

Research Article

Optimizing Train-Set Circulation Plan in High-Speed Railway Networks Using Genetic Algorithm

Yun Wang,^{1,2} Yu Zhou ,^{1,2} and Xuedong Yan ¹

¹MOT Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

²Department of Civil and Environmental Engineering, The Hong Kong University of Science & Technology, Clear Water Bay, Kowloon, Hong Kong

Correspondence should be addressed to Yu Zhou; ceyuzhou@ust.hk

Received 13 October 2018; Revised 26 December 2018; Accepted 22 January 2019; Published 14 February 2019

Academic Editor: Luca D'Acerno

Copyright © 2019 Yun Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As a sustainable transportation mode, high-speed railway (HSR) has been developing rapidly during the past decade in China. With the formation of dense HSR network, how to improve the utilization efficiency of train-sets (the carrying tools of HSR) has been a new research hotspot. Moreover, the emergence of railway transportation hubs has brought great challenges to the traditional train-sets' utilization mode. Thus, in this paper, we address the issue of train-sets' utilization problem with the consideration of railway transportation hubs, which consists of finding an optimal Train-set Circulation Plan (TCP) to complete trip tasks in a given Train Diagram (TD). An integer programming TCP model is established to optimize the train-set utilization scheme, aiming to obtain the one-to-one correspondence relationship among sets of train-sets, trip tasks, and maintenances. A genetic algorithm (GA) is designed to solve the model. A case study based on Nanjing and Shanghai HSR transportation hubs is made to demonstrate the practical significance of the proposed method. The results show that a more efficient TCP can be formulated by introducing train-sets being dispatched among different stations in the same hub.

1. Introduction

During the past decade, great developments have been achieved in the high-speed railway (HSR) in China. Being acknowledged as high efficiency, high capacity, and low energy, HSR has been one of the most influential travel modes [1, 2]. As the main carrier of HSR, train-sets play a vital role in the operation and management of HSR, and the train-sets should be utilized under complicated maintenance and inspection regulation. Due to the high cost of acquisition, train-sets are a long-term capital investment for railway operator which cannot be changed frequently [3]. Therefore, the efficiency of train-set utilization is one main objective in railway operation. When a Train Diagram (TD) is drawn up, train-sets are dispatched to complete the trip tasks in the given TD. To guide the train-sets' application, a Train-set Circulation Plan (TCP) should be formulated, which determines

the one-to-one correspondence relationships between trip tasks and train-sets as well as the connection relationships between trip tasks and maintenances. Once a TCP is drawn up, the amount of needed train-sets can be obtained. Thus, the TCP is a vital factor influencing the train-sets' utilization efficiency, which also has a great impact on operating costs of HSR. Besides, complex HSR networks have been gradually formed in China. On the one hand, a huge number of HSR stations have been, which means that the distances between stations have been shortened. On the other hand, many cities, especially metropolis, have built or planned to build more than one HSR stations. Generally, cities with several stations can be treated as railway transportation hubs, and these stations are usually connected by connection rails. With the emergence of railway transportation hubs, the traditional train-sets' utilization mode (i.e., for a set of trip tasks undertaking by a train-set, the departure station of

the latter trip should be the same with the arriving station of former trip) will greatly reduce the train-sets' utilization efficiency, which has been detail demonstrated in Section 3.1.

Thus, this paper addresses the issues of the TCP problem with the consideration of railway transportation hubs and maintenance requirements, which is one of the most significant aspects in the railway operation and management studies. Roughly speaking, it consists of finding the optimal assignment of train-sets so as to complete a set of trip tasks in the given TD with higher utilization efficiency, i.e., with less number of train-sets, and considering the maintenance requirements. The contribution of this study lies in the following aspects.

Firstly, an innovative train-set utilization mode is put forward to improve the train-sets' utilization efficiency. In previous studies, train-sets can only undertake two adjacent trips when the departure station of the latter trip is the same with the arriving station of the former trip. In this case, when the time interval between the connected two trip tasks is long, train-sets must wait at the station for undertaking the next trip. Obviously, it will cause huge waste of the train-set capacity. In this paper, we propose an innovative train-set utilization mode that train-sets can be dispatched among different stations in the same railway transportation hub. Thus, a train-set can be dispatched to other stations to undertake trip tasks, instead of waiting a long time at its arriving station for the next trip task. Clearly, by relieving the constraint that the departure station of latter trip must be the same with the arriving station of the former trip, the flexibility and the efficiency of train-set utilization can be significantly enhanced. This utilization mode turns out to be more suitable to be applied in the dense HSR network.

Secondly, an integer programming TCP model is proposed to optimize the train-set utilization scheme with the consideration of maintenance requirements. Accumulated variables have been introduced to represent the running distance and running time of train-sets. The model aims to obtain an optimal TCP that determines the one-to-one correspondence relationship among sets of train-sets, trip tasks, and maintenances. Thus, the dispatcher can determine each trip should be undertaken by which train-set, the sequence of the trips undertaking by the same train-set, and when and where each train-set should be maintained. The objective of the model is set to simultaneously minimize the number of using train-sets and the total maintenances costs.

Thirdly, a genetic algorithm (GA) is designed to solve the TCP optimization model. We use an innovative representation method to formulate the relationship between maintenance arcs and normal connection arcs. Furthermore, effective crossover and mutation processes have been designed. The TCP problem has been proved to be a traditional NP-hard problem, which cannot be solved efficiently or directly by ready-made software, especially in large-scale cases. Thus, in the process of model solution, the GA is applied to search for a near-optimal TCP. The results show that the proposed GA for TCP problem has obvious advantages over ant colony algorithm (ACA) and simulated annealing (SA) in the solution quality, and good performance can be found in both computational efficiency and stability.

Finally, a case study based on Nanjing and Shanghai HSR transportation hubs is carried out to demonstrate the practical significance of the proposed method. The results show that, by introducing train-sets being dispatched among different stations in the same hub, a more efficient TCP can be formulated compared with traditional utilization mode. To complete the same trip tasks, both the number of needing train-sets and the maintenance times can be reduced. Thus, such a mode is a feasible method to utilize train-sets with lower costs and high efficiency.

The remainder of this paper is organized as follows. Section 2 reviews relevant studies in the literature. Section 3 develops a modeling framework for obtaining an optimal TCP, including problem description, basic assumption, and model formulation. Section 4 introduces the solution approach. In Section 5, a case study based on Nanjing and Shanghai HSR transportation hubs is performed to illustrate the model application. Finally, Section 6 provides conclusions.

2. Literature Review

The rolling stock planning (RSP) problem is one of the most significant aspects in the application and assignment problem of transportation vehicles. How to improve the utilization efficiency of transportation vehicles has always been the research hotspot in transportation management, including not only RSP in railway management, but also Aircraft Routing Problem (ARP) in the airline operation, Vehicle Routing Problem (VRP) in the logistics filed, and so on. Specifically, extensive research on RSP problem has been carried out worldwide, leading to the development of various rolling stock utilization models and techniques. Interested readers can refer to Cacchiani et al. [4], who reviewed the main concepts and methods about railway (re)scheduling problem, especially about RSP problem, and summarized different models and algorithms for solving this problem.

Two main categories of RSP problem have been studied during different development stages of passenger railways. Before the emergence of HSR, the transportation vehicles used in the railways are carriages and locomotives. The former is a kind of vehicle without traction engines, which can be coupled individually and independently to a convoy. The latter is a necessary part of trains which can provide traction power and pull the convoy [3, 5]. Thus, most of previous research focused on the utilization planning of both locomotives and carriages. Ziarati et al. [6] solved the locomotive problem by using the method of modern branch-and-cut integer programming algorithms, and computational experiments were conducted using the actual data from the Canadian National Railway Company. Cordeau et al. formulated an integer programming model presenting the simultaneous assignment problem of locomotives and carriages and solved it by Benders decomposition [7]. An extend model was proposed by Cordeau et al. [8] with the consideration of real-life conditions, such as maintenance, and a heuristic branch-and-bound approach based on column generation was designed to solve the mode. Ahuja et al. [9] proposed a mixed integer programming model to guide the locomotive utilization and

solved the model by CPLEX. But, the real-time changes were not taken into consideration and the execution time was long exceeding 10 hours. Vaidyanathan, Ahuja, and Orlin [10] modified the model considering refueling and maintenance and improved the solving computational efficiency by using an aggregation and disaggregation technique. Most of above studies defined the locomotive utilization problem as a large-scale integer programming problem and model the problem based on a multicommodity flow formulation [6].

Then, with the rapid development of HSR, matched railway vehicles named train-sets have been introduced. Different from above traditional railway vehicles, the train-set is a kind of self-contained trains with an engine and passenger seats, which means self-propelled train units that are not required to be pulled by a locomotive. These units consist of a fixed number of carriages and have their own traction engines [5]. As a new and important branch of RSP problem, the topic of TCP problem has been receiving increasing attention. There are a few references considering the utilization of train-sets. Schrijver [11] firstly paid attention to the TCP problem. He proposed a mathematical model for a single-day workload based on integer linear programming (ILP), aiming to minimize the number of train-sets used, and solved the model by the software CPLEX. Different utilization modes of train-sets were discussed by Zhao et al. [12] and they found that the utilization efficiency of train-sets could be significantly improved when train-sets operated in unrestrained sections. Abbink et al. [13] explored the TCP problem in the peak period. He presented an ILP formulation with the objective of minimizing the number of shortage seats during the rush hours, and using CPLEX to obtain the solution. Yang et al. [14] built a train-set connection network model aiming to maximize the train-set utilization efficiency, and designed a genetic algorithm to solve the model. Ben-Khedher et al. [15] formulated a mathematical model considering a unique type of train-sets with the objective of maximizing the company's profit and solved it by techniques of stochastic optimization, branch-and-bound and column generation. Alfieri et al. [16] addressed the TCP problem on a single train line and a single day for the case of multiple train-set types. They proposed an ILP model aiming to minimize the number of units and the carriage-kilometers, and solved the problem by decomposing it into subproblems. Fiore et al. [17] put forward a mixed integer programming model based on widely absorbing previous research achievements, which can be seen as an extended version of the model proposed by Schrijver [11]. Several methods were applied to improve the continuous relaxation of the model, and the improved branch-and-bound algorithm was designed to obtain the optimization solution. Maróti and Kroon [18] proposed an integer programming model with the consideration of maintenance constraint, and they designed a heuristic algorithm to solve the model and obtain a feasible solution. Wang et al. [19] used the big M method to formulate the accumulated train-sets' running distances and running time, which is the most intractable part in the train-set utilization problem. To improve the model's solvability,

a strategy is proposed to reduce the scale of the connection network.

To ensure the operation safety of train-sets, maintenances should be carried out as long as a train-set has been utilized for certain time periods or accumulated running distances. Extensive research has focused on the rolling stock problem considering maintenance constraints. Maróti G and Kroon L are the pioneers in exploring the maintenance routing problem in railway management. A transition model was put forward [20] in order to formulate the maintenance constraint. Following the transition model, an improved [18] interchange model was proposed in 2007, and the numerical experiments proved that this approach is efficient when dealing with small-size and medium size problem. Hong et al. [21] presented a two-phased train-set routing algorithm to cover a weekly train timetable with minimal number train-sets. The maintenance requirements were firstly relaxed to obtain minimum cost routes by solving the polynomial relaxation, then, maintenance-feasible routes were generated from the crossovers of the minimum cost routes. Nishi et al. [22] formulated the rolling stock planning problem based on the set-partition problem with regular inspection constraints. Due to the characteristic of the formulation, a column generation based approach was put forward, and a number of experiments was carried out to verified the efficiency of this approach. It could be say that maintenance constraints are the vital requirements in rolling stock routing problem; a good circulation plan should lower the maintenance costs of the rolling stocks.

Another important issue which should be addressed in the train-set utilization procedure is how to estimate the time for two consecutive trips in terminal stations. In the operational stage, possible delays or fluctuations, which can occur during operations, will influence the connection between two trips. So compensating possible delays with the planned timetable is of great importance. Many scholars put forward approaches to estimate time rates involved in the preparation phase between two successive trips. D'Acerno et al. [23] proposed a method to estimate dwell time crowding level at platforms and related interaction between passengers and the rail service in terms of user behavior when a train arrives. Li et al. [24] put forward a dwell time estimation model in a generic condition and they coined a new predictor called "dwell time at the associated station" to evaluate the performance. Other factors also have great impact on the dwell time: gaps between the train and platform [25], interior layout of the carriages [26], and fare collection method [27]. Apart from dwell time, reverse time is another time rate which influences the procedure between two consecutive trips. D' Acerno et al. [28] provided an analytical approach for determining the reverse time, which exploits layover times for energy-saving purposes.

Recently, more and more scholars have shifted their attentions to the combination optimization of train-set utilization problem with other problems related to railway management. They argued that the train-set utilization problem should be simultaneously optimized with service demand prediction

problem, timetabling problem, crew scheduling problem, etc. Wang et al. [29] studied the integration of rolling stock circulation problem under the time-varying passenger demands for a rail transit line. Some practical issues were incorporated: capacity of trains, number of available rolling stocks, and the entering/exiting depot operations. Three designed solution approaches were implemented to cope with the multiobjective mixed integer nonlinear programming problem. Zhou and Teng [30] focused on simultaneous optimization of the passenger train routing and timetabling problem. The Lagrangian relaxation decomposition was applied to solve the complex model. In order to accelerate the solving process, a heuristic algorithm based framework was also introduced. The final experiments results demonstrate that the proposed approach has better performances in terms of minimizing both train travel time and computational times. Canca and Barrena [31] investigated the rolling stock plan by simultaneously considering the number and the location of depot facilities. A sequential solving approach was designed according to the problem structure. The test results showed the feasibility of proposed approach. These above-mentioned literatures revealed a new trend in the field of rolling stock utilization problem. It is really practical and useful to enhance the efficiency in railway management.

Although a comprehensive body of literature on RSP is available, there still exist limitations and gaps in the aspect of TCP problem. Firstly, most of existing researches on TCP are suitable for the train-set utilization in Europe, which is quite different from that in China. In Europe, scholars exploring the TCP problem mainly focused on train-sets' coupling and uncoupling so as to form trains to carry out trip tasks in a given TD. They sought to minimize the seat shortages as well as the empty train movements. But, in China, train-sets are fixed composition containing 8 train-sets or 16 train-sets, which are utilized as a whole. Thus, there is no need to consider the problem of train-sets' coupling and uncoupling in China. Secondly, the approach proposed in previous studies cannot be applied directly in this paper while the train-set utilization modes are quite different. Most scenarios of existing research on TCP are set under the condition that train-sets' departure station of the next trip should be the same with the arrival station of the former trip, while no studies are available for the utilization mode that train-sets can be dispatched among different stations in the same railway transportation hub. The former can obtain a feasible TCP only when the HSR network has not been formed and the distances between railway stations are a bit long. But, in China, a complex HSR network has been formed and the distribution density of railway stations is quite high, while there could be several railway stations in some metropolis. As thus, this study makes a new attempt about the issue on TCP.

3. Modeling Framework

3.1. Problem Description. Figure 1 gives the outline of TCP design process, which includes information preparation, model formulation, model solution, and results. In the first stage, some basic information should be prepared, including

the TD, the rules, and regulation of maintenance and station operation. In the second stage, a mathematical model should be developed based on integer programming (IP), mixed integer programming (MIP), travelling salesman problem (TSP), or other techniques. A specific optimization goal should be met and some constraints should be considered to find an optimal TCP. Generally, fewer train-sets, lower maintenance costs, or balance of train-set utilization are always chosen as the objective function, aiming to enhance the train-set utilization efficiency as well as reduce operational costs of HSR. As the TCP is designed to complete trip tasks in the given TD by utilizing limited train-sets, the obtained plan should satisfy some basic constraints, including diagram constraint, station constraint, passenger demand constraint, maintenance constraint, etc. In the third stage, when the TCP problem has been modeled, corresponding solution approaches need to be designed to solve the model according to the properties of the model. Based on previous research, there are mainly three kinds of solution methods, namely, solved by software directly, adopting mathematic technique, and designing heuristic algorithm. In the final stage, an optimized TCP can be obtained.

As previously mentioned, the expanded of HSR network and the construction of HSR transportation hubs have caused enormous complexity of train-set application, leading to a challenge for developing the optimal solution of algorithms. Compared with the traditional utilization mode that train-sets can only departure from stations they arrive, the train-set utilization efficiency can be considerably improved by introducing dispatching train-sets among different stations in the same transportation hub. Figure 2 illustrates a simple example. As shown, there are two transportation hubs, i.e., Transportation Hub M and N, which, respectively, include three and two HSR stations. Stations in the same hub are connected by connection railway lines, allowing train-sets to operate among these stations.

In the example, it is assumed that there are three trips needing to be undertaken, the basic information of which is as listed in Table 1. To complete the trip task j , a train-set should departure from Station D at 12:00 and will arrive at Station A at 15:00. To complete the trip task k , another train-set should departure from Station D at 15:00 and will arrive at Station B at 19:00. For different train-set utilization modes, different train-set dispatch scheme can be obtained to complete the trip task i . When train-sets cannot be dispatched between different stations in the same hub, to undertake two adjacent trips, the departure station of the latter trip should be the same with the arrival station of the former trip. Thus, for the traditional utilization mode, the train-set arriving at Station A at 15:00 needs to wait for nearly five hours at Station A to undertake its next trip task i . When introducing train-sets being dispatched between different stations in the same hub, the train-set arriving at Station B at 19:00 can be firstly dispatched to Station A and then departure from Station A at 20:00 to complete the trip task i . That is to say, the train-set arriving at Station A at 15:00 will not need to wait for nearly five hours, which releases its ability to undertake other

TABLE 1: Basic information of the sample.

	Departure Station	Arriving Station	Departure Time	Arriving Time
Trip i	Station A	Station D	20:00	23:00
Trip j	Station D	Station A	12:00	15:00
Trip k	Station D	Station B	15:00	19:00

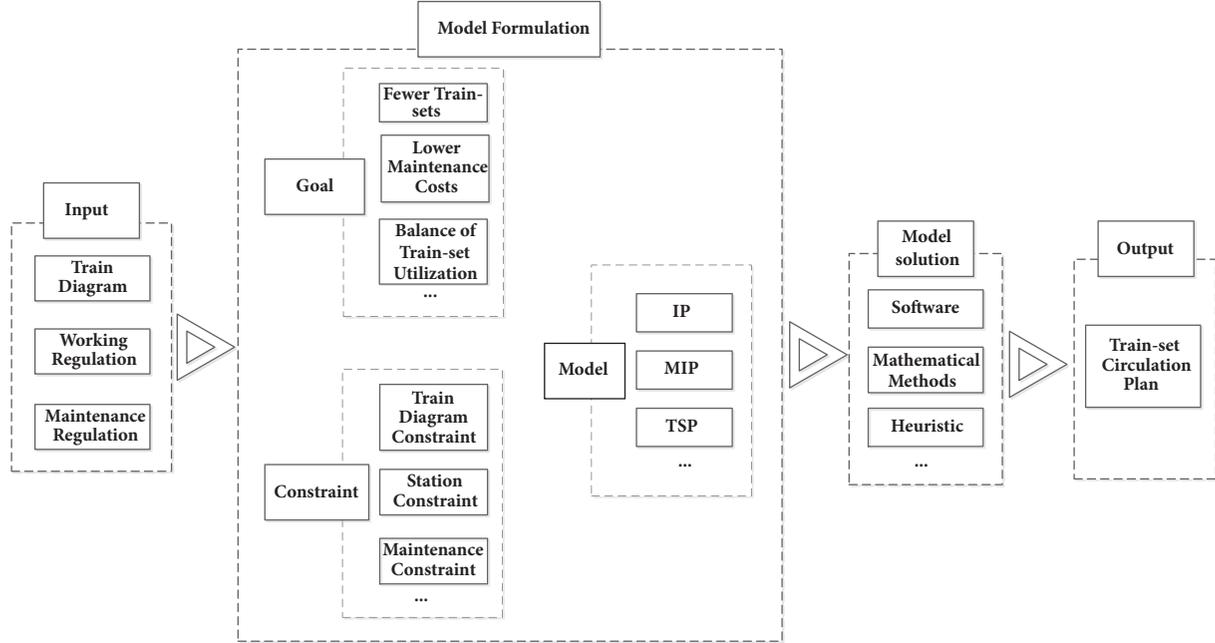


FIGURE 1: The outline of the TCP problem.

trip tasks in the hub. Obviously, the flexibility of train-set utilization has been significantly enhanced, as well as the utilization efficiency.

3.2. Basic Assumption. Several assumptions are made throughout the paper for simplicity of the model and are explained as follows:

Assumption 1. The TD of HSR are drawn up in pairs, which means that the number of arrival trains is equal to the number of departure trains in each station.

Assumption 2. We assume that only one type of train-set is taken into consideration.

Assumption 3. In practical operation, five levels of train-set maintenance standard are set to ensure the transportation safety according to train-sets' accumulated travelling time and distances. But, in this paper, only maintenance of Level One is taken into consideration. It is because of that, the TCP is usually drawn up by treating one day or several days as a cycle, which is much shorter than the travelling time standard of maintenance of Level Two (a week) to Level Five (several years). Moreover, when train-sets' accumulated travelling time and distances reach to the standard of maintenance of Level Two or above, a specific train-set maintenance plan will be formulated.

Assumption 4. We only consider a daily operation timetable; it means that means that the train timetable is the same every day.

Assumption 5. The time for passengers alighting and boarding at stations is determined.

3.3. Model Formulation. The following symbols are used in this paper:

- (i) S represents the set of railway stations, $S = \{s_i \mid i = 1, 2 \dots n\}$, where n is the total number of stations;
- (ii) $name_i$ represents the name of station i ;
- (iii) K represents the set of train-sets, $K = \{k_i \mid i = 1, 2 \dots m\}$, where m is the total number of train-sets;
- (iv) $bool_i^R$ is a binary variable identifying whether station i is near the Repair and Inspection Depot (RID) or not;
- (v) t_i^R represents the running time from station i to RID;
- (vi) t_{ik}^C represents the running time from station i to station k ;
- (vii) t_i^W represents the minimum time for preparation work in station i ;
- (viii) D_i^{INI} represents the accumulated travelling distances of train-set i ;

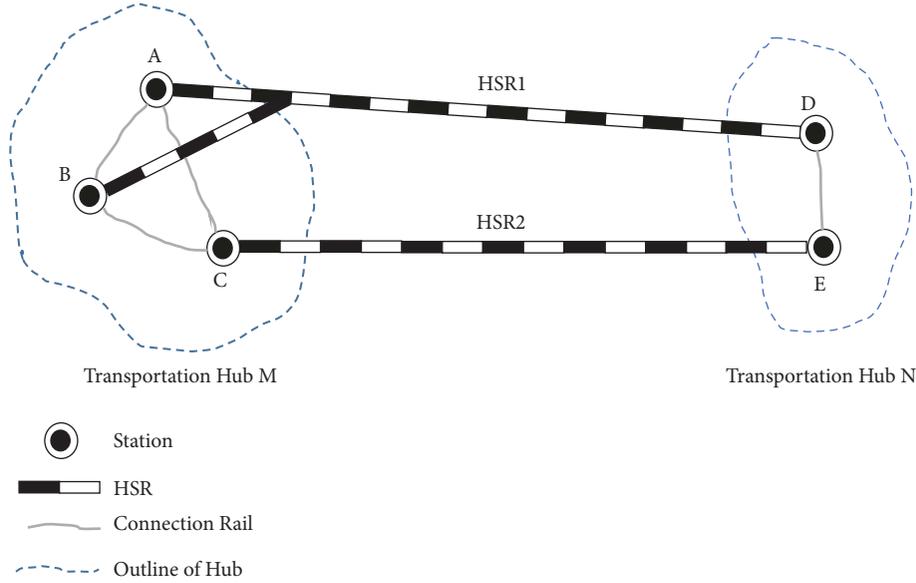


FIGURE 2: Topological structure of sample.

- (ix) T_i^{INI} represents the accumulated travelling time of train-set i ;
- (x) S_{ONE} represents the travelling distance standard of Level One maintenance;
- (xi) T_{ONE} represents the travelling time standard of Level One maintenance;
- (xii) T^M represents the time needed for maintenance;
- (xiii) Δ represents the fluctuation coefficient of maintenance standard.

3.3.1. Train-Set Utilization Network. The optimization model is based on a weighted directed graph $G(V, E)$. The node, arc, and weight are defined as follows.

Node. V denotes the set of all trips in the given TD. Let $v_i \in V$ represent a certain trip, which contains 9 attributes, including trip number n_i , departure station s_i^d , arrival station s_i^a , departure transportation hub p_i^d , arrival transportation hub p_i^a , departure time t_i^d , arrival time t_i^a , running distance D_i , and running time T_i .

Arc. E denotes the set of all arcs in the train-set utilization network, which also means the connection relationships between nodes. Specifically, let $e_{ij} \in E$ represent the arc from node v_i to node v_j . Variable c_{ij} denotes the weight of arc e_{ij} , which can be calculated by

$$c_{ij} = \begin{cases} t_i^a - t_i^d & t_i^a > t_i^d, p_i^a = p_j^d \\ 1440 + t_i^a - t_i^d & t_i^a < t_i^d, p_i^a = p_j^d \end{cases} \quad (1)$$

Decision Variable. To draw up a TCP, the planners need to solve the following three problems: the connection relationship between trips, when the train-sets should be overhauled,

and each trip should be undertaken by which train-set. Thus, three decision variables are defined here to obtain a TCP with a given TD. $x_{i,j}(i, j \in V)$ is a binary variable that represents the connection relationship between trip v_i and trip v_j . When trip v_i is connected to trip v_j , the value of $x_{i,j}$ is equal to one, otherwise zero. $y_{i,j}^k(i, j \in V)$ is a binary variable representing whether the train-set should be overhauled. Only when the train-set k is overhauled after completing the task of trip v_i and undertakes the task of trip v_j after completing maintenance, the value of $y_{i,j}^k$ is equal to one, otherwise zero. $z_i^k(i \in V, k \in K)$ is also a binary variable that represents the one-to-one corresponding relationship between train-sets and trips. When the train-set k undertakes trip v_i , the value of z_i^k is equal to one, otherwise zero.

3.3.2. Objective Function. When a TD is given, lots of different TCPs can be formulated to complete all the trip tasks. To obtain an optimal TCP, an objective needs to be proposed to improve the train-set utilization efficiency as much as possible on the premise of satisfying constraints. For a TCP, the number of needing train-sets is a crucial indicator to measure the quality of the plan. Planners always seek to complete the same number of trip tasks in the given TD with less train-sets. Moreover, the number of maintenance is also an important factor influencing the train-set utilization efficiency. When train-sets are overhauled, lots of time is spent on train-sets running between railway stations and inspection and repair depots, which will reduce the effective working time of train-sets. Thus, in this paper, both the number of using train-sets and the total number of maintenances are considered into the objective function. To deal with different situations in practical operation and expand the model's scope of application, variables σ_1 and σ_2

are introduced as weight values to the number of using train-sets and the number of maintenances, respectively. Moreover, the sum of σ_1 and σ_2 is equal to one.

$$\min \left(\sigma_1 \sum_k \left[1 - \prod_i (1 - z_i^k) \right] + \sigma_2 \sum_i \sum_j \sum_k y_{ij}^k \right) \quad (2)$$

$$k \in K, i \in V, j \in V$$

3.3.3. Constraints

Spatial Constraints. In the previous studies related to train-set utilization problem, the spatial constraints mean that the arriving station of the former trip should be the same with the departure station of the latter trip. It means that when train-sets complete a trip task, they have to stay at the station waiting for the next trip task. It is possible that train-sets need to wait for a long time. For example, in the off-peak hour, the time interval between the departure time of the latter trip and the arriving time of the former trip may be long due to the low frequency of train services. Thus, in this paper, we proposed a novel utilization mode that trips in the same transportation hub can be connected. It means that after completing a trip task, the train-set can be dispatched to undertake a new trip task departing from another station in the same hub. Obviously, there are two requirements for dispatching train-sets to another station in the same hub: firstly, there should be connection rail lines between these two stations; and secondly, the time interval between departure time of the latter trip and the arriving time of the former trip should satisfy corresponding requirements. Equation (3) means the spatial constraints, restraining that the departure transportation hub of the latter trip p_j^d should be the same with the arrival transportation hub of the former trip p_i^a .

$$p_i^a = p_j^d \quad i \in V, j \in V \quad (3)$$

Uniqueness Constraints. In this paper, the uniqueness constraints include two aspects: firstly, as shown in (4), each trip in the given TD must be undertaken by a train-set, and only one train-set can be allocated to complete the task; secondly, to avoid train-sets being idle after completing a trip task as well as formulate utilization circulation, (5) and (6) restrain that there must be one and only one trip task being assigned to train-sets; i.e., another trip must be connected with the former trip task.

$$\sum_k z_i^k = 1 \quad \forall k \in K \quad (4)$$

$$\sum_i x_{ij} = 1 \quad \forall j \in V \quad (5)$$

$$\sum_j x_{ij} = 1 \quad \forall i \in V \quad (6)$$

Maintenance Constraints. To ensure the safe operation, train-sets must be strictly complied with the maintenance rules and standards. As previously mentioned, only the maintenance

of Level One needs to be considered in this paper, which requires train-sets to be overhauled when their accumulated travelling time and distance reach up to 48 hours or 4,000 km. Additionally, in practical applications, a minor fluctuation compared with the standards are allowable. Equations (7) and (8) restrict that train-sets' accumulated travelling time, D_i^{INI} , and distances, T_i^{INI} , must be smaller than the maintenance standard in the fluctuation range, where S_{ONE} and T_{ONE} , respectively, represent travelling distance and time standards of Level One maintenance and Δ represents the fluctuation coefficient.

$$D_k^{INI} \leq S_{ONE} (1 + \Delta) \quad k \in K \quad (7)$$

$$T_k^{INI} \leq T_{ONE} (1 + \Delta) \quad k \in K \quad (8)$$

Basic Working Time Constraints. When a train-set arrives at the station after completing a trip task, some preparation works need to be carried out before departing for the next trip, such as the alighting and boarding of passengers, carriage cleaning, waste drainage, water and food supply, crew member change, shunting (if the next trip would departure from a different track), and so on [32]. As a crucial factor influencing the preparation time, the passenger flow should be considered when calculating the preparation time, especially when the flow is huge or operation fluctuation occur [33, 34]. But, due to the aim of this paper is to propose a novel utilization mode of TCP problem, we assume that the corresponding time for passenger alighting and boarding is determined and the fluctuation in the daily operation is not taken into consideration for simplification. In order to keep train-sets in good conditions and provide better services, the minimum preparation time should be guaranteed according to the official guidance. Thus, the time difference between the arrival time of the former trip and the departure time of the latter trip should be longer than the requirement of the minimal necessary working procedure duration (including the consideration time for passengers boarding and alighting), which is as shown in (9). Similarly, as shown in (10), when a maintenance needs to be conducted between two adjacent trip tasks, the time difference between the arrive time of the former trip and the departure time of the latter trip should be longer than the time duration for maintenance, including maintenance time and running time between stations and the Repair and Inspection Depot (RID).

$$x_{ij} (1 - y_{ij}^k) c_{ij} \geq x_{ij} (1 - y_{ij}^k) (t_n^W + t_{mn}^C) \quad (9)$$

$$x_{ij} y_{ij}^k c_{ij} \geq x_{ij} y_{ij}^k (2t_m^R + T^M + t_n^W) \quad (10)$$

$$i \in V, j \in V, k \in K, m = s_i^a, n = s_j^d \quad (11)$$

4. Solution Approach

The TCP problem is a traditional NP-hard problem. Due to the advantages over obtaining a feasible solution within a reasonable amount of time, intelligent algorithms are always

used to solve this problem. Genetic algorithm (GA) has been proved that it can efficiently and effectively solve a variety of real-world issues, e.g., train timetabling problems, timetable rescheduling problems, and network design optimization problems. Thus, in this paper, a GA is designed to search for an optimal TCP. In general, the procedure of GA can be described as follows.

(i) *Coding and Initialization.* For TCP problem, the results should clearly point out the utilization details of each train-set, including the trips it should undertake, the order of these undertaking trips, and the maintenance information. Thus, in this paper, the encoding style of the GA chromosome should represent the above information. As shown in Figure 3, v_i represents the trips the train-set should undertake, and the subscript i denotes the order of these trips. A binary variable is introduced to the maintenance information. When the train-sets are overhauled after completing a trip, the number 1 should be inserted before the next trip. Otherwise, the number 0 should be inserted between two trips. It means that the number 0 denotes that the connection between two adjacent trips is a normal trip connection, while the number 1 denotes the maintenance connection.

Generally, the initial population is generated randomly, allowing the entire range of possible solutions.

(ii) *Reproduction.* Reproduction is a process in which individual strings are copied according to the fitness function value, which can represent a measure of the utility or goodness related to what we want to minimize. Copying chromosomes according to the fitness function value means that chromosomes with a high value, indicating a higher probability of contribution to one or more offspring in the next generation. In this paper, the fitness function is as shown in (12), where $\sum_{i \in V} \sum_{j \in V} c_{ij} x_{ij}$ represents the total connection time. When the connection time is high, the fitness function value of the solution will be small, which means that the probability of reproducing the chromosomes is very low.

$$f = \frac{1}{\sum_{i \in V} \sum_{j \in V} c_{ij} x_{ij}} \quad (12)$$

(iii) *Crossover.* The members reproduced in the new mating pool are mated randomly and afterward each pair of chromosomes undergoes a cross-change. In order to do this, an integer position (cut point) is selected uniformly at random in the chromosome between the first gene and the last gene. Two new chromosomes are created swapping all genes (with the corresponding allelic values) between the selected position and the end of the gene.

In this paper, the crossover process can be demonstrated as Figure 4. Firstly, randomly select a trip v_k and two parents P1 and P2. Secondly, find the position of v_k in the two parents, and let v_l^1 and v_m^2 represent the positions in P1 and P2, respectively. Thirdly, calculate the connection costs between trip v_k and its next trip in P1 and P2, and, respectively, describe as $C_{l,l+1}$ and $C_{m,m+1}$. Then, add the trip with smaller connection cost into the offspring. For example, when $C_{l,l+1}$

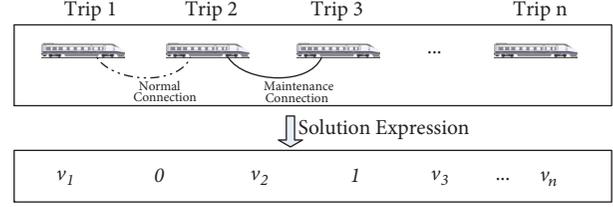


FIGURE 3: Coding and solution expression of GA chromosomes.

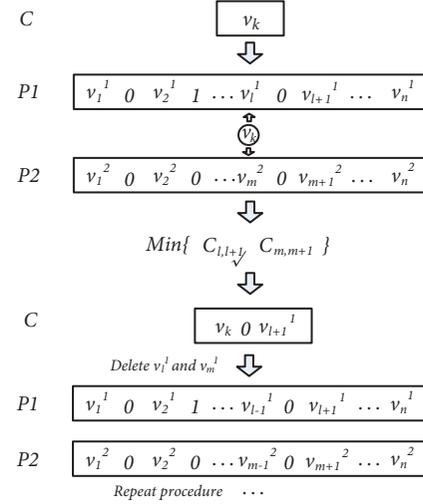


FIGURE 4: Crossover procedure.

is smaller than $C_{m,m+1}$, the trip after v_k in the offspring should be the trip in the position of v_{l+1}^1 . Finally, delete trip v_k from parents, and repeat above procedure with the new added trip in the offspring until the new offspring is fully generated.

(iv) *Mutation.* An irrecoverable loss of potentially useful information may occasionally occur in the process of reproduction and crossover. Moreover, in the process of iteration, there may be a high probability of finding a false peak. Therefore, the mutation should be conducted randomly to protect the useful information and guarantees the possibility of exploring the whole search space independently. Figure 5 shows the mutation process. Firstly, randomly choose a trip v_k , from the parent. Secondly, calculate the connection costs between trip v_k with those trips with the departure station s_k^a , but trip v_{k+1} should not be included. Then, find the trip v_p whose connection cost with trip v_k is lowest, and swap the position of trip v_p with trip v_{k+1} . Finally, a new solution has been generated.

(v) *Termination.* The process should be terminated when the difference between the best solutions in two generations is less than a given parameter or a maximum number of iterations are reached.

The processes of crossover and mutation are two important operations in GA, which decide the quality of solutions.

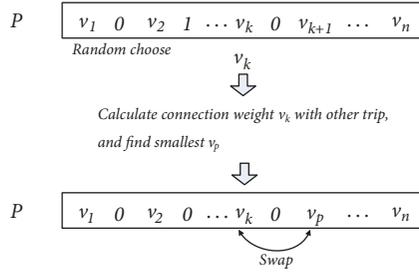


FIGURE 5: Mutation procedure.

5. Case Study

5.1. Basic Information. To evaluate the proposed model and algorithm, we used the HSR between Shanghai transportation hub and Nanjing transportation hub as a case study. Figure 6 illustrates the topological structure of the case study. As shown, there are two HSR lines between the two hubs, namely, Hu-Ning HSR and Jing-Hu HSR, the lengths of which are 301 kilometers and 295 kilometers, respectively. There are totally four railway stations in the two hubs, including Shanghai Station, Hongqiao Station, Nanjing Station, Nanjing, and South Station. All the four stations are near RID. Moreover, connection railway lines are built between Nanjing Station and Nanjing South Station as well as between Shanghai Station and Hongqiao Station. The train-sets of CRH380 series are utilized to complete the trip tasks. The number of trip tasks is listed in Table 2.

The parameters related to the TCP are set as follows:

- (i) Minimum time for preparation work in station i : $t_i^W = 24$ minutes.
- (ii) Travelling distance standard of Level One maintenance: $S_{one} = 4,000$ kilometers.
- (iii) Travelling time standard of Level One maintenance: $T_{one} = 2$ days.
- (iv) Fluctuation coefficient: $\Delta = 10\%$.
- (v) Running time from station i , to station k by the connection rail: the values of t_{ik}^C are listed in Table 3.
- (vi) Basic coefficients for GA: the population size is set as 80; the probability of crossover is set as 0.6; the probability of mutation is set as 0.01; and the maximum iteration times is set as 50.

5.2. Results. Based on the above information, we tested our proposed model and algorithm in the PC (WIN7, Intel Core i7-4779, 3.40GHz, and 16GRAM). Comparison experiments were conducted to evaluate whether introducing dispatching train-sets between different stations in the same hub can improve the train-set utilization efficiency. We tested 20 times for both conditions, and the best computational results are listed in Table 4.

From Table 4, it clearly shows that when train-sets can be dispatched between different stations in the same hub, the number of needing train-sets can be reduced. Particularly, to fulfill the same trip tasks, the best computational results

TABLE 2: Trips between different stations.

	Nanjing Station	Nanjing South Station
Shanghai Station	35 pairs	3 pairs
Hongqiao Station	14 pairs	4 pairs

TABLE 3: Running time between different stations in the same hub.

Departure Station	Arriving Station	Running Time
Shanghai Station	Hongqiao Station	17 minutes
Hongqiao Station	Shanghai Station	18 minutes
Nanjing Station	Nanjing South Station	23 minutes
Nanjing South Station	Nanjing Station	23 minutes

show that 2 train-sets can be saved. Moreover, the times of maintenance can be also reduced on average. The utilization efficiency of train-sets is an important indicator measuring the quality of TCP. In this paper, a variable η is introduced to quantitatively describe the train-set utilization efficiency. The value of η can be calculated by (13), where T_{total} denotes the total running time of all trips, T is equal to 1440 minutes representing one day, T_M is the maintenance time of railway lines in the midnight, usually being 240 minutes, and M denotes the number of train-sets. The computational results reveal that, compared with traditional utilization mode, the mode allowing train-sets being dispatched between different stations in the same hub can formulate a more efficient TCP. The average train-set utilization efficiency of our proposed utilization mode is 45.1%, which is 5.5% higher than that of traditional utilization mode.

$$\eta = \frac{T_{total}}{(T - T_M) \times M} \quad (13)$$

5.3. Convergence Test. To further illustrate the efficiency of the proposed approach, a convergence test of objective values is conducted. To reach the best found solution for the objective function, 30 iterations were performed. As shown in Figure 7, the objective values stop changing after 23 iterations. The results indicate that the algorithm can converge to a steady state within the given maximal iterations.

5.4. Comparison of Different Algorithms. In order to highlight the performance of this paper in improving the train-set utilization efficiency, we compared our algorithm with ACA [35] and SA [36], and detail comparison results are as listed in Table 5. Three important findings can be concluded.

- (1) GA can obtain a better TCP than ACA and SA. After 30 times computations, 25.3 train-sets are needed on average to finish the tasks, which are less than the results of ACA (28.8) and SA (30.1). The average train-set utilization efficiency by GA is much higher than that of ACA and SA, which means that the TCP obtained by GA can save more costs and serve more passengers with the same amount train-sets.
- (2) The computational time of GA is about 8 seconds, which is similar to SA (8 seconds) but longer than

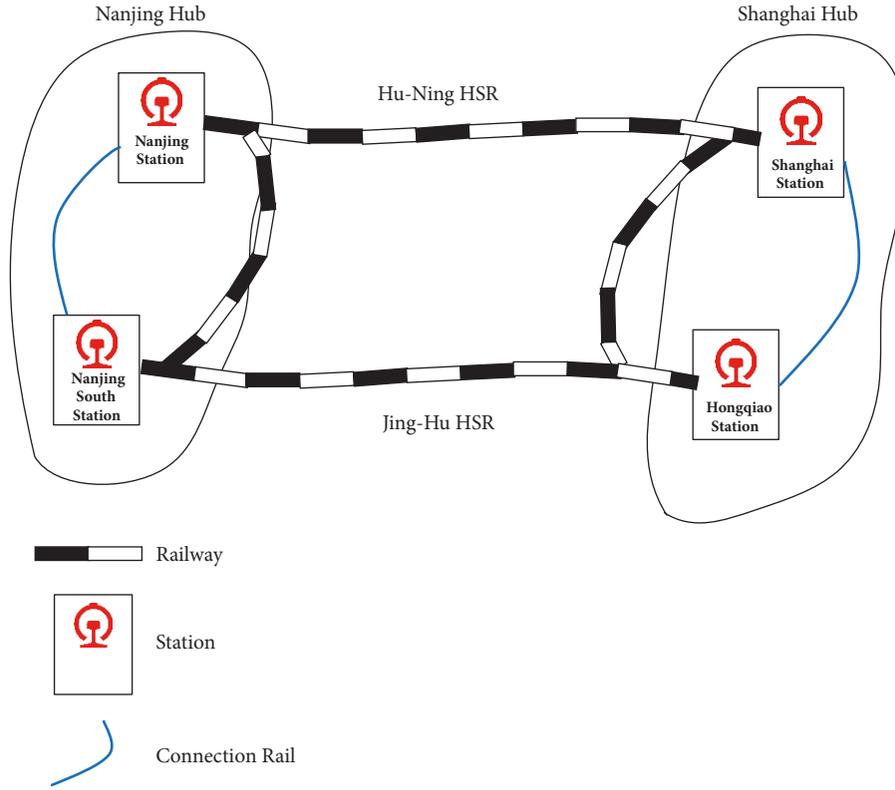


FIGURE 6: Topological structure of case study.

TABLE 4: Computational results between two different utilization modes.

Scenario	Train-set Amount	Maintenance Amount	Train-set Utilization Efficiency
Dispatching train-sets between different stations	24 (25.3)	10 (11.5)	51.9% (45.1%)
Not dispatching train-sets between different stations	26 (28.4)	10 (12.1)	47.9% (39.6%)

¹ The number in parentheses represents the average value for 20 times.

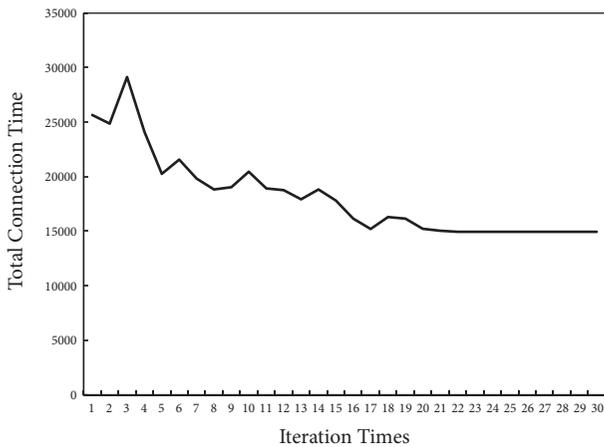


FIGURE 7: Convergence curve.

ACA (4 seconds). The reason is that the operations of crossover and mutation are time-consuming.

(3) Figure 8 demonstrates the convergence curves the three algorithms. The results reveal that ACA has the fastest convergence speed, followed by GA and finally

TABLE 5: Results of different algorithms.

Average Value	GA	ACA	SA
Train-sets Amount	25.3	28.8	30.1
Average Maintenance Amount	11.5	13.2	14.5
Efficiency	45.1%	39.4%	35.2%
Computational time (millisecond)	8321.5	4522.2	8421.5

SA. Moreover, the curves of GA and ACA are smooth, while the SA is fluctuating.

Based on the above analyses, it can be found that the GA for TCP has obvious advantages in the solution quality, and its computational efficiency and stability also have good performances. Thus, it can be said that the method proposed in this paper is a new way to utilizing train-sets with lower costs and high efficiency.

6. Conclusions

With the rapid development of HSR in China, the TCP problem, as a fundamental and vital part of railway management,

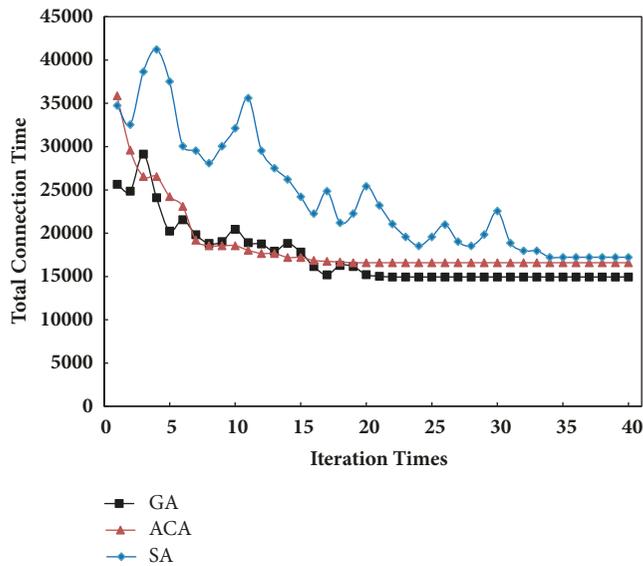


FIGURE 8: Comparison of convergence curves.

has been receiving increasing attention. How to efficiently utilize the train-sets in the complex HSR network becomes a hotspot. With the emergence of the transportation hub in China, the higher challenge is faced by all corresponding scholars and operators. To cope with this new development trend, this paper aims to put forward a novel utilization mode to enhance the utilization efficiency of the train-set. We modified the traditional train-sets' utilization mode that train-sets can only undertake trips that the departure stations are the same with the arriving stations of the trips the train-sets have just completed. Instead, an innovative train-set utilization mode is put forward that train-sets can be dispatched among railway stations in the same hub. We formulate an integer programming TCP model with the objective of simultaneously minimizing the number of using train-sets and the total maintenances times. In order to deal with the complicated maintenance constraints, accumulated variables have been introduced to represent the running distance and running time of the train-set. To obtain the optimal TCP, a genetic algorithm (GA) is designed. What makes the GA more unique is that our expression of the solution which considers the maintenance arc and connection arc and the crossover process makes good use of characteristics of maintenance constraints. In order to deal with the complicated maintenance constraints, accumulated variables have been introduced to represent the running distance and running time of the train-set. To obtain the optimal TCP, a genetic algorithm (GA) is designed. What makes the GA unique is our expression of the solution. The proposed GA considers the maintenance arc and connection arc, and the crossover process makes good use of characteristics of maintenance constraints. In order to verify the efficiency of the proposed utilization mode, numerical experiments and contrast experiments have been carried out based on the real data of Nanjing and Shanghai HSR transportation hubs. The results show that a high-quality train-set utilization scheme can be obtained. It can guide

dispatchers to determine each trip should be undertaken by which train-set, the sequence of the trips undertaking by the same train-set, and when and where each train-set should be maintained. Also, the flexibility and the efficiency of train-set utilization can be significantly enhanced, even though it may cause empty running. Additionally, the designed GA has obvious advantages over ACA and SA in both the solution quality and computational efficiency. Further research work is recommended in twofold: one is to take the fluctuation and perturbation in the real world into consideration when formulating the TCP problem and study the rescheduling problem of train-set utilization. The other is to consider the impacts of passengers flow on the TCP problem, which is a vital factor influencing the dwell time and travel time in the TCP problem.

Data Availability

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

Conflicts of Interest

Authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work is financially supported by the National Natural Science Foundation of China (91746201; 71621001).

References

- [1] Y. Wang, X. D. Yan, Y. Zhou, and Q. W. Xue, "Influencing mechanism of potential factors on passengers' long-distance travel mode choices based on structural equation modeling," *Sustainability*, vol. 9, no. 11, Article ID 1943, 2017.
- [2] H. R. Zhang, L. M. Jia, L. Wang, and X. Y. Xu, "Energy consumption optimization of train operation for railway systems: Algorithm development and real-world case study," *Journal of Cleaner Production*, vol. 214, pp. 1024–1037, 2019.
- [3] V. Cacchiani, A. Caprara, and P. Toth, "Solving a real-world train-unit assignment problem," *Mathematical Programming*, vol. 124, no. 1-2, pp. 207–231, 2010.
- [4] V. Cacchiani, D. Huisman, M. Kidd et al., "An overview of recovery models and algorithms for real-time railway rescheduling," *Transportation Research Part B: Methodological*, vol. 63, pp. 15–37, 2014.
- [5] J. T. Haahr, J. C. Wagenaar, L. P. Veelenturf, and L. G. Kroon, "A comparison of two exact methods for passenger railway rolling stock (re)scheduling," *Transportation Research Part E: Logistics and Transportation Review*, vol. 91, pp. 15–32, 2016.
- [6] K. Ziarati, F. Soumis, J. Desrosiers, and M.-M. Solomon, "A branch-first, cut-second approach for locomotive assignment," *Management Science*, vol. 45, no. 8, pp. 1156–1168, 1999.
- [7] J.-F. Cordeau, F. Soumis, and J. Desrosiers, "A benders decomposition approach for the locomotive and car assignment problem," *Transportation Science*, vol. 34, no. 2, pp. 133–149, 2000.

- [8] J.-F. Cordeau, G. Desaulniers, N. Lingaya, F. Soumis, and J. Desrosiers, "Simultaneous locomotive and car assignment at VIA Rail Canada," *Transportation Research Part B: Methodological*, vol. 35, no. 8, pp. 767–787, 2002.
- [9] R. Ahuja, C. Cunha, and G. Sahin, "Network models in railroad planning and scheduling," *Mathematics of Operations Research*, pp. 54–101, 2005.
- [10] B. Vaidyanathan, R. K. Ahuja, and J. B. Orlin, "The locomotive routing problem," *Transportation Science*, vol. 42, no. 4, pp. 492–507, 2008.
- [11] A. Schrijver, "Minimum circulation of railway stock," *CWI Quarterly*, vol. 6, pp. 205–217, 1993.
- [12] P. Zhao, H. Yang, and A. Z. Hu, "Research on usage of High-Speed passenger trains on uncertain railroad region," *Journal of the China Railway Society*, vol. 19, no. 2, pp. 15–19, 1997.
- [13] E. Abbink, B. van den Berg, L. Kroon, and M. Salomon, "Allocation of railway rolling stock for passenger trains," *Transportation Science*, vol. 38, no. 1, pp. 33–41, 2004.
- [14] J. Yang, H. Yang, and H. B. Lu, "Application of genetic algorithm in optimizing DMU circulating model," *Railway Transport and Economy*, vol. 26, no. 7, pp. 65–68, 2004.
- [15] N. Ben-Khedher, J. Kintanar, C. Queille, and W. Stripling, "Schedule optimization at SNCF: from conception to day of departure," *Interfaces*, vol. 28, no. 1, pp. 6–23, 1998.
- [16] A. Alfieri, R. Groot, L. Kroon, and A. Schrijver, "Efficient circulation of railway rolling stock," *Transportation Science*, vol. 40, no. 3, pp. 378–391, 2006.
- [17] P. Fioole, L. Kroon, G. Maróti, and A. Schrijver, "A rolling stock circulation model for combining and splitting of passenger trains," *European Journal of Operational Research*, vol. 174, no. 2, pp. 1281–1297, 2006.
- [18] G. Maróti and L. Kroon, "Maintenance routing for train units: the interchange model," *Computers & Operations Research*, vol. 34, no. 4, pp. 1121–1140, 2007.
- [19] Y. Wang, Y. Gao, X. Yu, I. Hansen, and J. Miao, "Optimization model for high-speed train unit routing problems," *Computers & Industrial Engineering*, 2018.
- [20] G. Maróti and L. Kroon, "Maintenance routing for train units: the transition model," CWI Research Report PNA-E0415, Amsterdam, 2004.
- [21] S. Hong, K. M. Kim, K. Lee, and B. Hwan Park, "A pragmatic algorithm for the train-set routing: the case of Korea high-speed railway," *Omega*, vol. 37, no. 3, pp. 637–645, 2009.
- [22] T. Nishi, A. Ohno, M. Inuiguchi, S. Takahashi, and K. Ueda, "A Combined column generation and heuristics for railway short-term rolling stock planning with regular inspection constraints," *Computers & Operations Research*, vol. 81, pp. 14–25, 2017.
- [23] L. D'Acerno, 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.
- [24] D. Li, Y. Yin, and H. He, "Testing the generality of a passenger disregarded train dwell time estimation model at short stops: Both comparison and theoretical approaches," *Journal of Advanced Transportation*, vol. 2018, Article ID 8521576, 16 pages, 2018.
- [25] S. Buchmuller, U. Weidmann, and A. Nash, "Development of a dwell time calculation model for timetable planning," *WIT Transactions on The Built Environment*, vol. 103, pp. 525–534, 2008.
- [26] N. G. Harris, "Train boarding and alighting rates at high passenger loads," *Journal of Advanced Transportation*, vol. 40, no. 3, pp. 249–263, 2006.
- [27] G. Fletcher and A. El-Geneidy, "The effects of fare payment types and crowding on dwell time: A fine-grained analysis," *Transportation Research Record*, vol. 2351, pp. 124–132, 2012.
- [28] L. D'Acerno, M. Botte, M. Gallo, and B. Montella, "Defining reserve times for metro systems: An analytical approach," *Journal of Advanced Transportation*, 2018.
- [29] Y. Wang, A. D'Ariano, J. Yin, L. Meng, T. Tang, and B. Ning, "Passenger demand oriented train scheduling and rolling stock circulation planning for an urban rail transit line," *Transportation Research Part B: Methodological*, vol. 118, pp. 193–227, 2018.
- [30] W. Zhou and H. Teng, "Simultaneous passenger train routing and timetabling using an efficient train-based Lagrangian relaxation decomposition," *Transportation Research Part B: Methodological*, vol. 94, pp. 409–439, 2016.
- [31] D. Canca and E. Barrena, "The integrated rolling stock circulation and depot location problem in railway rapid transit systems," *Transportation Research Part E: Logistics and Transportation Review*, vol. 109, pp. 115–138, 2018.
- [32] China Railway Corporation, *The Regulation of Railway Management (High Speed Railway Part)*, China Railway Press, Beijing, China, 2014.
- [33] L. D'Acerno, M. Gallo, B. Montella, and A. Placido, "Analysis of the interaction between travel demand and rail capacity constraints," *WIT Transactions on the Built Environment*, vol. 128, pp. 197–207, 2012.
- [34] L. D'Acerno, A. Placido, M. Botte, M. Gallo, and B. Montella, "Defining robust recovery solutions for preserving service quality during rail/metro systems failure," *International Journal of Supply and Operations Management*, vol. 3, no. 3, pp. 1351–1372, 2016.
- [35] Y. Zhou, L. Zhou, and Y. Wang, "Using improved ant colony algorithm to investigate EMU circulation scheduling problem," *Discrete Dynamics in Nature and Society*, vol. 2014, Article ID 767429, 13 pages, 2014.
- [36] H. Chen, *Research on the relative problems of motor train-set scheduling system [Master, thesis]*, Southwest Jiaotong University, Chengdu, China, 2007.



Hindawi

Submit your manuscripts at
www.hindawi.com

