

Received April 28, 2022, accepted June 3, 2022, date of publication June 9, 2022, date of current version June 16, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3181815

Optimizing Train Routing Problem in a Multistation High-Speed Railway Hub by a Lagrangian Relaxation Approach

YIDONG WANG^{ID}, RUI SONG, SHIWEI HE, ZILONG SONG, AND JUSHANG CHI

Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing 100044, China

Corresponding author: Rui Song (rsong@bjtu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 62076023, in part by the Key Issues of China Railway Corporation under Grant P2021X009, and in part by the Fundamental Research Funds for the Central Universities (No. 2022JBQY006).

ABSTRACT As the intersection of multiple high-speed railway lines, the multi-station high-speed railway hub is the key to improve the transport efficiency of the high-speed railway network. This paper focuses on the optimization of the multi-station high-speed railway hub and models it as a train routing problem (TRP). Considering the capacity of railway infrastructures and the demand of passengers, a mixed integer linear programming model is proposed to minimize the total cost of train routes and passenger routes. The optimized train routes include the macroscopic routes between stations and the microscopic track allocation inside stations and Electric Multiple Units (EMUs) depots. A Lagrangian relaxation (LR) approach is developed to dualize the hard constraints and decompose the origin model into train and passenger subproblems, then a shortest path algorithm is designed to solve the subproblems independently. Numerical experiments based on an illustrative railway hub network and a real-world network are implemented to demonstrate the effectiveness of the model and algorithm. The solution results prove that the LR approach can obtain high-quality solutions within an acceptable computational time. Compared with the existing fixed scheme, the optimization scheme can reduce the total cost by 37.18% and utilize the railway lines and tracks more reasonably.

INDEX TERMS Multi-station high-speed railway hub, train routing problem, EMUs depot, Lagrangian relaxation, passenger demand.

I. INTRODUCTION

As a key component of the transport system, high-speed railways provide an efficient, comfortable, punctual, and safety service for transport passengers, and the high-speed railway network has been gradually formed worldwide. In the railway network, the distribution of railway lines is closely related to the regional development, so the metropolises usually build multiple stations to connect various railway lines, thus forming the multi-station high-speed railway hubs [1]–[3]. A multi-station high-speed railway hub refers to a railway hub composed of multiple high-speed railway lines and passenger stations, such as Berlin railway hub in Germany, Paris railway hub in France and Zhengzhou railway hub in China. As the intersection of multiple railway lines, the stations in the hub need to operate a large number of trains from various

directions, so the operation efficiency of the multi-station hub would great affect the running state of the entire railway network [4], [5]. Due to the strategic significance, we focus on the operation of the multi-station hub by optimizing the utilization of station resources.

At present, the stations in the multi-station railway hub are independent of each other and only deal with trains from the fixed connecting directions [6]. Although this fixed scheme is clear and easy to formulate, it may lead to the imbalance of operation among different stations. For example, the stations on trunk lines are usually in busy and lack of capacity, while the stations on branch lines are idle and waste their capacity. In view of this weakness, some connecting railway lines are constructed between adjacent stations to transform some trains from busy stations to idle stations. However, the transformation of trains are only regarded as an temporary adjustment in the operational level [7] and not included the tactical plan. In order to make full use of these connecting

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

lines in the tactical level, we attempt to regard all the stations as a unified system and propose a flexible scheme to cooperate the operation of stations. Specifically, in the flexible scheme, the connecting directions of all the stations are flexible, which indicates each station can handle trains from all directions. And this scheme can make overall use of all stations and it is beneficial to play the scale effect of the multi-station hub.

Based on the flexible scheme, the optimization of the multi-station hub can be regarded as a network routing problem for trains from the connecting directions. Train routing problem (TRP) is a typical problem in the railway optimization and it has been well studied by many researchers [1], [5], [7]–[13]. According to the research scopes, the relevant researches include the macroscopic TRP and the microscopic TRP. The macroscopic TRP aims to optimize the train routes between stations in the railway corridors and networks, including the stop patterns and frequencies [4], [5], [14], [15]. In contrast, the microscopic TRP focuses on the train routes inside the individual station, including the track allocation and conflict resolution [16]–[19]. Different from previous researches, the TRP in the multi-station hub should optimize the macroscopic and the microscopic train routes simultaneously. On one hand, the macroscopic train routes between stations and connecting directions should be optimized to minimize the train running cost in the hub. On the other hand, the microscopic train routes inside stations also should be optimized to ensure the reasonable track allocation.

During the process of optimization, many kinds of railway infrastructure capacity will be considered as the constraints [20]. The capacity refers to the maximum number of trains that a railway infrastructure can deal with in a certain time period, such as line capacity and track capacity [21], [22]. Moreover, since we focus on the high-speed railway hub, the capacity of Electric Multiple Units (EMUs) depot is also vital for the optimization. The EMUs depot indicates the yard where the storage and maintenance services of EMUs are carried out [23]–[25], and the storage and maintenance capacity of tracks should be introduced into the formulation.

Furthermore, many recent studies have paid more attention to the passenger demand to improve the service quality of transportation [26]–[30]. As the main service target of high-speed railway hub is the urban residents, the passenger demand should be considered into the optimization. Thus, we attempt to introduce traffic zones to represent the demand of residents in different regions of the city.

Based on the above background, how to integrate the macroscopic TRP and the microscopic TRP considering the capacity of EMUs depots becomes the main challenge of this paper. Thus, the aim of our research is to optimize the train routes and track allocation with considering the passenger demand and capacity of high-speed railway infrastructures. The main contributions are summarized as follows: 1) A mixed integer linear programming model is proposed to minimize the train and passenger cost in the multi-station

railway hub. The macroscopic train routes between stations and the microscopic track allocation inside stations and EMUs depots are optimized simultaneously by considering the capacity of infrastructures; 2) A Lagrangian relaxation approach is developed to solve the model. We dualize the hard constraints into the objective function and decompose the model into train and passenger subproblems, which can be solved by a shortest path algorithm; 3) Numerical experiments based on an illustrative railway hub network and Zhengzhou high-speed railway hub network are implemented to demonstrate the effectiveness of the model and algorithm, and the optimization scheme is compared with the fixed scheme to examine the optimization effect.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literatures. Section 3 presents the problem description and Section 4 formulates the mixed integer linear programming model. The details of the LR approach are given in Section 5. Section 6 provides some numerical experiments to verify the model and algorithm. Conclusions and further work are summarized in Section 7.

II. LITERATURE REVIEW

The core problem of the multi-station high-speed railway hub optimization is a TRP combining the macroscopic train routes and the microscopic track allocation. In past decades, TRP has always been a research hotspot in the field of railway optimization, and we summarize relevant literatures as the macroscopic TRP and the microscopic TRP. Moreover, since the EMUs depot is a vital factor of the high-speed railway hub, we also conclude some related researches.

A. THE MACROSCOPIC TRP

The macroscopic TRP focuses on the train routes in the railway networks and corridors, which usually includes the stop patterns and frequencies of trains from the origins and the destinations. Szpigel [31] firstly used the job scheduling to study the TRP in single-track railway lines, then Carey and Lockwood [11] further extended to double-track railway lines. Lee and Chen [32] proposed a four-step heuristic algorithm to assign the trains to railway lines and generated feasible train routes according a threshold accepting rule. Jia *et al.* [2] constructed a bi-level programming model to optimize the train routes of express freight trains on a high-speed railway corridor. A heuristic approach combining the genetic algorithm (GA) and Frank-Wolfe algorithm was developed to solve the bi-level model.

Due to the tight relations between the train routes and timetables, many researchers usually integrated the TRP with the train timetabling problem (TTP) to acquire more reasonable train plans. Yang *et al.* [14] focused on the integration of the TRP and the TTP on a single-track high-speed railway corridor and proposed a multi-objective mixed integer programming model to minimize the total delay between the origin the actual timetable. Khan and Zhou [15] considered stochastic disturbances into the train timetables and

developed a heuristic algorithm to decompose the formulation into a series of train subproblems. Zhou and Teng [33] developed an integrated model based on an extended space-time-speed network to optimize the train routes and timetables simultaneously. In this formulation, energy consumption of trains with different speed were regarded as the part of train traveling arcs, and a Lagrangian relaxation solution with a dynamic programming algorithm was designed to solve the model. C. Zhang *et al.* [34] proposed a joint optimization of the TRP, TTP and maintenance planning on a double-track high-speed railway corridor, and a heuristic algorithm based on Lagrangian relaxation and dynamic constraint generation was developed. Tan *et al.* [9] proposed a multi-objective programming model to insert extra train routes into the existing timetable and developed a genetic algorithm to acquire high-quality solutions in the real railway network. Gao and Niu [35] focused on the TRP and TTP of multiple type trains with different speeds on a high-speed railway corridor. They relaxed the time windows to generate more flexible train routes and timetables, an improved algorithm based on the alternating direction method of multipliers (ADMM) was introduced for solving the formulation.

In general, since the scale of the macroscopic TRP usually covers a series of stations, the detailed layout of each individual station is ignored. As for the multi-station railway hub, not only the macroscopic train routes between adjacent stations, but also the detailed track allocation of each station should be optimized. Therefore, the existing macroscopic researches cannot be applied to the multi-station hub directly and we need to combine them with the microscopic researches.

B. THE MICROSCOPIC TRP

The microscopic TRP optimizes train routs inside the individual station, including the inbound route, outbound route and track allocation where the train enters and leaves. Compared with the macroscopic TRP, the microscopic TRP is more detailed and can obtain more accurate train routes and timetables. Lusby *et al.* [7] made an explicit literature review on the classical formulations and solutions of the microscopic TRP in previous studies. The most common approach is conflict graph methodology, which can be further divided into the node packing problem [36] and the graph coloring problem [37]. But when applied to the large-scale problem, the formulation based on the conflict graph would become complicated and cannot get high-quality solutions in an acceptable time. Therefore, some other methods were proposed to improve the applicability of the model, such as the heuristic manual method [38] and constraint programming model [39]. Afterwards, more researchers paid more attention to real-life operation and introduced disturbances into TRP. D'Ariano *et al.* [13] considered disturbances in train timetable and proposed an alternative graph model to resolve conflicts between the microscopic train routes inside stations. Pellegrini [40] also considered disturbances in complex stations and represented the capacities of infrastructure with fine granularity. A solution framework called

RECIFE-MILP was developed to obtain feasible solutions quickly. Meng and Zhou [12] developed an integer programming model using space-time network and cumulative flow variables to minimize the total deviation time of trains on an N-track network. A solution framework based on Lagrangian relaxation was designed to decompose the model into train least-cost subproblems, which could be solved by the dynamic programming algorithm. Qi *et al.* [41] integrated the layout of stations into train timetables and proposed a bi-level programming model to minimize the total travel time of all trains, where the upper level determined whether to design new siding tracks and the lower level routed trains. Z. Zhang *et al.* [42] considered the braking failure of trains and proposed a dynamic using strategy of successive routes.

In recent years, some researchers introduced the maintenance planning problem (MPP) into the microscopic TRP and TTP, which can make the train track allocation and timetable more practical for the railway dispatchers. Luan *et al.* [19] focused on the integration of the TRP, TTP and MPP on a railway hub. By introducing cumulative flow variables and representing the maintenance task as virtual trains, they proposed a mixed integer linear programming model based on a space-time network. A Lagrangian relaxation approach was developed to decompose the origin formulation. Y. Zhang *et al.* [17] also proposed a mixed integer linear programming model to minimize the total train travel time and maintenance cost on a railway line consisting of multiple stations, and further added speed restriction constraints into the model. A heuristic iteration algorithm was designed to decompose the model into train subproblems. Q. Zhang *et al.* [8] predefined the track maintenance task inside stations firstly, and then proposed a binary integer programming model to minimize the number of train cancellations and the weighted train travel times on several adjacent stations. An ADMM-based approach was developed to solve the model.

To sum up, the researches on microscopic TRP further detail the layout of stations and obtain the track allocation of trains inside stations. However, the previous researches just focus on a single station or several adjacent stations of a single railway line, as for the multi-station hub with multiple stations and railway lines, the current methods are not fully applicable. Furthermore, with the increasing importance of service quality, more attention should be paid for the passenger demand in the railway optimization. By referring to the methods of recent researches[29], [46]–[48], we attempt to represent the passenger demand of different regions in the city as several independent traffic zones.

C. THE EMUS DEPOT

The EMUs depot is the infrastructure where the EMUs are stored and maintained, which is essential for the hub and usually located next to the station. Recently, some researches have begun to introduce the storage and maintenance capacity of the EMUs depot into the optimization of train plans.

TABLE 1. The methods of the previous researches and this paper.

References	Macro TRP	Micro TRP	EMUs depot	Passenger demand	Model	Solution method
Carey and Lockwood [11]	√	×	×	×	MILP	A heuristic algorithm
Yang <i>et al.</i> [14]	√	×	×	×	Multi-objective MILP	GAMS with CPLEX
Khan and Zhou [15]	√	×	×	×	Two-stage stochastic-recourse model	A heuristic decomposition algorithm
Zhou and Teng [33]	√	×	×	×	ILP	LR with DP
C. Zhang <i>et al.</i> [34]	√	×	×	×	BILP	LR with dynamic constraint generation
Tan <i>et al.</i> [49]	√	×	×	×	Multi-objective MILP	GA
Gao and Niu [35]	√	×	×	×	BILP	ADMM with DP
D'Ariano <i>et al.</i> [13]	×	√	×	×	Alternative graph model	A heuristic iterative rescheduling algorithm
Pellegrini <i>et al.</i> [40]	×	√	×	×	MILP	RECIFE-MILP
Meng and Zhou [12]	×	√	×	×	ILP	LR with DP
Qi <i>et al.</i> [41]	√	√	×	×	Bi-level	GAMS with CPLEX
Luan <i>et al.</i> [19]	×	√	×	×	MILP	LR with DP
Y. Zhang <i>et al.</i> [17]	×	√	×	×	MILP	A heuristic iterative algorithm
Q. Zhang <i>et al.</i> [8]	×	√	×	×	MILP	LR with DP, ADMM with DP
Meng and Zhou [29]	√	×	×	√	MILP	LR with DP
Dong <i>et al.</i> [48]	√	×	×	√	MILP	LNS
H. Li <i>et al.</i> [25]	×	√	√	×	ILP	CPLEX
Shi and Li [24]	×	√	√	×	ILP	VNS
Wang [23]	√	×	√	×	MILP	Ant colony optimization algorithm
W. J. Li <i>et al.</i> [3]	√	×	√	×	BILP	Two-stage optimization with column generation
This paper	√	√	√	√	MILP	Improved LR with Floyd

Notes: MILP- Mixed integer linear programming; ILP- Integer linear programming; BILP- Binary integer linear programming; GA - Genetic algorithm; LR - Lagrangian relaxation; DP - Dynamic programming; ADMM - Alternating direction method of multipliers; LNS/VNS - Large-scale/Variable neighborhood search

Li *et al.* [25] proposed a model to optimize the track utilization of the storage, maintenance and washing area in the EMUs depot. Shi and Li [24] further studied the shunting operation plan among various areas in the EMUs depot. Two integer linear programming models were formulated to minimize the total delay time of trains and a variable neighborhood search algorithm was designed to solve the models. Wang [23] extended the problem scale to the multiple depots in the high-speed railway networks. The sharing strategy and EMUs' running data among different depots were introduced into the optimization of operation schemes. Li *et al.* [3] integrated the maintenance planning of EMUs into the microscopic TRP and proposed two related integrated programming models to optimize the circulation plans of EMUs in a high-speed railway network. A two-stage algorithm based on column generation was developed to solve the models. As for the multi-station high-speed railway hub studied in this paper, the EMUs depot is an important part of the hub and it has a great effect on the train operation scheme. Thus, we should refer the previous researches and introduce the capacity of the EMUs depot into the optimization.

In summary, although many existing literatures have paid much attention on TRP, there is few studies focusing on the multi-station high-speed railway hub. Due to the special characteristics of the multi-station hub, the methods of previous researches are not fully applicable. Specifically, the optimization of the multi-station high-speed railway hub should not only focus on the macroscopic train routes among stations in the hub, but also detail the microscopic track allocation inside each individual station. Besides, the EMUs depot capacity

should be considered, which is also significant to obtain more practical optimization scheme. Therefore, we need to integrate the macroscopic TRP, the microscopic TRP and the capacity of the EMUs depot in the optimization of the multi-station high-speed railway hub. The methods of the previous researches and this paper are listed in Table 1.

III. PROBLEM DESCRIPTION

In this section, the optimization of the multi-station high-speed railway hub is defined as a TRP under the capacity of railway infrastructures, and the detailed train routes include the macroscopic routes between stations and the microscopic track allocation inside stations and EMUs depots. Based on the above modeling ideas and actual railway structures, we construct a railway hub network $G = (V, E)$ to represent the multi-station high-speed railway hub, in which V refers to the set of nodes and E refers to the set of arcs. The layout of the railway hub network is shown in Figure 1 and the details are presented as follows:

A. SET OF NODES

The set of nodes $V = A \cup B$, includes: (a) set of stations A , indexed by a , representing the set of multiple passenger stations in the hub; (b) set of connecting directions B , indexed by b , representing the set of directions of the high-speed railway lines. Notably, each station has its corresponding EMUs depot serving for the trains arriving at the station. Thus, in order to reduce the scale, we do not set up independent nodes to express the depots, but regard them as a part of the stations.

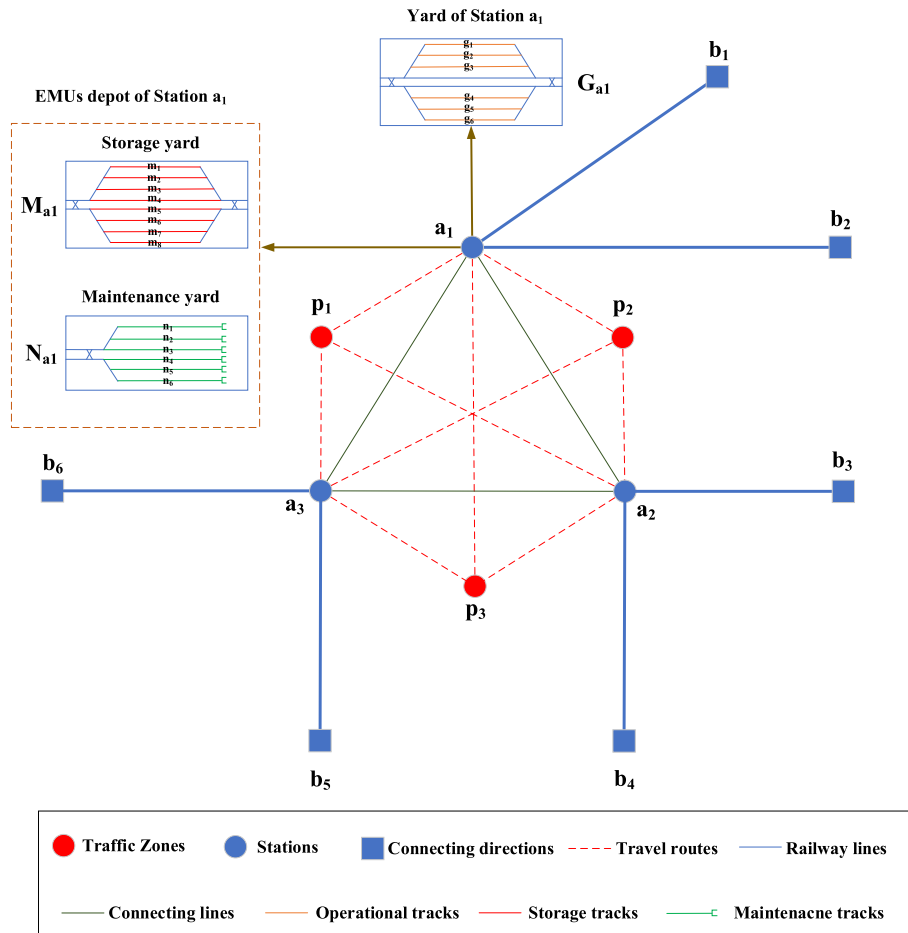


FIGURE 1. Illustration of the railway hub network.

B. SET OF ARCS

The set of arcs L refers to the double-track high-speed railway lines, indexed by (i, j) , includes: (a) the arcs between connecting directions and stations; (b) the arcs between stations.

C. SET OF TRACKS

The set of operational tracks G refers to the operational tracks for receiving and departing trains in the station, indexed by g , and $G = \sum G_a$, G_a represents the operational tracks of the station a .

Furthermore, as mentioned above, the track of the EMUs depot is also regarded as a part of the station. The set of storage tracks $M = \sum M_a$ indicates the storage tracks for the train in the depot of stations, indexed by m , and the set of maintenance tracks $N = \sum N_a$ indicates the maintenance tracks for the train maintenance in the depot, indexed by n .

D. SET OF TRAFFIC ZONES

The set of traffic zones P indicates the passenger aggregation nodes in different regions, indexed by p . These traffic zones can qualify the passenger demand of certain areas and are usually used in the urban transportation researches[50], [51].

E. SET OF TRAINS

The set of trains K refers to all EMUs operated at the stations of the railway hub, indexed by k , includes: (a) set of departure trains K_1 , representing the trains departing from the stations; (b) set of arrival trains K_2 , representing the trains arriving at the stations; (c) set of passing trains K_3 , representing the trains from the origin to the destination and passing through the stations.

Notably, the departure and passing trains $k \in K_1 \cup K_3$ just require the basic operations at the operational tracks $g \in G$. But the arrival trains $k \in K_2$ will be stored at the storage tracks $m \in M$ in the depot, and part of them will also be maintained at the maintenance tracks $n \in N$.

Based on the flexible scheme, all trains can be operated at any station of the railway hub, so we set up virtual origins and destinations, which can be shown in Figure 2. For the departure trains, by taking train k_1 as an example, since it may depart from any station, the origin of train k_1 is represented as a virtual node o_{k1} , which connects all possible stations in the hub, and the destination is connecting direction b_1 . Similarly, the destination of the arrival train k_2 is represented as a virtual node d_{k2} and the origin is connecting direction b_2 .

As for the passing train k_3 , its origin is connecting direction b_3 and the destination is connecting direction b_4 . Besides, to distinguish from the railway lines, the set of virtual arcs connecting virtual nodes with stations is represented as L_v .

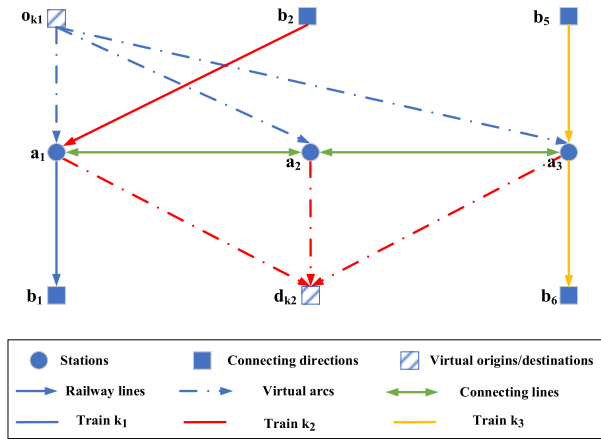


FIGURE 2. Illustration of virtual origins and destinations.

F. CAPACITY OF ARCS AND TRACKS

The capacity of arc refers to the maximum number of trains passing through a railway line within a certain time period. Specially, the capacity of virtual arcs is set to 1. The capacity of track refers to the maximum number of trains being operated at the track in a time period, includes the operation capacity of operational track, the storage capacity of storage track and the maintenance capacity of maintenance track.

Based on the high-speed railway hub network $G = (V, E)$, the TRP in the multi-station railway hub can be formulated as a multi-commodity flow problem. Given: (a) the origin o_k and the destination d_k of each train k , $k \in K$; (b) the length l_{ij} and capacity c_{ij} of each arc (i, j) , $(i, j) \in L \cup L_v$; (c) the operation capacity e_1^g of each operational track g , $g \in G$, the storage capacity e_2^m of each storage track m , $m \in M$ and the maintenance capacity e_3^n of each maintenance track n , $n \in N$; (d) the number of passengers q_{pb} from traffic zone p to direction b , $p \in P, b \in B$; (e) the distance d_{pa} from traffic zone p to station a , $p \in P, a \in A$. Considering all kinds of capacity, we aim to minimize the total cost of all trains and passengers, which consists of the train running cost on railway lines, train operation cost on operational tracks, train storage cost on storage tracks, train maintenance cost on maintenance tracks and passenger travel cost from traffic zones to stations. The related notations are listed in Table 2.

IV. MODEL FORMATION

Based on the multi-station high-speed railway hub network, we propose a mixed integer linear programming model by using the multi-commodity flow modelling method, which aims to minimize the total cost of trains and passengers.

TABLE 2. Notations.

Symbol	Definition
Sets	
A	Set of stations, indexed by a
B	Set of connecting directions, indexed by b
V	Set of nodes, $V = A \cup B$
L	Set of arcs, indexed by (i, j)
L_v	Set of virtual arcs, indexed by (i, j)
G	Set of operational tracks, indexed by g
G_a	Set of operational tracks in station a , $a \in A$
M	Set of storage tracks, indexed by m
M_a	Set of storage tracks in the depot of station a , $a \in A$
N	Set of maintenance tracks, indexed by n
N_a	Set of maintenance tracks in the depot of station a , $a \in A$
P	Set of traffic zones, indexed by p
K	Set of trains, indexed by k
K_1	Set of departure trains
K_2	Set of arrival trains
K_3	Set of passing trains
Parameters	
o_k	Virtual origin of train k , $k \in K$
d_k	Virtual destination of train k , $k \in K$
σ_k	Equals 1 if train k needs maintenance, and 0 otherwise, $k \in K$
c_{ij}	Capacity of arc (i, j) , $(i, j) \in L$, counted by trains/day
l_{ij}	Length of arc (i, j) , $(i, j) \in L$, counted by km
w_1	Train running cost on the railway lines, counted by kRMB/km
e_1^g	Capacity of operational track g , $g \in G$, counted by trains/day
w_2^g	Operation cost on the operational track g , $g \in G$, counted by kRMB
e_2^m	Capacity of storage track m , $m \in M$, counted by trains/day
w_3^m	Storage cost on the storage track m , $m \in M$, counted by kRMB
e_3^n	Capacity of maintenance track n , $n \in N$, counted by trains/day
w_4^n	Maintenance cost on the maintenance track n , $n \in N$, counted by kRMB
q_{pb}	Number of passengers from traffic zone p to direction b , $p \in P, b \in B$, counted by persons/day
d_{pa}	Distance from traffic zone p to station a , $p \in P, a \in A$, counted by km
w_5	Passenger travel cost from traffic zone to station, counted by kRMB/km
h	Passenger capacity of single train, counted by persons/train

A. ASSUMPTIONS

Without loss of generality, some principal assumptions are set up for the formulation:

1) Some information of the trains is pre-determined. Specifically, for the departure trains, the destinations are given and the origins are flexible. For the arrival trains, the origins are given and the destinations are flexible. For the passing trains, both the origins and the destinations are given.

2) The non-stopping trains are not considered, which indicates the trains do not stop at any station and pass through the hub directly. This is because these trains are not involved in

the track allocation and their occupation of railway lines is relatively small.

3) All the trains only stop at any station in the railway hub once. Besides, all the arrival trains will be stored and part of them need maintenance.

4) The cost of tracks is diversified, which is mainly based on the positions from the entrance of the yard. The cost rises as the distance from the entrance increases.

5) The route of the trains entering and leaving the EMUs depot is not included, which is relatively short for the entire route and can be ignored.

6) Only the passengers departing from the traffic zones of the hub are considered, while the passengers arriving at or passing through the hub are not included.

7) The building scale of each station is enough to accommodate passengers boarding at this station.

8) The capacity of other infrastructures in the high-speed railway hub is sufficient.

B. DECISION VARIABLES

In our formulation, we optimize the train and passenger routes simultaneously. The train routes include the macroscopic routes between stations and the microscopic track allocation inside stations and depots, while the passenger routes refer to the travel routes from the origin traffic zones to the boarding stations. Thus, we define five types of decision variables to represent the detailed routes and they are listed in Table 3.

TABLE 3. Decision variables.

Symbol	Definition
x_{ij}^k	Binary, $\{0,1\}$, equals 1 if train k occupies arc (i, j) , and 0 otherwise, $\forall k \in K, \forall (i, j) \in L \cup L_v$
y_g^k	Binary, $\{0,1\}$, equals 1 if train k stops at operational track g , and 0 otherwise, $\forall k \in K, \forall g \in G$
z_m^k	Binary, $\{0,1\}$, equals 1 if train k is stored at storage track m , and 0 otherwise, $\forall k \in K, \forall m \in M$
u_n^k	Binary, $\{0,1\}$, equals 1 if train k is maintained at maintenance track n , and 0 otherwise, $\forall k \in K, \forall n \in N$
r_{pab}	Integer, representing the number of passengers from traffic zone p to connecting direction b and boarding at station a , $\forall p \in P, \forall a \in A, \forall b \in B$

C. FORMULATION

By using the multi-commodity flow modelling method, a mixed integer linear programming model M1 is proposed to minimize the total cost of trains and passengers, the objective function is represented as follows:

$$\begin{aligned}
 \text{(M1) min } F &= \sum_{k \in K} \sum_{(i,j) \in L} x_{ij}^k l_{ij} w_1 + \sum_{k \in K} \sum_{g \in G} y_g^k w_2^g + \sum_{k \in K} \sum_{m \in M} z_m^k w_3^m \\
 &+ \sum_{k \in K} \sum_{n \in N} u_n^k w_4^n + \sum_{p \in P} \sum_{a \in A} \sum_{b \in B} r_{pab} d_{pa} w_5 \quad (1)
 \end{aligned}$$

In (1), the objective function includes the train cost and passenger cost. The train cost indicates the sum of train running cost, operation cost, storage cost and maintenance cost. The train running cost equals the length of train routes multiplied by train running cost per kilometer w_1 , and the other cost refers to the specific cost on each kind of track, which can be represented by w_2^g, w_3^m and w_4^n . The passenger cost equals the length of travel routes multiplied by passenger travel cost per kilometer w_5 .

Then, the constraints of the model are represented as follows:

$$\begin{aligned}
 &\sum_{j \in V, j \neq i} x_{ij}^k - \sum_{j \in V, j \neq i} x_{ji}^k \\
 &= \begin{cases} 1, & i = o_k \\ 0, & i \neq o_k \text{ and } i \neq d_k, \forall i \in V, \forall k \in K \\ -1, & i = d_k \end{cases} \quad (2)
 \end{aligned}$$

Constraint (2) is the flow balance constraint for nodes in the railway network, which is the basic constraint of the multi-commodity flow model.

$$\sum_{j \in V} x_{ij}^k \leq 1, \forall i \in V, k \in K \quad (3)$$

$$\sum_{j \in V} x_{ji}^k \leq 1, \forall i \in V, k \in K \quad (4)$$

Constraints (3)-(4) ensure there is no circle in each train route.

$$\sum_{g \in G} y_g^k = 1, \forall k \in K \quad (5)$$

Constraint (5) ensures each train can only stop once at any operational track of any station.

$$\sum_{m \in M} z_m^k = 1, \forall k \in K \quad (6)$$

$$\sum_{n \in N} u_n^k = \sigma_k, \forall k \in K \quad (7)$$

Constraint (6) ensures each arrival train can only be stored once at any storage track of any depot. Similarly, constraint (7) indicates the arrival train can only be maintained once at maintenance track if it needs maintenance.

$$\sum_{j \in V} x_{ij}^k + \sum_{j \in V} x_{ji}^k \geq \sum_{g \in G_i} y_g^k, \forall i \in A, k \in K \quad (8)$$

$$\sum_{j \in V} x_{ij}^k + \sum_{j \in V} x_{ji}^k \geq \sum_{m \in M_i} z_m^k, \forall i \in A, k \in K \quad (9)$$

$$\sum_{j \in V} x_{ij}^k + \sum_{j \in V} x_{ji}^k \geq \sum_{n \in N_i} u_n^k, \forall i \in A, k \in K \quad (10)$$

Constraints (8)-(10) indicate the station where the train is operated, stored and maintained must be on the train route.

$$\sum_{g \in G_a} y_g^k = \sum_{m \in M_a} z_m^k, \forall a \in A, k \in K \quad (11)$$

$$\sum_{m \in M_a} z_m^k \geq \sum_{n \in N_a} u_n^k, \forall a \in A, k \in K \quad (12)$$

Constraints (11)-(12) ensure the stations where each train are operated, stored or maintained must be same.

$$\sum_{k \in K} x_{ij}^k \leq c_{ij}, \forall (i, j) \in L \cup L_v \tag{13}$$

Constraint (13) is the arc capacity constraint for the railway lines, which ensures the number of trains passing through arc (i, j) is less than the capacity of arc (i, j) .

$$\sum_{k \in K} y_g^k \leq e_1^g, \forall g \in G \tag{14}$$

$$\sum_{k \in K} z_m^k \leq e_2^m, \forall m \in M \tag{15}$$

$$\sum_{k \in K} u_n^k \leq e_3^n, \forall n \in N \tag{16}$$

Constraints (14)-(16) are the track capacity constraints. Constraint (14) refers to the operation capacity constraint, which ensures the number of trains operated at operational track g is less than the capacity of track g . Similarly, constraint (15) refers to the storage capacity and constraint (16) refers to the maintenance capacity.

$$\sum_{a \in A} r_{pab} = q_{pb}, \forall p \in P, b \in B \tag{17}$$

Constraint (17) indicates the distribution of the passengers, which ensures the number of passengers boarding at all stations is equal to the passenger demand from traffic zone p to connecting direction b .

$$\sum_{p \in P} r_{pab} \leq \sum_{k \in K_1 \cup K_3} \sum_{g \in G_a} y_g^k h, \forall a \in A, b \in B \tag{18}$$

Constraint (18) is the coupling constraint for train variables and passenger variables, which indicates the passenger carrying capacity of departure and passing trains stopping at station a is larger than the number of passengers boarding at station a .

$$x_{ij}^k, y_g^k, z_m^k, u_n^k \in \{0, 1\}, r_{pab} \in N^+ \tag{19}$$

Constraint (19) indicates the domains of the decision variables.

V. SOLUTION METHOD

In previous researches, Lagrangian relaxation (LR) has been widely used in the railway optimization and proved to be efficient for the TRP. The basic principle of LR is to dualize the hard constraints of the model and move them into the objective function by multiplying non-negative Lagrange multipliers, then the dualized model can be decomposed into a series of independent subproblems which can be solved efficiently [4], [12], [33], [34].

According to the characteristics of the proposed model, we attempt to develop a LR approach to solve the model M1. In the approach, we first dualize the hard constraints of model M1 and penalize them into the objective function. Then, we further decompose the dualized model into two series of train and passenger subproblems and apply the shortest path

algorithm to solving these subproblems. Notably, sometimes LR may fall into the symmetry issues [4], [8], which will limit the algorithm into a loop and affect the improvement of the solution quality, so we try to diversify the subproblems by updating the cost of arcs and tracks.

A. THE SOLUTION FRAMEWORK OF LR

After analyzing the constraints of model M1, it can be conducted the capacity constraints (13)-(16) lead to a large number of train combinations, and the coupling constraint (18) integrates the train and passenger variables. Therefore, we regard these constraints as the hard constraints and dualize them into the objective function with five sets of non-negative Lagrange multipliers $\lambda_1^{ij}, \lambda_2^g, \lambda_3^m, \lambda_4^n, \lambda_5^{ab}$. The dualized model M2 is represented as follows:

$$\begin{aligned} \min F_{LR} &= \sum_{k \in K} \sum_{(i,j) \in L_1} x_{ij}^k l_{ij} w_1 + \sum_{k \in K} \sum_{g \in G} y_g^k w_2^g + \sum_{k \in K} \sum_{m \in M} y_m^k w_3^m \\ (M2) &+ \sum_{k \in K} \sum_{n \in N} y_n^k w_4^n + \sum_{p \in P} \sum_{a \in A} \sum_{b \in B} r_{pab} d_{pa} w_5 \\ &+ \sum_{(i,j) \in L_1} \lambda_1^{ij} (\sum_{k \in K} x_{ij}^k - c_{ij}) \\ &+ \sum_{g \in G} \lambda_2^g (\sum_{k \in K} y_g^k - e_1^g) + \sum_{m \in M} \lambda_3^m (\sum_{k \in K} z_m^k - e_2^m) \\ &+ \sum_{n \in N} \lambda_4^n (\sum_{k \in K} u_n^k - e_3^n) + \sum_{a \in A} \sum_{b \in B} \lambda_5^{ab} (\sum_{p \in P} r_{pab} \\ &- \sum_{k \in K} \sum_{g \in G_a} y_g^k h) \end{aligned} \tag{20}$$

Subject to constraints (2)-(12), (17), (19).

For the sake of simplification, the objective function of the dualized model M2 can be rewritten as follows:

$$\begin{aligned} \min F_{LR} &= [\sum_{k \in K} \sum_{(i,j) \in L_1} x_{ij}^k (l_{ij} w_1 + \lambda_1^{ij}) + \sum_{k \in K} \sum_{g \in G} y_g^k (w_2^g + \lambda_2^g) \\ &+ \sum_{k \in K} \sum_{m \in M} y_m^k (w_3^m + \lambda_3^m) + \sum_{k \in K} \sum_{n \in N} y_n^k (w_4^n + \lambda_4^n)] \\ &- \sum_{a \in A} \sum_{k \in K} \sum_{g \in G_a} y_g^k h \lambda_5^{ad_k} + \sum_{p \in P} \sum_{a \in A} \sum_{b \in B} r_{pab} (d_{pa} w_5 + \lambda_5^{ab}) \\ &- [\sum_{(i,j) \in L_1} \lambda_1^{ij} c_{ij} + \sum_{g \in G} \lambda_2^g e_1^g + \sum_{m \in M} \lambda_3^m e_2^m + \sum_{n \in N} \lambda_4^n e_3^n] \end{aligned} \tag{21}$$

In (21), we introduce the Lagrange multipliers into the cost of arcs and tracks. Specifically, λ_1^{ij} can be interpreted as the extra penalty cost of occupying arc (i, j) . $\lambda_2^g, \lambda_3^m, \lambda_4^n$ refer to the extra penalty cost of occupying operational track g , storage track m and maintenance track n . As the multiplier of the coupling constraint between train and passenger variables, λ_5^{ab} has an extra impact on both train and passenger cost.

Since the constant term of (21) is unconcerned with the decision variables, it can be ignored and the dualized model M2 can be decomposed into two series of independent train and passenger subproblems, which are represented as model M3 and M4 as follows:

$$\begin{aligned}
 & \min LR_k \\
 & = \sum_{(i,j) \in L_1} x_{ij}^k (l_{ij} w_1 + \lambda_1^{ij}) + \sum_{g \in G} y_g^k (w_2^g + \lambda_2^g) \\
 \text{(M3)} & + \sum_{m \in M} y_m^k (w_3^m + \lambda_3^m) + \sum_{n \in N} y_n^k (w_4^n + \lambda_4^n) \\
 & - \sum_{a \in A} \sum_{g \in G_a} y_g^k h \lambda_5^{ag} \quad (22)
 \end{aligned}$$

Subject to constraints (2)-(12), (19).

$$\text{(M4)} \min LR_p = \sum_{p \in P} \sum_{a \in A} \sum_{b \in B} r_{pab} (d_{pa} w_5 + \lambda_5^{ab}) \quad (23)$$

Subject to constraints (17), (19).

The train subproblem M3 aims to minimize the total cost of arcs and tracks for each independent train, which is a shortest path problem with specific constraints and can be solved by a shortest path algorithm based on the Floyd algorithm. At the meanwhile, the passenger subproblem M4 aims to minimize the total cost of passengers. After solving all the train and passenger subproblems, we substitute the train and passenger solutions into (20) to generate the lower bound of model M1.

As for the upper bound, we schedule the train and passenger routes according to the train profits by using a priority scheduling algorithm.

In the iterative part of the algorithm, we use the most common sub-gradient method to update the Lagrange multipliers iteratively, while the step size α is simultaneously updated with iteration. Denote q as the iteration number and the related formulas are expressed as (24)-(29):

$$\lambda_1^{ij}(q+1) = \max\{0, \lambda_1^{ij}(q) + \alpha(\sum_{k \in K} x_{ij}^k - c_{ij})\}, \quad \forall (i,j) \in L \cup L_v \quad (24)$$

$$\lambda_2^g(q+1) = \max\{0, \lambda_2^g(q) + \alpha(\sum_{k \in K} y_g^k - e_1^g)\}, \quad \forall g \in G \quad (25)$$

$$\lambda_3^m(q+1) = \max\{0, \lambda_3^m(q) + \alpha(\sum_{k \in K} z_m^k - e_2^m)\}, \quad \forall m \in M \quad (26)$$

$$\lambda_4^n(q+1) = \max\{0, \lambda_4^n(q) + \alpha(\sum_{k \in K} u_n^k - e_3^n)\}, \quad \forall n \in N \quad (27)$$

$$\begin{aligned}
 \lambda_5^{ab}(q+1) & = \max\{0, \lambda_5^{ab}(q) + \alpha(\sum_{p \in P} r_{pab} \\
 & - \sum_{k \in K_1 \cup K_3} \sum_{g \in G_a} y_g^k h)\} \quad \forall a \in A, b \in B \quad (28)
 \end{aligned}$$

$$\alpha = 1/(q+1) \quad (29)$$

The termination conditions of the algorithm include the following three rules: (1) the gap between best upper bound and best lower bound is within an acceptable range; (2) the values of best upper bound and best lower bound remain unchanged in consecutive iterations; (3) the value of iteration q exceeds the maximum q_{\max} .

Above all, the solution framework of LR is summarized in Algorithm 1 and the flowchart is shown in Figure 3.

Algorithm 1 The Solution Framework of LR

Step 1: Initialization

Initialize the maximum number of iterations $q_{\max} = 100$ and the initial iteration number $q = 1$;

Lagrange multipliers $\lambda_1^{ij} = 0, \lambda_2^g = 0, \lambda_3^m = 0, \lambda_4^n = 0, \lambda_5^{ab} = 0$;

Define current upper bound $UB(q)$ and lower bound $LB(q)$, best upper bound UB^* and lower bound LB^* .

Step 2: Solve the dualized model M2 and update the best lower bound

Step 2.1: Decompose the dualized model M2 into two series of subproblems, including train subproblems M3 and passenger subproblems M4.

Step 2.2: Apply the shortest path algorithm to the subproblems to generate train routes and passenger routes, then calculate the current lower bound $LB(q)$ by (20).

Step 2.3: Update the best lower bound $LB^* = \min\{LB(q), LB^*\}$.

Step 3: Find the feasible solution and update the best upper bound

Step 3.1: If the dual solution is feasible, compute the objective value of model M1 by substituting the dual solution into (1) and regard it as the current upper bound $UB(q)$. Otherwise, move to Step 3.2.

Step 3.2: If the dual solution is infeasible, apply a priority scheduling algorithm to the dualized model M2 and obtain the current upper bound $UB(q)$.

Step 3.3: Update the best upper bound $UB^* = \min\{UB(q), UB^*\}$.

Step 4: Update the Lagrange multipliers

Update the iteration $q = q + 1$;

Using the sub-gradient method to update the multipliers by (24)-(28);

Update the step size α by (29).

Step 5: Termination conditions test

As long as one of the three termination rules is satisfied, stop the algorithm and output the best upper bound UB^* , the best lower bound LB^* , the current gap and best feasible solution; Otherwise, loop back to Step 2.

B. THE SHORTEST PATH ALGORITHM

The subproblem M3 of train k can be regarded as a shortest path problem from the origin o_k to the destination d_k with operation, storage and maintenance constraints. Since there are multiple pairs of origins and destinations in the train subproblems, we choose the Floyd algorithm rather than the Dijkstra algorithm to calculate the train shortest routes[12]. Then, as for the constraints of tracks, we add the minimum track cost of each station into the cost of shortest train routes passing through the station, and select the train route with the least cost of arcs and tracks.

Furthermore, since the cost of arcs and tracks is the same for all trains in each iteration, some trains will concentrate on the same route, which is also called the symmetry issues of LR [4], [8]. The symmetry issues may make LR fall into local optimal solution and is not conducive to improve the solution quality. Therefore, we try to break the issues by diversifying the cost. Specifically, we update the occupation of arcs and tracks after solving each train subproblem, once an arc or

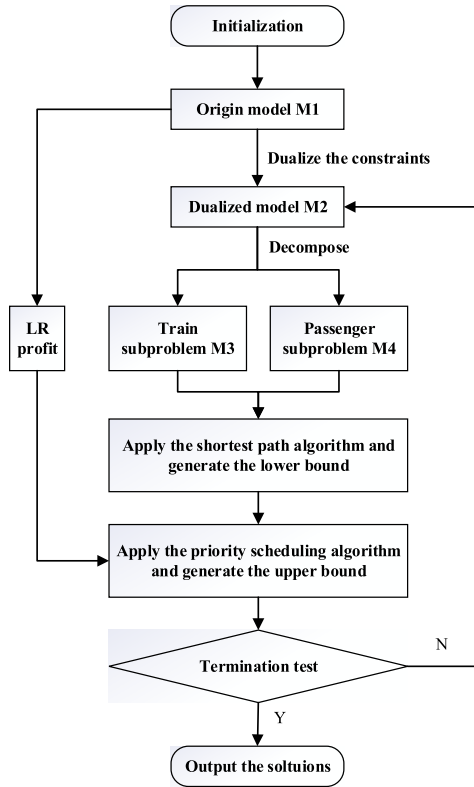


FIGURE 3. The flowchart of LR.

track is full of capacity, a penalty cost δ will be added into the cost of the arc or track to avoid more occupation of other trains.

As for the passenger subproblem M4, we compare the travel cost from the origin traffic zone to different stations and select the least cost travel routes.

The solution framework of the shortest path algorithm is summarized in Algorithm 2 as follows:

C. THE PRIORITY SCHEDULING ALGORITHM

As for the feasible solutions, referring to the methods of previous researches [12], [29], we also develop a heuristic algorithm based on the priority of trains. In the priority scheduling algorithm, we first sort the trains according to the Lagrangian profit in (30), which represents the violation of dualized constraints. Afterwards, we schedule the trains successively, once an arc or a track is full of capacity, it cannot be occupied by the following trains. Finally, we compute the passenger carrying capacity of each station and generate the optimized passenger routes.

$$\begin{aligned}
 LR_k^{profit} = & \sum_{(i,j) \in L_1} x_{ij}^k (\sum_{k \in K} x_{ij}^k - c_{ij}) + \sum_{g \in G} y_g^k (\sum_{k \in K} y_g^k - e_1^g) \\
 & + \sum_{m \in M} z_m^k (\sum_{k \in K} z_m^k - e_2^m) + \sum_{n \in N} u_n^k (\sum_{k \in K} u_n^k - e_3^n) \\
 & + \sum_{a \in A} (\sum_{p \in P} r_{pab} - \sum_{k \in K_1 \cup K_3} \sum_{g \in G_a} y_g^k h) \sum_{g \in G_a} y_g^k
 \end{aligned} \tag{30}$$

Algorithm 2 The Solution Framework of the Shortest Path Algorithm

Step 1: Apply the Floyd algorithm to generate the shortest routes between nodes

Step 1.1: Construct the adjacency matrix between any two nodes i, j in V . According to (22), set $cost(i, j)$, $cost(g)$, $cost(m)$ and $cost(n)$ to indicate the cost of arcs and tracks.

$$cost(i, j) = l_{ij}w_1 + \lambda_1^{ij}, cost(g) = w_2^g + \lambda_2^g, cost(m) = w_3^g + \lambda_3^g, cost(n) = w_4^g + \lambda_4^g$$

Set $dis(i, j)$ to represent the cost of the shortest route from i to j and $pre(i, j)$ to represent the previous node of (i, j) .

For $i \in V, j \in V$

$$pre(i, j) = -1$$

If $(i, j) \in L \cup L_v$ **Then**

$$dis(i, j) = cost(i, j)$$

Else

$$dis(i, j) = M, M \text{ indicates a large positive number}$$

End if

End for

Step 1.2: Update the shortest routes and previous nodes.

For $u \in V, i \in V, j \in V$

If $dis(i, j) > dis(i, u) + dis(u, j)$ **Then**

$$dis(i, j) = dis(i, u) + dis(u, j)$$

$$pre(i, j) = u$$

End if

End for

Step 1.3: Generate the occupied nodes of shortest routes, $route(i, j)$ indicates the set of occupied nodes in the shortest route from i to j .

For $i \in V, j \in V$

If $pre(i, j) \neq -1$ **Then**

If $pre(i, pre(i, j)) = -1$ **Then**

Add $pre(i, j)$ into $route(i, j)$

If $pre(pre(i, j), j) = -1$ **Then**

Add j to $route(i, j)$

Else

Add i to $route(i, j)$

End if

End for

Step 2: Find the train shortest routes from the origins to the destinations

Step 2.1: Generate the shortest routes from the origin to all the stations and from all the stations to the destination by using Floyd algorithm, the lengths of these routes can be represented as $dis(o_k, a)$ and $dis(a, d_k)$.

Step 2.2: Generate the train shortest routes, l_k refers to the length of the shortest route for train k and $train_route(k)$ refers to the set of occupied nodes. $cost_g(a)$ represents the minimum operation cost in the tracks G_a of station a , while $cost_m(a)$ indicates the minimum storage cost and $cost_n(a)$ indicates the minimum maintenance cost.

Initialize $l_k, cost_g(a), cost_m(a), cost_n(a) = M, M$ indicates a large positive number

For $a \in A$

$$cost_g(a) = \min\{e_1^g | g \in G_a\}$$

$$cost_m(a) = \min\{e_2^m | m \in M_a\}$$

$$cost_n(a) = \min\{e_3^n | n \in N_a\}$$

End for

For $k \in K$

If train k is a departure train or a passing train, $k \in K_1 \cup K_3$ **Then**

$$l_k = \min\{dis(o_k, a) + cost_g(a) + dis(a, d_k) | a \in A\},$$

the selected station is represented as a^* and the selected operational track is represented as g^* ;

$$train_route(k) = \{route(o_k, a^*), route(a^*, d_k), g^*\}$$

(Continued.) **Algorithm 2** The Solution Framework of the Shortest Path Algorithm

If train k is an arrival train and do not need maintenance, $k \in K_2$ and $\sigma_k = 0$ Then
 $l_k = \min\{dis(o_k, a) + cost_g(a) + cost_m(a) + dis(a, d_k) | a \in A\}$, the selected storage track is represented as m^* ;
 $train_route(k) = \{route(o_k, a^*), route(a^*, d_k), g^*, m^*\}$
 Else
 $l_k = \min\{dis(o_k, a) + cost_g(a) + cost_m(a) + cost_n(a) + dis(a, d_k) | a \in A\}$, the selected maintenance track is represented as n^* ;
 $train_route(k) = \{route(o_k, a^*), route(a^*, d_k), g^*, m^*, n^*\}$
 End if
End for
Step 2.3: Update the cost of arcs and tracks
 After solving each train subproblem, update the occupation of arcs $cap(i, j)$ and tracks $cap(g), cap(m), cap(n)$ based on the occupied arcs and tracks in $train_route(k)$. Then, according to the current occupation, update the cost of arcs and tracks.
For $(i, j) \in L \cup L_v, g \in G, m \in M, n \in N$
 If $cap(i, j) > c_{ij}$ Then
 $cost(i, j) = l_{ij}w_1 + \lambda_1^{ij} + \delta$
 If $cap(g) > e_1^g$ Then
 $cost(g) = w_2^g + \lambda_2^g + \delta$
 If $cap(m) > e_2^m$ Then
 $cost(m) = w_3^m + \lambda_3^m + \delta$
 If $cap(n) > e_3^n$ Then
 $cost(n) = w_4^n + \lambda_4^n + \delta$
 End if
End for
 If any cost of arcs and tracks is updated, call Floyd algorithm again and update the shortest routes $dis(i, j)$ and $route(i, j)$.
Step 3: Find the passenger shortest routes from the origin traffic zones to the stations
 According to (22), set $cost(p, b, a)$ to indicate the cost of travel routes from the traffic zone p to the connecting direction b boarding at station a , which can be represented as $cost(p, a) = d_{pa}w_5 + \lambda_5^{ab}$. l_{pb} refers to the length of the shortest route for the passengers from the traffic zone p to the connecting direction b , $a^*(p, b)$ refers to the boarding stations.
 Initialize $l_{pb} = M, passenger_route(p, b) = -1$, M indicates a large positive number
For $a \in A$
 If $cost(p, b, a) < l_{pb}$ Then
 $l_{pb} = cost(p, b, a)$ and $a^*(p, b) = a$
 End if
End for

The solution framework of the priority scheduling algorithm is summarized in Algorithm 3 as follows:

VI. NUMERICAL EXPERIMENTS

In this section, we design two numerical experiments to demonstrate the effectiveness of the model and algorithm. Firstly, an illustrative railway hub network is constructed and we compare the solutions obtained by the LR approach with the optimized results obtained by the commercial solver GUROBI. Secondly, we further apply the model and algorithm to a real-world network based on Zhengzhou high-speed railway hub in China, then evaluate the validity of the optimized scheme.

Algorithm 3 The Solution Framework of the Priority Scheduling Algorithm

Step 1: Train priority ranking
 Rank the trains according to the decrease of the train Lagrangian profits, which refers to the violation of dualized constraints (13-16) and (18).
Step 2: Generate the feasible train routes
Step 2.1: Choose the train k with the highest priority and obtain its route through using the shortest path algorithm (Step 2 of Algorithm 2). If there is no feasible route for train k , terminate the algorithm and output no feasible solutions.
Step 2.2: Update the occupation of arcs and tracks, once an arc or a track is full of capacity, it cannot be occupied by the following trains.
Step 2.3: If all trains have been scheduled, calculate the current passenger carrying capacity of each station for each connecting direction. Otherwise, move to Step 2.1.
Step 3: Generate the feasible passenger routes
 Considering the passenger carrying capacity of stations, assign passengers to the nearest station.

A. AN ILLUSTRATIVE RAILWAY HUB NETWORK

1) DESCRIPTION OF THE ILLUSTRATIVE NETWORK AND RELATED PARAMETERS

The illustrative multi-station high-speed railway hub network is shown in Figure 4, which consists of two passenger stations with the corresponding EMUs depots, four connecting directions and two traffic zones. Specifically, each station has two operational tracks, two storage tracks and two maintenance tracks, the capacity and cost of tracks is listed in Table 4. In Table 5, the capacity and lengths of arcs between nodes are listed. Except the double-track railway lines, the travel routes between traffic zones and stations are also included. The numbers of passengers from the origin traffic zones to the target directions are listed in Table 6. In order to make the effect of train cost and passenger cost on the objective function in an order of magnitude, the values of other basic parameters are set as follows: per train running cost on the railway lines w_1 is set to 1 RMB/(train-km), per passenger travel cost w_5 is set to 25 RMB/km and passenger carrying capacity per train is set to 100 persons/train.

2) RESULTS OF THE LR APPROACH

In order to demonstrate the effectiveness of the LR approach, we design four cases with different number of trains based on the illustrative network, the details of these cases are listed in Table 7. Furthermore, for testing the solution quality of the results, we regard results obtained by a commercial solver GUROBI 8.1.1 as benchmarks, which has been proved to be efficient in many optimization problems [17], [20].

The solution results of the LR approach are summarized in Table 8, where the best lower bounds (LB), the best upper bounds (UB), the gaps between the best upper and lower bounds (GAP1), the gaps between the best upper bounds and the optimal values (GAP2) and the computational times (CTs) are reported. The optimal values obtained by GUROBI (OP) are also included as benchmarks. Besides,

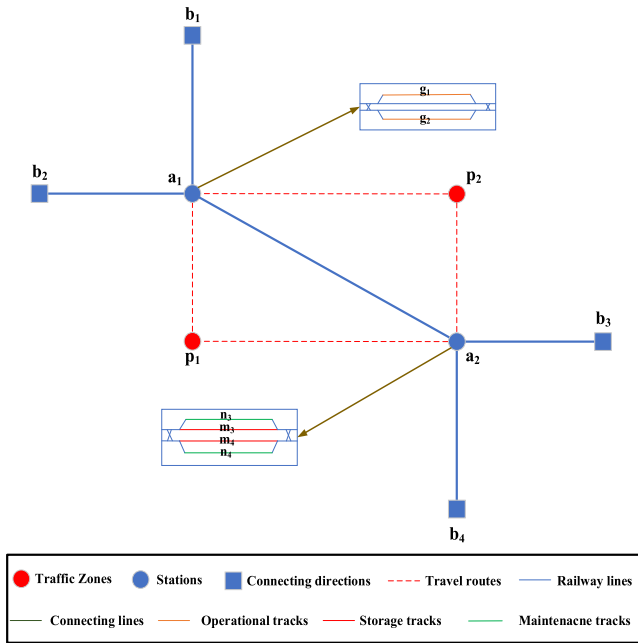


FIGURE 4. Illustrative multi-station high-speed railway hub network.

TABLE 4. The capacity and cost of tracks.

Track number	Station	Capacity (trains/day)	Cost (kRMB/train)	Type
g1	a1	10	10	G
g2	a1	10	5	G
g3	a2	10	10	G
g4	a2	10	5	G
m1	a1	6	10	M
m2	a1	6	5	M
m3	a2	6	10	M
m4	a2	6	5	M
n1	a1	4	10	N
n2	a1	4	5	N
n3	a2	4	10	N
n4	a2	4	5	N

Notes: G-operational track; M-storage track; N-maintenance track

TABLE 5. The capacity and lengths of arcs.

Origin	Destination	Capacity (trains/day)	Length (km)
b1	a1	80	5
b2	a1	80	5
b3	a2	80	5
b4	a2	80	5
a1	a2	80	5
a1	b1	80	5
a1	b2	80	5
a2	b3	80	5
a2	b4	80	5
a2	a1	80	5
p1	a1	-	0.05
p1	a2	-	0.1
p2	a1	-	0.1
p2	a2	-	0.05

in order to describe the iterative process of the algorithm, the convergence curves of LR are drawn in Figure 5. Both the LR approach and GUROBI 8.1.1 are implemented in

TABLE 6. The number of passengers.

Origin	Destination	Number (persons)
p1	b1	30
p1	b2	30
p1	b3	30
p1	b4	30
p2	b1	30
p2	b2	30
p2	b3	30
p2	b4	30

TABLE 7. The number of trains in different cases.

Case	Departure train	Arrival train	Passing train	Total
1	4	8 (4)	4	16
2	6	12 (6)	6	24
3	8	16 (8)	8	32
4	10	20 (10)	10	40

Notes: () -the number of arrival trains to be maintained. For example, 4 (2) indicates that there are 2 of 4 arrival trains to be maintained in the depots.

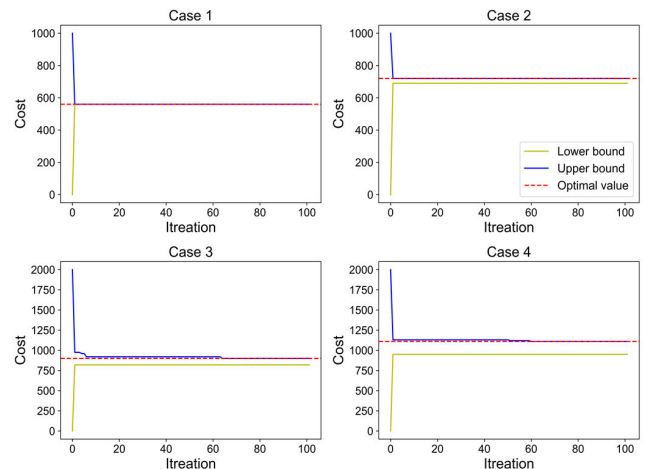


FIGURE 5. The convergence curves of LR in different cases.

TABLE 8. The solution results of LR in different cases.

Case	Number of trains	OP	LB	UB	GAP1	GAP2	CTs(s)
1	16	560	560	560	0%	0%	0.13
2	24	720	690	720	4.17%	0%	25.79
3	32	900	820	900	8.89%	0%	38.18
4	40	1100	950	1110	14.41%	0%	53.00

Python 3.6.3 on the PyCharm Community platform, and all the cases are carried out on a Windows 10 Home desktop computer with i7-7700@2.8Ghz CPU and 16GB RAM.

According to the results in Table 8 and Figure 5, the following conclusions can be drawn:

1) The LR approach can obtain solutions with good quality. In Table 8, it can be seen that the GAP2 of the four cases are all 0%, which also means the best feasible solutions generated by LR are the same to the optimal solutions generated by GUROBI. Furthermore, the average GAP1 of the four cases

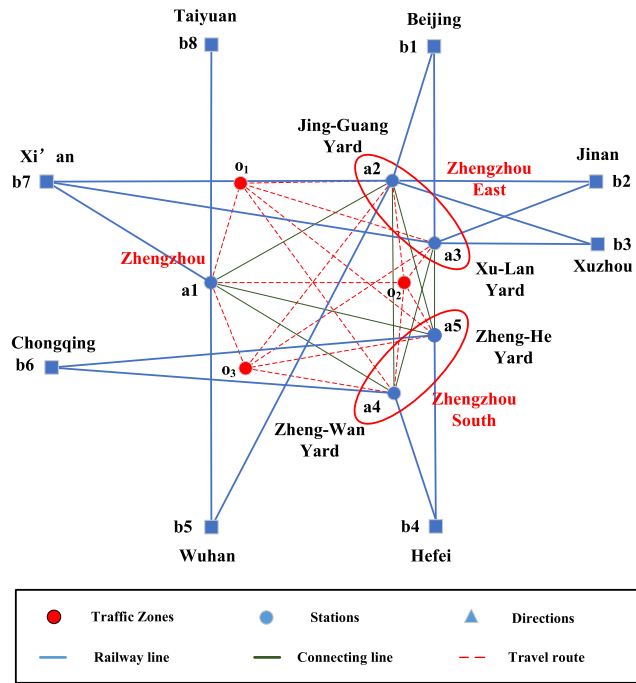


FIGURE 6. Illustration of the Zhengzhou high-speed railway hub network.

is 6.87%, the minimum is 0% and the maximum is 14.41%. These indicators are all in an acceptable range and prove the good convergence of the algorithm.

2) The LR approach has a good performance in solving efficiency. First, the computational times of LR in the four cases is all short, the minimum is just 0.13s and the maximum is 53.00s. Second, as shown in Figure 5, it can be seen that LR can obtain the best lower bounds quickly with several iterations. Besides, the numbers of iterations where the best upper bounds are generated are marked as red circles in Figure 5, the minimum is 1, the maximum is 64 and the average is 32, which indicate the LR can generate good feasible solutions in short iterations.

3) The LR approach has expansibility with the expansion of the problem scale. In the above cases, when the number of trains increases, the GAP1 becomes larger and the computational time is also longer, but the algorithm still can obtain optimal solutions within acceptable iterations, which proves the effectiveness for solving the large-scale problems.

B. A REAL-WORD NETWORK BASED ON THE ZHENGZHOU HIGH-SPEED RAILWAY HUB

1) DESCRIPTION OF THE ZHENGZHOU HIGH-SPEED RAILWAY HUB

The Zhengzhou railway hub is a typical multi-station high-speed railway hub, which has three passenger stations and eight high-speed railway lines. As the center of the Chinese high-speed railway network, there are various types of trains operated in the hub every day, and the passenger demand is growing with the development of the urbanization.

TABLE 9. The capacity and lengths of arcs.

O	D	Capacity (trains/day)	Length (km)	O	D	Capacity (trains/day)	Length (km)
b1	a2	263	179	a1	b8	287	111
a2	b1	263	179	a2	a4	331	36
b1	a3	263	179	a4	a2	331	36
a3	b1	263	179	a2	a5	331	36
b2	a2	263	237	a5	a2	331	36
a2	b2	263	237	a3	a4	331	36
b2	a3	263	237	a4	a3	331	36
a3	b2	263	237	a3	a5	331	36
b3	a2	164	250	a5	a3	331	36
a2	b3	164	250	a1	a4	331	43
b3	a3	164	250	a4	a1	331	43
a3	b3	164	250	a1	a5	331	43
b4	a4	287	212	a5	a1	331	43
a4	b4	287	212	a1	a2	331	15
b4	a5	287	212	a2	a1	331	15
a5	b4	287	212	p1	a1	-	5
b5	a1	150	102	p1	a2	-	10
a1	b5	150	102	p1	a3	-	10
b5	a2	263	91	p1	a4	-	15
a2	b5	263	91	p1	a5	-	15
b6	a4	273	282	p2	a1	-	15
a4	b6	273	282	p2	a2	-	5
b6	a5	273	282	p2	a3	-	5
a5	b6	273	282	p2	a4	-	10
b7	a1	150	331	p2	a5	-	10
a1	b7	150	331	p3	a1	-	10
b7	a2	120	201	p3	a2	-	15
a2	b7	120	201	p3	a3	-	15
b7	a3	221	201	p3	a4	-	5
a3	b7	221	201	p3	a5	-	5
b8	a1	287	111				

Therefore, how to make full use of the capacity of the hub and consider passenger demand into train plans becomes an urgent problem for the railway department. Under the above ground, we apply the proposed model and algorithm to the Zhengzhou railway hub to accomplish an optimization scheme Figure 6 illustrates the Zhengzhou high-speed railway network. There are Zhengzhou, Zhengzhou East and Zhengzhou South stations with corresponding EMUs depots in the hub. Notably, Zhengzhou East station has Jing-Guang yard and Xu-Lan yard, which usually operate trains independently, so we set up two nodes to represent them. Similarly, Zhengzhou South station has Zheng-He yard and Zheng-Wan yard.

Based on the technical data of China National Railway Group Zhengzhou Railway Bureau Co., Ltd., the related parameters are presented given. Table 9 shows the capacity and lengths of railway lines and travel routes.

Table 10 indicates the capacity and cost of tracks. Then, the numbers of trains between origins and destinations are shown in Table 11, the origins of the departure trains and the destinations of the arrival trains are represented as Z. The numbers of passengers from traffic zones to directions are shown in Table 12, and the values of other basic parameters are listed in Table 13.

2) OPTIMIZED RESULTS

After substituting the real data into the proposed model, we implement the LR approach to obtain solutions by

TABLE 10. The capacity and lengths of tracks.

Track	1	2	3	4	5	6	7	8	9	10	11	12	13
G	a1	36,6	38,6	40,8	45,8	40,10	38,10	38,12	38,12				
	a2	30,6	36,6	40,6	42,6	45,8	38,8	38,8	45,8	45,10	40,10	36,10	30,10
	a3	30,6	45,6	34,6	36,8	45,8	45,8	38,10	36,10	44,10	42,12	40,12	
	a4	30,6	36,6	38,6	45,6	44,8	40,8	36,10	38,10	34,12			
	a5	30,6	38,6	40,8	45,8	36,10	38,10	38,12	34,12				
M	a1	4,6	4,6	4,6	4,6	4,6	4,6						
	a2	4,6	4,6	4,6	4,8	4,8	4,8	4,8	4,8	4,8	4,10		
	a3	4,6	4,6	4,6	4,6	4,6	4,6	4,6	4,6	4,6	4,6	4,8	4,8
	a4	4,8	4,8	4,8	4,8	4,8	4,8	4,8	4,8	4,10			
	a5	4,6	4,6	4,6	4,6	4,6	4,6	4,6	4,8	4,8	4,8	4,8	4,10
N	a1	10,8	5,6										
	a2	25,6	25,6										
	a3	20,6	20,6										
	a4	15,6	15,6										
	a5	15,6	25,6										

Notes: “36, 6” refers to capacity is 36 trains/day and cost is 6 kRMB/train; G-operational track; M-storage track; N-maintenance track;

TABLE 11. The numbers of trains between origins and destinations (trains/day).

O \ D	D								Z
	b1	b2	b3	b4	b5	b6	b7	b8	
b1	0	0	8	4	64	20	20	0	19
b2	0	0	0	0	19	12	17	0	29
b3	0	0	0	0	20	13	36	4	24
b4	4	0	0	0	0	0	12	14	22
b5	64	19	20	0	0	0	15	10	25
b6	20	12	13	0	0	0	0	20	31
b7	20	17	36	12	15	0	0	0	26
b8	0	0	4	14	10	20	0	0	19
Z	19	29	24	22	25	31	26	19	

Notes: Z-any station in Zhengzhou railway hub.

TABLE 12. The numbers of passengers between origins and destinations (thousand persons/day).

O \ D	D								Total
	b1	b2	b3	b4	b5	b6	b7	b8	
p1	30	12	20	7	30	15	30	17	161
p2	30	12	20	7	30	15	30	17	161
p3	30	12	20	7	30	15	30	17	161
Total	90	36	60	21	90	45	90	51	483

Python 3.6.3 on PyCharm Community, which is carried out on a Windows 10 Home desktop computer with i7-7700@2.8Ghz CPU and 16GB RAM. The best upper bound is 8151020 kRMB and the best lower bound is 8089240 kRMB, the gap is 0.76%, which is within a reasonable range, so we can use the best feasible solution as the optimized results. Specifically, the objective value is 8151020 kRMB, in which the train running cost on railway lines is 3194320 kRMB, the train operation cost on tracks is 69980 kRMB, the train storage cost is 26720 kRMB, the train maintenance cost is 30000 kRMB and the passenger cost is 403000 kRMB.

Due to the large number of trains, it is troublesome to list all detailed train routes, so we take 74 trains of connecting direction b4 as examples and represent their routes in Table 14.

TABLE 13. The values of other basic parameters.

Parameter	Definition	Value
w_1	per train running cost on railway lines	10 kRMB/(train·km)
w_5	per passenger travel cost	2 RMB/(person·km)
h	per passenger capacity of the EMUs	1000 (persons/train)

As for the capacity utilization, the capacity utilization of arcs is shown in Figure 7, notably, since all the railway lines are double-track, we take the larger rates as the indicators of the coloring. Figure 8 shows the utilization of tracks of different stations, in which the purple bars represent the operational tracks, the pink bars refer to the storage tracks and the green bars refer to the maintenance tracks. Finally, the travel routes of passengers are presented in Table 15.

From the solution results in Table 13, 14 and Figure 7, 8, the following conclusions can be drawn:

1) The proposed model can effectively optimize the train cost. Specifically, in Table 14, almost all trains select the shortest routes. Then, from the track allocation of trains in Table 13, it can be seen the trains are operated, stored and maintained at the tracks with the minimum cost.

2) The utilization of arcs is relatively balanced in the whole network. In Figure 7, the average capacity utilization rate of arcs is 18.64%, and 91.30% of the arcs are with the rates less than 60%, which are all within a reasonable range and prove the arc capacity is sufficient. Moreover, the variance of the utilization rates is 0.0255, which indicates the gaps between utilization of different arc are relatively small and the distribution of train flows are balanced. The reasonable and balanced utilization shows the optimization scheme can allocate trains evenly to the railway lines of the hub network. It is beneficial to avoid congestion and operate further trains. Notably, the utilization rate of arc (a2, b7) reaches 71.67%, so more parallel railway lines need to be built between Xu-Lan yard and Xi'an.

TABLE 14. The detailed train routes of connecting direction b4.

Train type	Number	Route	G	M	N
Departure trains	1	a4-b4	g34		
	2	a4-b4	g33		
	3	a4-b4	g33		
	4	a4-b4	g37		
	5	a4-b4	g38		
	6	a5-b4	g43		
	7	a5-b4	g43		
	8	a4-b4	g38		
	9	a4-b4	g35		
	10	a5-b4	g43		
	11	a4-b4	g37		
	12	a5-b4	g44		
	13	a4-b4	g38		
	14	a5-b4	g44		
	15	a4-b4	g34		
	16	a4-b4	g37		
	17	a5-b4	g43		
	18	a4-b4	g34		
	19	a4-b4	g36		
	20	a4-b4	g38		
	21	a4-b4	g35		
	22	a4-b4	g38		
Arrival trains	1	b4-a5	g44	m51	n10
	2	b4-a4	g35	m48	n8
	3	b4-a5	g44	m51	
	4	b4-a4	g33	m45	
	5	b4-a4	g35	m39	
	6	b4-a5	g45	m54	
	7	b4-a5	g42	m56	
	8	b4-a4	g34	m38	
	9	b4-a4	g33	m39	
	10	b4-a4	g37	m43	
	11	b4-a5	g45	m51	
	12	b4-a4	g35	m40	
	13	b4-a5	g44	m60	
	14	b4-a5	g43	m52	
	15	b4-a4	g34	m40	
	16	b4-a5	g44	m55	
	17	b4-a4	g35	m40	
	18	b4-a5	g43	m54	
	19	b4-a5	g44	m52	
	20	b4-a5	g45	m52	
	21	b4-a4	g36	m41	
	22	b4-a5	g42	m53	
Passing trains (origin by b4)	1	b4-a4-a2-b1	g34		
	2	b4-a4-a3-b1	g33		
	3	b4-a5-a2-b1	g45		
	4	b4-a4-a2-b1	g34		
	5	b4-a4-a2-b7	g35		
	6	b4-a5-a3-b7	g45		
	7	b4-a5-a3-b7	g42		
	8	b4-a5-a3-b7	g45		
	9	b4-a4-a2-b7	g38		
	10	b4-a4-a2-b7	g36		
	11	b4-a4-a2-b7	g36		
	12	b4-a5-a2-b7	g44		
	13	b4-a4-a2-b7	g34		
	14	b4-a4-a2-b7	g36		
	15	b4-a4-a3-b7	g36		
	16	b4-a4-a2-b7	g35		
	17	b4-a4-a1-b8	g33		
	18	b4-a4-a1-b8	g38		
	19	b4-a4-a1-b8	g34		
	20	b4-a4-a1-b8	g34		
	21	b4-a4-a1-b8	g37		
	22	b4-a5-a1-b8	g42		
	23	b4-a4-a1-b8	g38		
	24	b4-a4-a1-b8	g33		
	25	b4-a4-a1-b8	g34		
	26	b4-a5-a1-b8	g42		
	27	b4-a4-a1-b8	g33		
	28	b4-a4-a1-b8	g34		
	29	b4-a5-a2-a1-b8	g11		
	30	b4-a5-a1-b8	g45		

TABLE 15. The travel routes of passengers.

Traffic zone	Direction	Boarding station	Number (thousand persons/day)
p1	b1	a1	30
p1	b2	a1	12
p1	b3	a1	20
p1	b4	a1	7
p1	b5	a1	30
p1	b6	a1	15
p1	b7	a1	30
p1	b8	a1	17
p2	b1	a2	25
p2	b3	a2	20
p2	b4	a2	3
p2	b5	a2	30
p2	b6	a2	12
p2	b7	a2	14
p2	b8	a2	17
p2	b1	a3	5
p2	b2	a3	12
p2	b4	a3	4
p2	b6	a3	3
p2	b7	a3	16
p3	b1	a4	24
p3	b2	a4	10
p3	b3	a4	16
p3	b4	a4	7
p3	b5	a4	19
p3	b6	a4	15
p3	b7	a4	14
p3	b8	a4	17
p3	b1	a5	6
p3	b2	a5	2
p3	b3	a5	4
p3	b5	a5	11
p3	b7	a5	16

3) The utilization of tracks in stations and depots is also reasonable. In Figure 8, the average capacity utilization rate of the operational tracks is 52.12% and the gaps among different stations are relatively small, which illustrates operation capacity of stations is sufficient and the division of labor is fair. Besides, the average utilization rate of maintenance tracks is 14.30%, which means the maintenance capacity of the hub is even affluent for more trains. However, the average utilization rate of storage tracks is 77.53% and many tracks are even full of capacity, so we need to construct more storage tracks in the depots, especially for the Zhengzhou East station.

4) In Table 15, it can be seen the passengers of different traffic zones can board at the nearest stations, which indicates the optimization scheme can meet the travel demand of residents in different regions of Zhengzhou.

3) DISCUSSION

At present, the fixed scheme is widely used in the multistation railway hub, in which the stations of the hub only deal with trains of fixed connecting directions. Due to the simplicity and clarity, this scheme is convenient for the railway operators, but it may limit the overall capacity of the hub and cannot make full use of all stations. Thus, we compare the optimized train routes with that of the fixed scheme to demonstrate the effectiveness of the optimization scheme.

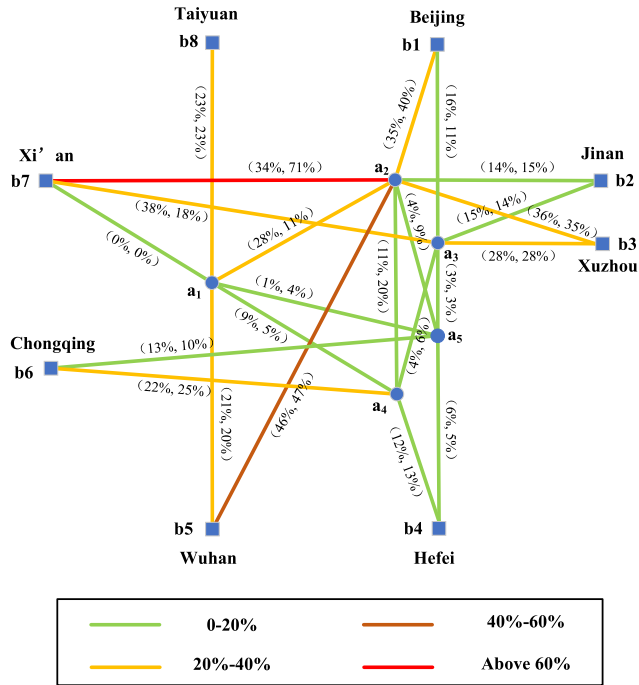


FIGURE 7. The utilization of arcs in the optimization scheme.

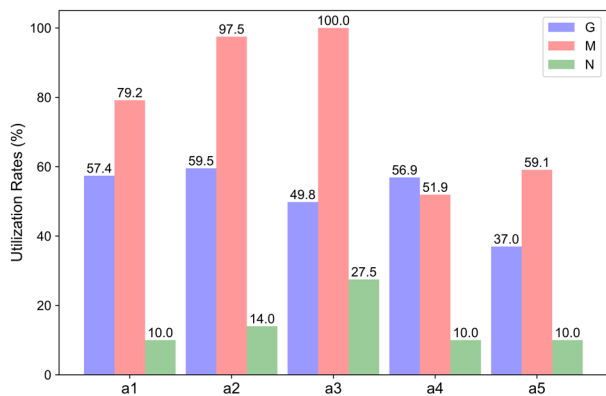


FIGURE 8. The utilization of tracks in the optimization scheme.

Based on the actual operation, we summarize the fixed scheme of the Zhengzhou high-speed railway hub as follows: 1) Zhengzhou station (a1) deals with the departure trains, arrival trains and passing trains of Taiyuan (b8); 2) For Zhengzhou East station, Jing-Guang yard (a2) operates the trains of Beijing (b1) and Wuhan (b5), while Xu-Lan yard (a3) handles the trains of Jinan (b2), Xuzhou (b3) and Xi'an (b7); 3) For Zhengzhou South station, Zheng-He yard (a4) is in charge of trains of Hefei (b4) and Zheng-Wan yard (a5) undertakes the trains of Chongqing (b6).

The results of the optimization scheme and the fixed scheme are presented in Table 16, including the total cost, train cost, passenger cost, average capacity utilization rates of arcs and operational tracks. The capacity utilization rates of the storage tracks and maintenance tracks in different depots

TABLE 16. The results of the optimized scheme and the fixed scheme.

Scheme	Total cost (kRMB)	Train cost (kRMB)	Passenger cost (kRMB)	Utilization rate of arc (%)	Utilization rate of operational track (%)
Fixed	12975820	3315820	966000	17.81	53.44
OP	8151020	3194320	403000	18.64	52.12

TABLE 17. The capacity utilization rates of the storage tracks and maintenance tracks.

Scheme	Utilization rate of storage track (%)					Utilization rate of maintenance track (%)				
	a1	a2	a3	a4	a5	a1	a2	a3	a4	a5
Fixed	79	110	94	42	70	10	10	32	7	13
OP	79	97	100	52	59	10	14	27	10	10

are shown in Table 17. Moreover, for the sake of visibility, we represent the utilization of arcs of the fixed scheme in Figure 9 and express the utilization of operational tracks in Figure 10. The distributions of the arc and operational track capacity utilization rates are represented as pie charts in Figure 11.

From the comparison results in Table 16, 17 and Figure 9-11, the following observations can be obtained:

1) In Table 16, the total cost of the optimization scheme is 37.18% lower than the cost of the fixed scheme. Specifically, the train cost decreases by 3.66% and the passenger cost decreases by 58.28%. These results illustrate the optimization scheme can effectively optimize the train routes and facilitate the passengers.

2) Compared with the fixed scheme, the utilization of arcs in the optimization scheme is more balanced. Although the average utilization rates of arcs are almost similar, the variance of the optimization scheme is 0.0255, which is lower than 0.0348 of the fixed scheme. Furthermore, in Figure 9, it can be seen there are more arcs with high utilization rates in the fixed scheme. For examples, the utilization rate of the arc (a3, b3) is 64%, which is difficult to operate new trains, but the tense situation is mitigated in the optimization scheme.

3) In the fixed scheme, the gaps between the operational track utilization rates of different stations are obvious and the variance reaches 0.2142. In Figure 10, Jing-Guang yard and Xu-Lan yard of Zhengzhou East station are much busier than the other stations, which will limit the introduction of more new trains and waste the capacity of other idle stations. On the contrary, the optimization scheme has a more balanced utilization of operational tracks, which is more conducive to playing the scale effect of the multi-station hub.

4) As for the EMUs depots, in Table 17, it can be seen the storage capacity of the fixed scheme is insufficient, such as the utilization rates of Jing-Guang yard even exceeds 100%. While in the optimization scheme, the trains to be stored in the depots are assigned more reasonably. Besides, since the capacity of maintenance is sufficient, the utilization rates of the two schemes are similar.

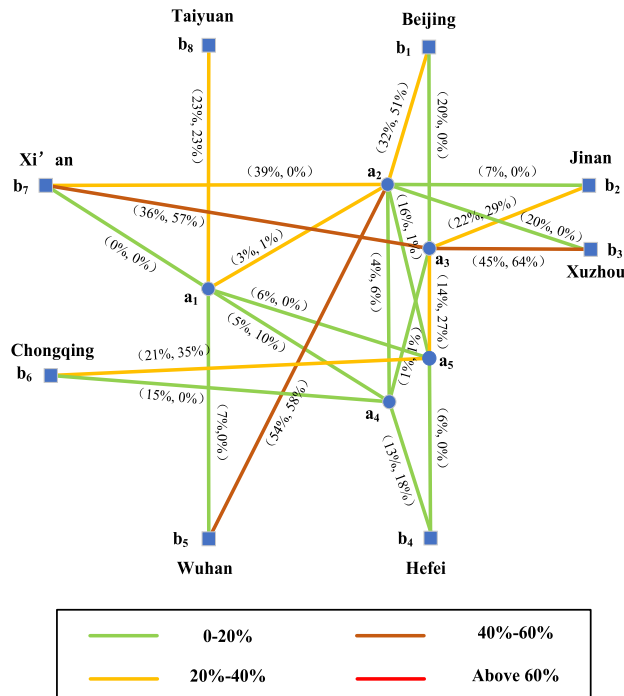


FIGURE 9. The utilization of arcs in the fixed scheme.

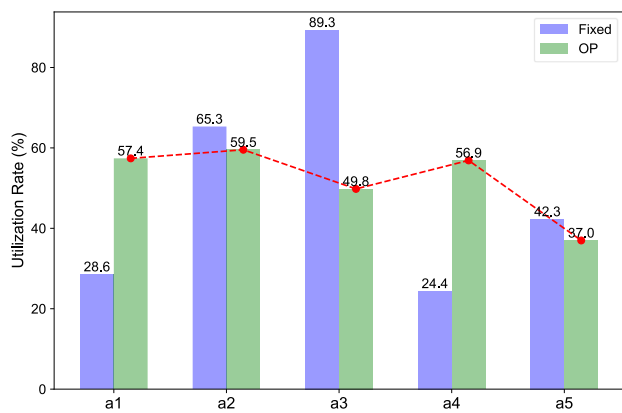


FIGURE 10. The utilization of tracks in different stations.

In summary, compared with the fixed scheme, the optimization scheme has a less cost and more balanced utilization of capacity, which is more suitable for the multi-station hub. In the optimization scheme, since the connecting directions of stations are flexible, all the capacity of stations can be used for operating trains of each direction. This flexible operation scheme diversifies the train routes and avoids the congestion on railway infrastructures, which is beneficial for the balanced distribution of train flows on the hub network. Moreover, the various train routes and stop plans also make it possible for passengers boarding at nearby stations, this is why the passenger cost is reduced. Therefore, in the real-life operation, the railway operators should make more flexible use of the resources and play the scale effect of multi station hubs.

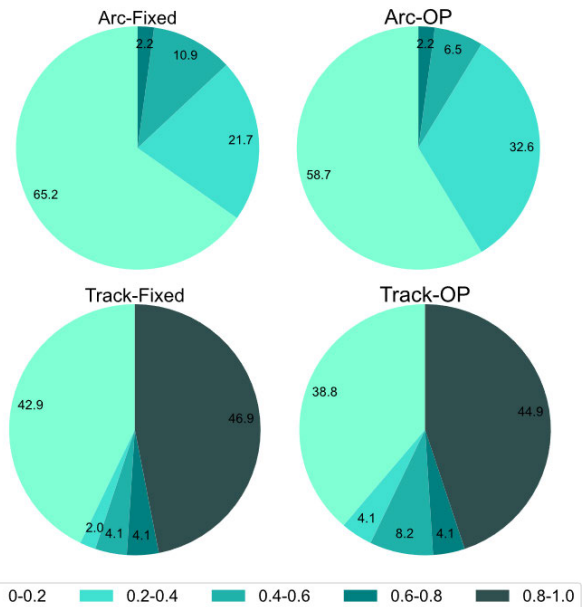


FIGURE 11. The distributions of capacity utilization rates in the fixed and optimization schemes.

VII. CONCLUSION AND FUTURE WORKS

In this paper, we focus on the train routing problem in a multi-station high-speed railway hub. Different from the independent optimization in previous researches, we propose a mixed integer linear programming to optimize the train routes and track allocation simultaneously. Furthermore, the capacity of EMUs depots and demand of passengers are also introduced into the model. According to the characteristics of the model, an Lagrangian relaxation approach with the shortest path algorithm is developed to dualize the hard constraints and decompose the model into two series of subproblems. The numerical experiments are demonstrated to verify the effectiveness of the model and algorithm. Specifically, in the illustrative network, the LR approach can obtain high-quality solutions within 60s and the gaps with the optimal solution is 0%. In the real-world network, the LR approach still can find high-quality solution with a reasonable gap 0.76%. Afterwards, we compare the optimization scheme with the existing fixed scheme, the results show that the total cost of the optimization scheme is decreased by 37.18% and the utilization of arcs and tracks is more reasonable, which prove the availability of our proposed method.

For the sake of simplicity, we just regard the EMUs depot as a part of the station and ignore the routes between the stations and the EMUs depot, which are usually the throat of the depot optimization. Thus, we will attempt to introduce these routes into the TRP of the high-speed railway hub in our future work. Furthermore, although we diversify the train subproblems to break the symmetry issues in LR, the convergence of the algorithm still can be further accelerated, and more techniques can be used to improve the lower bounds.

REFERENCES

- [1] G. Caimi, F. Chudak, M. Fuchsberger, M. Laumanns, and R. Zenklusen, "A new resource-constrained multicommodity flow model for conflict-free train routing and scheduling," *Transp. Sci.*, vol. 45, no. 2, pp. 212–227, May 2011, doi: [10.1287/trsc.1100.0349](https://doi.org/10.1287/trsc.1100.0349).
- [2] X. Jia, R. He, and H. Chai, "Optimizing the number of express freight trains on a high-speed railway corridor by the departure period," *IEEE Access*, vol. 8, pp. 100058–100072, 2020, doi: [10.1109/ACCESS.2020.2995176](https://doi.org/10.1109/ACCESS.2020.2995176).
- [3] W.-J. Li, Y.-T. Wang, L. Nie, Y. Wu, and L. Peng, "High-speed railway EMUs' circulation plan optimization: A two-stage optimization algorithm based on column generation," *IEEE Access*, vol. 8, pp. 15611–15631, 2020, doi: [10.1109/ACCESS.2020.2965601](https://doi.org/10.1109/ACCESS.2020.2965601).
- [4] H. Niu, X. Zhou, and X. Tian, "Coordinating assignment and routing decisions in transit vehicle schedules: A variable-splitting Lagrangian decomposition approach for solution symmetry breaking," *Transp. Res. B, Methodol.*, vol. 107, pp. 70–101, Jan. 2018, doi: [10.1016/j.trb.2017.11.003](https://doi.org/10.1016/j.trb.2017.11.003).
- [5] M. A. Shafia, M. P. Aghae, S. J. Sadjadi, and A. Jamili, "Robust train timetabling problem: Mathematical model and branch and bound algorithm," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 1, pp. 307–317, Mar. 2012, doi: [10.1109/TITS.2011.2169961](https://doi.org/10.1109/TITS.2011.2169961).
- [6] V. Cacchiani, F. Furini, and M. Kidd, "Approaches to a real-world train timetabling problem in a railway node," *Omega*, vol. 58, pp. 97–110, Jan. 2016, doi: [10.1016/j.omega.2015.04.006](https://doi.org/10.1016/j.omega.2015.04.006).
- [7] R. M. Lusby, J. Larsen, M. Ehrgott, and D. Ryan, "Railway track allocation: Models and methods," *OR Spectr.*, vol. 33, no. 4, pp. 843–883, Oct. 2011, doi: [10.1007/s00291-009-0189-0](https://doi.org/10.1007/s00291-009-0189-0).
- [8] Q. Zhang, R. M. Lusby, P. Shang, and X. Zhu, "Simultaneously re-optimizing timetables and platform schedules under planned track maintenance for a high-speed railway network," *Transp. Res. C, Emerg. Technol.*, vol. 121, Dec. 2020, Art. no. 102823, doi: [10.1016/j.trc.2020.102823](https://doi.org/10.1016/j.trc.2020.102823).
- [9] Y.-Y. Tan, Z.-B. Jiang, Y.-X. Li, and R.-X. Wang, "Integration of train-set circulation and adding train paths problem based on an existing cyclic timetable," *IEEE Access*, vol. 8, pp. 87142–87163, 2020, doi: [10.1109/ACCESS.2020.2988978](https://doi.org/10.1109/ACCESS.2020.2988978).
- [10] Y. Lu, M. Lang, Y. Sun, and S. Li, "A fuzzy intercontinental road-rail multimodal routing model with time and train capacity uncertainty and fuzzy programming approaches," *IEEE Access*, vol. 8, pp. 27532–27548, 2020, doi: [10.1109/ACCESS.2020.2971027](https://doi.org/10.1109/ACCESS.2020.2971027).
- [11] M. Carey and D. Lockwood, "A model, algorithms and strategy for train pathing," *J. Oper. Res. Soc.*, vol. 46, no. 8, pp. 988–1005, 1995, doi: [10.1057/jors.1995.136](https://doi.org/10.1057/jors.1995.136).
- [12] L. Meng and X. Zhou, "Simultaneous train rerouting and rescheduling on an N-track network: A model reformulation with network-based cumulative flow variables," *Transp. Res. B, Methodol.*, vol. 67, no. 3, pp. 208–234, Sep. 2014, doi: [10.1016/j.trb.2014.05.005](https://doi.org/10.1016/j.trb.2014.05.005).
- [13] A. D'Ariano, M. Pranzo, and I. A. Hansen, "Conflict resolution and train speed coordination for solving real-time timetable perturbations," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 208–222, Jun. 2007, doi: [10.1109/TITS.2006.888605](https://doi.org/10.1109/TITS.2006.888605).
- [14] L. Yang, J. Qi, S. Li, and Y. Gao, "Collaborative optimization for train scheduling and train stop planning on high-speed railways," *Omega*, vol. 64, pp. 57–76, Oct. 2016, doi: [10.1016/j.omega.2015.11.003](https://doi.org/10.1016/j.omega.2015.11.003).
- [15] M. B. Khan and X. Zhou, "Stochastic optimization model and solution algorithm for robust double-track train-timetabling problem," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 1, pp. 81–89, Mar. 2010, doi: [10.1109/TITS.2009.2030588](https://doi.org/10.1109/TITS.2009.2030588).
- [16] A. D'Ariano, L. Meng, G. Centulio, and F. Corman, "Integrated stochastic optimization approaches for tactical scheduling of trains and railway infrastructure maintenance," *Comput. Ind. Eng.*, vol. 127, pp. 1315–1335, Jan. 2019, doi: [10.1016/j.cie.2017.12.010](https://doi.org/10.1016/j.cie.2017.12.010).
- [17] Y. Zhang, A. D'Ariano, B. He, and Q. Peng, "Microscopic optimization model and algorithm for integrating train timetabling and track maintenance task scheduling," *Transp. Res. B, Methodol.*, vol. 127, pp. 237–278, Sep. 2019, doi: [10.1016/j.trb.2019.07.010](https://doi.org/10.1016/j.trb.2019.07.010).
- [18] C. Wen, P. Huang, Z. Li, J. Lessan, L. Fu, C. Jiang, and X. Xu, "Train dispatching management with data-driven approaches: A comprehensive review and appraisal," *IEEE Access*, vol. 7, pp. 114547–114571, 2019, doi: [10.1109/ACCESS.2019.2935106](https://doi.org/10.1109/ACCESS.2019.2935106).
- [19] X. Luan, J. Miao, L. Meng, F. Corman, and G. Lodewijks, "Integrated optimization on train scheduling and preventive maintenance time slots planning," *Transp. Res. C, Emerg. Technol.*, vol. 80, pp. 329–359, Jul. 2017, doi: [10.1016/j.trc.2017.04.010](https://doi.org/10.1016/j.trc.2017.04.010).
- [20] Z. Lan, S. He, Y. Xu, and Y. Wang, "Optimizing train formation problem with car flow routing and train routing by benders-and-price approach," *IEEE Access*, vol. 7, pp. 178496–178510, 2019, doi: [10.1109/ACCESS.2019.2958601](https://doi.org/10.1109/ACCESS.2019.2958601).
- [21] M. Abril, F. Barber, L. Ingolotti, M. A. Salido, P. Tormos, and A. Lova, "An assessment of railway capacity," *Transp. Res. E, Logistics Transp. Rev.*, vol. 44, no. 5, pp. 774–806, 2008, doi: [10.1016/j.tre.2007.04.001](https://doi.org/10.1016/j.tre.2007.04.001).
- [22] Y. Zhu, B. Mao, Y. Bai, and S. Chen, "A bi-level model for single-line rail timetable design with consideration of demand and capacity," *Transp. Res. C, Emerg. Technol.*, vol. 85, pp. 211–233, Dec. 2017, doi: [10.1016/j.trc.2017.09.002](https://doi.org/10.1016/j.trc.2017.09.002).
- [23] Z. Wang, "Optimization model for EMU collaborative scheduling problem among multiple China railway high-speed depots," in *Proc. Int. Conf. Electron. Inf. Eng. Comput. Sci. (EIECS)*, Sep. 2021, pp. 298–305, doi: [10.1109/EIECS53707.2021.9588008](https://doi.org/10.1109/EIECS53707.2021.9588008).
- [24] J. Shi and H. Li, "Optimization of the shunting operation plan at electric multiple units depots," *J. Adv. Transp.*, vol. 2019, pp. 1–15, Dec. 2019, doi: [10.1155/2019/9120239](https://doi.org/10.1155/2019/9120239).
- [25] H. Li, M. Jin, S. He, Z. Ye, and J. Song, "Optimal track utilization in electric multiple unit maintenance depots," *Comput. Ind. Eng.*, vol. 108, pp. 81–87, Jun. 2017, doi: [10.1016/j.cie.2017.03.031](https://doi.org/10.1016/j.cie.2017.03.031).
- [26] P. Shang, R. Li, J. Guo, K. Xian, and X. Zhou, "Integrating Lagrangian and Eulerian observations for passenger flow state estimation in an urban rail transit network: A space-time-state hyper network-based assignment approach," *Transp. Res. B, Methodol.*, vol. 121, pp. 135–167, Mar. 2019, doi: [10.1016/j.trb.2018.12.015](https://doi.org/10.1016/j.trb.2018.12.015).
- [27] X. Tian and H. Niu, "Optimization of demand-oriented train timetables under overtaking operations: A surrogate-dual-variable column generation for eliminating indivisibility," *Transp. Res. B, Methodol.*, vol. 142, pp. 143–173, Dec. 2020, doi: [10.1016/j.trb.2020.09.010](https://doi.org/10.1016/j.trb.2020.09.010).
- [28] C. Gong, J. Shi, Y. Wang, H. Zhou, L. Yang, D. Chen, and H. Pan, "Train timetabling with dynamic and random passenger demand: A stochastic optimization method," *Transp. Res. C, Emerg. Technol.*, vol. 123, Feb. 2021, Art. no. 102963, doi: [10.1016/j.trc.2021.102963](https://doi.org/10.1016/j.trc.2021.102963).
- [29] L. Meng and X. Zhou, "An integrated train service plan optimization model with variable demand: A team-based scheduling approach with dual cost information in a layered network," *Transp. Res. B, Methodol.*, vol. 125, pp. 1–28, Jul. 2019, doi: [10.1016/j.trb.2019.02.017](https://doi.org/10.1016/j.trb.2019.02.017).
- [30] X. Li, D. Li, X. Hu, Z. Yan, and Y. Wang, "Optimizing train frequencies and train routing with simultaneous passenger assignment in high-speed railway network," *Comput. Ind. Eng.*, vol. 148, Oct. 2020, Art. no. 106650, doi: [10.1016/j.cie.2020.106650](https://doi.org/10.1016/j.cie.2020.106650).
- [31] B. Szpigel, "Optimal train scheduling on a single track railway," *Oper. Res.*, vol. 72, pp. 344–351, 1973.
- [32] Y. Lee and C.-Y. Chen, "A heuristic for the train pathing and timetabling problem," *Transp. Res. B, Methodol.*, vol. 43, nos. 8–9, pp. 837–851, Sep. 2009, doi: [10.1016/j.trb.2009.01.009](https://doi.org/10.1016/j.trb.2009.01.009).
- [33] W. Zhou and H. Teng, "Simultaneous passenger train routing and timetabling using an efficient train-based Lagrangian relaxation decomposition," *Transp. Res. B, Methodol.*, vol. 94, pp. 409–439, Dec. 2016, doi: [10.1016/j.trb.2016.10.010](https://doi.org/10.1016/j.trb.2016.10.010).
- [34] C. Zhang, Y. Gao, L. Yang, Z. Gao, and J. Qi, "Joint optimization of train scheduling and maintenance planning in a railway network: A heuristic algorithm using Lagrangian relaxation," *Transp. Res. B, Methodol.*, vol. 134, pp. 64–92, Apr. 2020, doi: [10.1016/j.trb.2020.02.008](https://doi.org/10.1016/j.trb.2020.02.008).
- [35] R. Gao and H. Niu, "A priority-based ADMM approach for flexible train scheduling problems," *Transp. Res. C, Emerg. Technol.*, vol. 123, Feb. 2021, Art. no. 102960, doi: [10.1016/j.trc.2020.102960](https://doi.org/10.1016/j.trc.2020.102960).
- [36] P. J. Zwaneveld, L. G. Kroon, and S. P. M. Van Hoesel, "Routing trains through a railway station based on a node packing model," *Eur. J. Oper. Res.*, vol. 128, no. 1, pp. 14–33, 2001, doi: [10.1016/S0377-2217\(00\)00087-4](https://doi.org/10.1016/S0377-2217(00)00087-4).
- [37] D. D. L. Cardillo and N. Mione, "k L-list λ colouring of graphs," *Eur. J. Oper. Res.*, vol. 106, no. 1, pp. 160–164, 1998, doi: [10.1016/S0377-2217\(98\)00299-9](https://doi.org/10.1016/S0377-2217(98)00299-9).
- [38] M. Carey and S. Carville, "Scheduling and platforming trains at busy complex stations," *Transp. Res. A, Policy Pract.*, vol. 37, no. 3, pp. 195–224, 2003, doi: [10.1016/S0965-8564\(02\)00012-5](https://doi.org/10.1016/S0965-8564(02)00012-5).
- [39] J. Rodriguez, "A constraint programming model for real-time train scheduling at junctions," *Transp. Res. B, Methodol.*, vol. 41, no. 2, pp. 231–245, Feb. 2007, doi: [10.1016/j.trb.2006.02.006](https://doi.org/10.1016/j.trb.2006.02.006).
- [40] P. Pellegrini, R. Pesenti, and J. Rodriguez, "Efficient train re-routing and rescheduling: Valid inequalities and reformulation of RECIFE-MILP," *Transp. Res. B, Methodol.*, vol. 120, pp. 33–48, Feb. 2019, doi: [10.1016/j.trb.2018.12.008](https://doi.org/10.1016/j.trb.2018.12.008).

- [41] J. Qi, L. Yang, Y. Gao, S. Li, and Z. Gao, "Integrated multi-track station layout design and train scheduling models on railway corridors," *Transp. Res. C, Emerg. Technol.*, vol. 69, pp. 91–119, Aug. 2016, doi: [10.1016/j.trc.2016.05.020](https://doi.org/10.1016/j.trc.2016.05.020).
- [42] Z. Zhang, J. Wang, L. Guo, Y. Chen, and P. Xu, "Research on length design and dynamic usage strategy of railway successive routes," *IEEE Access*, vol. 9, pp. 82701–82712, 2021, doi: [10.1109/ACCESS.2021.3087025](https://doi.org/10.1109/ACCESS.2021.3087025).
- [43] G. Chandrasekaran, P. R. Karthikeyan, N. S. Kumar, and V. Kumarasamy, "Test scheduling of system-on-chip using dragonfly and ant lion optimization algorithms," *J. Intell. Fuzzy Syst.*, vol. 40, no. 3, pp. 4905–4917, Mar. 2021, doi: [10.3233/JIFS-201691](https://doi.org/10.3233/JIFS-201691).
- [44] G. Chandrasekaran, S. Periyasamy, and P. R. Karthikeyan, "Test scheduling for system on chip using modified firefly and modified ABC algorithms," *Social Netw. Appl. Sci.*, vol. 1, no. 9, pp. 1–12, Sep. 2019, doi: [10.1007/s42452-019-1116-x](https://doi.org/10.1007/s42452-019-1116-x).
- [45] G. Chandrasekaran, S. Periyasamy, and K. P. Rajamanickam, "Minimization of test time in system on chip using artificial intelligence-based test scheduling techniques," *Neural Comput. Appl.*, vol. 32, no. 9, pp. 5303–5312, May 2020, doi: [10.1007/s00521-019-04039-6](https://doi.org/10.1007/s00521-019-04039-6).
- [46] Z. Jing and X. Yin, "Neural network-based prediction model for passenger flow in a large passenger station: An exploratory study," *IEEE Access*, vol. 8, pp. 36876–36884, 2020, doi: [10.1109/ACCESS.2020.2972130](https://doi.org/10.1109/ACCESS.2020.2972130).
- [47] J. Yin, D. Chen, L. Yang, T. Tang, and B. Ran, "Efficient real-time train operation algorithms with uncertain passenger demands," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 9, pp. 2600–2612, Sep. 2016, doi: [10.1109/TITS.2015.2478403](https://doi.org/10.1109/TITS.2015.2478403).
- [48] X. Dong, D. Li, Y. Yin, S. Ding, and Z. Cao, "Integrated optimization of train stop planning and timetabling for commuter railways with an extended adaptive large neighborhood search metaheuristic approach," *Transp. Res. C, Emerg. Technol.*, vol. 117, Aug. 2020, Art. no. 102681, doi: [10.1016/j.trc.2020.102681](https://doi.org/10.1016/j.trc.2020.102681).
- [49] Y.-Y. Tan, Y.-X. Li, and R.-X. Wang, "Scheduling extra train paths into cyclic timetable based on the genetic algorithm," *IEEE Access*, vol. 8, pp. 102199–102211, 2020, doi: [10.1109/ACCESS.2020.2997777](https://doi.org/10.1109/ACCESS.2020.2997777).
- [50] Y. Chen, D. Yang, B. Jia, C. Zhou, D. Ma, W. Liu, and J. Gu, "Traffic wave models for the urban road with a work zone," *Transp. Lett.*, early access, pp. 1–10, Apr. 2021, doi: [10.1080/19427867.2021.1920250](https://doi.org/10.1080/19427867.2021.1920250).
- [51] S. A. Almasi and H. R. Behnood, "Exposure based geographic analysis mode for estimating the expected pedestrian crash frequency in urban traffic zones; case study of Tehran," *Accident Anal. Prevention*, vol. 168, Apr. 2022, Art. no. 106576, doi: [10.1016/j.aap.2022.106576](https://doi.org/10.1016/j.aap.2022.106576).



RUI SONG was born in Hebei, China, in 1971. She received the bachelor's degree in railway transportation and the Ph.D. degree in transportation planning and management from Southwest Jiaotong University, in 1992 and 1997, respectively. In 1997, she joined as a Postdoctoral Research Station with the School of Traffic and Transportation, Beijing Jiaotong University. She was worked as an Associate Professor, in 1999, a Professor, in 2005, and a Ph.D. Supervisor, in 2006.



SHIWEI HE received the Ph.D. degree in railway operations from Southwest Jiaotong University, Chengdu, China, in 1996. He is currently a Professor with the School of Traffic and Transportation, Beijing Jiaotong University. His research interests include railway transport organization theory and technology.



ZILONG SONG is currently pursuing the Ph.D. degree with Beijing Jiaotong University, Beijing, China. His research interests include operations management, supply chain management, optimization of e-commerce, and supply chain systems.



YIDONG WANG received the bachelor's and master's degrees in railway transportation from Southwest Jiaotong University, Chengdu, China, in 2014 and 2017, respectively. He is currently pursuing the Ph.D. degree with Beijing Jiaotong University, Beijing, China. His research interests include train routing, track allocation, and timetabling in high-speed railway networks and complex stations.



JUSHANG CHI received the bachelor's degree in railway transportation from Beijing Jiaotong University, Beijing, China, in 2019, where he is currently pursuing the Ph.D. degree. His research interests include vehicle routing and train timetabling.

...