Modeling and optimization for train-set utilization problem using a two-stage approach

Yu Zhou¹, Leishan Zhou²*, Yun Wang³

1. Yu Zhou, Ph.D., State Key Lab of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China. Tel: 152-1057-5587, E-mail: 14114231@bjtu.edu.cn.

2. Leishan Zhou, Ph.D. (Corresponding Author), Professor, State Key Lab of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China. Tel: 136-0101-2782, E-mail: lshzhou@bjtu.edu.cn.

3. Yun Wang, Ph.D. MOE Key Laboratory for Urban Transportation Complex System Theory and Technology, School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China. Tel: 152-1057-6691, E-mail: 13114248@bjtu.edu.cn.

- Number of words: 4957
- Number of figures and tables: 10
ABSTRACT
The utilization efficiency of train-sets (vehicles of HSR) is one of the most important factors influencing the transport capacity of HSR. Therefore, it is imperative to propose a theoretical method to guide efficient train-set utilization in the practical application. In this paper, an integer programming model is established to obtain a high-quality train-set utilization plan. A two-stage approach and a concept of segment are proposed to solve the model. The train graph in the Beijing-Tianjin passenger dedicated line is adopted to illustrate the model application, and a high-quality train-set utilization plan is obtained. Moreover, the comparison between the two-stage approach and the Ant Colony Algorithm (ACA) is conducted to evaluate advancements of the proposed method. The results show that the train-set utilization plan obtained by the two-stage approach is much better in the train-set utilization efficiency as well as the computation time.

KEYWORDS
High-speed-railway (HSR); Passenger railway transport; Train-set utilization plan; Train-set maintenance
1. Introduction

During the past few decades, extensive research has been carried out worldwide to optimize train-set utilization plans. The train-set utilization problem can be regarded as a typical NP hard problem, and it cannot be solved without considering the limitation of available train-sets.

During the past few decades, extensive research has been carried out worldwide to optimize train-set utilization plans. The train-set utilization problem can be regarded as a typical NP hard problem, and it cannot be solved without considering the limitation of available train-sets.

During the past few decades, extensive research has been carried out worldwide to optimize train-set utilization plans. The train-set utilization problem can be regarded as a typical NP hard problem, and it cannot be solved without considering the limitation of available train-sets.
directly by ready-made software. Characterized by easy-to-understand and fast, a large
number of heuristic algorithms have been applied to solve the train-set utilization problem. 
Yang et al. (15) built a train-set connection network model aiming to maximize the train-set
utilization efficiency, and designed a genetic algorithm to solve the model. Chen et al. (16)
proposed a multi-objective optimization model to maximize the equilibrium degree of
train-set utilization, considering the number of using train-sets, and a simulated annealing
algorithm was designed to solve the model. Hong et al. (17) put forward an integer
programming model considering the maintenance and train graph constraints, and a heuristic
approach was applied to solve this problem. In the past several years, some scholars presented
that the train-set utilization problem could be seen as the TSP problem with some special
constraints, which brings new methodologies to solve the train-set utilization problem. Zhang
(18) built a train-set utilization model in a railway network by taking both the running time
constraint and running distances constraint into consideration, and an ant colony algorithm
was designed to solve the model. Zhou (19) proposed an integer programming model to
obtain the train-set utilization plan, considering the constraint of Level 1 maintenance
standard, and an improved ant colony algorithm was designed to solve the model.

Although there is a comprehensive body of literature on the train-set utilization
problem, no results of previous research can be directly used into the practical production.
There are two major reasons. First, limited studies take the maintenance constraints into
consideration, which actually is a key factor influencing the compilation of the train-set
utilization plan. Second, even if the maintenance is considered, the solutions of models are
obtained by applying heuristic algorithms, which is unstable and inaccurate. Therefore, to
realize the objective that the train-set utilization plan can be generated by computers
automatically and used in practice, an integer programming model is built to maximize the
train-set utilization efficiency, considering the maintenance constraints. What makes this
study unique is that a two-stage method is proposed in the model solving process to obtain a
unique solution. With this method, the train-set utilization plan can be obtained by a
computer-generated system.

2. PROBLEM DESCRIPTION

The train-set utilization plan is drawn up to identify work arrangements of train-sets
according to a given train graph as well as the rules and regulations of maintenance. For
train-sets, except for some necessary operations in railway stations, the main work includes
two parts: one is to complete all the trip tasks in the given train graph; and the other is to be
carried out maintenances. The train-set utilization plan should figure out the following
aspects: the number of needed train-sets; 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. Different train-set utilization plans can be made to complete
the trip tasks in a given train graph. Therefore, in this study, the optimization problem of the
train-set utilization plan is set to complete all the trip tasks in a given train graph by utilizing
less train-sets on the premise of satisfying the requirements of the maintenance.

Figure 1 illustrates the framework of the train-set utilization optimization problem.
Firstly, some basic information should be prepared as the input parameters, including: the
train graph, the rules and regulations of maintenance and station working. Secondly, in order
to model the train-set utilization plan, a train-set utilization network is represented by a
directed graph. Thirdly, an integer programming model is established to minimize the total
connecting time, considering the spatial constraints, the time constraints, the maintenance constraints, and the unicity constrains. Finally, a two-stage approach is designed to solve the problem, and an optimal train-set utilization plan is obtained as the output.

**Input**

- Train diagram (Trips, Station, Time)
- Station Working (Operation time, Preparation time)
- Rules of Maintenance (Distance standard, Time standard)

**Utilization Network**

- Trip node
- Node
- Arc
- Connection arc
- Maintenance arc
- Arc after maintenance

**Model and Algorithm**

- Model
  - Goal
    - Minimum Connecting Time
  - Constraints
    - Spatial constraints
    - Time constraints
    - Maintenance constraints
    - Unicity constraints

- Algorithm
  - Two-stage Algorithm
    - Segment Generation
    - Segment Connection

**Output**

- Train-set utilization plan

**FIGURE 1** Framework of the train-set utilization optimization problem.
Figure 2 illustrates the relationship between the train graph and the train-set utilization plan. Figure 2(a) is a simple sample of a given train graph, including three stations and eight trips. After adding the train-sets’ connection relationships and maintenance information, Figure 2(b) represents the train-set utilization plan. In this sample, to complete the eight transportation tasks, it needs at least three train-sets. Particularly, Trip G82 and G89 are undertaken by one train-set; Trip G84, G87, and G88 are undertaken by one train-set; and Trip G83, G86, and G91 are undertaken by one train-set. The train-set undertaking Trip G83 should be maintained at the inspection and repair depot near Station B after completing the task of Trip G86.

3. TRAIN-SET UTILIZATION PLAN MODELING

3.1 Train-Set Utilization Network Representation
In order to model the train-set utilization plan, a train-set utilization network is represented by a directed graph \( G(V, E) \). The node set \( V \) is a union of subsets \( V_T \) and \( V_M \), which represent trips and maintenances, respectively. Each node \( i \in V_T \) is defined as a tuple \((n_i, s_i^d, s_i^a, t_i^d, t_i^a, d_i, t_i)\), where \( n_i \) indicates the train number, \( s_i^d \) indicates the departure station, \( s_i^a \) indicates the arrival station, \( t_i^d \) indicates the departure time, \( t_i^a \) indicates the arrival time, \( d_i \) indicates the running distance, and \( t_i \) indicates the running time. Each node \( k \in V_M \) is defined as a tuple \((\varphi_k, \tau_k^a)\), where \( \varphi_k \) indicates the inspection and repair depot, \( \tau_k^a \) indicates the time train-set departing from the inspection and repair depot to arrival station after maintenance. There are three kinds of arcs in the train-set utilization network, including connection arcs (i.e., trip→trip), maintenance arcs (i.e., trip→maintenance) and arcs after maintenance (i.e., maintenance→trip). Therefore, the arc set \( E \) consists of three disjoint sets \( E_C \), \( E_Q \) and \( E_H \), where \( E_C \) represents directional arcs between two trip nodes, i.e., connection arcs, \( E_Q \) represents directional arcs from trip node to maintenance node, i.e., maintenance arcs and \( E_H \) represents directional arcs from maintenance node to trip node, i.e., arcs after maintenance.

The weight of connection arcs \( \omega_i^C \) means the connection time between Trip \( i \) and Trip \( j \). Generally, when the arrival station of Trip \( i \), is the different with the departure
station of Trip \( j \), the connection time should be defined as infinite. It means that the two trips cannot be connected due to the huge waste of train-set empty running. As shown in Figure 3, when two trips can be connected, the values of \( \omega_{ij}^C \) is determined by the sequential order of the two trips’ departure time and arrival time. If the departure time of Trip \( j \) is behind the arrival time of Trip \( i \), the \( \omega_{ij}^C \) is equal to \( t_j^d - t_i^a \), else, the \( \omega_{ij}^C \) is equal to \( 1440 + t_j^d - t_i^a \), i.e., the train-set will undertake Trip \( j \) the next day. Based on above analyses, the values of \( \omega_{ij}^C \) can be calculated as Eq. (1).

The weight of maintenance arcs \( \omega_{ik}^O \) means the duration from the moment train-set arriving at its arrival station, \( t_i^a \), to the moment train-set returning back at its arrival station after maintenance, \( \tau_i^a \). As shown in Figure 4, when \( \tau_i^a - t_i^a \geq 0 \), the value of \( \omega_{ik}^O \) is equal to \( \tau_i^a - t_i^a \). When the \( \tau_i^a - t_i^a \leq 0 \), it means that the train-set can only undertake Trip \( j \) the next day, and the value of \( \omega_{ik}^O \) is equal to \( 1440 + \tau_i^a - t_i^a \). Therefore, the values of \( \omega_{ik}^O \) can be calculated as Eq. (2).
The weight of arcs after maintenance $\omega_{ki}^H$ means the duration from the moment train-set returning back at its arrival station after maintenance, $t_k^d$, to the moment train-set departing from the station, $t_i^d$. As shown in Figure 5, when $t_k^d - t_i^d \geq 0$, the value of $\omega_{ki}^H$ is equal to $t_k^d - t_i^d$. When the $t_k^d - t_i^d \leq 0$, it means that the train-set can only undertake Trip $i$ the next day, and the value of $\omega_{ki}^H$ is equal to $1440 + t_k^d - t_i^d$. Therefore, the values of $\omega_{ki}^H$ can be calculated as Eq. (3).
3.2 Analyses of Constraints

To undertake a trip task, the train-set should be satisfied four specific requirements, including the spatial constraint, the time constraint, the maintenance constraint and the unicity constraint. The detail analyses are as follows:

3.2.1 Spatial constraints

As previously mentioned, an empty running train-set needs to be dispatched to connect two trips, when the departure station of the latter trip is different from the arrival station of its former trip, which will cause huge waste. Therefore, as shown in Eq. (4), when a train-set undertake two adjacent trips, the departure station of the latter trip, $s_j^d$, is restricted to be the same with the arrival station of the former trip, $s_i^a$, where $i \leq j$.

$$s_i^a = s_j^d \quad i \leq j$$  \hspace{1cm} (4)

3.2.2 The time constraints

The time constraints that the duration of connection arcs, maintenance arcs and arcs after maintenance should respectively satisfy the requirement of the minimal necessary working
procedure duration between two tasks, respectively.

When a train-set arrives at the station after completing the trip task, it must be carried out some preparation work before departing for the next trip, such as cleaning, pollution discharge, and so on. The duration needed for such preparation work is defined as operation time, $T_o$. Therefore, to ensure that two trips can be undertaken by one train-set, the connection time $\omega_c^k$ of the two trips should be longer than the operation time $T_o$, as shown in Eq. (5).

$$\omega_c^k \geq T_o$$  \hspace{1cm} (5)

When a train-set completes a trip task and reaches to the maintenance standard, it must be maintained in the inspection and repair depot. As shown in Figure 6, four essential procedures will be followed by the train-set to complete the maintenance task. Table 1 lists the detail information about the four procedures along with their duration definitions.

![Diagram](image)

**FIGURE 6** Relationship between the process duration and the weight of maintenance arcs.

**TABLE 1 Four Essential Procedures**

<table>
<thead>
<tr>
<th>Procedures</th>
<th>Duration</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Some basic inspection will be carried out after the train-set arriving at the station.</td>
<td>Inspection time</td>
<td>$T_a$</td>
</tr>
<tr>
<td>(2) Train-set will depart from the arriving station to the inspection and repair depot.</td>
<td>Running time</td>
<td>$T_b$</td>
</tr>
<tr>
<td>(3) Maintenance will be carried out in the inspection and repair depot</td>
<td>Maintenance time</td>
<td>$T_m$</td>
</tr>
<tr>
<td>(4) Train-set will return back from the inspection and repair depot to the station.</td>
<td>Running time</td>
<td>$T_b$</td>
</tr>
<tr>
<td>Total</td>
<td>Duration of the process</td>
<td>$T_c = T_a + 2T_b + T_m$</td>
</tr>
</tbody>
</table>

To ensure a trip node can be connected to a maintenance node, there should be enough time for the train-set to go through these procedures. Therefore, Eq. (6) restrains that the
weight of maintenance arc $\omega_{ik}^l$ should be larger than the duration of process $T_c$.

$$\omega_{ik}^l \geq T_c$$  \hspace{1cm} (6)

Once completing the maintenance task, train-sets can be utilized to undertake the trip tasks. However, before departing, a series of preparation work should be carried out to ensure the train-set working normally. Therefore, to ensure a maintenance node can be connected to a trip node, the weight value of arc after maintenance $\omega_{ik}^H$ should be larger than the duration of preparation work $T_c$, as shown in Eq. (7).

$$\omega_{ik}^H \geq T_c$$  \hspace{1cm} (7)

3.2.3 The maintenance constraints

To ensure the safe operation, the train-sets must be strictly complied with the maintenance rules and standards. In China, the train-sets' maintenance standards are divided into five levels according to accumulated operation time and mileages. In this paper, it is assumed that the train-set utilization plan is drawn up by treating one day as a cycle. Generally, the operation time standard of Standard Level 2 to Standard Level 5 is more than a month. Therefore, in this paper, only the Standard Level 1 needs to be considered. The Standard Level 1 requires that when the accumulated operation time reaches up to 48 hours or the accumulated operation mileage reaches up to 4,000 km, the train-set must be maintained. Additionally, in practical applications, 10% fluctuation compared with the standards are allowable (e.g., Train-sets should be maintained when accumulated operation mileages are within the range of 3600 km to 4400 km).

When undertaking several trips, the accumulated operation time and mileages of the train-set are respectively $\sum_{j=1}^{n} T_i$ and $\sum_{i=1}^{n} d_i$, where $n$ is the number of undertaking trips. Let $\Delta S_m$ and $\Delta S_t$ represent the acceptable fluctuation ranges of operation time and mileages, respectively. Therefore, the accumulated operation time and mileages of each train-set should be within the range of $[S_t - \Delta S_t, S_t + \Delta S_t]$ and $[S_m - \Delta S_m, S_m + \Delta S_m]$, respectively.

$$S_t - \Delta S_t \leq \sum_{j=1}^{n} T_i \leq S_t + \Delta S_t$$  \hspace{1cm} (8)

$$S_m - \Delta S_m \leq \sum_{i=1}^{n} d_i \leq S_m + \Delta S_m$$  \hspace{1cm} (9)

3.2.4 The uniqueness constraints

The uniqueness involves two aspects. One is that each train-set, after completing a trip task, should either undertake another trip or be carried out maintenance. The other one is that for each trip in the given train graph, only one train-set can be allocated to complete the task.

The binary variable $x_{ij}^C$ is defined to describe whether the trip node $i$ and trip node $j$ are connected by the connection arc. The binary variable $x_{ik}^Q$ is defined to describe whether the trip node $i$ and maintenance node $k$ is connected by the maintenance arc. The binary variable $x_{ik}^H$ is defined to describe whether the maintenance node $k$ and trip node $i$ are connected by the arc after maintenance. If the two nodes are connected by arcs, the value
of the binary variable is one, otherwise zero, as shown in Eq. (10) to Eq. (12).

\[
x_{ij}^C = \begin{cases} 
1 & \text{trip node } i \text{ and trip node } j \text{ is connected} \\
0 & \text{by a connection arc} 
\end{cases} 
\] (10)

\[
x_{ik}^Q = \begin{cases} 
1 & \text{trip node } i \text{ and ma intenance node } k \text{ is connected} \\
0 & \text{by a maintenance arc} 
\end{cases} 
\] (11)

\[
x_{ki}^H = \begin{cases} 
1 & \text{ma intenance node } k \text{ and trip node } i \text{ is connected} \\
0 & \text{by a arc after maintenance} 
\end{cases} 
\] (12)

Eq. (13) restraints when a train-set completing Trip \( i \), it should either undertake Trip \( j \) or be carried out Maintenance \( k \). Eq. (14) restraints that Trip \( j \) should be undertaken by the train-set either just completing Trip \( i \) or Maintenance \( k \).

\[
\sum_{j=1}^{a} x_{ij}^C + \sum_{k=1}^{b} x_{ik}^Q = 1 \quad i \neq j \land k \neq i \quad (13)
\]

\[
\sum_{i=0}^{a} x_{ij}^C + \sum_{k=0}^{b} x_{ki}^H = 1 \quad i \neq j \land k \neq j \quad (14)
\]

3.3 Optimization Goal

Railway passenger transport corporations aim to realize profit maximization on the premise of ensuring passengers’ life and property security. Therefore, the cost is a key factor influencing the operation managers to make decisions. Due to the high acquisition cost and maintenance cost of train-sets, it is an effective measure to improve the utilization efficiency of train-sets. And so, to complete the same number of trip tasks in a given train graph, less train-sets are needed. In a train-set utilization plan, the required amount of train-sets is equal to the value that 1440 minutes, i.e., a cycle, divides the sum of the total connecting time and trips’ running time. The running time of all trips is a fixed value given in the train graph, so the required amount of train-sets is only determined by the total connecting time. Therefore, the optimization goal of train-set utilization plan can be set to minimize the total connecting time, which includes the following three parts: the connecting time of all connection arcs, that of all maintenance arcs, and that of all arcs after maintenance,

\[
\sum_{i,j\in[1,a],i\neq j} \omega_{ij}^C x_{ij}^C, \quad \sum_{i\in[1,a],k\in[1,b],i\neq k} \omega_{ik}^Q x_{ik}^Q, \quad \text{and that of all arcs after maintenance,} \\
\sum_{i\in[1,a],k\in[1,b],i\neq k} \omega_{ki}^H x_{ki}^H. 
\]

\[
\min \left( \sum_{i,j\in[1,a],i\neq j} \omega_{ij}^C x_{ij}^C + \sum_{i\in[1,a],k\in[1,b],i\neq k} \omega_{ik}^Q x_{ik}^Q + \sum_{i\in[1,a],k\in[1,b],i\neq k} \omega_{ki}^H x_{ki}^H \right) \quad (15)
\]

4. MODEL SOLUTION: A TWO-STAGE APPROACH

The train-set utilization planning model aims to identify whether arcs should be connected between trips and maintenances. Among the three kinds of arcs, the connection arcs between two trips is relatively easier to obtain due to the static characteristics that all trips are given in
the train graph. Inversely, when and where the maintenances should be carried out are indeterminate according to the accumulated operation time and mileages. Therefore, the difficulty of solving the train-set utilization planning problem lies in the solving process of maintenance arcs and arcs after maintenance. To solve the above difficulty, a two stage approach is proposed. In the first stage, all the trips in the given train graph are allocated into a set of segments and no arcs are connected between trip nodes and maintenance nodes. Then in the second stage, the maintenance arcs and arcs behind maintenance will be connected between different segments. By designed a solving algorithm in each stage, the optimal and unique solution of the train-set utilization planning model can be obtained, and so the compilation of train-set utilization plan can be realized in a computer-generated system.

4.1 The First Stage: Segment Generation

With the introduction of the segment concept, a directed graph \( \hat{G}(\hat{V}, \hat{E}) \) should be defined to represent the train-set utilization network. The node set \( \hat{V} \) is a union of ordered subsets \( \hat{V}_{SE} \) and \( \hat{V}_M \), which represent segments and maintenances, respectively. Each node \( i \in \hat{V}_{SE} \) is defined as a tuple \( (p_i, f_i^s, f_i^l, l_i^s, l_i^l, time_i, distance_i) \), where \( p_i \) indicates the number of trips in the Segment \( i \), \( f_i^s \) and \( f_i^l \) respectively indicate the departure station, and departure time of the first trip, \( l_i^s \) and \( l_i^l \) respectively indicate the arrival station and the arrival time of the last trip, \( time_i \) and \( distance_i \) indicates the running time and distance of all the trips, respectively.

The first stage aims to obtain the connection relationship between trip nodes and simultaneously prepare for the next stage. Therefore, when allocating trips into segments, the following requirements should be met.

- For each segment, to ensure no maintenance arcs and arcs after maintenance are included, the running time and distance, of all trips should be within the maintenance standards.
- For each segment, to ensure arbitrary two adjacent trips can be connected, the departure station of the latter trip should be the same with the arrival station of the former trip. Moreover, the dwell time between the two trips should satisfy the requirement of operation time.

Let \( ST = \{1, 2, \ldots, m\} \) represent the set of stations, where \( m \) is the total number of stations in the train graph. To obtain the segment sets more efficiently, all trips are classified into a group of new trip sets based on the trips’ departure station. Let \( DE = \{DE_1, DE_2, \ldots, DE_i, \ldots, DE_m\} \) represent the above new trip sets. Particularly, all the trips in set \( DE_i \in DE \) departure from station \( i \), where \( 1 \leq i \leq m \).

Based on above analyses, the segment generation algorithm is as follows.
Algorithm 1: Segment Generation

while \( DE \neq \emptyset \)
do generate segment \( \hat{v} \), initialize \( time = 0, distance = 0 \);
for \( i = 1 \) to \( 0 \)
  if \( DE_i \neq \emptyset \)
    add the first trip \( t_0 \) in the set \( DE_i \) into the segment \( \hat{v} \);
    \( distance = distance + d_0 \), \( time = time + (t_0^d - t_0^a) \);
    remove \( t_0 \) from \( DE_i \);break;
end
end
while \( distance < S_m + \Delta S_m \) and \( time < S_i + \Delta S_i \)
do find \( DE_j \) according to the arrival station \( s_0^a \) of trip \( t_0 \) in set \( DE_i \)
  if \( DE_j \neq \emptyset \)
    sort the set \( DE_j \) in ascending order based on the time interval between the arrival
    time of \( t_0 \) in set \( DE_i \) and the departure time of trips in the set \( DE_j \)
    for \( k = 0 \) to \( q_j \) ( \( q_j \) denotes the number of trips in the set \( ST_j \) )
      if \( t_k^d - t_k^a > T_o \)
        \( distance = distance + d_k \), \( time = time + (t_k^d - t_k^a) \)
        if \( S_m - \Delta S_m \leq distance \leq S_m + \Delta S_m \) & \( S_i - \Delta S_i \leq time \leq S_i + \Delta S_i \)
          add \( t_k \) into \( \hat{v} \), \( t_0 = t_k \), remove \( t_k \) from \( DE_j \);break;
        else: \( distance = distance - d_k \), \( time = time - (t_k^a - t_k^d) \)
      end
    end
  else:break;
end
record \( \hat{v} \);
4.2 The Second Stage: Segment Connection

There are three decision variables in the train-set circulation planning model, including $x^C_{ij}$, $x^O_{ik}$ and $x^H_{ki}$. In the first stage, the decision variable $x^C_{ij}$ has been solved since the connection relationships between trip nodes are obtained. Therefore, the objective of the second stage is to calculate the decision variables $x^O_{ik}$ and $x^H_{ki}$, i.e., obtain the connection relationship between trips and maintenance. With the concept of segment, the arcs between segments is equivalent to the combination of the maintenance arc and arc after maintenance. Therefore, the second stage aims to identify the segment connection relationship.

When the node $i \in \hat{V}_{SE}$ is connected to the node $j \in \hat{V}_{SE}$, it means that the train-set should be maintained after completing the last trip in the Segment $i$, and then undertake the first trip in Segment $j$. Therefore, to ensure the node $i \in \hat{V}_{SE}$ can be connected to the node $j \in \hat{V}_{SE}$, the following requirements should be satisfied:

- The arrival station of the last trip, $l_i^t$, in Segment $i$ should be the same with the departure station of the first trip, $f_i^{s_j}$ in Segment $j$.
- The time interval, $\alpha^S_{ij}$, between the last trip in the Segment $i$ and the first trip in the Segment $j$ should be longer than the duration of a series of working procedures, which is equal to the sum of $T_Z$ and $T_C$. Particularly, the value of $\alpha^S_{ij}$ can be obtained by Eq. (16).

$$\alpha^S_{ij} = \begin{cases} f_j^t - l_i^t & f_j^t - l_i^t \geq T_Z + T_C \\ 1440 + f_j^t - l_i^t & f_j^t - l_i^t < T_Z + T_C \end{cases} \quad (16)$$
Algorithm 2: Segment Connection

Step 1: Initial segment node set \( V_{SE} = \{ \hat{v}_1, \hat{v}_2, \ldots, \hat{v}_n \} \), arc set \( E = \phi \) and station set \( ST = \{ 1, 2, \ldots, m \} \), let \( st = 1 \).

Step 2: Classify the segment set \( \hat{V}_{SE} \) according to the last trip arrival station, \( l_i^a = st \), and the first trip departure station, \( f_j^d = st \); get arrival segment set \( \hat{V}_{st}^a = \{ \hat{v}_i | \hat{v}_i = (p_i, f_i^a, l_i^a, time_i, distan ce_i) \} \) and departure segment set \( \hat{V}_{st}^d = \{ \hat{v}_j | \hat{v}_j = (p_j, f_j^d, l_j^d, time_j, dis tan ce_j) \} \), sort the two sets by time ascending.

Step 3: Calculate \( \omega_{ij}^s \) according to the above rule, get the matrix \{ \( \omega_{ij}^s \) \}.

Step 4: Gain a set of optimal solutions by the Hungarian Algorithm, get the matrix \{ \( \omega_{ij}^s \) \}, add the zero corresponding arc into set \( E \).

Step 5: If \( st = m \), end the algorithm; else let \( st = st + 1 \) and go to step 1.

The segment connection algorithm is designed based on the Hungarian algorithm. The flow chart is shown as Figure 7.

FIGURE 7 Flow chart of the segment connection.
5. CASE STUDIES

5.1 Main Features of the Case Study

The length of Beijing-Tianjin passenger dedicated line is nearly 120 kilometers. There are five stations in the line, including two endpoint stations (Beijing South Station (BJSS) and Tianjin Station (TJS)), and three middle stations (Yizhuang Station (YLS), Yongle Station (YLS), and Wuqing Station (WQS)). Moreover, only the two endpoint stations BSS and TJS are located near the repair and inspection depot. Therefore, the maintenance of train-sets can only be carried out after arriving at BJSS and TJS.

The train graph used in this study was carried out realistically in December, 2014. There are total 174 trips in the train graph, among which 86 trips are from BJSS to TJS and 88 trips are from TJS to BJSS. The all the trips’ total running distances, $D_{all}$, is equal to 20,880 kilometers, and the total running time, $T_{all}$, is equal to 6,090 minutes.

The parameters related with the model and approach are set as follow:

- Basic operation time $T_O$: $T_O^{BJSS} = 14$ minutes; $T_O^{TJS} = 14$ minutes.
- Minimal duration of the four essential procedures if train-sets needs to be maintained $T_Z$: $T_Z^{BJSS} = 261$ minutes; $T_Z^{TJS} = 257$ minutes.
- Duration of preparation work before train-set departing in the station $T_C$: $T_C^{BJSS} = T_C^{TJS} = 28$ minutes.
- Distance range of maintenance standard $S_{one}$: $3,600 \text{ kilometers} \leq S_{one} \leq 4,400 \text{ kilometers}$.
- Time range of maintenance standard $T_{one}$: $1.8 \text{ day} \leq T_{one} \leq 2.2 \text{ day}$.

5.2 Computational Results of the Train-Set Utilization Plan

Figure 8 illustrates the computational results of the model, i.e., the train-set utilization plan of the train graph of the Beijing-Tianjin passenger dedicated line carried out in December, 2014.

Figure 8 Train-set utilization plan in Beijing-Tianjin passenger dedicated line.
The results show that to complete the above 174 trip tasks, the minimal amount of needed train-set $N$ is equal to 12. The total maintenance times of all the 12 train-sets $M$ equal to is 6 each day. The average running distances of each train-set per day $D_{average} = D_{all} / N$ is equal to 1740 kilometers. Each day, the rail line overhaul is carried out during the period 0:00-6:00. Therefore, the train-sets cannot work for six hours per day, and the maximal running time of each train-set can be calculated as follows: $1440-6*60=1080$ minutes. The average utilization efficiency of each train-set $U = T_{all} / (1080 * N)$ is equal to 46.9%.

In theory, the minimum connection time of this case is obtained by applying Hungarian algorithm, which is equal to 25,542 minutes when maintenance constraints are not considered. Moreover, the total maintenance times of all train-sets is 4.7 each day, which can be calculated as $D_{all} / (1+10%)S_{one}$. However, the theoretical results cannot be reached in practical production, because train-sets must be maintained according to the rules and regulations. Also, not all train-sets will be not maintained until their running distances reaching the upper limit of the standard. Therefore, the connection time 31,055 minutes and maintenance times 6 obtained by the proposed model and approach are very close to the theoretical values and satisfying. These results have proved that the proposed model and approach are feasible and the train-sets are utilized with high efficiency.

### 5.3 Comparison with the Ant Colony Algorithm

As what mentioned before, many researchers have applied the ant colony algorithm (ACA) to solve train-set utilization problem. In order to highlight the performance of this paper in improving the train-set utilization efficiency, the ant colony algorithm is applied to carry out same experiments in the same computer, and the detail comparison results are as listed in Table 2. The results obviously show that the two-stage approach has the advantages over the traditional ACA in improving the solution quality and shortening the computational time.

**TABLE 2 Results of The Comparison Between The Ant Colony Algorithm And The Two-Stage Approach**

<table>
<thead>
<tr>
<th></th>
<th>Two-stage approach</th>
<th>Traditional ACA</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>12</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>$M$</td>
<td>6</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>$D_{average}$ (kilometers/day)</td>
<td>1740</td>
<td>1606</td>
<td>134</td>
</tr>
<tr>
<td>$U$</td>
<td>46.9%</td>
<td>43.4%</td>
<td>3.5%</td>
</tr>
<tr>
<td>Computational time (millisecond)</td>
<td>2302</td>
<td>8032</td>
<td>5730</td>
</tr>
</tbody>
</table>

On the one hand, the utilization efficiency of train-sets is much higher by applying the two-stage approach. Firstly, the amounts of train-sets needed to complete the trip tasks are respectively 12 and 13 obtained by the two-stage method and traditional ACA. Secondly, the total maintenance times has been decreased from eight times to six time, saving the invalid running time from the railway stations to the inspection and repair depot. Thirdly, the average running distances of each train-set per day is 134 kilometers longer than that of the traditional ACA.

On the other hand, the computation time of the two-stage approach is 2302 milliseconds, which is almost one-third of the traditional ACA. What should be noted is that,
due to the convergence speed of the traditional ACA is indeterminate, we take the average
time of 20 computation times. Furthermore, the computation of the two-stage approach is
much more stable. Only one optimal train-set utilization plan can be obtained while
experiments will achieve different results by using the traditional ACA.

6. CONCLUSION

The train-set utilization plan is drawn up to identify the work arrangements of
train-sets, which is profoundly affected by to the given train graph as well as the rules and
regulations of maintenance. To optimize the train-sets’ utilization efficiency, in this study, an
integer programming model is proposed, considering the spatial constraints, the time
constraints, the maintenance constraints, and the unicity constraints. In the process of model
solution, a two-stage approach is designed, and then an optimal train-set utilization plan is
obtained as the output. Using the model, the case of the realistic train graph of the
Beijing-Tianjin passenger dedicated line was carried out. The results show that the model and
approach proposed in this study is practical and reasonable. By comparing the results
obtained by the two-stage approach and the ant colony algorithm, it could be found that to
complete the same trip tasks, less amount of train-sets is needed, the maintenance times is
decreased, and the average running distances of each train-set per day is much longer. In
short, the utilization efficiency of train-sets is much higher by applying the two-stage
approach. Moreover, the computation speed and stability are much better, which may
contribute to the realization of a train-set utilization plan computer-generated system.

ACKNOWLEDGMENT

This work is financially supported by “the Fundamental Research Funds for the
Central Universities” (2015YJS094 & 2015YJS091).

REFERENCE

(2) Abbink, E., V.D.B Bianca, L. Kroon, and M. Salomon. Allocation of railway rolling
(4) Fioole, P., L. Kroon, and A. Scrijver. A rolling stock circulation model for combining and
No.2, pp.1281-1297.
origin-destination integer multicommodity flow problems. Operational Research., 2000,
Vol.48, No.2, pp.318–326.
(7) Holmberg, K., and D. Yuan. A multi-commodity network flow problem with side
constraints on paths solved by column generation. Informs Journal on Computing, 2003,
(8) Dantzig, G.B., and P. Wolfe. Decomposition principle for linear programs. Operational


