

Research Article

Optimal Operation Scheme with Short-Turn, Express, and Local Services in an Urban Rail Transit Line

Tao Feng ^{1,2}, Siyu Tao ^{1,2} and Zhengyang Li ³

¹School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China

²National United Engineering Laboratory of Integrated and Intelligent Transportation, Southwest Jiaotong University, Chengdu 610031, China

³Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong, China

Correspondence should be addressed to Siyu Tao; taosiyu@swjtu.edu.cn

Received 6 February 2020; Revised 1 October 2020; Accepted 19 October 2020; Published 19 November 2020

Academic Editor: Yu Jiang

Copyright © 2020 Tao Feng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Flexible railway operation modes combining different operation strategies, such as short-turn, express, and local services, can significantly reduce operator and user costs and increase the efficiency and attractiveness of rail transit services. It is therefore necessary to develop optimization models to find optimal combinations of operation strategies for urban rail transit lines. In this paper, a model is proposed for solving the urban rail transit operation scheme problem. The model considers short-turn, express, and local services with the aim of minimizing the operator's and users' costs. The problem is first decomposed into two subproblems: the service route design problem and the passenger assignment problem. Then, a mixed-integer nonlinear program (MINLP) model is formulated, and linearization techniques are utilized to transform the MINLP model into a mixed-integer linear programming (MILP) model that can be easily solved by commercial optimization solvers. To accelerate the solution process, a heuristic search algorithm is proposed to obtain (nearly) optimal solutions based on the characteristics of the model. The two subproblems are solved iteratively to improve the quality of solutions. A real-life case study in Chengdu, China, is performed to demonstrate the effectiveness and efficiency of the proposed model and algorithm.

1. Introduction

Urban rail transit systems play an important role in meeting passenger demands in our daily lives because of their large passenger transport capacity, low pollutant emissions, and convenience. The increase in the number of long-distance passengers and imbalanced passenger demand due to urban expansion have brought increasing pressure to public transport. The elastic passenger demand cannot be satisfied by adopting only a single operation mode. As a means of improving the quality of public transport, urban rail transit has drawn much attention from operators and governments. Operators aim to improve the service efficiency and quality by adopting flexible service strategies such as short-turn and elastic services and elastic stopping patterns to satisfy the daily travel demand. A local service serves all the stations along the full length of the line, an express service runs from

a starting station to an ending station and visits only some intermediate stations, and a short-turn service is provided only on sections with larger passenger flow volumes. Flexible operation schemes integrating different service strategies can markedly improve the service level of urban rail transit. Full-length local services are commonly adopted in urban rail transit lines in China, but may lead to wasted train capacity due to the imbalanced distribution of passenger demand. In many big cities, such as Tokyo, New York, and Paris, routes with local, express, and short-turn services are used to cater to elastic passenger demand [1]. When short-turn, express, and local services are provided, different stations and sections are served at different frequencies so that the transport capacity matches the diversity of passenger demand on the line appropriately [2]. Hence, academic research on the design and financing of (nearly) optimal operation schemes is necessary for urban rail transit operators.

This study investigates the urban rail transit service design problem (URTSDP). The service routes, including which stations the trains turn back at, the stopping patterns, and the train frequencies need to be optimized to reduce the operator's operation cost and the users' time cost. Meanwhile, the capacity of each turn-back station is fixed and cannot be exceeded. After the starting and ending stations have been determined, a train departs from the starting station, serves some or all the intermediate stations, and runs in sections in a predetermined order. Express services can operate on any service route. Hence, the design of urban rail transit services is more complex and challenging compared to the design of bus service networks. The URTSDP can be divided into two subproblems: the service route design problem and the passenger assignment problem. The former determines the service routes and their frequencies, while the latter determines the optimal strategy for assigning users from their origins to their destinations [3].

Under these conditions, we propose an optimized model to solve the urban rail transit service design problem. The solution can be split into two aspects. Firstly, this model considers short-turn, express, and local services and the capacity limitations of the turn-back stations and sections in urban rail transit lines. Secondly, the passenger assignment problem is solved while meeting the capacity limitation constraints. A small-scale problem can be solved efficiently using the model and commercial solvers; however, the problem scale will increase dramatically because of the stopping patterns of the express and short-turn services if the problem is solved directly. To reduce the scale of the problem, a novel strategy called the *line pool* is introduced in the model formulation. Some important stations are designated in advance, and a service route must visit these stations if the route covers them. All the feasible service routes constitute the *line pool*. A mixed-integer nonlinear programming (MINLP) model is formulated. The model can be transformed into a mixed-integer linear programming (MILP) model by using linearization techniques such as the reformulation linearization technique (RLT). In addition, we design a search algorithm to solve the large-scale problem iteratively. Two groups of numerical studies are performed to verify the effectiveness of the model and the search algorithm.

The main contributions of this paper are as follows:

- (1) Formulation of a MILP optimization model to solve the URTSDP, which includes the service route design problem and the passenger assignment problem and considers the passenger transfers and the turn-back station capacity restriction
- (2) Proposal of a *line pool* strategy and novel search algorithm to solve the large-scale problem
- (3) Verification of the performance of the new search algorithm by comparing its performance with that of the MILP model
- (4) Demonstration of the effectiveness and efficiency of the model and the new search algorithm using real-life case studies

The remainder of this paper is organized as follows. In Section 2, an overview of the existing literature is presented.

Section 3 provides the problem description. Section 4 proposes a model for the URTSDP. Section 5 describes a search algorithm to solve the problem. In Section 6, a practical numerical study is used to verify the effectiveness of the model. Section 8 concludes the paper.

2. Literature Review

Most of the studies on the design of urban rail transit services have focused on the use of express and local services. Service optimization with the synchronous operation of short-turn, express, and local services is much less well studied. It is crucial to determine which stations to stop at or skip in express/local services. Chun et al. [4] proposed a synchronized express and local system implemented to optimize the urban transit system by reducing passengers' waiting times and transfers. Gao et al. [1] analyzed the relationship between the energy consumption and passenger waiting time in the operation of express and local services and formulated a biobjective model. Jamili and Pourseyed Aghaee [5] proposed a model for designing the stopping patterns of service routes under uncertain conditions to help reduce the train fleet size and energy consumption. Based on the assumption that the capacity of the train is unlimited, Abdelhafiez et al. [6] constructed a nonlinear integer programming model, designed a heuristic method to solve the model, and compared the heuristic method to the enumeration algorithm. Luo et al. [7] presented a mathematical model that minimizes the passenger travel time and optimizes the train stopping schedule for express and local services. The model analyzes passenger choice behavior using a logit model and evaluates the transport capacity and energy consumption efficiency. However, this work does not consider the operator's cost. Yang et al. [8] proposed a nonlinear programming model which optimizes the frequency, stopping pattern, and express/local trains on the cross-lines and verified the resulting operation scheme obtained by the genetic algorithm with the actual cases. The model can reduce the users' time cost and the operational investment. It is not the only method to address the imbalanced passenger demand by optimizing the express/local train services in urban rail transit lines. Some researchers have considered other factors wherein the positions of the turn-back stations, that is, the choices of the turn-back stations for the origin and destination, play a significant role in determining the service routes. Considering the capacities of the trains and turn-back stations and other operational constraints, Sun et al. [9] proposed a model to solve the complicated problem of optimizing the operation scheme to minimize the total cost to operators and users based on the actual problems faced in urban rail transit lines. In contrast, Ding et al. [10] constructed a mathematical model to minimize the average waiting time of passengers without considering the turn-back equipment. The resulting solution illustrates that the operation scheme based on this model can effectively alleviate congestion and avoid wasted transport capacity. Zhang et al. [11] presented an MINLP model for train scheduling in which short-turn services and the influence of other impact factors on the scheme formulation

are considered. Li et al. [12] proposed a collaborative optimization model in which all the service routes can use different train compositions, and the frequency and train composition of each service route and the positions of turn-back stations are optimized simultaneously. Cao et al. [13] formulated a model based on marshaling, the skip-stop pattern, and robustness to optimize the timetabling design to minimize the total cost to the operator and users and applied the model to real-life urban rail transit in Beijing.

The aforementioned studies only consider a single operation strategy. Multiple operation strategies can be integrated into a single model to obtain better schemes to relieve the transport pressure on urban rail transit lines. There are a number of related studies on bus network design which are worth learning from. The studies are plentiful, and they mainly focus on factors such as the stop patterns, express and short-turn services, multiple fleets, and bus capacities. Given the service design similarities between urban rail transit and bus corridors, theoretical methods for bus service design can be used as references for optimizing the stopping patterns of express services. Afanasiev and Liberman [14] defined an express service as a service with no intermediate stops between the origin and terminal stations to satisfy passenger demands. The passengers can achieve travel time savings of up to 30%. Vijayaraghavan and Anantharamaiah [15] presented a theoretical model to insert express as well as partial services for better fleet assignment. The model may allow fleet reduction and provide benefits for the operator and users in terms of saved travel time. Conlon et al. [16] described the introduction of express bus services to Chicago, which were preferred by the customers and achieved a customer satisfaction that exceeded the standard target mean score. Wang and Lo [17] proposed a ferry service design model with normal and express service. The problem was formulated as an MILP model, and a solution algorithm was designed to reduce the scale of the problem. Larrain et al. [18] presented an express selection model with capacity restriction and defined the parameters for determining which express service is attractive on a bus corridor. Meanwhile, Ulusoy et al. [19] proposed a model for further improving the optimal strategy by incorporating various factors and including all-stop, short-turn, and express services to meet the heterogeneous demand. Larrain et al. [20] proposed a methodology in which the express services skip all the intermediate stops except for the initial and final stops. The authors reported that the introduction of zonal services can save more than 6% of the social costs. Wang et al. [21] formulated a MINLP model to design bus services from the perspectives of operators and users.

From another perspective, short-turn operation strategies are commonly used in urban rail transit and bus transit systems. The short-turn points for buses can be at any stop, but urban rail transit faces more restrictions due to the operational rules. Hence, the consideration of the short-turn points in bus service design is less complicated. Ceder [22] presented a model to determine the feasible short-turn stops. Furth [23] found that it is necessary to coordinate the schedule with the short-turn pattern and proposed a model to obtain the optimal schedule offset. Delle Site and Filippi

[24] considered the short-turn location and vehicle size as variables and presented a framework to design the bus service scheme on a given corridor. In comparison, urban rail transit is more complicated because of the limited capacity of the turn-back stations. The selection of the turn-back stations is the main variable in designing a short-turn service. The direction and capacity of the turn-back station must be considered in urban rail transit. In order to analyze the advantages and disadvantages of the various models in more detail, a comparison will be performed between the proposed model and the previous studies (as shown in Table 1). Li et al. [12] proposed a model to optimize the operation scheme that considers short-turn services and the train composition. A flexible operation scheme can be achieved with a variable train composition to cope with varying passenger demand. However, this will increase the purchase cost of trains because the train composition is fixed and the expenditure increases with the size of the train composition. In contrast, express services can relieve the pressure of imbalanced passenger demand without changing the train composition.

The above literature review shows that (i) previous studies have only focused on express and local services or short-turn services, and few studies have integrated express and local services with short-turn services in urban rail transit operation schemes. There are more related studies on the design of bus service networks. These methods cannot be used directly to solve urban rail transit service design problems because of the unique restrictions and considerations that exist in operating urban rail transit services such as the capacities of the turn-back stations and the sections; (ii) there is usually an assumption that express services are operated on full-length services, which may also lead to wasted transport capacity. In this work, express services are operated in conjunction with short-turn services.

3. Problem Description

In this section, we describe the URTSDP, which includes two subproblems: the service route design problem (SRDP) and the passenger assignment problem (PAP). Without loss of generality, we assume that the layout of the urban rail transit line and the turn-back stations are given and the operational rules are known. The URTSDP decides the number of service routes, frequency, and train stopping pattern with the objective of minimizing operator and user costs.

3.1. Service Route Design Problem. An operation scheme that includes the number of service routes and the stopping pattern of the express service is determined given the urban rail line, the positions of the turn-back stations, and the condition that there are no route limitations on the express services, which can operate on both full-length and short-turn routes. In addition, there are two constraints to be satisfied:

- (1) Capacity limitation of the turn-back station: the service routes that end at these stations must not exceed the turn-back capacity of the station. Each

TABLE 1: Comparisons of this paper with previous studies.

Factors	Ulusoy et al. [19]	Larrain et al. [20]	Li et al. [12]	This paper
<i>Assumption</i>				
Demand	Fixed	Fixed	Fixed	Fixed
Transfer	Allowed	Allowed	Allowed	Allowed
Strategy	Short-turn services and one express service	Express service	Short-turn services	Short-turn services and express/local services
<i>Model</i>				
Single	✓	—	✓	✓
Bilevel	—	✓	—	—
<i>Objective</i>				
Operator cost	✓	✓	✓	✓
User cost	✓	✓	✓	✓
<i>Constraints</i>				
Transport capacity	✓	✓	✓	✓
Turn-back capacity	—	—	✓	✓
Minimum frequency	✓	—	✓	✓
<i>Solving method</i>				
Solution method	Exhaustive search	Heuristic + solver	Heuristic + solver	Heuristic + solver

service route incurs extra time when it turns back at the ending station. The time is called the turn-back time in urban rail transit lines.

- (2) Maximum frequency limitation of each section: each train needs to operate on its corresponding service route in a predetermined sequence. The total number of all trains in each section cannot exceed the maximum frequency because the rail transit has only one line in the inbound or outbound direction.

We take the urban transit line with 6 physical stations and 3 turn-back stations shown in Figure 1 as an example. The stations are denoted as n_1 to n_8 in the outbound direction of the corridor, while the sections are denoted as e_1 to e_7 . Stations n_1 , n_4 , and n_8 are the turn-back stations, and n_4 is designed for bidirectional turn-back.

Based on the given directions and positions of the turn-back stations and their associated capacities on the line, a turn-back station with an available direction can be used as the starting or ending station of the service route. A set of starting stations and a set of ending stations can thus be obtained in the initial stage. A service route consists of routes and nodes, where each node is associated with a physical station. A *line pool* is then obtained by varying the terminal stations among the starting or ending station sets. Figure 2 shows an operation scheme from the *line pool*. It consists of 3 service routes, where *route 1* is short turn with local service, *route 2* is short turn with express service, and *route 3* is full length with express service.

The SRDP is primarily concerned with the frequency, short-turn, and stop pattern of the express service. The operation scheme including the above factors can be obtained by solving the SRDP. The example in Figure 2 consists of three service routes. The terminal of each route is set in accordance with the available starting or ending station set and the condition that each station is served by at least one

service route. However, the operation scheme may be nonoptimal even after solving the SRDP because the frequency of each route was not determined in conjunction with the PAP to consider the total cost. To obtain (nearly) optimal operation schemes, the costs of the operator and users need to be considered. That is, the SRDP and PAP need to be solved simultaneously. This is a matter of great concern in academic studies.

3.2. Passenger Assignment Problem. Before determining the optimal passenger assignment strategy, the service network needs to be constructed in advance. A service network containing the arcs, nodes, and physical stations can be constructed depending on the operation scheme, as shown in Figure 2. A train starts from the origin node, goes through intermediate nodes, and ends at the destination node, wherein arcs are used to connect nodes with the associated physical stations, and passengers board (alight) trains operating on the routes served by these arcs. The arcs contain the specific routes of passenger travel and can be divided into three types of arcs, namely, the boarding, alighting, and riding arcs. A service network, as shown in Figure 3, includes a complete passenger traveling path that begins from his (her) origin station, goes through the boarding, riding, and alighting arcs, and eventually ends at his (her) terminal station.

For instance, a passenger travels from station n_1 to station n_5 along the travel path $n_1 \rightarrow 17 \rightarrow 18 \rightarrow 19 \rightarrow 20 \rightarrow n_5$. He (she) will pass through a boarding arc ($n_1 \rightarrow 17$), riding arcs ($17 \rightarrow 18 \rightarrow 19 \rightarrow 20$), and an alighting arc ($20 \rightarrow n_5$). Therefore, a connection is made by the boarding arc between physical station n_1 and node 17, while the alighting arc performs the same function along the opposite direction. The riding arc makes the connection between two route stations. If one physical station is not

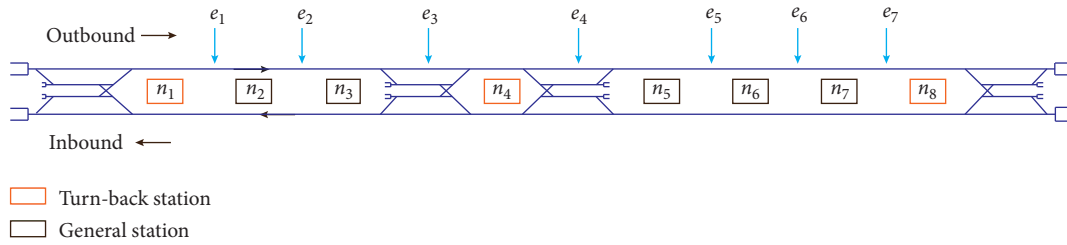


FIGURE 1: An urban rail transit corridor.

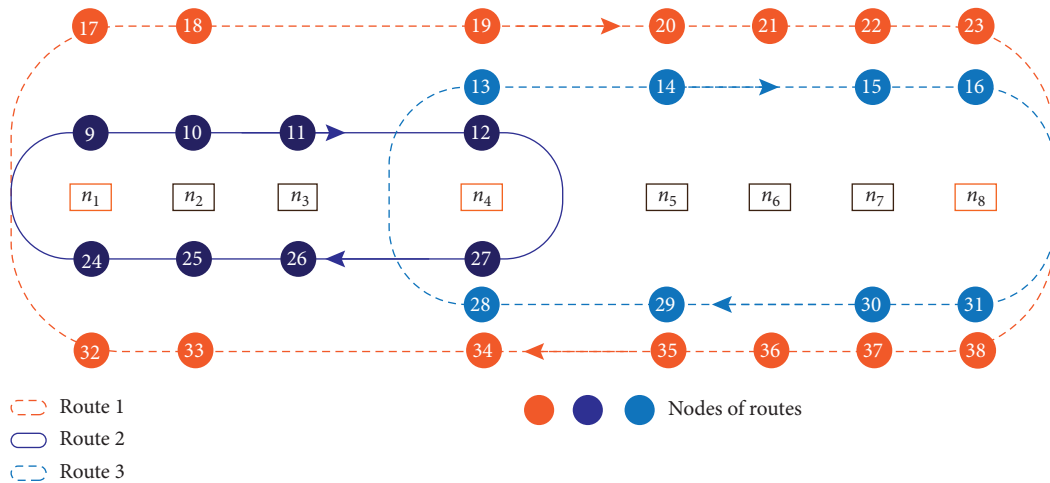


FIGURE 2: Service description of urban rail transit.

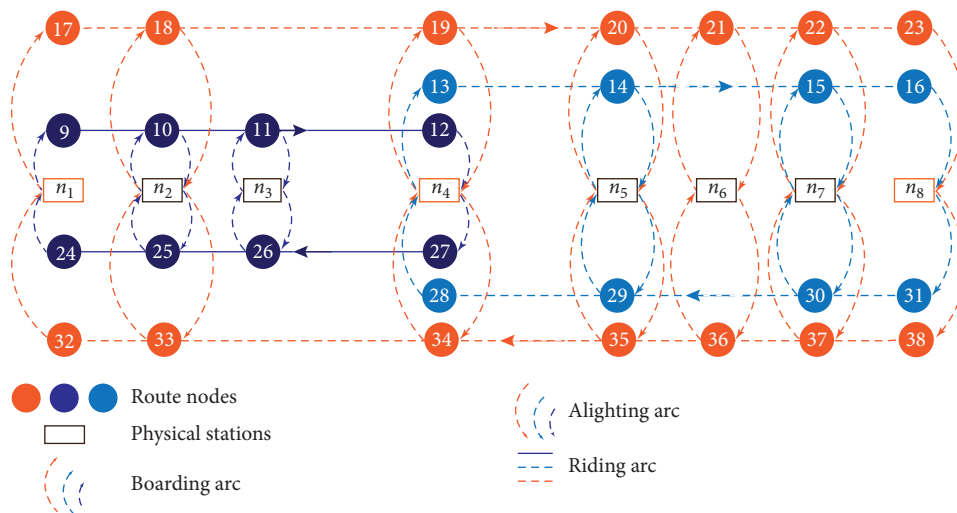


FIGURE 3: Description of directed service network.

served by a service route, then there is no connection between the physical station and the node. For n_3 shown in Figure 3, there is no boarding arc between n_3 and route 1.

As shown in Figure 3, there may be many boarding arcs between the physical stations and services. Therefore, it is important to analyze the service choice behavior of the passengers. To address this *common line* problem, Chriqui and Robillard [25] presented a solution to determine which

service passengers would choose by defining the attractiveness of each service based on the passengers' waiting time and in-vehicle time. This approach does not require the prior preparation of fixed service network. Spiess and Florian [3] proposed a strategy for a passenger to reach the terminal station based on the least cost service in a model with a fixed transit line. In this work, we construct a transit assignment model with a known service network for urban rail transit

services based on the optimal strategy and consider the effectiveness of transfers on the passenger distribution further. A detailed discussion is given in the next section.

4. Model Formulation for URTSDP

4.1. Notation Definitions. We formulate the URTSDP problem as an MINLP model and transform the model into an MILP model in this section. The notations to be used in the model in what follows are explained in Table 2.

4.2. Modelling Assumptions. The model should ideally provide station-to-station services for operators and users as quickly as possible. The services proposed for the given stations are determined by the operator's cost and users' cost. These two factors must be considered simultaneously for a more effective operation scheme. The main assumptions of this study are as follows:

- (1) The exogenous trip demand is considered to be fixed.
- (2) The train composition is an important factor that affects the transport capacity and service level of the urban rail transit line, which are predetermined and fixed. In addition, the fleet size is limited.
- (3) The passengers' arrival at each station is considered to be uniformly distributed, and the train arrivals follow an exponential distribution.
- (4) We consider the passenger travel behavior under the condition of no congestions.

4.3. Total Costs of URTSDP. The total cost of URTSDP in this study consists of the operator's cost and the users' cost. The former includes two contributions. First, the train ownership cost includes the purchasing cost and related expenses during the operation time of the vehicle. The operation period of a train is typically 30 years [12]. The train ownership cost can be obtained based on the operation time obtained from the model. This promotes sensitivity to the operation time by distributing the cost between the hours on each working day. A second contribution to the operating cost is costs such as the electricity, staff, and track maintenance cost. The user cost can be divided into three parts: the waiting cost, the riding cost, and the transfer cost. The first two are dependent on the time spent and the transfer cost depends on the number of transfers.

Based on the characteristics of the URTSDP, we can divide the model into two parts: (i) a submodel to solve the SRDP and (ii) an optimal strategy to solve the PAP. The generation of the service network requires both models to be solved simultaneously.

4.3.1. SRDP Model. The SRDP determines the service routes and operation frequency on the urban rail transit line. The objective function is formulated as follows:

$$(\text{SRDP}) \min z_1 = \sum_{f \in F} \sum_{l \in L} x_{fl} \cdot t_l \cdot c_0 + \sum_{f \in F} \sum_{l \in L} x_{fl} \cdot d_l \cdot c_1. \quad (1)$$

In the objective function (1), the first term is the ownership cost of the trains allocated to the express services and local services during the operation time, and the second term is the service distance-dependent operating cost. The major constraints in the model are as follows:

- (i) Service level: the number of service routes should not exceed the limits on the operator's cost and users' travel cost. In addition, the frequency should not be lower than a specified positive number to meet the passenger demand while considering the surplus capacity. The formulation of the service level is as follows:

$$\sum_{l \in L} \mu_l \leq r_{\max}, \quad (2)$$

$$\mu_l \cdot \lambda_{\min} \leq \sum_{f \in F} x_{fl} \leq \mu_l \cdot \lambda_{\max}, \quad \forall l \in L. \quad (3)$$

- (ii) Station coverage and passenger demand satisfaction: each station should be served by at least one service route, and the passenger demand at each station should be satisfied because passenger demand is fixed. Constraints (4) describe the variables for the stations:

$$\sum_{f \in F} \sum_{l \in L} x_{fl} \cdot \pi_{li} \cdot u \geq p_e, \quad \forall i \in N. \quad (4)$$

- (iii) Section coverage passenger flow volume satisfaction: each section should be covered by at least one service route, and the transport capacity should not be less than the volume of passenger flow in each section. Constraints (5) represents that the total number of trains should not be less than ϕ_{\min} and that the maximum frequency f_{\max} in each section is not exceeded because operational rules restrict the frequency in each section to maintain a safety distance between two adjacent trains along the urban rail transit line:

$$\phi_{\min} \leq \sum_{f \in F} \sum_{l \in L} x_{fl} \cdot k_{le} \leq \phi_{\max}, \quad \forall e \in E, \quad (5)$$

$$(1 - \gamma) \sum_{f \in F} \sum_{l \in L} x_{fl} \cdot k_{le} \cdot u \geq q_e, \quad \forall e \in E. \quad (6)$$

- (iv) Turning back capacity constraint: the operation scheme can be made more flexible by introducing short-turn services. Each short-turn service route must return at a specified station owing to the properties of the urban transit line. However, each train turning back occupies the track for a specified time. Therefore, the turning back capacity of a station is limited in the inbound and outbound directions, and the number of trains turning back cannot exceed the maximum limitation:

TABLE 2: Definition of sets, indices, parameters, and variables.

Notations	Description
<i>Sets</i>	
V	Set of service route nodes with index n
N	Set of physical stations in urban rail transit line with index $i, i, \kappa \in N$. $N = N_1 \cup N_2$, where N_1 is the set of low-priority stations with medium or small volume of passenger flow and N_2 is the set of high-priority stations with large volume of passenger flow
N_l	Set of physical stations that service route l visited
M	Set of the turn-back stations
L	Set of service routes in urban rail transit service network with index l . $L = L_1 \cup L_2$, where L_1 is the set of express services and L_2 is the set of local services
E	Set of sections in urban rail transit line with index e
A	Set of arcs in a service network with index a . $A = A_1 \cup A_2 \cup A_3$, where A_1 is the set of boarding arcs, A_2 is the set of riding arcs, and A_3 is the set of alighting arcs
A_i^+	Set of departing arcs in a service network in station i
A_i^-	Set of arriving arcs in a service network in station i
O	Set of inbound and outbound direction with index o
F	Set of the operational trains with index f , $F = \{1, 2, 3, \dots, f, \dots, F \}$, f denotes that f th train to be operating in service route
W	Set of O-D pairs with index $w(o, d) \in W$, where o and d denote the origin and destination
<i>Parameters</i>	
u	Capacity of each train
r_{\max}	The maximum number of service routes
ϕ_{\max}	Maximum operation frequency of the interval
ϕ_{\min}	Minimum operation frequency of the interval
λ_{\min}	Minimum number of trains operation on the service route
c_0	Train operation cost per unit time
c_1	Service route operation cost per unit distance with index l and e
c_2	Passenger waiting cost per unit time
c_3	Passenger riding cost per unit time
c_4	Passenger transfer penalty per person
α_i^o	The turn-back capacity of station i in the direction of o
β	The total number of available vehicles
γ	Train capacity surplus
b_l	Maximum number of serving physical stations for service route l in one direction
t_l	Round-trip time of service route l
d_l	Round-trip distance of service route l
t_a	Time required to pass arc a
p_i	The number of passengers at station i in urban transit line
q_e	The volume of passengers flow in section e in urban transit line
g_{κ}^w	The number of trips generated by O-D pair w from station κ
G^w	The number of trips generated by O-D pair w
π_{li}	Binary parameter, equals one if service route l visits station i
k_{le}	Binary parameter, equals one if service route l visits section e
σ_l^{io}	Binary parameter, equals one if the service l turns back at station i to direction o
<i>Definitional variables</i>	
μ_l	Binary variable, equals one if the l th service route is activated
θ_f	Binary variable, equals one if the f th train is activated
x_{fl}	Binary variable, equals one if the f th train operation is on service route l
v_a^w	Nonnegative continuous variable, flow of O-D pair w over arc a
γ_{κ}^w	Nonnegative continuous variable, waiting time of O-D pair w at station κ

$$\sum_{f \in F} \sum_{l \in L} x_{fl} \cdot \sigma_l^{io} \leq \alpha_i, \quad \forall i \in M, o \in O. \quad (7)$$

$$\sum_{f \in F} \sum_{l \in L} x_{fl} \leq \beta. \quad (8)$$

(v) Fleet-size constraint: in an urban rail transit line, the total number of available trains is given and limited, and the overall transport capacity of the trains should be able to meet the passenger demand with the minimum number of trains wherein the number of trains operated should not exceed the fleet size:

(vi) Constraint (10) is a validity inequality and imposes the limitation that the $(f + 1)$ th train can only be activated if and only if the f th train is activated. Constraint (10) indicates that only one service route can be selected and operated for a train in the activated state:

$$\theta_f \leq \theta_{f+1l}, \quad \forall f \in \frac{F}{|F|}, \quad (9)$$

$$\sum_{l \in L} x_{fl} = \theta_f, \quad \forall f \in F. \quad (10)$$

4.3.2. PAP Model. A complete trip of a passenger is composed of three portions: (i) waiting for a train in a station, (ii) riding in the vehicle, and (iii) alighting to egress. There may additionally be transfers. We present a model including the passenger waiting time cost, in-vehicle time cost, and transfer cost. In the optimal strategy, the passengers would board the first train they meet without preferring any other trains [3]. The PAP model is built assuming that there is no congestion and that the train arrivals follow an exponential distribution:

$$\begin{aligned} (\text{PAP}) \min z_2 = & \sum_{\kappa \in N} \sum_{w \in W} y_{\kappa}^w \cdot c_2 + \sum_{w \in W} \sum_{a \in A_2} v_a^w \cdot t_a \cdot c_3 \\ & + \sum_{w \in W} \left(\sum_{a \in A_1} v_a^w - G^w \right) \cdot c_4, \end{aligned} \quad (11)$$

$$\text{s.t. } v_a^w \leq y_{\kappa}^w \sum_{f \in F} x_{fa}, \quad \forall a \in A_{\kappa}^+, \kappa \in N, w \in W, \quad (12)$$

$$\sum_{a \in A_{\kappa}^+} v_a^w = \sum_{a \in A_{\kappa}^-} v_a^w + g_{\kappa}^w, \quad \forall \kappa \in N \cup V, w \in W, \quad (13)$$

$$\sum_{w \in W} v_a^w \leq \sum_{f \in F} x_{fa} \cdot u, \quad \forall a \in A. \quad (14)$$

In the objective function (11), the first term is the passenger waiting time cost and reflects the service frequency. The second term is the ride time cost in the vehicle. The third term is the transfer penalty put in place to ensure that passengers enjoy direct service as much as possible. Constraint (12) is a relaxation evolved from the original model which contains the waiting time, frequency, and arc flow from the nodes. It distributes the passenger demand from nodes to arcs. The volume of the arc flow is influenced by the service frequency. Constraint (13) imposes flow conservation on the nodes. Constraint (14) ensures that the flow on the arc does not exceed the arc capacity.

5. MINLP Model

The model of URTSDP including two submodels, i.e., model SRDP (20)–(22) and model PAP (11)–(14) which can be simplified as (15)–(19). Objective functions (1) and (15) constitute an integrated optimization model owing to that they are measured by the same dimension.

$$(\text{MINLP}) \min z = z_1 + z_2$$

$$\text{s.t. Express service line pool: (20)–(22)}$$

Solving route and frequency problem: (3)–(10)

Solving passenger assignment problem: (16)–(19)

The above model is a mixed-integer nonlinear program (MINLP) since constraints (12) and (16) are nonlinear.

5.1. Model Enhancement

5.1.1. Passenger Assignment Model Simplification. In the optimal strategy model, we obtain the volume of passenger flow by considering each OD pair. As the volume of passenger flow and the number of stations increase, the difficulty in solving the model increases correspondingly. It is a given property that the total passenger waiting time would not be affected if the OD pairs are aggregated. Consequently, we introduce the new variables v_a^i to denote the number of trips to station i over arc a , y_{κ}^i to denote the waiting time of passengers from station κ to station i , and G_i to denote the number of passengers at destination i . A new model is described as follows:

$$\begin{aligned} \min z_2 = & \sum_{\kappa \in N} \sum_{i \in N} y_{\kappa}^i \cdot c_2 + \sum_{i \in N} \sum_{a \in A_2} v_a^i \cdot t_a \cdot c_3 \\ & + \sum_{i \in N} \left(\sum_{a \in A_1} v_a^i - G^i \right) \cdot c_4, \end{aligned} \quad (15)$$

$$\text{s.t. } v_a^i \leq y_{\kappa}^i \sum_{f \in F} x_{fa}, \quad \forall a \in A_{\kappa}^+, i \in N, \kappa \in N, \quad (16)$$

$$\sum_{a \in A_{\kappa}^+} v_a^i = \sum_{a \in A_{\kappa}^-} v_a^i + g_{\kappa}^i, \quad \forall \kappa \in N \cup V, i \in N, \quad (17)$$

$$\sum_{i \in N} v_a^i \leq \sum_{f \in F} x_{fa} \cdot c, \quad \forall a \in A, \quad (18)$$

$$v_a^i, y_{\kappa}^i \geq 0, \quad \forall i \in N, a \in A, \kappa \in N. \quad (19)$$

5.1.2. Problem Reduction Based on Line Pool. As aforementioned, the service network must be completed before using the optimal strategy to tackle the assignment problem. The network is constructed based on the service routes, wherein a choice needs to be made if each station is to be served by an express service or not. Accordingly, for an urban rail transit line with numerous stations, the number of variables will increase exponentially. The binary parameter π_{li} indicates whether service route l stops at station i . There are 2^{n-2} possible service routes for express trains in the service network in a line with n stations. Additionally, many more variables will be generated for the v_a^i , which denotes the number of trips to station i over arc a because the three aforementioned types of arcs in the outbound and inbound directions for each service route must be considered. In fact, express service routes stop at special stations which attract large numbers of passengers in special areas or specific times

instead of at all the stations on the urban rail transit lines. In light of this, we define the set N_2 to account for the high-priority stations on the urban rail transit line. All express service trains must stop at these stations if the service route covers them. Constraint (21) is used to impose the relationship between the express service routes and the stations. Hence, the reduced service routes set can be obtained as the *line pool*. We will analyze the information on the stations as much as possible, especially the volume of the passenger flow. As a result, the *line pool* can be modeled by satisfying the following conditions:

$$\sum_{i \in N} \pi_{li} \leq b_l \cdot \mu_l, \quad \forall l \in L_1, \quad (20)$$

$$\pi_{li} = \mu_l, \quad \forall i \in N_2 \cap N_l, l \in L_1, \quad (21)$$

$$\pi_{li} \leq \mu_l, \quad \forall i \in N_1, l \in L_1. \quad (22)$$

Constraint (20) implies that the total number of stations served by service route l should not exceed the maximum number allowed. Equation (21) ensures that each high-priority station is served by at least one express service. Constraint (22) states that low-priority stations may be served by the express service routes.

5.2. Model Linearization. Constraint (12) is nonlinear because it is the product of a binary variable x_{fa} representing whether the f th train operates on arc a and a continuous variable y_i^w denoting the waiting time to station i for an OD pair w . Here, we transform the constraint into an MILP model and then solve it using CPLEX to obtain the (nearly) optimal solution. The reformulation linearization technique (RLT) replaces the nonlinear polynomials with new variables and adds linear constraints to the newly added variables to achieve the transformation [26]. We take the polynomial xy as an example where x is a binary variable and y is a nonnegative continuous variable. Given the assumption that $\underline{y} \leq y \leq \bar{y}$, the following formula can be obtained via the RLT:

$$\begin{aligned} 0 - \underline{x}y - x\underline{y} + \underline{x}\underline{y} &\geq 0, 0 - \bar{x}y - x\bar{y} + \underline{x}\underline{y} \geq 0, \\ \underline{x}y + x\bar{y} - \underline{x}\bar{y} - 0 &\geq 0, \bar{x}y + x\underline{x} - \bar{x}\underline{y} - 0 \geq 0. \end{aligned} \quad (23)$$

The nonlinearities in (12) and (16) result from the product of the service route frequency variable x_{fa} and the waiting time variable y_{κ}^i and can be transformed into equivalent linear conditions by using inequality (23):

$$v_{fa}^i \leq M \cdot x_{fa}, \quad \forall f \in F, \forall a \in A_{\kappa}^+, i \in N, \kappa \in N, \quad (24)$$

$$v_{fa}^i \leq y_{\kappa}^i, \quad \forall f \in F, \forall a \in A_{\kappa}^+, i \in N, \kappa \in N, \quad (25)$$

$$v_{fa}^i \geq y_{\kappa}^i - M(1 - x_{fa}), \quad \forall f \in F, \forall a \in A_{\kappa}^+, i \in N, \kappa \in N, \quad (26)$$

$$v_a^i \leq \sum_{f \in F} v_{fa}^i, \quad \forall a \in A_{\kappa}^+, i \in N, \kappa \in N, \quad (27)$$

$$v_{fa}^i \geq 0, \quad \forall f \in F, \forall a \in A_{\kappa}^+, i \in N, \kappa \in N. \quad (28)$$

In the above constraints, M is a large positive constant that can be calculated using the equation $M = \sum_{\kappa \in N, \kappa \neq i} g_{\kappa}^i$. Obviously, we have $v_{fa}^i = y_{\kappa}^i$ if and only if $x_{fa} = 1$. Constraint (28) limits the value of the variable v_{fa}^i . Constraints (4) and (5) can be solved by the textbook linearization method.

5.2.1. MILP Model. By the above linearization technique, a mixed-integer linear program (MILP) model can be obtained which is shown as follows. Commercial solvers like CPLEX can be used to solve the model:

$$\text{(MILP) min } z = z_1 + z_2$$

$$\text{s.t. Express service line pool: (20)–(22)}$$

$$\text{Solving route and frequency problem: (3)–(10)}$$

$$\text{Solving passenger assignment problem: (17)–(19) and (24)–(28)}$$

6. Solution Method

The MILP model can be solved by the CPLEX solver or other commercial solvers; however, preliminary computations show that the solution computational burden increases significantly as the number of variables and constraints increase. In this case, optimal schemes can be obtained by directly using commercial solvers only for small-scale problems, for example, an urban rail transit line with 10 stations. For a large-scale line, it would take more computational resources and time to obtain solutions than within a feasible limit. Therefore, it is essential to develop an algorithm to efficiently obtain high-quality solutions.

Based on the characteristics of the URTSDP, we design a search algorithm to solve large-scale problems. We first solve the SRDP model to obtain an optimal route operation scheme, wherein the cost of the operator is minimized (denoted as $z_{\text{operator}}^{\min}$) while the users' cost may not be extremal. We then introduce a new constraint for the value of the SRDP model. The (nearly) optimal solution is obtained iteratively by increasing the constraint gradually with a certain step based on $z_{\text{operator}}^{\min}$.

In summary, the search algorithm shown in Figure 4 includes two main parts: (1) determining the search range of the SRDP model and (2) obtaining the solution iteratively.

6.1. Operator Cost Screening Procedure. The passenger assignment problem model is solved based on frequency. Frequency of each route can be obtained by solving SRDP to optimal, wherein we can obtain the x_{fl} and π_{li} for operation scheme. The service network can be generated based on these two parameters. In this case, the PAP model is linear without transformation and can be solved by the commercial solver to obtain the user's cost. To do this, the total cost (denote as z') is easy to calculate, and the lowest value of the operator's cost is determined as $z_{\text{operator}}^{\min}$. The upper bound of

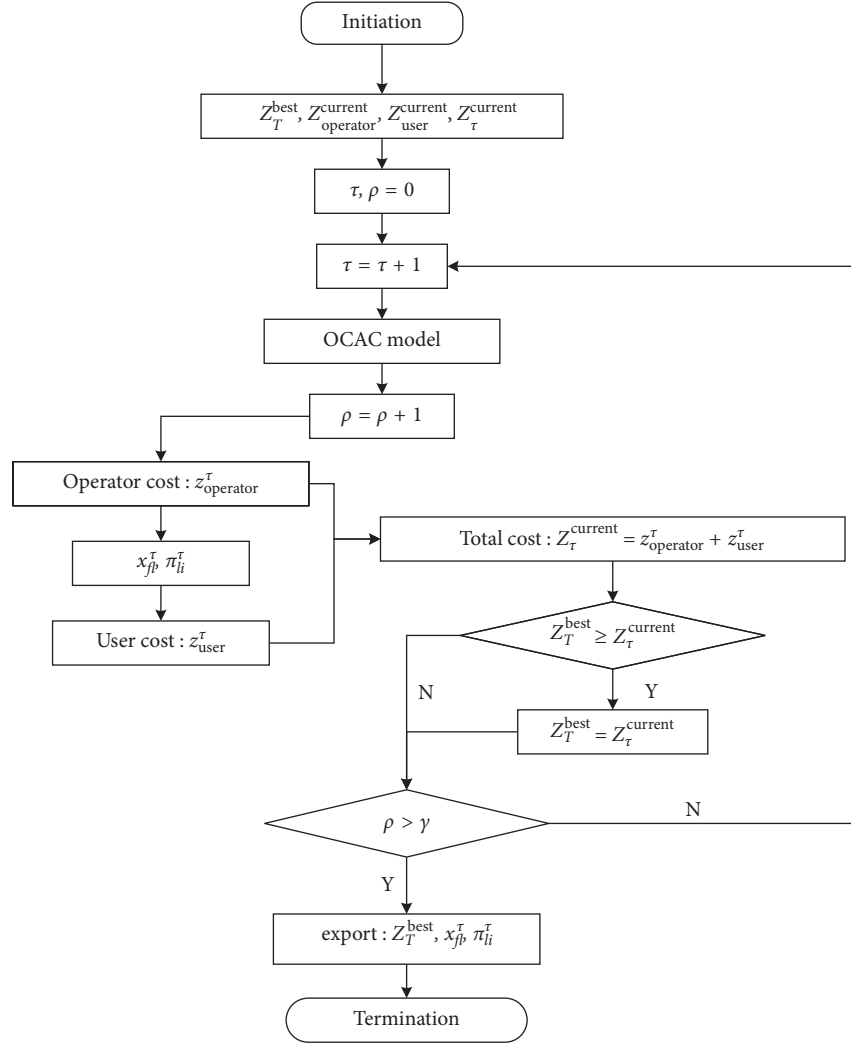


FIGURE 4: Flowchart of the search algorithm.

the SRDP model needs to be obtained by solving the PAP model to optimal. In general, passengers get better trips without the need for transfers when there is only one all-stop route with the maximal frequency. SRDP and PAP models can be solved based on the service network, which is constructed based on the one all-stop route with maximal frequency. The minimal user's cost hence can be obtained (denoted as z_{user}^{\min}). Finally, operator's screening range is generated, and $\Delta = z' - z_{\text{user}}^{\min} - z_{\text{operator}}^{\min}$. Here, we assume that there exists an operator cost z_1 , corresponding to a user's cost z_2 . If $z_1 \geq z' - z_{\text{user}}^{\min}$, then $z_1 + z_{\text{user}}^{\min} \geq z'$; i.e., the total cost based on z_1 and z_2 is greater than z' , and we therefore do not need to search the operator's cost which is beyond the range.

We can then introduce the operator cost augmented constraint (OCAC) in the following inequality shown in the SRDP model. This constraint is similar to the augmented ε -constraint method [27], and the inequality is a modified version based on it. Additionally, the parameter Ω is changeable and $\Omega = (\Delta \cdot \rho)/\eta$, wherein η is the number to

partition the operator cost range Δ and $\rho \in \{0, 1, 2, 3, \dots, \eta\}$. It is hence explicit to see that the right side of the inequality increases gradually as ρ increases:

$$(\text{OCAC})z_1 \geq z_{\text{operator}}^{\min} + \Omega. \quad (29)$$

6.2. Obtaining Solution in Iterative Algorithm

Step 1: initialize the parameters, including the station, and set the number of iterations, $i = 0$. Set the value of γ . Set $Z_T^{\text{best}} = \infty$, $Z_\tau^{\text{current}} = 0$.

Step 2: increase the number of iterations, i.e., $\tau = \tau + 1$.

Step 3: solve the SRDP model, and record the operator's cost which is assigned to z_{operator}^τ and the value of x_{fb}^τ , π_{li}^τ in iteration τ .

Step 4: generate passenger assignment service network in accordance with x_{fb}^τ and π_{li}^τ as input to the PAP model. Solve the model to obtain the value of

passenger's waiting cost which is assigned to z_{user}^{τ} and transfer cost in stage τ .

Step 5: set $Z_{\tau}^{\text{current}} = z_{\text{operator}}^{\tau} + z_{\text{user}}^{\tau}$, and record the value $Z_{\tau}^{\text{current}}$ in stage τ .

Step 6: check current value $Z_{\tau}^{\text{current}}$ and Z_T^{best} : if $Z_{\tau}^{\text{current}} \leq Z_T^{\text{best}}$, $Z_T^{\text{best}} = Z_{\tau}^{\text{current}}$; else turn to step 7.

Step 7: increase the value of ρ by 1

Step 8: check the stopping criteria. If $\rho > \gamma$, stop procedure and export Z_T^{best} , x_{jl}^{τ} , π_{jl}^{τ} ; otherwise, return to Step 2.

7. Numerical Studies

In this section, an urban rail transit line with 19 stations is used to verify the effectiveness of the model for urban rail transit in Chengdu, China. There are three service strategies: short-turn, express, and local services. We generate several instances to evaluate the solution algorithm by comparing the results with the MILP model.

The optimization results of the model coded in C# were obtained by invoking CPLEX 12.8 on the Visual Studio 2017 programming platform on a PC with an Intel Core i5-6200U 2.40 GHz CPU and 12 GB RAM. A limit on the total number of services was imposed throughout the entire optimization process.

7.1. Instance Generation. To verify the computational performance of the solution algorithm, the 5 OD matrices shown in Figure 5 were generated. The lines were generated randomly, and they each consist of six to ten stations with one turn-back station.

For convenience, we describe each line with the triplet (A, B, C) . Here, A denotes the line number, B denotes the number of stations, and C denotes the number of turn-back stations in the line. For example, $(1, 5, 1)$ denotes route 1 with five stations and one turn-back station.

7.2. Comparison of MILP Model and Search Algorithm. We analyzed the computational results of the MILP model. When we attempted to solve a line with 12 stations and 1 turn-back station, the optimal solution could not be obtained within the time limit. There was a significant difference in the computation time between 10 stations and 9 stations. The other details are displayed in Table 3.

The results obtained from the preliminary analysis of the two methods are compared in Table 3. These results highlight the differences between the MILP model and the search algorithm. First, the computational time increases exponentially in the MILP model. In contrast, it increases slowly in the search algorithm. This discrepancy can be attributed to the considerable scale of the arcs generated by all the feasible service routes. As Table 3 shows, the computational time of 2535.0 s for the MILP model is much larger than the 172.1 s of the search algorithm when there are 10 stations on the lines. It is apparent that the number of variables is the distinguishing feature between the two methods; there were up to 1754417 variables in the MILP model, while the search algorithm had only 2100

variables in the $(5, 10, 1)$ instance. Notably, small gaps in the results between the two methods were obtained despite the significant difference in the two aspects mentioned above. From the table, the maximum gap is 1.93% and the minimum gap is 0.19%, and there is an average gap of 1.31% in the five instances. In addition, the MILP model took considerable computational time to solve the line with 10 stations, and we could not even obtain a feasible solution when the line contained 12 stations with the time limit 12 h. The small gaps demonstrate that the search algorithm can be used to obtain high-quality operation schemes.

7.3. Practical Case and Results. A practical line with a length of 86.6 km and 5 turn-back stations is shown in Figure 6. $n_1, n_{12}, n_{16}, n_{17}$, and n_{19} are the turn-back stations at which only trains along one direction can turn back. For the purpose of verifying the search algorithm, we use the demand in the evening peak period shown in Figure 7(c).

The parameters are set as follows:

- (1) Total available train size = 42 trains
- (2) Train capacities: local service = 1520 passenger/train and express service = 1520 passenger/train
- (3) Turn-back capacities: total available turn-back stations are $n_1, n_{12}, n_{16}, n_{17}$, and n_{19} , and turn-back capacities of them are all 20 trains/h. One-way turning back time = 3 min
- (4) Operating cost per kilometer: local service = 185.44 RMB/km and express service = 185.44 RMB/km
- (5) Ownership cost per hour: local service = 467 RMB/h and express service = 467 RMB/h
- (6) Passengers cost: waiting cost = 31 RMB/h, vehicle time cost = 31 RMB/h, and transfer penalty cost = 5 RMB/h
- (7) Service level per route: maximum frequency = 20 trains/h, minimum frequency = 6 trains/h, and surplus capacity ratio = 26.3%

Practical case is introduced to determine the (nearly) optimal combinations of service strategies in previous sections, the operation scheme is shown in Figure 8, and details are obtained as follows:

- (1) Optimal express service pattern:

Express service route l_1 : 1 → 5 → 6 → 10 → 12

Express service route l_2 : 1 → 3 → 4 → 5 → 6 → 7 → 8 → 9 → 10 → 11 → 12 → 13 → 14 → 15 → 16 → 17

Local service route l_3 : 1 → 2 → 3 → 4 → 5 → 6 → 7 → 8 → 9 → 10 → 11 → 12

Local service route l_4 : 1 → 2 → 3 → 4 → 5 → 6 → 7 → 8 → 9 → 10 → 11 → 12 → 13 → 14 → 15 → 16 → 17 → 18 → 19

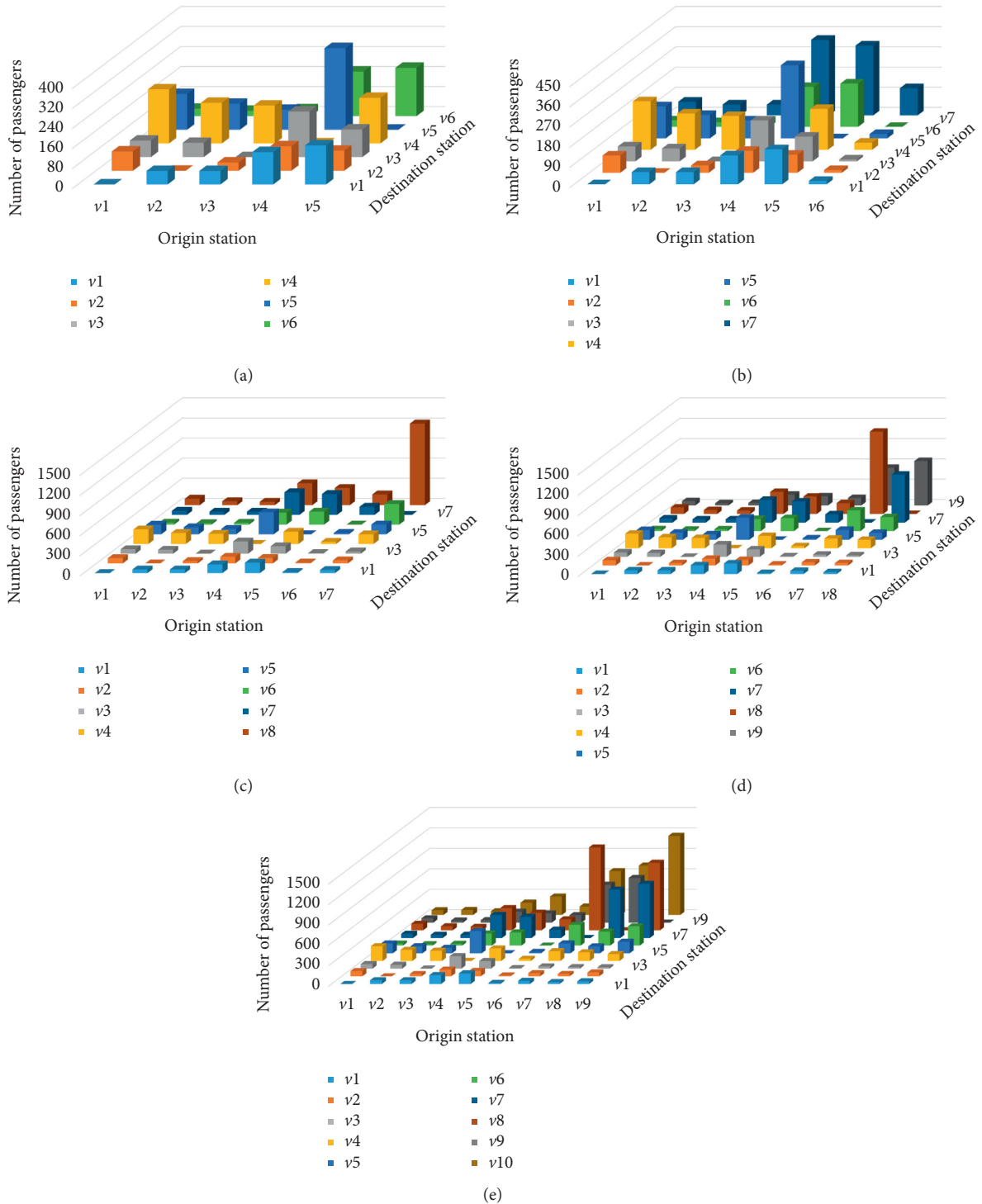


FIGURE 5: Passenger OD matrices of 5 lines. (a) Six-station OD demand. (b) Seven-station OD demand. (c) Eight-station OD demand. (d) Nine-station OD demand. (e) Ten-station OD demand.

(2) Optimal operation frequency for each service:

- l_1 : 10 trains/h
- l_2 : 8 trains/h
- l_3 : 8 trains/h
- l_4 : 10 trains/h

(3) Urban rail transit total cost: 825629 RMB

(4) Computation time: 3613.9 s

Figure 8 shows that the operation scheme contains four service routes consisting of two express service routes and two local service routes. One of the most notable aspects of

TABLE 3: Computational results of the models.

Instances	Variable size	MILP model				Search Algorithm				
		Objective value (RMB)	Best bound (RMB)	Gap (%)	Time (s)	Variable size	Objective value (RMB)	Gap (%)	Time (s)	
(1, 6, 1)	51978	46609	46609	0.00	2.8	1044	47253	1.38	74.2	
(2, 7, 1)	119147	55378	55374	0.01	5.3	1281	56442	1.93	77.7	
(3, 8, 1)	286494	79889	79887	0.00	44.3	1536	80043	0.19	125.8	
(4, 9, 1)	697875	124740	124734	0.00	300.3	1809	126401	1.34	127.5	
(5, 10, 1)	1754417	198218	198210	0.00	2535.0	2100	201629	1.72	172.1	

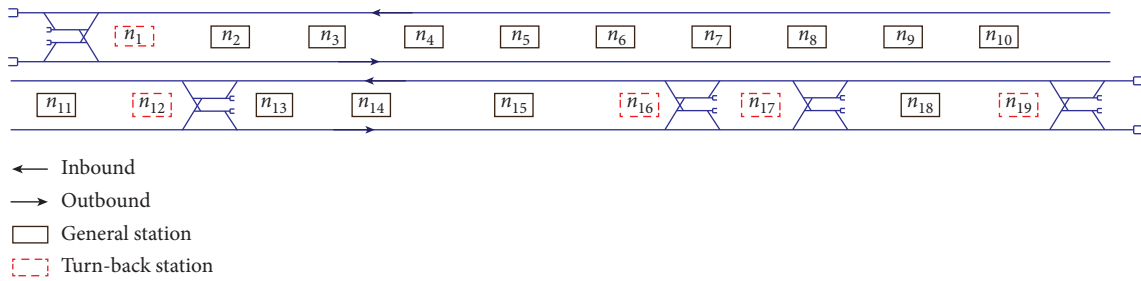


FIGURE 6: Layout of an urban transit line in Chengdu.

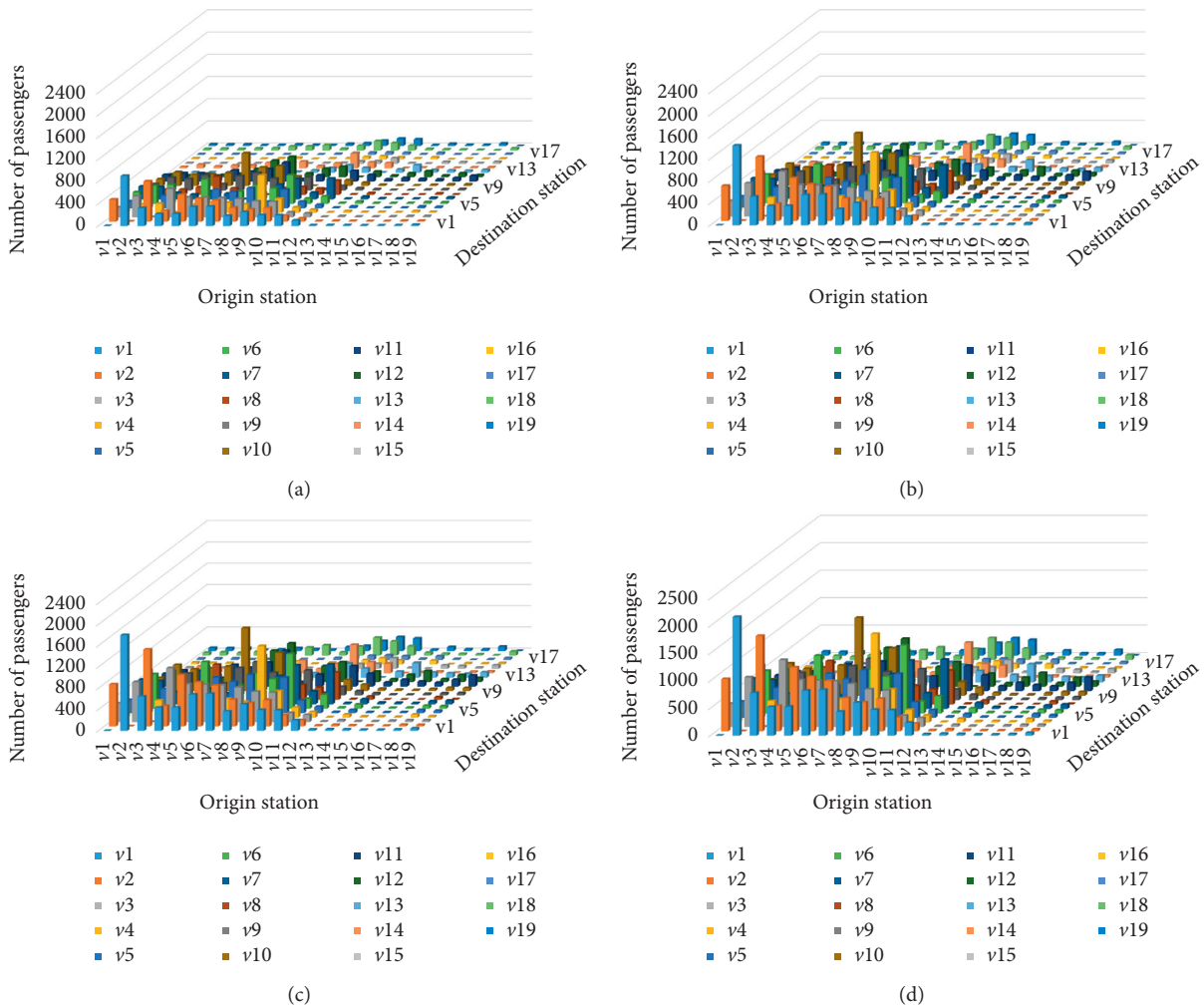


FIGURE 7: Multiple passenger demand levels. (a) 50% demand with respect to original demand. (b) 80% demand with respect to original demand. (c) Original demand. (d) 120% demand with respect to original demand.

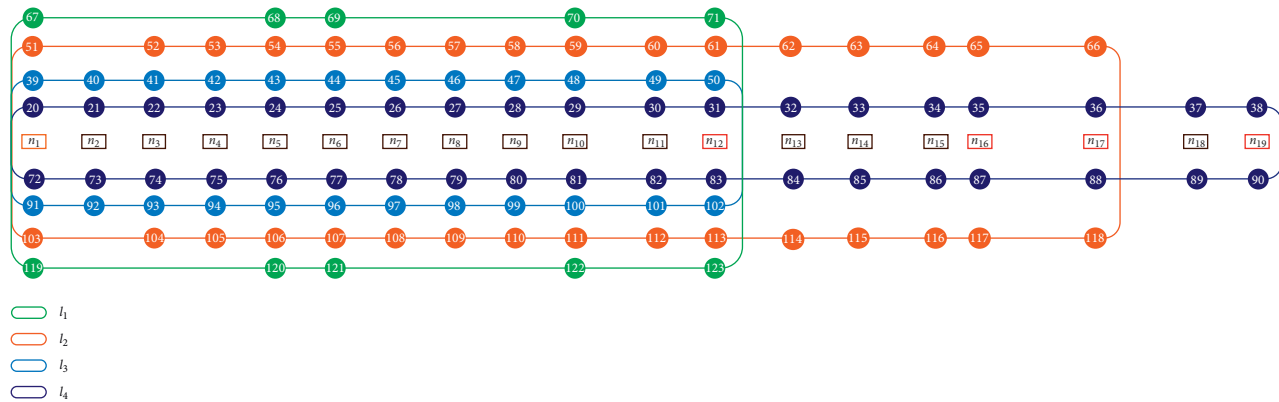


FIGURE 8: Optimized operation scheme for an urban rail transit line.

the operation scheme is the simultaneous existence of both express and local services in one route in which the express trains along service route l_1 stop at only five stations. Another express service route contains 16 stopping stations, except for station n_2 . A possible explanation for the difference in the number of stopping stations in the two express service routes is the passenger demand for medium- to long-distance travel. Skipping stations allows the express service routes to provide different transport capacities to all the stations covered while meeting the passenger demand at each station. Therefore, the operation scheme is different for different passenger demand levels. To verify the effectiveness of the algorithm, the effect of the demand level on the number of service routes and stopping pattern is further studied. Operation schemes were generated with different given levels of demand in the multiple-scenario analysis. Figure 7 shows the multiple demand levels used for evaluation.

7.4. Multiple-Scenario Analysis. Using the solution algorithm, different operation schemes can be obtained under different scenarios, such as varying levels of passenger demand. The impact of the passenger demand can be evaluated by inputting various levels of passenger demand into the solution algorithm. We evaluated the optimization capability of the solution algorithm at different demand levels and obtained distinguishing results with different express service routes, costs, and frequencies. The results for the multiple scenarios are shown in Table 4.

As can be seen from Table 4, high-quality operation schemes were generated for different demand levels by the search algorithm. There are distinctions in the total cost, transit line setting, solution time, and service frequency between the three passenger demand scenarios. The 120% demand scenario has the largest total cost, solution time, and frequency of 912395 RMB, 3620 s, and 38 trains/h, respectively, because of the high demand level in this period. It can be concluded that the higher the level of demand, the higher the cost and the solution time. It is hence necessary for operators and users to design suitable operation schemes.

To further analyze how the parameters influence the operation scheme and total cost, we varied the number of search range partitions. As Figure 9 shows, the number of grids determines the partitioning of the search range. This value is of great importance in exploring the variable space. In this experiment, the number of grids was gradually increased from 0 to 1800. A larger value indicates a more finely divided search range with a higher possibility of finding a higher-quality solution. However, more resources are required. When the number of grids is 0, the augmented model does not place a constraint on the objective function. When the number of grids is between 0–200 and 500–600, the total cost shows a rapid downward trend. When the number of grids reaches 1000, the value of the objective function changes only slightly as the number of grids is further increased. This shows that a rougher grid may miss the optimal solution, whereas an excessively high grid number does not significantly improve the search accuracy.

7.5. Comparison with Empirical Operation Scheme. We compared two schemes, namely, (i) a high-quality operation scheme based on the solution algorithm and (ii) an empirical scheme in operation in the urban rail transit line to verify the search algorithm. The empirical routes and the morning peak passenger volume were given and the differences in the total cost, frequencies, and line settings were determined. Table 5 shows a detailed comparison of the two schemes.

There are three service routes in the empirical scheme, while four service routes exist in the high-quality scheme. This implies that there are many more nodes in the high-quality scheme than in the empirical scheme. However, more trains are employed in the empirical scheme, and the total cost is higher. The cost of 825629 RMB in the high-quality scheme is 7.76% lower than the 895078 RMB in the empirical scheme. Additionally, the costs to the operator and the users in the high-quality scheme are both lower than those of the empirical scheme with relative cost reductions of 2.18% and 12.96%, respectively. There are two likely causes for the total cost differences between the two schemes. From the operator's perspective, the higher cost to the operator in the empirical scheme is attributed to the larger number of operating trains due to the positive correlation between the

TABLE 4: Scenarios analysis with respect to original demand levels.

Demand with respect to original demand (%)	Total cost (RMB)	Solution time (s)	Transit lines setting	Service frequency (trains/h)
120	912395	3620	1 → 4 → 5 → 6 → 7 → 8 → 9 → 10 → 11 → 12 → 13 → 14 → 15 → 16 → 17 1 → 2 → 3 → 4 → 5 → 6 → 7 → 8 → 9 → 10 → 11 → 12 1 → 2 → 3 → 4 → 5 → 6 → 7 → 8 → 9 → 10 → 11 → 12 → 13 → 14 → 15 → 16 → 17 → 18 → 19	8 20 10
80	630096	1681.1	1 → 3 → 4 → 5 → 6 → 7 → 8 → 9 → 10 → 12 1 → 2 → 3 → 4 → 5 → 6 → 7 → 8 → 9 → 10 → 11 → 12 1 → 2 → 3 → 5 → 6 → 9 → 10 → 11 → 12 → 13 → 14 → 15 → 16 → 17 → 18 → 19	8 8 10
50	421167	1305.5	1 → 2 → 3 → 4 → 5 → 6 → 7 → 8 → 9 → 10 → 11 → 12 1 → 2 → 3 → 4 → 5 → 6 → 7 → 8 → 9 → 10 → 11 → 12 → 13 → 14 → 15 → 16 → 17 → 18 → 19	8 8

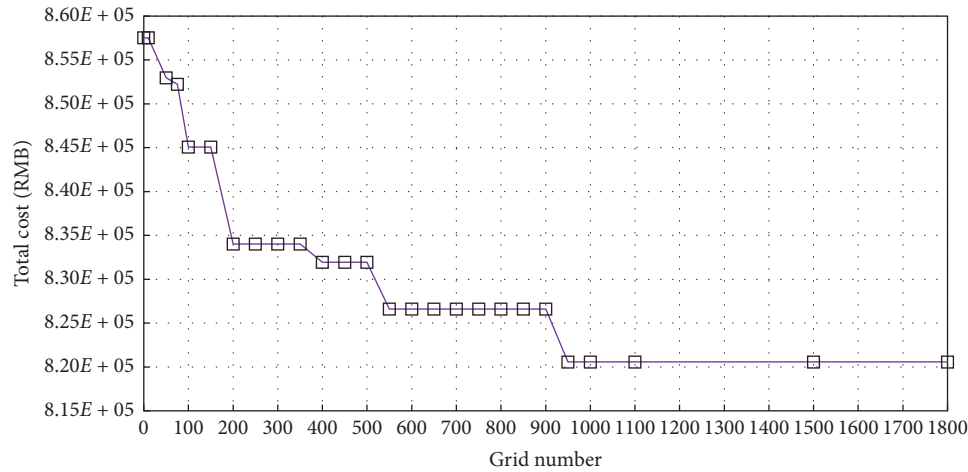


FIGURE 9: Sensitivity studies on total costs based on different partitions.

TABLE 5: Result comparison of the two schemes.

Scheme type	Total cost (RMB)	Operator cost (RMB)	User cost (RMB)	Transit lines setting	Service frequency (trains/h)
High-quality scheme	825629	422333	403296	1 → 5 → 6 → 10 → 12	10
				1 → 3 → 4 → 5 → 6 → 7 → 8 → 9 → 10 → 11 → 12 → 13 → 14 → 15 → 16 → 17	8
				1 → 3 → 4 → 5 → 6 → 7 → 8 → 9 → 10 → 11 → 12	8
				1 → 2 → 3 → 4 → 5 → 6 → 7 → 8 → 9 → 10 → 11 → 12 → 13 → 14 → 15 → 16 → 17 → 18 → 19	10
				1 → 2 → 3 → 4 → 5 → 7 → 10 → 15 → 16	8
Empirical scheme	895078	431728	463350	1 → 2 → 3 → 4 → 5 → 6 → 7 → 8 → 9 → 10 → 11 → 12	24
				1 → 2 → 3 → 4 → 5 → 6 → 7 → 8 → 9 → 10 → 11 → 12 → 13 → 14 → 15 → 16 → 17 → 18 → 19	8

cost and the number of trains. 40 trains are employed in the empirical scheme, while the employment of only 36 trains in the high-quality scheme can save considerable vehicle purchase costs. From the passengers' perspective, the provision of more service routes to passengers gives them the choice for flexible travel options without waiting for a specific service. For example, passengers can use the express service route from station 1 to station 12 or other stations, but they can also use the local service route when more waiting time is needed for the express train.

Accordingly, we carried out a detailed analysis of the section flows and capacity utilization ratios (CURs) in the two schemes. By comparing the passenger flow volumes in the sections, the influence of the service setting on the passenger flow accumulation can be understood. The CUR indicates whether the service design is reasonable. The empirical scheme and high-quality scheme are, respectively, denoted in the figure as ES and OS for short. ES and OS, respectively, signify the capacities of the empirical and optimized schemes in each section.

As shown in Figure 10, the CUR of the OS is higher than that of ES in most sections because the transport

capacity of most sections in the OS is higher than that of the ES, indicating that the trains are distributed more efficiently. The maximum OS CUR is 63.56%, which is considerably higher than the ES CUR of 57.20% at e_6 in the inbound direction. Moreover, the average OS CUR is 0.52% higher than the average ES CUR. In the outbound direction, the maximum OS CUR is 57.65% higher than the ES CUR of 51.89% at e_8 . The reasons for the differences are clear: the OS uses express services in conjunction with short-turn services, and service routes from station 1 to station 12 employ both express service and local services. The transport capacity can thus be fully utilized. In addition, the passenger demand is mainly concentrated at the first twelve stations, which account for 81.37% of the total passenger demand. These indicators explain the rationality and efficiency of the scheme based on the proposed model.

Regarding the waiting time, Figure 11 shows that the waiting time of the ES at most stations is lower than that of the OS, that is, passengers have to wait longer for the same service under the high-quality scheme. This is because the train frequency in the ES is higher than that in the OS. Although there is

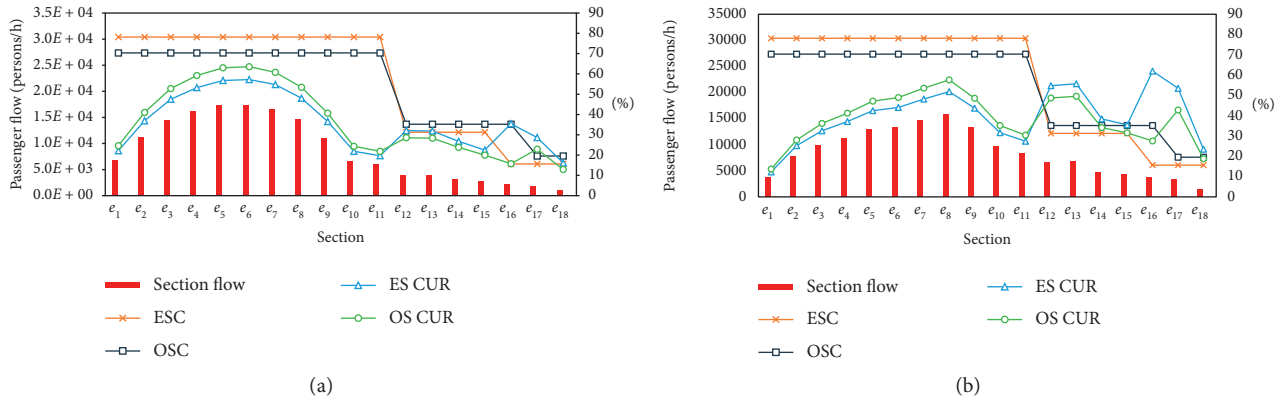


FIGURE 10: Section flow and capacity utilization ratio. (a) Inbound direction. (b) Outbound direction.

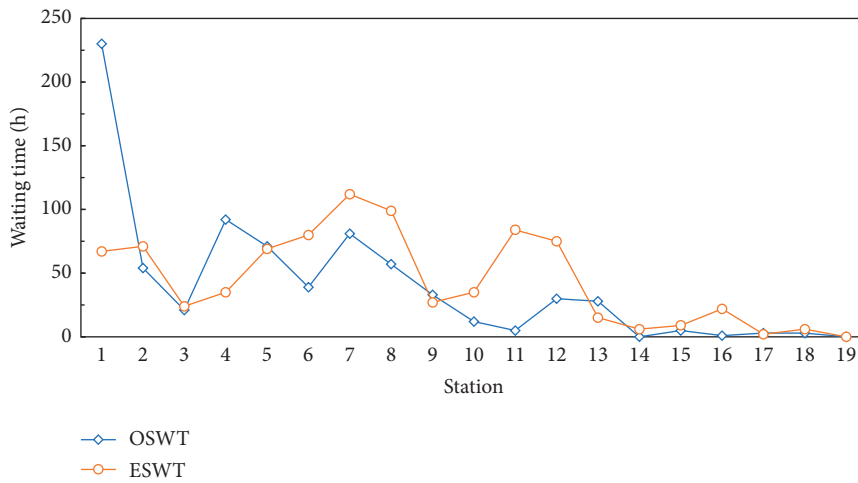


FIGURE 11: Passenger waiting time distribution.

a large difference in the waiting time in station 1 of up to 163 h, the total waiting time in the OS is lower than that of ES, and the reduction in waiting time is 73 h.

In summary, the present results are significant in two major respects. First, the operator’s cost can be reduced by reducing the number of vehicle purchases. In practice, the number of rolling stocks can be reduced through a reasonable and high-efficiency operation scheme as verified by the aforementioned example. Furthermore, a high-quality operation scheme also benefits the passengers and saves their waiting time and riding time by integrating multiple strategies.

8. Conclusions

In this paper, we proposed an MILP model for the URTSDP, wherein short-turn, express, and local services are provided. This problem can be analyzed from two aspects. Firstly, the SRDP determines which station to turn back at, the numbers and frequencies of the service routes, and the routes to run and stations to serve for each express service route. Secondly, an optimal strategy is applied to solve the PAP while considering the capacity restrictions. Because the MILP model

can be solved directly by commercial solvers for small-scale problems, we introduced the *line pool* strategy to reduce the scale of the problem and designed a search algorithm to solve the large-scale problem by minimizing the total cost. In addition, we applied the model and search method in a practical urban rail line under multiple scenarios. This differs from previous works that tested their models and methods only on small-scale scenarios.

Numerical studies were conducted to analyze the performance of the proposed model and the search algorithm. An instance comparison verified that the search algorithm could find high-quality solutions with a gap of 1.31% compared to the MILP model. In particular, compared to the empirical operation scheme, the high-quality scheme demonstrated practicability and effectiveness through a total cost saving of 7.76%. The performance of the search algorithm was tested under multiple scenarios and different parameter values. The multiple-scenario analysis illustrated the applicability of different operation schemes to meet different demands. A sensitivity experiment was performed to investigate the effects of different partitionings of the search range on the total cost and solution time.

This study contributes to academic research by introducing the practical strategy of the *line pool* and the solution algorithm to solve a large-scale problem. However, there are several future refinements that are worth studying. (i) The train size of each service route is not included into the model. It may be worthwhile to use trains of different sizes to cope with various demand levels to further conserve transport capacity. (ii) The local service stopping pattern is not addressed in this model. Addressing this issue can alleviate the demand imbalance and save the operator's cost. (iii) In the future, the model can be developed in conjunction with additional constraints, such as real-time adjustments [28], robustness, and other factors to create a better match between supply and demand.

Data Availability

The data are available by contacting Siyu Tao through taosiyu@swjtu.edu.cn.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the National Key Research and Development Program of China (Project no. 2017YF B1200701), the Research on Complete Technology of Integrated Transportation Hub in Complex Urban Environment (Grant no. cstc2014yykfb30003), and the Research on Functional Layout and Planning of Shapingba Integrated Transportation Hub (Grant no. 2015H01372).

References

- [1] Y. Gao, L. Yang, and Z. Gao, "Energy consumption and travel time analysis for metro lines with express/local mode," *Transportation Research Part D: Transport and Environment*, vol. 60, pp. 7–27, 2018.
- [2] 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 A: Transport Science*, vol. 15, no. 2, pp. 1459–1486, 2019.
- [3] H. Spiess and M. Florian, "Optimal strategies: a new assignment model for transit networks," *Transportation Research Part B: Methodological*, vol. 23, no. 2, pp. 83–102, 1989.
- [4] J. H. Chun, R. M. Anderson, and D. Paik, "The S-train system: synchronized express & local trains for urban commuter rail systems," in *Proceedings of the 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 1586–1591, IEEE, Washington, DC, USA, October 2011.
- [5] A. Jamili and M. Pourseyed Aghae, "Robust stop-skipping patterns in urban railway operations under traffic alteration situation," *Transportation Research Part C: Emerging Technologies*, vol. 61, pp. 63–74, 2015.
- [6] E. A. Abdelhafiez, M. R. Salama, and M. A. Shalaby, "Minimizing passenger travel time in URT system adopting skip-stop strategy," *Journal of Rail Transport Planning & Management*, vol. 7, no. 4, pp. 277–290, 2017.
- [7] Q. Luo, Y. Hou, W. Li, and X. Zhang, "Stop plan of express and local train for regional rail transit line," *Journal of Advanced Transportation*, vol. 2018, Article ID 3179321, 11 pages, 2018.
- [8] A. Yang, B. Wang, J. Huang, and C. Li, "Service replanning in urban rail transit networks: cross-line express trains for reducing the number of passenger transfers and travel time," *Transportation Research Part C: Emerging Technologies*, vol. 115, p. 102629, 2020.
- [9] Y. Sun, P. M. Schonfeld, Y. Lu, and M. Zhou, "Redesigning rail transit short-turn operations: case study of line 2 of the Shanghai metro in China," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2540, no. 1, pp. 46–55, 2016.
- [10] X. Ding, S. Guan, D. J. Sun, and L. Jia, "Short turning pattern for relieving metro congestion during peak hours: the substance coherence of Shanghai, China," *European Transport Research Review*, vol. 10, no. 2, 2018.
- [11] M. Zhang, Y. Wang, S. Su, T. Tang, and B. Ning, "A short turning strategy for train scheduling optimization in an urban rail transit line: the case of Beijing subway line 4," *Journal of Advanced Transportation*, vol. 2018, 2018.
- [12] Z. Li, J. Zhao, and Q. Peng, "Optimal train service design in urban rail transit line with considerations of short-turn service and train size," in *Proceedings of the 8th International Conference on Railway Operations Modelling and Analysis (ICROMA) RailNorrköping 2019*, pp. 665–687, Linköping University Electronic Press, Norrköping, Sweden, June 2019.
- [13] Z. Cao, A. Ceder, D. Li, and S. Zhang, "Robust and optimized urban rail timetabling using a marshaling plan and skip-stop operation," *Transportmetrica A: Transport Science*, vol. 16, no. 3, pp. 1217–1249, 2020.
- [14] L. L. Afanasiev and S. Y. Liberman, "Principles for organizing express bus services," *Transportation Research Part A: General*, vol. 17, no. 5, pp. 343–346, 1983.
- [15] T. A. S. Vijayaraghavan and K. M. Anantharamaiah, "Fleet assignment strategies in urban transportation using express and partial services," *Transportation Research Part A: Policy and Practice*, vol. 29, no. 2, pp. 157–171, 1995.
- [16] M. T. Conlon, P. J. Foote, K. B. O'Malley, and D. G. Stuart, "Successful arterial street limited-stop express bus service in Chicago," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1760, no. 1, pp. 74–80, 2001.
- [17] D. Z. W. Wang and H. K. Lo, "Multi-fleet ferry service network design with passenger preferences for differential services," *Transportation Research Part B: Methodological*, vol. 42, no. 9, pp. 798–822, 2008.
- [18] H. Larrain, R. Giesen, and J. C. Muñoz, "Choosing the right express services for bus corridor with capacity restrictions," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2197, no. 1, pp. 63–70, 2010.
- [19] Y. Y. Ulusoy, S. I.-J. Chien, and C.-H. Wei, "Optimal all-stop, short-turn, and express transit services under heterogeneous demand," *Transportation Research Record*, vol. 2197, no. 1, pp. 8–18, 2010.
- [20] H. Larrain, J. C. Muñoz, and R. Giesen, "Generation and design heuristics for zonal express services," *Transportation Research Part E: Logistics and Transportation Review*, vol. 79, pp. 201–212, 2015.
- [21] D. Z. W. Wang, A. Nayan, and W. Y. Szeto, "Optimal bus service design with limited stop services in a travel corridor," *Transportation Research Part E: Logistics and Transportation Review*, vol. 111, pp. 70–86, 2018.

- [22] A. Ceder, "Optimal design of transit short-turn trips," *Transportation Research Record*, vol. 1221, no. 557, pp. 8–22, 1989.
- [23] P. G. Furth, "Short turning on transit routes," *Transportation Research Record*, vol. 1108, pp. 42–52, 1987.
- [24] P. Delle Site and F. Filippi, "Service optimization for bus corridors with short-turn strategies and variable vehicle size," *Transportation Research Part A: Policy and Practice*, vol. 32, no. 1, pp. 19–38, 1998.
- [25] C. Chriqui and P. Robillard, "Common bus lines," *Transportation Science*, vol. 9, no. 2, pp. 115–121, 1975.
- [26] H. D. Sherali and W. P. Adams, *A Reformulation-Linearization Technique for Solving Discrete and Continuous Non-convex Problems*, Springer Science & Business Media, Berlin, Germany, 2013.
- [27] G. Mavrotas and K. Florios, "An improved version of the augmented constraint method (AUGMECON2) for finding the exact pareto set in multi-objective integer programming problems," *Applied Mathematics and Computation*, vol. 219, no. 18, pp. 9652–9669, 2013.
- [28] Z. Cao, A. Ceder, and S. Zhang, "Real-time schedule adjustments for autonomous public transport vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 109, pp. 60–78, 2019.