Research Article

A Branch and Bound Algorithm for the Exact Solution of the Problem of EMU Circulation Scheduling in Railway Network

Chao Lu, Lei-shan Zhou, Yi-xiang Yue, and Ran Chen

School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Lei-shan Zhou; lshzhou@bjtu.edu.cn

Received 19 June 2015; Revised 20 November 2015; Accepted 17 December 2015

Academic Editor: Francisco Chicano

Copyright © 2016 Chao Lu 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.

This paper is concerned with the scheduling of Electrical Multiple Units (EMUs) under the condition of their utilization on one sector or within several interacting sectors. Based on the introduction of the train connection graph which describes the possible connection relationship between trains, the integer programming model of EMU circulation planning is constructed. In order to analyzing the resolution of the model, a heuristic which shares the characteristics with the existing methods is introduced first. This method consists of two stages: one is a greedy strategy to construct a feasible circulation plan fragment, and another is to apply a stochastic disturbance to it to generate a whole feasible solution or get a new feasible solution. Then, an exact branch and bound method which is based on graph designing is proposed. Duetothecomplexity,thelowerboundiscomputedthroughapolynomialapproximation algorithm which is a modification from the one solving the degree constraint minimum 1-tree problem. Then, a branching strategy is designed to cope with the maintenance constraints. Finally, we report extensive computational results on a railway corridor in which the sectors possess the basic feature of railway networks.

1. Introduction

The assignment of transportation tools to fulfilling a group of tasks under certain conditions is one of the most important scheduling activities in many real world applications, such as the Aircraft Routing Problem (ARP) in airline operation, Vehicle Routing Problem (VRP) in logistics delivery, and the Locomotives Scheduling Problem (LSP) in railway system. As the mobile resources, the quality of the utilization of the transportation tools has a great effect on the efficiency and the economic benefits of the transportation enterprises. The operation of these resources also provides a foundation for the conduction of other works and has an interaction with them.

In high-speed railway network, Electrical Multiple Units (EMU) are the main tools used for the passenger transportation. They are significantly cost equipment, because not only is their acquisition or construction expensive but also they need power supply and regular maintenance. The scheduling of EMU falls in the category of rolling stock circulation problems and the operation plan of them determines the concrete activities they need to perform when the system is running. In order not to result in the waste of the precious resources, they need to be operated as reasonably as possible.

The EMU operation plan in each country differs in small points. In China, the EMU operation plan defines the arranging of the train trips and maintenance tasks of each level to each EMU within a certain period of time according to the given train timetable, the configuration of EMU, the condition of EMU maintenance facilities, and the rules of the operation of the stations. It is a comprehensive plan and could be generally divided into three parts, namely, the EMU circulation plan, the EMU allocation plan, and EMU maintenance plan. Among these plans, the EMU circulation plan mainly determines the connection relationship of the train trips and arranges the first level maintenance for EMUs. It is a basis for the other two plans.

The optimization of the rolling stock circulation scheduling problem has received a lot of attention from researchers. Some realistic problems such as the unit coupling or uncoupling when scheduling the EMU are also considered. Cacciani et al. [1] designed an effective heuristic procedure
for the train unit assignment problem that were able to find solutions significantly better than the “manual” solutions found by practitioners. Cacchiani et al. [2] proposed a heuristic which is based on the natural Lagrangian relaxation of a natural integer linear programming of the train unit assignment problem which turns out to be much faster in practice and still providing solutions of good quality. Abbink et al. [3] deal with the tactical problem of finding the most effective allocation of the train types, subtypes, and units of rolling stock to the train series, such that as many people as possible can be transported with a seat, especially during the rush hours. Lin and Kwan [4] proposed a two-phase approach for the train unit scheduling problem: the first phase assigns and sequences train trips to train units considering some real-world scenarios and the second phase focuses on satisfying the remaining station detailed requirements. Peeters and Kroon [5] focused on the efficient circulation of train units within a certain scope of railway line given the timetable and the passengers’ seat demand and a branch-and-price algorithm is described. Alfieri et al. [6] present a solution approach based on an integer multicommodity flow model to determine the appropriate numbers of train units of different types together with their efficient circulation on a single line. Cacchiani et al. [7] present two integer linear programming (ILP) formulations together with their relaxations (the linear programming (LP) relaxation and Lagrangian based approach, resp.) to assign the train units to the trips with minimum cost. When the high-speed railway system is running, many disturbances can occur and that would lead to the irregularities of the operations of the train units. In order to avoid such events to some extent, it is reasonable to find a plan that is insensitive to these disruptions; that is, the plan is able to cope with relatively small disruptions without structural changes, which is also known as the robustness scheduling, or to react immediately to those disruptions by applying a recovery strategy to the plan defined previously, which is also known as the rescheduling process (see, e.g., [8–10]). Cadarso and Marin [11] formulate a multicommodity flow model for the rolling stock problem in rapid transit networks. Empty movements and shunting operations are considered and the robustness is introduced by selectively avoiding empty train movements and these operations. Cadarso and Marin [12] presents a model to study the robust determining of the best sequence for each rolling stock in the train network. The method is based on an approach in which sequences are designed once the timetable and the rolling stock assignment have been done.

As for the particular EMU circulation plan in China, it is mainly based on a condition of the utilization of the EMU on one railway sector or within several interacting sectors, which is generally viewed as a more reasonable mode for the EMU operation (see, e.g., Zhao et al. [13]). Although the scheduling of the EMU shares something in common with the scheduling of other kinds of transportation tools, especially the scheduling of locomotive in ordinary railway, it has its own characteristics. Locomotives are often scheduled on a fixed sector and could be viewed as an assignment problem which could be solved by Hungarian algorithm. Ahuja et al. [14] formulated the locomotive-scheduling problem as a multicommodity flow problem with side constraints on a weekly space-time network. Each locomotive type defines a commodity in the network. However, the EMUs are scheduled within several linked sectors and some complicated constraints such as the EMU maintenance constraints must be taken into account; therefore, more applicable methods should be developed for solving the construction of the EMU circulation plan. Zhao and Tomii [15] transformed the original problem into the Traveling Salesperson Problem and introduced a probability based local search algorithm, whose key points are about the connection of the trains and the generation of the maintenance arcs. On the basis of that, Miao et al. [16] transformed the original problem into a multiple Traveling Salesperson Problem with replenishment and designed a hierarchical optimization heuristic algorithm. Shi et al. [17] designed a simulated annealing algorithm by introducing the penalty function and 3-opt neighborhood structure, which is based on the circular permutation of all trains. Li et al. [18] introduces the optimized EMU connection graph, based on which the improved particle swarm optimization algorithm is designed for solving the problem.

Through the analysis of the studies concerning EMU circulation plan, it can be seen that the problem could be transformed into some classic optimization problems, and due to the complexity of these problems, most of the existing solution generation methods belong to the range of probability based searching heuristics. The motivation of this paper is to propose an exact algorithm for solving the EMU circulation scheduling problem in high-speed railway network. The algorithm is based on the graph theory and could be able to deal with the problems of practical size within a reasonable time. Furthermore, we propose a heuristic, which shares the characteristics with the existing methods and is based on local search strategy. We also make a comparison between these methods.

The paper is organized as follows. In Section 2, we describe the details of the EMU circulation scheduling problem we study. In Section 3, we introduce the concept of the train connection graph, based on which the model of EMU circulation scheduling in high-speed railway network is constructed. In Section 4, we design the neighbor structure of the solution of the problem and the local search method is illustrated. In Section 5, the exact branch and bound algorithm for solving the problem is outlined, and the details of the algorithm such as the calculation of the approximate lower bound and the branching strategy is illustrated. In Section 6, we report the computational results of the comparison of the proposed methods in a test case.

2. Problem Description

In high-speed railway network, EMUs are utilized on one railway sector or within several interacting sectors; see Figure 1. There are a set of train trips of up and down direction in each sector. The EMUs circulation scheduling problem needs to assign all the tasks of the train trips to a set of EMUs.
In each sector, the train trips of the up direction and the train trips of the down direction are connected at the endpoint stations by certain EMUs. At the stations which are linked with the EMU maintenance base, the EMUs need to have maintenance if either the accumulative running distance or running time reaches the upper bound which is provided in the first level maintenance document of the EMU.

Given a train timetable, the circulation plan could be shown as the format in Figure 2. During a period of one day, we use the term circulation plan fragment to indicate a series of tasks that an EMU should fulfill, which includes several train trips and train connections at stations. In Figure 2, a circulation plan fragment is represented by a polygonal line that links the tasks within one day. Each two train trips could be connected at the station if the duration between the arrival time of one train trip and the departure time of another train trip is not less than the provided minimum duration time. If the EMU needs to have maintenance, then the duration of the connection should also exceed the minimum duration of the maintenance. For example, if an EMU conducts the tasks which are provided in fragment ① of the circulation plan in the first day, it starts from station A and finally arrives at station B after all the tasks have been conducted; it stays at station B in the night. In the second day it starts from station B and conducts the tasks which are provided in fragment ② of the circulation plan. After it finishes serving as train 3102, it has maintenance at station A. Finally, after conducting all the tasks it returns to station A and stays there in the night. In the following days it should conduct the tasks which are provided in another fragment of the circulation plan.

The whole circulation plan is formed by connecting all the fragments through the overnight connection. When the first EMU conducts the fragment ① of the circulation plan, other EMUs conduct the fragments ②, ③, ④, ..., of the circulation plan according to the same rules respectively, then the number of the utilized EMU is known. The maintenance in the circulation is of the basic level and the extensive maintenance which needs much longer time, therefore they are not considered in the circulation plan.

It is worth mentioning that one of the objectives pursued in the scheduling process of the rolling stock circulation would consider maximization of service to the passengers by minimizing the seat shortages. However, under the Chinese railway organization mode the types of trains are viewed as the same and the number of travelers of each sector is dealt with when working out the train operation scheme which provides the train service frequency. Therefore, the consideration of the number of travelers when solving the EMU circulation scheduling problem is out of the scope of this paper.

3. Model of the EMU Circulation Planning

3.1. Train Connection Graph. Given a timetable of several interacting railway sectors, the corresponding EMU connection graph could be constructed correspondingly, which is a useful tool to describe the essentials of the problem of EMU circulation scheduling. The notations with regard to the problem are listed in Notations section.

The parameter \( q(i,j) = 1 \) if and only if there could exist a connection relationship between trains \( i \) and \( j \) under the condition of \( z_{i,d} = z_{i,a} \) and \( c_{i,j} \geq \varepsilon \). \( c_{i,j} \) is calculated as follows:

\[
c(i, j) = \begin{cases} 
\pi_{j,d} - \pi_{j,a}, & \pi_{j,d} - \pi_{j,a} \geq \varepsilon, \\
\pi_{j,d} - \pi_{j,a} + P, & \pi_{j,d} - \pi_{j,a} < \varepsilon. 
\end{cases}
\]

Parameter \( \mu_{i,j} = 1 \) if and only if \( c_{i,j} \geq \varepsilon \) and \( z_{i,d}, z_{i,a} \in Z \). Figure 3 shows the connection graph of a certain group of routes, in which we only display the arcs from the up-bound trains to the down-bound trains. Each row of nodes represents the set of trains that belong to the same route. The solid arcs are the connection arc whose duration time is within one period, while the dotted arcs are the connection arc in which the departing arc and the arriving arc exist. The EMU circulation planning could be formulated as follows:

\[
\min \ Z = \omega_1 \sum_{i \in V} \sum_{j \in V} c(i, j) x_{i,j} + \omega_2 \sum_{i \in V} \sum_{j \in V} y_{i,j} 
\]

s.t. \( \sum_{i \in V} x_{i,j} = 1 \), \( j \in V \)  
(3)

\( \sum_{j \in V} x_{i,j} = 1 \), \( i \in V \)  
(4)

\( x_{i,j} \leq q_{i,j} \), \( i, j \in V \)  
(5)

\( \sum_{i \in [d(e) - a(e')] \cup V} x_{i,j} c_{i,j} + \sum_{i \in [d(e) - a(e')] \cup V} t_i \leq L, \)
(6)

\( e, e' \in W \)

\( \sum_{i \in [d(e) - a(e')] \cup V} x_{i,j} d_{i,j} \leq M, \quad e, e' \in W \)
(7)

\( y_{i,j} \leq \mu_{i,j} x_{i,j}, \quad i, j \in V \)
(8)

\( N(i) + N(j) = |V|, \quad i, j \in V \)
(9)

\( \sum_{i \in V} \sum_{j \in V} y_{i,j} \geq 1 \)
(10)
Objective function (2) consists of two terms, the first term is the total connection time, which is closely related to the number of EMUs needed and this is the main objective of this problem, the second term is the total times of the maintenance. Constraints (3) and (4) impose that each train should be used as the starting node and the ending node of only one arc in $G$; that is, each node of train trip should be visited only once in the circulation plan. Constraints (5) impose that the connection between trains should be reasonable. Constraints (6) and (7) impose that the cumulative running time and running distances of an EMU between two maintenance arcs must not exceed the upper bound of the prescriptive distance and time, where $W = \{(i, j) : y_{i,j} = 1\}$ is the set of the maintenance arcs. The set $\{d(e) \cdot a(e)\}$ contains all the trains between maintenance arcs $e$ and $e'$. Constraints (8) impose that an arc could be selected as the maintenance arc if and only if that arc is used as a connection arc and is able to provide the maintenance. Constraints (9) avoid the occurrence of the subloop. Given a certain arc set connecting all the nodes in $G$ (see Figure 4), when arbitrarily deleting two arcs (represented by the hidden line), then with the starting or ending node of the deleted arc as the root node and searching all the connected nodes using the adjacency matrix of the existing arcs the number of the nodes that can be found is denoted as $N(i)$, which is linked by the dotted line in Figure 4. Obviously, (a) is subjected to constraints (9) while (b) is not subjected to constraints (9). Constraints (10) avoid the existence of none maintenance loop for the circulation plan.

4. Heuristic

4.1. Generation of the Initial Solution. Through the analysis of the constraints of the model, the format of the solution could be written as $s = \{e_1, e_2, \ldots, e_k\}$ which satisfies constraints (3)–(10) and the condition of $a(e_i) = d(e_j)$ for each pair of adjacent arcs $e_i, e_j$ in $s$. The construction of the solution starts from the first selected arc, which could be any one of $E$, and then keeps finding arcs and adds them into the solution. Assuming the current selected arc is $e_i$, then the basic rules used to select the next arc $e_j$ are

$$z_{a(e_i),a} = z_{d(e_j),d},$$

$$\sum_{meW,m \neq a(e_i)} x_{m,d(e_i)} = 0,$$

$$\varphi_{i,j} = 1.$$ (11)

The set of nodes that satisfy condition (11) is represented as $V_{d(e_i)}$. In the process of selecting the connection arcs, the maintenance tasks should be considered simultaneously. Since the existence of constraint (10), then the first arc $e_1$ must be any $e_m \in E$ that satisfies $\mu_m = 1$. Let $I[a(e_i)]$ and $P[a(e_i)]$ be the current accumulative running distance and running time when the EMU arrives at the station $z_{a(e_i),a}$, whose value...
Mathematical Problems in Engineering could be computed using (12), where $\overline{e}$ is the last maintenance arc before $e_i$, and $\overline{e}$ is the next arc after $\overline{e}$:

$$
\begin{align*}
I[a(e_i)] &= \begin{cases} 
    d_{a(\overline{e})} + \sum_{m \in \overline{E} - \overline{e}} d_{d(m)}, & y_i = 0, \\
    d_{a(e_i)}, & y_i = 1,
\end{cases} \\
 p[a(e_i)] &= \begin{cases} 
    t_{a(\overline{e})} + \sum_{m \in \overline{E} - \overline{e}} (c_m + t_{a(m)}), & y_i = 0, \\
    t_{a(e_i)}, & y_i = 1.
\end{cases}
\end{align*}
$$

(12)

If an EMU arrives at station $z_{a(e_i),a}$ after taking on train $a(e_i)$ of current arc $e_i$ in $s$ and the relationship of $d(e_i) \neq a(e_i)$ holds (otherwise, the solution is finished), then due to the existence of the constraints (6) and (7) the next arc to be selected should also to make it be subjected to $I[a(e_i)] \leq L$, $p[a(e_i)] \leq M$. Therefore, $e_j$ is selected as follows:

$$
c(a(e_i),d(e_i)) = \min_{m \in V_{a(e_i)}} \{c_{a(e_i),m} : I[a(m)] \leq L, \ p[a(m)] \leq M, \ \mu_{a(e_i),m} = 0 \}.
$$

(13)

We define the set of arcs that satisfies the condition defined in (13) as the first arc set. If the first arc set does not contain any elements, then the EMU needs maintenance and the alternative arc to be selected is a random one that is subjected to the following constraint:

$$
\min_{m \in V_{a(e_i)}} \{c_{a(e_i),m} : \mu_{a(e_i),m} = 1 \} \leq c(a(e_i),d(e_i))
\leq \max_{m \in V_{a(e_i)}} \{c_{a(e_i),m} : \mu_{a(e_i),m} = 1 \}.
$$

(14)

We then define the set of arcs that satisfy the condition mentioned in (14) as the second arc set. When such an arc is found then continue to implement the above procedure until meeting the node of $a(e_i)$. If the second arc set dose not contain any elements either, then the next arc subjected to all the constraints does not exist in the process of constructing solution $s$. As a result, it is necessary to improve the structure of the already found part of the solution by means of the strategy which will be introduced in next section.

4.2. Local Improvement. In order to change the structure of a feasible solution or part of an unfinished solution, a disturbance should be brought into them. In case that a solution needs to continue to be constructed, which is a basic condition in the generation of different feasible solutions, let $\xi$ be the already existing feasible part of solution $s$; $E_{m,\xi}$ is the set of the maintenance arcs of $\xi$, for a randomly selected element $\overline{e} \in E_{m,\xi}$; change the property $d(\overline{e})$ as a random node $v$ which satisfies the following conditions:

$$
\begin{align*}
\sum_{c_{a(e_i),v}} x_{d(\overline{e}),v} &= 0, \\
\sum_{c_{a(e_i),v}} z_{d(\overline{e}),v} &= z_{v,d}, \\
\mu_{a(e_i),v} &= 1.
\end{align*}
$$

(15)

The elements not belonging to the fragment of $\xi$ are deleted from the temporary solution $s$. Then, set the newly found arc $\overline{e}$ to be the current arc and continue to conduct the solution construction procedure to finish $s$.

In case that a feasible solution $s$ needs to be improved to get another one, the basic idea is to randomly combine the maintenance arcs with connection arcs, and between each pair of adjacent maintenance arcs, the set of connection arcs do not violate all the constraints.

Firstly, the component of $\xi$ derived from $s$ should be determined, which would be kept temporarily in the next generated solution. Since $s$ is a feasible solution, then any part of it is also feasible. Therefore, in order to ensure the randomness of the new solution, $\xi$ is formed through the random selection of two maintenance arcs $e'$ and $\overline{e}$ from $s$ and has the format of $\xi = \{e', \ldots, \overline{e}\} \in s$ which satisfies the condition of $a(e_i) = d(e_i)$ for any adjacent arcs $e_i, e_j$ in $\xi$. Note that $e'$ could not necessarily be $e_1$. Secondly, change the property $d(\overline{e})$ as the one satisfying the conditions described by (15). Then, set the newly found arc $\overline{e}$ to be the current arc and continue to conduct the solution construction procedure to finish $s$.

When a feasible solution is found, function (2) will be used to evaluate the quality of the solution. The whole process of the heuristic used to find the optimal solution of the problem is shown in Figure 5.

![Figure 4: Illustration of the notation $N(j)$.](image-url)
5. An Exact Branch and Bound Algorithm

5.1. Computing the Lower Bound. Given a graph with \( n + 1 \) nodes, a 1-tree is defined to be a set of \( n + 1 \) arcs that span the graph (see, e.g., Fisher [19]). The Degree Constraint Minimum 1-Tree (DCM1-T) is a 1-tree such that the total weight of the set of the arcs is minimum and that the degree of each node does not exceed the limitation of the degree constraint.

It can be seen from the analysis of the model that, under the condition of the utilization of the EMU within several interacting sectors and given the connection graph \( G \), finding the optimal circulation plan which satisfies all constraints is a Degree Constraint Minimum 1-tree Designing Problem. Since DCM1-T is NP-complete (see, e.g., Alexandre and Abilio [20]), there are only methods to find a satisfactory optimal solutions. In order to solve the problem, we draw on the experience of a polynomial approximation algorithm which is introduced in Subhash and Cesar [21]. However, Subhash and Cesar [21] focused on the undirected graph, whereas, according to Figure 3, the multisector train connection graph is a directed, incomplete graph; besides, the constraints of the degree of a node include both the part of in-degree constraints and the part of out-degree constraints. Obviously, the maximum value of the in-degree and the out-degree of each node is 1. All these characteristics about the train connection graph are defined as NCDC rules. Therefore, the algorithm used in this paper is an improved edition of the original one, which could not be applied directly.

(1) The Construction of the Basic Connection Tree. If none of the arcs associated with a node has been selected in \( G \), then we call the node an idle node. The construction process starts with an initial set of nonidle nodes \( C \), which could include arbitrary number of nodes. For example, we could select a certain arc of overnight connection in \( G \) at the beginning and...
then $C$ contains two elements, namely, the departure train trip and the arrival train trip of the arc. Let $U$ be the initial set of idle nodes. The construction process constitutes continuously selecting arcs which connect the nodes in $C$ with the nodes in $U$. After an arc is selected the associated node in $U$ is transformed into the set $C$.

For each node $i \in U$, its in-degree $d_i^-$ and the out-degree $d_i^+$ are initialized to 0. Let $vr(i)$ denote the node in $C$ such that there is an arc from $i$ to it and that
\[ c(i, vr(i)) = \min_{j \in C} \{c(i, j)\}. \]

Let $vt(i)$ denote the node in $C$ such that there is an arc from it to node $i$ and that
\[ c(vt(i), i) = \min_{j \in C} \{c(j, i)\}. \]

Let $e_i^*$ denote the arc whose cost satisfies the condition of $c_{e_i^*} = \min_{e \in U} \{c(i, vr(i)), c(vt(i), i)\}$. If $e_i^*$ has the form $(i^*, vr(i^*))$, then the out-degree of node $i^*$ is updated as $d_{i^*}^- = d_{i^*}^- + 1$, and the in-degree of the node $vr(i^*)$ is updated as $d_{vr(i^*)}^+ = d_{vr(i^*)}^+ + 1$. If $d_{i^*}^- > 1$ or $d_{vr(i^*)}^+ > 1$, then the arc $e_i^*$ could not be selected and the node $vt(i)$ is modified as the node $vt(i')$ such that
\[ c(vt(i'), i) = \min_{j \in C \setminus \{i\}} \{c(j, i)\}. \]

Similarly, If $e_i^*$ has the form $(vt(i^*), i^*)$, then the out-degree of the node $vt(i^*)$ is updated as $d_{vt(i^*)}^- = d_{vt(i^*)}^- + 1$, and the in-degree of the node $i^*$ is updated as $d_{i^*}^+ = d_{i^*}^+ + 1$. If $d_{vt(i^*)}^- > 1$ or $d_{i^*}^+ > 1$, then the arc $e_i^*$ could not be selected either and the node $vt(i)$ is modified as the node $vt(i')$ such that
\[ c(vt(i'), i) = \min_{j \in C \setminus \{i\}} \{c(j, i)\}. \]

The arc $e_i^*$ is reselected according to the new property of the $c(i, vr(i'))$ and $c(vt(i'), i)$. After selecting the arc $e_i^*$, the set of $C$ is updated as $C = C \cup \{i^*\}$, and the node $i^*$ is removed from the set $U$ simultaneously. In addition, as the components of both set $C$ and set $U$ are changed, therefore $vr(i)$ and $vt(i)$ for each $i \in U$ need to update again; specifically, if $c(i, vr(i)) < c(i, i^*)$, set $vr(i) = i^*$, and if $c(vt(i), i) < c(i^*, i)$, set $vt(i) = i^*$. If all the nodes are checked with the procedure mentioned before, then the basic connection tree (BCT) is gotten. However, due to the incompleteness of the train connection graph, there are still elements in the set $U$, and the supplementary measures need to be applied to continue the construction of the DCMI-T.

(2) Finishing the Construction of 1-Tree. If set $U$ is still not empty after the BCT is constructed then it means no nodes could be found in the set of $U$ under the NCDC rules. Therefore, the remaining nodes in $U$ should first be incorporated into the BCT and the BCT is transformed into an ordinary tree; then, some exchanging method should be applied to the ordinary tree and transform it into a degree constraint minimum 1-tree which contains all the nodes in $G$.

In order to transform as many nodes as possible from set $U$ into set $C$, repeat the procedure introduced before and do not consider the in-degree or out-degree constraints of the nodes. During this process, for $\forall i \in U$, the corresponding $vr(i)$ or $vt(i)$ should not belong to the initial $C$, which is the set of the nodes before any other node being added into it. When all the remaining elements in $U$ are transformed to the set $C$, the structure of BCT could not be subjected to constraints (3) and (4).

Under the NCDC rules, if an arc connecting nodes $i$ and $j$ could not be selected and the node $i$ and $j$, which could be the form of $(i, j)$ or $(j, i)$. We then introduce two methods for exchanging the tree's structure.

The dual exchanging method for the ordinary tree (exm-1): let $ST(i)$ denote the full subtree of BCT rooted at node $i$, $SV(i)$ denote the set of trains which are connected with the node $i$ of the BCT through the already selected arcs. Then, the arcs could be exchanged as follows:
\[ [m,n] \Rightarrow [k, r(k) \lor t(k)], \quad \forall d_{k \in C}, d_{k \in C} > 1. \]

In (20), $[m,n]$ is the arc which is to be added into the BCT and $[k, r(k) \lor t(k)]$ is the arc which is to be removed from the BCT, $r(k) \lor t(k)$ is some node which is connected with node $k$ through an arc and they are determined according to the following conditions:
\[ c[k, r(k) \lor t(k)] = \min_{i \in SV_k} c[k, i], \]
\[ m \in ST(k), \]
\[ n \in ST(r(k)) \cup ST(t(k)), \]
\[ m \equiv n. \]

The second exchanging method for the 1-tree (exm-2): after applying exm-1 for the BCT and adding the edge $[i, j]$ to it, where $i$ and $j$ are the nodes whose in-degree and out-degree are not equal, then the structure of the modified BCT becomes 1-tree. The exchange is as follows:
\[ [i,j] + [j,n] \Rightarrow [m,n], \quad \forall i, j \in U : i \equiv j. \]

In (22), the arc $[m,n]$ is the edge which is to be deleted from the BCT, and the other three edges are to be added into the BCT. These edges could be determined through the following condition:
\[ c[i,m] + c[i,j] + c[j,n] - c[m,n] = \min_{[h,k] \in \text{BCT}} \{c[i,h] + c[i,j] + c[j,k] - c[h,k]\}. \]
Initialize the set of $C$, $U$, and $\nu r(i)$, $\nu t(i)$ for each $i \in U$, $m = |U|$

**Figure 6:** Flowchart of the computation of DCM1-T.

5.2. The Branching Strategy. Since the DCM1-T might violate constraints (6) and (7), then we need first to fix part of the DCM-T in the circulation plan and build some branches to, respectively, determine some values of the maintenance variables; then, the DCM1-T are calculated for each branch and new parts of the DCM-T are fixed. Repeat this process until all the parts of the DCM-T are feasible with respect to the constraints (6) and (7). After each branch is calculated, the best feasible circulation plan could be obtained simultaneously.

Let $EN$ denote the set of arcs that are included when recomputing the DCM1-T. Let $EX$ denote the set of arcs that are excluded when recomputing the DCM1-T. At the
beginning, there is only one element in \( EN \) which is the basic maintenance arc in \( DCMI-T \) while the \( EX = \emptyset \). At each level of branch, if the objective value of the \( DCMI-T \) computed based on the subset \( (EN, EX) \) is smaller than the current upper bound, then the circulation fragment \( \epsilon = (e_k, e_{k+1}, \ldots, e_{k+n}) \) is fixed, which satisfies that \( d(e_i) = a(e_{i+1}) \) for each pair of adjacent arcs \( e_i, e_{i+1} \in \epsilon \) in the \( DCMI-T \) where \( e_i \) means the adjacent arc of the last arc in \( EN, k = |EN| + 1 \), and \( n \) is an integer used to limit the number of the arcs which would violate the maintenance constraints. Then, the set of maintenance arcs \( EM = \{\omega_1, \omega_2, \ldots, \omega_\lambda\} \) is constructed, in which the arrival train of each arc is \( d(e_{\lambda+1}) \) and \( i \) is the parameter indicating the level of the branching process. \( \lambda \) means the number of the elements of the set \( EM \), and if the number of arcs in \( G \) that could be selected into the set \( EM \) is bigger than \( \lambda \), then the actual elements held in \( EM \) are the first \( \lambda \) lower cost arcs; otherwise, \( EM \) contains all the arcs. Based on \( EM \) and \( EN \), the list of the new subsets could be constructed as follows:

\[
EN_{j(0)} = EN_{f(0)} \cup \{\omega_{j(0)}\},
\]

\[
EX_{j(0)} = EX_{f(1)} \cup \{\omega_1, \omega_{j−1(0)}, \omega_{j+1(0)}, \ldots, \omega_{\lambda(0)}\},
\]

For a subset \( (EN_{j(0)}, EX_{j(0)}) \) which is also the branch \( j \) at level \( i \), a \( DCMI-T \) which includes the arcs in \( EN_{j(0)} \) and excludes the arcs in \( EX_{j(0)} \) is calculated and its objective value is obtained as a lower bound. In the whole branching process, we apply strategy like the one mentioned in [21]: always branch to the left subnode and at each node, keep a “left” bound as the current lower bound of the subset \( (EN_{j(0)}, EX_{j(0)}) \), and a “right” bound as the lower bound of the subset \( (EN_{j−1(0)}, EX_{j−1(0)}) \), as shown in Figure 7.

At beginning, let the time of iteration \( d = 0 \), \( f(C) \) represent the value of the objective function for a given EMU circulation \( C \), \( R(d) \) and \( L(d) \) are the right and the left bounds of a subset at level \( i \) of the branching tree, respectively. Then, the algorithm could be summarized as follows.

**Step 1.** Generate an EMU circulation plan \( TC \) which considers the maintenance using an arbitrary heuristic; set the upper bound \( Z = f(TC) \). Initialize set \( EN_{j(0)} \) and set \( EX_{j(0)} \). Implement the lower bound calculation algorithm to generate the circulation plan \( C_0 \), which is a \( DCMI-T \). Set the left bound \( L(d) = f(C_0) \), the right bound \( L(d) = \infty \), and \( d = d + 1 \).

**Step 2.** Determine the subset \( (EN_{j(d)}, EX_{j(d)}) \), based on which we implement the lower bound calculation algorithm to generate the circulation plan \( C_{\lambda(d)} \). Set the left bound \( L(d) = f(C_{\lambda(d)}) \), \( m(d) = \lambda(d) − 1 \). Determine the subset \( (EN_{md(d)}, EX_{md(d)}) \). Implement the lower bound calculation algorithm to generate the circulation plan \( C_{md(d)} \), set the right bound \( R(d) = f(C_{md(d)}) \).

**Step 3.** If \( L(d) < Z \), go to Step 4; otherwise, go to Step 5.

**Step 4.** If the circulation plan corresponding to the left bound \( L(d) \) does not violate the maintenance constraints, then a feasible solution is found. Update the upper bound. Set \( Z = L(d) \), and go to Step 5; otherwise, set \( d = d + 1 \) and go to Step 2.

**Step 5.** If \( R(d) < Z \), then update the left bound and set \( L(d) = R(d) \); go to Step 6; otherwise, go directly to Step 6.

**Step 6.** Set \( m(d) = m(d) − 1 \); if \( m(d) = 0 \), go to Step 7. Otherwise, if \( m(d) > 1 \), determine the subset \( (EN_{md(d)}, EX_{md(d)}) \), implement the lower bound calculation algorithm to generate the circulation plan \( C_{md(d)} \), and update the right bound \( R(d) = f(C_{md(d)}) \). Otherwise, set \( R(d) = \infty \). When the \( R(d) \) is determined, if \( R(d) > Z \), repeat Step 6. Otherwise, go to Step 3.

**Step 7.** Go back to the previous level; set \( d = d − 1 \); if \( d > 1 \), go to Step 5; otherwise, stop.

It is worth mentioning that the solution made of the combination of the branch constructed with the lowest cost arc in \( EM \) of each level under the condition that these branches are able to formalize a feasible solution could be viewed as the result of another greedy strategy, namely, the usually mentioned Tight Connection, which is also a way to quickly generate a feasible EMU circulation plan.

**5.3. Improvement of the Heuristic.** During the generation of a feasible solution introduced in Section 4, due to the incompleteness of the train connection graph there is an interactive process between the selecting of arcs to construct the solution and the exchanging of the arcs for the part of the solution which has been already obtained in order to complete the generation of a feasible solution. A lot of detection procedures for the arc exchanging may appear until a whole feasible solution is found in the interactive process; therefore, it accounts for the majority of the running time of the algorithm.

However, according to the illustration of Section 4.2, let \( \xi = [\xi, \ldots, \xi] \in \xi \) be the randomly selected part of either the unfinished solution or a feasible solution \( s \) when applying the local improvement, after property \( d(\xi) \) is changed and the newly found arc \( \xi \) is set to be the current arc, if the branch strategy with a limitation on the levels of the branches is
incorporated into the remaining process of the construction of the feasible solution, then a more reliable way could be obtained to generate the feasible solution faster and with a probably better objective value.

6. Numerical Experiments

Without loss of generality, we test our method on the set of trains on a corridor, in which we use the Jing-Hu railway line between Beijing and Shanghai as reference to setting the parameters when working out the EMU circulation plan. Figure 8 shows which segments a sector contains. This example has general features in common with other high-speed railway networks that consist of several different train sectors and these sectors interact with each other.

Table 1 shows the data of the train operation scheme used in the following experiments referring to the corridor. $T_1$ ($T_2$) is the set of high-speed trains (medium-speed trains) in each sector and is twice as much as the trains of each direction of the sector. $d$ is the length of each sector. The travel time of each train is different and could be obtained from the timetable.

Stations BJN, QFD, BBN, and SHHQ are able to fulfill the task of first grade maintenance. Other basic parameters are as follows: $L = 2880$ min, $M = 4000$ km, $\varepsilon = 15$ min [18], and $\tilde{\varepsilon} = 120$ min [16]. Given a train timetable which provides the condition for EMUs operated within these sectors, the number of arcs in the train connection graph is 10228. The circulation plan computed using the branch and bound algorithm is as shown in Figure 9.

In order to show further computational results, we test the method on other instances of the problem in which the timetable of the trains on the corridor is randomly generated and other parameters remain the same. Figures 10–13 compare the convergence curves generated by the heuristic and the branch and bound algorithm on four different instances of the problem. And for the heuristic we run it twice and get two curves. The modified heuristic is also tested.

It can be seen from these figures that the time needed by the branch and bound algorithm to generate an optimal solution is 2 min or so in the test case, whose number of trains corresponds to the scale of a majority of practical cases, and the time could be viewed as acceptable. At the same time, the convergence time needed by the heuristic reflects a fluctuation to some extent. In addition, either the value of the total connection cost of the heuristic is bigger than that of the branch and bound or the convergence time of the former algorithm is longer than that of the latter algorithm. However, as the heuristic is modified, the corresponding solutions appear to be better to some extent and the convergence time is shorter compared with the initial heuristic.

We also compare among the heuristic, branch and bound, and tight connection methods mentioned in this paper based on 20 different instances of the problem. The results are listed in Table 2. The second column is the total connection time of the initial DCM1-T, which is a circulation plan not considering the maintenance, and columns 2–8 are the total connection cost (TCC) and the times of maintenance (MT) of

| Sector | From | To | $|T_1|$ | $|T_2|$ | $d$ (km) |
|--------|------|----|-------|-------|----------|
| 1      | BJN  | SHHQ| 32 x 2| 10 x 2 | 1680     |
| 2      | BBN  | SHHQ| 19 x 2| 6 x 2  | 600      |
| 3      | BJN  | QFD | 19 x 2| 6 x 2  | 695      |

The three algorithms. Among them, the TTC of the heuristic is the value of the best solutions generated by the heuristic.

There are a few observations to be made from this table. First, there is not a clear proportional relation between the total connection cost and the maintenance times; that is, the maintenance times of a solution with a larger total connection cost might not be more than the maintenance times of a solution with a smaller total connection cost. This is because, under a given structure of the train timetable, the more times of EMUs staying at a station at night the larger the accumulative running time of them after a maintenance, while the accumulative running distance may not reach the limitation. Second, almost all solutions (associated with the notation “*”) computed by the branch and bound method reach the lower bound of the problem which is the total cost of the initial DCM1-T not considering the maintenance constraints except for two solutions, but the total connection cost of these two solutions is very close to their lower bound and smaller than that of the solutions computed by the other two methods too. Furthermore, there are even some solutions (associated with the notation “**”) computed by the branch and bound method whose total connection cost is smaller than the one of the initial DCM1-T. This is because the outputs obtained by the lower bound algorithm and the circulation plans obtained by branch and bound all belong to degree constraint 1-tree. Since the determining of the DCM1-T is NP-hard and the DCM1-T calculating algorithm could only find an approximate solution whose objective value only provides an approximate lower bound on the optimal circulation plan as mentioned before, therefore through the continuous applying of the DCM1-T calculating algorithm in the process of the branch and bound there would exist the probability of finding a degree constraint 1-tree whose objective is better than the one found by directly applying the DCM1-T calculating algorithm once. Third, it can be inferred from the table that for, the same case, even a slight adjustment of the train timetable may produce completely different optimal EMU circulation plans with different total connection costs. Therefore, the proposed branch and bound
method could be used for the problem of the optimization of network train timetable, which is incorporated with the EMU circulation scheduling problem. Finally, by comparison, the mean value of the total connection costs of the solutions generated by the tight connection method is \((70521.5 - 65506.1)/65506.1 = 7.7\%\) more than the one of the solutions generated by the heuristic and \((70521.5 - 59002.1)/59002.1 = 19.5\%\) more than the one of the solutions generated by the branch and bound method, while mean value of the total connection costs of the solutions generated by the heuristic is \((65506.1 - 59002.1)/59002.1 = 11.0\%\) more than the one of the solutions generated by the branch and bound method.

7. Conclusion

Based on the introduction of the train connection graph the of EMU circulation scheduling model is constructed. Through the analyzing of the features of the structure of the
feasible EMU circulation plan, this paper mainly designs an exact branch and bound algorithm for solving the problem.

We first put the initial problem down to a graph designing problem; that is, the degree constraint minimum 1-tree problem. Due to the complexity of the problem we use a polynomial approximation algorithm to compute the DCM1-T which does not consider the maintenance constraints and provides an approximately lower bound of the problem. This algorithm is modified in order to adjust to the characteristics of the train connection graph. Then, a branch strategy is proposed to cope with the maintenance constraints and to generate an optimal circulation plan.

Through the analysis of the computational results of the numerical experiments of the comparison of the proposed branch and bound method with the heuristics which shares the features with the existing methods, it can be seen that the running time needed by branch and bound method is more reasonable when dealing with the instance of the problem.
with the scale of most practical cases, the objective value of the solutions generated by the branch and bound method is better than the one of the solutions generated by other heuristics, and the branch and bound method could be extended to other problems concerning the optimization of the train timetable.

**Notations**

- **G:** The EMU connection graph
- **V:** Set of nodes of the connection graph, means all train trips involved in the planning time horizon
- **i, j:** Indexes of node, \( i, j \in V \)
- **\( z_{i,d} \), \( z_{i,a} \):** the first station the train \( i \) departs from and the final station the train arrives at along its route, respectively
- **\( \pi_{i,d}, \pi_{i,a} \):** the departure time of the train \( i \) at the station \( z_{i,d} \) and the arrival time of the train \( i \) at the station \( z_{i,a} \)
- **d, t:** distance between station \( z_{i,d} \) to station \( z_{i,a} \) and time consumed of the train \( i \) running from station \( z_{i,d} \) to station \( z_{i,a} \), respectively
- **E:** Set of arcs of the connection graph, means all the potential connection relationship between train trips
- **\( q_{i,j} \):** Binary parameter indicating whether there exist an arc from train trip \( i \) to \( j \)
- **\( e, e' \), (i, j):** Indexes of the arc \( e, e' \), \( (i, j) \in E \)
- **c(i, j):** Connection cost of the arc \( (i, j) \) which is measured as the duration time of the arc
- **\( \varepsilon \):** Minimum duration time necessary for EMU to connect trains at stations
- **\( \mu_{i,j} \):** Binary parameter indicating whether the arc \( (i, j) \) could be used as an maintenance arc
- **\( \bar{\varepsilon} \):** Minimum duration time of the maintenance for EMU at stations
- **\( \Xi \):** Set of stations which are able to fulfill the maintenance task
- **L, T:** Upper bounds of accumulative running distance and time for a EMU under the first level maintenance, respectively
- **\( a(e), d(e) \):** Arrival train and departure train the of the arc \( e \), respectively
- **\( \omega_1, \omega_2 \):** Weights of the terms in the objective function
- **P:** the time of the period of one day.

**Conflict of Interests**

The authors declare that there is no conflict of interests regarding the publication of this paper.

**Acknowledgment**

This work was financially supported by The Fundamental Research Funds for the Central Universities (no. 2014JBZ008).

**References**


