | 47 | 5476931 | 30 | 5441326 | 45 | 34 |
| (7, 0) | 5528983 | 42 | 5477391 | 31 | 5453357 | 46 | 32 |
| (8, 3) | 5590469 | 44 | 5507152 | 27 | 5485273 | 46 | 29 |
| (9, 5) | 5620982 | 41 | 5506495 | 31 | 5495959 | 56 | 42 |

## 6. Conclusions

This paper studies first-train timetable synchronization in metro networks under passenger OD demand conditions. These passengers only take the first train on each line and expect to complete their journeys in the minimum travel times. To minimize the passenger travel times, the flexible selection of travel routes in the network is permitted for passengers. Furthermore, the train-connection-based route is introduced to calculate the passenger travel times. In addition, to guarantee the service quality of nontransfer passengers, we minimize the deviation between the actual schedule and the optimized schedule on each line. Thus, a biobjective nonlinear integer programming model, which is subjected to train operation and connection constraints, is established. Then, NSGA-II is used to solve the proposed model, and an effective preprocessing method is proposed to speed up the computational times. Taking the Xi'an metro network as the background, several experiments are implemented to demonstrate the effectiveness and efficiency of the model and algorithm.

As for this complex first-train timetable synchronization problem, our future studies will focus on the following aspects. First, a strict train capacity will be considered to model passenger travel behaviors. Second, it is necessary to establish some tractable linear optimization models and develop high-performance solution algorithms, as well as mathematical models under uncertain train operation and passenger travel conditions. Finally, it may be a better research direction to expand the research object from partial trains to all trains for the construction of full-day network-wide train timetables.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 72161023 and 71771109) and the Department of Education of Gansu Province: Youth Doctoral Fund Project (Grant No. 2021QB-054).

## References

[1] L. Kang, X. Zhu, H. Sun, J. Puchinger, M. Ruthmair, and B. Hu, "Modeling the first train timetabling problem with minimal missed trains and synchronization time differences in subway networks," *Transportation Research Part B: Methodological*, vol. 93, pp. 17–36, 2016.

[2] X. Guo, J. Wu, H. Sun, H. Sun, and Z. Gao, "Timetable coordination of first trains in urban railway network: a case study of Beijing," *Applied Mathematical Modelling*, vol. 40, pp. 8048–8066, 2016.

[3] C. Fleurent, R. Lessard, and L. Séguin, "Transit timetable synchronization: evaluation and optimization," *International Conference om Computer-Aided Scheduling of Public Transport*, vol. 9, pp. 9–11, 2004.

[4] F. Cevallos and F. Zhao, "Minimizing Transfer Times in Public Transit Network with Genetic Algorithm," Transportation Research Record," *Journal of the Transportation Research Board*, vol. 1971, no. 1, pp. 74–79, 2006.

[5] Y. Chen, B. Mao, Y. Bai, T. K. Ho, and Z. Li, "Optimal coordination of last trains for maximum transfer accessibility with heterogeneous walking time," *Journal of Advanced Transportation*, vol. 2019, Article ID 9692024, 13 pages, 2019.

[6] Y. Chen, B. Mao, Y. Bai, T. K. Ho, and Z. Li, "Timetable synchronization of last trains for urban rail networks with maximum accessibility," *Transportation Research Part C: Emerging Technologies*, vol. 99, pp. 110–129, 2019.

[7] X. Li, T. Yamamoto, T. Yan, L. Lu, and X. Ye, "First train timetabling for urban rail transit networks with maximum passenger transfer satisfaction," *Sustainability*, vol. 12, no. 10, p. 4166, 2020.

[8] C. M. Chung Min Kwan and C. S. Chang, "Timetable synchronization of mass rapid transit system using multiobjective evolutionary approach," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 38, no. 5, pp. 636–648, 2008.

[9] X. Tian and H. Niu, "A bi-objective model with sequential search algorithm for optimizing network-wide train timetables," *Computers & Industrial Engineering*, vol. 127, pp. 1259–1272, 2019.

[10] X. Guo, J. Wu, H. Sun, X. Yang, J. G. Jin, and D. Z. W. Wang, "Scheduling synchronization in urban rail transit networks: trade-offs between transfer passenger and last train operation," *Transportation Research Part A: Policy and Practice*, vol. 138, pp. 463–490, 2020.

[11] L. Kang, H. Sun, J. Wu, and Z. Gao, "Last train station-skipping, transfer-accessible and energy-efficient scheduling in subway networks," *Energy*, vol. 206, Article ID 118127, 2020.

[12] W. Nie, H. Li, N. Xiao, H. Yang, Z. Jiang, and N. Buhigiro, "Modeling and solving the last-shift period train scheduling problem in subway networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 569, Article ID 125775, 2021.

[13] H. Niu and X. Zhou, "Optimizing urban rail timetable under time-dependent demand and oversaturated conditions," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 212–230, 2013.

[14] E. Barrena, D. Canca, L. C. Coelho, and G. Laporte, "Exact formulations and algorithm for the train timetabling problem with dynamic demand," *Computers & Operations Research*, vol. 44, no. 3, pp. 66–74, 2014.

[15] H. Niu, X. Zhou, and R. Gao, "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*, vol. 76, pp. 117–135, 2015.

[16] X. Tian and H. Niu, "Optimization of demand-oriented train timetables under overtaking operations: a surrogate-dual-variable column generation for eliminating indivisibility," *Transportation Research Part B: Methodological*, vol. 142, pp. 143–173, 2020.

[17] P. Shang, R. Li, and L. Yang, "Demand-driven timetable and stop pattern cooperative optimization on an urban rail transit line," *Transportation Planning and Technology*, vol. 43, no. 1, pp. 78–100, 2020.

[18] R. C. W. Wong, T. W. Y. Yuen, K. W. Fung, and J. M. Y. Leung, "Optimizing timetable synchronization for rail mass transit," *Transportation Science*, vol. 42, no. 1, pp. 57–69, 2008.

[19] Y. Wang, B. Ning, T. Tang, T. J. van den Boom, and B. De Schutter, "Efficient real-time train scheduling for urban rail transit systems using iterative convex programming," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 6, pp. 3337–3352, 2015.

[20] Y. Shafahi and A. Khani, "A practical model for transfer optimization in a transit network: model formulations and solutions," *Transportation Research Part A: Policy and Practice*, vol. 44, no. 6, pp. 377–389, 2010.

[21] H. Niu, X. Tian, and X. Zhou, "Demand-driven train schedule synchronization for high-speed rail lines," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2642–2652, 2015b.

[22] W. Li, R. Xu, Q. Luo, and S. Jones, "Coordination of last train transfers using automated fare collection (AFC) system data," *Journal of Advanced Transportation*, vol. 50, no. 8, pp. 2209–2225, 2016.

[23] Y. Zhou, Y. Wang, H. Yang, and X. Yan, "Last train scheduling for maximizing passenger destination reachability in urban rail transit networks," *Transportation Research Part B: Methodological*, vol. 129, pp. 79–95, 2019.

[24] L. Yang, Z. Di, M. M. Dessouky, Z. Gao, and J. Shi, "Collaborative optimization of last-train timetables with accessibility: a space-time network design based approach," *Transportation Research Part C: Emerging Technologies*, vol. 114, pp. 572–597, 2020.

[25] J. Yin, A. D'Ariano, Y. Wang, L. Yang, and T. Tang, "Timetable coordination in a rail transit network with time-dependent passenger demand," *European Journal of Operational Research*, vol. 295, no. 1, pp. 183–202, 2021.

[26] K. Yang, Y. Lu, L. Yang, and Z. Gao, "Distributionally robust last-train coordination planning problem with dwell time adjustment strategy," *Applied Mathematical Modelling*, vol. 91, no. 3, pp. 1154–1174, 2021.

[27] Y. Zhou, Y. Bai, J. Li, B. Mao, and T. Li, "Integrated optimization on train control and timetable to minimize net energy consumption of metro lines," *Journal of Advanced Transportation*, vol. 2018, Article ID 7905820, 19 pages, 2018.

[28] Y. Wu, H. Yang, J. Tang, and Y. Yu, "Multi-objective re-synchronizing of bus timetable: model, complexity and solution," *Transportation Research Part C: Emerging Technologies*, vol. 67, pp. 149–168, 2016.

[29] J. F. Guan, H. Yang, and S. C. Wirasinghe, "Simultaneous optimization of transit line configuration and passenger line assignment," *Transportation Research Part B: Methodological*, vol. 40, no. 10, pp. 885–902, 2006.

[30] S. Li, R. Xu, and K. Han, "Demand-oriented train services optimization for a congested urban rail line: integrating short turning and heterogeneous headways," *Transportmetrica: Transportation Science*, vol. 15, no. 2, pp. 1459–1486, 2019.