Metro Train Timetable Rescheduling Based on Q-learning Approach

Boyi Su, Zhikai Wang, Shuai Su* and Tao Tang

Abstract— **In metro system, unpredictable disturbances influence the normal operation and bring much inconvenience to passengers. This paper focuses on train timetable rescheduling (TTR) problem with considering practical operations in the metro management. At first, an optimization model that takes the deviation of rescheduled timetable, the total delay time of passengers and energy consumption as objective is developed. Meanwhile, the constraints, together with some practical rescheduling rules (e.g., the preprogrammed speed profiles, train detention strategy) are introduced. Secondly, the model is reformulated into an Markov decision process (MDP) with well defining the state, action and reward function, which is then solved by the proposed Q-learning approach. Finally, some case studies using the operational data of Beijing Yizhuang Subway Line are carried out to demonstrate the effectiveness of the proposed approach. The results indicate that a tradeoff solution among the optimization objectives can be obtained within a short time.**

Keywords: Train timetable rescheduling; Train detention strategy; Q-learning.

I. INTRODUCTION

With the rapid development of railway transportation in China, the passengers benefit a lot from the fast and convenient service of metro systems. However, the large passenger demand also brings a great pressure to operation management in some inevitable disturbance scenarios (e.g., infrastructure failures, equipment faults). If the headway time is short, the primary delay will easily knock-on from one train to others, resulting in unavailability of planned timetable. More importantly, the occurrence of train delays will influence service quality and passengers satisfaction. In this case, the planned timetable needs to be rescheduled in time to reduce the negative impact on passengers. The operations of the dispatcher including calling the drivers, noticing the staff in the station, recording the operation process are very complex, which may lower the efficiency of the rescheduling strategy. Therefore, it is necessary to design an algorithm to assist the dispatcher to deal with the train delays so that the normal operations can be restored from a disturbance quickly.

The study of train timetable rescheduling originated from the single-line operation management in the trunk railway network. In the 1970s, B.Szpigel [1] first used the linear programming method to determine the optimal position of overtaking and crossing operations when disturbance occurs. Cacchiani et al. [2] summarized the essence of train timetable rescheduling (TTR) problem as job shop scheduling model. Because this model presents to be a NP-hard problem, the complexity of computation increases rapidly with the increasing of model scale. In order to acquire the rescheduled timetable in real time, many scholars are devoted to develop an algorithm that can lead to shorter computation time. D'Ariano et al. [3] proposed a truncated branch and bound algorithm for real-time management. Then the computational experiments that include dozens of block sections and platforms were carried out to verify the performance of the algorithm. A tabu search algorithm that can solve the compound train rerouting and rescheduling problem within a few seconds was developed in [4]. The results of large instances showed that computation time was reduced by at least 15% compared with the previous algorithm in ROMA.

In recent years, urban metro train timetable rescheduling have attracted extensive attention owing to the short headway time, uncertain passenger flow and vulnerability to disturbances of metro system. Xu et al. [5] described the state transition of train positions as a series of discrete events, and proposed an efficient train rescheduling strategy for optimization model to balance service performance after disturbances. An an optimization model for a metro line in the case of disruptions was introduced in [6], where the overcrowded and time-variant passenger flow were considered. Moreover, the effectiveness and efficiency of a proposed iterative algorithm was verified by the numerical experiments. In [7], Yin et al. designed a stochastic programming model for train rescheduling, and the objective is to minimize the operational cost and delay time of passengers. Then an approximate dynamic programming (ADP) algorithm was proposed to solve the model. Hou et al. [8] further considered pre-programmed recommended speed profiles and the impact of the passenger flow on train dwell time, and developed a mixed integer programming model, which is solved on CPLEX.

When primary train delay caused by the disturbance arises, it may cause the following train to stop or run slowly in the section if they are not detained at the subsequent stations. From the perspective of passengers, it will bring much discomfort in case of lasting for a period of time. Hence, the dispatcher will try to arrange the trains to stop at stations instead such that the unnecessary panic is avoided. In addition, the adjustment of the train running and dwell time is another most commonly used rescheduling strategy with the preprogrammed speed profile in the automatic train operation (ATO) system, which corresponds to the running time one by one. Taking these practical rescheduling strategies into account, the detention strategy and operation level are firstly introduced in the a multi-objective TTR model in this paper.

B. Su, Z. Wang, S. Su and T. Tang are with the State Key Laboratory of Traffic Control and Safety, Beijing Jiaotong University, Beijing, P. R. China (e-mail:{2016112281@my.swjtu.edu.cn, wzk98417@163.com, shuaisu@bjtu.edu.cn, ttang@bjtu.edu.cn}). Corresponding author: S. Su.

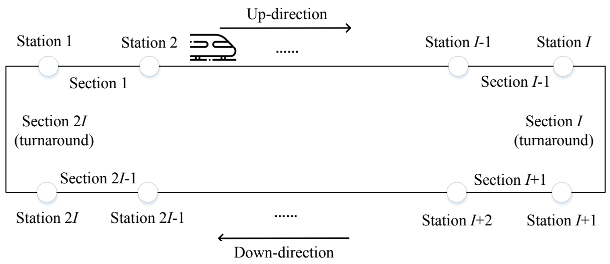


Fig. 1: The considered metro line in this paper

From a decision perspective, the model is formulated into a MDP model and a corresponding reinforcement learning approach is proposed.

The rest of the paper is organized as the following. In Section II, the metro train timetable rescheduling model is developed, which is reformulated into an MDP model in Section III, and a Q-learning approach is also proposed to solve the problem. In Section IV, some case studies based on the data of Beijing Yizhuang Subway Line are carried out. Some conclusions and future research are made in Section V.

II. Problem Statement

The structure of metro line that we consider is illustrated in Fig. 1, which contains up-direction, down-direction and a turnaround station. The stations and sections are denoted by $1, 2, ..., 2I$.

Some assumptions are made according to the characteristics of metro systems and practical operations.

1) Due to the single layout of metro systems, the station usually has no sidings. In this sense, Overtaking and crossing strategies are not allowed in this paper. In other words, all the trains run in a first-in-first-out manner and stop at every station.
2) The running time of a particular section corresponding to an operation level is a fixed value, since the speed profiles are preprogrammed in ATO system.
3) Regardless of external factors such as extreme weather, earthquakes and other situations that leading to complete interruption of operation.

A. The objective function

In this paper, we consider optimization objectives from three aspects, i.e., the deviation between the planned timetable and rescheduled timetable, the energy consumption and the total delay time of passengers.

The TTR problem firstly aims to resume normal operation order and planned timetable as soon as possible. $T_{\text{deviation}}$ is used to denote the sum of deviation between the planned and rescheduled timetable, which can be described as follows

$$T_{\text{deviation}} = \sum_{k=1}^{K} \sum_{i=1}^{2I} \left(|\tilde{T}_{k,i}^{\text{arrive}} - T_{k,i}^{\text{arrive}}| + |\tilde{T}_{k,i}^{\text{depart}} - T_{k,i}^{\text{depart}}| \right) \tag{1}$$

where $\tilde{T}_{k,i}^{\text{arrive}}$ and $T_{k,i}^{\text{arrive}}$ represent the planned and rescheduled time of train k arriving at station i, respectively; $\tilde{T}_{k,i}^{\text{depart}}$ and $T_{k,i}^{\text{depart}}$ represent the planned and rescheduled time of train k departing from station i, respectively; K is the sum of trains need to be rescheduled and $2I$ is the sum of stations and sections.

Secondly, the energy consumption of trains in a fixed section depends on the trip time which is uniquely determined by the operation level. Once the operation level is given, the energy consumed in this section is able to be confirmed. Then the total energy consumption E_{consume} can be given by

$$E_{\text{consume}} = \sum_{k=1}^{K} \sum_{i=1}^{2I} E_{k,i} \tag{2}$$

where $E_{k,i}$ is the energy consumption of train k in section i.

Train delays under disturbance will reduce the quality of service, which influences passenger satisfaction thereby. The delay time of onboard passengers arriving at destination T_{delay} is calculated as

$$T_{\text{delay}} = \sum_{k=1}^{K} \sum_{i=1}^{2I} \left[N_{k,i}^{\text{arrive}} \left(\tilde{T}_{k,i}^{\text{arrive}} - T_{k,i}^{\text{arrive}} \right) \right] \tag{3}$$

where $N_{k,i}^{\text{arrive}}$ is the number of passengers on train k that get off at station i.

In conclusion, the objective function was formulated to minimize the weighted sum of these indicators so as to reach a tradeoff between operational cost and service quality, i.e.,

$$\min Z = \omega_d T_{\text{deviation}} + \omega_t T_{\text{delay}} + \omega_e E_{\text{consume}} \tag{4}$$

where ω_d, ω_t and ω_e are weight coefficients, representing the importance of different indicators.

B. Constraints

1) Headway constraints: Certain headway time of trains must be met in order to ensure safety, which are guaranteed by the lower limits. These constraints will be satisfied as long as the arrival and departure time of adjacent trains at the same station are restricted [9], which can be written as

$$\tilde{T}_{k+1,i}^{\text{arrive}} - \tilde{T}_{k,i}^{\text{arrive}} \geq h_{\min}, \forall 1 \leq k < K, 1 \leq i \leq 2I \tag{5}$$

$$\tilde{T}_{k+1,i}^{\text{depart}} - \tilde{T}_{k,i}^{\text{depart}} \geq h_{\min}, \forall 1 \leq k < K, 1 \leq i \leq 2I \tag{6}$$

where $h_{\min}$ is the minimum headway time.

2) Running and dwell time constraints: As stated above, the preprogrammed speed profiles in the ATO system are considered here. Therefore, the running time is uniquely determined by operation level. Let $\delta_{k,i}^{l}$ and $\eta_{k,i}^{n}$ be the binary variables that represent whether the operation level l or n is chosen, the constraints for the running and dwell time are formulated as

$$\begin{gathered} \tilde{T}_{k,i+1}^{\text{arrive}} - \tilde{T}_{k,i}^{\text{depart}} = \sum_{l \in \mathbb{L}} \delta_{k,i}^{l} R_{k,i}^{l}, \\ \forall 1 \leq k \leq K, (1 \leq i < I) \cup (I+1 \leq i < 2I) \end{gathered} \tag{7}$$

$$\tilde{T}_{k,i}^{\text{depart}} - \tilde{T}_{k,i}^{\text{arrive}} = \sum_{n \in \mathbb{N}} \eta_{k,i}^{n} D_{k,i}^{n}, \forall 1 \leq k \leq K, 1 \leq i \leq 2I \tag{8}$$

$$\sum_{l\in\mathbb{L}} \delta^l_{k,i} = 1, \sum_{n\in\mathbb{N}} \eta^n_{k,i} = 1 \tag{9}$$

where $R^l_{k,i}$ denotes the running time for the operation level l and $D^n_{k,i}$ denotes the dwell time for the operation level n. Equation (9) indicates only one operation level can be chosen for train k in section i or at station i.

3) Turnaround constraints: After the train arriving at station I, the train will turnaround and change the running direction. Therefore, we use turnaround constraints to specify the time required for the train to return, which is described by

$$\tilde{T}^{\text{arrive}}_{k,I+1} - \tilde{T}^{\text{depart}}_{k,I} = t_{\text{turn}} \tag{10}$$

where t_{turn} is the turnaround time at station I.

4) Train capacity constraints: In this paper, we establish a passenger demand origin-destination (OD) matrix to predict the dynamic passenger flow and calculate the negative impact of train delays [10], which is defined as

$$\begin{cases} p_{i,j}(t) \geq 0, \text{if } (1 \leq i \leq I, 1 \leq j \leq I, i < j) \cup \\ \qquad\qquad (I+1 \leq i \leq 2I, I+1 \leq j \leq 2I, i < j) \\ p_{i,j}(t) = 0, \text{otherwise} \end{cases} \tag{11}$$

The train delays will inevitably lead to the overcrowded at platforms. In this case, the passengers on the platform may not be able to board entirely at once because of capacity constraints. Based on an OD matrix, the number of waiting passengers who can board train k at station i can be obtained. It is not only related to the arrival rate of passengers at station i, but also depends on whether there are passengers who cannot board the previous train. $T^w_{k,i}$ is defined as the time that when last passenger who board train k at station i enters this station, then the train capacity constraints can be formulated as follows

$$N_{k,i} = N_{k,i-1} - N^{\text{arrive}}_{k,i-1} + N^{\text{board}}_{k,i-1} \leq C, \forall 1 \leq i \leq 2I \tag{12}$$

with

$$N^{\text{board}}_{k,i} = \sum_{j=i+1}^{2I} \int_{T^w_{k-1,i}}^{T^w_{k,i}} p_{i,j}(t)dt, N^{\text{arrive}}_{k,i} = \sum_{j=1}^{i-1} \int_{T^w_{k-1,j}}^{T^w_{k,j}} p_{j,i}(t)dt \tag{13}$$

$$T^w_{k,i} = \min\left\{T^{\text{depart}}_{k,i}, \max\left\{t \middle| C \geq N_{k,i-1} - N^{\text{arrive}}_{k,i} + \sum_{j=1}^{i-1} \int_{T^w_{k-1,j}}^{t} p_{j,i}(t)dt \right\}\right\} \tag{14}$$

where C is the capacity of train; $N_{k,i}$ is the number of passengers onboard when train k arrives at station i; $N^{\text{arrive}}_{k,i}$ is the number of passengers that board train k at station i. In particular, $N_{k,0} = N^{\text{arrive}}_{k,0} = N^{\text{depart}}_{k,0} = 0$.

4) Kinetic constraints: For the purpose of facilitating calculation of energy, it is necessary to establish some kinetic constraints. The train operations in the section are generally divided into four phases, i.e., accelerating, cruising, coasting and braking, which are represented by $t^{l,a}_{k,i}$, $t^{l,c}_{k,i}$, $t^{l,o}_{k,i}$ and $t^{l,b}_{k,i}$. It is generally believed that only the first two phases have energy consumption. There are two kinds of resistances according to the generation method: basic resistance and additional resistance. In practical application, the basic resistance basically follows Davis formula [11]. Regarding additional resistance, the gradient factor is considered. So the traction force at time t can be obtained by

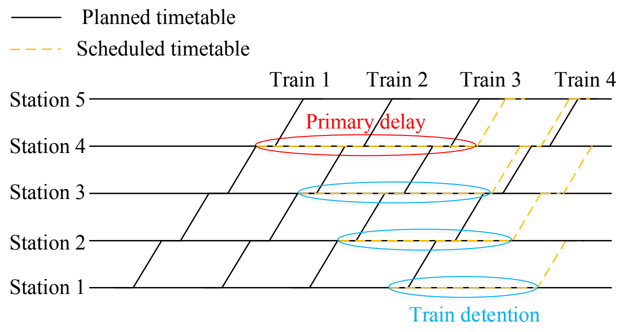


Fig. 2: An example of train detention

$$\begin{aligned} F^l_{k,i}(t) =& M_{k,i}\left(a^l_{k,i}(t) + g \cdot \sin\theta_{k,i}(t)\right) + R_0 + R_1 v^l_{k,i}(t) \\ &+ R_2 v^l_{k,i}(t)^2, \forall\ 0 \leq t \leq t^{l,a}_{k,i} + t^{l,c}_{k,i} \end{aligned} \tag{15}$$

where $M_{k,i}$ represents the train mass, which is equal to the sum of the net weight of the vehicle and the mass of the passengers on board; R_0, R_1 and R_2 are the coefficients of Davis formula; $a^l_{k,i}(t)$, $v^l_{k,i}(t)$ is the accelerated speed and speed of train k with speed level l at time t, respectively; $\theta_{k,i}(t)$ is the route gradient in section i when train k is running at time t.

Based on the mechanical power equation, the energy consumption can be calculated by the integral of traction force times speed [12], which is shown as

$$E_{k,i} = \delta^l_{k,i} \int_0^{t^{l,a}_{k,i}+t^{l,c}_{k,i}} F^l_{k,i}(t) v^l_{k,i}(t)dt, \forall 1 \leq k \leq K, 1 \leq i \leq 2I \tag{16}$$

C. Train Detention Strategy

The current methods of the dispatcher for dealing with delays include: (1) shorten running time in sections; (2) detain the tracking trains; (3) make use of backup trains; (4) let train turn back at the intermediate station. In this paper, we focus on TTR problem based on method (1) and (2). Through the above analysis, the running time can be reduced by adjusting operation levels. Next the train detention strategy is introduced.

In practical metro system, the tracking trains cannot adjust their speed in real time according to the primary delay. Thus, the following trains need to be detained by dispatcher at subsequent stations if the primary delay caused by the disturbance violates the security constraints. Fig. 2 describes an example that applies train detention strategies, in which the planned timetable is shown as the black solid line and the rescheduled timetable is shown as the orange dotted line. When train 1 is unable to depart from station 4 as the

TABLE I: Description of elements for MDP

Element	Notation	Definition
State	$S_i(k)$	The state variable that describes train k when it arrives at station i.
Action	$a_i(k)$	The action variable made by train k when it arrives at station i.
Reward function	$\mathrm{R}(S_i(k),a_i(k))$	The reward function for state $S_i(k)$ with action $a_i(k)$
State transition function	$\mathrm{S}^{\mathrm{M}}(S_i(k),a_i(k))$	The state transition function that depicts the next state after state $S_i(k)$ selects action $a_i(k)$

vehicle failure, train $2,3,4$ need to be detained at station $3,2,1$ correspondingly until the minimum headway time is satisfied.

As for how to determine whether the train needs to be detained, we pay attention the departure time of the previous train. A binary variable $\mu_{k,i}$ is defined as the flag of using train detention strategy, which is expressed as

$$\mu_{k,i} = \begin{cases} 1, \text{if } \tilde{T}_{k,i}^{\text{arrive}} + \max\left(D_{k,i}^{n} + R_{k,i}^{l}\right) < \tilde{T}_{k-1,i+1}^{\text{depart}} + h_{\min} \\ 0, \text{else} \end{cases} \tag{17}$$

If the train should be detained after judgment, the dwell time will be regulated accordingly. The tracking interval between trains should be maintained to be uniform in daily operation, even under the disturbance. Thus, the dwell time can be changed as follows

$$D_{k,i}^{n} = D_{k,i}^{n} + \mu_{k,i}\frac{K-k+1}{K}t_p, \forall 1 < k \leq K, 1 \leq i \leq 2I \tag{18}$$

In addition, if there is a train detained at the station 1, it is no longer possible to implement the train detention strategy, the time for the subsequent trains to arrive at this station should be adjusted as follows

$$\tilde{T}_{k,i}^{\text{arrive}} = T_{k,i}^{\text{arrive}} + \frac{K-k+1}{K}t_p \tag{19}$$

III. Train Timetable Rescheduling With Q-learning Approach

Q-learning approach is the content of reinforcement learning (RL), which is suitable to handle the multi-stage decision. To solve the above problem, we first transform the model into a Markov decision process (MDP) in this section for the convenience of understanding. Then a algorithm based lookup table is proposed to acquire the rescheduled timetable.

A. Fundamental Elements of MDP

MDP is a framework that describes the problem of learning from interactions. Decisions will be changed through trial and error under the interaction of the agent and environment. To be specific, after the agent selects the action, the environment updates and returns the rewards. The task of agent in this paper is to maximum these rewards. Next, the definitions of each element in MDP are introduced in detail, which are shown in Table I.

1) State and action: To approach the calculation of rewards and transition function, each state should have corresponding record information, which is called state variable. The state variable for train k at station i is defined as

$$S_i(k) = \left\{N_i^{\text{onboard}}(k), T_i^{\text{arrive}}(k)\right\}, \forall 1 \leq k \leq K, 1 \leq i \leq 2I \tag{20}$$

where the state variable $N_i^{\text{onboard}}(k)$ denotes the number of passengers onboard; $T_i^{\text{arrive}}(k)$ denotes the time when train k arrives at station i.

The actions correspond to decision variables of optimization model, i.e. the dwell time and running time of trains. Therefore, the action of each state is defined as

$$a_i(k) = \{r_i(k), d_i(k)\}, \forall 1 \leq k \leq K, 1 \leq i \leq 2I \tag{21}$$

where $r_i(k)$ represents the dwell time of train i at station k, $d_i(k)$ represents the running time of train i before station k. Specifically, $r_I(k)$ and $r_{2I}(k)$ are the turnaround time.

2) Reward function: Energy consumption, timetable deviation and delay time of passengers considered as the objective of train timetable rescheduling problem, formulated by equation (4). As a result, the reward for state $S_i(k)$ and action $a_i(k)$ is given by

$$\begin{gathered}\mathrm{R}(S_i(k),a_i(k)) = -\left(\omega_d T_{k,i}^{\text{deviation}} + \omega_t T_{k,i}^{\text{delay}} + \omega_e E_{k,i}\right), \\ \forall S_i(k) \in \mathbb{S}, a_i(k) \in \mathbb{A}\end{gathered} \tag{22}$$

where $\mathbb{S}$ is the set of all possible states and $\mathbb{A}$ is the set of all feasible actions can be selected.

3) State transition function: The state transition function S^M involves the change in number of passengers onboard and nodes of trains. The rule of passenger transition is described by equation (12). The nodes of trains refer to not only the change in indexes of stations, but also the time that train k arrives at station i, which is written as

$$\begin{gathered}T_i^{\text{arrive}}(k) = T_{i-1}^{\text{arrive}}(k) + r_{i-1}(k) + d_{i-1}(k), \\ \forall 1 \leq k \leq K, 1 < i \leq 2I\end{gathered} \tag{23}$$

B. Solution methodology

Based on Q-learning approach, we design an effective algorithm to solve the train timetable rescheduling problem proposed in this paper. Q-learning is a method of approximate dynamic programming, which uses an approximate structure to approach the value function in the dynamic programming equation so that Bellman optimality principle is satisfied. Here, we introduce an approximate technique using lookup table iteratively and an off-policy, that is, the learned value function is independent of the policy being followed.

A matrix is built to express lookup table, whose column is the different state and row is one of the permutations of all possible actions. Then a ε-greedy policy is used to select actions. Once the agent selects an action, the corresponding location of the lookup will be updated as follows

$$\begin{aligned}Q(S_i(k),a_i(k)) = & Q(S_i(k),a_i(k)) + \alpha \Big[\mathrm{R}(S_i(k),a_i(k)) + \\ & \gamma \cdot \max_{a \in \mathbb{A}} Q(S_{i+1}(k),a_i(k)) - Q(S_i(k),a_i(k))\Big]\end{aligned} \tag{24}$$

Algorithm 1 Train timetable rescheduling algorithm based on Q-learning approach

Step 1. Input planned timetable; system parameter initialization; set $n = 1$.
Step 2. Repeat for $k = 1, 2, \ldots, K$.
Step 2.1. Do for $i = 1, 2, \ldots, 2I$.
Step 2.2. Find feasible actions according to constraints.
Step 2.3. With the probability ε,choose the action $a_i^n(k)$ from the set of feasible actions randomly. With the probability $1-\varepsilon$, choose $a_i^n(k)$ by

$$a_i^n(k) = \arg\max_{a \in \mathbb{A}} Q\left(S_i^n(k), a_i^n(k)\right).$$

Step 2.4. Update the lookup table by equation (24).
Step 2.5. Calculate the state of the next stage according to $\mathrm{S}^{\mathrm{M}}\left(S_i(k), a_i(k)\right)$.
Step 3. Set $n \leftarrow n+1$. If $n \leq N$, return to step 2.
Step 4. Output the lookup table and optimal action.

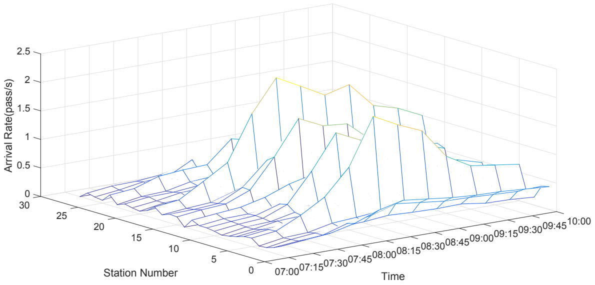


Fig. 3: The profile of passenger OD data

where α is the learning rate; γ is the discount factor, which can be understood as the importance of future steps.

The complete summary of algorithm procedure is illustrated in algorithm 1.

IV. CASE STUDY

In this section, the data of Beijing Yizhuang Subway Line are used to conduct some case and verify the effectiveness and efficiency of algorithm designed. Due to the large traffic of passenger in early peak hour, the impact of train rescheduling is significant during that time. The horizon of timetable that we consider is from 7:00 to 10:00. There are five operation levels in ATO system and five recommended dwell times for train to choose. The profile of passenger origin-destination data is shown in Fig. 3.

For the simulation, the parameters are given as follows: net weight of train $M_n = 2 \times 10^5$ kg / train; average weight of passenger $M_p = 60$ kg / passenger; maximum capacity $C = 1500$; coefficients of Davis formula $R_0 = 3.48$, $R_1 = 0.144$, $R_2 = 0.085$; learning rate $\alpha = 0.02$; discount factor $\gamma = 0.88$;

A. Computational Results With Different Delay Scenarios

In this part, we fix weight coefficients and observe the performance of algorithm. In order to ensure that all indicators are at the same order of magnitude, weights are set as: $\omega_d = 1 \times 10^5$, $\omega_t = 1 \times 10^3$, $\omega_e = 5$.

TABLE II: The objective with different weight coefficients

Weight value	Deviation(s)	Delay time(s)	Energy consumption(J)
$\omega_d = 3 \times 10^5$	25689	2.3896×10^6	6.8827×10^8
$\omega_d = 5 \times 10^5$	24439	2.2466×10^6	7.4199×10^8
$\omega_d = 7 \times 10^5$	23640	2.1378×10^6	7.8827×10^8
$\omega_t = 2 \times 10^3$	25568	2.4062×10^6	6.5559×10^8
$\omega_t = 3 \times 10^3$	24974	2.3324×10^6	6.9460×10^8
$\omega_t = 4 \times 10^3$	24213	2.2724×10^6	7.2487×10^8
$\omega_e = 6$	26431	2.4159×10^6	6.8598×10^8
$\omega_e = 7$	27238	2.5370×10^6	6.5544×10^8
$\omega_e = 8$	27604	2.6149×10^6	6.5047×10^8

The delay scenarios conducted are the 100s primary delay of train 5 in station 5 (Jinghailu - Tongjinanlu) and the 400s primary delay time of train 5 at station 5 (Tongjinanlu). The simulation results are presented in Fig. 4, where only one train is rescheduled in the first scenario, while several following trains are detained and rescheduled in the second scenario. It is verified that normal operation order and planned timetable can be resumed. As shown in Fig. 5, the objective can converge to a related stable value after about 1000-3000 iterations. Meanwhile, the computation time is within 1 minute.

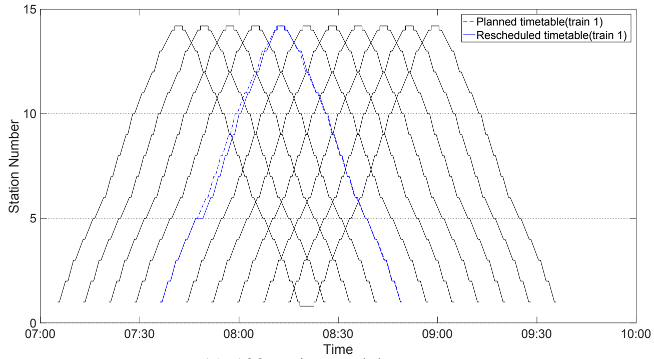
B. Computational Results With Different Weight Coefficients

Since the solution is tradeoff among the deviation between rescheduled timetable and planned timetable, the delay time of passengers and energy consumption, the different weight coefficients will influence the importance of factors. Thus, comparative experiments are made to prove this idea. Note that only test parameter is changed at a time, which means other parameters are fixed. The second delay scenario and initial weight in section IV-A are used to test.

Table II shows the results of variable controlled experiments. With the increase of the weight coefficient, the objective value of its corresponding term decreases to varying degrees. That is to say, the dispatcher can set different weight coefficients according to demand. For example, there are more passengers during peak hours, if disturbance occurs at this time, dispatcher can increase appropriately to improve service quality and reduce the negative impact on passengers.

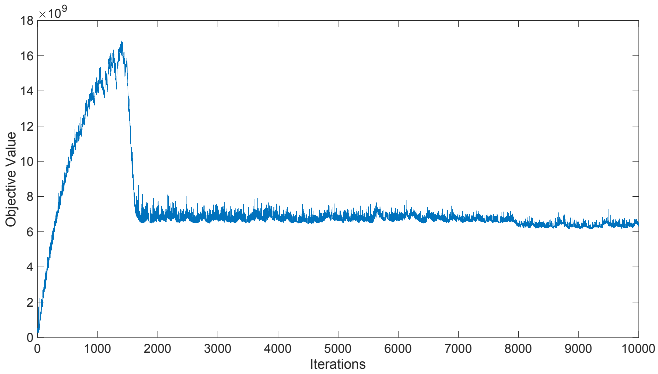
V. CONCLUSION

In this paper, we have studied the train timetable rescheduling problem with train detention strategy under disturbance. By taking speed profiles in ATO system, time-varying passenger flow into consideration, we developed a optimization model to generate a tradeoff between operational cost and service quality. A Q-learning approach is proposed to solve the multi-stage decision optimization problem. Some case studies based on Beijing Yizhuang Subway Line were carried out, which shows that a good solution can be obtained within a short time. Meanwhile, the weight coefficients can be changed according to actual



(a) 100s primary delay case (b) 400s primary delay case

Fig. 4: Planned and rescheduled timetable



(a) 100s primary delay case (b) 400s primary delay case

Fig. 5: Convergence process of the proposed algorithm

demand so that the algorithm can be applied to different delay scenarios.

In future research, other strategies used by the dispatcher to deal with disturbance like making use of backup trains and letting train turn back at the intermediate station will be considered.

Acknowledgment

This work was supported by the State Key Laboratory of Rail Traffic Control and Safety, Beijing Key Laboratory of Urban Rail Transit Automation and Control, National Natural Science Foundation of China (No. U1734210,U1734211), the State Key Laboratory of Rail Traffic Control and Safety (Contact No. RCS2020ZZ004).

References

[1] B. Szpigel, "Optimal train scheduling on a single track railway," *Operations Research*, vol. 72, pp. 343–352, 1972.
[2] V. Cacchiani, D. Huisman, M. Kidd, L. Kroon, P. Toth, L. Veelenturf, and J. Wagenaar, "An overview of recovery models and algorithms for real-time railway rescheduling," *Transportation Research Part B: Methodological*, vol. 63, pp. 15–37, 2014.
[3] A. D'Ariano, D. Pacciarelli, and M. Pranzo, "A branch and bound algorithm for scheduling trains in a railway network," *European Journal of Operational Research*, vol. 183, no. 2, pp. 643–657, 2007.
[4] F. Corman, A. D'Ariano, D. Pacciarelli, and M. Pranzo, "A tabu search algorithm for rerouting trains during rail operations," *Transportation Research Part B: Methodological*, vol. 44, no. 1, pp. 175–192, 2010.
[5] X. Xu, K. Li, and L. Yang, "Rescheduling subway trains by a discrete event model considering service balance performance," *Applied Mathematical Modelling*, vol. 40, no. 2, pp. 1446–1466, 2016.
[6] Y. Gao, L. Kroon, M. Schmidt, and L. Yang, "Rescheduling a metro line in an over-crowded situation after disruptions," *Transportation Research Part B: Methodological*, vol. 93, pp. 425–449, 2016.
[7] J. Yin, T. Tang, L. Yang, Z. Gao, and B. Ran, "Energy-efficient metro train rescheduling with uncertain time-variant passenger demands: An approximate dynamic programming approach," *Transportation Research Part B: Methodological*, vol. 91, pp. 178–210, 2016.
[8] Z. Hou, H. Dong, S. Gao, G. Nicholson, L. Chen, and C. Roberts, "Energy-saving metro train timetable rescheduling model considering ato profiles and dynamic passenger flow," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 7, pp. 2774–2785, 2019.
[9] S. Su, X. Wang, Y. Cao, and J. Yin, "An energy-efficient train operation approach by integrating the metro timetabling and eco-driving," *IEEE Transactions on Intelligent Transportation Systems*, 2019. doi:10.1109/TITS.2019.2939358.
[10] J. Yin, L. Yang, T. Tang, Z. Gao, and B. Ran, "Dynamic passenger demand oriented metro train scheduling with energy-efficiency and waiting time minimization: Mixed-integer linear programming approaches," *Transportation Research Part B: Methodological*, vol. 97, pp. 182–213, 2017.
[11] S. Su, T. Tang, and X. Li, "Driving strategy optimization for trains in subway systems," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 232, no. 2, pp. 369–383, 2018.
[12] S. Su, X. Li, T. Tang, and Z. Gao, "A subway train timetable optimization approach based on energy-efficient operation strategy," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 2, pp. 883–893, 2013.