IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 11 August 2022, accepted 24 September 2022, date of publication 28 September 2022, date of current version 21 October 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3210578

RESEARCH ARTICLE

Optimization of High-Speed Railway Line Planning With Passenger and Freight Transport Coordination

SHIQI LI[1,2], MAOXIANG LANG[1,2], (Member, IEEE), SIYU LI[1,2],
XINGHAN CHEN[1,2], (Graduate Student Member, IEEE), XUEQIAO YU[3],
AND YIXUAN GENG[1], (Graduate Student Member, IEEE)
[1]School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
[2]Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing 100044, China
[3]Transportation and Economics Research Institute, China Academy of Railway Sciences Corporation Limited, Beijing 100081, China

Corresponding authors: Maoxiang Lang (mxlang@bjtu.edu.cn) and Xinghan Chen (chenxinghan@bjtu.edu.cn)

This work was supported by the Joint Funds of the National Natural Science Foundation of China under Grant U2034208.

ABSTRACT This paper studies the line planning optimization problem based on the coordinate mode of high-speed railway (HSR) passenger trains and freight trains. The multi objective nonlinear mixed integer programming model of HSR passenger train and freight train line planning with passengers and freight is designed on the basis of comprehensive consideration of passenger and freight transport demand. Then, in order to simultaneously determine the types, origin and destination stations, operation sections, stop schemes, operation frequencies, and demand allocation of HSR passenger trains and freight trains, the model is solved iteratively using a hybrid heuristic algorithm combining a column generation algorithm and a genetic algorithm. Finally, a numerical experiment based on the operation data of China's Dalian-Harbin HSR line is implemented to verify the effectiveness of the proposed model and algorithm, and the solution performance of the CPLEX solver and the hybrid heuristic algorithm is compared. The results show that both the CPLEX solver and the hybrid heuristic algorithm can obtain the global optimal solution set. With the expansion of the scale of the problem, the solution quality and convergence efficiency of the hybrid heuristic algorithm have significantly improved, and it can solve large-scale problems and obtain satisfactory solutions within a shorter time.

INDEX TERMS Generation algorithm, genetic algorithm, high-speed railway, passenger and freight transport coordination, the line planning, train candidate set.

I. INTRODUCTION

Since the opening of the Shinkansen system in Japan in 1964, the high-speed railway (HSR) has been serving passenger transport. The HSR has the outstanding advantages of excellent timeliness, high punctuality, safety and reliability, energy conservation, and environmental protection. It has not only met the needs of diversified and personalized passenger flow but also brought significant economic and social benefits. For example, since the opening of the Beijing-Tianjin Intercity Railway in 2008, China's HSR has developed rapidly. By the end of 2021, the construction of the HSR network will have been basically completed, with an operating mileage of more than 40.000 kilometers, covering 93% of cities with a population of more than 500,000. At the same time, the passenger volume of HSR has grown rapidly, becoming the core driving force behind the growth of railway passenger volume worldwide. Taking China as an example, the passenger volume of China's HSR in the initial stage of 2008 was only 7.34 million. With explosive growth in more than a decade, at the end of 2019, the passenger volume of HSR was 2.36 billion. Although affected by COVID-19, the passenger volume of HSR has declined in the past two years, the overall proportion is still growing.

The construction of HSR passenger transport network is increasingly improved and the level of passenger transport

The associate editor coordinating the review of this manuscript and approving it for publication was Jesus Felez.

service is constantly improving. In order to give full play to the utilization efficiency of HSR resources in regions with sufficient transport capacity and create more operating income, HSR has also begun to explore new service modes in the field of freight transport. The first attempts to use HSR to transport goods were made in France in 1984. During the idle night time of the HSR line, a TGV special post train with a net load of 61 tons and an average speed of 270 km/h operated on the 425 km HSR line between Paris and Lyon [21], transporting letters, express mail, and parcels, which has significant advantages over ordinary freight trains and postal aircraft. Since then, Germany, China [22] and other countries have also explored the HSR freight business to transport mail, express parcels and other high-value-added items. In 2012, China railway companies initially tried to cooperate with express companies to carry out the exploration of HSR freight transportation. Since 2014, China Railway Express Company has gradually expanded the scope of HSR freight transportation in the country by using the spare capacity of the passenger multiple units (MU).

However, there is a significant gap in the development of HSR passenger and freight transportation worldwide. The development of HSR freight transport has been nearly 40 years. The exploration of the express logistics market has only been explored on the consumer side. For example, at this stage, China only transports scattered express parcels by HSR inspecting train and passenger MU trains on some lines. And the French TGV postal train was completely taken offline in 2015 due to the high speed of the Internet, which has led to a decline in the number of letters each year. Meanwhile, in the face of the current rapidly expanding multidimensional freight market with a modern production, circulation, and consumption cycle, the advantages of large-scale, punctual, and efficient HSR trains cannot be fully used. The transportation capacity has not been fully developed, the development scale still needs to be comprehensively expanded, and the transportation organization needs to be optimized and improved. At this stage of the development of HSR, passenger and freight transport are extremely unbalanced in terms of demand expansion, resource allocation, organizational optimization, etc. Based on the development of HSR passenger transport, it is necessary to put forward higher requirements for the growing and diversified demand types of HSR passenger and freight transport, as well as unbalanced and inadequate supply.

Therefore, this paper takes China's HSR, which has the largest scale of HSR network in the world and the most complex railway operation circumstances and external environment, as the research context, and takes the HSR passenger train and freight train running on the same line as the research object. The aim is to improve the development status of HSR freight, to make full use of the capacity resources of the HSR network, to explore the HSR freight demand market, and to balance the organizational structure of HSR passenger and freight transportation. This paper focuses on the optimization of HSR train line planning with passenger and freight transport coordination.

II. LITERATURE REVIEW

While the research system of HSR passenger transport has been perfected so far, HSR freight transport is still in its initial stage. Therefore, this paper first sums up the current situation of HSR freight transport research. Then it discusses the feasibility of HSR passenger train and freight train coordinated organization with reference to the passenger and freight coordinated transport of conventional railways and urban rail transit, and finally summarizes the research related to the railway passenger train and freight train line planning problem.

A. RESEARCH ON HSR FREIGHT

Scholars' research on HSR freight first focused on the feasibility of independent freight trains. Such as Plotkin [36], Gao and Lang [10] analyzed the relevant technical parameter setting requirements of freight MU according to the basic technical requirements of HSR train operation and the technical characteristics of passenger MU. And from the operating speed, loading capacity the technical feasibility and economic feasibility of HSR freight trains are discussed in terms of economic benefits.

After that, the specific application of HSR freight is considered to be mainly in transport organization planning. Scholars have conducted relevant research on HSR freight station location [14], demand prediction and share calculation [46], network design [2], [17], [35], line planning [15], [42], [45], freight flow allocation [7], [41], train timetable preparation [6], [47], etc. Overall, the research is relatively extensive, but the research literature is relatively few, with the line planning part of this paper focusing on the subsequent development of the research.

B. RESEARCH ON OPTIMIZATION OF PASSENGER AND FREIGHT TRANSPORT COORDINATION

This paper considers the organizational coordination of passenger trains and freight trains in the study of line planning problems. The initial concept of passenger and freight coordinated transportation is the idea of urban passenger and freight integration, proposed by the European Commission in 2007, which simply means using urban public transport to carry out freight services while meeting passenger transport needs. On the basis of summarizing the practical experience of many cities in using the idle transport capacity of public transport to carry out urban logistics services, Elbert [39] proposed passenger and freight integration at three levels: shared rail, shared vehicle, and shared carriage.

According to the pilot experiences of HSR freight business in many countries and their relevance to HSR passenger transport, as well as the research results of Lin [26], and others, this paper proposes four operation modes with different degrees of coordination, as shown in Figure 1.

a. One HSR train transports both passengers and freight

b. HSR passenger and freight trains running on common HSR lines at same time

c. HSR passenger and freight trains running on common HSR lines at different times

d. HSR passenger and freight trains running independently on completely different HSR lines

FIGURE 1. **Four operation modes with different degrees of passenger and freight transport coordination.**

The HSR passenger and freight coordinated transportation studied in this paper is mainly the transportation resource sharing at the rail line level, as in Figure 1-c.

The most common realistic scenario in established transport modes is the mixed passenger and freight transportation on common lines of the conventional railway. The key problem and decisive factor of that is the speed difference between various trains and the line carrying capacity. Li [16] made a detailed study on this aspect. The value of this type of research is ultimately reflected in transport scheduling arrangements (i.e. line planning and timetable development), so many scholars have conducted research on the planning of train lines and timetables in the context of mixed passenger and freight lines. Following the organization principle of giving priority to passenger transport of conventional railway, Brännlund [3], Talebian [40], Liu [27], [28] considered the scheduling problem of mixed passenger and freight lines of single track railway, proposed the idea of drawing passenger train lines first and then freight train lines. M. H. Miandoab [33] constructed an optimization model to solve for passenger train and freight train schedules simultaneously while following the constraints of passenger transport priority.

Qi [38] took the single-line, one-way, multi-station urban rail transit system with idle capacity as the research object, comprehensively considered the competitive relationship between passenger flow and freight flow, and proposed adopting the flexible train formation mode to carry out freight services on rail transit to realize the coordinated transportation of passengers and freight. Li [20] took the compound airport rail line as the research object, and discussed the design of train line planning for running freight trains and passenger-freight mixed trains on the passenger airport line.

It can be seen from the above conventional railway and urban rail transit related research basis that HSR passenger and freight coordinated transportation on common lines is feasible in the transportation organization. It is necessary to fully consider the interaction between HSR passenger and freight transportation, so as to ensure the high-quality development of HSR passenger and freight coordinated transportation.

C. RESEARCH ON LINE PLANNING PROBLEM

Overall, transportation organization planning is a comprehensive problem. Learning from Lusby [32] combined the three-level Planning Division of strategy, tactics and operation proposed by Assad [1] with the hierarchical sequential planning strategy proposed by Bussieck [4], this paper describes the planning sequence and corresponding time cycle of sub-tasks in each stage, as shown in Figure 2. The problem of HSR passenger and freight coordinated transportation organization planning studied in this paper is mainly aimed at the line planning stage of the strategic level of railway transportation organization planning.

FIGURE 2. **The whole process of transportation organization planning.**

The earliest research literature on route planning was published by Patz [34], which generates transportation routes according to each origin and destination (OD) demand pair and selects the route set through a greedy algorithm.

The railway line planning problem can be simplified as a problem of selecting a set of paths and the number of different forms of trains on the path according to a given transport demand, which is based on certain optimization objectives and considers the relevant constraints. From the perspective of decision-making content, it mainly involves five parts: train formation structure, route selection, frequency setting, stopping mode, and demand allocation.

Railway line planning problems are usually expressed by optimization models, which belong to NP-hard problems, involving many decision variables, and there are relatively complex logical relationships between decision variables, which require many complex mathematical formulas. According to the differences in rail network structure, it can be divided into network level and line-level rail train line planning optimization problems.

Most of the early line planning problems are focused on the network level. Generally, it is assumed that only trains with a fixed stopping mode, such as through (i.e., no stopping at intermediate stations) or all-stop (i.e., stopping at all intermediate stations), will be operated on the train route. For example, Bussieck [5], López-Ramos [29], Li [19], who used passenger railway trains as their research subjects, and Li [18], Lin [23], Lu, and Lang [30], [31], who used conventional railway freight as their research subjects, investigated the problem of selecting train routes with fixed stopping modes and setting train frequencies and demand allocations based on the link paths between nodes in the network and passenger demand.

With the in-depth study of line planning problems at the network level, scholars have gradually focused on the actual train running mode of each sub-route in the network and no longer assume a single fixed train operation mode on the line but consider the diversity of operation modes [12] to provide more travel options for passengers who choose this route. Based on this, scholars focus their research perspective on the line planning problem at the line level. It has a pre-determined physical structure of the line with the start and end points and the connection sequence of intermediate stations, so the train does not need to select the shortest operation route. Usually, the diversity of train operation modes makes the solution scale of the line planning problem at the line level with the shortest route known larger than that at the network level with the stop scheme known.

Therefore, many scholars, such as Goossens [11], [12] and Fu [8], [9], pre-generated the train candidate set with a feasible stopping mode according to the problem situation and followed the relevant optimization principles before solving. Some scholars, such as Ye [44], assumed train running routes (i.e., train origin and destination and operating sections) and stop schemes are known, and made decisions on train frequency and passenger flow allocation based on passenger demand between nodes on the line. Some scholars, such as Lin [24], [25] and Han [13], determined the train stop scheme and passenger flow allocation based on the known train running route and frequency.

In order to more fully consider the optimization problem more comprehensively, some scholars have conducted comprehensive research on the relatively complete line planning problem. Qi [37] comprehensively summarized the decision-making contents of various parts involved in the railway line planning problem, explored the integration of railway operation sections, stations selection, and passenger flow service processes into a single basic optimization model, and sought the optimal solution to the problem from a global perspective. Zhang and Qi [48] further considered passenger flow demand uncertainty and introduced slack decision variables with the objective of minimizing unmet additional passenger demand to build an uncertain, robust model.

The transportation organization planning of HSR freight is similar to that of HSR passenger transport. Therefore, according to the differences in HSR network structure, it can be similarly divided into network-level and line-level HSR freight train line planning optimization problems. At the network level, usually such as the research of Pazour [35], Bi [2], Li [17] and other scholars, firstly solves for the feasible path or even the shortest path for HSR goods trains between each OD demand, and secondly, based on the optimization objectives, establishes a network-level optimization model and designs corresponding algorithms to decide on train path selection, frequency setting, and freight flow allocation schemes. At the line level, considering multiple stopping modes for trains running on the same line, based on the given train running route and stop scheme, Jia [42] developed a two-level planning model and designed a hybrid algorithm to solve for train operation mode selection, frequency setting, and freight flow allocation; Yu [45] designed a two-stage method for line planning based on the candidate set of HSR freight trains to determine the selection of trains in the candidate set, frequency setting, and freight flow allocation. Jin [15] conducted an expanded study on Yu [46] from the perspective of multiple HSR freight train operation modes and multiple product freight flow demands.

D. CONTRIBUTION

In the summary of the above aspects, the research on HSR freight at this stage focuses on the demand analysis, the prediction of share rate, and the line planning at the network level or line level. From the perspective of transport organization with passenger and freight transport coordination, scholars have studied transport modes such as conventional railways and urban railways, while there is still a blank in the field of HSR. In terms of railway line planning, the theoretical basis of the research on the line planning of HSR freight trains is mainly drawn from railway passenger trains, which provides the theoretical feasibility for this paper to study HSR line planning with passenger and freight transport coordination.

Based on the above summary, this paper mainly references Yu [45] on the research basis of HSR freight including train operation mode selection, frequency setting, and freight flow allocation decision; Qi [37] on the research basis of line planning for railway passenger trains including train origin and destination, passing sections, stop schemes, and passenger flow allocation decision; and Qi [38] on the research basis of passenger and freight coordination transportation schemes for urban rail, including the selection of formation types and departure interval decision, taking HSR passenger trains and freight trains as the research object, and studies the line planning problem of HSR trains with passenger and freight transport coordination. This paper is helpful in the following aspects:

Firstly, considering the future development trend of HSR freight and the market potential after demand expansion, according to passenger and freight demand, this paper studies the coordinated optimization of the line planning of HSR passenger trains and freight trains with the principle of passenger transport priority and sufficient line capacity, and fills the research blank of HSR train coordinated optimization.

Secondly, this paper proposed a nonlinear mixed integer programming model with a single line HSR passenger train and freight train as the research object. The origin and destination, passing sections, stop scheme, operating frequency, and passenger and freight flow allocation of HSR passenger trains and freight trains are solved simultaneously.

Thirdly, this paper introduces the idea of training a candidate set, and the column generation algorithm is combined with a genetic algorithm to construct a hybrid heuristic algorithm for iterative solution.

Finally, taking the Dalian-Harbin line of China's HSR as the experimental environment, the effectiveness of the model and algorithm is illustrated by numerical examples. By comparing the CPLEX solver with the iterative solution of the hybrid heuristic algorithm, the adaptive solution scale and the solution efficiency are evaluated.

The rest of this paper is arranged as follows. In Chapter 3, the detailed problem statements and assumptions are given, the optimization model is constructed, and the trained candidate set is introduced for linearization. In Chapter 4, a hybrid heuristic algorithm combining a column generation algorithm and a genetic algorithm is designed to solve this problem. In Chapter 5, taking the Dalian-Harbin line of China's HSR as an example, the feasibility and solution efficiency of the model and algorithm are evaluated. Finally, Chapter 6 introduces some conclusions and further work.

III. PROBLEM DESCRIPTION AND MODEL DESIGN

The HSR train line planning needs to coordinate and determine the train type, operation section, operation frequency, stop scheme, demand allocation, and other elements. The optimization model structure is complex, the solution scale is huge, and its convexity is difficult to guarantee, so it is difficult to find its global optimal solution or local optimal solution, even a feasible and satisfactory solution. On the basis of existing research on line planning, this part innovatively coordinates and unifies passenger and freight flow demand, and simultaneously solves the line planning of passenger trains and freight trains through the model.

A. PROBLEM DESCRIPTION

In order to clearly express this problem, a four-layer network of ''physics-demand-train-service'' is constructed [43].

The first level is railway infrastructure, called the physical network, including stations and tracks connecting stations.

The second layer is demand origin and destination (OD), which is called the demand network, including passenger flow OD demand and freight flow OD demand between nodes.

The third layer, the demand OD, provides service opportunities with train line planning with different operation modes (OD, route sections, stop schemes), which is called train network.

The fourth layer, which provides a service network for the demand OD. A service arc is defined as a direct transport service from one route to one OD without considering the need to transfer during the journey. The service network consists of service arcs from all train lines. It describes how to allocate passenger and freight demand to each train.

Figure 3 shows an HSR network with four stations. Sub figures 3-a (railway infrastructure), 3-b (OD demand), 3-c (line system) and 3-d (service network) correspond to four layers, respectively.

On the given railway line A to D, the OD demand between nodes is known, and the passenger carrying capacity of HSR trains is 600 people per train. It is assumed that three train lines with different origin and destination stations, operation sections, and stopping modes are provided. Train type 1 departs from A, stops at B and arrives at C; train type 2 departs from A, stops at C and arrives at D; train type 3 departs from A, stops at B and arrives at D. In the service network, two different scenarios for the allocation of the three train line service arcs can be observed.

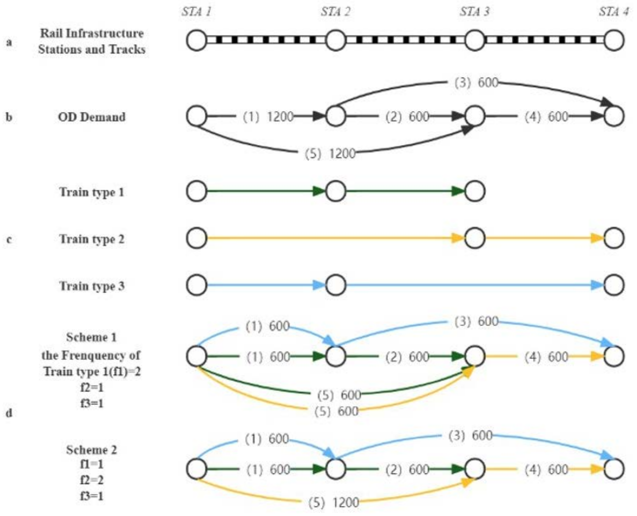


FIGURE 3. The four-layer network of ''physics-demand-train-service.''

Based on the information given in the four-layer network diagram above, in the service network, scheme 1 is served by 2 trains of train type 1, 1 train of train type 2, and 1 train of train type 3. While Scheme 2 is served by 1 train of train type 1, 2 trains of train type 2 and 1 train of train type 3, As the operation cost of trains is directly related to the mileage travelled, Scheme 1 is better from the perspective of the operation cost of HSR trains. In terms of service arc allocation, Scheme 1 is served by one train type 1 and one train type 2, while Scheme 2 is served by two trains types 2. As passenger travel time is directly related to the running time, stopping waiting time, and starting-stopping additional time between trains, Scheme 2 is better from the perspective of passenger travel time.

The above simple cases can more intuitively express the problem of HSR train line planning. In order to more clearly describe the problem and propose a model, the following assumptions are put forward for this problem:

For HSR line, this paper assumes that the HSR network studied is a known one-way line, in which the running direction of HSR trains, the total number of stations on the line, the sequence of stations in the direction, and the running distance between stations are known.

For HSR station, this paper assumes that the stations where the HSR train stops have the operating capacity and conditions that match the passenger transport demand and freight demand of the HSR.

For HSR train, this paper assumes that the HSR freight multiple units operating on the HSR network can satisfy the basic requirements of HSR operation in terms of transportation conditions, operation speed, etc. At the same time, considering the formation structure selection of the existing HSR passenger trains, it is assumed that the passenger and freight HSR trains adopt the 8-car formation structure as the basic formation unit. In order to simplify the speed change in the process of train operation, it is assumed that the whole running speed of the train is a fixed value, the running time of the train section is determined by the section length and the fixed speed of the train, and the additional starting-stopping time of the train and the waiting time at the station are taken as fixed values.

For the demand, this paper assumes that the passenger flow demand and freight flow demand between the OD of each train station within the research time range on the line are known and stable, and there is no transfer and transshipment in the whole transportation process.

B. SYMBOL DESCRIPTION

In order to facilitate the establishment of the mathematical optimization model of this problem, the symbols such as relevant parameters and decision variables involved in the model are described below.

The HSR train line planning model with passenger and freight transport coordination aims to solve the set of HSR passenger trains and freight trains that serve all passenger and freight flow needs in the HSR network, in which the origin and destination stations, operation sections, stop schemes, operation frequencies, and flow allocation of any passenger trains and freight trains should be clear. Therefore, the decision variables in the model constructed in this part include three types: the decision variables related to train line operation sections and stops; the decision variables related to train operation frequency; and the decision variables related to the demand allocation of train services.

C. MATHEMATICAL OPTIMIZATION MODEL

1) OBJECTIVE FUNCTION

In real world operation, the railway department often hopes to provide transportation services satisfactory to customers through low operating costs to obtain the highest operating income, while customers expect to pay less transportation costs to ensure travel demand and obtain transportation services with the best travel efficiency. Therefore, this part takes the maximum operating profit of HSR and the minimum generalized cost of passenger and freight transportation as the objective function.

The operating revenue of HSR trains consists of the sum of passenger and freight transport revenue. The passenger transport revenue is determined by the total passenger flow served by the train and the transportation ticket price of each passenger. The transportation ticket price is calculated according to the fixed ticket price and the variable ticket price related to distance. The calculation formula of HSR passenger train revenue is as follows:

$$\sum_{k_p\in K^P}\sum_{p_{ij}\in P} q_{p_{ij}}^{k_p}\cdot\left(C_{p_1}+C_{p_2}\cdot\sum_{i}^{j-1} zl_i\right) \tag{1}$$

Freight revenue is similar to passenger transport revenue, which is composed of the total freight flow and unit freight

TABLE 1. Model related parameters.

Symbol	Meaning
S	Set of HSR stations, which e, i, j represents the index of S, and $\mid S\mid$ represents the number of stations.
Ori	Starting station of HSR line.
Des	Ending station of HSR line.
Z	Set of line section, which z_i represents the line section from station i to station i+1.
ZL	Set of line section distance, which zl_i represents the line section distance from station i to station i+1, kilometer.
P	Set of passenger demand, which p_{ij} represents the passenger flow demand from station i to station j.
G	Set of freight demand, which g_{ij} represents the freight flow demand from station i to station j.
K	Set of HSR train, which k represents the index of K, $K^P\cup K^G$=K, and $\mid K\mid$ represents the number of trains.
K^P	Set of HSR passenger train, which k_p represents the index of K^P, and $\mid K^P\mid$ represents the number of passenger trains.
K^G	Set of HSR freight train, which k_g represents the index of K^G, and $\mid K^G\mid$ represents the number of freight trains.
C_{p_1}	Fixed fee of passenger transportation, ¥ / passenger.
C_{p_2}	Variable fee of passenger transportation related to distance, ¥ / passenger-kilometer.
$C_{k_p}^1$	Fixed cost of passenger train k_p operation, ¥ / train.
$C_{k_p}^2$	Variable cost of passenger train k_p related to distance, ¥ / train-kilometer.
$C_{k_p}^3$	Variable cost of passenger train k_p related to stop, ¥ / train stop.
C_{g_1}	Fixed fee of freight transportation, ¥ / ton.
C_{g_2}	Variable fee of freight transportation related to distance, ¥ / ton-kilometer.
$C_{k_g}^1$	Fixed cost of freight train k_g operation, ¥ / train.
$C_{k_g}^2$	Variable cost of freight train k_g related to distance, ¥ / train-kilometer.
$C_{k_g}^3$	Variable cost of freight train k_g related to stop, ¥ / train stop.
$C_{k_g}^4$	Variable cost of freight trains k_g related to loading and unloading, ¥ / ton.
N_{z_i}	Upper limit of train passing capacity of line section z_i in unit cycle, number of trains.
N_{S_i}	Upper limit of train operation capacity of station i in unit period, number of trains.
N_P	Seating capacity of passenger train, number of passengers
N_G	Fixed load capacity of freight train, tons.
$N_{stop}^{lbk_p}$	Lower limit of passenger train k_p stopping times in the whole process, times.
$N_{stop}^{ubk_p}$	Upper limit of passenger train k_p stopping times in the whole process, times.
$N_{stop}^{lbk_g}$	Lower limit of freight train k_g stopping times in the whole process, times.
$N_{stop}^{ubk_g}$	Upper limit of freight train k_g stopping times in the whole process, times.
$N_{dis}^{lbk_p}$	Lower limit of passenger train k_p mileage in the whole process, kilometers.
$N_{dis}^{ubk_p}$	Upper limit of passenger train k_p mileage in the whole process, kilometers.
$N_{dis}^{lbk_g}$	Lower limit of freight train k_g mileage in the whole process, kilometers.
$N_{dis}^{ubk_g}$	Upper limit of freight train k_g mileage in the whole process, kilometers.
v_{k_p}	Average running speed of passenger train k_p, km / h.
v_{k_g}	Average running speed of freight train k_g, km / h.
$t_{stop}^{k_p}$	Stopping time of passenger train k_p at intermediate stop, minutes.
$t_{stop}^{k_g}$	Stopping time of freight train k_g at intermediate stop, minutes.
$t_{ss}^{k_p}$	Additional starting-stopping time caused by intermediate stop of passenger train k_p, minutes.
$t_{ss}^{k_g}$	Additional starting-stopping time caused by intermediate stop of freight train k_g, minutes.
M	Positive number of infinity.

TABLE 2. Model decision variables.

Symbol	Meaning
$x_i^{k_p}$	Whether the passenger train k_p stops at the station i, 0-1 variable, 1 for stop and 0 for non-stop.
$x_i^{k_g}$	Whether the freight train k_g stops at the station i, 0-1 variable, 1 for stop and 0 for non-stop.
$s_{z_i}^{k_p}$	Whether the passenger train k_p passes through the section z_i, 0-1 variable, is 1, else 0.
$s_{z_i}^{k_g}$	Whether the freight train k_g passes through the section z_i, 0-1 variable, is 1, else 0.
$O_i^{k_p}$	Whether the station i is the departure station of passenger train k_p, 0-1 variable, is 1, else 0.
$D_i^{k_p}$	Whether the station i is the terminal station of passenger train k_p, 0-1 variable, is 1, else 0.
$O_i^{k_g}$	Whether the station i is the departure station of freight train k_g, 0-1 variable, is 1, else 0.
$D_i^{k_g}$	Whether the station i is the terminal station of freight train k_g, 0-1 variable, is 1, else 0.
f_{k_p}	Number of passenger trains k_p in the study period, integer variable.
f_{k_g}	Number of freight trains k_g in the study period, integer variable.
$w_{p_{ij}}^{k_p}$	Whether the passenger train k_p can serve the passenger transport demand p_{ij}, 0-1 variable, is 1, else 0.
$w_{g_{ij}}^{k_g}$	Whether the freight train k_g can serve the freight transport demand g_{ij}, 0-1 variable, is 1, else 0.
$q_{p_{ij}}^{k_p}$	Passenger flow allocated to passenger train k_p from station i to station j, integer variable.
$q_{g_{ij}}^{k_g}$	Freight flow allocated to freight train k_g from station i to station j, integer variable.

transportation fee served by the train. The calculation formula of HSR freight train revenue is as follows:

$$\sum_{k_g \in K^G} \sum_{g_{ij} \in G} q_{g_{ij}}^{k_g} \cdot (C_{g_1} + C_{g_2} \cdot \sum_{i}^{j-1} zl_i) \tag{2}$$

The passenger transport cost is composed of the actual operation cost of each passenger train, in which the passenger train operation cost is composed of the fixed cost of train operation, the variable cost related to distance and the variable cost related to stop. The calculation formula of HSR passenger train operation cost is as follows:

$$\sum_{k_p \in K^P} f_{k_p} \cdot (C_{k_p}^1 + \sum_{z_i \in Z} C_{k_p}^2 \cdot s_{z_i}^{k_p} \cdot zl_i + \sum_{i \in S} C_{k_p}^3 \cdot x_i^{k_p}) \tag{3}$$

In addition to the actual operation cost of each freight train, the freight cost also includes the loading and unloading cost of goods. The calculation formula of HSR freight train operation cost is as follows:

$$\sum_{k_g \in K^G} f_{k_g} \cdot (C_{k_g}^1 + \sum_{z_i \in Z} C_{k_g}^2 \cdot s_{z_i}^{k_g} \cdot zl_i + \sum_{i \in S} C_{k_g}^3 \cdot x_i^{k_g}) + \sum_{k_g \in K^G} \sum_{g_{ij} \in G} q_{g_{ij}}^{k_g} \cdot C_{k_g}^4 \tag{4}$$

To sum up, the optimization objectives of maximizing the operating profit of HSR passenger trains and freight trains expressed by Z$_1$ is as follows:

$$\begin{aligned} Z_1 = \max & \sum_{k_p \in K^P} \sum_{p_{ij} \in P} q_{p_{ij}}^{k_p} \cdot (C_{p_1} + C_{p_2} \cdot \sum_{i}^{j-1} zl_i) \\ & + \sum_{k_g \in K^G} \sum_{g_{ij} \in G} q_{g_{ij}}^{k_g} \cdot (C_{g_1} + C_{g_2} \cdot \sum_{i}^{j-1} zl_i) \\ & - \sum_{k_p \in K^P} f_{k_p} \cdot (C_{k_p}^1 + \sum_{z_i \in Z} C_{k_p}^2 \cdot s_{z_i}^{k_p} \cdot zl_i + \sum_{i \in S} C_{k_p}^3 \cdot x_i^{k_p}) \\ & - \sum_{k_g \in K^G} f_{k_g} \cdot (C_{k_g}^1 + \sum_{z_i \in Z} C_{k_g}^2 \cdot s_{z_i}^{k_g} \cdot zl_i + \sum_{i \in S} C_{k_g}^3 \cdot x_i^{k_g}) \\ & - \sum_{k_g \in K^G} \sum_{g_{ij} \in G} q_{g_{ij}}^{k_g} \cdot C_{k_g}^4 \end{aligned} \tag{5}$$

In this paper, due to the existence of demand satisfaction constraints and the assumption that the supply capacity of HSR trains is sufficient, HSR passenger trains and freight trains must serve all passenger and freight transport needs. In the objective function, passenger transport revenue, freight revenue, and freight variable costs related to loading and unloading are constants in the solution, so the objective function Z$_1$' can be simplified as:

$$\begin{aligned} Z_1' = \min & \sum_{k_p \in K^P} f_{k_p} \cdot (C_{k_p}^1 + \sum_{z_i \in Z} C_{k_p}^2 \cdot s_{z_i}^{k_p} \cdot zl_i + \sum_{i \in S} C_{k_p}^3 \cdot x_i^{k_p}) \\ & + \sum_{k_g \in K^G} f_{k_g} \cdot (C_{k_g}^1 + \sum_{z_i \in Z} C_{k_g}^2 \cdot s_{z_i}^{k_g} \cdot zl_i + \sum_{i \in S} C_{k_g}^3 \cdot x_i^{k_g}) \end{aligned} \tag{6}$$

In the generalized cost composition of passenger and freight transportation, ticket expenses and freight expenses have been analyzed and calculated in the operating revenue of HSR, so the generalized cost of passenger and freight transportation does not need to be calculated repeatedly. In addition, considering the assumption of sufficient transportation capacity and direct transportation, ignoring the passenger congestion cost, cargo damage and freight difference cost, passenger and cargo transfer related costs, etc., the generalized cost of passenger travel and freight transportation is only based on the train operation time, the sum of stop time and additional starting-stopping time. According to the system optimization principle of demand allocation, Z$_2$ represents the minimum objective of total travel time for all passengers, and Z$_3$ represents the minimum objective of total transportation time for all goods, as follows:

$$\begin{aligned} Z_2 = \min & \sum_{\forall p_{ij} \in P} \sum_{k_p \in K^P} q_{p_{ij}}^{k_p} \cdot w_{p_{ij}}^{k_p} \\ & \cdot (\sum_{i}^{j-1} s_{z_i}^{k_p} \cdot zl_i / v_{k_p} \cdot 60 + (\sum_{i}^{j-1} x_i^{k_p} - 2) \cdot t_{stop}^{k_p} \\ & + (\sum_{i}^{j-1} x_i^{k_p} - 1) \cdot t_{ss}^{k_p}) \end{aligned} \tag{7}$$

$$
\begin{aligned}
Z_3 = \min \sum_{\forall g_{ij} \in G} \sum_{k_g \in K^G} q_{g_{ij}}^{k_g} \cdot w_{g_{ij}}^{k_g} \\
\cdot (\sum_{i}^{j-1} s_{z_i}^{k_g} \cdot zl_i / v_{k_g} \cdot 60 + (\sum_{i}^{j-1} x_i^{k_g} - 2) \cdot t_{stop}^{k_g} \\
+ (\sum_{i}^{j-1} x_i^{k_g} - 1) \cdot t_{ss}^{k_g})
\end{aligned}
\tag{8}
$$

It should be noted that although the generalized travel cost of total passengers and the generalized transportation cost of total cargoes are expressed by time length. There are differences between the time value of passengers and cargoes, so they cannot be directly added and summed.

2) CONSTRAINTS

The constraints of HSR train line planning mainly include meeting passenger and freight demand, transportation capacity constraints and basic train operation conditions. Meeting the passenger and freight flow demand is the core goal of the scheme design. Transportation capacity constraints mainly refer to the carrying capacity of stations and sections, train carrying capacity, etc. The basic operation conditions of the train include constraints such as the range of train operation, train stop requirements, the number of train stop, train operation mileage, etc. In order to ensure that all passenger and freight flow requirements are met, and to ensure the safe operation of the train, the constraints for the model are analyzed in detail below.

Meeting passenger and freight transport demand constraints. According to the basic requirements of the operation plan, all demands can be served by the train, so it is necessary to ensure that the flow demands on any OD pair can be met and arranged on the designated passenger trains and freight trains, that is

$$
\sum_{k_p \in K^P} q_{p_{ij}}^{k_p} \cdot w_{p_{ij}}^{k_p} = p_{ij} \quad \forall p_{ij} \in P \tag{9}
$$

$$
\sum_{k_g \in K^G} q_{g_{ij}}^{k_g} \cdot w_{g_{ij}}^{k_g} = g_{ij} \quad \forall g_{ij} \in G \tag{10}
$$

Train carrying capacity constraints. In order to ensure the accuracy of passenger and freight flow allocation scheme, the carrying capacity of each passing train on any section shall be greater than the actual carrying capacity of the train, that is

$$
\sum_{\forall p_{ej} \in P, e \leq i, j > i, e \in S, j \in S} q_{p_{ej}}^{k_p} \cdot w_{p_{ej}}^{k_p} \leq N_P \cdot f_{k_p} \quad \forall k_p \in K^P, \quad \forall i \in S/Des \tag{11}
$$

$$
\sum_{\forall g_{ej} \in G, e \leq i, j > i, e \in S, j \in S} q_{g_{ej}}^{k_g} \cdot w_{g_{ej}}^{k_g} \leq N_G \cdot f_{k_g} \quad \forall k_g \in K^G, \quad \forall i \in S/Des \tag{12}
$$

The operation of HSR passenger trains and HSR freight trains are independent of each other in many aspects, such as train stop, operation section, frequency, flow allocation, etc. Their relevance is mainly reflected in the joint use of HSR network resources, so the constraints of section passing capacity and station operation capacity is the key of passenger and freight transport coordination.

Section passing capacity constraint. The section passing capacity is not only related to the section fixed equipment, but also affected by the transportation organization mode. Therefore, the passing capacity of each section is also different due to the difference of infrastructure and equipment and the difference of demand density. The number of HSR trains passing through any section is less than its section passing capacity, that is

$$
\sum_{k_p \in K^P} f_{k_p} \cdot s_{z_i}^{k_p} + \sum_{k_g \in K^G} f_{k_g} \cdot s_{z_i}^{k_g} \leq N_{z_i} \quad \forall z_i \in Z \tag{13}
$$

Station operation capacity constraint. The constraint of station operation capacity is similar to the constraint of section passing capacity, which are related to the conditions of station facilities and equipment and the passenger and freight demand density of the station. There are certain differences in the operation capacity of each station. The number of trains stop at any station for relevant operations is less than its station operation capacity, that is

$$
\sum_{k_p \in K^P} f_{k_p} \cdot x_i^{k_p} + \sum_{k_g \in K^G} f_{k_g} \cdot x_i^{k_g} \leq N_{S_i} \quad \forall i \in S \tag{14}
$$

The number of train stop constraints. In order to ensure that any train can fully serve the flow demand in the network, and at the same time, there will not be too many stops, which will affect the whole running time and transportation efficiency of the train, the number of train stop constraints are set for any passenger train and freight train in the whole operation process, so as to ensure that the number of train stop of passenger trains and freight trains are in a certain interval, that is

$$
N_{stop}^{lbk_p} \leq \sum_{i \in S} x_i^{k_p} \leq N_{stop}^{ubk_p} \quad \forall k_p \in K^P \tag{15}
$$

$$
N_{stop}^{lbk_g} \leq \sum_{i \in S} x_i^{k_g} \leq N_{stop}^{ubk_g} \quad \forall k_g \in K^G \tag{16}
$$

Train mileage constraints. If the origin and destination stations of the train are too close, the total fixed cost of train operation will increase. If the origin and destination stations of the train are too far, the competitiveness of HSR train transportation beyond the economic haul distance will be difficult to reflect. Therefore, it is necessary to ensure that passenger and freight trains operate within a reasonable mileage, that is

$$
N_{dis}^{lbk_p} \leq \sum_{z_i \in Z} s_{z_i}^{k_p} \cdot zl_i \leq N_{dis}^{ubk_p} \quad \forall k_p \in K^P \tag{17}
$$

$$
N_{dis}^{lbk_g} \leq \sum_{z_i \in Z} s_{z_i}^{k_g} \cdot zl_i \leq N_{dis}^{ubk_g} \quad \forall k_g \in K^G \tag{18}
$$

Train service OD stop constraints. The OD transportation service can be completed only when the HSR train stops at any OD pair start and end points within the operation route, that is

$$w_{p_{ij}}^{k_p} = \begin{cases} 1 & \text{if } x_i^{k_p} = 1 \text{ and } x_j^{k_p} = 1 \\ 0 & \text{else} \end{cases} \quad \forall k_p \in K^P, \ \forall p_{ij} \in P \tag{19}$$

$$w_{g_{ij}}^{k_g} = \begin{cases} 1 & \text{if } x_i^{k_g} = 1 \text{ and } x_j^{k_g} = 1 \\ 0 & \text{else} \end{cases} \quad \forall k_g \in K^G, \ \forall g_{ij} \in G \tag{20}$$

Quantity constraints of train origin and destination stations. The train can only have one departure station and one terminal station on its running line, that is

$$\sum_{i \in S} O_i^{k_p} = 1 \quad \forall k_p \in K^P \tag{21}$$

$$\sum_{i \in S} D_i^{k_p} = 1 \quad \forall k_p \in K^P \tag{22}$$

$$\sum_{i \in S} O_i^{k_g} = 1 \quad \forall k_g \in K^G \tag{23}$$

$$\sum_{i \in S} D_i^{k_g} = 1 \quad \forall k_g \in K^G \tag{24}$$

Sequence constraints of train origin and destination stations. In order to express that the running direction of the train is unified and fixed in the running route, it is stipulated that any train will run from its departure station to the terminal station, and the departure station of the train on its running line must be before the terminal, that is

$$\sum_{i \in S} O_i^{k_p} \cdot i \le \sum_{i \in S} D_i^{k_p} \cdot i \quad \forall k_p \in K^P \tag{25}$$

$$\sum_{i \in S} O_i^{k_g} \cdot i \le \sum_{i \in S} D_i^{k_g} \cdot i \quad \forall k_g \in K^G \tag{26}$$

The train origin station and destination station stop constraints. The train finishes the departure and arrival operations at the origin and destination stations, and any train must stop at the origin and destination stations of the running line, that is

$$x_i^{k_p} = 1 \quad \text{if } O_i^{k_p} = 1 \text{ or } D_i^{k_p} = 1 \ \forall k_p \in K^P, \ \forall i \in S \tag{27}$$

$$x_i^{k_g} = 1 \quad \text{if } O_i^{k_g} = 1 \text{ or } D_i^{k_g} = 1 \ \forall k_g \in K^G, \ \forall i \in S \tag{28}$$

Train stop range constraints. The origin and destination stations of trains determine the operation sections of the train on the line, and the train cannot stop outside the origin and destination stations of its running line, that is

$$x_e^{k_p} = 0 \quad \text{if } O_i^{k_p} = 1 \text{ and } e < i \ \forall k_p \in K^P, \ \forall e, i \in S \tag{29}$$

$$x_j^{k_p} = 0 \quad \text{if } D_i^{k_p} = 1 \text{ and } j > i \ \forall k_p \in K^P, \ \forall i, j \in S \tag{30}$$

$$x_e^{k_g} = 0 \quad \text{if } O_i^{k_g} = 1 \text{ and } e < i \ \forall k_g \in K^G, \ \forall e, i \in S \tag{31}$$

$$x_j^{k_g} = 0 \quad \text{if } D_i^{k_g} = 1 \text{ and } j > i \ \forall k_g \in K^G, \ \forall i, j \in S \tag{32}$$

Train passing range constraints. The origin and destination stations of the trains determine the operation sections of the train on the line. The sections between the origin and destination stations of the train running line are passing sections, and do not pass through the sections outside the origin and destination stations, that is

$$s_{z_i}^{k_p} = \begin{cases} 1 & \text{if } O_e^{k_p} = 1, D_j^{k_p} = 1, e \le i < j \\ 0 & \text{else} \end{cases} \quad \forall k_p \in K^P, \ \forall e, i, j \in S \tag{33}$$

$$s_{z_i}^{k_g} = \begin{cases} 1 & \text{if } O_e^{k_g} = 1, D_j^{k_g} = 1, e \le i < j \\ 0 & \text{else} \end{cases} \quad \forall k_g \in K^G, \ \forall e, i, j \in S \tag{34}$$

3) DECISION VARIABLES

The decision-making information such as train operation frequency, flow allocation scheme, stop selection, service OD, passing sections, origin and destination stations are written into the model as the decision variables of the HSR train line planning model in this paper. The specific value range of each decision variable is as follows:

Decision variables of train operation frequency.

$$f_{k_p} \in \mathrm{N} \quad \forall k_p \in K^P \tag{35}$$

$$f_{k_g} \in \mathrm{N} \quad \forall k_g \in K^G \tag{36}$$

Decision variables of passenger and freight flow.

$$q_{p_{ij}}^{k_p} \in \mathrm{N} \quad \forall k_p \in K^P, \ \forall p_{ij} \in P \tag{37}$$

$$q_{g_{ij}}^{k_g} \ge 0 \quad \forall k_g \in K^G, \ \forall g_{ij} \in G \tag{38}$$

Decision variables of train stop.

$$0 \le x_i^{k_p} \le 1 \quad \forall x_i^{k_p} \in \mathrm{N} \ \forall k_p \in K^P, \ \forall i \in S \tag{39}$$

$$0 \le x_i^{k_g} \le 1 \quad \forall x_i^{k_g} \in \mathrm{N} \ \forall k_g \in K^G, \ \forall i \in S \tag{40}$$

Decision variables of train demand service.

$$0 \le w_{p_{ij}}^{k_p} \le 1 \quad \forall w_{p_{ij}}^{k_p} \in \mathrm{N} \ \forall k_p \in K^P, \ \forall p_{ij} \in P \tag{41}$$

$$0 \le w_{g_{ij}}^{k_g} \le 1 \quad \forall w_{g_{ij}}^{k_g} \in \mathrm{N} \ \forall k_g \in K^G, \ \forall g_{ij} \in G \tag{42}$$

Decision variables of train passing through the section.

$$0 \le s_{z_i}^{k_p} \le 1 \quad \forall s_{z_i}^{k_p} \in \mathrm{N} \ \forall k_p \in K^P, \ \forall z_i \in Z \tag{43}$$

$$0 \le s_{z_i}^{k_g} \le 1 \quad \forall s_{z_i}^{k_g} \in \mathrm{N} \ \forall k_g \in K^G, \ \forall z_i \in Z \tag{44}$$

Decision variables of train OD stations.

$$0 \le O_i^{k_p} \le 1 \quad \forall O_i^{k_p} \in \mathrm{N} \ \forall k_p \in K^P, \ \forall i \in S \tag{45}$$

$$0 \le D_i^{k_p} \le 1 \quad \forall D_i^{k_p} \in \mathrm{N} \ \forall k_p \in K^P, \ \forall i \in S \tag{46}$$

$$0 \le O_i^{k_g} \le 1 \quad \forall O_i^{k_g} \in \mathrm{N} \ \forall k_g \in K^G, \ \forall i \in S \tag{47}$$

$$0 \le D_i^{k_g} \le 1 \quad \forall D_i^{k_g} \in \mathrm{N} \ \forall k_g \in K^G, \ \forall i \in S \tag{48}$$

4) OVERALL MODEL STRUCTURE

To sum up, the complete mathematical optimization model of the HSR train line planning with the passenger and freight transport coordination as Model1 is sorted as follows:

Objective function: Equation (6), (7), (8).

Constraints: Equation (9) – (34).

Decision variables: Equation (35) - (48).

The model sets the optimization objectives of minimizing the operating costs of HSR passenger trains and freight trains, the total travel time of passengers and the total transport time of goods. The normal and safe operation of trains is ensured by establishing constraints on train's service OD, the number of train origins and destinations, the sequence of train origins and destinations, train origins and destinations stop, the range of train stop, the range of sections trains pass through, etc. The necessary transport services for passengers and freight owners are ensured by meeting passenger and freight transport demand constraints. The design of the scheme to meet line and train capacity requirements is ensured by constraints on train carrying capacity, section passing capacity and station operating capacity. Economic and competitive train operation is ensured by specifying the number of passenger train and freight train stop and the travel mileage. The objectives and constraints of the model achieve the aim of designing a coordinated solution for the line planning of HSR trains to meet passenger and freight transport demand.

D. NONLINEAR OBJECTIVE AND CONSTRAINT LINEARIZATION BASED ON CANDIDATE SET

Although the above HSR train line planning model with passenger and freight transport coordination has multiple items generated by the multiplication of decision variables, and there are conditional constraints with decision variables, the model belongs to a multi-objective nonlinear mixed integer programming model. The existence of nonlinear objectives and constraints in the model greatly increases the difficulty of using accurate solution algorithms (such as the simplex method, branch and bound method) to solve the problem. The line planning model constructed in this paper is a typical NP-hard problem, and its shortcomings are particularly obvious.

In order to make the model of HSR train line planning with passenger and freight transport coordination be accurately solved by classical mathematical programming software, This part linearizes the nonlinear objectives and constraints in the model, respectively, so as to obtain a linear programming model equivalent to the original model.

Referring to the idea of "train candidate set" adopted by Fu [8] and other scholars to deal with the linearization of nonlinear objectives and constraints, this paper defines the "HSR passenger train and freight train candidate set" as an HSR train with known passenger and freight transportation objects, origin and destination stations and operating sections, train types, train formation, and stop schemes. It needs to be solved is only the operation frequency of the trains to be selected in the candidate set on the line. If the operation frequency is 0, it means that the candidate train is not selected to run on the line at last. If the operation frequency is positive integers that are non-zero, it means that the candidate train is selected to run on the line and serve the corresponding passenger and freight demand.

Therefore, these four types of decision variables that train stop, train demand service, train passing sections, and train origin and destination stations are transformed into the subsidiary information which are known and determined of the operation frequency. The operation frequency can be written as $f_{k_p}(\underset{i\in S}{\exists} O_i^{k_p}, \underset{i\in S}{\exists} D_i^{k_p}, \underset{i\in S}{\exists} x_i^{k_p}, \underset{z_i\in Z}{\exists} s_{z_i}^{k_p}, \underset{p_{ij}\in P}{\exists} w_{p_{ij}}^{k_p})$, $f_{k_g}(\underset{i\in S}{\exists} O_i^{k_g}, \underset{i\in S}{\exists} D_i^{k_g}, \underset{i\in S}{\exists} x_i^{k_g}, \underset{z_i\in Z}{\exists} s_{z_i}^{k_g}, \underset{g_{ij}\in G}{\exists} w_{g_{ij}}^{k_g})$.

Therefore, based on the "train candidate set", the nonlinear multiplicative multiple terms associated with these four types of decision variables in the model are transformed into linear objectives and constraints.

At the same time, constraints such as train's service OD, number of train origin and destination stations, sequence of train origin and destination stations, train origin station and destination station stop, train stop range, train passing section range, etc. in the model can also be omitted as known conditions. Furthermore, non-linear constraints in the model containing conditional constraints on decision variables no longer affect the linearization of the model.

But the matching relationship between the train to be selected and passenger and freight flow is still unknown. There are multiple terms similar to $q_{p_{ij}}^{k_p} \cdot w_{p_{ij}}^{k_p}$, $q_{g_{ij}}^{k_g} \cdot w_{g_{ij}}^{k_g}$ in passenger and freight demand satisfaction constraints and train loading capacity constraints, among which the train demand service is one of the determining parameters of the operation frequency $f_{k_p}(\underset{i\in S}{\exists} O_i^{k_p}, \underset{i\in S}{\exists} D_i^{k_p}, \underset{i\in S}{\exists} x_i^{k_p}, \underset{z_i\in Z}{\exists} s_{z_i}^{k_p}, \underset{p_{ij}\in P}{\exists} w_{p_{ij}}^{k_p})$, $f_{k_g}(\underset{i\in S}{\exists} O_i^{k_g}, \underset{i\in S}{\exists} D_i^{k_g}, \underset{i\in S}{\exists} x_i^{k_g}, \underset{z_i\in Z}{\exists} s_{z_i}^{k_g}, \underset{g_{ij}\in G}{\exists} w_{g_{ij}}^{k_g})$ of the decision variables, so these multiple items can be expressed as $q_{p_{ij}}^{k_p} \cdot f_{k_p}(w_{p_{ij}}^{k_p})$, $q_{g_{ij}}^{k_g} \cdot f_{k_g}(w_{g_{ij}}^{k_g})$ in essence, which are still multiple items.

In order to clarify the relationship between train operation and flow allocation, the following linearization is made for the nonlinear constraints with multiple items of meeting passenger and freight demand and train loading capacity.

Add train flow and passenger and freight flow coupling constraints. For a HSR train passing through an OD, if the train k_p or k_g cannot serve the OD demand p_{ij} or g_{ij}, then$w_{p_{ij}}^{k_p} = 0$ or $w_{g_{ij}}^{k_g} = 0$. At this time, the flow $q_{p_{ij}}^{k_p}$ or $q_{g_{ij}}^{k_g}$of the train k_p or k_g loading OD demand p_{ij} or g_{ij} is 0. If the train k_p or k_g can serve the OD demand p_{ij} or g_{ij}, then $w_{p_{ij}}^{k_p} = 1$or$w_{g_{ij}}^{k_g} = 1$ also exists:

$$q_{p_{ij}}^{k_p} \le w_{p_{ij}}^{k_p} \cdot f_{k_p} \cdot \mathrm{M} \quad \forall k_p \in K^P,\ \forall p_{ij} \in P \tag{49}$$

$$q_{g_{ij}}^{k_g} \le w_{g_{ij}}^{k_g} \cdot f_{k_g} \cdot \mathrm{M} \quad \forall k_g \in K^G,\ \forall g_{ij} \in G \tag{50}$$

Combined with the above coupling constraints, the meeting passenger and freight demand constraints (9), (10) are transformed into

$$\sum_{k_p \in K^P} q_{p_{ij}}^{k_p} = p_{ij} \quad \forall p_{ij} \in P \tag{51}$$

$$\sum_{k_g \in K^G} q_{g_{ij}}^{k_g} = g_{ij} \quad \forall g_{ij} \in G \tag{52}$$

Train carrying capacity constraints (11), (12) are transformed into

$$\sum_{\forall p_{ej}\in P, e\leq i, j>i, e\in S/Des, j\in S/Ori} q_{p_{ij}}^{k_p} \leq N_P \cdot f_{k_p} \quad \forall k_p \in K^P,\ \forall i \in S/Des \tag{53}$$

$$\sum_{\forall g_{ej}\in G, e\leq i, j>i, e\in S/Des, j\in S/Ori} q_{g_{ij}}^{k_g} \leq N_G \cdot f_{k_g} \quad \forall k_g \in K^G,\ \forall i \in S/Des \tag{54}$$

The decision variables (35) and (36) of the operation frequency are transformed into:

$$f_{k_p}(\underset{i\in S}{\exists} O_i^{k_p}, \underset{i\in S}{\exists} D_i^{k_p}, \underset{i\in S}{\exists} x_i^{k_p}, \underset{z_i\in Z}{\exists} s_{z_i}^{k_p}, \underset{p_{ij}\in P}{\exists} w_{p_{ij}}^{k_p}) \in \mathrm{N} \quad \forall k_p \in K^P \tag{55}$$

$$f_{k_g}(\underset{i\in S}{\exists} O_i^{k_g}, \underset{i\in S}{\exists} D_i^{k_g}, \underset{i\in S}{\exists} x_i^{k_g}, \underset{z_i\in Z}{\exists} s_{z_i}^{k_g}, \underset{g_{ij}\in G}{\exists} w_{g_{ij}}^{k_g}) \in \mathrm{N} \quad \forall k_g \in K^G \tag{56}$$

The selection of trains in the candidate set of HSR trains is determined by the number of train stop constraints (15), (16) and train mileage constraints (17), (18) in the original model, and all trains that meet the constraints are added to the "candidate set" as candidate trains.

Thus, the original model is transformed into Model 2 through relevant processing:

Objective function: Equation (6), (7), (8).

Constraints: Equation (49) - (54), (13), (14).

Decision variables: Equation (55), (56), (37), (38).

The above the Model 2 is the equivalent linear solution model based on the train candidate set of Model 1 that HSR train line planning mode with the passenger freight coordination. This model mainly solves the feasible train selection based on the train candidate set, the operation frequency of the selected train and the passenger and freight demand flow allocation.

IV. ALGORITHM DESIGN

The linearized HSR train line planning model with passenger and freight transport coordination still belongs to a large-scale multi-objective mixed integer programming problem, and its solution scale is mainly reflected in the large scale of the global solution space of the feasible candidate set train. The solution efficiency is low and the solution is difficult. Therefore, this part further uses the idea of "candidate set" and reduces the solution space of the candidate set based on the column generation algorithm. And a hybrid heuristic algorithm combining column generation and genetic algorithms is designed to solve the model.

A. DESIGN OF HYBRID HEURISTIC ALGORITHM

The column generation algorithm is a very efficient algorithm for solving large-scale linear optimization problems. In essence, it is a form of the simplex method, which is widely used to solve the famous NP-hard optimization problems.

From the perspective of structure, the large scale of linear programming problems is mainly reflected in two aspects: one is the large scale of decision variables, and the other is the extremely large number of constraints, which makes the number of poles on the polyhedron describing the problem model extremely large. According to the correlation differences between constraints and decision variables, Danzig and Wolfe designed an effective decomposition method for large-scale linear programming problems, which is called the DW decomposition principle. Its essence is to decompose the original linear programming problem into a master problem (MP) and several sub-problems (SPs); that is, linear programming problem constraints are allocated to the main layer and subsidiary layer.

The traditional simplex method is to test the poles of the problem polyhedron one by one. The drawback is that with the exponential growth of the number of poles, the complexity of the solution also increases exponentially. The idea of the column generation algorithm is to jump out of the pole to test one by one, and directly select the sub polyhedron in the polyhedron of the model (composed of the selected pole set) to test. Its solution steps are as follows. Firstly, based on the DW decomposition principle, the decomposed MP is used to represent the original model of the problem, and a part of the feasible column set (sub polyhedron, which is composed of the selected set of poles that pass the test) is extracted from the MP to construct a restricted master problem (RMP). At this time, the solution space of the RMP is relatively small. The simplex method can be used to obtain the optimal solution of RMP and the corresponding simplex multiplier (also called the shadow price). The simplex multiplier is used to guide the optimization goal of forming several SPs, and the effective addition column of RMP is generated by solving SP. After repeating the above process to make all the test numbers meet the optimal test, the column generation algorithm ends.

The key to the implementation of the column generation algorithm lies in:

1) To determine the initial solution, intelligent algorithms such as genetic algorithm can be selected.
2) RMP and several SPs can be solved directly by using the linear programming method.
3) To determine the integer solution, the branch and bound or cut plane techniques are generally used to obtain the integer solution of the problem. This paper calls the mathematical optimization engine CPLEX and uses the branch and bound method to get the integer solution.

The first condition for solving the column generation algorithm is to identify a group of RMP's initial feasible solutions. Because the essence of solving the column generation algorithm is linear programming, the RMP optimal solution of any group of initial feasible solutions iteratively solved by the column generation algorithm is unique. And the MP optimal solution of any group of RMP optimal solutions iteratively solved by the branch and bound method is also unique. That is, a set of selected initial feasible solutions uniquely corresponds to a set of MP optimal solutions. On the contrary, different initial feasible solutions may solve the same MP optimal solution.

Due to the selection of the initial feasible solution directly affects the acquisition of the MP optimal solution to the problem, it is easy to fall into the local optimal solution only starting from a certain feasible initial solution. In order to obtain the global optimal solution or a satisfactory solution efficiently, it is crucial to choose a suitable heuristic algorithm to compile the generation of the initial solution. Furthermore, the MP optimal solution will not be uniquely limited by the selection of the initial feasible solution, so it is not necessary to traverse all the feasible initial solutions globally.

A genetic algorithm takes all individuals in a population as objects, and uses randomization technology to search a coded parameter space. Usually, it starts with a potential population representing the problem. After the initial generation of the population, according to the principle of survival of the fittest, generation-by-generation evolution generates increasingly better approximate solutions. In each generation, individuals are selected according to their fitness in the problem domain with the help of genetic operators of natural genetics. The population representing the new solution set is generated by combining crossover and mutation. This process will lead to the offspring population of species groups like natural evolution being more suitable for the environment than the previous generation. The optimal individual in the last generation population can be used as the approximate optimal solution to the problem after decoding.

The column generation algorithm needs a certain scale of random initial solution generation to ensure a large enough search solution space coverage, and the random search characteristics of genetic algorithm can adapt to it. At the same time, the optimization ability and convergence efficiency of genetic algorithm are limited by the fitness setting of randomly generated chromosomes. Solving the train candidate set by column generation algorithm is a further optimization of the initial train candidate set. From the perspective of overall algorithm design, it is also an optimization of the chromosome fitness of genetic algorithm, which significantly enhances the optimization ability and convergence efficiency of genetic algorithm. Therefore, this paper combines genetic algorithm with column generation algorithm, compiles the initial solution of column generation algorithm as chromosome, and designs a hybrid heuristic algorithm.

B. PROCESS OF HYBRID HEURISTIC ALGORITHM

In order to linearize the HSR train line planning model with passenger freight coordination, the idea of ‘‘candidate set’’ is used. Before the model is directly solved, the ‘‘train candidate set’’ is screened through the number of train stop constraints (15), (16) and train mileage constraints (17), (18). But the scale of the ‘‘train candidate set’’ screened only through these two constraints is still large. With the increase in the number of stations, it will increase exponentially.

Therefore, in order to further screen the optional trains in the ‘‘train candidate set’’, this part constructs a heuristic algorithm mixed with column generation and genetic algorithm, which is applied to the model solution. It constructs the initial candidate set through genetic algorithm coding and solves the trained candidate set based on the column generation algorithm (including the RMP of the candidate set and the train generation SP). According to the obtained train candidate set, Model 2 is used to solve the HSR passenger train and freight train line planning and demand flow allocation problems. The approximate optimal solution is selected through continuous iteration, which reduces the difficulty of solving the original model. The hybrid heuristic algorithm is shown in Figure 4:

The overall structure of the hybrid heuristic algorithm is based on the genetic algorithm framework design. First, the chromosome structure form is clearly coded, and the chromosome list with the attributes of the initial candidate set is randomly generated according to the population size. The train candidate set solution module is constructed by using the column generation algorithm, and the current optimal train candidate set is substituted into the linear equivalent Model 2 to obtain the fitness value of the chromosome of the initial candidate set. The fitness value of each chromosome in the initial population is calculated one by one. Then the next generation of a new population is generated according to the rules of selection, crossover, and mutation of the genetic algorithm, and then the fitness value is calculated, the population genetic evolution and termination are determined. Finally, the global optimal solution or a satisfactory approximate optimal solution is output, and the algorithm ends.

C. RELEVANT MODEL FOR SOLVING TRAIN CANDIDATE SET BASED ON COLUMN GENERATION ALGORITHM

According to the flow design of the hybrid heuristic algorithm, the overall algorithm needs to apply a column generation algorithm to solve the train candidate set and then substitute it into Model 2 to obtain the chromosome fitness value, train line planning and flow allocation scheme. Therefore, this part describes the solution of the test case set in detail.

For the solution of the train candidate set problem, each ‘‘column’’ represents the trains of different operation sections and stop schemes. In the iteration process of the column generation algorithm, if there is a more suitable ‘‘column’’ to make the objective function of the candidate set better, it will ‘‘add’’ a column. Each ‘‘addition of column’’ means that trains of different types from the previous feasible solution are added to the candidate set, and the ‘‘addition of column’’ will continue with the iteration process. Then, the train candidate set of the initial solution will be continuously optimized until the SP cannot be ‘‘addition of column’’, that is, there is no better kind of train. The proposed candidate set solution is the current optimal training candidate set solution.

Based on the principle of the column generation algorithm, the train candidate set solving problem is decomposed into the train candidate set RMP model as Model 3, the passenger train candidate set generation SP model as Model 4, and the

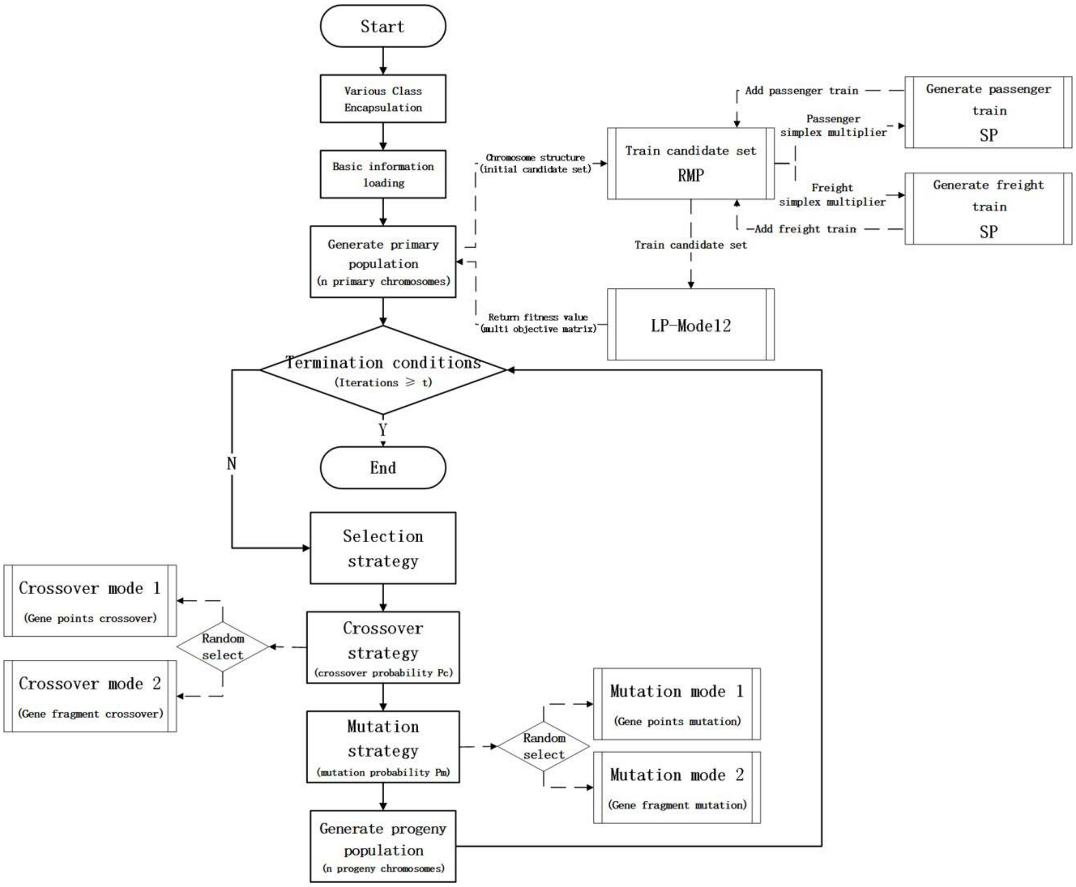


FIGURE 4. The hybrid heuristic algorithm flow.

freight train candidate set generation SP model as Model 5. Through iterative solution, the current optimal training candidate set based on the given initial candidate set is obtained.

Next, several models of the solution part of the training candidate set are discussed in turn.

1) RMP MODEL OF TRAIN CANDIDATE SET

According to the principle of the column generation algorithm, Model 3, based on the column generation algorithm, is constructed. The solution to Model 3 is mainly to select a more appropriate candidate set train for Model 2. Therefore, based on Model 2, the transformation is carried out. The part only related to the decision variables of train operation frequency in Model 2 is retained, and the part related to the decision variables of passenger and freight flow allocation is deleted. The specific structure is as follows.

a: OBJECTIVE FUNCTION PART

Among the three objective functions of Model 2, the HSR train operation cost objective (6) is expressed as the multiplier value of the passenger train and freight train operation frequency and the actual operation cost of a single passenger train and freight train. Only the train operation frequency is related, so it can be directly used as the optimization objective of Model 3. The other two objectives (7) and (8) are related to passenger and freight demand, so they are not considered.

b: CONSTRAINTS SECTION

The section passing capacity constraint (13) and station operation capacity constraint (14) in Model 2 are only related to the decision variables of train operation frequency, so they are directly used in Mode 3.

The meeting passenger and freight demand constraints (51) and (52) in Model 2 are only related to the passenger and freight flow allocation decision variables, so they are directly discarded.

The train carrying capacity constraints (53) and (54) in Model 2, the train passenger and freight flow coupling constraints (49) and (50) are related to the decision variables of train operation frequency and passenger and freight flow allocation, so certain transformation processing is required in Model 3.

① The meeting section flow constraints are transformed from the train carrying capacity constraints.

The mathematical meaning of the train carrying capacity constraints (53) and (54) is that the carrying capacity of each passing train on any section is greater than the actual carrying volume of the train. Since the passenger and freight flow allocation decision of the train is not considered, constraints (53) and (54) are transformed into the meeting section flow constraints.

The size of the section flow depends on the carrying capacity of the train and the section flow demand. That is, for any section, the total OD passenger and freight demand covering the section is less than the sum of the carrying capacity of all trains that can serve the section. The meeting section flow constraints as follows:

$$\sum_{\forall p_{ej}\in P, e\le i, j>i, e\in S/Des, j\in S/Ori} p_{ij} \le \sum_{k_p\in K^P} f_{k_p}\cdot N_P\cdot s_{z_i}^{k_p} \quad \forall z_i\in Z \tag{57}$$

$$\sum_{\forall g_{ej}\in G, e\le i, j>i, e\in S/Des, j\in S/Ori} g_{ij} \le \sum_{k_g\in K^G} f_{k_g}\cdot N_G\cdot s_{z_i}^{k_g} \quad \forall z_i\in Z \tag{58}$$

In terms of mathematical expression, the meeting section flow constraint are to relax the relationship between the flow satisfaction of each train and the carrying capacity of each train, and transform it into the relationship between the flow satisfaction of all passing through the section and the total carrying capacity of all trains serving the section. From the point of view of the constraints on the solution space, the meeting section flow constraints, which reduces the constraints on the solution space.

② The coupling constraints of train flow and passenger and freight flow is transformed into train assignment constraints.

The mathematical meaning of the coupling constraints of train flow and passenger and freight flow is the matching relationship between all HSR trains passing through an OD demand and whether the OD demand flow is allocated to the HSR train. In order to remove the passenger and freight flow allocation decision, the constraints (49), (50) are transformed into train assignment constraints.

The train assignment constraint expresses that when an OD demand exists, there must be a train serving the OD. The expression is as follows:

$$p_{ij} \le \sum_{k_p\in K^P} f_{k_p}\cdot w_{p_{ij}}^{k_p}\cdot \mathrm{M} \quad \forall p_{ij}\in P \tag{59}$$

$$g_{ij} \le \sum_{k_g\in K^G} f_{k_g}\cdot w_{g_{ij}}^{k_g}\cdot \mathrm{M} \quad \forall g_{ij}\in G \tag{60}$$

In terms of mathematical expression, similar to the meeting section flow constraints, the train assignment constraints relax the coupling constraints of train flow and passenger and freight flow to a certain extent, and relax the coupling relationship between the original service OD of each train and the allocated OD on the train to the relationship between an OD demand and the assignment of all trains to that OD service. From the perspective of constraints on solution space, train assignment constraints reduce the constraints on solution space.

c: DECISION VARIABLES SECTION

Only the train operation frequency decision variables (55) and (56) in Model 2 are reserved in the decision variable part.

Complete train candidate set RMP model (Model 3) as follows:

Objective function: Equation (6).

Constraints: Equation (57) - (60), (13), (14).

Decision variables: Equation (55), (56).

The objective of solving the SP of passenger train and freight train candidate set generation is to select trains that meet the conditions for the train candidate set and can further optimize the objective function of the RMP. These added trains also need to have certain attributes such as the origin and destination stations of trains, train operation sections and train stop schemes. It should be noted that in the train generation SP, passenger trains and freight trains are added to their respective train candidate sets, so the SP can be divided into passenger train generation SP and freight train generation SP.

2) PASSENGER TRAIN GENERATION SP MODEL

According to the principle of the column generation algorithm, the inspection number and simplex multiplier obtained by the solution of RMP are taken as the constituent elements of the optimization objective of the passenger train candidate set generation SP, and the constraints related to the passenger train operation conditions before the linearization of the original Model 1 are taken as the constraints of the train candidate set generation SP, and Model 4 is constructed.

a: OBJECTIVE FUNCTION PART

The objective function of the passenger train candidate set generation SP is determined by the inspection number and the simplex multiplier of the RMP. First, it is necessary to clarify the correlation between the RMP and the train candidate set generation SP.

From the perspective of economics, the constraints in RMP are the application constraints on different resources, so each constraint can be written as the following equation:

① Section passing capacity constraint

$$\sum_{k_p\in K^P} f_{k_p}\cdot s_{z_i}^{k_p} + \sum_{k_p\in K^P} \alpha_{z_i}^{k_p}\cdot s_{z_i}^{k_p} + \sum_{k_g\in K^G} f_{k_g}\cdot s_{z_i}^{k_g} + \sum_{k_g\in K^G} \alpha_{z_i}^{k_g}\cdot s_{z_i}^{k_g} = N_{z_i} \quad \forall z_i\in Z \tag{61}$$

② Station operation capacity constraint

$$\sum_{k_p\in K^P} f_{k_p}\cdot x_i^{k_p} + \sum_{k_p\in K^P} \beta_i^{k_p}\cdot x_i^{k_p} + \sum_{k_g\in K^G} f_{k_g}\cdot x_i^{k_g} + \sum_{k_g\in K^G} \beta_i^{k_g}\cdot x_i^{k_g} = N_{S_i} \quad \forall i\in S \tag{62}$$

③ The meeting section flow constraint

$$\sum_{k_p\in K^P} f_{k_p}\cdot N_P\cdot s_{z_i}^{k_p} - \sum_{k_p\in K^P}\rho_{z_i}^{k_p}\cdot N_P\cdot s_{z_i}^{k_p} = \sum_{\forall p_{ej}\in P, e\le i, j>i, e\in S/Des, j\in S/Ori} p_{ij} \quad \forall z_i\in Z \tag{63}$$

④ Train assignment constraint

$$\sum_{k_p\in K^P} f_{k_p}\cdot w_{p_{ij}}^{k_p}\cdot \mathrm{M} - \sum_{k_p\in K^P}\sigma_{p_{ij}}^{k_p}\cdot w_{p_{ij}}^{k_p}\cdot \mathrm{M} = p_{ij} \quad \forall p_{ij}\in P \tag{64}$$

When constructing the passenger train candidate set generation SP, the terms related to freight trains can be regarded as constants.

Among them, $\alpha_{z_i}^{k_p}$ refers to the simplex multiplier of the section passing capacity constraint for passenger train, that is, the "resource" price referred to by the section capacity;

$\beta_i^{k_p}$ refers to the simplex multiplier of the station operation capacity constraint for passenger train, that is, the "resource" price referred to by the station operation capacity;

$\rho_{z_i}^{k_p}$ refers to the simplex multiplier of the meeting section flow constraint for passenger train, that is, the "resource" price referred to by the meeting section flow;

$\sigma_{p_{ij}}^{k_p}$ refers to the simplex multiplier of train assignment constraint for passenger train, that is, the "resource" price referred to by train assignment;

Based on the above simplex multiplier and RMP objective function structure, the objective function of the passenger train candidate set generation SP model is constructed as:

$$\begin{aligned}\min\ (C_{k_p}^1 &+ \sum_{z_i\in Z} C_{k_p}^2\cdot s_{z_i}^{k_p}\cdot zl_i + \sum_{i\in S} C_{k_p}^3\cdot x_i^{k_p}) \\ &+ \sum_{z_i\in Z}\alpha_{z_i}^{k_p}\cdot s_{z_i}^{k_p} + \sum_{i\in S}\beta_i^{k_p}\cdot x_i^{k_p} - \sum_{z_i\in Z}\rho_{z_i}^{k_p}\cdot N_P\cdot s_{z_i}^{k_p} \\ &- \sum_{p_{ij}\in P}\sigma_{p_{ij}}^{k_p}\cdot w_{p_{ij}}^{k_p}\cdot \mathrm{M}\end{aligned} \tag{65}$$

b: CONSTRAINTS SECTION

As mentioned above, the constraints of the train candidate set generation SP select the relevant constraints on the train operation conditions in Model 1. Specifically, it includes constraints on the number of train stops; train mileage; train service OD stop; the number of train origin and destination stations; the sequence of train origin and destination stations; train origin station and destination station stop; train stop range; and train passing section range.

When linearizing the original problem, due to the application of the idea of "candidate set", the constraints and decision variables related to the train operation conditions are regarded as known conditions and not added to the model. Some of these constraints contain the condition constraints of decision variables, which belong to non-linear constraints, so they need to be linearized.

① The train service OD stop constraints (19) are linearized into

$$w_{p_{ij}}^{k_p} \le x_i^{k_p} \quad \forall k_p\in K^P,\ \forall p_{ij}\in P \tag{66}$$

$$w_{p_{ij}}^{k_p} \le x_j^{k_p} \quad \forall k_p\in K^P,\ \forall p_{ij}\in P \tag{67}$$

$$x_i^{k_p} + x_j^{k_p} - w_{p_{ij}}^{k_p} \le 1 \quad \forall k_p\in K^P,\ \forall p_{ij}\in P \tag{68}$$

② The train origin station and destination station stop constraints (27) are linearized into

$$x_i^{k_p} + O_i^{k_p} \le 2 \quad \forall k_p\in K^P,\ \forall i\in S \tag{69}$$

$$O_i^{k_p} \le x_i^{k_p} \quad \forall k_p\in K^P,\ \forall i\in S \tag{70}$$

$$x_i^{k_p} + D_i^{k_p} \le 2 \quad \forall k_p\in K^P,\ \forall i\in S \tag{71}$$

$$D_i^{k_p} \le x_i^{k_p} \quad \forall k_p\in K^P,\ \forall i\in S \tag{72}$$

③ The train stop range constraints (29), (30) are linearized into

$$\sum_{e<i, e\in S} x_e^{k_p} \le (1-O_i^{k_p})\cdot \mathrm{M} \quad \forall k_p\in K^P,\ \forall i\in S \tag{73}$$

$$\sum_{j>i, j\in S} x_j^{k_p} \le (1-D_i^{k_p})\cdot M \quad \forall k_p\in K^P,\ \forall i\in S \tag{74}$$

④ The train passing section range constraints (33) are linearized into

a. The sections between the origin and destination stations of the train line are passing sections

$$s_{z_i}^{k_p} - \sum_{e\le i, e\in S} O_e^{k_p} - \sum_{j>i, j\in S} D_j^{k_p} \ge -1 \quad \forall k_p\in K^P,\ \forall i\in S/Des \tag{75}$$

b. The train must not pass through the section outside the origin and destination stations

$$s_{z_i}^{k_p} \le \sum_{e\le i, e\in S} O_e^{k_p} \quad \forall k_p\in K^P,\ \forall i\in S/Des \tag{76}$$

$$s_{z_i}^{k_p} \le \sum_{j>i, j\in S} D_j^{k} \quad \forall k_p\in K^P,\ \forall i\in S/Des \tag{77}$$

Other constraints related to train operation conditions continue to be used.

c: DECISION VARIABLES SECTION

The decision variables of passenger train candidate set generation SP are directly selected from the stop decision, demand service decision, passing section decision, origin and destination stations decision of passenger train in Model 1.

Complete passenger train candidate set generation SP model as Model 4:

Objective function: Equation (65).

Constraints: Equation (15), (17), (21), (22), (25), (66)-(77).

Decision variables: Equation (39), (41), (43), (45), (46).

3) FREIGHT TRAIN GENERATION SP MODEL

According to the principle of the column generation algorithm, the inspection number and simplex multiplier obtained by the solution of RMP are taken as the constituent elements of the optimization objective of the passenger train candidate set generation SP, and the constraints related to the passenger train operation conditions before the linearization of the original Model 1 are taken as the constraints of the train candidate set generation SP, and Model 4 is constructed.

The freight train candidate set generation SP model, Model 5, is essentially similar to Model 4. It also determines the objective function of the SP by the objective function, inspection number, and simplex multiplier of RMP, and finally confirms the constraint conditions after linearization of the constraints related to the freight train operation conditions. Similar to the passenger train part, the objective function of the SP of freight train candidate set generation is expressed as:

$$\begin{aligned}\min\, (C^1_{k_g} &+ \sum_{z_i\in Z} C^2_{k_g}\cdot s^{k_g}_{z_i}\cdot zl_i + \sum_{i\in S} C^3_{k_g}\cdot x^{k_g}_i) \\ &+ \sum_{z_i\in Z}\alpha^{k_g}_{z_i}\cdot s^{k_g}_{z_i} + \sum_{i\in S}\beta^{k_g}_i\cdot x^{k_g}_i - \sum_{z_i\in Z}\rho^{k_g}_{z_i}\cdot N_G\cdot s^{k_g}_{z_i} \\ &- \sum_{g_{ij}\in G}\sigma^{k_g}_{g_{ij}}\cdot w^{k_g}_{g_{ij}}\cdot \mathrm{M}\end{aligned} \tag{78}$$

Among them, $\alpha^{k_g}_{z_i}$ refers to the simplex multiplier of the section passing capacity constraint for freight train, that is, the "resource" price referred to by the section capacity;

$\beta^{k_g}_i$ refers to the simplex multiplier of the station operation capacity constraint for freight train, that is, the "resource" price referred to by the station operation capacity;

$\rho^{k_g}_{z_i}$ refers to the simplex multiplier of the meeting section flow constraint for freight train, that is, the "resource" price referred to by the meeting section flow;

$\sigma^{k_g}_{p_{ij}}$ refers to the simplex multiplier of train assignment constraint for freight train, that is, the "resource" price referred to by train assignment;

Similar to passenger trains, the linearization of relevant constraints in the generation SP of freight train candidate set is as follows:

① The train service OD stop constraints (20) are linearized into

$$w^{k_g}_{g_{ij}} \le x^{k_g}_i \quad \forall k_g\in K^G,\ \forall g_{ij}\in G \tag{79}$$

$$w^{k_g}_{g_{ij}} \le x^{k_g}_j \quad \forall k_g\in K^G,\ \forall g_{ij}\in G \tag{80}$$

$$x^{k_g}_i + x^{k_g}_j - w^{k_g}_{g_{ij}} \le 1 \quad \forall k_g\in K^G,\ \forall g_{ij}\in G \tag{81}$$

② The train origin station and destination station stop constraints (28) are linearized into

$$x^{k_g}_i + O^{k_g}_i \le 2 \quad \forall k_g\in K^G,\ \forall i\in S \tag{82}$$

$$O^{k_g}_i \le x^{k_g}_i \quad \forall k_g\in K^G,\ \forall i\in S \tag{83}$$

$$x^{k_g}_i + D^{k_g}_i \le 2 \quad \forall k_g\in K^G,\ \forall i\in S \tag{84}$$

$$D^{k_g}_i \le x^{k_g}_i \quad \forall k_g\in K^G,\ \forall i\in S \tag{85}$$

③ The train stop range constraints (31), (32) are linearized into

$$\sum_{e<i,e\in S} x^{k_g}_e \le (1-O^{k_g}_i)\cdot M \quad \forall k_g\in K^G,\ \forall i\in S \tag{86}$$

$$\sum_{j>i,j\in S} x^{k_g}_j \le (1-D^{k_g}_i)\cdot M \quad \forall k_g\in K^G,\ \forall i\in S \tag{87}$$

④ The train passing section range constraints (34) are linearized into

a. The sections between the origin and destination stations of the train line are passing sections

$$\begin{aligned} s^{k_g}_{z_i} - \sum_{e\le i,e\in S} O^{k_g}_e - \sum_{j>i,j\in S} D^{k_g}_j \ge -1 \quad \forall k_g\in K^G,& \\ \forall i\in S/Des& \end{aligned} \tag{88}$$

b. The train must not pass through the section outside the origin and destination stations

$$s^{k_g}_{z_i}\ \text{ł}\ e \sum_{e\le i,e\in S} O^{k_g}_e \quad \forall k_g\in K^G,\ \forall i\in S/Des \tag{89}$$

$$s^{k_g}_{z_i} \le \sum_{j>i,j\in S} D^{k_g}_j \quad \forall k_g\in K^G,\ \forall i\in S/Des \tag{90}$$

Other constraints related to train operation conditions continue to be used.

Complete freight train candidate set generation SP model that Model 5:

Objective function: Equation (78).

Constraints: Equation (16), (18), (23), (24), (26), (79)-(90).

Decision variables: Equation (40), (42), (44), (47), (48).

D. MULTI OBJECTIVE SOLVING PROCESS AND FITNESS SETTING

According to the principle of the column generation algorithm, the inspection number and simplex multiplier obtained by the solution of RMP are taken as the constituent elements of the optimization objective of the passenger train candidate set generation SP, and the constraints related to the passenger train operation conditions before the linearization of the original Model 1 are taken as the constraints of the train candidate set generation SP, and Model 4 is constructed.

The Priority method is a common solution method for multi-objective mathematical programming model. Its solution idea is to solve multiple objective functions one by one according to the degree of importance or priority. The Model 1 and Model 2 studied in this paper include three objective functions: minimizing the operation cost of HSR passenger trains and freight trains; total passenger travel time; and total freight transport time. With the assumption of sufficient transportation capacity and direct transportation, the difference in passenger and freight transportation times is mainly reflected in the number of stops on the way of the train. The difference is relatively small in essence. The optimization of the line planning of high-speed passenger trains and freight trains studied in this paper aims to fully attract the passenger and freight demand between service nodes by running high-speed passenger trains and freight trains at the same time on the basis of meeting the passenger and freight demand; maximizing the utilization efficiency of high-speed rail resources; and then improving the operating income of high-speed rail trains. Therefore, it is considered that the goal of minimizing the operating cost of high-speed rail, as well as the goal of maximizing the operating income of high-speed rail, is the highest priority.

At the same time, considering the principle of passenger transport priority, it is necessary to ensure that passenger travel services are as little affected by high-speed rail freight services as possible. Therefore, the minimum passenger travel time goal can be considered a priority over the minimum cargo transport time goal. Finally, after meeting the minimum high-speed rail operation cost and the minimum passenger travel time in turn, the minimum cargo transport time goal can be met. Therefore, this paper takes the candidate set based Model 2 as the multi objective processing object and uses the priority method to decompose the multi objective candidate set based Model 2 into three single objective optimization models as follows, which are solved one by one.

Model 6:

Objective function: Equation (6).

Constraints: Equation (49) - (54), (13), (14).

Decision variables: Equation (55), (56), (37), (38).

Model 7:

Objective function: Equation (7).

Constraints: Equation (49) - (54), (13), (14), and

$$\begin{aligned}&\sum_{k_p\in K^P} f_{k_p}\cdot(C^1_{k_p}+\sum_{z_i\in Z} C^2_{k_p}\cdot s^{k_p}_{z_i}\cdot zl_i+\sum_{i\in S} C^3_{k_p}\cdot x^{k_p}_i)\\&\quad+\sum_{k_g\in K^G} f_{k_g}\cdot(C^1_{k_g}+\sum_{z_i\in Z} C^2_{k_g}\cdot s^{k_g}_{z_i}\cdot zl_i+\sum_{i\in S} C^3_{k_g}\cdot x^{k_g}_i)\\&\leq Z_1'*\end{aligned}\tag{91}$$

Decision variables: Equation (55), (56), (37), (38).

Model 8:

Objective function: Equation (8).

Constraints: Equation (49) - (54), (13), (14), (91) and

$$\begin{aligned}&\sum_{\forall p_{ij}\in P}\sum_{k_p\in K^P} q^{k_p}_{p_{ij}}\cdot w^{k_p}_{p_{ij}}\cdot(\sum_i^{j-1} s^{k_p}_{z_i}\cdot zl_i/v_{k_p}\cdot 60\\&\quad+(\sum_i^{j-1} x^{k_p}_i-2)\cdot t^{k_p}_{stop}+(\sum_i^{j-1} x^{k_p}_i-1)\cdot t^{k_p}_{ss})\leq Z_2*\end{aligned}\tag{92}$$

Decision variables: Equation (55), (56), (37), (38).

Z_1'∗ and Z_2∗ are the optimal target values obtained from Model 6 and Model 7 respectively. Using the priority method to solve the multi objective Model 2, the steps are as follows: First, solve the objective function (6), that is, optimize the Model 6, and get the minimum operating cost Z_1'∗ of HSR passenger trains and freight trains; Then, with the condition of considering the new constraints (91), the objective function (7), that is Model 7, is solved separately to obtain the minimum value Z_2∗ of the total travel time of passengers; Finally, with the condition of considering adding constraints (92), solve the objective function (8), that is, Model 8, and obtain the optimal target value Z_3∗ of Model 8.

The three objective function values of Model 2 obtained by the priority method are expressed as the objective matrix solution [Z_1'∗, Z_2∗, Z_3∗], which constitutes the fitness values of chromosomes in genetic algorithm for subsequent comparison and selection.

E. THE ALGORITHM STEPS OF TRAIN CANDIDATE SET SOLUTION AND FITNESS VALUE CALCULATION

The core part of the hybrid heuristic algorithm is the train candidate set solution and fitness value calculation. The key solution models of these two parts are described in front, and the algorithm design details of these two parts are described in detail here.

By constructing the train candidate set based on column generation algorithm to solve the RMP model and the passenger train and freight train generation SP model, a group of screened and scaled candidate train sets with operation sections, origin and destination stations, OD service conditions, and stop schemes can be obtained. Next, taking this train candidate set as the known condition, the train operation frequency, passenger and freight flow allocation scheme and three objective function values are further determined through Model 6, Model 7 and model8 decomposed by the priority method, and the three objective function values of the model are formed into a matrix form according to the priority as the fitness value of the chromosome in the genetic algorithm.

The algorithm steps of train candidate set solution and fitness matrix calculation are shown in the following Figure 5:

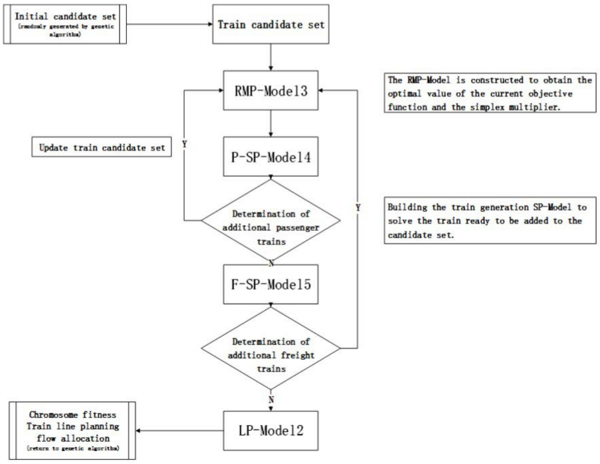


FIGURE 5. The algorithm steps of train candidate set solution and fitness matrix calculation.

Step 1 (Initial Train Candidate Set Generation of Passenger and freight): The trains in the initial candidate set randomly generated by the chromosome of the genetic algorithm are deduplicated, and the trains that do not repeat and have the attributes of stop scheme, origin and destination stations, operation section, mileage, stop numbers, and operation cost are added to the initially generated passenger train and freight train candidate sets. Turn to Step 2.

Step 2 (Train Candidate Set RMP Solving): The current passenger train and freight train candidate sets are substituted into Model 3, and the integer programming model is solved by using the branch and bound method by calling CPLEX to obtain the current optimal objective function value, the constrained simplex multiplier, and the related inspection number of the model. Turn to Step 3.

Step 3 (Passenger Train Generation SP Solving): The simplex multipliers related to passenger trains in the constraints obtained by Step 2 are substituted into Model4, and the 0-1 integer programming model is solved by using CPLEX. Turn to Step 4.

Step 4 (Judgment of Adding Train to Passenger Train Candidate Set): Judge the solution result of Step 3. If "① there is no solution to Step 3, ② the value of the objective function obtained from the solution is greater than or equal to 0 (the original problem is a minimization problem), that is, the inspection number of the original problem meets the conditions of the optimal solution, and ③ the passenger train with the attributes of stop scheme, origin and destination stations, operation section, mileage, stop numbers, operation cost, etc. obtained from the solution exists in the original train candidate set," one of the three judgment conditions is satisfied. Then it shows that the train candidate set can no longer optimize the objective function value of RMP by adding new passenger train types. At this time, turn to Step 5. If the three determination conditions are not met, add the new passenger train solved by Step 3 to the passenger train and freight train candidate sets, update the current train candidate set, and return to Step 2.

Step 5 (Freight Train Generation SP Solving): The simplex multipliers related to freight trains in the constraints obtained by Step 2 are substituted into Model 5, and the 0-1 integer programming model is solved by using CPLEX. Turn to Step 6.

Step 6 (Judgment of Adding Train to Freight Train Candidate Set): Judge the solution result of Step 5. If "① there is no solution to Step 5, ② the value of the objective function obtained from the solution is greater than or equal to 0, that is, the inspection number of the original problem meets the conditions of the optimal solution, and ③ a group of freight trains with stop scheme, origin and destination stations, operation section, mileage, stop numbers, operation cost and other attributes obtained from the solution exist in the original train candidate set", one of the three judgment conditions is satisfied, Then it shows that the train candidate set can no longer optimize and the objective function value of RMP by adding new freight train types. At this time, turn to Step 7; If the three determination conditions are not met, add the new freight train solved by Step 5 to the passenger train and freight train candidate set, update the current train candidate set, and return to Step 2.

Step 7 (Line Planning, Flow Allocation Scheme and Fitness Value Solution): At this time, the update of the passenger train and freight train candidate set is over, and the objective function of the RMP cannot be further optimized by adding new passenger or freight train types. Therefore, the current optimal passenger train and freight train candidate set starting from the initial passenger train and freight train candidate set is obtained. The current optimal passenger train and freight train candidate set is substituted into Model 6, Model 7, model 8 in turn, and the branch and bound method is called to solve it, the priority matrix of the current optimal objective function value, the passenger train and freight train line planning and flow allocation scheme are obtained. Turn to Step 8.

Step 8 (Return Fitness Matrix, Line Planning and Flow Allocation Scheme): The optimal target value matrix obtained from the solution is assigned to the fitness matrix of the chromosome of the initial candidate set, and the line planning and flow allocation scheme are saved in the chromosome attribute. The train candidate set, line planning, flow allocation scheme and fitness value solution are completed. It should be noted that due to the fitness matrix represents the multi objective function value matrix with priority, the three objective functions have a strong correlation, and the high priority objective function directly affects the solution of the low priority objective function, when comparing and selecting the best fitness value in the genetic algorithm, the objective-1 function can be directly selected for comparison until the final objective-1 function is no longer further optimized or reaches the maximum number of iterations, the subsequent objective function are solved.

F. HYBRID HEURISTIC ALGORITHM STRUCTURE

The hybrid heuristic algorithm based on column generation and genetic algorithm as a whole includes 8 modules: algorithm main module, chromosome coding module, initial candidate set generation module, train candidate set solution module, train line planning and flow allocation scheme solution module, selection operation module, crossover operation module, mutation operation module, etc. The algorithm main module is the structure of the overall solution process of the algorithm, which is designed according to the algorithm flow. The following describes the seven modules respectively.

1) CHROMOSOME CODING MODULE

The purpose of chromosome coding module is to construct chromosome classes and clarify the structural characteristics of each chromosome. The problem of passenger train and freight train line planning solved in this paper requires that all passenger and freight OD demand are met. When the number of HSR stations on the line is $|S|$, assuming that there is OD demand between any two stations, there are a total of $|S| \cdot (|S| - 1)/2$ OD pairs that need train service. Without considering that the same train serves multiple OD, a total of

$|S| \cdot (|S| - 1)/2$ trains need to be selected to correspond to OD one by one.

Combined with the requirements of train service OD stop constraint, the train corresponding to the OD demand must stop at the OD start and end station. Therefore, this paper takes "OD-train, station" as the coding object to construct a two-dimensional chromosome structure. The scale of the two-dimensional matrix is $(|S| \cdot (|S| - 1)/2) \cdot |S|$. "OD-train, station" two-dimensional matrix represents the stop scheme that can serve the trains corresponding to different OD demands. The specific value represents the stop selection of a "OD-train" at the corresponding station. Take the two-dimensional matrix of Figure 6 as an example.

$$\begin{array}{c|cccc} & \text{Sta0} & \text{Sta1} & \text{Sta2} & \text{Sta3} \\ \hline \text{0-1} & 1 & 1 & 0 & 0 \\ \text{0-2} & 1 & 1 & 1 & 0 \\ \text{0-3} & 1 & 0 & 1 & 1 \\ \text{1-2} & 0 & 1 & 1 & 0 \\ \text{1-3} & 1 & 1 & 0 & 1 \\ \text{2-3} & 0 & 1 & 1 & 1 \end{array} \Rightarrow$$

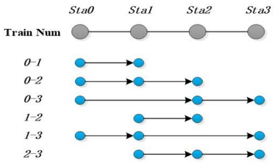


FIGURE 6. An example of "OD-train, station" two-dimensional matrix.

Among them, the two-dimensional matrix of chromosome structure is expressed as the parking scheme matrix of 6 "OD-trains" at 4 stations. Take the first behavior as an example, OD demand "0-1" is served by "OD-train" "0-1" which the train stop scheme is [1,1,0,0].

The main structure of chromosome coding module is a two-dimensional matrix composed of OD demand number and station number. Because the problems solved by the line planning include passenger trains and freight trains, and the passenger and freight demands are different in the types of service trains, this paper takes the three-dimensional "passenger and freight, OD-train, station" as the coding object to construct the chromosome main structure. According to the number of stations on the solved line $|S|$, the size of the three-dimensional matrix array is determined to be $2 \cdot (|S| \cdot (|S| - 1)/2) \cdot |S|$, In addition, it also includes the expression of the initial candidate set of the chromosome, the candidate set of the train after chromosome solution, the fitness value of the chromosome, the line planning and flow allocation scheme corresponding to the chromosome, and other attributes.

Among them, the specific "OD-train" stop scheme value and chromosome initial candidate set in the three-dimensional matrix should be determined by the initial candidate set generation module, the train candidate set corresponding to the chromosome should be determined by the train candidate set solution module, and the chromosome fitness value and line planning and flow allocation scheme corresponding to the chromosome should be determined by the line planning and flow allocation solution module.

2) INITIAL CANDIDATE SET GENERATION MODULE

The initial candidate set generation module is mainly to compile the specific value of the three-dimensional "passenger and freight, OD-train, station" matrix for the new chromosome generated in the population, and determine the initial candidate set information for the chromosome to further obtain the fitness value. Therefore, the module structure can be divided into two parts: chromosomal structure assignment and initial candidate set determination.

a: CHROMOSOME STRUCTURE ASSIGNMENT PART

The construction mechanism of the three-dimensional "passenger and freight, OD-train, station" matrix has been described in the chromosome coding module, so the assignment operation of the three-dimensional matrix is essentially to determine the stop scheme of the "OD-train". Firstly, considering the differences between passenger trains and freight trains in train attributes, the three-dimensional matrix is decomposed into two-dimensional passenger matrix and two-dimensional freight matrix. Secondly, judge whether the demand exists, that is, whether there is passenger demand or freight demand between the two stations. If the OD demand does not exist, it is not necessary to compile the stop scheme of the service train for the OD, and set all the values of the train stop scheme column corresponding to the OD as "0". Then, following the requirements of train service OD stop constraint, the current "OD-train" must stop at the current OD start and end station, and the corresponding station location is taken as "1", and other stations are randomly taken as "0" or "1", and so on, the stop scheme value for each "OD-train" of the passenger and freight two-dimensional matrix can be completed respectively.

b: INITIAL CANDIDATE SET DETERMINATION PART

The determination of the initial candidate set is different from the train list after the chromosome assignment. After the chromosome assignment, there are the number of OD demand the stop scheme of trains determined. In order to ensure the integrity of the chromosome structure, the generated trains are not processed. However, the candidate set only saves the train sequences with different stop schemes. Therefore, the chromosome train list needs to be de-duplicated.

3) TRAIN CANDIDATE SET SOLUTION MODULE

The specific train candidate set solution module compilation method is Step 1 to Step 6 of Chapter 4.E above, which is not repeated here.

4) TRAIN LINE PLANNING AND FLOW ALLOCATION SCHEME SOLUTION MODULE

The specific solution module of line planning and flow allocation scheme module compilation method is Step 7 of Chapter 4.E above, which is not repeated here.

5) SELECTION STRATEGY MODULE

The selection operation is to determine the relevant information of each chromosome in each generation population, take the current population as the parent population, follow the set selection rules, and select a certain number of chromosomes from the parent population to pass on to the offspring, so as to form the offspring population after crossover, mutation and other evolution.

In this algorithm, the optimal reservation and roulette are used to select the parent population.

a: SELECTION STRATEGY MODULE 1

First, the chromosomes with the optimal fitness matrix in the parent population are saved into the offspring population list.

b: SELECTION STRATEGY MODULE 2

Secondly, the roulette method is adopted. The number of comparisons and selections of each Roulette is set to m ($m \geq 4$), and a uniformly distributed pseudo-random r in the [0, 1] range is generated. According to the value of r, the corresponding chromosome is selected to enter the roulette waiting list, and the two chromosomes with the best fitness value are selected from the list of m waiting chromosomes to enter the waiting list of crossover and mutation.

c: SELECTION STRATEGY MODULE 3

After the two chromosomes that waiting for crossover and mutation are selected to complete the crossing and mutation operations, the newborn chromosomes are saved into the list of candidate chromosomes of offspring, and the chromosomes in the list are called train candidate set solution module and line planning and flow allocation scheme solution module for fitness calculation, and the chromosome with the best fitness value is selected to be saved into the population of offspring.

6) CROSSOVER STRATEGY MODULE

The two chromosomes waiting for crossover selected by roulette in the parent population that meet the determination of crossover probability are crossed. According to the chromosome structure, two crossover methods, random gene site crossover and random gene fragment crossover, are designed.

a: RANDOM GENE SITE CROSSOVER

For two chromosomes to be crossed, set n gene sites to be crossed, in which the gene sites are the number positions of ''OD-train'', and the n gene site of the two chromosomes are crossed and exchanged correspondingly.

b: RANDOM GENE FRAGMENT CROSSOVER

For two chromosomes to be crossed, set two gene fragment endpoints to be crossed, in which the gene fragment is the ''OD-train'' gene position contained in the endpoint number position of the two ''OD-trains'', which corresponds to the cross exchange of the gene fragments of the two chromosomes.

Crossover result judgment. In order to ensure that the chromosomes of the two offspring after crossing are different from those of the parent, it is necessary to call the initial candidate set generation module to determine the initial candidate set of the two chromosomes after crossing.

If the initial candidate sets of chromosomes of the two offspring after crossing are different from those of the parent, it indicates that two new offspring chromosomes have been generated, and the chromosomes of the two offspring are saved in the list to be mutated; If there is only one initial candidate set of chromosome of two offspring after crossing, which is different from the two parent chromosomes, it means that only one new offspring chromosome is generated, which is saved in the list to be mutated, and the other chromosome that is the same as the initial candidate set of the parent is eliminated. If the initial candidate sets of chromosomes of the two offspring after crossing are the same as that of the parent, the chromosomes of the two offspring will be eliminated. The selected cross gene sites at this time will be saved in the cross failure list, and the cross gene sites or gene fragments will be re-selected. When the randomly generated cross gene sites or fragments are not in the cross failure list, the cross will be carried out according to the gene sites or fragments.

If the crossover failure list reaches the maximum limit, exit the crossover operation module and return to the selection operation module to select two new parent chromosomes to be crossed by roulette and re-enter this module.

If there is at least one chromosome to be mutated in the list to be mutated after the crossover operation, the crossover is successful, and the crossover operation module ends.

7) MUTATION STRATEGY MODULE

The chromosomes in the list to be mutated generated by crossover that meet the determination of mutation probability are mutated. According to the chromosome structure, two mutation methods, random gene site mutation and random gene fragment mutation, are designed.

a: RANDOM GENE SITE MUTATION

For the chromosome to be mutated, set n gene sites to be mutated, follow the basic stop requirements of gene site ''OD-train'', and regenerate a different train stop scheme from that before mutation.

b: RANDOM GENE FRAGMENT CROSSOVER

For the chromosome to be mutated, set two endpoints of the gene fragment to be mutated, follow the basic stop requirements of ''OD-train'' in the gene fragment, and regenerate a train stop scheme different from that before the mutation.

Mutation result judgment. Similar to the crossover operation module, it is also necessary to judge the mutation results of the offspring chromosomes after mutation, and will not be repeated here.

G. SOLVING STEPS OF HYBRID HEURISTIC ALGORITHM

According to the flow of hybrid heuristic algorithm and the design structure of each module, the overall solution steps of hybrid heuristic algorithm are summarized as follows:

Step 1 (Various Structural Packaging): Package station, section, OD demand, passenger train, freight train and chromosome.

Step 2 (Basic Information Loading): Through data reading, load station information (station name, station serial number), section information (front-end station, back-end station, section distance), OD demand information (demand front-end station, demand back-end station, demand passenger flow, demand cargo flow, OD distance).

Step 3 (Generation of Primary Population): Initialize various data of genetic algorithm, determine population size N, overall iteration times T, crossover probability P_c, mutation probability P_m, roulette selection scale m.

According to the size of the primary population, N initial chromosome classes are randomly generated. For each chromosome, the ''initial candidate set generation module'', ''train candidate set solution module'' and ''line planning and allocation scheme solution module'' are called respectively to obtain N primary chromosomes with three-dimensional ''passenger and freight, OD-train, station'' matrix and fitness values to form the primary population.

Step 4 (Determination of Termination Conditions): Judge whether the current iteration number reaches the maximum iteration number T.

If the maximum number of iterations T is reached, the optimal chromosome output of the algorithm is selected from the optimal chromosomes of each generation, and the algorithm ends.

If the maximum number of iterations T is not reached, turn to Step 5.

Step 5 (Select Strategy Compilation): Select the current population as the parent population.

Call ''selection strategy module 1'' to save the parent's optimal chromosome into the offspring population, turn to Step 5-1.

Step 5-1 (Select Chromosomes to Cross and Mutate): Call ''selection strategy module 2'', select the two parent chromosomes to be crossed and mutated through roulette, and substitute them into Step 6.

Step 5-2 (Offspring Population Addition): Call ''selection strategy module 3'' and add it into the offspring population from the list of candidate chromosomes of the offspring according to the fitness value, turn to Step 5-3.

Step 5-3 (Determination of Progeny Population): If the number of chromosomes in the progeny population reaches the population size, turn to Step 8.

If the number of chromosomes in the progeny population has not reached the population size, return to Step 5-1.

Step 6 (Crossover Strategy Compilation): Cross the chromosomes of two parents to be crossed.

According to the crossover probability P_c, if it is necessary to crossover, call the ''crossover operation module'' to cross the chromosomes of the current two parents, and add the chromosomes to be mutated of the offspring after successful crossover to the list to be mutated (also called the list of successful crossover). If there is no need to cross, the chromosomes of the two parents will remain unchanged, and the chromosomes of the two parents will be added to the list of to be mutated, turn to Step 7.

Step 7 (Mutation Strategy Compilation): The chromosomes in the list to be mutated are mutated.

According to the mutation probability P_m, if it is necessary to mutate, call the ''mutation operation module'' to mutate the current chromosome to be mutated, and add the mutated chromosome to the offspring candidate chromosome list (also called list of successful mutations); If there is no need to mutate, the chromosome to be mutated will be kept unchanged and directly added to the list of candidate chromosomes of offspring (also called list of successful mutations). Finally, get the list of candidate chromosomes of offspring, and return to Step 5-2.

Step 8 (Generation of Offspring Population): Update the current offspring population to the current population, and update the number of iterations. Return to Step 4.

Through the above steps, the whole process of designing and solving the hybrid heuristic algorithm based on column generation and genetic algorithm is completed. This method combines the classical genetic algorithm with column generation algorithm. Following the principle that the passenger and freight demand OD must be fully served, the method randomly generates a certain number of initial candidate sets through the genetic algorithm, uses column generation algorithm to solve the train candidate set, and call the CPLEX to solve the optimal target value matrix, train line planning and flow allocation scheme of the HSR train line planning with the passenger and freight transport coordination, The optimal target value matrix is used as the fitness value of each chromosome in the genetic algorithm, and then according to the iterative evolution of the genetic algorithm, the optimal chromosome is selected as the output of the final approximate optimal solution after generation by generation. Then the HSR train line planning with passenger and freight transport coordination can be obtained.

V. CASE DESIGN AND SOLUTION COMPARISON

In order to verify the effectiveness of the various models and algorithms constructed in this paper, this part provides validation and comparative analysis for the above mathematical optimization model and solution algorithm design examples.

A. CASE DESIGN

As shown in Figure 7, this part takes a one-way Dalian-Harbin HSR line consisting of six stations connected in northeast China as the background. The distance between stations on the line is known, and the passenger and freight OD demand between stations in the unit time cycle is statistically derived from the 12306 China Railway passenger ticket service system and the express market transportation demand forecast, both of which are given as initial known conditions.

At the same time, various parameters such as the operating line conditions, the parameters related to train operation and the cost of passenger and freight transport are pre-defined. Refer to Appendix 1. In particular, in order to facilitate the example solving calculations, it is assumed that the capacity of passing operations per unit cycle is the same for each section and station, and that the technical parameters for each

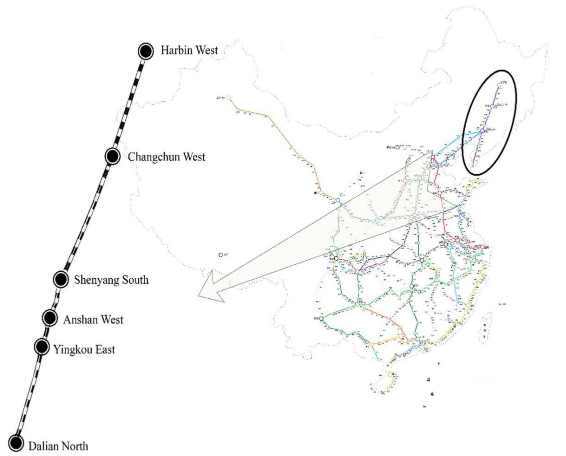


FIGURE 7. The one-way Dalian-Harbin HSR line.

passenger train of the HSR are the same for goods trains. The same applies to freight trains.

This part solves the above example on a computer configured with an Intel(R) Core(TM) i7-8750H CPU @ 2.20 GHz 2.21 GHz, 8 GB RAM and Windows 10 by calling the CPLEX 12.8.0.0 engine in Python.

B. ANALYSIS OF DIRECT SOLUTION OF THE MODEL

In the solution of this example, set the gap termination range to 0.01% and the time termination range to 600 seconds, call the CPLEX solver through Python language to solve, and get the results including passenger train line planning, freight train line planning, passenger train flow allocation scheme, freight train flow allocation scheme and so on. The specific contents of the results and scheme are as follows.

Based on the initial train set generated above, the final train line planning can be obtained directly by solving the model. The passenger train line planning scheme is shown in Table 3 and Figure 8.

TABLE 3. Train set of HSR passenger trains.

PT No.	Train OD	Intermediate stops	Frequency
p_1	*Yingkou-Changchun*	*Anshan*	*1*
p_2	*Dalian-Harbin*	—	*1*
p_3	*Dalian-Harbin*	*Changchun*	*1*
p_4	*Dalian-Harbin*	*Shenyang*	*2*
p_5	*Dalian-Harbin*	*Yingkou, Changchun*	*1*
p_6	*Yingkou-Harbin*	*Shenyang*	*1*
p_7	*Dalian- Changchun*	*Anshan, Shenyang*	*2*

According to the above passenger train operation diagram, the operation train set based on passenger OD demand includes 7 kinds of trains, and a total of 9 passenger trains are required to operate.

The freight train line planning scheme is shown in Table 4 and Figure 9.

As shown in the above freight train operation diagram, the operation train set based on freight OD demand includes 6 kinds of trains, and a total of 6 freight trains are required to operate.

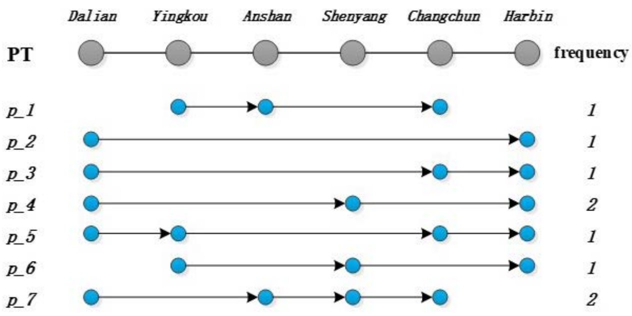


FIGURE 8. Train set of HSR passenger trains.

TABLE 4. Train set of HSR freight trains.

FT No.	Train OD	Intermediate stops	Frequency
f_1	*Yingkou-Harbin*	*Shenyang*	*1*
f_2	*Anshan-Shenyang*	—	*1*
f_3	*Yingkou-Harbin*	*Anshan*	*1*
f_4	*Dalian- Changchun*	*Shenyang*	*1*
f_5	*Dalian-Harbin*	*Anshan, Changchun*	*1*
f_6	*Dalian-Harbin*	*Yingkou, Changchun*	*1*

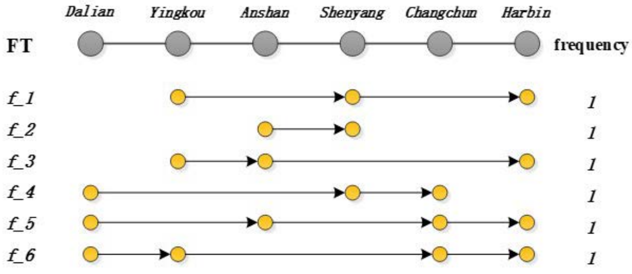


FIGURE 9. Train set of HSR freight trains.

The corresponding allocation of HSR passenger trains is shown in Figure 10 below.

For passenger trains with the same train operating frequency greater than 1, the service passenger flow is reasonably allocated to each train. The passenger flow of each line section is shown in Figure 11.

It can be seen from Figure x that the passenger flow of each line section reflects the saving of passenger train operation cost and the satisfaction of OD demand. The operation plan of HSR passenger trains can ensure that most passenger trains have a high occupancy rate (more than 88%), so as to ensure passenger transport revenue on the basis of meeting passenger demand.

The flow allocation scheme of HSR freight trains is shown in Figure 12.

The freight flow of each line section is shown in Figure 13.

Therefore, the solution of the HSR train line planning model with passenger and freight transport coordination is completed, and its objective function value matrix is [1997704, 718338.9, 58770], that is, through the above line

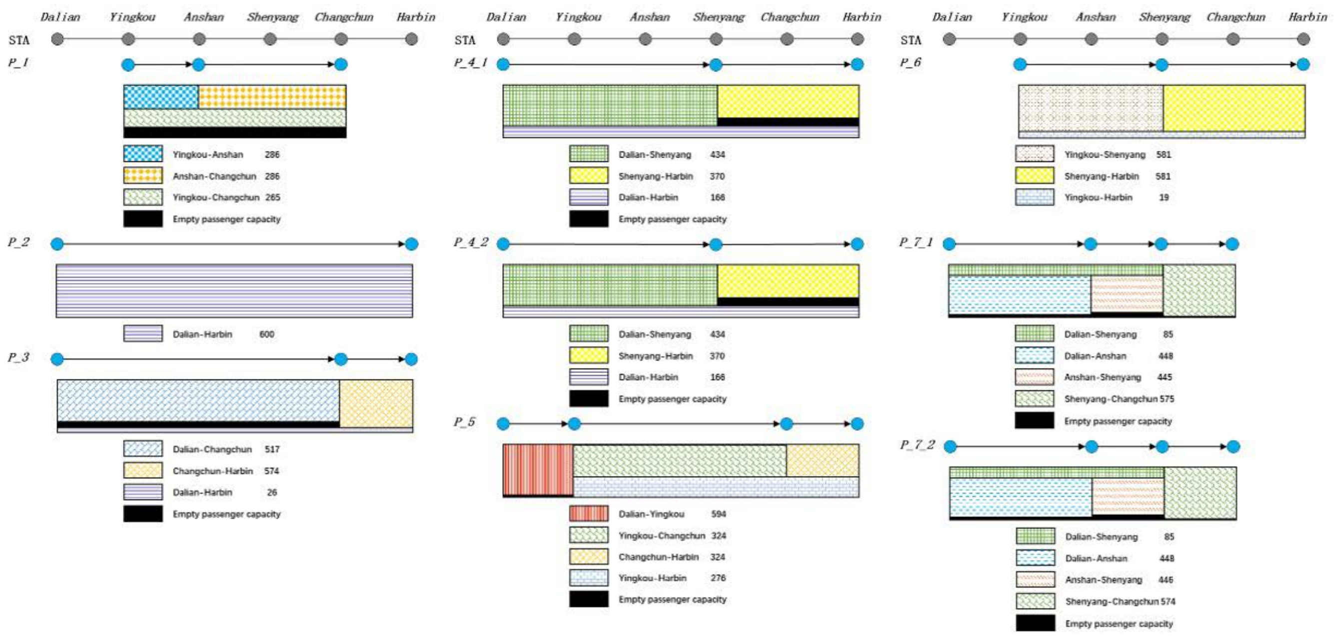


FIGURE 10. The flow allocation of HSR passenger trains.

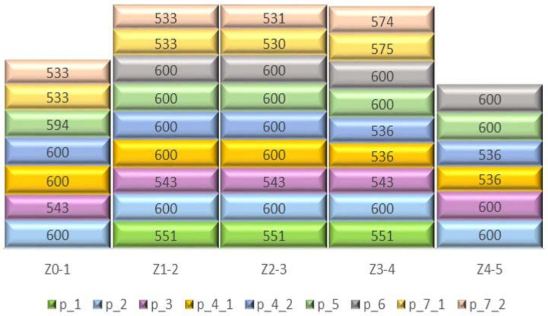


FIGURE 11. The passenger flow of each line section.

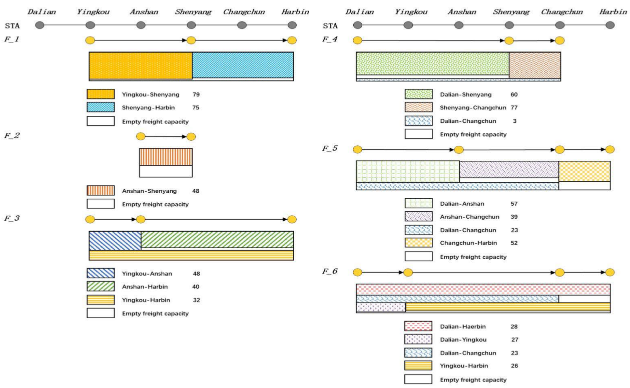


FIGURE 12. The flow allocation of HSR freight trains.

planning and flow allocation scheme, all passenger and freight transport demand on the case line in the current research cycle will be served. The total operating cost of HSR passenger trains and freight trains (excluding freight loading and unloading cost) is 1,997,704 ¥/ unit cycle, and the total travel time of all passengers is 718,338.9 man-minute / unit cycle, the total transportation time of goods is 58,770 ton-minute / unit cycle.

More specifically, the total operating profit of HSR passenger train and freight train service per unit cycle can reach 4,233,382.8 ¥, of which the total revenue of passenger and freight demand service is 6,261,646.8 ¥. The total operating cost of passenger train is 1,267,200 ¥, the average operating profit of a single HSR passenger train is about 310,000 ¥, and the total operating cost of freight train is 759,984 ¥, the average operating profit of a single HSR freight train is about 230,000 ¥.

C. HYBRID HEURISTIC ALGORITHM SOLUTION ANALYSIS

In this part, the hybrid heuristic algorithm based on column generation and genetic algorithm is used to solve the HSR train line planning with passenger and freight transport coordination, and the solution is carried out according to the solution process described above.

The algorithm initialization settings are shown in Table 5.

Through the iterative calculation of the hybrid heuristic algorithm, based on the objective-1 function value, the generation process of each chromosome and optimization trends of all generated chromosomes are obtained, as shown in the Figure 14 below.

The optimal chromosome objective-1 function value in the 32nd generation population reached 1,997,704, which is the

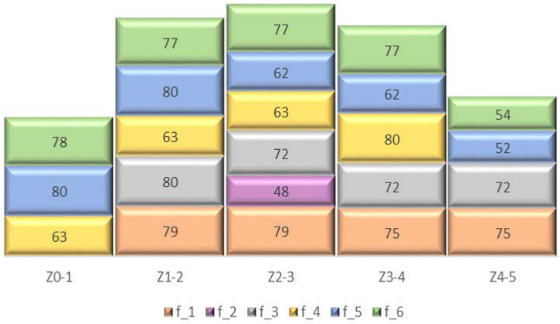


FIGURE 13. The freight flow of each line section.

TABLE 5. Algorithm initial parameters.

N // Population size	10
T // Total iterations	50
P_c // Crossover probability	0.8
P_m // Mutation probability	0.5
m // Roulette selection scale	4

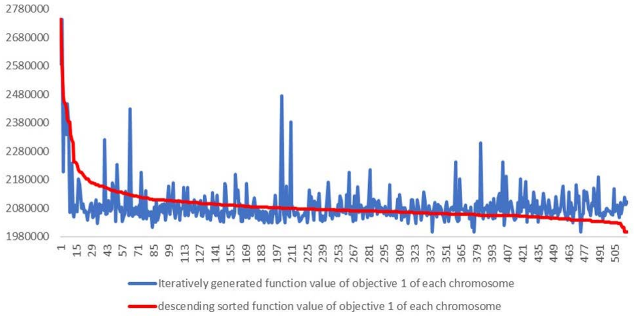


FIGURE 14. The variation trend of the objective-1 function value of all generated chromosomes.

same as the direct solution of the model, but the specific line planning is different. The information about the approximate optimal solution finally generated is as follows:

1) INITIAL PASSENGER TRAIN AND FREIGHT TRAIN SET

In the initial train set, there are 14 passenger trains and 12 freight trains shown in Table 6 and Table 7.

2) TRAIN CANDIDATE SET AFTER COLUMN GENERATION SOLVING

Through the solution of column generation algorithm, the updated candidate set of passenger and freight trains is obtained, including 21 passenger trains and 18 freight trains, 7 passenger trains and 6 freight trains are added to the train candidate set. The added trains are shown in Table 8.

3) TRAIN LINE PLANNING AND FLOW ALLOCATION SCHEME

At this time, 39 passenger trains and freight trains are substituted into the linear models Model 6, Model 7 and Model 8 in turn. After excluding the invalid stops from trains, the approximate optimal train line planning and flow allocation scheme are obtained as follows.

TABLE 6. The stop scheme of initial passenger train set.

PT ID	*Dalian*	*Yingkou*	*Anshan*	*Shenyang*	*Changchun*	*Harbin*
P0	1	1	1	1	0	0
P1	1	1	0	1	0	0
P2	1	1	1	1	1	1
P3	1	0	0	1	1	0
P4	1	1	1	0	1	1
P5	0	1	1	1	1	0
P6	1	0	0	0	1	1
P7	1	1	0	1	1	1
P8	1	0	1	1	1	0
P9	1	1	1	0	1	0
P10	1	0	1	1	1	1
P11	0	0	1	1	1	1
P12	0	1	0	1	0	1
P13	1	1	1	1	0	1

TABLE 7. The stop scheme of initial freight train set.

FT ID	*Dalian*	*Yingkou*	*Anshan*	*Shenyang*	*Changchun*	*Harbin*
F0	1	1	1	1	1	0
F1	1	0	1	0	1	1
F2	1	0	1	1	1	1
F3	1	0	1	1	1	0
F4	0	1	1	1	1	0
F5	1	1	1	1	1	1
F6	0	1	1	0	1	1
F7	0	0	1	0	1	1
F8	0	0	1	1	1	1
F9	1	1	0	1	1	1
F10	0	0	1	1	0	0
F11	0	1	0	1	1	1

TABLE 8. The stop scheme of added train.

Train ID	*Dalian*	*Yingkou*	*Anshan*	*Shenyang*	*Changchun*	*Harbin*
P14	0	1	0	0	0	1
P15	1	0	1	1	1	0
P16	0	1	0	0	1	1
P17	1	0	0	0	0	1
P18	0	1	0	0	1	0
P19	1	0	0	1	0	1
P20	1	0	0	0	1	0
F12	1	1	0	0	1	1
F13	1	0	0	0	0	1
F14	0	1	0	1	0	1
F15	0	1	0	0	1	0
F16	0	1	0	0	0	1
F17	0	0	1	0	0	1

The passenger train line planning scheme is shown in Table 9 and Figure 15.

According to the above passenger train operation diagram, the operation train set based on passenger OD demand includes 9 kinds of trains, and a total of 9 passenger trains are required to operate.

The freight train line planning scheme is shown in Table 10 and Figure 16.

As shown in the above freight train operation diagram, the operation train set based on freight OD demand includes

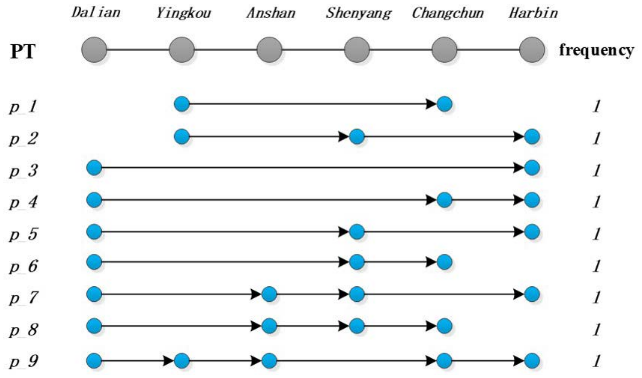


FIGURE 15. The train set of HSR passenger trains.

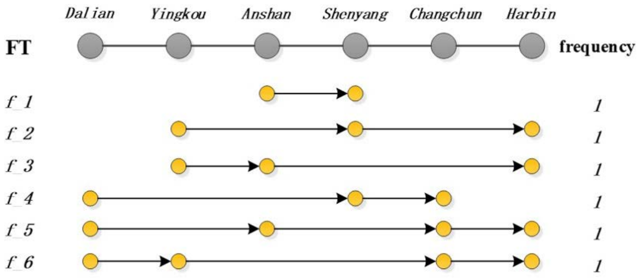


FIGURE 16. The train set of HSR freight trains.

TABLE 9. Train set of HSR passenger trains.

PT No.	Train OD	Intermediate stops	Frequency
p_1	*Yingkou-Changchun*	—	*1*
p_2	*Yingkou-Harbin*	*Shenyang*	*1*
p_3	*Dalian-Harbin*	—	*1*
p_4	*Dalian-Harbin*	*Changchun*	*1*
p_5	*Dalian-Harbin*	*Shenyang*	*1*
p_6	*Dalian- Changchun*	*Shenyang*	*1*
p_7	*Dalian-Harbin*	*Anshan, Shenyang*	*1*
p_8	*Dalian- Changchun*	*Anshan, Shenyang*	*1*
p_9	*Dalian-Harbin*	*Yingkou, Anshan, Changchun*	*1*

TABLE 10. Train set of HSR freight trains.

FT No.	Train OD	Intermediate stops	Frequency
f_1	*Anshan-Shenyang*	—	*1*
f_2	*Yingkou-Harbin*	*Shenyang*	*1*
f_3	*Yingkou-Harbin*	*Anshan*	*1*
f_4	*Dalian- Changchun*	*Shenyang*	*1*
f_5	*Dalian-Harbin*	*Anshan, Changchun*	*1*
f_6	*Dalian-Harbin*	*Yingkou, Changchun*	*1*

6 kinds of trains, and a total of 6 freight trains are required to operate.

The corresponding allocation of HSR passenger trains is shown in Figure 17 below.

The passenger flow of each line section is shown in Figure 18.

The operation plan of HSR passenger trains can ensure that most passenger trains have a high occupancy rate (more than 76%), so as to ensure passenger transport revenue on the basis of meeting passenger demand.

The flow allocation scheme of HSR freight trains is shown in Figure 19.

The freight flow of each line section is shown in Figure 20.

Therefore, the solution of the HSR train line planning model with passenger and freight transport coordination is completed, and its objective function value matrix is [1997704, 718382.9, 58770], that is, through the above line planning and flow allocation scheme, all passenger and freight demand on the case line in the current research cycle will be served. The total operating cost of HSR passenger trains and freight trains is 1,997,704 ¥/unit cycle, and the total travel time of passengers is 718,382.9 person-minute/unit cycle, the total transportation time of goods is 58,770 ton-minute/unit cycle.

D. COMPARATIVE ANALYSIS OF DIRECT SOLUTION AND HYBRID HEURISTIC ALGORITHM SOLUTION

From the solution results in the previous two part, it can be seen that the model and algorithm of HSR train line planning with passenger and freight transport coordination constructed in this paper are effective and feasible, and feasible passenger train and freight train line planning and flow allocation scheme can be obtained. There are differences between the direct solution of the model and the iterative solution of heuristic algorithm in the solution process, solution efficiency and solution results. The following compares and analyzes the two methods.

1) SOLUTION PROCESS

In terms of the solution process, the core step of the two methods is to solve the linearized HSR train line planning model with passenger and freight transport coordination by decomposing the objective function according to priority and using the CPLEX solver in turn.

The main difference between the two methods lies in the composition of train candidate set based on the idea of "candidate set". The direct solution of the model is to include all trains that meet the conditions of basic stop numbers and mileage into the candidate set. The number of trains stored in the candidate set of trains is $2 \cdot (2^{|S|} - |S| - 1) - C$, where C is the number of stopping modes that do not meet the conditions of stop numbers and mileage. In the iterative solution of hybrid heuristic algorithm, the initial train candidate set is composed of randomly generated trains that meet the conditions of basic stop numbers and mileage and can serve all OD. The number of trains stored in the initial train candidate set is $|S| \cdot (|S| - 1)$ (equivalent to OD pairs) at most. Since various randomly generated trains often serve multiple OD at the same time, after removing the initial trains of the same category, the number of trains in the initial train set is greatly reduced, and then the updated train candidate set is obtained through column generation algorithm, and the number of stored trains is far less than the number of the direct solution of the model.

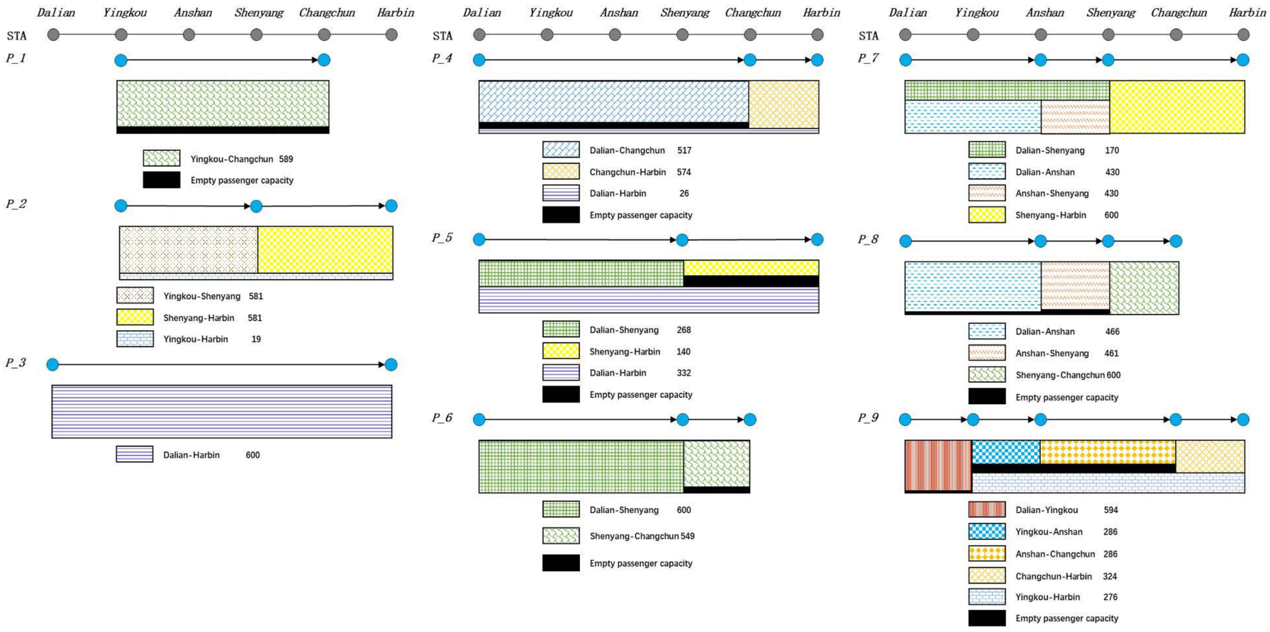


FIGURE 17. The flow allocation of HSR passenger trains.

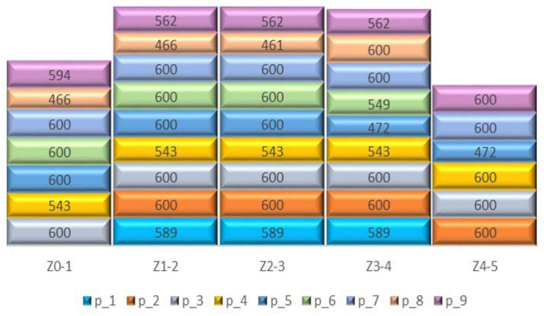


FIGURE 18. The passenger flow of each line section.

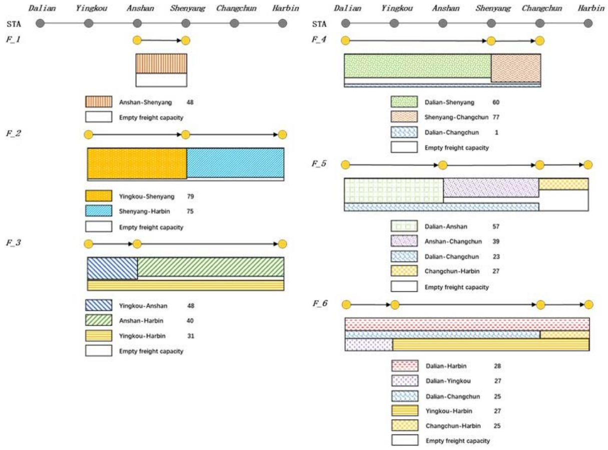


FIGURE 19. The flow allocation of HSR freight trains.

According to the solution of the case, in the hybrid heuristic algorithm, different initial train candidate sets can be solved to get different line planning. Therefore, it needs many

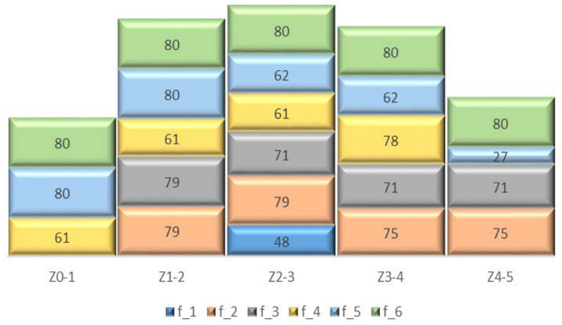


FIGURE 20. The freight flow of each line section.

iterations to gradually find the initial train candidate set that can meet the optimization requirements of the final line planning. The step-by-step iterative process is the more complex part of the hybrid heuristic algorithm compared with the direct solution of the model.

2) SOLUTION EFFICIENCY

In terms of solution efficiency, both methods are solved by Python calling CPLEX 12.8 engine programming. When the number of stations in the line is 6, both methods can obtain the optimal solution and near optimal solution with the certain convergence accuracy (gap=0.005) and convergence time (time=600). With the increase of the number of stations, the solution efficiency of the two methods gradually differs.

It should be noted that when calling the CPLEX 12.8.0.0 engine to program and solve through Python language, for the hybrid heuristic algorithm, if any initial train candidate set can be solved, the problem must be solvable

TABLE 11. Solution results for different methods.

		Direct solution of model		Random initial train candidate sets		iterative solution of heuristic algorithm:
Station Number	Solving times[a]	Time/s	Gap	Time/s	Gap	(8, 4, 10)[b]
7 stations	Linear	1.07	—	1.01	—	
	First	603.767	0.54%	1.93	0.39%	137 chromosomes were solved
	Second	605.82	0.87%	2.01	0.43%	Time=652.68s
	Third	603.47	0.57%	1.98	0.20%	Once every 4.76 seconds on average
	Fourth	603.43	0.52%	1.55	0.42%	
8 stations	Linear	4.91	—	2.14	—	
	First	626.19	0.66%	2.94	0.46%	137 chromosomes were solved
	Second	610.24	0.52%	2.68	0.32%	Time=680.50s
	Third	104.65	0.48%	2.66	0.47%	Once every 4.96 seconds on average
	Fourth	622.09	0.56%	3.73	0.45%	
9 stations	Linear	26.94	—	2.15	—	
	First	623.51	0.57%	3.84	0.48%	137 chromosomes were solved
	Second	623.77	0.65%	3.03	0.41%	Time=699.72s
	Third	623.13	0.59%	3.28	0.42%	Once every 5.11 seconds on average
	Fourth	623.31	0.65%	4.11	0.48%	
10 stations	Linear	196.62	—	4.29	—	
	First	749.68	0.69%	4.68	0.47%	145 chromosomes were solved
	Second	419.53	0.48%	5.96	0.43%	Time=866.12s
	Third	746.97	0.55%	3.86	0.44%	Once every 5.97 seconds on average
	Fourth	764.53	0.56%	16.20	0.41%	
11 stations	Linear	1767.74	—	5.42	—	
	First	2182.77	2.51%	6.90	0.45%	145 chromosomes were solved
	Second	2266.89	3.31%	5.10	0.36%	Time=870.93s
	Third	2201.85	3.44%	5.49	0.50%	Once every 6 seconds on average
	Fourth	2358.24	1.40%	7.10	0.43%	

[a](Linear, First, Second, Third, Fourth) means linear optimal solution and 4-integer optimal solution; [b](8, 4, 10) means population size 8, selection size 4, iteration times 10.

after setting the population size and iteration times of the genetic algorithm. Therefore, the solution result of chromosome fitness values of a group of initial train candidate sets randomly generated in the hybrid heuristic algorithm is measured simultaneously. With the increase of the number of stations, the initial conditions such as section distances and OD demand are randomly set, and the solution efficiency of different methods is compared with the certain convergence accuracy (gap=0.005) and convergence time (time=600 seconds). The solution results are shown in the following Table 11.

Combined with the calculation results of the case and the calculation results of randomly setting the initial conditions after adding the stations in the table 9, it can be seen that taking the six HSR stations in the example as the line background, the direct solution of the model and the iterative solution of hybrid heuristic algorithm can meet the convergence accuracy and convergence time limits, and with the increase of the number of stations, the change in solving efficiency between the two is shown in Figure 21 below.

For the direct solution of the model method, when the number of stations reaches 7, the time to solve the linear optimal solution with relaxed integer constraints is very short, only about 1 second. After adding integer constraints, it is impossible to obtain the approximate integer optimal solution with compound convergence accuracy (gap= 0.005) with the limitation of convergence time (time=600) like 6 stations, and the convergence speed of its branch and bound is significantly reduced. When the convergence time reaches about 600s, its convergence accuracy is between [0.52%, 0.87%]; When the number of stations is 11, the output time of the linear optimal solution increases to about 1768 seconds. It can be seen that the solution efficiency of the linear optimal solution obtained by the direct solution of the model increases exponentially with the increase of the number of stations, and the convergence speed of the integer solution within the convergence time range is also significantly reduced, and the convergence accuracy range is between [1.40%, 3.44%].

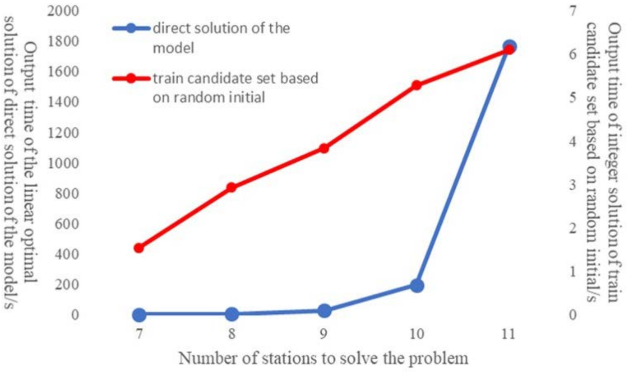


FIGURE 21. Solving efficiency of the two methods.

Compared with the direct solution of the model method, the method which the initial train candidate set generated randomly and updated through column generation algorithm, that offers significant improvements in solution efficiency. When the number of stations reaches 7, the output linear optimal solution takes only about 1 second. After adding integer constraints, the branch and bound convergence speed is extremely fast. The total solution takes about 1.5 to 2 seconds to obtain an integer solution with convergence accuracy between [0.20%, 0.43%]; When the number of stations is 11, the output time of the linear optimal solution is about 5 seconds, and the time of solving the linear optimal solution by the column generation method is also

TABLE 12. **The final solution results of different solutions.**

	Direct solution of model	Initial train set of all stops	Initial train set of all direct train	Algorithm (10, 6, 50)
Solving times	1	1	1	516
Total stops of PT	29	34	34	29
Total mileage of PT	7175	7615	7615	7175
Total stops of FT	19	22	21	19
Total mileage of FT	4059	4342	4342	4059
Train operation cost	1997704	2108112	2099312	1997704
Total travel time of passengers	718338.9	725126.9	746069.1	718382.9
Total transportation time of goods	58770	58650	58410	58770

slowly increasing, but the convergence speed of the branch and bound solution is still very fast. The total solution time is [5.10, 7.10] seconds, and the output integer solution convergence accuracy can reach [0.36%, 0.50%]. It shows that the sharp increase in the number of stations will significantly improve the solution time of the linear optimal solution, but the convergence speed of using branch and bound method to solve integer solution is still fast.

For the iterative solution of heuristic algorithm method, the solution time is mainly related to the number of solutions and solution time of column generation. The change of the solution time of column generation with the increase of the number of stations is as described above, and the number of solutions are mainly related to the population size, selection number and iteration times initially set by the genetic algorithm. Therefore, through the control of the number of solutions, the generation by generation optimization integer solution can be obtained within a certain solution time. According to the calculation, when the number of stations is 7, the fitness value can be solved once in about 4.76 seconds in the genetic algorithm structure. With the increase of the number of stations, the calculation time increases slightly. When the number of stations is 11, the average fitness value can be solved once in 6 seconds.

Therefore, from the perspective of solution efficiency, when the number of stations on the line is small, the approximate integer optimal solution can be quickly obtained by using the direct solution of the model or the iterative solution of hybrid heuristic algorithm. With the increase of the number of stations, the solution efficiency of the iterative solution of hybrid heuristic algorithm is significantly better than that of the direct solution of the model. Therefore, facing the design problem of train line planning on HSR lines with large scale and large number of stations, using hybrid heuristic algorithm can get the train line planning more efficiently.

3) SOLUTION RESULTS

There are also some differences in the final solution results of different solutions. This part compares the results of the direct solution of the model and the hybrid heuristic algorithm in the case context, as shown in the following Table 12.

It can be seen from the above table 10 that for the line planning design problem with the scale of the example, the direct solution of the model results are the best, while the difference between the direct solution results and the results of using column generation algorithm to solve the initial train set composed of stop trains at each station and all direct trains is only [5.53%, 0.94%, −0.20%] and [5.09%, 3.86%, −0.61%], The error of the target matrix is small and the efficiency of solution is much higher than that of direct solution.

The final objective-1 function value of hybrid heuristic algorithm (10, 6, 50) is consistent with the direct solution result in the 32nd generation; The difference between the value of the objective-2 function and the direct solution result is only 44 person-minutes / unit period, and the error is within 0.006%; The value of the objective-3 function is also consistent with the direct solution.

Although the optimal target value matrix is basically the same, from the results of the optimal line planning and flow allocation scheme obtained by the above two methods, there are many different line planning with the nearly same target value matrix results. Therefore, it can be seen that the current line planning and flow allocation scheme are the optimal but not the only solution.

4) COMPARISON AND SUMMARY OF TWO METHODS

Through the comparison between the direct solution of the model and the iterative solution of hybrid heuristic algorithm in the aspects of solution process, solution efficiency, solution results, etc., when the scale of model solution is small, the direct solution after linear transformation of the problem model by using the solution engine such as CPLEX can get a better development plan. With the gradual increase of solution scale, the efficiency of direct solution is significantly reduced, and even a group of feasible solutions cannot be output in a very long solution time; The hybrid heuristic algorithm based on column generation and genetic algorithm designed in this paper can quickly and efficiently solve the HSR train line planning problem. When the scale of the problem is small, the line planning with minimal error compared with the direct solution can be solved for any given initial train set. With the expansion of the scale of the problem, the efficiency advantage of its solution gradually highlights, compared with the direct solution, the hybrid heuristic algorithm can output the approximate integer optimal solution with the optimal solution within the set convergence accuracy and the certain solution time, and its applicability is stronger.

VI. CONCLUSION

This paper analyzes in detail the optimization problem of HSR train line planning based on the passenger and freight transport coordination. Firstly, a multi objective nonlinear mixed integer programming model of HSR train line planning based on the passenger and freight transport coordination is proposed, and the train candidate set is introduced to linearize the model. Secondly, a hybrid heuristic algorithm combining column generation algorithm and genetic algorithm is designed to solve the problem. Then the experimental case is designed to verify the validity of the model, the efficiency of the algorithm, and the adaptability of the two solution methods in this paper.

The experimental case study of the Dalian-Harbin line of the China's HSR shows that, within a certain scale of the problem, the direct solution of the equivalence model based on train candidate set can obtain a globally optimal solution for the line planning and flow allocation scheme. The hybrid heuristic algorithm can obtain a number of different globally optimal solutions for the line planning and flow allocation scheme by setting the appropriate population size, selection number, and iteration time. As the scale of the problem increases, the efficiency advantage of the hybrid heuristic algorithm gradually becomes more and more evident, with its optimization-seeking capability and convergence efficiency ensuring that a satisfactory set of solutions can be generated within a certain time limit.

The design problem of HSR train line planning with passenger and freight transport coordination studied in this paper belongs to a part of the HSR train transportation organization planning problem with passenger and freight transport coordination. Taking this problem as a starting point, we can further study the train timetable problem and vehicle scheduling problem with passenger and freight transport coordination. At the same time, the research scenario in this paper is set for HSR lines, where passenger and freight demand can be met with sufficient capacity. For HSR lines where capacity resources are tight and the demand for passenger and freight transport cannot be fully satisfied. The corresponding model and algorithm for the optimization problem in the scenario of ''following the principle of giving priority to passenger transport''. It can be further developed by adopting the basic conditions that meet all passenger demands and considering the maximum satisfaction of freight demands. In addition, optimization of the actual HSR network can be considered in future application scenarios.

APPENDIX

Line data, demand data and parameter data are shown in Table 13, Table 14, Table 15 and Table 16.

TABLE 13. **The station and section information.**

	Dalian	*Yingkou*	*Anshan*	*Shenyang*	*Changchun*	*Harbin*
Sta Dis	0km	206km	283km	383km	687km	921km
Sec Dis		206km	77km	100km	304km	234km

TABLE 14. **Passenger flow OD demand (persons/unit cycle).**

Demand	*Yingkou*	*Anshan*	*Shenyang*	*Changchun*	*Harbin*
Dalian	594	896	1038	517	958
Yingkou		286	581	589	295
Anshan			891	286	0
Shenyang				1149	1321
Changchun					898

TABLE 15. **Freight flow OD demand (persons/unit cycle).**

Demand	*Yingkou*	*Anshan*	*Shenyang*	*Changchun*	*Harbin*
Dalian	27	57	60	49	28
Yingkou		48	79	0	58
Anshan			48	39	40
Shenyang				77	75
Changchun					52

TABLE 16. **Relevant parameters of the calculation example.**

Symbol	Value	Symbol	Value	Symbol	Value
C_{p_1}	80	$C_{k_g}^4$	40	$N_{dis}^{ubk_p}$	1000
C_{p_2}	0.8	N_{z_i}	30	$N_{dis}^{lbk_g}$	100
$C_{k_p}^1$	64000	N_{S_i}	30	$N_{dis}^{ubk_g}$	1000
$C_{k_p}^2$	64	N_P	600	v_{k_p}	350
$C_{k_p}^3$	8000	N_G	80	v_{k_g}	300
C_{g_1}	600	$N_{stop}^{lbk_p}$	2	$t_{stop}^{k_p}$	3
C_{g_2}	6	$N_{stop}^{ubk_p}$	10	$t_{stop}^{k_g}$	5
$C_{k_g}^1$	56000	$N_{stop}^{lbk_g}$	2	$t_{ss}^{k_p}$	1
$C_{k_g}^2$	56	$N_{stop}^{ubk_g}$	10	$t_{ss}^{k_g}$	1
$C_{k_g}^3$	8800	$N_{dis}^{lbk_p}$	100	M	7000000

REFERENCES

[1] A. A. Assad, ''Modelling of rail networks: Toward a routing/makeup model,'' *Transp. Res. B, Methodol.*, vol. 14, nos. 1–2, pp. 101–114, 1980.

[2] M. Bi, S. He, and W. Xu, ''Express delivery with high-speed railway: Definitely feasible or just a publicity stunt,'' *Transp. Res. A, Policy Pract.*, vol. 120, pp. 165–187, Feb. 2019.

[3] U. Brännlund, P. O. Lindberg, A. Nõu, and J. E. Nilsson, ''Railway timetabling using Lagrangian relaxation,'' *Transp. Sci.*, vol. 32, no. 4, pp. 358–369, 1998.

[4] M. R. Bussieck, T. Winter, and U. T. Zimmermann, ''Discrete optimization in public rail transport,'' *Math. Program.*, vol. 79, no. 1, pp. 415–444, Oct. 1997.

[5] M. R. Bussieck, T. Lindner, and M. E. Lübbecke, ''A fast algorithm for near cost optimal line plans,'' *Math. Methods Oper. Res.*, vol. 59, no. 2, pp. 205–220, Jun. 2004.

[6] X. Chen, T. Zuo, M. Lang, S. Li, and S. Li, ''Integrated optimization of transfer station selection and train timetables for road–rail intermodal transport network,'' *Comput. Ind. Eng.*, vol. 165, Mar. 2022, Art. no. 107929.

[7] X. Chen, P. Zhou, M. Lang, X. Yu, and S. Li, ''HSR express cargo flow allocation and operation organization optimization under varying demand,'' *J. Transp. Syst. Eng. Inf. Technol.*, vol. 8, no. 18, pp. 1–18, 2022.

[8] H. Fu, L. Nie, H. Yang, and L. Tong, ''Research on the method for optimization of candidate train set based train operation plans for HSRs,'' *J. China Railway Soc.*, vol. 32, no. 6, pp. 1–8, 2010.

[9] H. Fu, L. Nie, L. Meng, B. R. Sperry, and Z. He, ''A hierarchical line planning approach for a large-scale high speed rail network: The China case,'' *Transp. Res. A, Policy Pract.*, vol. 75, pp. 61–83, May 2015.

[10] L. Gao, M. Lang, Q. Wang, and X. Li, "Technical and economic feasibility for operating dedicated HSR mail trains in China," *Logistics Technol.*, vol. 30, no. 9, pp. 4–6, 2011.

[11] J.-W. Goossens, S. van Hoesel, and L. Kroon, "A branch-and-cut approach for solving railway line-planning problems," *Transp. Sci.*, vol. 38, no. 3, pp. 379–393, Aug. 2004.

[12] J.-W. Goossens, S. van Hoesel, and L. Kroon, "On solving multi-type railway line planning problems," *Eur. J. Oper. Res.*, vol. 168, no. 2, pp. 403–424, Jan. 2006.

[13] B. Han and S. Ren, "Optimizing stop plan and tickets allocation for high-speed railway based on uncertainty theory," *Soft Comput.*, vol. 24, no. 9, pp. 6467–6482, May 2020.

[14] W. Huang and B. Shuai, "Using improved entropy-cloud model to select high-speed railway express freight train service sites," *Math. Problems Eng.*, vol. 2017, pp. 1–13, Jan. 2017.

[15] W. Jin, X. Li, L. Zhou, and X. Yu, "Research on optimization of HSR freight transportation organization scheme based on column generation algorithm," *J. China Railway Soc.*, vol. 42, no. 9, pp. 26–32, 2020.

[16] H. Li, C. Tian, S. Zhang, and Y. Jiang, "Calculation method for carrying capacity of mixed passenger and freight railway based on improved rotor model," *China Railway Sci.*, vol. 42, no. 3, pp. 144–155, 2021.

[17] L. L. Li, H. Yan, and H. Liu, "Operation scheme optimization of HSR freight EMU," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 17, no. 1, pp. 94–100, 2019.

[18] S. Li, H. Lv, M. Lv, C. Xu, and S. Ni, "Daily dynamic freight train service optimization," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 20, no. 5, pp. 177–184, 2020.

[19] 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.

[20] Z. Li, A. Shalaby, M. J. Roorda, and B. Mao, "Urban rail service design for collaborative passenger and freight transport," *Transp. Res. E, Logistics Transp. Rev.*, vol. 147, Mar. 2021, Art. no. 102205.

[21] X.-H. Liang, K.-H. Tan, A. Whiteing, C. Nash, and D. Johnson, "Parcels and mail by high speed rail—A comparative analysis of Germany, France and China," *J. Rail Transp. Planning Manage.*, vol. 6, no. 2, pp. 77–88, Sep. 2016.

[22] X.-H. Liang and K.-H. Tan, "Market potential and approaches of parcels and mail by high speed rail in China," *Case Stud. Transp. Policy*, vol. 7, no. 3, pp. 583–597, Sep. 2019.

[23] B. Lin, Z. Wang, S. Ni, and Y. Zhao, "Research on influence of locomotive routing in train formation plan optimization," *J. China Railway Soc.*, vol. 43, no. 10, pp. 1–11, 2021.

[24] D.-Y. Lin and Y.-H. Ku, "Using genetic algorithms to optimize stopping patterns for passenger rail transportation," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 29, no. 4, pp. 264–278, Apr. 2014.

[25] D.-Y. Lin and Y.-H. Ku, "An implicit enumeration algorithm for the passenger service planning problem: Application to the Taiwan railways administration line," *Eur. J. Oper. Res.*, vol. 238, no. 3, pp. 863–875, Nov. 2014.

[26] Z. Lin and Q. Yu, "Discussion on fast freight transportation by HSR," *Railway Econ. Res.*, vol. 2012, no. 4, pp. 1–5, 2012.

[27] L. Liu and M. Dessouky, "A decomposition based hybrid heuristic algorithm for the joint passenger and freight train scheduling problem," *Comput. Oper. Res.*, vol. 87, pp. 165–182, Nov. 2017.

[28] L. Liu and M. Dessouky, "Stochastic passenger train timetabling using a branch and bound approach," *Comput. Ind. Eng.*, vol. 127, pp. 1223–1240, Jan. 2019.

[29] F. López-Ramos, E. Codina, Á. Marín, and A. Guarnaschelli, "Integrated approach to network design and frequency setting problem in railway rapid transit systems," *Comput. Oper. Res.*, vol. 80, pp. 128–146, Apr. 2017.

[30] Y. Lu, M. Lang, X. Yu, and S. Li, "A sustainable multimodal transport system: The two-echelon location-routing problem with consolidation in the Euro–China expressway," *Sustainability*, vol. 11, no. 19, p. 5486, Oct. 2019.

[31] 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.

[32] 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.

[33] M. H. Miandoab, V. Ghezavati, and D. Mohammaditabar, "Developing a simultaneous scheduling of passenger and freight trains for an inter-city railway considering optimization of carbon emissions and waiting times," *J. Cleaner Prod.*, vol. 248, Mar. 2020, Art. no. 119303.

[34] A. Patz, "Die richtige auswahl von verkehrslinien bei groβen straβenbahnnetzen," *Verkehrstechnik*, vol. 50, no. 51, pp. 977–983, 1925.

[35] J. A. Pazour, R. D. Meller, and L. M. Pohl, "A model to design a national high-speed rail network for freight distribution," *Transp. Res. A, Policy Pract.*, vol. 44, no. 3, pp. 119–135, Mar. 2010.

[36] D. Plotkin, "Carrying freight on HSR lines," *J. Transp. Eng.*, vol. 123, no. 3, pp. 199–201, 1997.

[37] J. Qi, L. Yang, Z. Di, S. Li, K. Yang, and Y. Gao, "Integrated optimization for train operation zone and stop plan with passenger distributions," *Transp. Res. E, Logistics Transp. Rev.*, vol. 109, pp. 151–173, Jan. 2018.

[38] J. Qi, L. Yang, Z. Di, S. Li, K. Yang, and Y. Gao, "Optimization methods of combined passenger and freight transportation based on flexible train composition mode," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 22, no. 2, pp. 197–205, 2022.

[39] R. Elbert and J. Rentschler, "Freight on urban public transportation: A systematic literature review," *Res. Transp. Bus. Manage.*, vol. 2021, Jun. 2021, Art. no. 100679.

[40] A. Talebian and B. Zou, "Integrated modeling of high performance passenger and freight train planning on shared-use corridors in the US," *Transp. Res. B, Methodol.*, vol. 82, pp. 114–140, Dec. 2015.

[41] Y. Wang and X. Ma, "Two-stage shipping scheme of HSR logistics considering the change of the train piggyback capacity," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 20, no. 2, p. 9, 2022.

[42] 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.

[43] F. Yan and R. M. P. Goverde, "Combined line planning and train timetabling for strongly heterogeneous railway lines with direct connections," *Transp. Res. B, Methodol.*, vol. 127, pp. 20–46, Sep. 2019.

[44] Y. Ye, Y. Zhou, and L. Yang, "Optimization research on the operation frequency of intercity trains under different targets," *J. Tongji Univ., Natural Sci.*, vol. 46, no. 4, pp. 472–477, 2018.

[45] X. Yu, M. Lang, Y. Gao, K. Wang, C.-H. Su, S.-B. Tsai, M. Huo, X. Yu, and S. Li, "An empirical study on the design of China high-speed rail express train operation plan—From a sustainable transport perspective," *Sustainability*, vol. 10, no. 7, p. 2478, Jul. 2018.

[46] X. Yu, M. Lang, L. Zhou, D. Wang, and X. Yu, "Calculation method of the freight flow sharing rate of HSR express products," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 19, no. 1, pp. 40–44, 2019.

[47] X. Yu, M. Lang, W. Zhang, S. Li, M. Zhang, and X. Yu, "An empirical study on the comprehensive optimization method of a train diagram of the China high speed railway express," *Sustainability*, vol. 11, no. 7, p. 2141, Apr. 2019.

[48] C. Zhang, J. Qi, L. Yang, Y. Gao, and Z. Gao, "Robust train operation plan based on uncertain passenger demands for HSR corridors," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 22, no. 1, pp. 115–123, 2022.

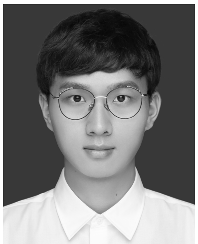
SHIQI LI received the Bachelor of Engineering degree in engineering from Beijing Jiaotong University, Beijing, China, where he is currently pursuing the Ph.D. degree.

His research interests include the area of multi modal transportation and logistics management. He has done a lot of in-depth research in these areas.

MAOXIANG LANG (Member, IEEE) is currently a Professor, a Doctoral Tutor, and the Deputy Director of the Department of Logistics Engineering, Beijing Jiaotong University, and a Special Researcher of the China Logistics Association. He has been engaged in research in the fields of transportation logistics management, and modern freight transportation technology and management.

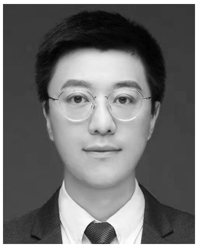
SIYU LI received the Bachelor of Engineering degree in engineering from Dalian Jiaotong University, Shenyang, China. She is currently pursuing the Ph.D. degree with Beijing Jiaotong University, Beijing, China.

Her research interests include the area of multi modal transportation and logistics management. She has done a lot of in-depth research in these areas.



XINGHAN CHEN (Graduate Student Member, IEEE) is currently pursuing the Ph.D. degree with Beijing Jiaotong University, Beijing, China.

His research interests include transportation planning and management, optimization of logistics networks, and intelligent transportation systems.



XUEQIAO YU received the Ph.D. degree from Beijing Jiaotong University, Beijing, China.

His research interests include transportation planning and management, and modern freight transportation technology and management.

YIXUAN GENG (Graduate Student Member, IEEE) is currently pursuing the Ph.D. degree with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, China.

His research interests include machine learning, fault diagnosis, and point clouds processing.

• • •