Journal of Computers Vol. 36 No. 2, April 2025, pp. 263-279
https://doi.org/10.63367/199115992025043602018

Optimization Method for Urban Rail Transit Transportation Organization Mode Based on Virtual Grouping

Xiu-Juan Chang[1], Fei An[2], and Dong-Mei Guo[1*]

[1] HeBei Vocational College of Rail Transportation,
Shijiazhuang City 052165, Hebei Province, China
{changxiujuan9587, anfei2021888, gdm88765}@163.com

[2] HeBei Jiaotong Vocational and Technical College,
Shijiazhuang City 050035, Hebei Province, China

Received 29 March 2025; Revised 7 April 2025; Accepted 16 April 2025

Abstract. The continuous development of urban rail transit networks has highlighted the mismatch between passenger demand and transportation capacity, necessitating more flexible transportation organization models. To achieve a more agile train operation organization, this paper employs "virtual grouping" technology to reduce running intervals and increase density, thereby maximizing line capacity, enhancing the service level of urban rail transit, and facilitating greater flexibility in various transportation organization modes, such as long and short routes, fast and slow trains, and Y-shaped routes. Initially, a transportation organization mode for virtual grouping trains was devised, leading to the establishment of an optimization model for the operation plan of large and small routes based on virtual grouping. Subsequently, considering the operation of large and small routes, virtual grouping, and stop-stop modes, the peak period of a specific urban rail transit line was selected as the research object. The objective was to minimize the weighted sum of passenger waiting time, vehicle mileage, and the number of online train sets. Decision variables included the starting and ending points of large and small routes, as well as their departure frequency. Taking into account constraints such as departure frequency, line capacity, virtual grouping intervals, full load rate, and vehicle inventory, a bi-level optimization model for the train operation plan was formulated. A nested genetic algorithm was devised to determine the optimal solution for the model. Lastly, using the operating conditions of Shijiazhuang Metro Line 1, the optimal results for passenger waiting time, vehicle mileage, and the number of train sets in the optimal train operation plan were obtained.

Keywords: virtual coupling, long-short route, transportation organization mode, train operation plan, optimization

1 Introduction

As a major artery of the comprehensive transportation economy, rail transit is characterized by safety, reliability, comfort, efficiency and environmental friendliness. It has effectively driven the rapid development of regional economic construction and the economic growth of industries along the lines. With the rapid economic development and the expansion of the international market, high-quality development of railways has been promoted, and the demand for passenger transport has been increasing. High-speed and high-density tracking operation of trains has become a new normal in railway operation management. According to the "Completion of Major Indicators of National Railways in September 2024" and "Completion of Major Indicators of National Railways in October 2024" released by the National Railway Administration, the number of railway passengers sent across the country in the two months was 341 million and 372 million respectively, increasing by 5.1% and 6.0% compared with the same period of the previous year. At the same time, a comprehensive analysis of the total number of passengers sent in 2024 shows that it has increased compared with each month of 2023.

Against the backdrop of the rapid development of urban rail transit, the imbalance in the spatio-temporal distribution of passenger flow has become increasingly prominent, and passengers' travel demands have shown a trend of diversification. This requires us to explore more flexible and efficient transportation organization models

---

* Corresponding Author

to adapt to the constantly changing demands. With the continuous growth of travel demands, the rail transport sector needs to continuously enhance its capacity. However, in terms of railway construction, due to resource and environmental constraints as well as construction cost limitations, solving the problem of insufficient transportation capacity supply in China through the construction of new lines faces significant resistance, and its role in alleviating transportation pressure is relatively limited. Therefore, researching how to utilize advanced train control systems to shorten train operation intervals and enhance the capacity of existing lines has become one of the important approaches. In this process, it is necessary to ensure safety and that the train operation speed is not affected, while minimizing the operation intervals between trains as much as possible.

The traditional fixed formation and single route mode is difficult to effectively match the current passenger flow characteristics, leading to excessive differences in the full-load rate of different sections or time periods, long passenger waiting times, and other issues, thereby causing waste of transportation capacity and reducing passenger satisfaction. In contrast, the virtual formation [1] and large-small route mode can dynamically adjust the number of formations and operation intervals according to actual passenger flow demands, achieving more precise passenger flow matching. This mode not only can increase the full-load rate of trains, reduce passenger waiting times, and lower travel costs, but also effectively reduce enterprise operating costs, thereby ensuring the service level and attractiveness of urban rail transit, promoting the economic benefits of operating enterprises, and ensuring the sustainable development of urban rail transit.

Virtual grouping technology can flexibly formulate virtual grouping trains based on the travel needs of passenger flow within urban rail transit stations through vehicle to vehicle communication, optimizing grouping strategies and operation plans without occupying additional line resources. Virtual grouping technology effectively breaks through the limitation of the minimum tracking interval of trains in traditional signal control systems, and increases the total number of train operations. At the same time, its flexible grouping and operation strategy can better adapt to changes in passenger flow in the network, solve the problem of mismatch between transportation capacity and demand, especially in dealing with passenger flow fluctuations during emergencies, demonstrating higher scheduling efficiency. In addition, virtual grouping technology breaks through the limitations of fixed and variable grouping on grouping operation locations and other conditions, and can flexibly and quickly realize the disassembly and grouping operations in urban rail transit operations. Different from the operating characteristics of heavy-duty and high-speed railways, urban rail trains need to face problems such as large track slopes, small turning radii, and high real-time performance when controlling multiple trains in a coordinated manner. At the same time, it is necessary to consider the collaborative interference caused by factors such as uneven tracks and wind resistance, which leads to poor compactness of tracking intervals between train groups, low speed matching, and inability to adapt to high-density operation requirements. Therefore, studying the multi train collaborative tracking and operation control method under the virtual formation operation organization form in the train to train communication technology is of great significance for achieving intelligent dynamic scheduling, so that trains can run stably at the required speed and position [2].

2 Related Work

The concept of virtual train formation originated from the idea proposed by Bock et al., who briefly introduced the operation process of virtual train formation by comparing road traffic modes. Yan Li and others from Lanzhou Jiaotong University proposed an optimization method for urban rail transit timetable based on virtual grouping, considering the spatiotemporal distribution characteristics of urban rail transit passenger flow, traffic oversaturation state, and the limitation of the number of cars in the yard, to address the problem of mismatch between passenger flow and capacity during peak hours [3]. Jia Liu and others analyzed the structure and key technologies of the train control system based on virtual grouping technology, and designed a virtual grouping system scheme based on CBTC system [4]. Yuan Cao et al. proposed a dynamic train grouping and scheduling scheme based on virtual grouping technology, which optimizes train scheduling based on urban rail transit passenger flow dynamics. The verification results show that virtual grouping technology can significantly improve the transportation efficiency of urban rail transit trains [5]. Ying Fan, in response to the current situation where traditional positioning technologies struggle to meet the safety and integrity requirements of level 4 virtual formation, based on traditional train positioning technology, divided the positioning process into three parts: train speed measurement, distance measurement, and initial positioning. Considering the error factors existing in each part of the positioning process, she proposed a new method for calculating the train positioning error envelope. Through mathematical derivation, she derived and calculated various error factors to obtain the upper and lower bounds of each part of the positioning process [6]. Bin Shuai from Southwest Jiaotong University, based on the concept of virtual formation, introduced train dynamics and kinematics theories. According to the speed and distance relationship between two adjacent trains, he proposed a new acceleration adjustment strategy for the trailing train and established a corresponding virtual formation acceleration control model for the train group. This model ensures that all trains in the train group have the same speed and the distance between any two adjacent trains is equal to the expected distance when achieving virtual formation. At the same time, it guarantees train safety and passenger comfort. The model was verified using the CRH380A EMU as a background. Simulation results show that the acceleration adjustment strategy can effectively achieve virtual formation of the train group, reducing the time for train group coordination by 9.7% and the distance between trains by 10.1% [7]. In summary, scholars have achieved corresponding research results in the field of virtual formation. However, the models constructed have not considered a comprehensive train operation model and constraint equations. Moreover, research on the operation of virtual formation by trains is almost non-existent.

Gkiotsalitis et al. optimized the specific location, departure interval, and comprehensive resource allocation of small and large routes for turning back with the goal of minimizing passenger waiting time and operating costs when constructing the model [8]. Wei Li from Shenzhen University proposed an optimization method for flexible grouping and operation of urban rail transit trains under online grouping and decoding (reconnection or uncoupling) mode to cope with the increasing diversity and imbalance of passenger flow and travel demands, resulting in capacity waste [9]. Lei Zhang from Beijing Jiaotong University has developed a brake control algorithm design and optimization analysis based on multi vehicle collaboration to address the issue of uneven spacing control and increased safety risks during emergency parking of various train units during virtual train formation operation [10]. Anan Yang from Beijing Jiaotong University analyzed the train operation plan based on full long and short transfer routes under virtual grouping technology, and aimed to address the problem of uneven passenger flow distribution in urban rail transit. With the goal of minimizing train operating mileage, he constructed an optimization model for urban rail transit capacity configuration, and finally achieved a good balance between full load rate and departure interval, improving passenger travel efficiency and reducing operating costs [11].

When seeking the optimal operation plan for trains, Ai Ren from Beijing Jiaotong University focused on the operation of light rail between and within cities. With the objective of minimizing passenger travel costs and enterprise operation costs, a bi-objective nonlinear integer programming model was constructed, taking into account constraints such as passenger travel demand and line capacity. A genetic algorithm was designed for solving the model. Taking a certain line in Chongqing as an example, after optimization by the model, passenger travel time was reduced by 2.92%, and the total operation cost of the enterprise was reduced by 4.88% [12]. Yao Chen designed a customized decomposition algorithm based on the search of train routes and operation frequencies and the coordination of operation intervals. By constructing a passenger travel spatio-temporal network diagram, the passenger flow behavior under complex train routes was accurately depicted. A mixed integer programming model was established with the objective of minimizing train running kilometers and passenger travel time costs. With train route plans, operation ratios and frequencies as decision variables, the model optimized the efficiency of the operation network [13].

In summary, the main work of this article is as follows:

1) Research on virtual grouping focuses on the basic concepts and safety of this technology, the requirements for train control systems to implement this technology, and its value in improving the line’s throughput and supply-demand matching capabilities. There is relatively little research on the operation plan of virtual grouping trains.

2) The research on train operation plans usually starts from the perspectives of passengers and enterprises, with passenger travel time cost and enterprise operation cost as the objective functions. Some studies take train full load rate as the objective function, and most of them take train departure frequency and small route turnaround station position as decision variables. There are few studies that consider the cost of passenger travel time, enterprise operating costs, and train full load rate in a coordinated manner, and there are few studies that use the number of train formations as a decision variable. Even if the number of train formations is used as a decision variable, it remains fixed during operation.

There is almost no research on the background of online reconnection or disassembly operations, and there are even fewer studies on the variable size and routing of train formations during operation. This article takes the operation plan of virtual grouped trains as the research object, proposes a transportation organization mode for trains with different sizes from the traditional mode, and then constructs an optimization model for the operation plan of virtual grouped trains with different sizes.

3 Construction of Virtual Marshalling Train Model

Virtual formation train refers to a train that no longer relies on physical couplers to complete coupling, but is based on wireless communication, automatic control, and information exchange technology between trains. Two or more trains are coupled and virtually run as the same train, jointly undertaking the same transportation task. According to the direction of train operation, the train located in front is defined as the leading car, and the train located behind is defined as the following car. The composition of the virtual formation train is shown in Fig. 1. From a macro perspective, virtual formation trains maintain synchronous operation of train units through collaborative control methods; From a microscopic perspective, the tracking interval between each train unit is dynamically adjusted based on the line information and the kinematic parameters of the preceding train. To avoid collisions, the minimum allowable interval distance should be maintained at all times.

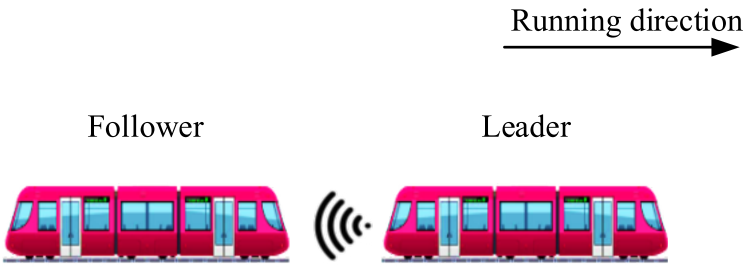

Fig. 1. Schematic diagram of virtual coupling train

3.1 Train Tracking Time Interval Model

The Communication based Train Control System (CBTC), which applies virtual grouping technology, adds “train-train” communication on the basis of the original “train ground train” communication structure, that is, direct communication between trains, real-time exchange of train position, speed and other information, and active calculation of train movement authorization. Compared to traditional CBTC systems, CBTC systems using virtual grouping technology also include synchronization of acceleration or braking command information for trains, with more advanced communication technology. The information transmission delay during train operation is smaller, the distance is farther, and the anti-interference ability is stronger, making it safer and more efficient [14].

This article considers the coupling and uncoupling of two train units at a station. The schematic diagram of train station tracking in traditional mode is shown in Fig. 2. The schematic diagram of station tracking for multiple trains under virtual grouping technology is shown in Fig. 3. Under the traditional CBTC system, the train tracking interval only considers the position of the preceding train, without considering its speed, treating the preceding train as a static obstacle (also known as the “hard wall collision” mode). In virtual grouping, due to the consideration of real-time changes in the position and speed of the preceding train, the preceding train will not be immediately regarded as a static obstacle when encountering an emergency situation, but will brake urgently for a certain distance with braking deceleration (also known as the “soft wall collision” mode), so the braking time of the following train can be later. The time interval for tracking train stations during virtual grouping is:

$$T_{track}=\begin{cases}\dfrac{2a_1a_2\left(2L_{train}+L_{interval}+L_s\right)+\left(a_1+a_2\right)v^2}{2a_1a_2v}+T_{reaction}+T_w+\dfrac{v}{a_2},v<\sqrt{\dfrac{2a_1a_2\left(2L_{train}+L_{interval}+L_s\right)}{a_1+a_2}}\\ \sqrt{\dfrac{2a_2\left(2L_{train}+L_{interval}+L_s\right)}{a_1a_2+a_1^2}}+T_{reaction}+T_w+\dfrac{v}{a_2},v\geq\sqrt{\dfrac{2a_1a_2\left(2L_{train}+L_{interval}+L_s\right)}{a_1+a_2}}\end{cases}\quad(1)$$

The parameter explanations in the above formula are shown in Table 1.

Table 1. Parameter description list

Parameter symbols	Parameter description
T_{track}	Train station tracking time interval during virtual grouping (unit: s)
L_{train}	Train length (unit: m)
$L_{interval}$	The distance between two trains inside a virtual formation train (unit: m)
v	The running speed of the train in the section (unit: m/s)
L_s	Reserved train safety protection distance under virtual grouping technology (unit: m)
a_1	Train startup acceleration(unit: m/s^2)
a_2	Train braking deceleration (unit: m/s^2)
$T_{reaction}$	Response time of drivers and equipment (unit: s)
$L_{reaction}$	Reaction distance between driver and equipment
T_w	Train stopping time at the station (unit: s)

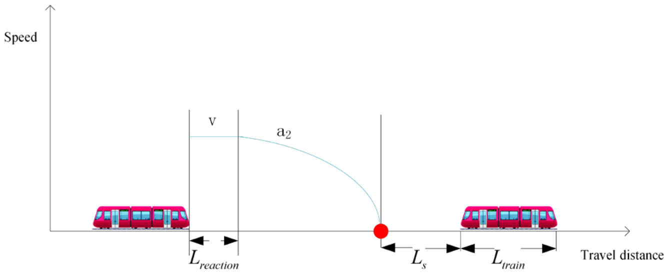

Fig. 2. Train station tracking in traditional mode

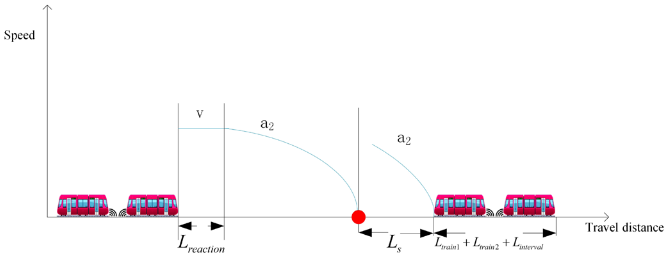

Fig. 3. Train station tracking based on virtual coupling technology

3.2 Quantification of Virtual Grouping Problem

In order to match the uneven distribution of urban rail transit passenger flow in time and space, this paper proposes a virtual grouping based transportation organization model for large and small routes. As shown in Fig. 4, a certain urban rail transit line has a total of N stations, with the downward direction from Station 1 to Station N, and the upward direction vice versa. Among them, station x and station y are turnaround stations for small routes. The route includes large trains running back and forth from station 1 to station N, as well as small trains running back and forth from turnaround station x to y. Assuming that all major trains will operate in conjunction with minor trains in minor sections, minor trains can be divided into two types: individual minor trains and “virtual formation trains” that are connected to major trains.

The process of uncoupling and uncoupling between large and small trains is as follows: assuming that the number of formations for large trains is $b1$ and the number of formations for small trains is $b2$, when the descending large train runs from station 1 to station x, it is coupled with the small train as a “virtual formation train”; When continuing to run to station y, the two trains with the number of formations $b1$ and $b2$ are disassembled. Among them, the train with the number of formations $b1$ continues to run on the main route to station N, while the train with the number of formations $b2$ passes through the turnaround line of station y and reaches the up section of the small route. It can continue to be connected with the up main route train as a “virtual formation train”, or run separately on the small route section. The same applies to the up section [15].

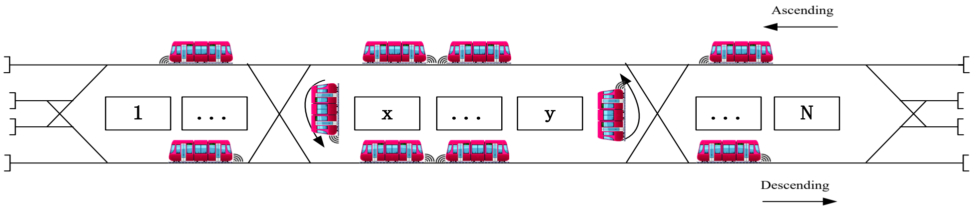

Fig. 4. Organizational model of long and short routing based on virtual coupling

3.3 Construction of Organizational Operation Model for Rail Trains

Taking the peak hours of a certain city’s rail transit line as the research object, the OD distribution table of passenger flow and related equipment and facility parameters are known. Under the constraints of meeting passenger demand, existing number of carriages, line capacity, section full load rate, number of carriages, departure frequency, and virtual carriage interval, the departure frequency of trains with different sizes of carriages, the location of virtual coupling/uncoupling stations, and the number of carriages with different sizes of carriages are determined to minimize the comprehensive goals of passenger travel costs and enterprise operating costs.

Basic assumption:

1) Convert the exchange passenger flow into the inbound passenger flow;

2) Adopting the method of paired up and down operation and station stop;

3) Assuming that there are only large and small routes, no other route forms are set, and the trains with large and small routes use their own carriages independently;

4) Assuming that all passengers can board the next train after arriving at the station, without considering the situation where passengers are stranded on the platform, and the time value of passengers is the same;

5) Not considering the impact of small route trains turning back on the operation of large route trains;

6) The distance between two adjacent stations in the up and down directions, as well as the running time between sections, are the same, and the stopping time of trains with different sizes in the up and down directions at each station is the same;

7) When a virtual multiple unit train stops at a station for boarding and alighting operations, the train stops based on the middle position of the platform, and the excess carriages stop outside the platform for boarding and alighting operations. Passengers first board the train and then move towards the empty carriages parked outside the platform;

8) All stations on the line do not have the conditions for overtaking, and train overtaking is not considered;

9) All trains depart evenly, meaning that the departure intervals of all trains are equal.

3.4 Establishment of Objective Function

All parameters and their descriptions in the objective function are shown in Table 2.

Table 2. Parameter list and parameter description

Parameter symbols	Parameter definition and description
i, j, m	Station index, $i, j, m \in \{1, 2, ... N\}$
x, y	Index of the starting and ending stations for the downward movement, as well as the turning back stations, $x < y$
$e(m, m+1)$	Interval index, representing the interval or section between two adjacent stations
c	Upstream and downstream indexes, $c \in (1, 2)$, where 1 represents downstream and 2 represents upstream
h	Small and large intersections, $h \in (1, 2)$ where 1 represents a small intersection and 2 represents a large intersection.
O_{ij}	The passenger flow from boarding at station i to alighting at station j in that hour.
$L_{m,m+1}$	The distance between the center of Station m and the center of Station $m+1$, unit: *km*
$T_{m,m+1}$	Train running time for section $e(m, m+1)$, unit: *s*
l_m^{zf}	The length from the center of Station m to the endpoint of the turnaround line, unit: *km*
S_m	When station m stops, unit: *s*
t_m^{zf}	Train turnaround time at station m, unit: *s*
U	The seating capacity of a car.
β	Train full load rate.
$\beta_{\max}$	Maximum train loading rate.
$\beta_{\min}$	Minimum train load factor.
b	The current number of train formations
$b_{\max}$	Maximum number of cars in a train formation
$b_{\min}$	Maximum number of cars in a train formation
b_1	The number of cars in a short routing train formation
b_2	The number of cars in the formation of the short routing train
Y_1	Train operating frequency for small routes, unit: pairs/hour
Y_2	The frequency of train operation on the main route, in pairs per hour
x	The starting and turning station of the small route
y	The final destination of the small route is the turnaround station

The objective function considers both passenger travel costs and enterprise operating costs. Passenger travel costs are represented by passenger waiting time, while enterprise operating costs are represented by the total number of vehicle kilometers traveled and the number of online train sets.

The waiting time of passengers is equal to the total number of passengers multiplied by the average waiting time of passengers, mainly related to the number of passengers, arrival distribution, and departure interval. Research in reference shows that when the travel interval is less than 10 minutes, passengers taking the subway will not pay attention to the departure interval, and the arrival pattern follows a random normal distribution. Half of the departure interval can be used to represent the average waiting time of the overall passenger flow. Total waiting time for descending passengers:

$$T_{\text{x}} = \frac{1}{2} \times \frac{60}{Y_1 + Y_2} \times \sum_{i=x}^{y-1} \sum_{j=x+1}^{y} O_{ij} + \frac{1}{2} \times \frac{60}{Y_2} \times \left(\sum_{i=1}^{N-1} \sum_{j=2}^{N} O_{ij} - \sum_{i=x}^{y-1} \sum_{j=x+1}^{y} O_{ij} \right), i < j \tag{2}$$

Total waiting time for up passengers:

$$T_{\text{s}} = \frac{1}{2} \times \frac{60}{Y_1 + Y_2} \times \sum_{i=x+1}^{y} \sum_{j=x}^{y-1} O_{ij} + \frac{1}{2} \times \frac{60}{Y_2} \times \left(\sum_{i=2}^{N} \sum_{j=1}^{N-1} O_{ij} - \sum_{i=x+1}^{y} \sum_{j=x}^{y-1} O_{ij} \right), i > j \tag{3}$$

The total waiting time of passengers is expressed as:

$$\min Z_1 = T_{\mathrm{s}} + T_{\mathrm{x}} \tag{4}$$

The minimum representation of the total mileage traveled by the vehicle is:

$$\min Z_2 = b_2 \times Y_2 \times 2 \times \left(\frac{\sum_{m=1}^{N-1} L_{m,m+1}}{1000} + l_1^{zf} + l_N^{zf} \right) + b_1 \times Y_1 \times 2 \times \left(\frac{\sum_{m=x}^{y-1} L_{m,m+1}}{1000} + l_x^{zf} + l_y^{zf} \right) \tag{5}$$

Minimum number of online vehicle groups:

$$\min Z_3 = \sum_{h=1}^{2} \left\lceil \frac{T_h^{turnover} \times Y_h}{60} \right\rceil \tag{6}$$

$T_h^{turnover}$ represents the turnaround time of route h, in this paper, the weight coefficients of the three objectives of passenger waiting time, vehicle kilometers and the number of on-line vehicle groups are transformed into a single objective, which is expressed as:

$$\min Z = \min\left(\lambda_1 Z_1 + \lambda_2 Z_2 + \lambda_3 Z_3\right) \tag{7}$$

Where λ_1 a is the weight dilution of Z_1, λ_2 is the weight coefficient of Z_2, and λ_3 is the weight coefficient of Z_3.

Because most subways adopt the operation scheme of 6-car large marshalling and single routing, this operation scheme is selected as the benchmark to calculate the weight coefficient. In this train operation scheme, the relationship between passenger waiting time T, vehicle running kilometers L and on-line train sets Z can be expressed as:

$$\begin{cases} \lambda_1 T = \lambda_2 L = \lambda_3 Z \\ \lambda_1 = 1 \\ \lambda_2 = \dfrac{\lambda_1 T}{L} \\ \lambda_3 = \dfrac{\lambda_1 T}{Z} \end{cases} \tag{8}$$

3.5 Constraint Condition

1) Departure frequency limit, Article 3.2.2 of code for design of Metro (GB 50157-2013) stipulates that the maximum long-term capacity of system design shall meet the requirement that the traffic density shall not be less than 30 pairs/h:

$$\sum_{h=1}^{2} Y_h \leq 30 \tag{10}$$

According to GB 50490-2009 technical specification for urban rail transit, the maximum running interval shall not be greater than 10 min when the train operates normally during the operation period:

$$Y_2 \geq 6 \tag{11}$$

The departure frequency shall also consider the restriction of turn back capacity of turn back station:

$$Y_2 \leq \frac{3600}{\min\left(t_1^{\mathrm{zf}}, t_N^{\mathrm{zf}}\right)} \tag{12}$$

$$Y_1 \leq \frac{3600}{\min\left(t_x^{\mathrm{zf}}, t_y^{\mathrm{zf}}\right)} \tag{13}$$

The line adopts virtual marshalling technology, and all trains depart evenly, so the train tracking interval is:

$$\frac{3600}{Y_1 + Y_2} \geq L_{track} \tag{14}$$

All large routing trains will be coupled with small routing trains in small routing sections, while small routing trains can operate independently:

$$Y_1 = nY_2, n \in N^+ \tag{15}$$

The train departure frequency of each route shall be an integer, namely:

$$Y_h \in N^+, h = 1,2 \tag{16}$$

2) Restrictions on passenger flow at each section

Section passenger flow of section $e(m, m+1)$ in direction c:

$$D_{e(m,m+1),c} = \begin{cases} \sum_{i=1}^{m} \sum_{j=m+1}^{N} O_{ij}, c = 1, i < j \\ \sum_{i=m+1}^{N} \sum_{j=1}^{m} O_{ij}, c = 2, i > j \end{cases} \tag{17}$$

In any running direction and on any section, the passenger flow that the train can bear should be greater than or equal to that section. Maximum section passenger flow in independent operation section of long routing:

$$D_{e(m,m+1),c,\max}^{long} = \max\{D_{e(m,m+1),c}, 1 \leq m \leq x\text{-}1 \cup y \leq m \leq N-1, c = 1, 2\} \tag{18}$$

Maximum section passenger flow in mixed operation section of large and small routes:

$$D_{e(m,m+1),c,\max}^{long-short} = \max\{D_{e(m,m+1),c}, m < N-1, c = 1, 2\} \tag{19}$$

The train operation scheme meets the passenger flow demand:

$$D_{e(m,m+1),c,\max}^{long} \leq Y_2 \times U \times b_2 \times \beta_{\max} \tag{20}$$

$$D_{e(m,m+1),c,\max}^{long-short} \le \sum_{h=1}^{2} Y_h \times U \times b_h \times \beta_{\max} \tag{21}$$

3) The full load rate meets the condition limit

$$\beta_{\min} \le \beta \le \beta_{\max} \tag{22}$$

4) Limit on the number of train formation

$$b_{\min} \le b_h \le b_{\max}, h = 1, 2 \tag{23}$$

5) Constraints of small turn back station

$$1 \le x \le y \le N \tag{24}$$

6) Total number of trains

Urban rail transit has three types of trains, i.e. operation train, maintenance train and standby train. The standard for construction of urban rail transit projects stipulates that the number of vehicles under repair is generally 10% -15% of the number of trains in use, and the number of standby vehicles is controlled at about 10%, the total number of trains is:

$$N_{total} = N_{using} + N_{overhaul} + N_{spare} \tag{25}$$

N_{total} represents the total number of trains, N_{using} represents the number of trains in use, $N_{overhaul}$ represents the number of maintenance trains, and N_{spare} represents the number of standby trains.

Then the number of on-line trains shall be less than or equal to the number of trains in use:

$$\sum_{h=1}^{2} \left\lceil \frac{T_h^{turnover} \times Y_h \times b_h}{60} \right\rceil \le \lceil 0.75 N_{total} \times b \rceil \tag{26}$$

$\lceil\ \rceil$Indicates rounding up, the turnaround time of long routing is expressed as:

$$T_2^{turnover} = \frac{2\sum_{m=1}^{N-1} T_{m,m+1} + 2\sum_{m=1}^{N} S_m + t_1^{zf} + t_N^{zf}}{60} \tag{27}$$

Turnover time of short routing trains:

$$T_1^{turnover} = \frac{2\sum_{m=x}^{y-1} T_{m,m+1} + 2\sum_{m=x}^{y} S_m + t_x^{zf} + t_y^{zf}}{60} \tag{28}$$

4 Optimization Algorithm Design

The virtual marshalling scheme has many optimization factors, and the problem scale is large. When solving non single objective and there are multiple constraints, this paper uses nested genetic algorithm to solve the model. A double-layer optimization model for the operation scheme of large and small routes of virtual marshalling trains is constructed. The model takes the passenger travel time cost and enterprise operation cost as the upper objective function, the equilibrium of train load factor as the lower objective function, and the train departure frequency, the location of small route turn back station and the number of train marshalling vehicles as the decision variables [16]. At the same time, a double-layer improved genetic algorithm is designed to solve the optimal solution of the virtual marshalling scheme [17]. The overall structure model of the algorithm is shown in Fig. 5.

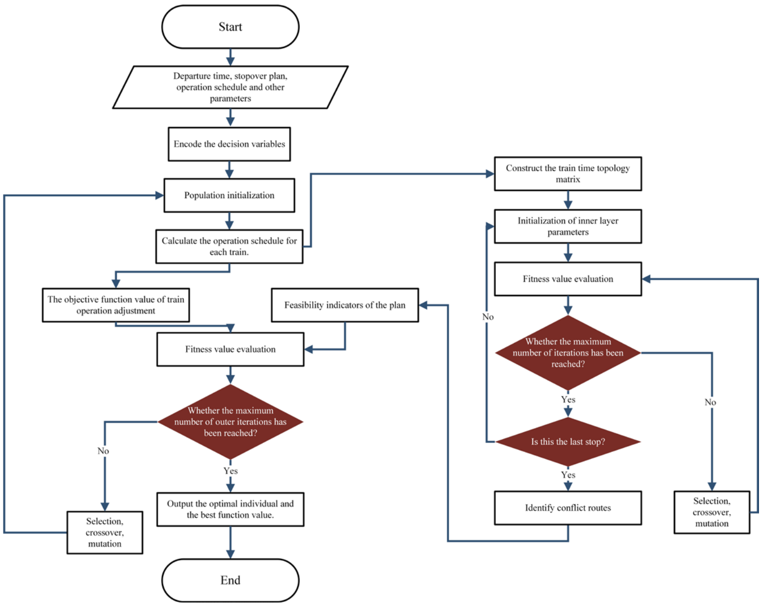

Fig. 5. Flow chart of genetic algorithm

Some pseudo codes are as follows:

```
#Initialization parameters
Population_size=100
Max\u generations=200
Mutation_rate=0.01
Crossover\u rate=0.8
#Define decision variables
#Train departure frequency
#Turnround_station
#Train_cars
#Initialize population
def initialize_population(population_size):
population = []
for _ in range(population_size):
frequency = random_frequency()
turnaround_station = random_turnaround_station()
train_cars = random_train_cars()
population.append((frequency, turnaround_station, train_cars))
return population
#Upper level objective function: passenger travel time cost+enterprise operation
cost
def upper_objective_function(individual):
frequency, turnaround_station, train_cars = individual
#Calculate passenger travel time cost
passenger_cost = calculate_passenger_cost(frequency, turnaround_station, train_
cars)
#Calculate enterprise operating costs
operation_cost = calculate_operation_cost(frequency, turnaround_station, train_
cars)
return passenger_cost + operation_cost
```

```
#Lower level objective function: equilibrium of train load factor
def lower_objective_function(individual):
frequency, turnaround_station, train_cars = individual
#Calculate the equilibrium of train load factor
load_balance = calculate_load_balance(frequency, turnaround_station, train_cars)
return load_balance
#Fitness function
def fitness_function(individual):
upper_cost = upper_objective_function(individual)
lower_balance = lower_objective_function(individual)
#Comprehensive consideration of upper and lower level objectives
return upper_cost + lower_balance
#Select action
def selection(population, fitness_scores):
#Use roulette selection
selected = roulette_wheel_selection(population, fitness_scores)
return selected
```

5 Real Case Solution and Result Analysis

This paper takes Shijiazhuang Metro Line 1 as an example. The route map of Shijiazhuang Metro Line 1 is shown in Fig. 6, starting from Xiwang Station in the West and from east to north to Fuze station, with a total of 26 stations and 25 sections. Stations with turn back lines include fuze, convention and Exhibition Center, Xizhuang, Juhe Avenue, Shijiazhuang east station, Liucun Station, Tanggu Station, Beiguo mall Station, Xinbai square Station, Heping hospital Station and Xiwang Station. The length from the center of the station to the end of the turn back line is about 0.4 km. Assuming that there is no turn back line at present, if the turn back line needs to be rebuilt, the length from the station center to the terminal of the turn back line is 0.4 km.

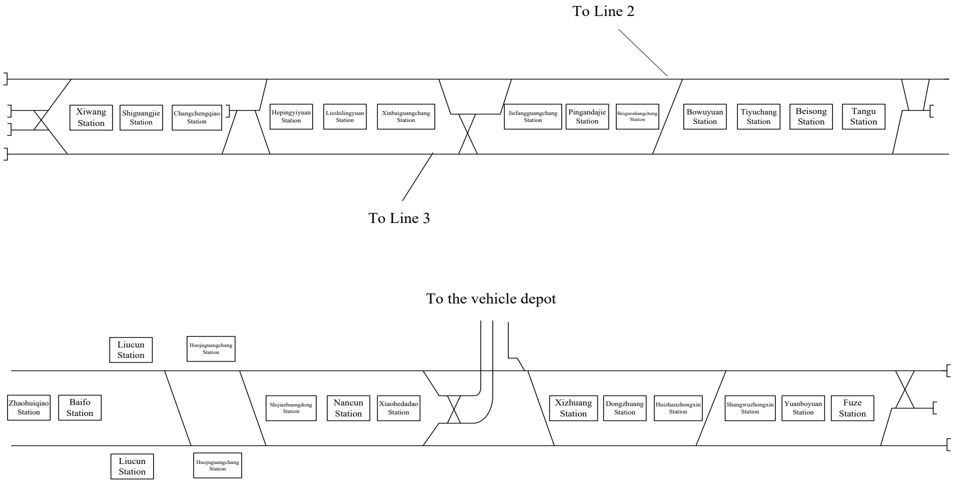

Fig. 6. Subway map

At present, 6-car marshalling of type a car is adopted, with a seating capacity of 310 cars per car, the maximum and minimum load factor of 1.2 and 0, and 36 cars under the car. In order to make full use of the existing resources and save costs, this paper assumes that the existing type a vehicle is still used, and the number of virtual marshalling train units for both large and small routes is 3, so the number of virtual marshalling train units is 72. Use the real value of passenger flow distribution from 8:00 to 9:00 in the morning peak. The passenger flow distribution during the morning rush hour from 8:00 to 9:00 is shown in Table 3. Among them, the sections with a larger number of passengers are between Shijiazhuang East Railway Station and Xiwang Station. Therefore, only the passenger flow of the sections with a larger number of people is displayed.

Table 3. Passenger flow OD table of some stations on Shijiazhuang Metro Line 1 during the morning peak from 8:00 to 9:00

Station	Shijiazhuang East Railway Station	Torch Square Station	Liucun Station	Baifo Station	Chaohui Bridge Station	Tanggu Station	Beisong Station	Sports Field Station	Museum Station	Beiguo Mall Station	Ping'an Street Station	Jiefang Square Station	Xinbai Square Station	Martyrs' Cemetery Station	Heping Hospital Station	Great Wall Bridge Station	Shiguang Street Station	Xiwang Station
Shijiazhuang East Railway Station	0	5	9	11	11	11	11	25	77	245	97	109	659	61	263	49	23	41
Torch Square Station	17	0	5	9	13	13	15	21	145	354	217	213	977	173	287	127	33	161
Liucun Station	65	57	0	5	11	15	17	23	387	1246	497	813	1129	265	333	137	37	165
Baifo Station	31	115	31	0	9	15	18	21	27	121	65	57	297	89	257	71	5	91
Chaohui Bridge Station	31	149	63	11	0	5	9	13	543	1557	555	703	1335	125	243	77	5	79
Tanggu Station	39	113	71	15	9	0	11	13	383	1464	495	690	1597	169	253	85	7	99
Beisong Station	27	65	69	17	23	27	0	9	263	1187	81	153	1037	93	259	91	9	83
Sports Field Station	29	75	71	23	25	65	29	0	9	209	47	57	129	77	93	33	5	39
Museum Station	39	189	131	19	27	81	67	39	0	17	21	29	33	25	27	23	11	55
Beiguo Mall Station	143	619	507	111	71	355	343	443	303	0	15	63	111	75	83	75	39	131
Ping'an Street Station	39	201	197	39	31	93	107	99	235	173	0	7	35	35	37	39	9	43
Jiefang Square Station	45	255	231	35	37	83	87	93	243	449	67	0	47	35	37	37	17	71
Xinbai Square Station	71	289	253	59	57	195	119	133	775	1409	697	397	0	15	37	41	13	65
Martyrs' Cemetery Station	25	215	177	37	29	83	87	63	213	665	253	155	49	0	7	35	7	43
Heping Hospital Station	33	285	275	45	47	125	109	101	427	1141	443	221	129	49	0	7	7	39
Great Wall Bridge Station	53	273	255	49	55	179	117	113	455	877	335	257	105	85	23	0	5	31
Shiguang Street Station	39	167	159	45	37	129	59	65	121	269	125	135	57	43	25	13	0	19
Xiwang Station	197	475	465	147	119	437	541	623	681	1857	825	823	117	147	99	75	43	0

The parameter values of virtual marshalling train tracking interval are shown in Table 4. Among them, when the running speed of urban rail transit trains is high, the distance $L_{\text{int}erval}$ between two single trains in the train with large marshalling is about $110m$. When the running speed of trains decreases, $L_{\text{int}erval}$ may be reduced accordingly, which may be $10m$ at low speed. In order to ensure safety, this paper takes the train spacing $L_{\text{int}erval}$ =$110m$ to calculate the tracking interval at various speeds. T_w takes the maximum dwell time of each station as $55s$.

Table 4. Related parameter settings of virtual marshalling headway

Parameter symbols	Parameter value
L_{train}	66m
$L_{\text{int}erval}$	110m
v	15m
L_s	$0.8\,m/s^2$
a_1	$0.8m/s^2$
a_2	$80\,km/h$
$T_{reaction}$	3s
T_w	55s

By substituting the above parameters into formula (1), the tracking time interval of train station during virtual marshalling is $T_{track} = 103.71s$.

Table 5 shows the distance between stations of Shijiazhuang Metro Line 1 and the running time of trains in the section. Table 6 shows the turn back time and dwell time of stations of Shijiazhuang Metro Line 1. If there is no turn back line at the start and end stations of the small routing, it is necessary to relay the turn back line, which is very expensive. Therefore, this article stipulates that the turn back time of stations without turn back line is set to 20 million s, and the turn back time of stations with turn back line is 90 s. Assuming that the time for virtual coupling and de marshalling at the turn back station is less than the dwell time, it is not necessary to reserve the virtual coupling and de marshalling time separately.

Table 5. Distance between stations and interval train operation time of Shijiazhuang Metro Line 1

Interval name	Station spacing(km)	Train operation time(s)	Interval name	Station spacing(km)	Train operation time(s)
$e(1,2)$	1.064	91	$e(14,15)$	1.441	107
$e(2,3)$	1.316	101	$e(15,16)$	1.146	95
$e(3,4)$	1.174	100	$e(16,17)$	0.898	80
$e(4,5)$	3.607	221	$e(17,18)$	0.876	84
$e(5,6)$	1.403	108	$e(18,19)$	0.984	85
$e(6,7)$	1.404	105	$e(19,20)$	1.267	105
$e(7,8)$	1.030	89	$e(20,21)$	0.862	80
$e(8,9)$	1.195	95	$e(21,22)$	1.083	91
$e(9,10)$	1.622	122	$e(22,23)$	1.156	94
$e(10,11)$	1.285	102	$e(23,24)$	1.337	103
$e(11,12)$	2.345	165	$e(24,25)$	1.220	95
$e(12,13)$	0.810	77	$e(25,26)$	1.290	102
$e(13,14)$	1.270	102			

Table 6. Dwell time and turn back time of each station of Shijiazhuang Metro Line 1

Station No	Dwell time(s)	Turn back time(s)	Station No	Dwell time(s)	Turn back time(s)
1	45	90	14	40	90
2	40	20 000 000	15	40	20 000 000
3	40	20 000 000	16	40	20 000 000
4	35	90	17	40	20 000 000
5	35	20 000 000	18	45	90
6	35	90	19	40	20 000 000
7	45	90	20	40	20 000 000
8	30	20 000 000	21	55	90
9	35	90	22	35	20 000 000
10	40	20 000 000	23	40	90
11	40	90	24	35	20 000 000
12	35	20 000 000	25	35	20 000 000
13	40	20 000 000	26	45	90

When Shijiazhuang Metro Line 1 adopts a single routing and 6-car large marshalling, the departure frequency should not only meet the requirements of maximum section passenger flow, but also meet the requirements of running interval greater than or equal to 10 min. The minimum running frequency is:

$$Y = \max\left\{\left\lceil \frac{D_{\max}}{B \times U \times \beta_{\max}} \right\rceil, 6\right\} = \max\left\{\left\lceil \frac{26609}{6 \times 310 \times 1.2} \right\rceil, 6\right\} = 12\,pairs/h \tag{29}$$

The total passenger volume from 8:00 to 9:00 in peak period is $\sum O_{ij} = 60471$ person times. Therefore, the waiting time of passengers can be expressed as:

$$T = \frac{1}{2} \times \frac{60}{Y} \times \sum O_{ij} = \frac{1}{2} \times \frac{60}{12} \times 60471 = 151177.5 \min \tag{30}$$

Vehicle kilometers can be expressed as:

$$L = B \times Y \times 2 \times 33.9 = 6 \times 12 \times 2 \times 33.9 = 4881.6km \tag{31}$$

The number of on-line trains can be expressed as:

$$Z = \left\lceil \frac{T_2^{turnover} \times Y}{60} \right\rceil = \left\lceil \frac{124.8 \times 12}{60} \right\rceil = 25 \tag{32}$$

After calculation, $\lambda_1 = 1$, $\lambda_2 = 30.97$, $\lambda_3 = 6047.1$, to verify the effectiveness of the model and algorithm, the Python programming solution, the running environment is windows 11, 64 bit host, and the computer frequency is Intel (R) core (TM) i9-13900HX, the running memory is 32GB, therefore, the final objective function is:

$$\min Z = \min\left(Z_1 + 30.97Z_2 + 6047.1Z_3\right) \tag{33}$$

After calculation, the following results are obtained: the small routing operation section is set between the 9th station and the 26th station, that is, from Shijiazhuang east station to Xiwang station, and the departure frequency is 18 pairs/hour; Dajiao road starts from fuze station to Xiwang station, and the departure frequency is 6 pairs/hour. The optimized train operation scheme is compared with the actual train operation scheme of Shijiazhuang Metro Line 1. At present, Shijiazhuang Metro Line 1 adopts the 6-car marshalling and single routing operation scheme at peak hours, with the frequency of 12 pairs/h. The optimized scheme is compared with the existing one, as shown in Fig. 7.

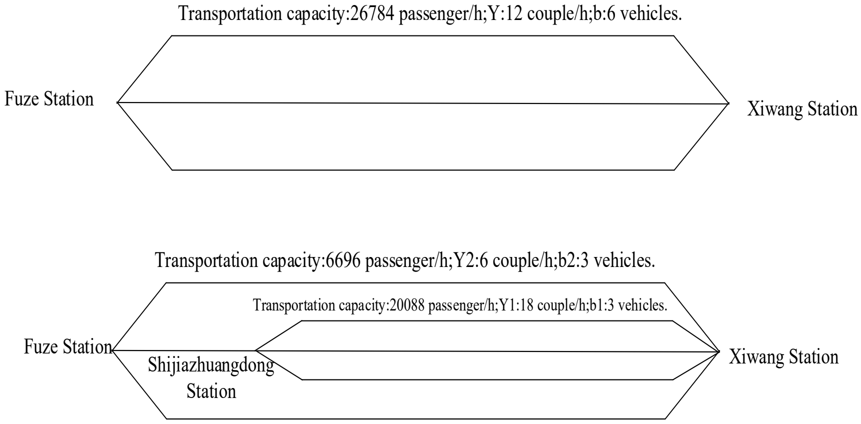

Fig. 7. Comparison of different train operation schemes

Table 7 shows the comparison of the evaluation indexes between the optimized and the actual operation plan.

Table 7. Comparison of evaluation indexes of different routing schemes

	Total waiting time of passengers(min)	Vehicle running kilometers(*km*)	Number of on-line trainsets
Actual operation plan	151177.5	4881.6	25
Optimize the development plan	89216.25	3563.136	38
Variation	-61961.25	-1318.464	13
Rate of change	-40.99	-27.01	52.00

After calculation, the total objective function value of the optimized operation scheme is 429356.3719 from 8:00 to 9:00 in the morning rush hour on weekdays, while the total objective function value of the actual operation scheme is 453538.152, a decrease of 24181.7801, indicating that the optimized train operation scheme in the morning rush hour is better than the actual train operation scheme. The optimized train operation scheme has less waiting time for passengers and less vehicle kilometers, which is obviously beneficial to passengers and enterprises. Because the virtual marshalling number of the optimized operation scheme is 3 and the actual marshalling number of the operation scheme is 6, it will inevitably lead to an increase in the number of on-line trains.

6 Conclusion

Taking an urban rail transit line in peak hours as the research object, considering the interests of passengers and enterprises, this paper establishes an optimization model of large and small routing train operation scheme based on virtual marshalling. By inputting the passenger OD distribution table and relevant parameters into the model, the optimal train operation scheme under the current conditions can be calculated.

Taking Shijiazhuang Metro Line 1 as an example, the optimized operation scheme is compared with the actual operation scheme. The comprehensive objective is smaller and the comprehensive cost is lower, which verifies the feasibility and correctness of the model. When this paper studies the operation scheme, the number of virtual marshalling of large and small routes is set as known. If the number of virtual marshalling is unknown, it is necessary to continue to study.

Virtual formation technology is gradually becoming an important development direction for the next-generation train control technology in urban rail transit, and it is also one of the key means to achieve energy conservation, carbon reduction and efficiency improvement. By dynamically adjusting the capacity of virtual formation trains based on time-varying passenger flow demands, efficient utilization of vehicle resources and precise matching of capacity and volume can be achieved, thereby providing passengers with higher-quality passenger transport services and further enhancing the attractiveness and share of rail transit travel. When studying the operation plan of virtual formation trains, if the number of virtual formations in large and small routes is set as a known condition, it can help simplify the problem analysis; however, when the number of virtual formations is unknown, further in-depth research is still needed to determine the optimal configuration plan. In addition, when constructing an optimization model, increasing the number of trains in the train schedule will lead to a significant increase in total calculation time. Therefore, designing more lightweight and efficient solution algorithms will be one of the key research directions in the future. Moreover, in line with the future trend of urban rail transit network development, it is necessary to further improve the optimization methods for multi-line interconnection and interconnection operation between railways and urban rail transit, and deeply analyze the new changes and their impacts brought about by different performance train units forming virtual formation trains under cross-line transportation conditions.

7 Acknowledgement

This research was funded by Science Research Project of Hebei Education Department (No. QN2025413).

References

[1] Z.-Q. Li, L.-Y. Zhong, H. Yang, Distributed cooperative predictive control of virtual coupled trains, Journal of Traffic and Transportation Engineering 24(5)(2024) 362-378.

[2] H.-S. Zhou, J.-G. Qi, L.-X. Yang, J.-G. Shi, H.-M. Zhang, An Optimization Method for Train Unit Flexible Scheduling Using Virtual Coupling Technology, Journal of Transportation Systems Engineering and Information Technology 24(5)(2024) 140-147.

[3] Y. Li, L. Gong, D.-J. Xu, X. Pan, C.-H. Hu, A Timetable Optimization Method for Urban Train Transit Based on Virtual Coupling, Journal of Transport Information and Safety 42(3)(2024) 74-84.

[4] J. Liu, Research on Virtual Coupling Train Control Solution in CBTC System, Railway Signalling & Communication Engineering 19(1)(2022) 73-77.

[5] Y. Cao, J.-K. Wen, L.-C. Ma, Dynamic marshalling and scheduling of trains in major epidemics, Journal of Traffic and Transportation Engineering 20(3)(2020) 120-128.

[6] Y. Fan, B.-W. Wen, Z.-Z. Wang, A Train Positioning Error Estimation Method Suitable for Multiple Trains Virtual Marshalling Control System, Urban Mass Transit 27(6)(2024) 270-275.

[7] B. Shuai, J.-N. Luo, X.-Y. Feng, W.-C. Huang, A Train Group Control Method Based on Car Following Model Under Virtual Coupling, Journal of Transportation Systems Engineering and Information Technology 24(3)(2024) 1-11.

[8] K. Gkiotsalitis, Z. Wu, O. Cats, A cost-minimization model for bus fleet allocation featuring the tactical generation of short-turning and interlining options, Transportation Research Part C: Emerging Technologies 98(2019) 14-36.

[9] W. Li, S.-Q. Chen, J. Zhou, Q. Luo, Y.-X. He, Optimization Method of Flexible Train Composition Scheduling in Urban Rail Transit under Online Coupling/Decoupling Mode, China Railway Science 45(1)(2024) 203-214.

[10] L. Zhang, Z.-M. Li, Y.-Y. Yan, F. Dou, H.-J. Liu, Brake control algorithm for virtually coupled trains based on multi vehicle cooperation, Journal of Jilin University (Engineering and Technology Edition) 54(10)(2024) 3027-3036.
[11] A.-A. Yang, J.-Y. Sun, B. Wang, S.-K. Chen, W. Ming, Optimization of virtual-coupling-orientated train operation plan based on full-length and short-turn routing, Journal of Beijing Jiaotong University 46(4)(2022) 9-14.
[12] A. Ren, L.-Y. Meng, Y.-H. Wang, W.-Z. Huang, R. Ren, Optimization of express/slow train operations in integrated suburban and urban rail systems based on passenger demand, Journal of Beijing Jiaotong University 48(4)(2024) 153-163.
[13] Y. Chen, Y. Bai, B.-H. Mao, F. Wen, Optimization of Train Routing Plan for Interconnected Lines in Urban Rail Transit, Journal of the China Railway Society 46(8)(2024) 21-29.
[14] W.-M. Wang, X.-F. Du, X. Wan, Research on LTE-M technology planning of CBTC system in Dalian metro Line 13, Wireless Internet Science and Technology 21(2)(2024) 13-15.
[15] D.-J. Xu, B.-H. Mao, S.-K. Chen, L. Gong, J.-W. Zeng, Optimization of operation scheme for full-length and short-turn routings considering operation proportion, Journal of Traffic and Transportation Engineering 21(2)(2021) 173-186.
[16] W.-X. Zha, J.-Y. Zhou, J. Li, L.-X. Yan, Two-layer optimization model of train rescheduling for high-speed railway based on preferred penalty value of arrival-departure track allocation, Journal of Beijing Jiaotong University 47(6)(2023) 13-20.
[17] T. Li, S.-J. Zhang, J.-L. Sun, Y. Zhang, F. Jia, A Research on Optimization of Passenger Station Arrival & Departure Tracks Utilization based on IGA, Railway Transport and Economy 42(6)(2020) 30-37.