Optimization of the Electrical Multiple Unit Circulation Plan Considering Train Connection Time and Under-Repair Mileage

Yongzhi YANG, Xinfeng YANG, and Ke YANG

Abstract—**Reasonable operation optimization of Electric Multiple Units (EMU) can effectively reduce the maintenance costs of high-speed railways and enhance the level of transportation organization. With the premise of ensuring that EMUs can arrive at the maintenance depot promptly for necessary maintenance, a closed-loop connection network is constructed based on the operational requirements of EMUs in China. A nonlinear integer programming model is proposed to minimize the total cost of EMU mileage loss, total connection time, and the number of routes, which overcomes the shortcomings of previous studies that did not sufficiently coordinate EMU maintenance. A linearization technique transforms the original nonlinear state function into a linear form, and the Gurobi solver is used through Python to solve the model. The modeling and solution methods are then applied to practical Chinese high-speed railway system examples. The results show that compared to manual design and simulated annealing algorithm approaches, the proposed optimization method increases the average cumulative mileage utilization by 12.8% and 8.9%, respectively. This method can significantly enhance the quality of EMU circulation planning, mitigating the impractical mileage and time costs associated with traditional optimization methods. Moreover, it demonstrates excellent adaptability to the dual maintenance cycle requirements of EMUs. The proposed approach can achieve more realistic and efficient EMU circulation planning by effectively considering these factors.**

Index Terms—**Connection Network, EMU Circulation Planning, High-speed Railway, Maintenance Constraints, Nonlinear Integer Programming**

I. Introduction

In order to handle the heavy passenger transportation tasks on the high-speed railway system and ensure that each train can be operated by an appropriate EMU (Electric Multiple Unit), it is necessary to enhance the utilization efficiency of EMUs while keeping the number of EMUs unchanged. Over time, railway managers have begun to pay more attention to this issue, but it has yet to be effectively resolved. Furthermore, the links between any two train services must adhere to the precise guidelines of the train timetable, including restrictions on travel time and distance. This ensures that every train set receives maintenance throughout its maximum maintenance cycle, providing an essential basis for allocation and planning. If this issue is effectively addressed, it can reduce the number of EMUs required, lower the costly procurement expenses for EMUs, improve the operational efficiency of high-speed railways, provide greater transportation capacity to meet the growing demand for transportation, and prevent the situation where EMUs exceed the maintenance period yet continue to operate. This research problem holds significant theoretical significance and practical value.

Manuscript received June 03, 2023; revised November 22, 2023. This work was supported in part by Double-First Class Major Research Programs, Educational Department of Gansu Province (No. GSSYLXM-04), Projects Funded by the Central Government to Guide Local Scientific and Technological Development (No. 22ZY1QA005).

Yongzhi Yang is a postgraduate student at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (e-mail: 17361600434@163.com).

Xinfeng Yang is a professor at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (Corresponding author, e-mail: xinfengyang@mail.lzjtu.cn).

Ke Yang is an engineer of Beijing Rail Transit Design & Research Institute, Beijing 100037, China. (e-mail: Y93yk@126.com).

In existing studies, the optimization of EMU utilization primarily involves two aspects: mathematical problem formulation and algorithm design for solving. Regarding modeling, most current research transforms the problem into multi-commodity flow or multi-traveling salesman problems. For example, Abbink et al. [1] developed a bi-objective multi-commodity flow optimization model to minimize the number of EMUs and the total traveled distance in a single-line periodic timetable by optimizing the rotation plan that allows for EMU coupling and decoupling. Giacco et al. [2] studied EMU operations and maintenance planning in the short term, considering time-based maintenance constraints and utilizing the minimum-cost Hamiltonian cycle approach for modeling. Wang Chao et al. [3] analyzed the optimization model of EMU operations and existing train timetables, and building on the Vehicle Routing Problem (VRP), they constructed a coordinated optimization model for train timetables and EMU operations. The model was addressed using GAMS software, resulting in satisfactory EMU operation plans that met the operating requirements. Miao Jianrui et al. [4] transformed the EMU operation problem into M-TSP issues with resupply, established a corresponding multiple-goal model, and designed a layered optimization heuristic algorithm. Wang Ying et al. [5] developed a branching pricing algorithm to tackle the characteristic of a large number of decision variables compared to the number of constraints in the optimization model. Zhong Qingwei et al. [6] analyzed multi-commodity network flow theory and set up a two-stage optimization model for EMU operations. Cadarso et al. [7] aimed to minimize the comprehensive cost and considered passenger demand, train composition, train formation, and empty movements in the train plan. They developed a robust optimization model without considering maintenance constraints. Li Jian et al. [8] formulated a 0–1 integer programming model with the objectives of cutting connection time and increasing the cumulative mileage of EMUs. According to the trip sequence graph, Gao et al. [9] suggested arc modeling and path modeling and used a branch and pricing algorithm to solve them. Regarding solution algorithms, the methods for solving EMU circulation planning include exact analytical and intelligent algorithms. The former includes commercial solvers [10–14] and branch-and-price algorithms [15–16]. The latter mainly includes heuristic algorithms such as ant colony optimization [17–18], genetic algorithm optimization [19], and simulated annealing [20–21].

In summary, the majority of current research focused on optimizing Electric Multiple Unit (EMU) circulation plans is geared toward reducing the required number of EMUs within an existing railway timetable and minimizing the overall connection time of EMUs. These studies consider different constraints, such as maintenance cycle constraints based on mileage and time, EMU coupling and decoupling scenarios, relationships between trains and EMUs, and empty EMU movements. However, research about the optimization of circulation plans tends to concentrate on optimizing the routes of a single high-speed line or a few EMUs between several stations within a day, and some routes extend beyond one day in duration. Limited research considering multi-day optimization of circulation planning and a lack of studies simultaneously optimizing route mileage and the number of routes. Additionally, there is insufficient coordination of EMU maintenance plans, with inadequate consideration of accumulated mileage and time for maintenance under a dual maintenance cycle. Therefore, there is a need for an accurate mathematical model to describe the cumulative changes in mileage and time. This study further adopts a transfer approach to calculate cumulative mileage and time, considering the complex and numerous interconnection relationships between train services. These interconnections are abstracted into a connection network to provide a more precise and intuitive description of EMU operations.

The remaining sections of this article are organized as follows: Section 2 explores the daily operational requirements of high-speed trains in China. It proposes an analysis of train connections and maintenance tasks within a closed-loop high-speed train network. Drawing inspiration from the processing approach used in the Traveling Salesman Model, a multi-day operational planning optimization model is developed. This model takes into consideration train connection time and accumulated maintenance distance. To represent train connections and maintenance, binary variables (0–1) are utilized, and nonlinear constraints are linearized. Section 3 presents the specific algorithm design that uses Python to invoke commercial solvers for solving the optimization model. Section 4 conducts a case study to validate the effectiveness of the proposed method. Finally, Section 5 provides a comprehensive summary of the article and offers a preview of possible future research directions.

II. PROBLEM DESCRIPTION AND MODELING

A. Optimization of EMU Circulation Planning Based on Closed-Loop Networks

1) Description of the problem

In the actual operational process, the circulation planning of train units, also known as EMU circulation or train route planning, plays a crucial role. It involves organizing and coordinating train service tasks and other related tasks. Circulation planning is part and parcel of operational planning, as it combines train services from the timetable, adhering to connection rules and maintenance requirements, to create train unit routes. These routes are commonly referred to as EMU circulation plans or train route plans. From this perspective, the EMU circulation can be seen as the smallest unit in preparing the EMU operational plan, and preparing the circulation plan is a preprocessing process to form the operational plan. Therefore, the primary concern of the article is how to optimize the circulation plan.

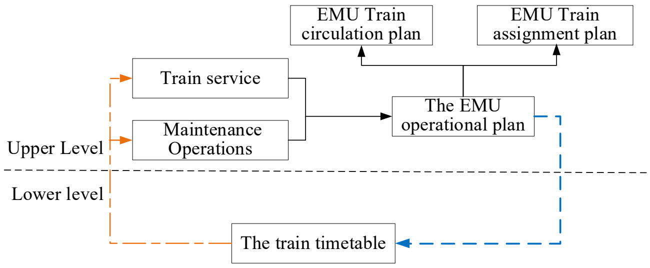


Fig. 1. Relationship Graph of Planning Formulations

The EMU circulation is typically seen as a sequential succession of train unit operational tasks, with each train unit undergoing maintenance after completing a transportation task. Since maintenance needs to be scheduled at the train depot where the train unit corresponds, the same station, adjacent to the train depot, serves as the starting point for the initial train service and the finishing point for the final train service. Furthermore, there are many different types of train units in practice, and train services and train units must be matched appropriately. Additionally, multiple train depots collectively undertake transportation missions. It is clear that making the EMU circulation plans involves several situations and is extraordinarily challenging. This complex issue can be decomposed into several simpler subproblems, each containing train units of the same type and belonging to the same train depot. By decomposing the problem, we can apply the same optimization methods to solve each subproblem. This study created a circulation planning optimization approach for EMUs with the premise that every train unit is a part of the same train depot. Furthermore, this study does not permit deadhead movements as they are rare in practice and incur high scheduling costs.

In conclusion, the compilation of EMU circulation plans involves determining the connections between train services based on the train timetable within one cycle. It aims to achieve certain planning objectives and satisfy specific constraints, ensuring the rationality of the train service connections.

2) Construction of Closed-Loop Circulation Networks for High-Speed Train Units

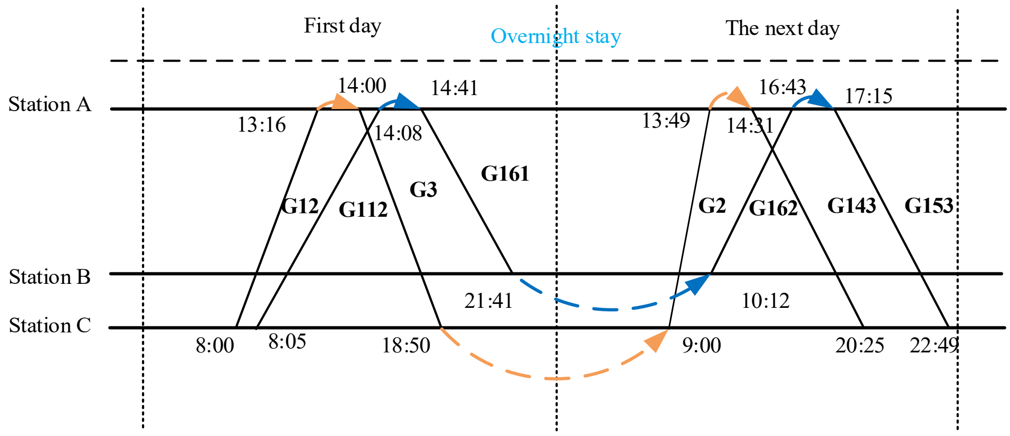


Fig. 2. Example of the Train Unit Route Diagram

Based on the theory of optimization networks, the train services are set as the nodes of the train unit connection network, and the connections between any two train services are defined as arcs. Additionally, the train unit connection network also includes arcs for maintenance work. Each path can be considered a closed loop where the maintenance or connection arcs can only be connected to one preceding and one succeeding train service for a particular train service.

An EMU connection network is built using the example of the two loops in Fig. 2. C is the station connected to the EMU depot, serving as the maintenance node. Loop route I in Fig. 2 starts at station C, stops at stations A and B, and returns to station A the following day. With only a stop at station C, Loop Route II lasts for two days. They are both two-day circulations and can be reduced to the connection network described below.

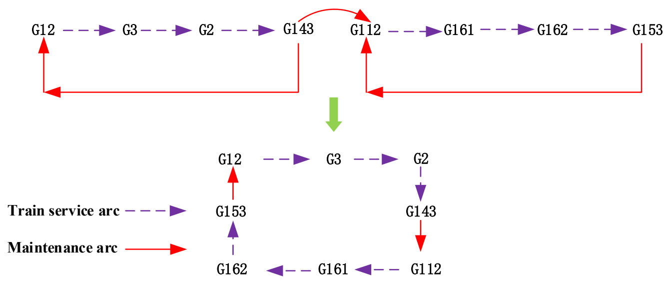


Fig. 3. Schematic diagram of the connection network of train services and maintenance tasks

The red arc in Fig. 3, from G143 to G112, represents the need for Level 1 maintenance (since the connection network only contains connections between train services, timing information for maintenance jobs cannot be acquired from it, and whether a train service is suitable for connection depends on time and should be determined based on constraints). Each train unit cycle can be defined as a closed loop consisting of train services connected in sequence and train unit maintenance. To decrease the level of complexity of EMU circulation planning, since the originating and terminating stations of the train unit route are the same stations, the two closed loops are connected into one closed loop through the arc representing train unit maintenance. As shown in Fig. 3, it includes all train service nodes, train service connection arcs, and train unit maintenance arcs. Based on the idea of establishing a single closed loop, the primary task of EMU circulation planning is to determine the train service connections and reasonably arrange maintenance, thereby generating a network with only one closed loop. Then, the single closed-loop arcs are decomposed from the arcs representing train unit maintenance, generating a sequence of arcs connecting multiple train services.

B. Optimization modeling based on connection networks

1) Modeling assumptions

To facilitate the construction of the model, several assumptions are made in this study, as follows:

Assumption 1: The high-speed train unit is considered an indivisible entity, and the coupling and decoupling of train units are not considered.

Assumption 2: Crew scheduling issues are not considered. In this study, it is assumed that all of the high-speed train units are of the same type and come from the same depot. Train services in the timetable are paired, and deadhead runs are not permitted.

Assumption 3: The limitations on the first-level maintenance capacity at the depots are not considered. In the case of high-speed train units arriving at the depot collectively, the maintenance and shunting plans within the depot are optimized to balance the maintenance of high-speed train units.

Building upon existing research findings and drawing from the concept of constructing connection networks, this study primarily focuses on optimizing the turnover of high-speed train units. The objectives include reducing the overall train unit connection time, increasing mileage within each stage, and minimizing the number of circulations.

2) Definition of symbols

A mathematical optimization model is proposed for making high-speed train unit circulation plans based on analyzing high-speed unit circulation problems and creating high-speed train unit connection networks. A group of train services included in a deterministic train schedule is taken into account by the model. Moreover, recognizing the need for train connections within each route to adhere to maintenance requirements, we have meticulously incorporated maintenance operations into our considerations. The symbols and parameters involved in the model are shown in TABLE I below.

TABLE I
SYMBOLS IN THE CONNECTION NETWORK

Symbol	Definition
V	Set of Train Services: A collection of train services.
m,n	Train Service Index: An index used to identify a specific train service.
v_m	Train Service m: A specific train service within the set of train services.
k	Number of Train Services: The total count of train services in the set.
$S^{station}$	Set of Stations: A collection of stations.
S^0	Set of Stations Connected to the Depot: The stations that are connected to the depot.
s_m^{Arr}	Arrival Station of Train Service m: The station where train service m arrives.
t_m^{Arr}	Arrival Time of Train Service m: The time when train service m arrives.
s_m^{Dep}	Departure Station of Train Service m: The station from where train service m departs.
t_m^{Dep}	Departure Time of Train Service m: The time when train service m departs.
d_m^{Travel}	Current Mileage of Train Service m: The current distance traveled by train service m.
t_m^{Travel}	Current Journey Time of Train Service m: The current duration of the journey for train service m.
u_m	Position Index of Train Service m in a Unique Closed Loop: The numbering of train service m's position within a unique closed loop.
t_0	Operation Time for First-Level Maintenance: The time required for performing first-level maintenance.
L^{Cycle}	Maintenance Mileage Cycle: The cycle for first-level maintenance mileage.
T^{Cycle}	Maintenance Time Cycle: The cycle for first-level maintenance time.

3) Accumulated mileage and time of connection

Based on the EMU connection network, the connection between EMU train services is considered a 0–1 decision variable. With the help of binary decision variables, the formula for the auxiliary decision variable cumulative mileage is described as:

$$l_n = d_n^{Travel} + \sum_{m=1, m\neq n}^{k} x_{m,n}\left(1-y_{m,n}\right) l_m \tag{1}$$

If train m is connected to train n, *i.e.*, $x_{m,n}=1$, and there is no overhaul between train n and train m, the accumulated mileage is the same as the total of the accumulated mileage of the immediately preceding train m, *i.e.*, l_m, and the mileage of the current train n, *i.e.*, d_n^{Travel}. If there is an overhaul between train m and train n, *i.e.*, $y_{m,n}=1$, the value of the accumulated mileage is only equal to the mileage of the current train n.

In the same way, the formula for the total time of train n can be written as follows:

$$t_n = t_n^{Travel} + \sum_{m=1, m\neq n}^{k} x_{m,n}\left(1-y_{m,n}\right)\left(t_m + \Delta t_{mn}^{connect}\right) \tag{2}$$

Similarly, when there is no first-level maintenance scheduled between two interconnecting train services m and n, the cumulative running time, *i.e.*, t_n, is equal to the running time of the current train service n plus the cumulative running time of the immediately preceding train service, *i.e.*, t_m, and the connecting time between the two, *i.e.*, $\Delta t_{mn}^{connect}$. If first-level maintenance is scheduled, then it is only equal to the running time of the train service, *i.e.*, t_n^{Travel}.

4) Consideration of train connection times for two-day

Since empty train deployment is not permitted in this paper, every connection between any two trains must adhere to particular restrictions. As a result, train service m and train service n can only be connected if train service n leaves from the same station as the arrival of train service m. In summary, a parameter $\Delta t_{mn}^{connect}$ is defined and used as a generic cost for the connection between train services m and n. The parameter represents the connection time between train services m and n. If the arrival station of train m is different from the departure station of train n, the connection time takes on the cost of infinity, as shown in equation (3).

$$\Delta t_{mn}^{connect} = \begin{cases} t_n^{Dep} - t_m^{Arr}, & s_m^{Arr} = s_n^{Dep}, (t_n^{Dep} - t_m^{Arr}) \geq \tau^{connect} \\ t_n^{Dep} - t_m^{Arr} + t^{Day}, & s_m^{Arr} = s_n^{Dep}, a_{mn}(t_n^{Dep} - t_m^{Arr}) \leq \tau^{connect} \\ M, & other \end{cases} \tag{3}$$

Where $\tau^{connect}$ represents the shortest time requirement (typically 15 minutes) for connecting train services generated by passenger boarding and alighting, as well as the train cleaning time, and M is a sufficiently large positive number. There are two cases when train m and train n connect at a station. First, if the departure time t_n^{Dep} and arrival time t_m^{Arr} differ by more than the minimum allowed amount of time, then train m can be directly connected to train n. Otherwise, it can be considered to connect to the train service on the next day, typically after $t^{Day} = 1440$ min, as shown in the red arc in Fig. 4.

For example, in Fig. 4, Station C is located near the train depot. The alternative train service G6' will be available for Train G4 to connect to if the time gap between the arrival of Train G4 and the departure of Train G6 in Fig. 3 is less than the regular connecting time. This two-day connection method allows for more flexibility in the design of the EMU train circulation, but it also requires more connection time.

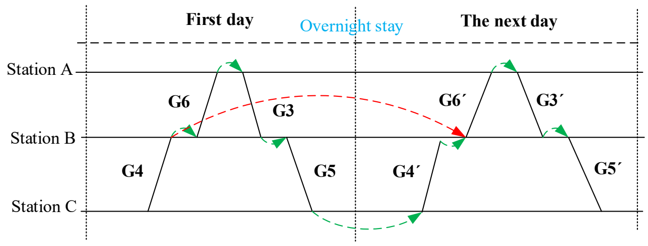


Fig. 4. Illustration of connection time

In addition, when the terminal station of train m and the starting station of train n are both connected to the EMU operation depot, *i.e.*, $s_m^{Arr} = s_n^{Dep} = 0 \in S^0$, maintenance work can be scheduled between v_m and v_n. Therefore, a binary auxiliary decision variable $\delta_{m,n}$ is defined to determine whether maintenance work can be scheduled, and its value can be determined according to equation (4).

$$\delta_{m,n} = \begin{cases} 1, & s_m^{Arr} = s_m^{Dep} = 0,\ \Delta t_{mn}^{connect} \geq \tau^{connect} + t_0 \\ 0, & other \end{cases} \tag{4}$$

Based on the above problem analysis and the definition of relevant parameters and variables, an integer programming model for the optimization of EMU train scheduling has been established.

5) Optimization objectives

The EMU circulation plan aims to find a closed loop of train connections for a specified number of EMUs, i.e., to increase the EMU trains' usage effectiveness. Therefore, the objective function is formulated to minimize connection time, maximize the utilization of mileage cycles, and reduce the number of EMUs in service.

a) Minimization of idle mileage

The overall difference between the cumulative mileage during maintenance and the upper bound on the maximum maintenance mileage is minimized. This entails raising each EMU train's cycle mileage as much as feasible while staying within the maintenance mileage maximum. On the one hand, this improves the utilization efficiency of EMU trains and reduces the number of EMU trains required. On the other hand, it reduces the frequency of maintenance operations and lowers maintenance costs.

$$Z_1 = \sum_{m=1}^{k} \sum_{n=1, n \neq m}^{k} y_{m,n}((1+\lambda)L^{cycle} - l_m) \tag{5}$$

b) Minimization of connection time

The EMU train connection network makes an effort to reduce the total connection time between any two train services connected by arcs. This lessens the need for EMU trains and forces closer links between train services within each circulation, providing more idle time for maintenance operations. This objective excludes the consideration of connection time between train services connected by maintenance arcs, as these train services belong to different circulations.

$$Z_2 = \sum_{m=1}^{k} \sum_{n=1, n \neq m}^{k} x_{m,n} t_{m,n} (1 - y_{m,n}) \tag{6}$$

c) Minimization of the number of circulations:

On the one hand, the number of circulations (routes) directly determines the frequency of primary maintenance. Minimizing the number of circulations can significantly reduce the maintenance costs of EMU trains. On the other hand, once the train timetable is determined, the overall operational tasks of the EMU trains remain constant. The number of circulations affects the number of EMU trains required when assigning tasks to the EMU depot. Therefore, reducing the number of circulations in optimization can decrease the number of EMU trains needed and reduce maintenance costs. The corresponding objective function is as follows:

$$Z_3 = \sum_{m=1}^{k} \sum_{n=1, n \neq m}^{k} y_{m,n} \tag{7}$$

Based on the definition of circulations, it can be determined that the primary maintenance operations of EMU trains serve as the starting and ending points of circulations. Therefore, the frequency of primary maintenance can determine the number of circulations.

d) Conversion to single-objective optimization:

Three weighting factors for the optimization objectives are defined: β , γ , and η . This reduces the model's complexity. With this, the multi-objective optimization issue is reduced to a single-objective optimization. The values of these three weighting coefficients can be determined according to the actual situation in the application.

$$Min \quad Z = \beta Z_1 + \gamma Z_2 + \eta Z_3 \tag{8}$$

6) Binding conditions

a) Uniqueness constraint of train connections:

The EMU circulation plan is developed after the train timetable is determined. Therefore, all the train services in the diagram must be connected and can only be connected once. In the connection network, for a certain train service, there is exactly one train service connected to it in the preceding position and exactly one train service connected to it in the succeeding position, i.e.,

$$\sum_{m=1, m \neq n}^{k} x_{m,n} = 1 \quad n = 1, 2, \cdots, k \tag{9}$$

$$\sum_{n=1, n \neq m}^{k} x_{m,n} = 1 \quad m = 1, 2, \cdots, k \tag{10}$$

b) Constraints on the relationship between train number and maintenance succession variables

First-class repair work can only be arranged when train m and train n are connected and the conditions for repair work are met between them.

$$y_{m,n} \leq x_{m,n} \delta_{m,n} \quad m, n = 1, 2, \cdots, k, \ m \neq n \tag{11}$$

c) Maintenance cycle constraints for EMU

EMU trains usually need to return to the EMU depot for maintenance after a cumulative mileage of 5000 km or 48 hours of running time after the last maintenance. But the cumulative mileage following the most recent maintenance can't exceed the corresponding maintenance cycle within a certain upper bound, often not more than 10%, in the actual operation and administration of EMUs. Therefore, the EMU Class I maintenance cycle constraint can be expressed as:

$$l_n \leq (1+\lambda) L^{cycle} \quad n = 1, 2, \cdots, k \tag{12}$$

$$t_n \leq (1+\lambda) T^{Cycle} \quad n = 1, 2, \cdots, k \tag{13}$$

d) Avoid generating subloop constraints.

The classical constraint to avoid making subloops in the TSP problem is used to make sure that the final generated connection network is a closed loop and to stop subloops from being made in the connection network. Namely:

$$u_m - u_n + k x_{m,n} \leq k - 1 \quad m = 1, 2, \cdots, k \quad n = 2, \cdots, k \tag{14}$$

$$0 \leq u_m \leq k - 1 \quad m = 1, 2, \cdots, k \tag{15}$$

e) Decision variable constraints

$$x_{m,n}, y_{m,n} \in \{0,1\} \quad m, n = 1, 2, \cdots, k, \ m \neq n \tag{16}$$

7) Model processing

Through the analysis of the above model, it is found that there are two 0–1 variables multiplied in the model, and for the convenience of the solution, it is necessary to linearize the corresponding parts of its objectives and constraints. According to the derivation, the auxiliary decision variable $z_{m,n}$ is introduced such that $z_{m,n} = x_{m,n} * (1 - y_{m,n})$, and $z_{m,n} \in \{0,1\}$, the following auxiliary constraints are introduced:

$$z_{m,n} \leq x_{m,n} \tag{17}$$

$$z_{m,n} \leq 1 - y_{m,n} \tag{18}$$

$$z_{m,n} \geq x_{m,n} - y_{m,n} \tag{19}$$

After the above process, the non-linear model is reconstructed into a linear programming model that can be solved using a mathematical solver.

III. SOLVING ALGORITHMS

The above model shows that the EMU circulation optimization problem is a non-linear, constrained integer programming optimization problem. To facilitate the solution, the mathematical solver is invoked to solve it after linearizing it.

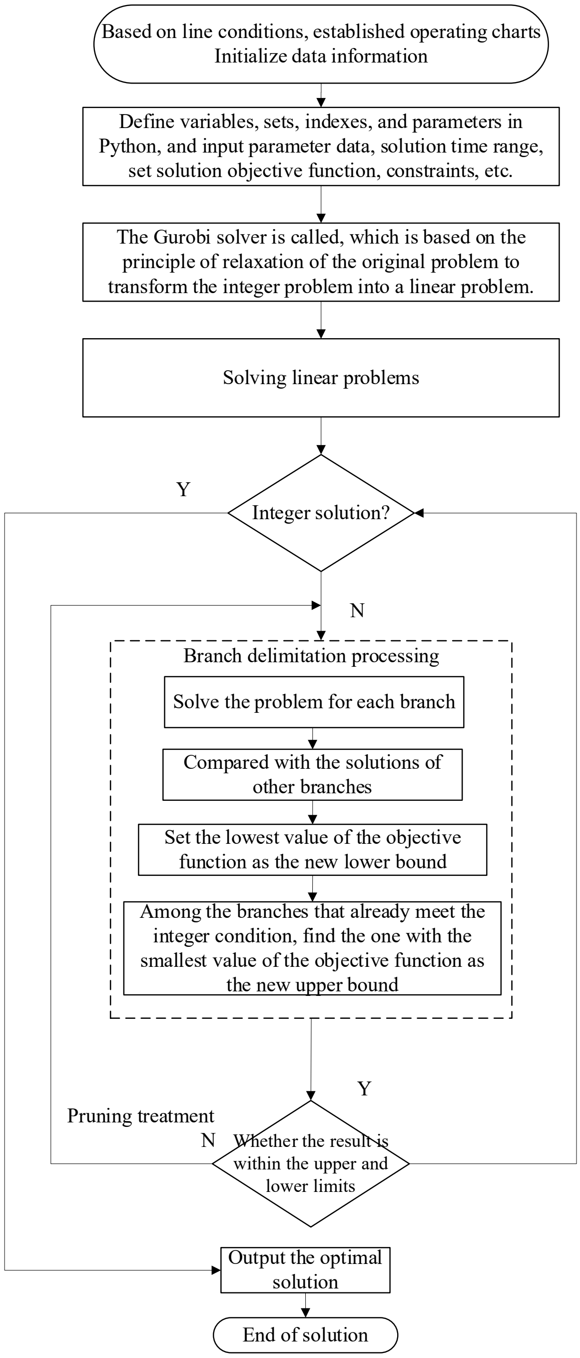


Fig. 5. Flow chart of solution

Based on the train operation diagram, the spatiotemporal connection network of the EMU circulation plan is built, the arc segment data of each node is set, and the Gurobi solver is used to solve it. Compared with other purely linear problems, this model is an integer programming problem, so the solver for this problem adopts a branching delimitation algorithm; that is, it is solved as a series of LP subproblems. The solver solution steps are shown in Fig. 5.

IV. SIMULATIONS AND ANALYSIS OF RESULTS

This section conducts extensive practical numerical research to evaluate the viability and efficiency of the suggested model and solution. Initially, assess the model's performance in a small-scale scenario comprising 16 train journeys and 4 stations. Upon successful validation on this limited dataset, proceed to test the model on a larger dataset to ensure its generalization capabilities. This step is crucial for evaluating how well the model performs in broader and more diverse scenarios.

A. Small-scale Case Study Analysis

To further validate the model's accuracy and assess the mathematical programming software's ability to solve the model within an acceptable time frame, we performed a verification on a scenario involving 16 trains and 4 stations. In the test case, detailed information for each train is presented in Table II. After 24 seconds of optimization computation using the software, we obtained the globally optimal solution for the EMU circulation scheme.

TABLE II
TRAIN BASIC INFORMATION TABLE

No	Departure station	Arrival station	Departure time	Arrival time	Distance	Travel time
G1	Ningbo	Changsha	6:56	12:59	1079	363
G2	Ningbo	Shanghai	7:25	9:52	314	147
G3	Shanghai	Ningbo	8:05	9:52	314	107
G4	Shanghai	Changsha	9:00	13:41	1083	281
G5	Shanghai	Changsha	9:51	15:17	1083	326
G6	Guiyang	Shanghai	10:09	19:31	1790	562
G7	Ningbo	Shanghai	10:28	12:37	314	129
G8	Shanghai	Changsha	10:50	16:14	1083	324
G9	Guiyang	Shanghai	11:47	20:56	1790	549
G10	Shanghai	Guiyang	11:50	20:51	1790	541
G11	Changsha	Shanghai	13:25	19:11	1083	346
G12	Shanghai	Guiyang	13:43	22:39	1790	536
G13	Changsha	Shanghai	15:00	19:39	1083	279
G14	Changsha	Ningbo	15:36	21:34	1079	358
G15	Changsha	Shanghai	16:40	22:10	1083	330
G16	Shanghai	Ningbo	20:43	22:58	314	135

Utilized Gurobi for the optimization of a model with 1763 rows, 815 columns, and 3729 nonzeros. The model encompasses 256 quadratic objective terms and 32 quadratic constraints, featuring variable types of 32 continuous variables and 783 integers (768 binary). Explored 3522 nodes through 65060 simplex iterations, achieving optimization within a rapid 0.77 seconds. The results are detailed in Table III and Figure 6.

TABLE III
OPTIMAL CIRCULATIONS PLAN

NO.1	L^{Cycle} /km	T^{Cycle} /min	Connection time/min	Components of circulation
1	4960	2230	734	G4→G13→G16→G2 →G8→G15
2	3580	1901	798	G10→G6
3	4208	2211	890	G3→G7→G12→G9
4	5244	2114	716	G5→G14→G1→G11

Based on the validation of the aforementioned small-scale data, both the model and the algorithm have proven to be reliable and accurate. The next section will proceed with calculations on larger-scale cases.

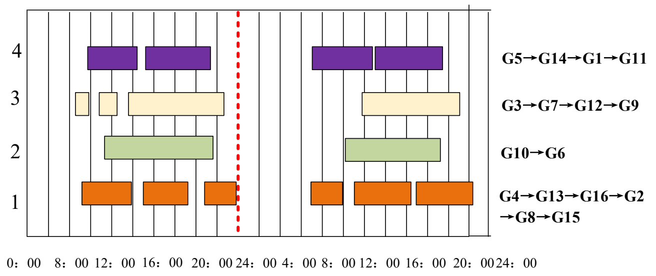


Fig. 6. Train Network Gantt Chart

B. *Large-scale Case Study Analysis*

The data comes from a railway EMU depot in China that is responsible for a total of 92 CRH380B series EMUs. The physical network of the 92 trains is shown in Fig. 7. There are 22 stations involved, with Shanghai Hongqiao station being in proximity to the depot. The basic information on these 92 trains and the manual simulated annealing method to prepare the EMU circulation scheme is shown in the literature [21].

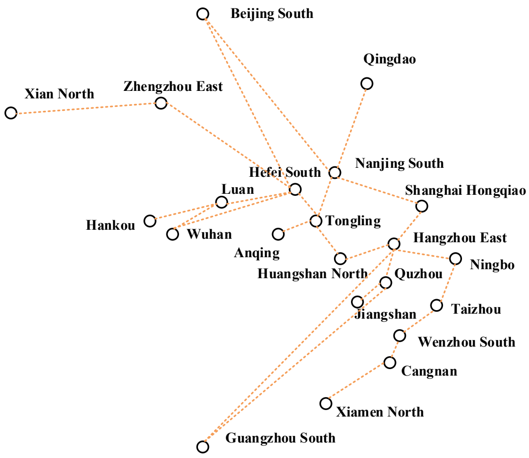


Fig. 7. High-speed railway network diagram

In this example, the mileage cycle for the maintenance of trains is 5,000 km, and the maximum value of the mileage cycle is 5,500 km; the time cycle is set to 2,880 min; the operation time for first-level maintenance is 240 min; and according to the information on 92 trains in the example, there are 23 circulations in the manual sketching scheme. The average mileage of these 23 manually outlined circulations is 4,257 km, and the average travel time is 1,815 min.

Gurobi was invoked via Python to solve the EMU circulation plan, and the following settings were made for the model's and the algorithm's parameters: The optimization objective function's weight factor was set to 1.0, and the minimum successive time between adjacent running lines was set to 15 minutes. The method was implemented on a laptop with an AMD Ryzen 5 5500U with Radeon Graphics clocked at 2.10 GHz and 16.0 GB of RAM. After 600 seconds, a nearly optimal circulation plan was found. According to the results of the calculations, 20 circulations were generated with an average cumulative travel distance $\bar{L}^{Cycle}$ and travel time $\bar{T}^{Cycle}$ of 4895 km and 2140 min, and the approximate optimal plan found is given in TABLE IV.

Tables V and VI comprise crucial performance metrics pertaining to the results of the simulated annealing algorithm and manually devised calculations. Table VII shows a comparison of the three calculation methods' indicators. When using Gurobi's approximate optimal solution, the average mileage is 639 km higher than when using the manual sketching method and 445 km higher than when using the simulated annealing algorithm. The average mileage of each EMU circulation is also closer to the ceiling. The average cumulative mileage utilization is 12.8% higher than the manual sketching scheme and 8.9% higher than the simulated annealing method. The average train connection time is 574 min for the manual scheme, 876 min for the simulated annealing algorithm, and 712 min for the model's nearly optimal loop scheme. Given that the objective function seeks to achieve a balance between minimizing connection time, maximizing mileage, and minimizing the number of circulations, the connection time of the optimal circulation plan may not be smaller than the connection time of the manual plan.

TABLE IV
APPROXIMATE OPTIMAL CIRCULATIONS PLAN

NO.1	L^{Cycle} /km	T^{Cycle} /min	Connection time/min	Components of circulation
1	5272	2217	929	G110→G3→G2→G143
2	4542	2028	593	G1806/7→G1868/5→G1866/7→G1828/5
3	5112	1961	550	G128→G267→DJ8693→DJ8694→D5495→G266→G17
4	5232	2194	560	G102→G133→G7535→G590/1→G600/597
5	5144	2309	797	G114→G161→G162→G21
6	5094	2146	508	G1655→G1658→G1912/3→G1930/27
7	4596	1778	517	G132→G269→G262→G129
8	4744	2357	806	G7503→G7542→G7545→G7504→G126→G157
9	5272	2139	735	G112→G145→G12→G139
10	5442	2238	534	G230/1→G236/3→G1916/7→G1934/1
11	5442	2109	629	G360/1→G1938/5→G222/3→G228/5
12	5244	2114	716	G14→G351→G352→G19
13	5144	2334	710	G598/9→G592/89→G7330→G122→G153
14	5272	2059	654	G118→G159→G108→G141
15	4589	1997	550	G16→G149→DJ7725→G7361→G7364→G7509→G7514→G7519 →G7524
16	3580	1781	763	G1303→G1302
17	4760	2250	682	G1509/8→G1507/10→G7381→G7382→G116→G147
18	4815	2350	995	G85→G1306→G7501→G7536
19	4682	1848	591	G130→G201→G202→G131
20	3928	2588	1430	G7521→G7516→G8→G5

TABLE V
THE RESULTS OF THE SIMULATION ALGORITHM FOR SOLVING

NO.1	L^{Cycle} /km	T^{Cycle} /min	Connection time/min
1	3840	2181	1126
2	4682	1938	666
3	3720	1890	747
4	5272	2266	870
5	4511	2377	974
6	4815	2415	1060
7	5272	2369	1066
8	2636	1778	1082
9	4542	2028	593
10	4234	2285	862
11	4820	2233	742
12	4792	2188	835
13	5272	2163	796
14	5144	2369	873
15	4916	1888	516
16	3928	2588	1430
17	3538	2050	943
18	5064	2407	775
19	3580	1781	763
20	5244	2036	656
21	4020	2118	673
22	4064	2473	1225

TABLE VI
THE RESULTS OF MANUALLY DEVISING CIRCULATION PLANS

NO.1	L^{Cycle} /km	T^{Cycle} /min	Connection time/min
1	5272	2028	782
2	3333	1074	133
3	5272	2053	665
4	5244	2146	702
5	5442	2238	534
6	5442	2317	643
7	5272	2120	750
8	5144	2324	816
9	5272	1954	550
10	5070	2258	614
11	4020	2118	673
12	2636	1015	432
13	2527	2420	1347
14	2910	701	30
15	4308	2307	684
16	2636	734	35
17	4542	2028	593
18	3580	905	48
19	4682	1824	576
20	4792	2031	677
21	4916	1802	421
22	3580	1781	763
23	2014	1562	724

The above data leads to the conclusion that the two-day EMU circulation connection model allows for more flexibility in the design of the circulation plan and reduces the number of circulations to a certain extent, but it also increases the cost of the connection time and requires a combination of the two optimization objectives of the connection time and the number of circulations to be considered during the optimization. In addition, from Table IV, we can see that the accumulated mileage and time do not exceed the upper limit requirements, and the model satisfies the maintenance constraints in terms of both mileage and time cycles, so the mathematical optimization model and solution procedure proposed in this paper have good adaptability to the double maintenance requirements of the EMU.

TABLE VII
COMPARATIVE ANALYSIS OF PERFORMANCE INDICATORS FOR THE PROPOSED SOLUTIONS

NO.1	$\overline{L}^{Cycle}$ /km	$\overline{T}^{Cycle}$ /min	Connection time/min	Cumulative mileage efficiency
Manual Circulation Plan	4256.78	1814.78	573.57	0.851
Simulated Annealing Scheme	4450.27	2173.68	876.05	0.890
Python+Gurobi Scheme	4895.3	2139.85	712.45	0.979
Difference from manual plan	638.52	325.07	138.89	0.128
The difference with the SA Scheme	445.03	-33.83	-163.6	0.089

Note: The accumulated mileage efficiency in the table indicates the utilization of the mileage of the first level of maintenance.

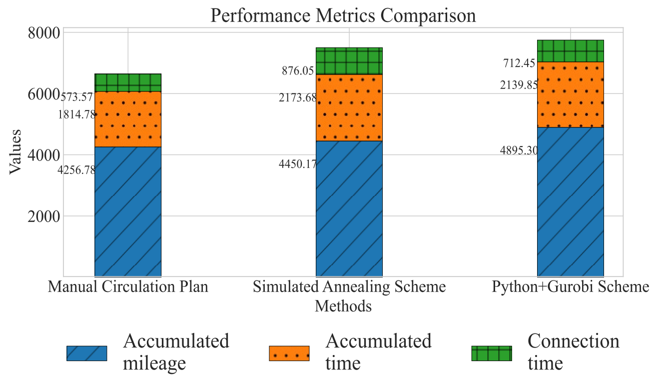
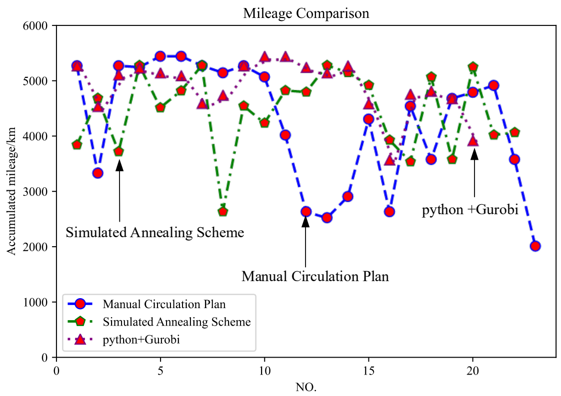
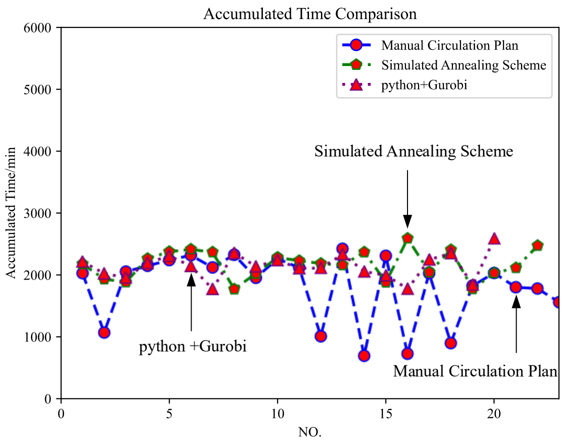


Fig. 8. Comparative analysis of solution results

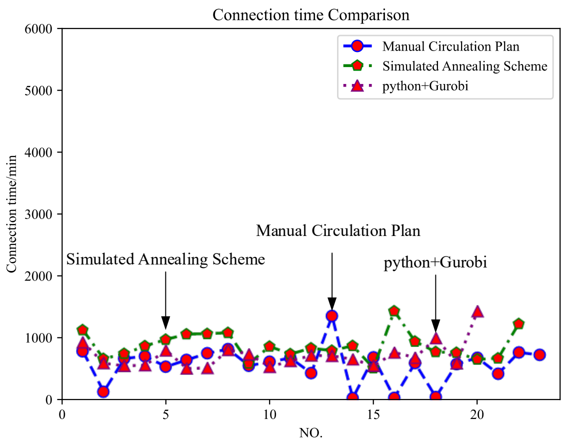
Figs. 8, 9, and 10 present a comparison analysis of the results obtained from the three algorithms mentioned above. From the figure, it is more obvious that the simulated annealing algorithm and the Python+Gurobi solution result in greater cumulative mileage of the EMU and more efficient use of the EMU compared to the manual plan. In addition, the model and solution method used result in an optimized solution with a total of 20 circulations, all of which are below the maximum value of the mileage cycle, which is fewer than the number of circulations found by the Simulated Annealing Scheme (22). This shows that the model and algorithm are feasible for reducing the number of circulations. Since reducing the number of circulations can decrease the number of trains and the maintenance cost of trains, based on the above index analysis, the model and solution method adopted is relatively better.



(a) Accumulated mileage



(b) Accumulated time



(c) Connection time

Fig. 9. Comparative analysis of solution results

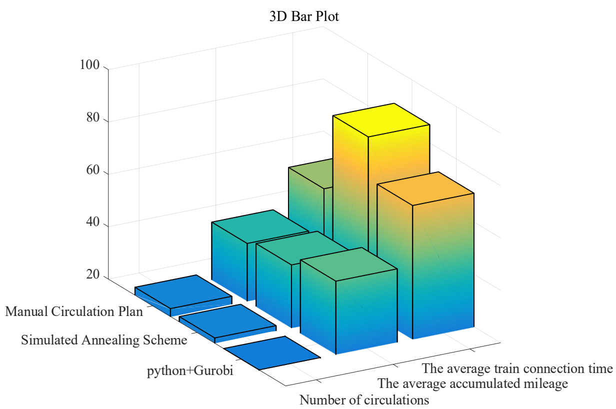


Fig. 10. Key indicator comparison chart

In Figure 11, three distinct datasets are showcased: Data 1 represents the cumulative mileage distribution resulting from manual planning; Data 2 illustrates the cumulative mileage distribution derived through the simulated annealing algorithm; and Data 3 presents the cumulative mileage distribution obtained by implementing the Python+Gurobi approach. Upon a meticulous comparison of these datasets, a conspicuous pattern emerges: the distribution of Data 3 is predominantly concentrated around the range of 5000, with a notable reduction in instances where the values significantly deviate from this threshold. This discernment signifies that, compared to alternative methodologies, the model and algorithm employed in this study have adeptly minimized the divergence from 5000 to a remarkable degree. This accomplishment in effectively harnessing the potential of feasible mileage has led to a substantial curtailment in unused mileage, thereby yielding highly commendable results in the realm of scheme formulation.

Furthermore, it is of paramount importance to acknowledge that the cumulative mileage results for all three methodologies comfortably adhere to the predetermined maximum mileage limit of 5500. This resoundingly underscores the real-world viability and applicability of these approaches.

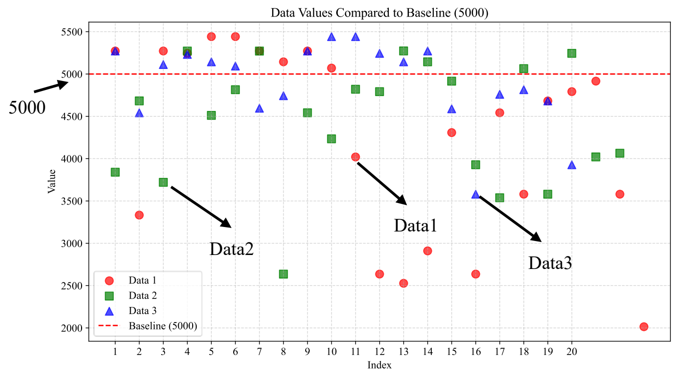


Fig. 11. Comparison chart of cumulative mileage disparities

Moreover, the results obtained also reveal that the manual circulation plan exhibits excessively lengthy connection durations. For instance, Figure 12 demonstrates that the EMU train necessitates 1317 minutes of connection time after serving train G7521 before proceeding to serve train G7536. This prolonged connection time results in wasted operational time and significantly diminishes the carrying capacity of the EMU train, indicating a low quality of the manual circulation plan. Conversely, the plan obtained using Python+Gurobi, as depicted in Figure 14, only requires a wait time of 485 minutes after serving train G7501, thereby significantly reducing the connection time. Hence, this model also demonstrates a positive impact from the perspective of connection time.

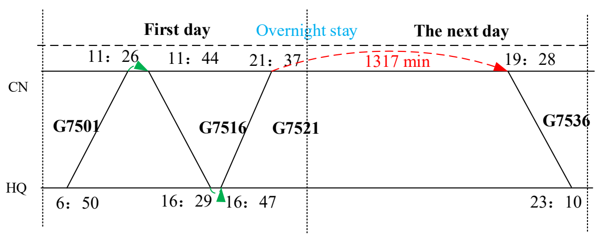


Fig. 12. Illustration of connection time in the manual circulation plan

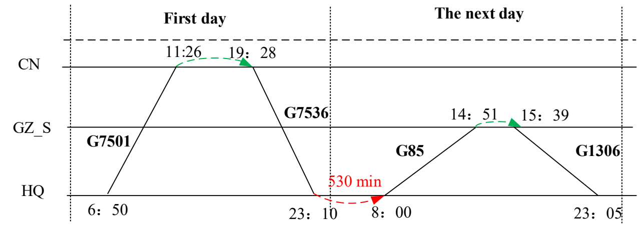


Fig. 13. Illustration of connection time in the Simulated Annealing scheme

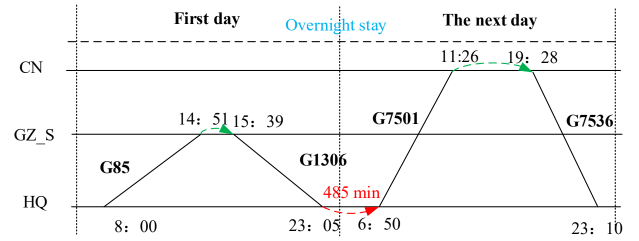


Fig. 14. Illustration of connection time in the Python+Gurobi scheme

As for cumulative mileage, cumulative time, and connection time, as shown in Table VIII, it can be observed that not only is the cumulative mileage substantial, but the turnaround time is also significantly reduced. In summary, the model constructed performs well in terms of both mileage and time.

TABLE VIII
COMPARATIVE ANALYSIS RESULTS OF INDICATORS IN THE EXAMPLE

NO.1	L^{Cycle} /km	T^{Cycle} /min	Connection time/min
Manual Circulation Plan	4815	2350	995
Simulated Annealing scheme	4815	2415	1060
Python+Gurobi	2527	2420	1347

V. CONCLUSION

This study analyzes train unit connections and maintenance operations by adopting the processing approach of the Traveling Salesman Problem. It comprehensively considers the interrelationships and constraints between decision variables within the dual maintenance system, such as maintenance constraints, uniqueness constraints, and connection condition constraints. A nonlinear integer programming mathematical model based on the train unit connection network is constructed to optimize train unit operations. The constraints have been linearized to allow for direct solving by mathematical solvers. A large-scale real-world case study involving 92 trains of the CRH380B series train units has been conducted to demonstrate the reliability and effectiveness of the solution method. The main conclusions obtained from the case study are as follows:

(1) The two-day connection mode allows for a more flexible design of the train route plan, thereby reducing the number of circulations and lowering maintenance costs to some extent. This approach is well-suited for train unit operations under the dual maintenance system.

(2) The model and design algorithm yield an average travel distance that is closer to the mileage cycle, with less under-maintenance mileage. The average travel distance increases by 639 km compared to manually drawn schemes and by 445 km compared to simulated annealing methods, indicating higher efficiency in train unit utilization. The comparison shows that the proposed optimization method can effectively optimize the EMU circulation plan under maintenance constraints.

This study primarily focuses on the optimization problem of the EMU circulation plan for a single train type and a single-train depot. In the future, further consideration will be given to optimizing train unit operations involving multiple train types, train depot maintenance capacity, and the presence of train unit coupling and decoupling operations.

REFERENCES

[1] E. Abbink, B. V. D. Berg, L. Kroon, and M. Salomon, “Allocation of railway rolling stock for passenger trains,” Transportation Science, vol. 12, no. 3, pp. 33-41, 2004.
[2] G. L. Giacco, A. D’Ariano, and D. Pacciarelli, “Rolling stock rostering optimization under maintenance constraints,” Journal of Intelligent Transportation Systems, vol. 18, no. 1, pp. 95-105, 2014.
[3] C. Wang, S. S. Wang, P. Wang, X. J. Wang, L. Chen, “A Study on the Optimization Model and Algorithm for the Integration of High-Speed Train Operating Diagram and EMU Utilization,” Railway Transport and Economy, vol. 40, no. 9, pp. 92-97, 2018.
[4] J. R. Miao, Y. Wang, and Z. X. Yang, “Research on the Optimization of EMU Circulation Based on Optimized Connecting Network,” Journal of the China Railway Society, vol. 32, no. 2, pp. 1-7, 2010.
[5] Y. Wang, J. Liu, and J. R. Miao, “Column generation algorithms based optimization method for maintenance scheduling of multiple units,” China Railway Science, vol. 31, no. 2, pp. 115-120, 2010.
[6] Q. W. Zhong, R. M. Lusby, J. Larsen, Y. X. Zhang, Q. Y. Peng, “Rolling stock scheduling with maintenance requirements at the Chinese High-Speed Railway,” Transportation Research Part B: Methodological, vol. 126, pp. 24-44, 2019.
[7] L. Cadarso, and Á. Marín, “Improving Robustness of Rolling Stock Circulations in Rapid Transit Networks,” Computers & Operations Research, vol. 51, pp. 146-159, 2014.
[8] J. Li, and B. Lin, “Optimizing Electric Multiple Unit Circulation Plan within Maintenance Constraints for High-Speed Railway System,” Mathematics. Optimization and Control, 2019. DOI: https://doi.org/10.48550/arXiv.1908.10099
[9] Y. Gao, M. Schmidtet, L. X. Yang, and Z.Y Gao, “A branch-and-price approach for trip sequence planning of high-speed train units,” Omega, vol. 92, pp. 102150, 2020.
[10] M. Y. Chen, Y. J. Zhang, S. S. Zhang, and H. F. Yan, “Operation Plan Optimization of Electric Multiple Units Based on the Minimum Number of Rolling Stock,” Railway Transport and Economy, vol. 44, no. 8, pp. 52-57, 2022.
[11] Y. Wang, Y. Gao, X. Y. Yu, I. A. Hansen, and J. R. Miao, “Optimization models for high-speed train unit routing problems,” Computers & Industrial Engineering, vol. 127, pp. 1273-1281, 2019.
[12] J. T. Haahr, J. C. Wagenaar, L. P. Veelenturf, L. G. Kroon, “A comparison of two exact methods for passenger railway rolling stock (re) scheduling,” Transportation Research Part E: Logistics and Transportation Review, vol. 91, pp. 15-32, 2016.
[13] B. Lin, and Y. A. Zhao, “Synchronized optimization of EMU train assignment and second-level preventive maintenance scheduling,” Reliability Engineering & System Safety, vol. 215, pp. 107893, 2021.
[14] J. R. Wei and Y. Gao, “Optimization of EMU Assignment and Maintenance Plan Based on Train Circulation Plan,” 2021 2nd International Conference on Urban Engineering and Management Science (ICUEMS), Sanya, China, pp. 219-222, 2021. DOI: 10.1109/ICUEMS52408.2021.00054
[15] Y. Gao, J. Xia, A. D’Ariano, and L. X. Yang, “Weekly rolling stock planning in Chinese high-speed rail networks,” Transportation Research Part B: Methodological, vol. 158, pp. 295-322, 2022.
[16] W. J. Li, Y. T. Wang, L. Nie, Y. H. Wu, and L. Peng, “High-Speed Railway EMUs’ Circulation Plan Optimization: A Two-Stage Optimization Algorithm Based on Column Generation,” in IEEE Access, vol. 8, pp. 15611-15631, 2020.
[17] S. Tong, S. K. Chen, G. H. Liu, S. Liu, B. H. Mao, “Optimum Circulation Model for Electrical Multiple Units Considering Reliability of Their Operational Connections,” Journal of the China Railway Society, vol. 42, no. 4, pp. 27-34, 2020.
[18] Y. Zhou, L. S. Zhou, and Y. Wang, “Using improved ant colony algorithm to investigate EMU circulation scheduling problem,” Discrete Dynamics in Nature and Society, vol. 2014, pp. 1-14, 2014.

[19] Y. Zhou, L. S. Zhou, Y. Wang, Z. Yang, and J. W. Wu, "Application of multiple-population genetic algorithm in optimizing the train-set circulation plan problem," Complexity, vol. 2017, pp. 1-14, 2017.

[20] Z. K. Wang, T. Y. Shi, W. J. Zhang, and H. Zhang, "Model and Algorithm for the Integrative Scheduling of EMU Utilization Plan and Maintenance Plan," China Railway Science, vol. 33, no. 3, pp. 102-108, 2012.

[21] B. L. Lin, Y. A. Zhao, J. Li, and R. X. Lin, "Fuzzy Programming Approach for the Electric Multiple Unit Circulation Planning Problem Using Simulated Annealing," Transportation Research Record, vol. 2676, pp. 456-467, 2022.