Application of Multiple-Population Genetic Algorithm in Optimizing the Train-Set Circulation Plan Problem

Yu Zhou, Leishan Zhou, Yun Wang, Zhuo Yang, and Jiawei Wu

1Department of Transportation Management Engineering, School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China
2Ministry of Education (MOE) Key Laboratory for Urban Transportation Complex System Theory and Technology, School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China
3Department of Civil, Environmental, and Infrastructure Engineering, Volgenau School of Engineering, George Mason University, Fairfax, VA, USA
4Center for Advanced Transportation System Simulation, Department of Civil Environment Construction Engineering, University of Central Florida, Orlando, FL, USA

Correspondence should be addressed to Leishan Zhou; lshzhou@bjtu.edu.cn

Received 2 January 2017; Revised 5 April 2017; Accepted 28 May 2017; Published 2 July 2017

Academic Editor: Jose Egea

Copyright © 2017 Yu Zhou 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.

The train-set circulation plan problem (TCP) belongs to the rolling stock scheduling (RSS) problem and is similar to the aircraft routing problem (ARP) in airline operations and the vehicle routing problem (VRP) in the logistics field. However, TCP involves additional complexity due to the maintenance constraint of train-sets: train-sets must conduct maintenance tasks after running for a certain time and distance. The TCP is nondeterministic polynomial hard (NP-hard). There is no available algorithm that can obtain the optimal global solution, and many factors such as the utilization mode and the maintenance mode impact the solution of the TCP. This paper proposes a train-set circulation optimization model to minimize the total connection time and maintenance costs and describes the design of an efficient multiple-population genetic algorithm (MPGA) to solve this model. A realistic high-speed railway (HSR) case is selected to verify our model and algorithm, and, then, a comparison of different algorithms is carried out. Furthermore, a new maintenance mode is proposed, and related implementation requirements are discussed.

1. Introduction

During the last few decades, high-speed railway (HSR) has developed rapidly in China, with a total length of 19,000 kilometers [1], accounting for approximately 60% of all HSR in the world. Due to the characteristics of high-capacity, high efficiency, and low-energy, HSR has been a major transportation mode to satisfy passengers’ travel demands. In HSR, the train-set functions as the means to carry passengers, which is a collection of passenger cars. Generally, the passenger cars can be classified into two types: (1) self-propelled cars, which means that the engine is installed in the cars and can be treated as power units, and (2) cars without engines. In China, a train-set is usually comprised of 8 passenger cars, including either 4 or 6 self-propelled cars. Moreover, the length of a train can only be increased by coupling train-sets.

As the carrying tools of HSR, train-sets play a vital role in the HSR management, the utilization efficiency of which greatly influences the HSR operating cost. How to maximize the train-sets’ utilization efficiency has been a research hotspot for many years. On the one hand, due to the high cost of purchasing a train-set, it is crucial to reduce the number of required train-sets, meaning to fulfill the passengers’ travel demands by using as few train-sets as possible. On the other hand, due to the high train-sets’ maintenance expense, how to reduce the times of maintenance has been a great challenge for HSR operating company.

When a timetable is published, train-sets must be assigned to specific routes to satisfy passengers’ travel demands. The train-set circulation plan (TCP) is a technical scheme that guides the train-set assignments, which determines the connecting order and maintenance time of
The remainder of this study is organized as follows. Section 2 reviews relevant studies in the literature. Section 3 develops a modeling framework for the TCPP, including a problem statement and optimization model construction. Section 4 introduces the MPGA to solve this model. In Section 5, a case study based on realistic HSR is performed, and the comparison with other algorithms is carried out. In Section 6, a new maintenance mode is put forward, and relevant preparation work is discussed. Finally, Section 7 provides conclusions and suggestions for future research directions.

2. Literature Review

Many scholars and researchers have studied the ARP and the VRP with maintenance constraints. Compared with the VRP, the ARP with maintenance is much more similar to the TCPP. The ARP with maintenance is to determine which aircraft should fly which segment and when and where each aircraft should undergo different levels of maintenance checks required by the Federal Aviation Administration [2]. Related studies usually consider type-A checks, and the maintenance standard is approximately 3–4 days. It is assumed that the maintenance can be done only at night [3–5]. The maintenance standard of the TCPP is scheduled by the running time and running distance. However, due to the high utilization efficiency, the running distance usually first reaches the maintenance standard, and the maintenance can be performed during the entire day, in addition to the evening. These factors make the TCPP more flexible, and, therefore, the approaches to solve the ARP with maintenance cannot be directly implemented in the TCPP. However, the approaches to solve the ARP with maintenance have inspired scholars and researchers to solve the TCPP. The solution methodologies for the ARP with maintenance were categorized into three approaches by Liang et al. [6]. The most common approach is to model a sequence of aircraft rotations as connecting flight strings and find the optimal routing by solving a set partitioning problem [7–9]. The second approach is to model the ARP as Euler tour problem or asymmetric traveling salesman problem with side constraints [10–12]. The last approach is to convert the problem into a network flow problem [13, 14]. In related studies, the flight leg can be carried out by any type of aircraft, but, in the TCPP, different trips need different type of train-sets. For example, there are two HSR lines between Beijing and Shanghai. One is the new-built HSR (design speed: 300 km/h) and the other is the rebuilt HSR from the conventional railway (design speed: 200 km/h). If one trip runs on the new HSR, only the CRH3 series train-sets or other higher level train-sets can be used to carry out the trip. If the trip runs in the rebuilt HSR, most of the train-set types can be used. The differences in infrastructures (signal, power supply, train control systems, etc.) between lines make the TCPP slightly different from the ARP with maintenance. However, the idea of modeling and the solution approaches of the ARP have inspired corresponding studies of the TCPP. Zhao et al. [15] modeled the TCPP as the TSP with side constraints similar to the second approach mentioned above. In brief, the VRP...
and the ARP with maintenance have inspired researchers to study the TCPP, but these approaches cannot be directly used for the TCPP due to the different situations between aircraft utilization and train-set utilization.

The train-set utilization problem has been a hot research topic for several decades and belongs to the field of railway routing and scheduling. Extensive studies have been conducted on this problem worldwide. HSR first developed rapidly in Europe; consequently, many European scholars and researchers have focused on this problem. Schrijver [16] first studied the train-set scheduling problem and proposed a basic model based on the minimum cost flow theory. Then, the model was solved using CPLEX software. His model and solution approaches inspired later researchers. Abbink et al. [17] studied the marshaling of train-sets during morning rush hours; the goal of the model was to minimize shortages of train seats. This study contributed to passenger service during rush hours. Peeters and Kroon [18] developed a model that addressed train-set marshaling and rolling stock utilization. Their work was unique because they applied D-W decomposition and used branch and bound techniques to solve their model, which provided ideas for later studies. Arianna et al. [19] addressed the train-set utilization problem on a single train line and for a single day using an integer programming model to obtain the rolling stock circulation while considering the order of the train units in the compositions. Fioole et al. [20] proposed a mixed integer programming model based on widely adopted previous research achievements and applied an improved branch and bound algorithm to obtain an optimized solution. Together, these studies exploring the train-set utilization problem represent great contributions to this research field. Most of the studies refer to previous locomotive and car assignment problems [21–24]. However, these studies have rarely considered maintenance constraints.

In practical utilization, maintenance must be conducted on train-sets after running for a certain time or distance. As train-sets are the main carrying mechanism, carrying out maintenance is one of the most important aspects to ensure HSR operation safety. Moreover, proper maintenance can keep the train-sets in good operation status, which will effectively reduce the possibility of perturbations and disruptions and, thus, contribute to a high level of punctuality and high service availability. Therefore, these previous studies have some limitations that make them difficult to apply in practice.

The research focus later shifted to the train-set utilization problem under maintenance constraints. There are two standards for maintenance, namely, running distances and running time. After a train-set has traveled a certain distance or has been operated for a specific length of time, maintenance must be undertaken to keep the train-set in good condition. When these maintenance constraints are taken into consideration, the train-set utilization problem becomes NP-hard, which has been proven in studies such as Cacchiani et al. [25]. There are no polynomial-time algorithms that can solve NP-hard problems [26]; therefore, the development of efficient algorithms for solving such problems has gained increasing attention. Heuristic algorithms are both suitable and efficient for NP-hard problems; their solutions are usually near-optimal, and their computational time is acceptable. These characteristics have popularized the use of heuristic algorithms for solving the train-set utilization problem. Maróti and Kroon [27] designed an interchange strategy with a shortest-path heuristic algorithm to solve the Netherland Passenger Rail rolling stock problem. This solution was put into practice on the Nederlandse Spoorwegen (NS) lines and turned out to be efficient. Cadarso et al. [28, 29] divided the train-set utilization problem into two subproblems, namely, the train-set assignment problem and the train-set routing problem, and then proposed a heuristic algorithm based on Bender’s decomposition. The RENFE (the main Spanish operator of suburban passenger trains) in Madrid, Spain, was used as a case study to verify the proposed approach. The results of this model, which can be solved in approximately 1 minute, were received positively by RENFE planners. Similar studies have been carried out around the world. For example, Hong et al. [30] formulated a two-stage heuristic algorithm for the Korea Train Express (KTX) case, and Thorlacius et al. [31] proposed a hill climbing heuristic to improve the existing rolling stock plans of the DSB S-tog in Copenhagen, Denmark. It is worth mentioning that the passenger demand constraint of these studies has usually been satisfied by considering maximization of passenger service and minimizing seat shortages. The inputs are passenger demand and train-sets, and the train-sets can be reconnected and decomposed. The papers in European aim to provide services as much as possible to reduce the shortage of seats. Thus, the goals and the constraints of the train-set utilization problem are different between China and Europe.

In contrast, in China’s HSR situation, train-set compositions are usually fixed, containing 8 or 16 passenger cars; consequently, the capacity for each trip is also fixed. The passenger demand is dealt with when working out the train diagram, which provides the train service frequency. In the process of formulating the train-set utilization plan, the inputs are all the trips in the train diagram and the train-sets. Over the past two decades, many scholars and researchers have studied the TCPP in China’s HSR situation. Table 1 provides a systematic comparison of the key model components and solution methods in the existing studies in China. Most of these studies took maintenance constraints into consideration, and their objectives focus primarily on minimizing total connection time, minimizing maintenance costs, and balancing the utilization of train-sets. A variety of self-organization heuristic approaches have been designed to solve this problem. Zhao et al. [15] took the TSP as an example and introduced the train-set utilization network. Then, they used a TSP-based heuristic algorithm to solve the model, which turned out to be an efficient solution. Many scholars later tried to optimize the train-set utilization problem based on the approach of Zhao et al. Various types of bionic heuristic algorithms including the GA (Genetic Algorithm) [32], the ACA (Ant Colony Algorithm) [33], the PSO (Particle Swarm Optimization) [34], and an improved ACA [35] have been proposed. Although these algorithms have made great contributions to the problem, they are difficult to apply in practice due to their poor local optimization ability and slow computational speeds. As the scale of the problem...
Complexity

<table>
<thead>
<tr>
<th>Maintenance constraints</th>
<th>Objective function</th>
<th>Multi. type</th>
<th>Solution approach</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>Distance</td>
<td>Min connecting time</td>
<td>Min maintenance cost</td>
<td>TSP heuri. + random choice</td>
</tr>
<tr>
<td>√</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td>GA heuri.</td>
</tr>
<tr>
<td>×</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>ACA heuri.</td>
</tr>
<tr>
<td>√</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>PSO heuri.</td>
</tr>
<tr>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>Hierarchical heuri.</td>
</tr>
<tr>
<td>×</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>Improved ACA heuri.</td>
</tr>
<tr>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>MPG A heuri.</td>
</tr>
</tbody>
</table>

In addition, these algorithms become time-consuming and easily become stuck in local optimal solutions. In addition, the previous studies only considered one type of train-set, but there are many types of train-sets in practice. Only train-sets of the same type can connect with each other. After reviewing the shortcomings of these algorithms, we realized that the performance could be further improved by adopting more intelligent search mechanisms. The MP GA is based on partitioning a population into several semi-isolated subpopulations. Each subpopulation is associated with an independent GA and explores different promising regions of the search space. Therefore, we can propose a train-set utilization model that considers different types of train-sets. We design an efficient version of MP GA to solve the TCPP, and then, we compare the MP GA with other algorithms. Finally, a realistic case study was carried out to verify our model and algorithm. The results showed that MP GA is efficient and obtains better solutions than previous approaches.

3. Modeling Framework

3.1. Problem Statement. Figure 1 shows a simple example of train diagram (TD), which is formulated based on passenger demands. The task in the train diagram undertaken by a train-set is called a trip, which might cover multiple stations. For example, in Figure 1, trip G1 goes from station A to station C via station B. Importantly, the train diagram involved in this paper is scheduled with trains in pairs, indicating that, for each station in the train diagram, the number of inbound trips is equal to the number of outbound trips. The horizontal axis is time and the vertical axis represents the stations. The connecting order in our paper means the order of the trips undertaken by the same train-set. As shown in Figure 2, there are totally 4 trips, and the connecting order is 1→4→3→2, which means that a train-set should undertake trip 1 firstly, followed by trip 4, trip 3, and finally trip 2.

Figure 3 shows the corresponding TCP, while Figure 4 shows the train-set utilization process. On the first day, a train-set performs trips G1, G2, and G3 and stays overnight at station D. On the second day, the same train-set starts from station D and performs the task sequence of G4, G5, G6, and G7. On the third day, after completing trips G9 and G10, the train-set returns to its original departure station A and undergoes the maintenance task. A complete train-set circulation has been carried out. Note that there may be many different TCPPs based on the same TD. The goal of the TCPP is to find a better solution that needs fewer train-sets and reduces maintenance costs.

Train-set safety is of primary importance. After a train-set has been utilized for a certain time or traveled a certain distance, maintenance must be performed to maintain them in good condition. In China, train-set maintenance is divided into 5 levels by the CRC based on the maintenance tasks involved, and different types of train-sets have different maintenance standards. Table 2 lists the maintenance standards.
Table 2: Maintenance standard of train-sets in China.

<table>
<thead>
<tr>
<th>Train-Set Series</th>
<th>Distance Standard</th>
<th>Time Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRH1 series</td>
<td>4,000 km/48 h</td>
<td>Uncertain</td>
</tr>
<tr>
<td>CRH5A</td>
<td>5,000 km/48 h</td>
<td>Uncertain</td>
</tr>
<tr>
<td>CRH380B series</td>
<td>4,000 km/48 h</td>
<td>Uncertain</td>
</tr>
<tr>
<td>CRH380A series</td>
<td>4,000 km/48 h</td>
<td>Uncertain</td>
</tr>
</tbody>
</table>

Carried out in the DIR but must be carried out in the train-set factories. Specific train-set maintenance plans for level 2 to 5 standard maintenance should be formulated. Thus, when formulating a TCP, only level 1 standard maintenance needs to be taken into consideration.

As discussed above, many types of train-sets exist in China. Train-sets are bound to a TUB. A TUB is a place with comprehensive train-set operation and management facilities: a DIR belongs to a TUB. Train-sets can undergo maintenance only in the DIR corresponding to its TUB. For example, train-set A is bound to TUB M. When train-set A reaches a maintenance standard, that maintenance task can be undertaken only at the corresponding DIR for TUB M. The reasons are as follows.

1. Different types of train-sets require different equipment and mechanical parts during maintenance. Only their corresponding TUB satisfies those maintenance requirements; therefore, the train-sets must return to the TUB to which they are bound when they require maintenance.

2. The second reason is the safety responsibility problem. Suppose that train-set A is bound to TUB M but undergoes maintenance at TUB N for some reason. Unfortunately, a maintenance accident occurs to train-set A, and then it is unclear which TUB should be responsible for this accident. Obviously, both bases would not be responsible for this problem.

Despite the maintenance constraint discussed above, there are many other constraint aspects in the TCPP. (1) When train-sets arrive at a station, some basic work should be done before the train-set embarks on another trip. When a train-set arrives at the station after completing a trip task, some preparation work must be carried out before undertaking the next trip, such as cleaning, pollution discharge, and shift changes of crew members. The minimal duration of these necessary working procedures between two adjacent trip tasks is defined as the minimum connection time. Therefore, connection schedules must consider the time to perform this basic work. As shown in Figure 5, when two trips
are undertaken by the same train-set and the arrival station of the former trip is the same as the departure station of the latter trip, the time interval between the departure time of the latter trip and the arrival time of the former trip is defined as the connection time. After the train-set has completed trip 1 and arrived at its destination, the basic work is performed, and the schedule must accommodate that time, which means the connection time should be longer than the minimum connection time. The train-set will carry out trip 2 after the basic work is complete. (2) When the train-set goes to the DIR for maintenance, the connection time between the two adjacent trips must be longer than the maintenance time. (3) Every trip in the TD should be performed by the train-sets and can be conducted only once. (4) When a train-set reaches a maintenance limit, the last trip's arrival station should be near the binding base so that the maintenance can be easily performed. All the above factors restrict the utilization of train-sets and make TCPP complicated. In this paper, we propose a train-set utilization model and design an MPGA to solve the TCPP and obtain a better solution.

3.2. Optimization Model Construction. We assume that our optimization model is formulated based on a directed complete multigraph $G(V, E)$, where the nodes $v_i \in V$ represent the trips that contain trip number $n_i$, departure station $s_i^d$, arrival station $s_i^a$, departure time $t_i^d$, arrival time $t_i^a$, distance $s_i$, running time $T_i$, and train type $k$—a total of $8$ attributes. $n$ denotes the number of trips. $E$ is the arc set representing all the connection relationships between departure trips and arrival trips. Set $E$ is partitioned into $p$ subsets, namely, $E^1, \ldots, E^p$, where $p$ is the number of train-set types. Let $E^k \in \{E^1, \ldots, E^p\}$ denote the connection relationships of departure trips and arrival trips undertaken by the train-sets of type $k$. The simple directed graph $G^k = (V, E^k)$ is complete. The sequencing constraints are implicitly represented by the costs of the arcs. In particular, if trip $v_i$ can be performed immediately after trip $v_j$ by train-sets of type $k$ on the same day, the cost $c_{ij}^k$ of arc $(i, j)^k \in E^k$ is given by the time in minutes that elapse between the arrival of trip $v_i$ and the departure of trip $v_j$ (i.e., $c_{ij}^k = t_i^a - t_j^d$). Otherwise, if trip $v_j$ can be performed immediately after trip $v_i$ by train-sets of type $k$ on the following day, the cost becomes $c_{ij}^k = t_i^a - t_j^d + 1440$.

We adopt an integer linear programming (ILP) formulation of TCPP with arc variables based on the graph representation of the problem above. The decision variables for our model are as follows: $x_{ij}^k$ is a binary variable that indicates the connection relationship between trips. If $x_{ij}^k = 1$, then trip $v_i$ is connected with trip $v_j$ for a train-set of type $k$ and vice versa. Similarly, $y_{ij}^k$ is a binary decision variable that indicates maintenance. When $y_{ij}^k = 1$, the connection arc $(i, j)^k$ is a maintenance arc, while 0 means the arc is a normal connection arc. Figure 6(a) depicts a normal connection: after trip 1, the train-set continues to perform trip 2, while (b) represents a maintenance situation: after trip 1, the running distance or time has reached the maintenance threshold. Consequently, the train-sets will go to the nearby DIR to undergo maintenance. After maintenance, the train-set returns to the original station and fulfills trip 2.

The goal of the objective function is to minimize the overall cost of the arcs and the amount of maintenance required, as shown in (1), where $\sigma_1$ and $\sigma_2$ are weight coefficients of the total connecting time and the amount of maintenance.

$$\min \sigma_1 \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{p} x_{ij}^k c_{ij}^k + \sigma_2 \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{p} y_{ij}^k.$$ (1)

Constraint (2) requires that the connection time between trip $v_i$ and trip $v_j$ must be longer than the minimum connection time in a station when the connection arc is a normal connection arc. Train-sets must meet this constraint to be able to undertake the next trip.

$$x_{ij}^k (1 - y_{ij}^k) c_{ij}^k \geq x_{ij}^k (1 - y_{ij}^k) T_i^k.$$ (2)

Constraint (3) means that, on a maintenance arc, the connection time between two adjacent trips must be longer than the time required for maintenance at level $1$.

$$x_{ij}^k y_{ij}^k c_{ij}^k \geq x_{ij}^k y_{ij}^k T_i^k M.$$ (3)
Constraints (4) are maintenance restrictions, which ensure that the accumulated running distance and accumulated running time do not exceed the maintenance thresholds.

\[ S_{ij}^{AC} \leq S_{one} (1 + \Delta), \]
\[ T_{ij}^{AC} \leq T_{one} (1 + \Delta). \]  

(4)

Constraints (5) are uniqueness constraints, which ensure that every trip is connected with another trip.

\[ \sum_{j=1}^{n} \sum_{k=1}^{p} x_{ij}^k = 1 \quad \forall j \in V, \]  
\[ \sum_{j=1}^{n} \sum_{k=1}^{p} x_{ij}^k = 1 \quad \forall i \in V. \]  

(5)

Constraint (6) requires that when a train-set undergoes maintenance, its arrival station must be the same as the departure station for the next trip after maintenance. This constraint guarantees that the train-sets undergo maintenance in the DIR corresponding to the binding TUB.

\[ s_i^a = s_j^{INI}, \]
\[ j_{ij}^k = 1, \]  
\[ i \in V, \quad j \in V, \quad i \neq j. \]  

(6)

3.3. Notation

- \( T^k \): minimum connection time for train-sets of type \( k \).
- \( T_{AC}^k \): minimum maintenance time for train-set of type \( k \).
- \( S_{one} \): running distance threshold for maintenance level 1.
- \( T_{one} \): running time threshold for maintenance level 1.
- \( \Delta \): fluctuation coefficient for maintenance thresholds.
- \( S_{ij}^{AC}, T_{ij}^{AC} \): represent accumulated running distance and running time; \( S_{ij}^{AC} = 0 \), and \( T_{ij}^{AC} = 0 \) immediately after maintenance has been performed.

4. Solution Approach

The TCPP is generally an NP-hard problem. There is no precise algorithm to solve it in polynomial time. Researchers typically adopt heuristic algorithms to solve this problem to ensure that the execution time allows the operation managers to obtain a solution within a reasonable amount of time. Genetic algorithms are robust optimum search techniques that find high-quality solutions for complex problems based on principles inspired by natural genetic and evolution mechanisms. The problem solution is replaced by strings called individuals. Multiple individuals comprise a population. Within the population, the fittest individuals are selected based on crossover and mutation operations, which are the basic mechanisms in the high-quality solution search process [37]. However, because these algorithms often exhibit slow convergence and are easily trapped in local optimal solutions, a multiple-population genetic algorithm (MPGA) is proposed in this paper to solve the TCPP, which has been proved to be an efficient optimization technique for highly nonlinear problems [38]. MPGAs are based on partitioning the population into several semi-isolated subpopulations. Each subpopulation is associated with an independent standard genetic algorithm (SGA) and has a different probability of crossover. Thus, different promising regions of the search space can be expanded. Due to the excellent optimization ability and high efficiency, MPGAs have been widely applied in the solving process of various research fields, such as the transportation scheduling [39], the wind turbine layout optimization [40], estimating snow properties [41], and dynamic facility layout problem [42]. Searching among the good solutions in the multipopulation approach avoids the elite individuals of every population and makes searching more effective. As shown in Figure 7, the evolution mechanism of populations 1 to \( n \) is the same as the SGA. Immigrant individuals entering the population achieve coevolution across multiple populations. The good solution is the comprehensive result of the multiple populations.

4.1. Representation Scheme. To avoid the waste of capacity, the “deadhead,” that is, empty trains, is not taken into consideration in this paper. Thus, for each station, the number of arrival trips is the same as that of departure trips. When the number of arrival trips in station \( s \) is \( n_s \), that is, the number of departure trips, a \( n_s \times n_s \) matrix can be built to represent the connection relationship between arrival trips and departure trips. As shown in Figure 8, the horizontal axis represents departure trips, and the vertical axis represents arrival trips. When the value of the \( i \)th line and the \( j \)th column is “1,” the arrival trip \( i \) is connected with the departure trip \( j \). Then, a row matrix can be obtained according to the results of the \( n_s \times n_s \) matrix. The \( i \)th element in the row matrix specifies to which departure trip the arrival trip \( i \) should be connected to. For example, the matrix in Figure 8 can be converted into the row matrix, \( n_s, 1, \ldots, 2 \). In particular, the first element in the row matrix is \( n_s \), which means that the first arrival trip, that is, arrival trip 1 in the example, is connected with the departure trip \( n_s \); the second element in the row matrix is 1, which means that the second arrival trip is arrival trip 2, and arrival trip 2 is connected with departure trip 1; finally, the last element in the row matrix is 2, which means that the last arrival trip, that is, arrival trip \( n_s \) in the example, is connected with departure trip 2.

As previously mentioned, maintenance must be carried out in the DIR near the station, when either the running distances or the running time of train-sets reaches the maintenance standard. Thus, the solution representation of stations near the DIR should include the information on maintenance. Figure 9 illustrates an example of the solution representation of stations near the DIR. When the maintenance is conducted between two connected trips, the number in parentheses is defined as 1, representing a maintenance arc. Otherwise, the number is defined as 0 representing a connecting arc. In the example, the maintenance is conducted between the first arrival trip and the last departure trip, that is, arrival trip 1 and departure trip \( n_s \).
4.2. Initial Populations. With the introduction of the representation scheme, the process of creating initial populations becomes easy. First, a sequence of numbers needs to be generated randomly according to the number of trips in the station. For example, when a station involves 4 arrival trips and 4 departure trips, a sequence of "2, 1, 4, and 3" can be generated randomly. Similarly, the representations of all stations can be generated. Second, the trips should be connected by arcs. Then, the accumulated running time and distances of each connection should be calculated. When the accumulated running time or distances reach the maintenance standard, the arc between the corresponding two adjacent trips should be the maintenance arc (number "(1)"); otherwise, the arc should be the connection arc (number "(0)"). Moreover, when the connecting time between two trips is less than the minimum connection time, the arrival trip should be connected to the departure trip of the next day. Finally, initial populations can be obtained.

4.3. Crossover and Mutation Operators. After generating the representation of the solutions (the individuals), two genetic operators, called crossover and mutation, are used to evolve these individuals. The crossover operator combines two solutions to generate new individuals. It should avoid generating infeasible solutions; otherwise, it loses computational time from searching an infeasible space. The proposed crossover operator involves the following steps. Figure 11 shows an example of a crossover operation.

(1) Two strings are randomly selected from the parent pool.
(2) A crossover point is randomly selected for each station.
(3) For the first station, the numbers before cross point 1 of Parent 1 are directly copied into the offspring. The remaining numbers are put into empty positions according to their relative positions in Parent 2.
(4) Evaluate the feasibility of the generated offspring. If the connection time cannot meet the minimum...
connection time or maintenance time, then move the crossover point backward and check the feasibility again. When no crossover point meets the requirements, then the original solution is selected directly into the offspring.

(5) This procedure is repeated for subsequent periods.

The second genetic operator is mutation. Its purpose is to make distinct changes to offspring to prevent the algorithm from becoming trapped in a local optimal solution. We apply the following procedure as the mutation operator. First, two numbers are randomly selected. Then, the two numbers are swapped within the chromosome. Then, a feasibility check (whose procedure is similar to the crossover operator) is performed. Figure 12 shows an example of a mutation operation.

4.4. Immigrant Operator and Fitness Function. All populations are independent, but they are connected by the immigrant operator. The immigrant operator introduces the best individual found within some predefined time period into other populations, thus exchanging chromosomes among the populations. In our paper, the best individual from each iteration will replace the worst individual in all populations. This process continues until the best individual remains the same for some number of iterations. At that point, the algorithm is complete, and the best solution has been found.

The fitness function is shown in (7), where the denominator represents the total connection time. Its reciprocal is good for evaluating the quality of the solution: a larger fitness function value indicates a better solution.

$$f = \frac{1}{\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{p} x_{ijk} c_{ij}}. \tag{7}$$
5. Case Study

To test our new approach, an HSR network containing 3 HSRs was selected as a study case. As shown in Figure 13, this HSR network contains 3 HSR lines, namely, the Beijing-Tianjin HSR, the Jinan-Qingdao HSR, and the Beijing-Shanghai HSR. The Beijing, Tianjin, Jinan, Qingdao, Nanjing, and Shanghai Stations are the departure and arrival stations that are near the DIRs. Consequently, when train-sets reach the running time or running distance maintenance threshold, they should arrive at these stations and go to the corresponding DIR to undergo maintenance.

The trips between all stations are as follows. Table 3 lists the departure and arrival stations, the train-set type, and the distance for each tested trip.

<table>
<thead>
<tr>
<th>Amount</th>
<th>Train-set type</th>
<th>Distance (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BJ-SH</td>
<td>38 pairs</td>
<td>CRH380BL</td>
</tr>
<tr>
<td>BJ-TJ</td>
<td>70 pairs</td>
<td>CRH3C</td>
</tr>
<tr>
<td>JN-QD</td>
<td>6 pairs</td>
<td>CRH5</td>
</tr>
<tr>
<td>BJ-NJ</td>
<td>3 pairs</td>
<td>CRH380BL</td>
</tr>
<tr>
<td>SH-QD</td>
<td>4 pairs</td>
<td>CRH380BL</td>
</tr>
</tbody>
</table>

5.1. Comparison with Other Algorithms. To test the efficiency of MPGA, we compared the obtained result with that of the traditional GA and the Ant Colony Algorithm (ACA) [35]. What needs to be stated here is that, to ensure the fairness and reliability of the comparisons, the basic information was set to be the same with MPGA, namely, the crossover and mutation processes of GA, the way of generating the initial population, and corresponding parameters except for the possibility of crossover. In this paper, there were 10 populations in the MPGA, and each population had a unique crossover possibility shifting from 0.5 to 0.8. The crossover possibility of the GA was 0.65. In order to make the comparison more reliable, the initial population of GA was 800, and the initial population of MPGA was $10 \times 80 = 800$. The ACA used in this paper was from [35], as well as the corresponding parameters, which were as shown in Table 5. $\alpha$ was the coefficient of track attraction, $\beta$ was the coefficient of vision attraction, $\rho$ was the pheromone evaporation coefficient and $Q$ was a constant.

All the three algorithms were conducted for 50 runs using the same case study. Figure 14 shows the results of the different algorithms. The MPGA and GA both achieved the best values (63 train-sets and 54 maintenance tasks), and the average values of GA and MPGA were quite close. To fulfill the same transportation tasks, the results of ACA were poorer than the MPGA and GA, while 2 more train-sets were needed and 3 more maintenance tasks should be conducted. Figure 14(c) shows the computational time (the period from the program start to when the results are obtained) of different algorithms. The best computational times of MPGA and GA are both about 1200 ms, while the average computational time of GA is 3824.3 ms, which is much longer than the MPGA and ACA. Among 50 times of
GA experiments, only 2 experiments had a relatively short computational time (one is 1241 ms, and the other is 1549 ms), while the others were almost all beyond 4000 ms.

In order to find the reason why there exists such a huge difference in GA, we changed the crossover possibility of GA from 0.5 to 0.8 (0.53, 0.56, 0.59, 0.62, 0.65, 0.68, 0.71, 0.74, 0.77, and 0.80), and 50 computations were conducted under each possibility. As shown in Figure 15, when the possibility was between 0.62 and 0.74, the average computational time would be obviously lower. Moreover, it was found that the optimal results of GA required 63 or 64 train-sets, which was almost the same with MPGA. But the average computational speed of GA was slower than MPGA. The reasons could be as follows. The computational speed is related to both the initial
Table 6: Comparison of two maintenance modes.

<table>
<thead>
<tr>
<th></th>
<th>Train-sets amount</th>
<th>Maintenance amount</th>
<th>Efficiency of train-sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>63</td>
<td>54</td>
<td>33.6%</td>
</tr>
<tr>
<td>New</td>
<td>46</td>
<td>37</td>
<td>43.3%</td>
</tr>
</tbody>
</table>

population and crossover possibility. The initial population of MPGA was the same with GA, but it had more kinds of crossover possibility than GA. The important operator "immigration operator" can accelerate the iteration process while high-quality chromosomes and genes can be inherited and spread among offsprings and different populations. Indeed, sometimes GA might work out faster than MPGA. It was because the initial population was created randomly, and some high-quality individuals might be generated in the initial population. But, generally, MPGA were more likely to have a faster speed than GA by statistics when conducting our case study.

6. Further Discussion

As discussed above, the train-sets should return to the corresponding DIR to undergo maintenance. However, this constraint prevents further improvement of train-set utilization efficiency. As shown in Figure 16, in our case study, in the Beijing-Shanghai HSR, many train-sets bound to the Beijing base perform only two trips, namely, Beijing to Shanghai and back (situation C in Figure 16) and then go to the DIR to undergo maintenance. The running distance maintenance threshold for the CRH380BL-series train-sets is 4000, with a fluctuation of 10%. In our case study, these train-sets reach only $1,318 \times 2 = 2,636$—far below the standard. In situation A, the train-sets perform 4 trips, but the running distance exceeds the maintenance threshold, while, in situation B, the running distance is $1,318 \times 3 = 3,954$—close to the threshold. However, the train-sets cannot return to Beijing for maintenance in this situation, which is not allowed in our model.

Therefore, the cause of low efficiency in the TCP is constraint (6). The reasons are discussed in the Modeling Framework. However, with the rapid development of HSR in China, the HSR network has formed gradually, and, thus, many train-set utilization bases have been built and equipped. If train-set maintenance could be conducted at any DIR, train-set utilization efficiency could be greatly enhanced. In our model, we removed constraint (6) and repeated the case study to test this new utilization mode. Table 6 shows a comparison of the traditional maintenance mode with this new maintenance mode. Under the new maintenance mode, only 46 train-sets are required to fulfill the same transportation tasks—17 fewer train-sets than in the traditional maintenance mode. The maintenance amount decreased from 54 to 37, and the train-set efficiency rose from 33.6% to 43.3%. Obviously, the new maintenance mode is more efficient than the traditional one. However, the premises of this new maintenance mode are as follows:

1. In order to make all types of train-sets able to be utilized in different HSR networks, some upgrade and unified standard work should be done; for example, the signaling system should be upgraded to suit different types of train-sets. Furthermore, tracks, platforms, and other infrastructure elements may also need to be updated. Corresponding station work procedures should be unified. China recently has designed standard train-sets, which will greatly
enhance the utilization efficiency. In addition, the train-sets should not be bound to a particular TUB; all train-sets should be dispatched by the CRC, and the CRC should be responsible for the safety of all train-sets.

(2) The CRC should improve the effectiveness of management of all TUBs by enhancing training and supervision. In addition, a train-set utilization and management system should be established so that the CRC can manage information and monitor condition for all train-sets.

(3) Corresponding maintenance equipment for different train-sets should be replicated among all TUBs. The maintenance capacity should be increased in some busy TUBs to meet maintenance demands at any time.

Additionally, we found that the distance of input trips greatly influences the TCP. When all trips are long distance trips, train-set utilization is inefficient. In contrast, if the trips are all short, almost every train-set can closely approach the maintenance thresholds before undergoing maintenance, which greatly improves the efficiency. These results revealed that the TCP can be formulated by the TD: when the TD contains both short and long trips, the train-sets can perform combinations of long and short trips that make full use of the maintenance thresholds to improve the utilization efficiency. To accomplish this, the corresponding departments should cooperate with each other under the command of the CRC. This is not an easy solution, but it is the future trend in the operation and management of train-sets.

7. Conclusion
A train-set utilization plan created to identify the working arrangements of train-sets is profoundly affected by the given train graph as well as by maintenance rules and regulations. To improve the utilization efficiency of train-sets, this paper proposed an integer programming model that considers maintenance constraints. Because the TCP is NP-hard, an MPMA was designed to solve it. To verify the proposed model and algorithm, a realistic case study was conducted, and our approach was compared with the ACA and the GA. The results showed that the proposed approach is both efficient and reasonable. Furthermore, we proposed a new maintenance mode and discussed its ramifications. If implemented, this new maintenance mode could greatly improve train-set utilization efficiency.

Conflicts of Interest
The authors declare that they have no conflicts of interest.

Acknowledgments
This work is financially supported by Railway Ministry Science and Technology Projects (Grant 2016X005-D) and the Fundamental Research Funds for the Central Universities (Grant 2017JBZ001).

References


Submit your manuscripts at https://www.hindawi.com