WILEY | Hindawi

Research Article

Multiobjective Collaborative Optimization Method for the Urban Rail Multirouting Train Operation Plan

Lianbo Deng [ID],[1] Qi Peng [ID],[1] Li Cai,[1] Junhao Zeng,[1] and Nava Raj Bhatt[1,2]

[1]*School of Traffic and Transportation Engineering, Rail Data Research and Application Key Laboratory of Hunan Province, Central South University, Changsha 410075, China*
[2]*Department of Railways, Government of Nepal, Bishalnagar, Kathmandu 999098, Nepal*

Correspondence should be addressed to Lianbo Deng; lbdeng@csu.edu.cn

The Train Operation Plan (TOP) of urban rail transit (URT) is a comprehensive plan for the operation of trains, the use of facilities and equipment, and the organization of other operational tasks. The TOP should not only be formulated in terms of time-varying passenger flow periods, but it should also be arranged to consider the substitutability of trains between multiple routes combined with the passenger choice. Based on the principle of "operating by the flow" and the requirement for precise allocation of transport capacity for multiple routes, this article constructs a multiobjective nonlinear integer programming model by taking the minimized generalized travel cost of passengers, total running mileage of trains, fluctuation of trains for each route (as optimization targets), and the combination of requirements of both headways and fully loaded rates as constraints. A multiobjective genetic-based algorithm is designed to simultaneously optimize the TOP and the two-way train stopping time in each period. Finally, the proposed model and algorithm are validated with the real data from the Guangzhou Metro Line 2. The results show that the Pareto optimal TOP and dynamic train stopping time are significantly improved compared to the original values.

1. Introduction

With the continuous advancement of urbanization in China and the rapid development of urban rail transit networks, the network effects of rail transit have emerged. Compared with the buses, URT is more attractive, stable, and punctual. The daily traffic volume of the URT system has gradually increased and its share rate has continued to rise compared with various transit modes [1, 2]. The Train Operation Plan (TOP) is an important intermediate link between passenger flow demand and transportation supply, of which "operating by the flow" is the basic principle. Reasonable development of the train operation plan can effectively allocate the capacity resources of URT, making full use of the transport capacity of the network, meeting the differentiated spatio-temporal demands of passengers, and providing passengers with safe, punctual, and comfortable commuting services, while reducing the operating costs of transport companies. Due to the spatiotemporal transportation of passenger flow across the network, the imbalance of traffic volumes in-line levels is exacerbated. Compared with the previous single routing mode, i.e., between two endpoints, the operational method of multirouting is considered better for the un-balanced passenger flow requirements, while improving organizational efficiency of transportation.

The formulation of the URT Train Operation Plan is a hierarchical process of top-to-bottom preparation and bottom-to-top feedback [3], in which: (a) the strategic planning layer includes passenger flow analysis, the TOP for network, and the TOP for line; (b) the tactical development layer includes train diagram optimization, rolling stock circulation, and crew planning; (c) the operational organization layer includes train rescheduling and a corresponding adjustment scheme for emergencies and delays, as shown in Figure 1. In general, the upper layer in the planning process is the input of the lower one, and the lower layer in the feedback process is the input of the upper one. The TOP for a specific line determines the routing plan, train marshalling
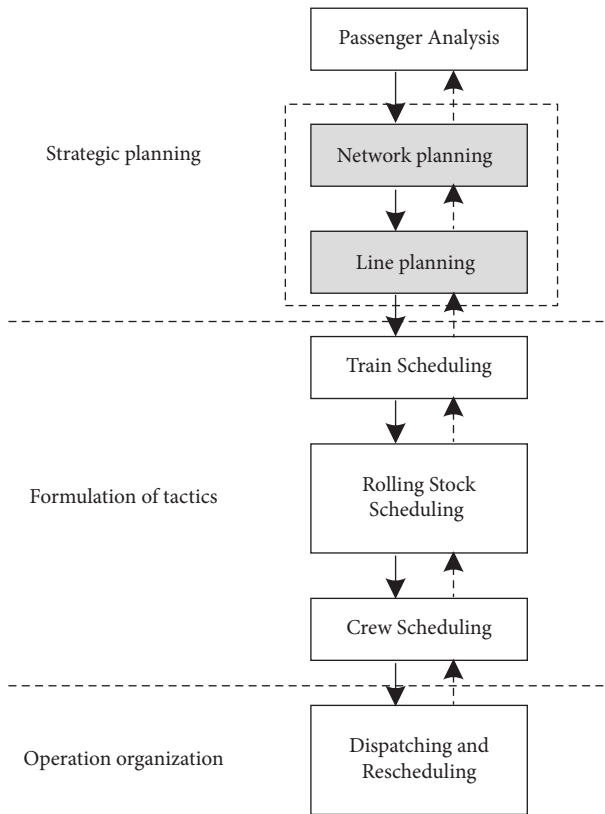
FIGURE 1: URT planning process.

plan, and service frequency, which is the focus of this article. The network TOP is developed after the coordinated optimization of the train operation plans for each line, where the interline interchanges are mostly considered at the high-volume transfer stations.

The train stopping time is a very important parameter for shortening the travel time and passengers' waiting time [4]. The train stopping time is not only the key factor in restricting the transport capacity of the URT system, it is also an important link which affects the quality of the passenger service, including whether the alighting and boarding movement is smooth and attractive to potential passengers. The formulation of the URT Train Operation Plan should not only consider the operational safety, technical requirements of network operation organization, transportation capacity, and operational benefit of the transport enterprise from the supply side, it should also take into account the distribution of passenger flow from the demand side to improve the transport service quality and reduce the passenger travel cost.

With the URT network established, the TOP is also a networked arrangement. Since the URT in China mainly adopts the operational method of passengers transferring at the transfer station, the formulation of the TOP for each line is relatively independent. Based on the results of the passenger flow assignment of the entire network within one day, this article aims to propose a TOP optimization model that is suitable for the multirouting mode. It focuses on a single URT line and an operational organization scheme that is consistent with most URT systems in China, including fixed vehicle size, invariant train routing, and an all-stop mode. According to the riding O-D (origin-destination) of the line obtained by the passenger flow pushing the assignment method [5], a multiobjective nonlinear integer programming model is constructed, with the optimization goal of minimizing the passenger's generalized travel cost, total travel mileage of trains, and fluctuation of trains for each routing in each period. By comprehensively considering the constraints of both the maximum and minimum headways and fully loaded rates, the TOP and the dynamic train stopping time in each period are collaboratively optimized. Using the model and algorithm proposed in the article, the TOP of Guangzhou Metro Line 2 is solved and the running process of the algorithm is analyzed. Finally, the obtained Pareto optimal solution and the output dynamic train stopping time are compared with the actual TOP and the original fixed stop time, respectively.

The rest of the article is organized as follows: First, the relevant studies regarding the formulation of the TOP of the multirouting mode and the calculation of stop time is summarized. Then, a multiobjective collaborative optimization model of the TOP and train stopping time is established, based on the analysis of generalized travel costs to passengers. Thirdly, a multiobjective genetic algorithm is designed for the model to solve the Pareto optimal solution, achieving the collaborative optimization of the multirouting TOP and train stopping time. The actual data from Guangzhou Metro Line 2 is then used as an example to verify the model and the algorithms. The final section presents the conclusions.

2. Literature Review

Existing studies have comprehensively taken factors such as the network structure, cost benefit, line type, transport service level, and organizational difficulty into account, on the basis of passenger flow, which were contained in the constructed planning models. In earlier studies, a single-objective mathematical programming model was usually established to generate the TOP, with the maximum direct passenger volume, the minimum cost, and minimum delay as the optimization goals [6, 7]. The TOP, nevertheless, is an organization plan for the distribution of passenger flows by transport enterprises. Essentially, it is a multiobjective optimization problem in which the stakeholders, represented by the supplier and the demander, play games on their respective benefits and costs. Therefore, considering the interests of multiple aspects in the model will mean that the decision-making plan will not be biased towards any particular party's position [8, 9].

Multirouting is a significant method of networking the operation and organization of URT when dealing with the large surging volume of passengers in certain periods. Vast numbers of passenger are unevenly distributed in some sections of the line during peak hours, so some trains are dedicated for special routing, e.g., the short-turning mode, skip-stop mode, and Y-type mode, which can accelerate the circulation of trains in those sections and reduce the

continuous accumulation of passengers at heavily crowded stations [10]. A multiobjective optimization model for a multiroute train plan is established, considering factors such as transport capacity, organizational requirements, business benefits, passenger demands, and selection behavior. It was proposed that a three-stage algorithm, based on the characteristics of the model and practical experience, should be used to solve the problem. It is worth noting that, in the current design of the bus network, passenger flow assignment and bus service frequency can be synchronously optimized and the balance between the passenger demand and transport supply can be achieved through interactions; lessons can be drawn from this. However, the passenger flow assignment, which only considers the service frequency, is difficult to directly apply to the formulation of the URT TOP [11].

Taking a high-speed railway without a feeder in Taiwan as an example, Chang [12] established a model with minimum operating costs and the lowest total travel time as the optimization goal; it used fuzzy mathematical programming to solve the best-compromise TOP. Ceder [13] and Canca et al. [14] proposed the application of long-short turning routing in different scenarios of the passenger flow demand. The former used the method of cyclic iteration to automatically generate the TOP of the long-short turning routing with the goal of minimizing the number of rolling stocks under a given train schedule diagram. The latter proposed a strategy of long-short turning routing to improve the delivery capacity of certain sections under the disturbance of the surge of passengers in some stations, thereby reducing the passenger waiting time and ensuring service levels.

The skip-stop operating mode is also of great importance, with respect to formulating the TOP for URT lines that link the suburbs with downtown. Freyss et al. [4] divided the passing stations into three categories and set the stop plan for the AB, A, and B stations to reduce the train stopping time, thereby increasing the travel speed of trains. They proposed that by coordinating parking stations and passing stations to build a skip-stop operating model, a genetic algorithm (including many actual scenes) can be designed to solve the model. The Seoul Metro example shows that the reduction in travel time also increases passenger waiting and transfer time [15]. In terms of the impact of unexpected disturbances on the actual operations, the multirouting mode (e.g., skip-stop route and long-short turning route) can be integrated in the model of timetable adjustment, so as to reduce the impact on passengers [16]. Jiang and Guo [17] analyzed the spatial distribution and passenger flow of the rural road network in China and showed that combining the trunk lines and branch lines into "Y", double "Y," or "8" types for integrated transportation can reduce operating costs and exhaust emissions and improve operating efficiency. Zhao et al. [18] proposed a routing planning model for URT Y-type lines, aiming to minimize passenger travel time and train operating distance, determine the turnaround station of the routing, and the departure frequency of multiple routes. They found that by converting multiple targets into a single target and solving them by a genetic algorithm, the results show that three train routes are more suitable for the characteristics of passenger flow on a Y-type line.

The train stop time is a critical control parameter for laying out the train diagram and it has an important influence on the delivery capacity of the diagram and the passenger service level. For most URT systems in China, the train stop time is usually set to a fixed value throughout the day [4]. This approach can simplify the difficulty of drawing a train diagram by facilitating the parameter management of the diagram, but is not conducive to improving the passenger service level. Normally, the fixed train stopping time is set to meet the passenger demand during peak hours. Therefore, for the smaller intensity of passenger flow during off-peak hours, it will increase the extra in-vehicle waiting time for passengers. Lin and Wilson [19] compared the fare payment methods of the light rail and bus systems. According to the actual data of the green line under Massachusetts Bay Transportation Authority (MBTA), the boarding passengers, alighting passengers, and in-vehicle passengers were used as the three analytical variables to explain the proposed train stop model, which indicated that the nonlinear form of congestion in the vehicle had a greater impact on the model. With the development of Tehran's URT network, Aashtiani and Iravani [20] believed that accurate estimation of the train stop time can yield more precise results of the passenger flow assignment. Therefore, a stop time model based on boarding and alighting passenger flow, congestion degree, and number of train doors was established. Karekla and Tyler [21] explored whether the reduction of train stop time in a busy subway system will improve the passenger service level and convenience, while reducing operating costs. Taking the step height between the train and the platform, passenger boarding and alighting time, door width, and the combination of the above as variables, four models were established. The results showed that the functional combination of the step height and the door width could effectively reduce the stop time and the train turnaround time, but it required higher reconstruction costs. D'Acierno et al. [22] defined the stop time as an analytical formula related to passenger flow and estimated the stop time based on the platform congestion degree and the interaction between the train and passenger to realize the dynamic setting of the stop time. Then, a strongly robust timetable was designed to improve the service quality and system attractiveness.

At present, although there has been much research on the network train operation plan and train stop time in urban rail transit, there are still several topics to be considered.

(1) The TOP is formulated according to the passenger flow demand. The service frequency mostly considers the matching of transport capacity and traffic volume, but the cost of entry-exit depot and the difficulty of transport organization are considered less.

(2) In case analysis, the situation during peak hours is usually taken as an example, and the trade-off between the differentiated setting of time segments and

the stability of the train operation plan is rarely considered.

(3) Some researchers realize that this is a multiobjective optimization problem but most of them use weighted linearization to simplify it.

(4) Much of the research on the stopping time mainly focuses on the relationship between hardware facilities, such as stations and vehicles, and the passenger flow on and off the train but ignores the relationship between the stop time and service frequency.

3. Description of TOP Optimization

The study of the TOP in this article is based on the following assumptions:

(1) During all passenger travel periods throughout the day, the basic train routing that runs between two endpoints must be operated and the nonbasic train routing that runs within a shorter range should be operated on the basis of the passenger demand. All trains of two routes stop at each station (the all-stop mode) and adopt a single vehicle type and a fixed formation, which is universal for URT systems.

(2) For a specific urban rail line, passengers generally do not transfer between the train routings but only select the train routing that includes passengers' origin-destination (O-D). This assumption is in-line with the selection pattern of urban rail passenger flow and the characteristics of operational organization.

(3) According to the assumption of "first come first service" (FCFS), passengers choose the train with the closest departure time, regardless of boarding delay. This assumption is applicable to the TOP that emphasizes the arrangement of transport capacity.

China's urban rail transit lines mainly use the nonjoint operation mode and the transport capacity of the TOP is also independently allocated according to the traffic demand of each line. We define that $N = (S, E)$ represents one URT line, where the station set $S = \{1, 2, ..., H_S\}$ and its increasing order implies the up direction. The segment set $E = \{e_i^j | i, j \in S\}$ and segment $e_i^j \in E$ represent the $O - D$ from the station i to the station j, of which the station spacing is w_i^j and the total running time is τ_i^j, including section traveling time and additional accelerating/decelerating time. It should be noted that $\varpi = 1$ indicates the up direction and $\varpi = 0$ indicates the down direction. Also, the onward section of the station i in the direction ϖ is $e_i^{i*} \in E$, where $i^* = i + 1$ when $\varpi = 1$ and $i \neq H_S$; $i^* = i - 1$ when $\varpi = 0$ and $i \neq 1$.

The service time of an URT line is expressed as $[T_s, T_e]$ during which the travel demand of passenger flow has a strong volatility, obviously indicating the converge-disperse and time-varying characteristics. In order to accurately reflect the spatiotemporal fluctuation characteristics of passenger travel demand, especially the description of it in key periods (e.g., peaks and troughs), and facilitate the development of TOPs at the same time, a step function can be used on the basis of collecting data from the automatic fare collection (AFC) system to describe the intensity of passenger travel demand. By counting the traffic demand in all periods throughout the day, the characteristics of turning points can be obtained, including peaks and troughs that have a great impact on train operation. The step function is used to divide the service time span of URT $[T_s, T_e]$ into H_t periods with relatively stable passenger flow. It can be considered that in each period, the passenger demand is evenly distributed with equal intensity, which is called the passenger flow travel period (PFTP). The set of PFTP for a whole day can be expressed as $T = \{T_K = (t_k^a, t_k^b) | k = 1, 2, ..., H_t\}$, where t_k^a and t_k^b are the start and end of T_k, respectively, and $t_1^a = T_s$, $t_{H_t}^b = T_e$. $|T_k|$ defines the length of one PFTP.

The demand basis of the TOP of an URT line is the set of full-day passenger flow boarding and alighting within the station, S: $W = \{f_{kxy\varpi} | k = 1, 2, ..., H_t; x, y \in S; \varpi \in \{0, 1\}\}$, where $f_{kxy\varpi}$ represents the passenger volume originating from station x to station y during the PFTP T_k.

The TOP of an URT line can be expressed as the service frequency of each train routing during each PFTP in a day. In general, a train operation period is a period during which trains depart at regular intervals and contain more than one PFTP [10]. As the degree of spatiotemporal differences in passenger flow continues to increase under networking conditions, refined distribution of transport capacity can effectively boost efficiency [23, 24]. Thereby the train operation period is expressed as PFTP in this article. In each PFTP, the corresponding number of trains is determined and a reasonable time division is selected to ensure the accuracy of passenger flow description and to facilitate the formulation of TOP.

According to assumption (1), the train capacity is V, the train formation is b, the number of doors of each coach is c, the average travel speed is v, and the upper and lower limits of the fully loaded rates and headways of each train are φ_{max}, φ_{min} and $\overline{\eta}$, η, respectively.

URT trains, in most parts of the world, generally do not run across lines. The train routing is set on one URT line $U = \{u_m = e_{\underline{s}_m}^{\overline{s}_m} | m = 1, 2, ..., H_u; \underline{s}_m, \overline{s}_m \in S\}$, where $\overline{s}_m$ and $\underline{s}_m$ are the endpoints of the routing and u_1 is the basic routing that turns back at both ends of the line.

Among all train routings, the train routing that must be operated during all PFTPs throughout the day is called "basic routing" and the other routings are called nonbasic routings. For example, in the operation mode of long and short routing, the short-turning routing usually serves the sections where the cross-section ridership is sharply higher than other parts of the line and is called nonbasic routing, which could be replaced by the basic routing. The TOP of the line can be defined as $D = \{d_k^m | k = 1, 2, ..., H_t; m = 1, 2, ..., H_u\}$, where d_k^m is the number of trains running on routing u_m during the PFTP T_k and the optimization object of this article.

The stopping time of the train at each station is set according to the demands of each PFTP and is divided into two parts: one is the fixed time for opening and closing the

door λ_0 and the other is the effective time for passengers to get on and off the train. The fixed time includes door opening time, door closing time (including notice time), unbalanced delay time (for boarding and alighting by each door), and start response time, after closing the door. In addition, when the platform is equipped with screen doors, the delay time of opening and closing the door will be increased. The effective time required for passenger boarding and alighting λ_1 is related to the congestion on the platform and in the coach. The effective boarding and alighting time for passengers is determined by the number of passengers getting on and off at the station during each PFTP and is calculated based on the average of the number of boarding and alighting passengers by each door.

4. Analysis of Generalized Travel Cost of Passengers

During the PFTP T_k, the cross-section passenger flow of the section $e_i^{i^*} \in E$ in the ϖ direction is shown in the following equation:

$$g^S\left(e_i^{i^*}, k, \varpi\right) = \left\{ \sum_{y=i^*}^{H_S} \sum_{x=1}^{i} f_{kxy\varpi}\varpi = 1 \sum_{y=1}^{i} \sum_{x=i^*}^{H_S} f_{kxy\varpi}\varpi = 0. \right.$$

(1)

During the PFTP T_k, the boarding passenger flow of station $i \in S$ in the ϖ direction is shown in the following equation:

$$g^B(i, k, \varpi) = \sum_{x=i} f_{kxy\varpi} \quad y \in S; k = 1, 2, \ldots, H_t; \varpi = 0, 1.$$

(2)

The boarding passenger flow at the turning station of two directions is 0. In other words, for the up direction $\varpi = 1$ at the station $x = H_S$, the boarding passenger flow is $g^B(H_S, k, \varpi) = 0$; for the down direction $\varpi = 0$ at the station $x = 1$, the boarding passenger flow is $g^B(1, k, \varpi) = 0$.

During the PFTP T_k, the alighting passenger flow of station $i \in S$ in the ϖ direction is shown in the following equation:

$$g^A(i, k, \varpi) = \sum_{y=i} f_{kxy\varpi} \quad x \in S; k = 1, 2, \ldots, H_t; \varpi = 0, 1,$$

(3)

where the alighting passenger flow at the starting station of two directions is 0. In other words, for the up direction $\varpi = 1$ at the station $y = 1$, the alighting passenger flow is $g^A(1, k, \varpi) = 0$; for the down direction $\varpi = 0$ at the station $y = H_S$, the alighting passenger flow is $g^A(H_S, k, \varpi) = 0$.

Generalized travel cost of passengers mainly includes the cost of fares, travel time, and congestion expenses. China's URT mainly adopts the mileage pricing strategy [25]. The passenger's fare is determined by the O-D and is independent of the routes, even though there may be more than one route for a particular O-D pair on the URT network. This means that the passengers who choose to take a train on a particular line are not motivated by the fare, which can be regarded as a constant C in the generalized cost perceived by passengers and is irrelevant to the number of trains in each PFTP.

The passenger travel time cost includes the in-vehicle travel time cost and waiting time cost. The in-vehicle travel time is determined by the running time of the journey sections and the waiting time of stopping stations. The total travel time of loaded trains in each PFTP T_k throughout the day is the product of the cross-sectional passenger flow and the running time between sections, as shown in the following equation:

$$\Gamma^{run} = \sum_{\varpi=0}^{1} \sum_{k=1}^{H_t} \sum_{i \in S} g^S\left(e_i^{i^*}, k, \varpi\right)\tau_i^{i^*}.$$

(4)

The total waiting time caused by the train stopping at each station for each PFTP T_k throughout the day is given by the following equation:

$$\Gamma^{stop} = \sum_{\varpi=0}^{1} \sum_{k=1}^{H_t} \sum_{i \in S} \left(g^S\left(e_i^{i^*}, k, \varpi\right) - g^A\left(i^*, k, \varpi\right)\right)\lambda_{k,\varpi}^i.$$

(5)

The waiting time refers to the additional period that results from the inconsistency between the arrival of passengers and the arrival of trains. It is generally considered that the waiting time of passengers is related to the headway between trains ϕ in the onward section $e_i^{i^*} \in E^l$ of station i. The waiting time is a uniformly distributed random variable $(0, \phi)$. In the direction ϖ of the PFTP T_k, all trains whose routes cover the onward sections of the station i are combined to calculate the service frequency, which directly affects the headway between trains. Assuming that the operation interval of the PFTP T_k is $\phi(e_i^{i^*}, k, \varpi) = |T_k|/ \sum_{m \in \{m|e_i^{i^*} \subseteq u_m\}} d_k^m$, the average waiting time is $0.5\phi(e_i^{i^*}, k, \varpi)$. The total passenger waiting time of each PFTP T_k throughout the day is the product of the boarding passenger flow at each station and the average waiting time, as shown in the following equation:

$$\Gamma^{wait} = 0.5 \sum_{\varpi=0}^{1} \sum_{k=1}^{H_t} \sum_{i \in S} g^B(i, k, \varpi)\phi\left(e_i^{i^*}, k, \varpi\right).$$

(6)

For the station $i \in S$ in the direction ϖ, the congestion degree of the train's onward section $e_i^{i^*} \in E$ is affected by the train capacity V and the overlap of train routes that cover this section. The total congestion cost function caused by train crowding in each PFTP T_k throughout the whole day is shown in the following equation:

$$\Gamma^{jam} = \begin{cases} 0 & g^S\left(e_i^{i^*}, k, \omega\right) \leq \sum_{m \in \underset{\sim}{M}} V d_k^m \\ \\ \alpha \sum_{\omega=0}^{1} \sum_{k=1}^{H_t} \sum_{i \in S} \left(\tau_i^{i^*} + \lambda_{k,\omega}^i\right) \left(\dfrac{g^S\left(e_i^{i^*}, k, \omega\right)}{\sum_{m \in \widetilde{M}} V d_k^m}\right)^{\beta} & g^S\left(e_i^{i^*}, k, \omega\right) > \sum_{m \in \underset{\sim}{M}} V d_k^m, \end{cases} \tag{7}$$

where α and β are the parameters to be calibrated and the set of $\widetilde{M}$ means $\widetilde{M} = \{m | e_i^{i^*} \subseteq u_m\}$. The calibration values recommended by the US Federal Highway Administration [26] are $\alpha = 0.15$ and $\beta = 4$.

5. Establishment of a Multiobjective Collaborative Optimization Model

The TOP of URT requires comprehensive consideration of transport capacity, transport organization requirements, spatiotemporally differentiated passenger demand, and transport costs. The constraints include the train's fully loaded rate and the headway between train departures. The optimization objective mainly includes the generalized travel cost of passengers, the cost of operating the train, and the fluctuation of the train service frequency, which can be considered as a multiobjective optimization problem. As mentioned earlier, the TOP number of trains per route in each PFTP directly affects the stopping time. In order to improve the overall optimization quality, this article adopts the dynamic stopping time to meet the passenger demand, and collaboratively optimizes the TOP and bidirectional stopping time at each station in each PFTP.

Based on the requirements of transport capacity and the service level, transportation enterprises usually use the upper limit of the load rate $\varphi_{\max}$ and the lower limit of the load rate $\varphi_{\min}$ to control the train running cost and allocate transport capacity. During the PFTP, the load rate of the train varies from section to section due to the spatial imbalance in the passenger flow distribution. In general, the train capacity requirement is limited only to the maximum cross-section passenger flow to ensure that the capacity rate is within a reasonable range.

It is assumed that during PFTP T_k, the section with the largest ridership in the ω direction of the line is $\overline{e}_i^{i^*}$. The corresponding maximum cross-section passenger flow is given by the following equation:

$$g^S\left(\overline{e}_i^{i^*}, k, \omega\right) = \max\left\{g^S\left(e_i^{i^*}, k, \omega\right) | k = 1, 2, \ldots, H_t\right\}. \tag{8}$$

In the case of multiple crossing routes, the total number of trains in each crossing route should not exceed the capacity determined by the signal system, the switch of turnout, and the turn-back mode. Therefore, when considering the maximum loaded rate and the minimum departure interval at the same time, the constraint should satisfy the following equation:

$$\max\left\{\min\left\{\left\lceil\dfrac{g\left(\overline{e}_i^{i^*}, k, \omega\right)}{\varphi_{\max} V}\right\rceil, \left\lfloor\dfrac{|T_k|}{\eta}\right\rfloor\right\}, \left\lceil\dfrac{|T_k|}{\overline{\eta}}\right\rceil\right\} \leq \sum_{m=1}^{H_u} d_k^m \leq \min\left\{\max\left\{\left\lfloor\dfrac{g\left(\overline{e}_i^{i^*}, k, \omega\right)}{\varphi_{\min} V}\right\rfloor, \left\lceil\dfrac{|T_k|}{\overline{\eta}}\right\rceil\right\}, \left\lfloor\dfrac{|T_k|}{\eta}\right\rfloor\right\} \tag{9}$$

$$i \in S; k = 1, 2, \ldots, H_t.$$

According to the policy requirements and the need to improve the passenger service level, the number of trains d_k^1 on the basic routing during the PFTP T_k should meet equations (10) and (11) when considering the minimum loaded rate and the maximum departure interval.

$$\left\lceil\dfrac{|T_k|}{\overline{\eta}}\right\rceil \leq d_k^1 \leq \min\left\{\max\left\{\left\lfloor\dfrac{g\left(\overline{e}_i^{i^*}, k, \omega\right)}{\varphi_{\min} V}\right\rfloor, \left\lceil\dfrac{|T_k|}{\overline{\eta}}\right\rceil\right\}, \left\lfloor\dfrac{|T_k|}{\eta}\right\rfloor\right\} \quad i \in S; k = 1, 2, \ldots, H_t, \tag{10}$$

$$d_k^m \geq 0, d_k^m \in Z \ k = 1 \ldots, H_t; m = 1, \ldots H_u, \tag{11}$$

where Z means the integer set.

The passenger is the service object of the TOP, so the optimization objective is to minimize the total generalized travel cost of the passenger, which is given by the following equation:

$$\min Z_1 = C + \beta\left(\Gamma^{run} + \Gamma^{stop} + \Gamma^{wait}\right) + \Gamma^{jam}. \tag{12}$$

In the above formula, the passenger's fare expenditure and in-vehicle time are independent of the number of trains d_k^m running in each PFTP T_k and routing u_m of the TOP, which is equivalent to the constant in the objective function,

so they can be ignored. The objective function related to the TOP can be simplified as follows:

$$\min Z_1 = \beta\left(\Gamma^{\text{stop}} + \Gamma^{\text{wait}}\right) + \Gamma^{\text{jam}}. \tag{13}$$

The low-cost and high-efficiency operation of URT is an effective measure to alleviate the pressure of government financial subsidies. It must not only meet the goals of social interests but also promote the efficiency of transport enterprises from the level of government subsidies mechanism [27, 28]. It is of great significance to reduce the train operating cost at the level of TOP, i.e., reduce the total train mileage while meeting the passenger flow demand. The optimization objective is to minimize the travel distance of trains crossing each route during all passenger travel periods, as shown in the following equation:

$$\min Z_2 = \sum_{k=1}^{H_t} \sum_{m=1}^{H_u} d_k^m w_{\underline{S}_m^m}^S. \tag{14}$$

In order to meet the needs of different passenger travel periods, the number of trains on each transit route can be determined during the passenger travel period to achieve a refined capacity allocation. However, when the number of trains on the same route is different during adjacent passenger flow periods, the insufficiency of rolling stock needs to be addressed for the exit-depot operation in advance or the excess of rolling stock needs to be sent back to the depot, i.e., there are additional running miles of rolling stock for exit-depot and entry-depot. When the number of trains in several adjacent periods is not the same, it will cause the frequent operation of train entering and leaving the depot, which not only produces unnecessary additional operating costs but also increases the difficulty of transport organization. Therefore, in order to facilitate the transport organization of the URT operation department, the optimization objective should be to minimize the fluctuation in rolling stock for exit-depot and entry-depot during the adjacent passenger flow periods to avoid unnecessary running miles of rolling stock for exit-depot and entry-depot during the periods. For the multirouting transport mode, taking the long-short route strategy as an example, when the train departure intervals of the long and short routes do not meet the integer multiple relation, the train running diagram will generate empty time and waste the line passing capacity [29, 30].

In the PFTP T_k, the operating ratio of short (nonbasic routing) and long (basic routing) routes is 2 : 1, as shown in Figure 2.

The departure interval h_L of the long-routing train is an integer multiple of the departure interval h_S of the short-routing train, where T_C^L and T_C^S are the total circulation durations, respectively. The rolling stock in use on the long route is $n_k^L = \left\lceil T_C^L d_k^L / |T_k| \right\rceil$, while the rolling stock in use on the short route is $n_k^S = \left\lceil T_C^S d_k^S / |T_k| \right\rceil$. The total rolling stock can be expressed as $n_k = n_k^L + n_k^S$. Since the length $|T_k|$ of the PFTP cannot be divided exactly by the departure interval, both n_k^L and n_k^S need to be rounded up.

In general, if the total circulation durations of the route u_m is t_z^m, then the rolling stock of each route can be calculated by the following equation:

$$n_k^m = \left\lceil \frac{t_z^m d_k^m}{T_k} \right\rceil \quad k = 1, 2, \ldots, H_t; m = 1, 2, \ldots, H_u. \tag{15}$$

In the transition of the passenger travel time, the number of trains on different routes is usually inconsistent, due to the fluctuation of passenger flow, especially in peak hours. When the number of trains in the current and later time periods fluctuates greatly, it will cause additional running miles of rolling stock for exit-depot and entry-depot, which will increase the difficulty of the rolling stock turnover plan and make the TOP less feasible and practical. Therefore, the third objective function is to encourage the replacement of the basic route with the nonbasic route when the constraints are met, while ensuring the number of trains for each route in the adjacent periods is as balanced as possible, which can be expressed as follows:

$$\min Z_3 = \sum_{k=1}^{H_t-1} \sum_{m=1}^{H_u} w_m \exp\left[-\left(\overline{n}^m - \left(\left|n_k^m - \overline{n}^m\right| + \left|n_{k+1}^m - \overline{n}^m\right|\right)\right)\right], \tag{16}$$

where w_m is the round-trip distance of the train from the depot or parking lot to the starting station of the train routing u_m and $\overline{n}^m = 1/2\left(n_k^m + n_{k+1}^m\right)$ is the average number of trains on the same route in the adjacent passenger flow periods. Thus, it can be seen that the value of the objective function is obviously affected by the number of the trains in the adjacent passenger flow periods and their balance.

6. Design of the Multiobjective Genetic Algorithm

Although the three objective functions of this article have cost considerations, their weights are difficult to determine accurately due to the different economic factors behind them. At the same time, in the case of more than three objective functions, the game relations are too complex to solve the problem. The constraints of the problem take into account the train load rate, the running interval, the number of large and small trains, etc. The solution variables of the problem for the number of trains are all integers, and its solution space is not continuous and analytical, and through the analysis of the total number of allocated trains in the TOP during each PFTP, it can be found that the number of trains on each route is subject to the total number, while there is a strong coupling relationship between the objective functions. Changing the number of trains on any route in any PFTP independently causes the objective functions to change in different directions, thus constituting a multi-objective optimization problem for train operation planning. It is difficult to directly solve the model by combinatorial optimization and analytical methods, and even more difficult to obtain the multiple Pareto optimal solution for the multiobjective problem. Since the 1990s, genetic algorithms have been widely used and continuously
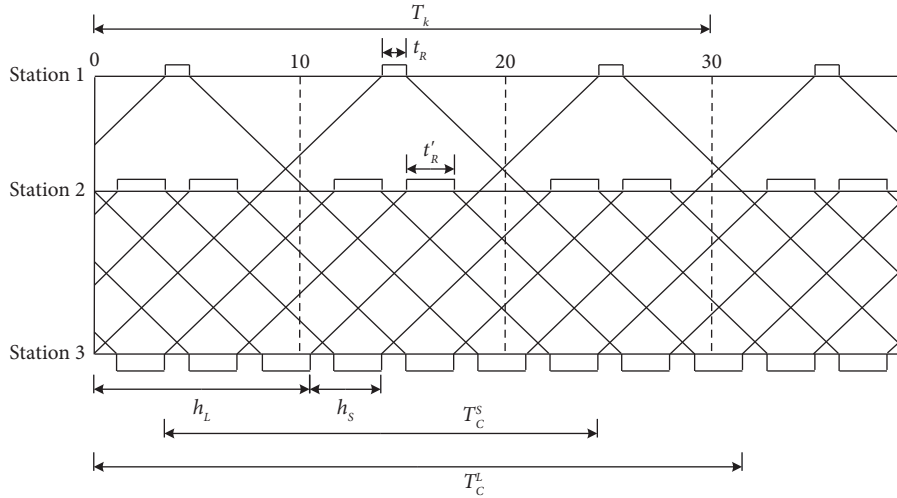
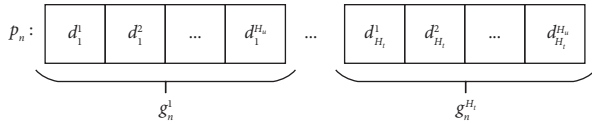Figure 2: Schematic train diagram of the long-short route.

Figure 3: Example of gene grouping.

developed in the field of multiobjective optimization due to the characteristics of multipoint and multidirectional searches [31–34]. This article takes advantage of the general applicability of the genetic algorithm to complex programming problems to solve complex optimization problems around the features of the model and design the adaptive weight genetic algorithm (AWGA) to solve the Pareto optimal solution set of the model.

6.1. Adaptive Weight Genetic Algorithm. The fitness allocation mechanism of the adaptive weight multiobjective genetic algorithm automatically adjusts individual weights by effectively using individual information in each generation of the population in the genetic process, so that Pareto optimal solutions are constantly approaching the ideal point. Since the better individuals in the current population are more likely to be retained in the next generation under the effect of adaptive weights, this ability to actively seek optimization has made the AWGA algorithm widely used in the field of multiobjective optimization after Gen and Cheng proposed it in 2000 [35, 36]. The steps are as follows:

Step 0: Set the population size *popSize*, crossover rate p_C, mutation rate p_M, maximum generation max *Gen*, and initial evaluation function value min *Eval*.

Step 1: Generate a chromosome that satisfies the constraints of the multiobjective optimization problem (coding).

Step 2: The objective function $f_k(x)$ of the multiobjective optimization problem is calculated according to the chromosome (decoding).

Step 3: According to the value of each objective function $f_k(x)$, a Pareto optimal solution is generated.

Step 4: The adaptive weight method is used to evaluate and select each chromosome.

Step 5: Let *gen*+ = 1.

Step 5.1: Crossover
The following single-point crossover method is used to perform the crossover action.
(1) Let *cCnt* = 0 (the number of chromosomes generated by the crossover).
(2) Generate a random number list r_k ($k = 1, 2, ..., popSize$) in the range of [0, 1] and select the chromosome $v_k[\cdot]$ that satisfies $r_k < p_C$.
(3) Pair the selected chromosomes and make *cCnt*+ = 2.
(4) Randomly determine the position and part of the crossover, so that the newly generated chromosomes are $v'_{cCnt-1}[\cdot]$, $v'_{cCnt}[\cdot]$, respectively.
Step 5.2: Mutation
Perform the mutation as follows:
(1) Let *mCnt* = 0 (the number of chromosomes generated by mutation). Generate a random number list r_k ($k = 1, 2, ..., popSize$) in the range of [0, 1], select the chromosome that satisfies $r_k < p_M$, and mutate according to certain rules.
(2) Let *mCnt* = 1 and the newly generated chromosome be $v'_{cCnt+mCnt}[\cdot]$.
Step 5.3: Decode the newly generated chromosomes and update the Pareto optimal solution *E*.
Step 5.4: The adaptive weighting method is used to evaluate and select each chromosome to form a new population of the next generation.

Step 6: If $gen < \max Gen$, return to Step 5; otherwise, output the Pareto optimal solution $E(\cdot)$ and terminate the process when the termination condition is satisfied.

6.2. AWGA Design of the Optimal Model for the TOP

6.2.1. Chromosome Design.
Considering the dynamic fluctuation of passenger flow during each PFTP of the whole day, the corresponding feasible domains of the number of trains d_k^m on each routing are also different. Under the effect of the integer constraint of decision variables in the model, compared with binary coding, the coding space formed by the decimal chromosomes is more suitable for the dynamic correspondence of the solution space. The efficiency of the crossover and mutation actions of the algorithm is significantly improved by using decimal coding of the chromosomes to represent the decision variable $d_{k,m}^i$. Chromosome p_n consists of $H_t \times H_u$ genes, as shown in the following equation:

$$p_n: \left[d_1^1, d_1^2, \ldots, d_{H_t}^{H_u} \right] \quad n = 1, 2, \ldots, p_S, \tag{17}$$

where p_S is the size of the population. $p_n[j]$ represents the jth gene of chromosome p_n; $p_n[i: j]$ represents the gene segment from the ith to the jth gene of chromosome p_n; $p_n[i] + p_n[j]$ represents the merged i^{th} and j^{th} genes of chromosome p_n, $i, j \in [1, H_t \times H_u]$.

Gene values are randomly distributed in the initial state and the primary goal of optimization is to make all genes satisfy the constraint conditions (9)–(11). The advantage of decimal encoding is that the encoding space is more intuitive. When the constraint conditions change dynamically, it can avoid recorrection due to the noncorrespondence of the decoding space with the solution space. For PFTP T_k, the number of trains on each routing is $d_k^1, \ldots, d_k^{H_u}$. Due to the restriction of the above constraints, the resulting gene values are not completely independent and embody the interdependence and mutual restriction between genes, which has a high degree of coupling. Therefore, the chromosome p_n can be divided into H_t genomes $g_n^k = \left\{ d_k^1, d_k^2, \ldots, d_k^{H_u} \right\}$ according to the PFTP T_k, $n \in [1, p_S]$, $k \in [1, H_t]$. In the genetic process such as crossover and mutation, operations are performed in units of genome, as shown in Figure 3.

6.2.2. Single Genome Crossover.
Crossover is the most effective operation to improve genetic coding.

Among the many crossover methods, single-point crossover has wide applicability due to its simple operation and effectiveness. When solving the optimization model of the TOP, a classical single-point crossover is likely to break the interdependence of genes within the genome and cannot meet the constraints. Therefore, in order to completely retain the coupling relationship of the number of trains on each routing during a PFTP, crossover operations can be performed with the genome as the smallest unit. This means that for any two chromosomes of the current population, such as p_{n1} and p_{n2}, randomly select a PFTP T_k ($k = 1, 2, \ldots, H_t$), of which the genome $g_n^k = \left\{ d_k^1, d_k^2, \ldots, d_k^{H_u} \right\}$ is the starting position and the gene d_k^1 is the breakpoint, to perform the single genome crossover. In this process, the right part of the breakpoint is exchanged to form the offspring chromosome, as shown in Figure 4.

6.2.3. Directed Mutation.
In the early stages of the algorithm search, some highly adaptive chromosomes may dominate the roulette selection process, making the local optimal solution in the population become the main body. In the later searching stage, the difference between chromosomes is small. If the mutation rate is small, it is difficult to generate new chromosomes, but increasing it may produce the concussion of objective functions.

The initial population for solving the optimization model of TOP is randomly generated. In the early stage of the algorithm iteration, it is mainly to search for solutions that satisfy the constraints. Because the number of constraints on the model is proportional to the number of PFTPs, i.e., the finer the characteristics of the temporal distribution of passenger flow, the more constraints there are. Therefore, when the crossover and mutation operations are completely random, it is difficult for chromosomes to evolve at an exact position and in an accurate direction, and local optimal solutions easily dominate the population. For the entire multiobjective genetic algorithm, each iteration is an opportunity for evolutionary optimization. If the population stays in the local optimal state, it is a waste of evolutionary opportunities. Therefore, evolution of the population can continue through artificial intervention.

Based on the above conception, on the basis of the classical stochastic mutation operation, the characteristics of directed mutation are added. The specific method is to monitor the state of the Pareto optimal solution in the population and set a threshold n_c^*. When the generation counts for the state of Pareto optimal solution not being improved exceed the threshold, namely, $n_c > n_c^*$, this indicates that the local optimal solution is dominant in the population. The PFTP $T_{k'}$ of the genome that does not meet the constraints in the current Pareto optimal solution are put into the set $\widetilde{T} = \left\{ k' | 1 \leq k' \leq H_t \right\}$ and the direction of mutation $\mu_{k'}$ of those genomes, $k' \in \widetilde{T}$, are determined. As for the mutation operation of the next generation, according to the set $\widetilde{T}$ and the corresponding mutation direction $\mu_{k'}$, the location and direction of mutation of labeled genomes in the Pareto optimal solution are explicitly determined, while the stochastic mutation is still adopted for the nonlabeled genome. Directed mutation can make the genome $k' \in \widetilde{T}$ satisfy the constraint (9) directly and constraint (10) indirectly, thus breaking the local optimal deadlock and the genetic evolution can continue effectively. The expression of $\mu_{k'}$ is shown in the following equation:
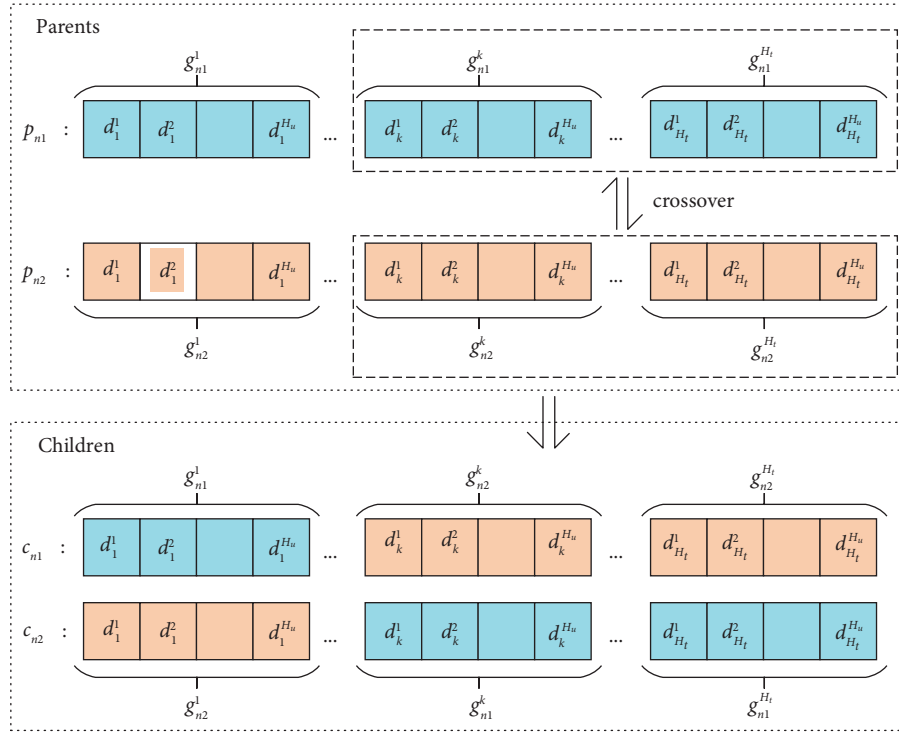
FIGURE 4: Single genome cross-over.

$$\mu_{k'} = \left\{ 1 \quad \sum_{m=1}^{H_u} d_{k'}^m < \max \left\{ \min \left\{ \left\lceil \frac{g\left(\bar{e}_i^{i*}, k', \omega\right)}{\varphi_{\max} V} \right\rceil, \left\lfloor \frac{|T_{k'}|}{\eta} \right\rfloor \right\}, \left\lceil \frac{|T_{k'}|}{\bar{\eta}} \right\rceil \right\} - 1 \sum_{m=1}^{H_u} d_{k'}^m > \min \left\{ \max \left\{ \left\lfloor \frac{g\left(\bar{e}_i^{i*}, k', \omega\right)}{\varphi_{\min} V} \right\rfloor, \left\lceil \frac{|T_{k'}|}{\bar{\eta}} \right\rceil \right\}, \left\lfloor \frac{|T_{k'}|}{\eta} \right\rfloor \right\} 0 \text{else}, \right.$$

(18)

where $\mu_{k'} = 1$ means that the total number of trains running in PFTP $T_{k'}$ should be increased, i.e., increasing mutation; $\mu_{k'} = -1$ means that the total number of trains running in PFTP $T_{k'}$ should be reduced, i.e., reducing mutation; and $\mu_{k'} = 0$ means that the genome k' satisfies constraint (9) but no other constraints, so only the nondirectional stochastic mutation is required.

6.2.4. Adaptive Weight Evaluation Method.

When using the AWGA to solve the multiobjective optimization problem, one of the most important points is to correctly evaluate and select the Pareto optimal solution obtained in each generation and keep it for the next generation. The adaptive-weight evaluation method proposed by Gen et al. [35, 37] can effectively use the positive ideal points obtained in each generation of population and actively adjust the weights to make the Pareto optimal solution close to the ideal point to search for the solution.

After decoding feasible chromosomes of a certain generation, the values of three objective functions are calculated and the maximum and minimum ideal points are defined as follows:

$$z^+ = \left[z_1^{\max}, z_3^{\max}\right],$$
$$z^- = \left[z_1^{\min}, z_2^{\min}, z_3^{\min}\right],$$

(19)

where $z_k^{\max}$ and $z_k^{\min}$ represent the maximum and minimum values of the kth objective function, $k \in [1, 3]$, defined as

$$z_k^{\max} = \max \left\{z_k(x) | x \in P\right\}, k = 1, 2, 3,$$
$$z_k^{\min} = \min \left\{z_k(x) | x \in P\right\}, k = 1, 2, 3,$$

(20)

where P is the set of feasible solutions. Therefore, the evaluation function formed by the sum of the adaptive weights of any chromosome x can be obtained by equation (21). This evaluation function realizes the non-dimensionalization of each objective function, so they can be directly added together.

$$z(x) = \sum_{k=1}^{3} w_k \left(z(x) - z_k^{\min}\right) = \sum_{k=1}^{3} \frac{z(x) - z_k^{\min}}{z_k^{\max} - z_k^{\min}},$$

(21)

where the adaptive weight w_k of the kth objective function is calculated by equation (22). In this article, the model seeks the minimization of objective functions. So, when the value of the kth objective function approaches its minimum, the evaluation function can reflect the fact that the chromosome has a high fitness value, which is to say that the probability of the elite individuals retained to the next generation is higher.

$$w_k = \frac{1}{z_k^{\max} - z_k^{\min}}, k = 1, 2, 3. \tag{22}$$

The AWGA for the TOP optimization model designed in this article fully considers the characteristics of the problem and makes corresponding adjustments on its original framework. The main flow of AWGA is shown in Figure 5.

7. Case Study

The model and algorithm proposed in this article are verified by the network and passenger data from Guangzhou Metro Line 2 in China. Line 2, the second of the 15 lines used at Guangzhou Metro, starts from Guangzhou South Station and ends at Jiahewanggang Station and has a roughly "S-shaped" north-south trend, running through four administrative regions: Panyu, Haizhu, Yuexiu, and Baiyun in Guangzhou. This line directly connects Guangzhou South Station and Guangzhou Railway Station and can reach Baiyun International Airport in one transfer; therefore, it is a core line linking major transportation hubs in Guangzhou.

The Guangzhou Metro Line 2 has a total length of 31.8 km, with 24 underground stations set up and train formations of 6 A-type coaches, as shown in Table 1. The stations with turn-back capabilities along the line include Jiangxia, Sanyuanli, Gongyuanqian, Jiangtai Road, and Nanpu. In the first seven months of 2019, Line 2 topped the list for all lines with an average daily passenger volume of 1.413 million passengers. This implies a high amount of pressure to ensure the transport security and necessity of finely organizing the transport capacity.

According to the results obtained from the assignment of the passenger flow of the URT network [5], sections of Line 2 during peak hours (when surging traffic volumes are mainly caused by commuters in the morning and evening) have the following characteristics: (1) the passenger flow is obviously centripetal, (2) the peak values of passenger flow between the morning and evening periods are close and the corresponding sections are the same; (3) the distribution trends of the passenger flow in the morning peak and evening peak in both the up and down direction are similar. Maximum cross-sectional passenger flow in the morning peak is slightly larger than the amount in the evening peak, which shows a certain symmetry, as shown in Figure 6. In 2019, Line 2 adopted a mixed mode of long and short routing in the morning peak periods and single-turning routing in other periods. The coverage of short-turning routing is the section between station 8 and station 17.

The optimization model of the URT Train Operation Plan is to be solved by the AWGA proposed in this article. The basic parameter settings of the algorithm are given in Table 2.

As described in the section for formulating the model, a large number of parameters should also be assigned values, as shown in Table 3.

The parameters in Tables 2 and 3 are put into the TOP optimization model and AWGA and an iterative calculation is started for the desirable plan. The running environment of the algorithm proposed in this article is the Windows 10 OS
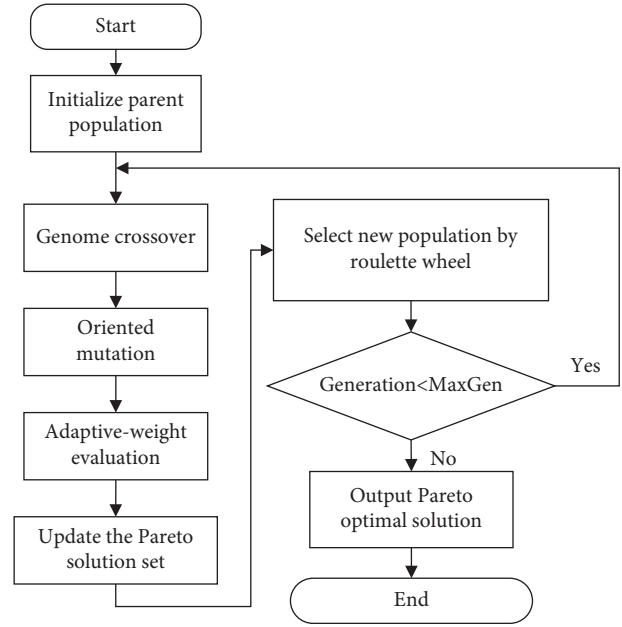

Figure 5: AWGA algorithm overall process.

Table 1: Station list of Line 2.

Stations	Number
Guangzhou South Railway Station	1
Shibi	2
Huijiang	3
Nanpu	4
Luoxi	5
Nanzhou	6
Dongxiaonan	7
Jiangtai Lu	8
Changgang	9
Jiangnanxi	10
The 2nd Workers' Cultural Palace	11
Haizhu Square	12
Gongyuanqian	13
Sun Yat-sen Memorial Hall	14
Yuexiu Park	15
Guangzhou station	16
Sanyuanli	17
Feixiang Park	18
Baiyun Park	19
Baiyun culture square	20
Xiao-gang	21
Jiangxia	22
Huangbian	23
Jiahewanggang	24

system, with an Intel (R) Core (TM) i5-6300U CPU, 8 GB RAM, SQL Server 2012 as a database and Microsoft IIS 6.0 as a network server. Because the AWGA is a random algorithm, the Pareto optimal solution obtained is not necessarily the same each time, but they are all noninferior solutions that satisfy the constraints. According to the random characteristics of the algorithm, there is little possibility that there are feasible chromosomes in the population at the initial
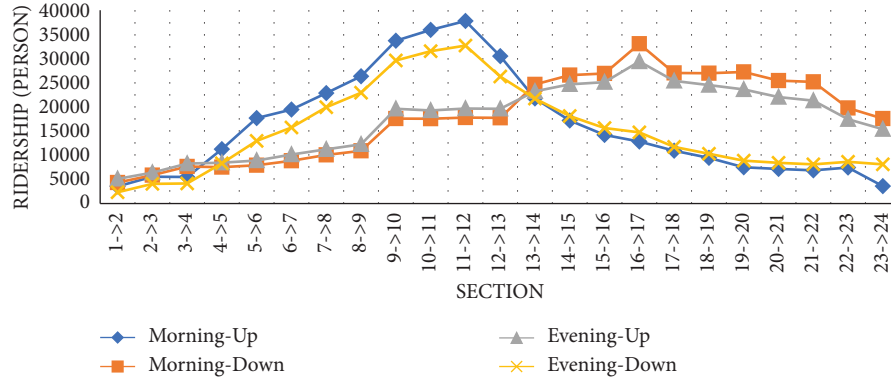
FIGURE 6: Cross-sectional ridership during the morning and evening peaks.

TABLE 2: Basic parameters of AWGA.

Parameters	Values
Population size p_S	10
Crossover rate p_C	0.4
Mutation rate p_M	0.05
Maximum generation $p_G^{\max}$	1000
Threshold of unchanged state n_c^*	10

TABLE 3: Model parameters.

Parameters	Values
Line l	2
Average time value of passengers	30¥/h
Start and end time of URT service T_s, T_e	T_s = 6: 00 and T_e = 23: 30
Duration of a PFTP $|T_k|$	30 min
Train capacity V	1860 persons, A-type train
Train formation b	6 coaches
Number of doors on each coach c	5
Effective boarding and alighting rate λ_1	0.6 s/person
Upper limit of fully loaded rate $\varphi_{\max}$	120%
Lower limit of fully loaded rate $\varphi_{\min}$	80%
Upper limit of operating headway $\overline{\eta}$	360 s
Lower limit of operating headway η	132 s
Circulation duration of long-turning routing t_z^L	107 min 24 s
Circulation duration of short-turning routing t_z^S	45 min 11 s
Distance of the round trip between the depart station and depot for long-turning routing w_L	3.2 km
Distance of the round trip between depart station and depot for short-turning routing w_S	27.85 km

phase, i.e., where all genes in the chromosome meet the constraint. Therefore, the genetic evolution direction of the algorithm is to first enable the chromosomes to meet the constraints, where the number of trains on each routing of each PFTP not only satisfy the distribution characteristics of passenger flow, but also meet the requirements of fully loaded rates and departure intervals. As shown in Figure 7, the dark blue curve represents the population size of each generation of chromosomes, which is composed of the parent chromosomes and their offspring chromosomes generated by crossover and mutation operations. The maximum and minimum population sizes are 29 and 17 chromosomes, respectively. The red curve represents the

chromosomes that satisfy the constraints, where the maximum size contains 21 feasible chromosomes in the 775th generation. The absolute number of feasible chromosomes is significantly less than the total population size in each generation. When inherited to the 177th generation, the first feasible chromosome meets the constraint which appears, of which the fitness value in that generation is obviously the largest. Then, the chromosome takes the dominant position and produces more feasible ones, from which Pareto optimal solutions are generated.

The trends for three objective functions in the model are shown in Figure 8, where the curves evolve synchronously and simultaneously. Since the exponential function is used
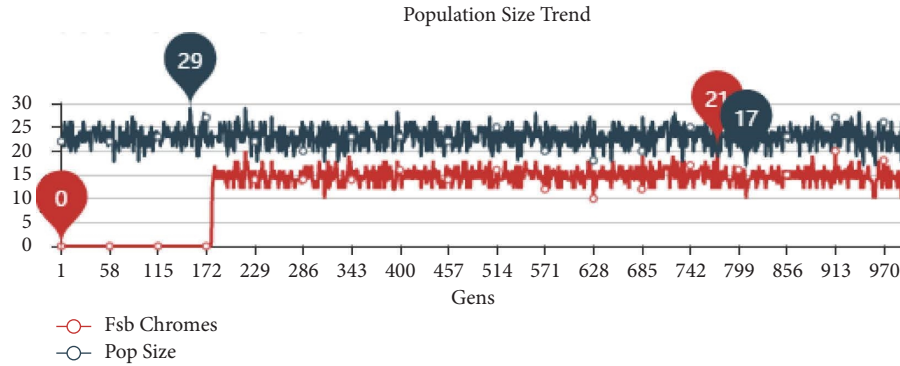
Population Size Trend

FIGURE 7: Population size trend.

as a penalty in the MOGA, the three objective function values are meaningless when the chromosomes in the population are not feasible. So, the curves in the figure begin where feasible chromosomes appear, i.e., at the 177th generation.

The red curve in Figure 8 represents the objective function values of the feasible chromosomes in each generation and the dark blue curve represents the Pareto optimal value of that generation. From its evolutionary trend, it can be seen that the feasible objective function value is highly volatile, while the Pareto optimal value is relatively stable. In the process of genetic evolution, the Pareto optimal value fluctuated in the 186th, 356th, 781th, 821th, and 835th generations, respectively. The fluctuation tempo of the three objective functions is exactly the same but the trends are different. Among them, objective functions 2 and 3 are related to the interests of operating companies and their fluctuation trends are, basically, the same, while objective function 1 represents the generalized travel cost of passengers, which is contrary to the trend of objective function 2, i.e., when the Pareto optimal value of objective function 1 rises, the objective function 2 drops. When iterating to the 835th generation, the three objective functions undergo their last fluctuations and the Pareto optimal values are stabilized at 2,703,433 people·¥, 6254 train·km, and 1083 km, respectively. From the evolutionary trend and contradictory characteristics of the three objective functions, it can be concluded that the Pareto optimal solution of the multiobjective optimization problem of URT is a result of mutual compromise and trade-off, which not only considers the generalized travel cost of passengers, but also takes into account the efficiency, cost, and organizational difficulty of the transportation enterprises.

According to the algorithm description of the TOP optimization model in the above section, Pareto optimal values of the previous generation are compared with the function values of the current generation after each evolution of the algorithm, so as to realize the dynamic update of the Pareto solution set. In the algorithm parameter setting of this case, the size of the parent population is 10 chromosomes. However, after the evolution of 177 generations, the number of feasible chromosomes is at least 10, which can be seen in Figure 7. Thus, for feasible chromosomes which exceed the population size, those for the next generation, randomly selected by the roulette wheel selection method, have relatively high fitness values. The chromosomes that are not selected for subsequent evolution are not necessarily inferior and there must be some chromosomes whose vectors are more prominent in many dimensions, i.e., there is a set where each solution is not worse than any other solution. In other words, each solution in this set cannot dominate or affect others, namely, Pareto optimal solutions. Although it cannot be said that they are superior to any other solutions, there are no better ones either.

These feasible solutions are preserved in the whole evolution process. The range of feasible solutions of the entire model can be obtained, where the boundary value is the frontier of the Pareto optimal value, as shown in the 3D grid in Figure 9, in which the diamond represents the Pareto optimal value of the entire evolution process and the circle represents the feasible value. The Pareto optimal value set forms a Pareto frontier at the boundary of all feasible solutions.

The Pareto optimal solution of the multiobjective optimization model of TOP, which is a set of multiple nondominated solutions, is obtained by balancing the optimal solutions of the three objective functions. In order to optimize the objective function Z1, the highest transport capacity should be allocated to minimize the generalized travel cost of the passenger, i.e., the higher the service frequency, the more it will help reduce the passengers' waiting and congestion costs. In order to optimize the objective function Z2, it is necessary to allocate the minimum number of trains to meet the passenger demands and reduce the travel mileage of trains, thereby reducing the operating costs of enterprises. In order to optimize the objective function Z3, the number of trains on the same routing in the adjacent passenger flow periods should fluctuate as little as possible to ensure the continuity and feasibility of the TOP. Taking the minimum of each objective function as the boundary, four Pareto optimal schemes and their compromised rate of each objective function value, relative to the boundary, are listed in Table 4.

The transport capacity configurations corresponding to the four schemes above are shown in Figure 10. It can be seen that the number of trains in each scheme is the same during off-peak hours and the differences mainly arise in peak periods. For schemes with less compromise in the

Obj Func Z1 Trend

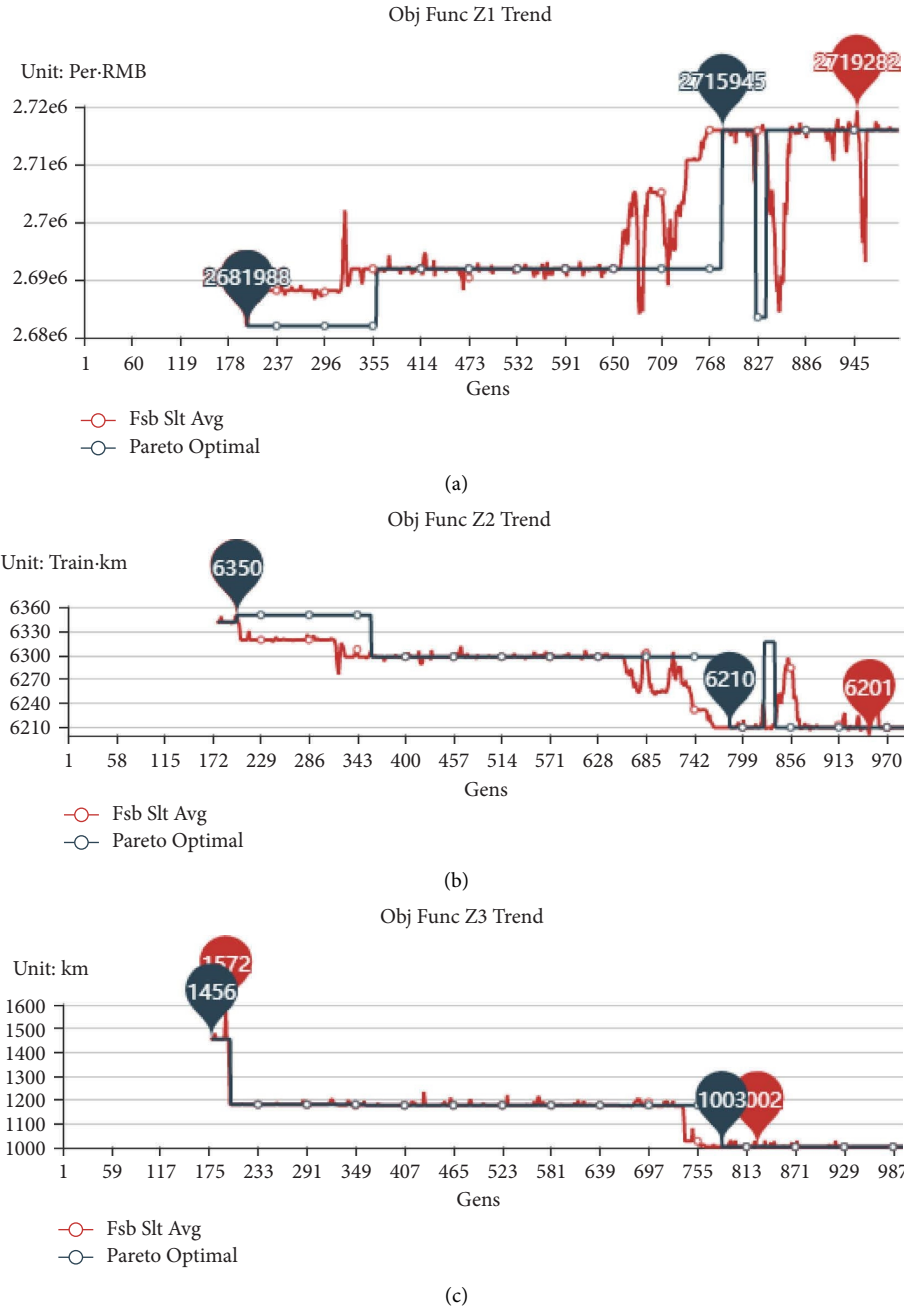(a)

Obj Func Z2 Trend

(b)

Obj Func Z3 Trend

(c)

FIGURE 8: Trend of 3 objective functions: (a) Trend of objective function Z1, (b) trend of objective function Z2, and (c) trend of objective function Z3.

objective function Z1, such as Pareto scheme 2 and 3, more trains are allocated during peak hours; for schemes with less compromise in the objective function Z2, such as Pareto scheme 1 and 4, less trains are allocated in peak hours; and for schemes with less compromise in the objective function Z3, such as Pareto scheme 3 and 4, the number of trains on each routing is more evenly distributed.

For the multiple Pareto optimal schemes of TOP, this article takes Pareto scheme 1 as an example for detailed analysis, which is represented by a triangle in Figure 9 and located in the lower middle part of the Pareto front. Due to

the long travel distance of the entry-exit depot of the short-turning rolling stock, in order to make full use of them, the number of short-turning trains in morning and evening peak hours are maintained as consistently as possible, so that the number of the entry-exit depot rolling stock is reduced during the transition of PFTPs. Therefore, one pair of short-turning trains run in the morning peak hours of 07:00–07:30 and 09:00–09:30 and two pairs run from 07:30 to09:00. During the evening peak hours, one pair of trains run from 16:30 to17:00 and 18:30 to19:00 and three pairs run from 17:00 to 18:00, as shown in Table 5. Five pairs of trains are
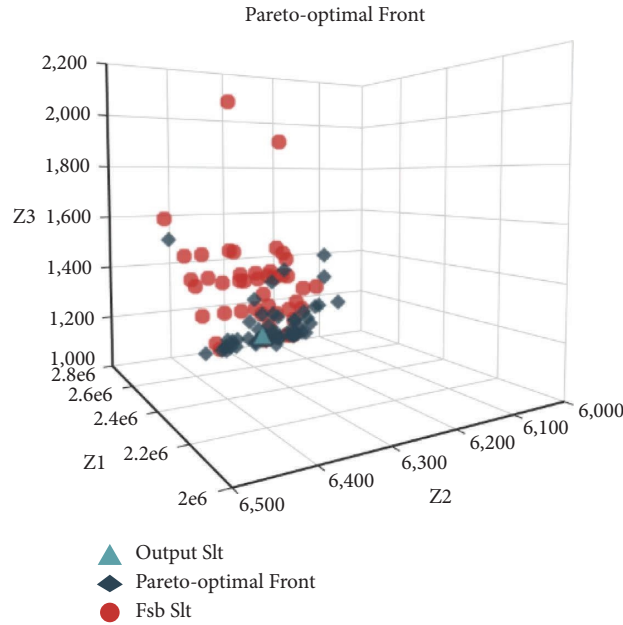
Figure 9: Pareto optimal front.

Table 4: Comparison of Pareto optimal schemes for TOP.

Pareto Opt. Schemes	Obj. Func. Z1 (Per·¥)	Obj. Func. Z2 (Train·km)	Obj. Func. Z3 (km)
Min Z1 scheme	**2,659,422**	6430	1521
Min Z2 scheme	2,771,763	**6069**	1168
Min Z3 scheme	2,715,345	6223	**998**
Pareto scheme 1	2,703,433 (1.65%)	6254 (3.05%)	1083 (8.52%)
Pareto scheme 2	2,667,943 (0.32%)	6459 (6.43%)	1046 (4.81%)
Pareto scheme 3	2,674,193 (0.56%)	6385 (5.21%)	1030 (3.21%)
Pareto scheme 4	2,714,684 (2.08%)	6279 (3.46%)	1038 (4.01%)

These bold values represent the minimum values of each objective function and are used for Pareto optimization comparison.

operated during the off-peak PFTPs, mainly considering the constraint of the maximum departure interval of 360 s. In the morning peak period, the number of trains on the basic routing from 07:30 to 08:00 is the largest with eight pairs, plus two pairs of short-turning ones, so the departure interval of overlapping sections between long and short-turning routing is 180 s, which is still above the lower limit of headway and conforms to the constraints.

To verify whether the calculated transport capacity configuration meets the passenger demands, it should be compared with the actual passenger flow curve, as shown in Figure 11. The figure shows the periods of the morning and evening super peaks, as well as the transport capacity calculated by the 100% fully loaded rate. The light blue curve is the maximum cross-sectional flow for each PFTP. The transport capacity of most PFTPs is higher than the passenger flow curve during the entire day but, for some periods, the maximum cross-section passenger flow exceeds the transport capacity at a fully loaded rate of 100%. The orange curve shows the actual fully loaded rates during each PFTP. It can be seen that, during the PFTP from 18:00 to 18:30, the maximum fully loaded rate of a train reached 111%,

which is still less than the upper limit of $\varphi_{max}^{l} = 120\%$, controlled by the operating department. Those in the rest of the PFTPs are all less than that limit. Therefore, it can be considered that the TOP not only meets the needs of actual operational safety, it also ensures the passengers' riding comfort to a greater extent, which improves the level of URT passenger service.

In 2019, Guangzhou Metro Line 2 adopted the routing plan of the long and short turning mode for the morning peak hour only. TOP indicators of the scheme, generated by using the method described in this article, are compared with those of the actual plan in Table 5. It can be seen that the number of trains in the optimized plan is reduced by six pairs, compared to the actual one. The loaded running mileage is reduced by 520 train·km, which is 4.01% of the original, due to the implementation of the long and short turning mode in the evening peak. At the same time, the average fully loaded rate of trains during the whole day increased from 35.7% to 37.2%, which improved the efficiency of train utilization. By making full use of the train capacity, the fully loaded rates of cross sections with maximum ridership, in both the up and down directions,

(a)

(b)

(c)

(d)

FIGURE 10: Capacity allocation of Pareto optimal schemes. (a) Pareto scheme 1, (b) Pareto scheme 2, (c) Pareto scheme 3, and (d) Pareto scheme 4.

increased by 10.2% and 12.2%, respectively. Because Line 2 has been operating since 2002, the organizational experience of operation and passenger dependency tend to be mature and there is not much room for adjustment in operation organization. Therefore, the model and algorithm proposed in this article can improve the current TOP, which still has a certain reference value and theoretical significance.

The stop time of trains optimized by the algorithm, according to the upstream and downstream passenger flows, changes dynamically with PFTP. The results of the

TABLE 5: Trains of long and short turning routing.

| PFTPs	Long-turning routing	Short-turning routing
05:30–06:00	5	0
06:00–06:30	5	0
06:30–07:00	5	0
07:00–07:30	7	1
07:30–08:00	9	2
08:00–08:30	8	2
08:30–09:00	5	2
09:00–09:30	5	1
09:30–10:00	5	0
10:00–10:30	5	0
10:30–11:00	5	0
11:00–11:30	5	0
11:30–12:00	5	0
12:00–12:30	5	0
12:30–13:00	5	0
13:00–13:30	5	0
13:30–14:00	5	0
14:00–14:30	5	0
14:30–15:00	5	0
15:00–15:30	5	0
15:30–16:00	5	0
16:00–16:30	5	0
16:30–17:00	5	1
17:00–17:30	5	3
17:30–18:00	5	3
18:00–18:30	7	1
18:30–19:00	5	1
19:00–19:30	5	0
19:30–20:00	5	0
20:00–20:30	5	0
20:30–21:00	5	0
21:00–21:30	5	0
21:30–22:00	5	0
22:00–22:30	5	0
22:30–23:00	5	0
23:00–23:30	5	0
23:30–00:00	5	0
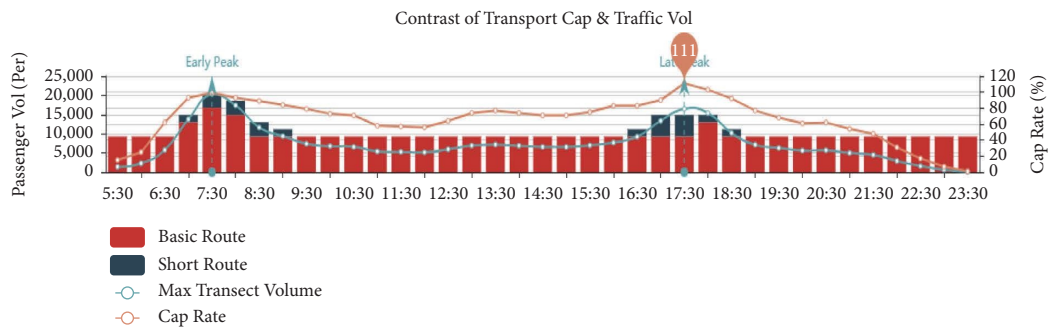
FIGURE 11: Contrast of transport capacity and traffic volume.

optimized stop time at all stations during the period 07:00–08:30, in the morning peak, and 17:00–18:30 in the evening peak, are compared with the original fixed stop time presented in Table 6 The optimized stop time values of each station in each PFTP dynamically change according to the boarding and alighting passenger volume, which fluctuates around the original stop time. Among these, the stop times between the section from station 8–station 17 are longer and the longest time (68 s) occurs in the down direction of station 13, during PFTP 17:00–17:29, which is consistent with the peak passenger flow distribution in Figure 6.

TABLE 6: Comparison of train stop times.

Station No./PFTP	Optimized stop time (s)						Original stop times (s) full day
	07:00–07:29	07:30–07:59	08:00–08:29	17:00–17:29	17:30–17:59	18:00–18:29	
1	50/—	50/—	50/—	50/—	50/—	50/—	50/—
2	30/32	32/31	33/30	30/33	30/32	31/31	34/34
3	30/32	31/35	32/33	31/32	32/31	33/30	34/34
4	38/32	41/33	42/32	31/38	32/39	32/37	34/34
5	38/31	42/32	43/32	31/39	31/39	32/38	34/34
6	43/39	47/39	46/36	41/46	39/45	39/41	34/34
7	40/34	41/35	41/34	35/43	35/42	35/39	33/33
8	37/31	39/31	38/32	34/37	33/37	32/35	37/37
9	53/46	56/45	56/44	51/61	49/58	47/52	50/50
10	37/34	37/34	37/34	38/42	37/41	36/38	35/35
11	37/33	37/33	35/33	35/38	34/38	33/36	35/35
12	47/39	46/40	45/41	47/50	44/48	41/44	35/35
13	57/50	61/52	59/53	64/68	59/64	53/58	55/55
14	35/34	37/35	36/35	37/38	35/36	32/33	35/35
15	34/33	34/34	34/34	35/35	34/35	31/33	35/35
16	44/50	43/56	42/56	65/53	62/50	55/48	50/50
17	33/41	34/47	33/47	44/39	44/38	42/37	37/37
18	32/31	33/34	32/35	37/33	37/33	34/33	35/35
19	33/32	35/36	33/39	39/34	39/34	36/34	34/34
20	31/33	32/35	31/36	35/31	36/31	33/32	34/34
21	29/30	29/31	29/31	31/29	30/29	30/29	34/34
22	35/37	37/43	33/48	43/33	45/33	41/35	34/34
23	34/33	33/36	31/38	38/32	37/32	34/33	34/34
24	—/50	—/50	—/50	—/50	—/50	—/50	—/50

The numerator and denominator represent the up and down stop time, respectively. The stop time of the terminal station is indicated by "—".

TABLE 7: Comparison of indicators.

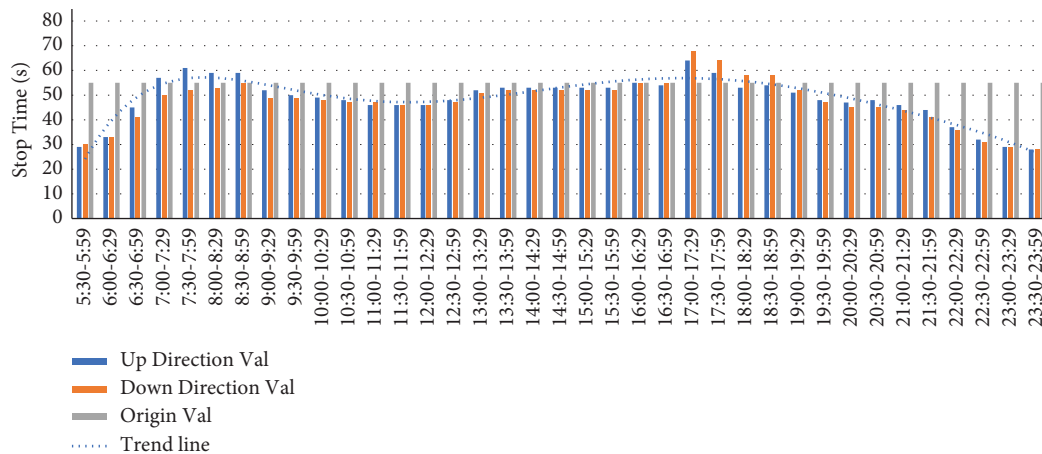| Schemes	Train pairs	Loaded mileage (trainkm)	Train capacity (Person)	Avg. FLR (%)	Max cross-section FLR (Up) (%)	Max cross-section FLR (down) (%)
Actual	219	12,966	814,680	35.70	99.50	99.10
Optimized	213	12,446	788,640	37.20	109.70	111.30

FLR is short for fully loaded rate.

FIGURE 12: The stop time at station 13.

It should be noted that the fixed stop time of station 13 in Table 7 is the longest, which is 55 seconds. Taking the stop time of that station as an example, the dynamic stop times in both directions throughout the day are compared with the original fixed ones, as shown in Figure 12. The trend line indicates that the length of the optimized stop time is basically consistent with the tendency of passenger flow in the up and down directions during each PFTP. In the four PFTPs of the morning peak, the stop time for the up direction is higher than that of the original value and the down direction; in the four PFTPs of the evening peak, the stop time for the down direction is higher than that of the original value and the up direction. The summation of optimized stop time in all PFTPs is only 87.3% of the original totality. It can be seen that the dynamic setting of the stopping time can effectively compress the whole travel duration of the train and reduce the unnecessary waiting time of passengers, which is of practical importance for improving the passenger service level.

8. Conclusion

The formulation of the TOP needs to consider the balance of the relationship between the supply of the transportation department and passenger demand. In this article, a multi-objective collaborative optimization model for train operation planning and dynamic train stopping time is constructed to achieve the goals of minimizing the generalized travel costs of passengers, the travel distance of trains, and the fluctuations of the rolling stock exit/entry-depot in the adjacent passenger flow periods. Based on the framework of MOGA, combined with the characteristics of the model, the coupling relationship between the numbers of trains between routings is described in the form of a genome and the genetic operation of changing the genome at a definite position and in a certain direction is implemented by using directed mutation. Finally, we get the Pareto optimal solution of the TOP and the dynamic train stop time to meet the time-varying passenger demand. Analysis of the multirouting mode of Guangzhou Metro Line 2, as an example, shows the correctness of the model and the effectiveness of the algorithm, as follows:

(1) The Pareto optimal solution set is generated and the multiple solutions constitute the Pareto frontier. Each nondominating solution has superiority in each dimension of objective function, which is reflected in the differences of the totality of trains, the number of short-turning trains and the balance of allocation of trains between PFTPs.

(2) By comparing the optimized TOP with the actual one, the former plans to operate short-turning trains during the evening peak, which makes it possible to use less trains to complete the transportation tasks of the existing plan, thereby reducing operating costs, improving the fully loaded rate of trains within a reasonable range and alleviating the strain of train turnover during peak hours.

(3) The number of trains on long and short-turning routings tends to be balanced in multiple consecutive PFTPs, which facilitates the formulation of train turnover plans and operations of the rolling stock exit/entry-depot, ensuring the operability of TOP.

(4) According to the spatiotemporal distribution characteristics of passenger flow, the optimized dynamic train stopping time can better meet passenger demands and the safety requirements of passenger organization in stations, than the fixed one, which is reduced by 12.7%. Therefore, it is conducive to shorten travel time and improve the transport capacity.

The model and algorithm proposed in this article can quickly generate TOPs with high service levels that meet the needs of passenger flow under certain routing plans, provide effective decision support for transportation departments, and have broad applicability to the development of URT TOPs. The next step of the research is to realize the integrated optimization of the TOP, transport mode, and train stopping plan.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

References

[1] X. Gong and C. Liu, "Research on traffic mode division based on Markov model - a case study of bus and rail transit in Beijing," *Journal of Qingdao University of Technology*, vol. 40, no. 3, pp. 107–111, 2019.

[2] H. Sun, B. Si, and J. Wu, "Combined model for flow assignment and mode split in two-modes traffic network," *Journal of Transportation Systems Engineering and Information Technology*, vol. 8, no. 4, pp. 77–82, 2008.

[3] K. Ghoseiri, F. Szidarovszky, and M. J. Asgharpour, "A multi-objective train scheduling model and solution," *Transportation Research Part B: Methodological*, vol. 38, no. 10, pp. 927–952, 2004.

[4] M. Freyss, R. Giesen, and J. C. Muñoz, "Continuous approximation for skip-stop operation in rail transit," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 419–433, 2013.

[5] L. Deng, J. Zeng, and H. Mei, "Passenger flow pushing assignment method for an urban rail network based on hierarchical path and line decomposition," *Sustainability*, vol. 11, no. 22, p. 6441, 2019.

[6] M. R. Bussieck, P. Kreuzer, and U. T. Zimmermann, "Optimal lines for railway systems," *European Journal of Operational Research*, vol. 96, no. 1, pp. 54–63, 1997.

[7] A. Higgins, E. Kozan, and L. Ferreira, "Optimal scheduling of trains on a single line track," *Transportation Research Part B: Methodological*, vol. 30, no. 2, pp. 147–161, 1996.

[8] Y. Wang, T. Tang, B. Ning, and L. Meng, "Integrated optimization of regular train schedule and train circulation plan for urban rail transit lines," *Transportation Research Part E: Logistics and Transportation Review*, vol. 105, pp. 83–104, 2017.

[9] D. Canca, E. Barrena, A. De-Los-Santos, and J. L. Andrade-Pineda, "Setting lines frequency and capacity in dense railway rapid transit networks with simultaneous passenger assignment," *Transportation Research Part B: Methodological*, vol. 93, pp. 251–267, 2016.

[10] L. Deng, Q. Zeng, W. Gao, and S. Bin, "Optimization of train plan for urban rail transit in the multi-routing mode," *Journal of modern transportation*, vol. 19, no. 4, pp. 233–239, 2011.

[11] M. Nikolić and D. Teodorović, "A simultaneous transit network design and frequency setting: computing with bees," *Expert Systems with Applications*, vol. 41, no. 16, pp. 7200–7209, 2014.

[12] Y. H. Chang, C. H. Yeh, and C. C. Shen, "A multiobjective model for passenger train services planning: application to Taiwan's high-speed rail line," *Transportation Research Part B: Methodological*, vol. 34, no. 2, pp. 91–106, 2000.

[13] A. Ceder, "Designing transit short-turn trips with the elimination of imbalanced loads," *Computer-Aided Transit Scheduling*, vol. 308, pp. 288–303, 1988.

[14] D. Canca, E. Barrena, G. Laporte, and F. A. Ortega, "A short-turning policy for the management of demand disruptions in rapid transit systems," *Annals of Operations Research*, vol. 246, no. 1-2, pp. 145–166, 2016.

[15] Y. J. Lee, S. Shariat, and K. Choi, "Optimizing skip-stop rail transit stopping strategy using a genetic algorithm," *Journal of Public Transportation*, vol. 17, no. 2, pp. 135–164, 2014.

[16] Y. Zhu and R. M. P. Goverde, "Railway timetable rescheduling with flexible stopping and flexible short-turning during disruptions," *Transportation Research Part B: Methodological*, vol. 123, pp. 149–181, 2019.

[17] X. Jiang and X. Guo, "Integrated operation of trunk routes and branches of rural transit," *Procedia - Social and Behavioral Sciences*, vol. 138, pp. 501–509, 2014.

[18] X. Zhao, Q. Sun, Y. Zhu, Y. Ding, C. Ma, and Z. Chen, "Multi-routing planning design of Y-type urban rail transit," *Advances in Mechanical Engineering*, vol. 8, no. 8, pp. 168781401666738–12, 2016.

[19] T. Lin and N. H. M. Wilson, "Dwell time relationships for light rail systems," *Transportation Research Record*, vol. 1361, pp. 287–295, 1992.

[20] H. Z. Aashtiani and H. Iravani, "Application of dwell time functions in transit assignment model," *Transportation Research Record*, vol. 1817, no. 1, pp. 88–92, 2002.

[21] X. Karekla and N. Tyler, "Reduced dwell times resulting from train-platform improvements: the costs and benefits of improving passenger accessibility to metro trains," *Transportation Planning and Technology*, vol. 35, no. 5, pp. 525–543, 2012.

[22] L. D'Acierno, M. Botte, A. Placido, C. Caropreso, and B. Montella, "Methodology for determining dwell times consistent with passenger flows in the case of metro services," *Urban Rail Transit*, vol. 3, no. 2, pp. 73–89, 2017.

[23] C. Liebchen and R. H. Möhring, "A case study in periodic timetabling," *Electronic Notes in Theoretical Computer Science*, vol. 66, no. 6, pp. 18–31, 2002.

[24] Y. Li, W. Xu, and S. He, "Expected value model for optimizing the multiple bus headways," *Applied Mathematics and Computation*, vol. 219, no. 11, pp. 5849–5861, 2013.

[25] Z. Xueyu and Y. Jiaqi, "Research on the Bi-level programming model for ticket fare pricing of urban rail transit based on particle swarm optimization algorithm," *Procedia - Social and Behavioral Sciences*, vol. 96, pp. 633–642, 2013.

[26] Transportation Research Board Business Office, *Highway Capacity Manual 2008*, National Research Council, Washington, DC, USA, 2009.

[27] Z. C. Li, W. H. K. Lam, S. C. Wong, and A. Sumalee, "Design of a rail transit line for profit maximization in a linear transportation corridor," *Transportation Research Part E: Logistics and Transportation Review*, vol. 48, no. 1, pp. 50–70, 2012.

[28] Q. Wang and L. Deng, "Integrated optimization method of operational subsidy with fare for urban rail transit," *Computers and Industrial Engineering*, vol. 127, pp. 1153–1163, 2019.

[29] Y. Yue, J. Han, S. Wang, and X. Liu, "Integrated train timetabling and rolling stock scheduling model based on time-dependent demand for urban rail transit," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 10, pp. 856–873, 2017.

[30] Á. Marín, L. Cadarso, and Á. Marín, "Integration of timetable planning and rolling stock in rapid transit networks," *Annals of Operations Research*, vol. 199, no. 1, pp. 113–135, 2012.

[31] D. Powell and M. M. Skolnick, "Using genetic algorithms in engineering design optimization with non-linear constraints," in *Proceedings of the 5th International Conference on Genetic Algorithms*, San Francisco, USA, June 1993.

[32] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: nsga-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[33] T. Murata, H. Ishibuchi, and H. Tanaka, "Multi-objective genetic algorithm and its applications to flowshop scheduling," *Computers and Industrial Engineering*, vol. 30, no. 4, pp. 957–968, 1996.

[34] P. C. Chang, J. C. Hsieh, and C. Y. Wang, "Adaptive multi-objective genetic algorithms for scheduling of drilling operation in printed circuit board industry," *Applied Soft Computing*, vol. 7, no. 3, pp. 800–806, 2007.

[35] M. Gen, W. Zhang, L. Lin, and Y. Yun, "Recent advances in hybrid evolutionary algorithms for multiobjective manufacturing scheduling," *Computers and Industrial Engineering*, vol. 112, pp. 616–633, 2017.

[36] H. Hu, X. Li, Y. Zhang, C. Shang, and S. Zhang, "Multi-objectivelocation-routing model for hazardous material logistics with traffic restriction constraint in inter-city roads," *Computers and Industrial Engineering*, vol. 128, pp. 861–876, 2019.

[37] M. Gen and R. Cheng, *Genetic Algorithms and Engineering Optimization*, John Wiley and Sons, New York, NY, USA, 2000.