Applied Mathematical Modelling 38 (2014) 894–909

Contents lists available at SciVerse ScienceDirect

Applied Mathematical Modelling

journal homepage: www.elsevier.com/locate/apm

Balanced train timetabling on a single-line railway with optimized velocity

Xiaoming Xu [a], Keping Li [a], Lixing Yang [a,*], Jingjing Ye [b]

[a] State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China
[b] School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China

ARTICLE INFO

Article history:
Received 27 September 2012
Received in revised form 20 March 2013
Accepted 4 July 2013
Available online 23 July 2013

Keywords:
Train scheduling
Discrete event
Genetic algorithm
Travel advanced strategy

ABSTRACT

This paper aims to find an optimal balanced schedule with the least delay-ratio (i.e., the ratio of the total delay time and the total free-run time) by considering the impacts of the train velocity. A rigorous optimization model is proposed under the consideration of feasible speed constraint for finding the optimal velocity for each train on the railway line. To obtain an approximately optimal scheduling strategy, a combination of the improved TAS (ITAS) method and the genetic algorithm (GA), called GA-ITAS method, is in particular proposed to effectively solve the proposed model. The results of numerical experiments demonstrate the efficiency and effectiveness of the proposed approaches.

1. Introduction

Train scheduling, which focuses on specifying the departure and arrival times at each station, is required to satisfy a set of operational constraints, including the platform capacity constraints, headway constraints, velocity constraints, etc. Due to the complexity of these constraints, train scheduling problem then becomes a challenging and difficult problem for the real-world applications, especially for real-time scheduling. In general, two types of approaches can often be adopted to generate a feasible train timetable in the real-world applications, i.e., the manual method and computer-based method. In comparison, computer-based method has attracted more attentions in recent decades from the researchers and engineers with the development of computer techniques, and two scheduling techniques, namely mathematical optimization and simulation methods, have been proposed for generating a feasible or near-optimal schedule.

The majority of the researches in literature focus the scheduling and rescheduling methods on the mathematical optimization method, which was firstly applied to train scheduling problem by Amit and Goldfarb [1] in 1971. Three types of rail traffic environments, i.e., the single-line railway [2–6], multiline railway and railway network [7–10] have been often taken into account. For scheduling trains, for instance, on the single-line railway, Szpigel [4] firstly formulated the train scheduling problem as a mixed-integer program and applied a branch-and-bound algorithm to solve the problem. Zhou and Zhong [6] developed a branch-and-bound solution procedure to obtain feasible schedules with guaranteed optimality considering the constraints of the railway resource. Higgins et al. [2] presented a two-objective optimization model which minimizes the delay time and fuel consumption cost. And the model was developed as a decision support tool for dispatchers to optimize train schedules on a single line in real-time. Besides, an expected value programming model proposed by Yang et al. [5] to minimize the train delay time and the total passenger-time, in which the number of passengers getting on/off the train at each station is assumed to be a fuzzy variable. In recent years, some researchers turn to a more complex situation with

* Corresponding author.
E-mail addresses: xuxiaomingnch@126.com (X. Xu), kpli@bjtu.edu.cn (K. Li), lxyang@bjtu.edu.cn (L. Yang).

0307-904X/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.apm.2013.07.023

the development of computer techniques, namely, scheduling on the railway network. For this topic, Carey [7], Carey [11] and Carey and Lockwood [8] developed an iterative decomposition approach for solving the train timetable and path problem on the rail lines. In this approach, some solution heuristics and strategies, with which dispatchers can plan large-scale complex rail systems in an acceptable time, were proposed. Brannlund et al. [12] presented a heuristic optimization approach for the timetabling problem where Lagrangian relaxation solution approach was designed. Moreover, Ghoseiri et al. [9] developed a train scheduling model for minimizing the fuel consumption cost and the total passenger-time. Likewise, considering the global information of the railway network, Yang et al. [10] proposed an efficient mathematical model to find the optimal train movements strategy, which aims to minimize both energy consumption and travel time simultaneously.

It is worth pointing out that the train scheduling problem is essentially an integer programming problem known to be NP hard (see Cai and Goh [13]). Then, the above-mentioned mathematical programming models for train scheduling problem are usually solved by some special heuristics or modified algorithms. According to the statement by Dorfman and Medanic [14], two drawbacks mainly exist in the mathematical programming methods: (1) despite the tremendous speed of today's supercomputers, an integer programming problem with constraints for every train and siding in a railway network still cannot be resolved in an acceptable time; (2) once such a solution is implemented, whenever a single train is unable to keep the schedule, the entire problem must be rescheduled from the current state of the network. Due to these drawbacks, Medanic and Dorfman [15] proposed an effective simulation approach to schedule trains by considering the combination of local, state-dependent, travel-advance strategy (TAS) and discrete-event model, which can produce sub-optimal time-efficient and energy-saving schedules rapidly. In their recent works, they also extended the aforementioned approach to solve the large-scale real-world train scheduling problem. To further improve the quality of the generated schedule, Li et al. [16] proposed a modified TAS method by introducing some modified rules which aim to characterize train's micro operation based on global conflicts information. Also, they incorporated the simulation of train's acceleration and deceleration into their work after analyzing the train's micro movement features. To search for a good train schedule, Pudney and Wardrop [17] proposed a probabilistic search technique, called problem space search, with which hundreds of different train schedules can be quickly generated. As for the other simulation techniques, interested readers may refer to Frank [18], Peat et al. [19], Rudd and Storry [20], Petersen et al. [21], Sahin [3], Sauder and Westerman [22] and Jovanovic and Harker [23].

In literature, few researches allowed different train speeds on different sections such as Mills et al. [24]. The majority of the researches on train scheduling problem usually assume that the velocity of the train on each railway link is a pre-specified constant (see [14,16,5,6]). In general, although this assumption is much easy to be handled in the scheduling process, the generated timetable may probably correspond to some non-balance characteristics, e.g., causing more waiting time at stations or more delays on the railway line. For instance, according to the original timetable, a train is scheduled to arrive at an intermediate station and then wait a long time for another train's meeting and crossing. For this case, it is generally expected to further reduce this train's actual waiting time as much as possible if its velocity is allowed to be reduced, and what's more, this modification essentially produces two advantages, i.e., (1) decreasing the delay time at the stations and (2) saving the energy consumption according to Davis Equation because of the reduction of the velocity (see [9]). On the other hand, this non-balance feature (waiting time) can also be canceled by enhancing velocity of the opposite train at the cost of more energy consumption. Practically, this type of non-balance feature essentially can be attributed to the coupling effects of the non-optimal velocities of trains. With this concern, we are particularly interested in generating a balanced schedule through optimizing train's velocity with the least delay-ratio. To the best of our knowledge, this is a new idea in literature.

The remainder of this paper is organized as follows. For the completeness of this research, the discrete-event model proposed by Dorfman and Medanic [14] and Medanic and Dorfman [15], is firstly reviewed in Section 2. In Section 3, train scheduling problem with speed constraints is described in detail. To solve the model, a new scheduling algorithm, which is combined by our improved TAS and genetic algorithm and denoted by GA-ITAS, is proposed to find an approximate optimal schedule. The efficiency and the effectiveness of the proposed approaches are demonstrated in Section 4. Finally, a conclusion is made in Section 5.

2. Reviews of the discrete-event model

As we aim to propose a genetic algorithm based on our improved TAS to generate train schedules, a brief review of TAS will be given in this section for the completeness of this research.

Based on the concept of travel advanced strategy (TAS), an efficient train scheduling generation approach was proposed by Dorfman and Medanic[14] and Medanic and Dorfman [15]. In this approach, the involved trains on the railway network are firstly classified into two categories, i.e., trains on sections and trains at stations. A discrete event is defined through the arrival of a train at a station. That is, a discrete event occurs once a train arrives at a station. At the time that an event occurs, a capacity checking algorithm will then be employed to detect the state of the traffic system to avoid the possible future deadlock (A segment is said to be in a state of deadlock if no train on a segment of railway line can advance without causing a collision). In the capacity checking algorithm, for each train at a station, local and non-local information will be used to determine the capacity of the rest rail line of the focus train's trip. In the case of enough capacity, the focus train can be scheduled to continue its journey; otherwise, it is required to stop at its current station until the next discrete event occurs.

Table 1
The comparison of different characteristics of closely related methods.

Characteristic	Dorfman and Medanic [10]	Li et al. [18]	This paper
Status of train velocity	A constant	A constant, considering the acceleration and deceleration	A decision variable
Feature of TAS	Original rules	Improved rules with global information	Improved rules on traversing order
Algorithm	TAS (Heuristic method)	ETAS (Heuristic method)	GA-ITAS (Genetic Algorithm + Heuristic method)
Feature of solution	Not consider coupling effects of trains	Not consider coupling effects of trains	Consider coupling effects of trains

In general, the train on a section is always allowed to continue its travel at the current section until it arrives at the next station. Integrating with the operational rules above, the TAS method can schedule trains quickly and efficiently on a railway network under the consideration of three performance criteria, i.e., the time to clear the line, the delay of all trains, and the maximal delay. Additionally, TAS has been proved to be much faster than exact methods such as branch and bound algorithm [16].

It is interesting to mention that the TAS method has two implicit drawbacks according to the statements in Li et al. [16]: (1) the local Greedy TAS cannot obtain the mathematically optimal solution and can only produce a feasible train scheduling plan; (2) the collisions among the trains can be categorized into two classes: relative deadlocks and absolute deadlocks, which should be given the separate treatments. For the first drawback, it is easy to explain. As mentioned above, TAS is a greedy algorithm which ensures that there is always at least one free track at each station at the end of every movement for trains. This mechanism may probably prevent some optimal solutions. To further decrease the potential delay of the trains, Li et al. [16] focused their researches on the second problem, and proposed an improved approach, namely effective travel-advance strategy (ETAS) method based on the global scheduling information. ETAS method is proved to be a more efficient approach to schedule trains in a railway network. However, these two deterministic methods (TAS and ETAS) for generating a train schedule are based on a nominal speed for each train. In practice, trains can be driven faster or slower than the nominal speed. If we utilize this fact, we may be able to find a better train schedule.

In this research, based on the genetic algorithm (GA) and our improved TAS (ITAS) method, we shall design a new algorithm, called GA-ITAS, to search for a better train timetable with the minimal delay-ratio through varying the speeds of trains within reasonable limits. The proposed algorithm has the following advantages: (1) The TAS method used in GA-ITAS is essentially a new version that is improved based on the TAS proposed in Ref. [14], and it is more efficient than the original one. (2) To the best of our knowledge, relaxing some system constraints in the railway system can be expected to take advantages of the railway resources as much as possible (for instance, decreasing the waiting time). Therefore, focusing on decreasing the total delay-ratio, we particularly treat the train velocity as a decision variable to further reduce the coupling effects among the trains. Compared with some closely related researches in literature, we give some detailed characteristics analysis with respect to different works in Table 1 to clarify the contributions of our research.

3. Problem statements

Consider a railway line composed of a sequence of two-way single-track railway links and stations, in which two adjacent stations are linked by a railway link (e.g., the link *j* links the station C and station D, as shown in Fig. 1). On the railway, the involved stations can be categorized into "intermediate" stations (see station C and station D in Fig. 1) and terminal stations (see station A and station B in Fig. 1), where the former allow three trains to dwell (the structure is illustrated in Fig. 2) and the latter have multiple platforms at which trains can originate and terminate.

To clear show the formulation of this research, an illustration in Fig. 3 is especially given to state the importance of the velocity optimization. As shown in Fig. 3a, train 1 is scheduled to travel in outbound direction and meet with other two

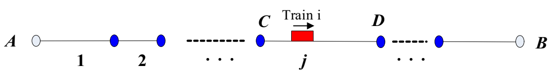

Fig. 1. An illustration of a two-way single-line railway.

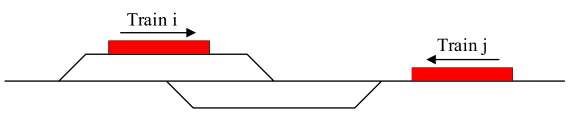

Fig. 2. Two trains meet at a station.

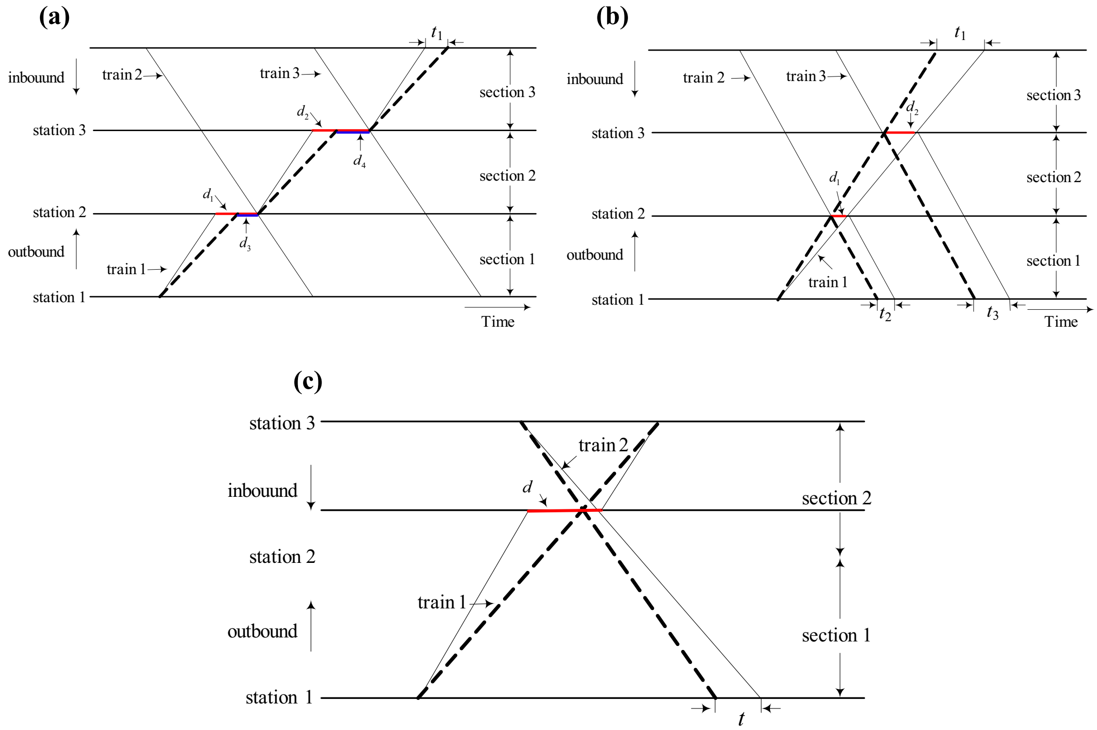

Fig. 3. (a) Case of reducing a train's velocity; (b) case of enhancing a train's velocity; and (c) case of changing velocities of different trains (the original schedule is represented by solid lines).

inbound trains, i.e., trains 2 and 3, at the second and third stations. Due to the pre-specified velocities, train 1 has to wait at these two stations for crossing operations, which leads to the extra delay time d_1 and d_2. In practice, these delays can be expectedly cancelled through adjusting the velocity of train 1. Specifically, if we can change the trajectory of train 1 to the dotted line through decreasing its velocity, fewer waiting time, i.e., d_3 and d_4, appears in the rescheduled train timetable, and meanwhile, the energy consumption cost of train 1 further decreases due to the reduction of the velocity according to Medanic and Dorfman [15]. Likewise, after enhancing the velocity of train 1 in Fig. 3b, delay time d_1 and d_2 can be eliminated automatically with respect to two inbound trains. In this illustration, it is easy to see that unnecessary delay time (i.e., d_1 and d_2) can be effectively reduced by adjusting the pre-specified speed on the links. Correspondingly, the total delay on the railway network can also be potentially optimized in comparison to the original schedule with the fixed speed on each link.

Hereinafter, it is interesting and necessary to discuss the relationship between the change of the total delay time and the change of the total travel time. As shown in Fig. 3a, with the velocity adjustment, the reduction of the total delay time (i.e., $(d_1 + d_2) - (d_3 + d_4)$) is a positive number. However, the total travel time is increased by t_1. For this case, train 1 reduces its speed to decrease the waiting time at stations, but potentially increase its total travel time. In Fig. 3b, comparing the adjusted and original schedules, the reduction of the total delay time is $d_1 + d_2$, and the change of the total travel time is $t_1 + t_2 + t_3$. Clearly, $d_1 = t_2$ and $d_2 = t_3$. Therefore, the reduction of the total travel time is larger than that of the total delay time. In Fig. 3c, velocities of trains 1 and 2 are reduced and enhanced, respectively. Since the values of t and d are unknown. So it is unclear that whether the reduction of the total travel time is larger than that of the total delay time. Note that, in Fig. 3c, though train 1 reduces its velocity, its total travel time still remains unchanged.

As a conclusion, there is no specific relationship between the total travel time and the total delay time. In fact, from the perspective of the psychology, rather than wait in trains, passengers would be more willing to travel with slow velocity.

3.1. Formulation

As stated above, velocity optimization plays an important role in generating a high-performance schedule for the railway traffic. A well-designed velocity planning can not only decrease the potential conflicts of the journey but also reduce the coupling effects between different trains. In view of this aspect, we shall particularly consider the train scheduling problem with velocity optimization to obtain a balanced timetabling plan. For formulating convenience, several assumptions are firstly made as follows:

(A1) We are intended to find a constant velocity for each train to minimize the delay-ratio. For simplicity, such velocity is called as *optimal velocity* in our paper.
(A2) Each intermediate station has only one main track and two side tracks, which states at most three trains can stop at each station simultaneously.
(A3) Trains can pass or overtake at any station, provided that there is sufficient capacity at the station.
(A4) The first train starts at time $t_0 = 0$.
(A5) In our research, no dwell time and time headway will be scheduled for simplicity.

The following discussion focuses on setting up a rigid formulation for the train scheduling problem such that the least delay-ratio can be achieved through seeking an optimal velocity for all the trains on their trips. To this end, the decision variables, the system constraints and the objectives, respectively, will be specified below.

(1) Decision Variable

As the aim of this research is to find an optimized velocity for each train on its journey, the velocity variable is then treated as the decision variable. That is,

v_i: the velocity of the ith train on its journey,

$V = (v_i)_{1\times N}$: velocity vector consisting of velocity v_i for each train.

(2) System Constraints

To find an optimal velocity in the rail traffic system, this paper will relax the fixed velocity constraint in TAS (denoted by v_i^c) to a selection interval $[v_i^c - \Delta v, v_i^c + \Delta v]$ identifying the range of the possible optimal velocity. Therefore, the system constraint can be formulated as:

$$v_i^c - \Delta v \leqslant v_i \leqslant v_i^c + \Delta v, \tag{1}$$

where Δv is referred to as a pre-specified degree of velocity relaxation. In practice, different types of trains might be scheduled on the rail line, for instance, slow trains and fast trains.

(3) Objective Function

In the real-world operations, the delay-ratio, which is defined as the ratio of the total delay time and the total free-run time, can be adopted to reflect the reduction level of the unnecessary delay time. Obviously, the lower the delay-ratio value is, the more balanced the schedule is. For instance, if its corresponding delay-ratio is zero, the schedule will be a balanced one with the highest level. Then, in the process of scheduling, it is desirable that the delay-ratio be minimized with the optimal speed settings in order to enhance the service level of the rail traffic. This research formulates the delay-ratio by the following objective function $F(V)$,

$$F(V) = \frac{\sum_{i=1}^{N}\left(T_i^{ob}(v_i, V) - T_i^f(v_i, V)\right)}{\sum_{i=1}^{N} T_i^f(v_i, V)} \tag{2}$$

where

$F(V)$: the delay-ratio in the optimal schedule with the velocity vector V;

$T_i^{ob}(v_i, V)$: the time that the train i arrives at the end of its journey in the schedule with the velocity vector V;

$T_i^f(v_i, V)$: the time that the train i arrives at the end of its unobstructed journey with the velocity v_i, i.e., free-run time of train i;

N: the number of trains under the consideration.

For further clarifying the objective function $F(V)$, we shall calculate the corresponding delay-ratio of the scheduling plan in Fig. 3a. We assume that the pre-specified velocity vector $V^c = \{20, 20, 20\}$ (unit: m/s). Lengths of Section 1, 2 and 3 are denoted by s_1, s_2 and s_3, and $s_1 = s_2 = s_3 = 5.4$ km. Train 1 departs from station 1 at time 120 s. Trains 2 and 3 depart from station 4 at time 0 s and 600 s, respectively. If we reduce the velocity of train 1 to 18 m/s, the adjusted velocity vector will be $V^a = \{18, 20, 20\}$ (unit: m/s). After calculating, $d_1 = 150$ s, $d_2 = 60$ s, $d_3 = 120$ s, and $d_4 = 30$ s, and we have

$$F(V^c) = \frac{\sum_{i=1}^{3}(T_i^{ob}(v_i^c, V^c) - T_i^f(v_i^c, V^c))}{\sum_{i=1}^{3} T_i^f(v_i^c, V^c)} = \frac{d_1 + d_2}{\sum_{i=1}^{3}\left(\sum_{j=1}^{3} s_j / v_i^c\right)} = \frac{210}{810} \approx 0.259,$$

and

$$F(V^a) = \frac{\sum_{i=1}^{3}(T_i^{ob}(v_i^a, V^a) - T_i^f(v_i^a, V^a))}{\sum_{i=1}^{3} T_i^f(v_i^a, V^a)} = \frac{d_3 + d_4}{\sum_{i=1}^{3}\left(\sum_{j=1}^{3} s_j / v_i^a\right)} \approx \frac{150}{838} \approx 0.179.$$

According to the definition of "balanced schedule" in this paper, it is clear that adjusted schedule is more balanced than the original one, since $F(V^a) < F(V^c)$.

Similarly, in Fig. 3b and c, delay-ratios of the original schedules are positive numbers. While delay-ratios of the adjusted schedules are zero, since there is no delay time appears.

(4) Mathematical Model

The problem of searching for optimal velocities planning with the minimum delay-ratio can be formulated as

$$\begin{cases} \min F(V) \\ s.t. \\ v_i^c - \Delta v \leqslant v_i \leqslant v_i^c + \Delta v. \end{cases} \tag{3}$$

In this model, the objective function specifies the delay-ratio in the optimal train schedule with velocity vector V, and the constraint states the interval of allowable velocity in the scheduling process. The aim is to generate optimal velocities for different trains such that the corresponding optimal train schedule has the least delay-ratio.

3.2. Solution algorithm

In this section, we shall combine the improved TAS (ITAS) method with the GA into a new heuristics, called GA-ITAS method, to solve the proposed model, in which ITAS is applied to scheduling trains with the given velocity selections and GA is employed to search for an optimal set of trains' velocities. The following discussion aims to give a detailed description for the proposed solution methodology.

3.2.1. Improved TAS method

In formulation (3), an implicit optimization must be solved to determine an approximate optimal train timetable for any given velocity vector V. To this end, a time-efficient schedule algorithm is expected to be designed for computing the value of $F(V)$. In what follows, an improved TAS method is proposed on the basis of the analysis of the train operational rules, which ensures to generate an approximately optimal schedule with the given velocity variable within a reasonable computational time.

In the real-world applications, due to the constraints of the rail tracks and stations, stopping at the track is generally prohibited for each train, and some important operations, such as Meeting & Passing (M&P), Meeting & Overtaking (M&O), are required to be carried out at stations. In M&P operations, the first train to arrive at a common station should stop at a siding track at the station and then wait for the opposite train to pass, which is intended to avoid the deadlocks between the trains in opposite directions. Likewise, the M&O operation can be used to avoid the collision produced by two trains with different maximum velocities (e.g., a slow train runs ahead of the fast one). For this operation, the slow train can be required to stop at a siding track of the station and wait for the fast train to overtake on the condition of enough station capacity. For more details of these two operations, we can refer to Dorfman and Medanic [14].

In this research, we shall modify an overtaking rule proposed by Dorfman and Medanic [14] to further reduce the total delay time based on the analysis of train operational rules. In particular, we focus on the case (b) of the M&O event in Ref. [14], because it is considered as the only case where an overtake event might take place at a considered discrete event by Dorfman and Medanic [14]. As shown in Fig. 4a, two trains i and $i-1$, in which train i is a fast train and train $i-1$ is a slow train, are traveling in the same direction with the marked position at the considered discrete event. Due to the difference of the speed between two trains, it is obvious that a conflict will occur in their rest trips (see Fig. 4b).

For this conflict, whether an overtaking event takes place or not is determined in Ref. [14] by the speed of these two trains. In detail, if the velocity at which train i travels on its current section $S(i)$ (denoted by $v(i, S(i))$) is greater than that of train $i-1$ on the section $S(i-1)$ (denoted by $v(i - 1, S(i - 1))$), i.e., $v(i, S(i)) > v(i - 1, S(i - 1))$, an overtaking event will probably occur at station $m + 1$. An illustration of scheduling process can be found in Fig. 5a. However, we need particularly mention that an improved overtaking rule for this case can be designed to further decrease the waiting time, described below:

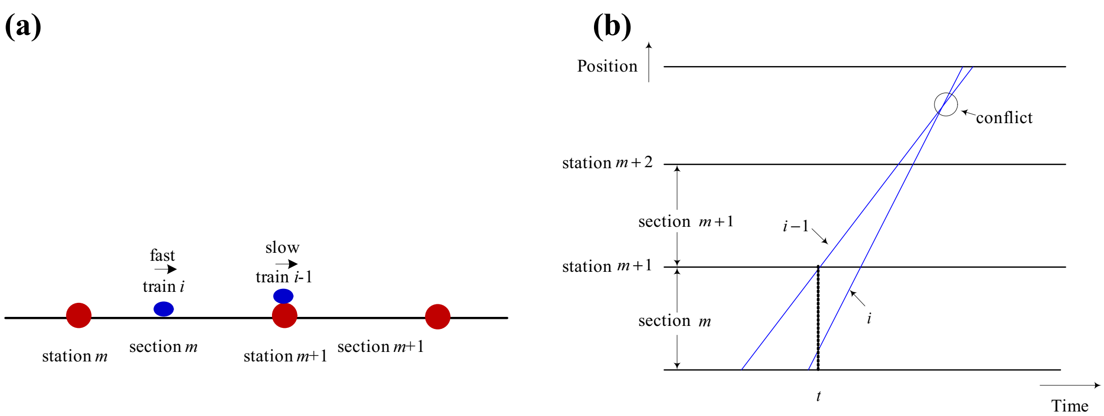

Fig. 4. Potential overtaking operation for two trains with different speeds: (a) positions of the train i and train $i-1$ at time t; (b) a simple timetable.

- **Improved rule**: *If and only if* $t_1 < t_2$, train $i-1$ is scheduled to leave the station $m+1$ and will stop at station $m+2$ where an overtaking event will occur.

In this specific rule, t_1 denotes the required travel time for train $i-1$ to reach station $m+2$, t_2 represents the travel time that train i needs to arrive at station $m+2$. Clearly, compared with the rule of Dorfman and Medanic [14], the new rule can further reduce delay time (see t_{delay} in Fig. 5a and b). With the improved rule, we finally produce a more effective method, called improved travel advanced strategy (ITAS), in our research.

For analyzing the performance of the TAS method, three criteria, namely (a) the time to clear the line, (b) the delay of all trains, and (c) the maximal delay, are proposed in Ref. [14]. Specifically, the time to clear the line can be represented by the following function:

$$J_1 = T_M^{ob}(v_M, V) - t_0 \tag{4}$$

where T_M^{ob} is the arrival time at the end of the journey of the last train M on the schedule. The total delay of the trains is given by

$$J_2 = \sum_{i=1}^{N}(T_i^{ob}(v_i, V) - T_i^f(v_i, V)). \tag{5}$$

The maximal delay criterion is defined by

$$J_3 = \max_i\{T_i^{ob}(v_i, V) - T_i^f(v_i, V)\}. \tag{6}$$

In addition, Dorfman and Medanic [14] especially proposed a ratio to measure the time-efficiency of the schedule obtained by the TAS, formulated by

$$\eta = \frac{(T_M^f(v_M, V) - t_0)}{(T_M^{ob}(v_M, V) - t_0)}. \tag{7}$$

These four criteria will be used as extra methods to further analyze the performance of the near-optimal train schedule in the formulation (3).

3.2.2. The genetic algorithm

In this section, genetic algorithm is designed to search for the optimal velocity of each train, where the corresponding train schedule generated by the improved TAS can be used to evaluate the quality of the velocity selections.

As a stochastic search algorithm, genetic algorithm (GA) was first initiated by Holland [25] in 1975 based on biology evolutionary theory. In recent decades, GA has been well-developed by a lots of researchers [26–29], and successfully applied to a variety of optimal models [30–35]. In this algorithm, the population and chromosome are the two most important parts in the searching process, where the population consists of a set of chromosomes randomly generated at the start of GA. The fitness of a chromosome in the population is assessed by the evaluation function $F(V)$. High-fitness chromosomes can be selected with higher probability to produce the next generation via crossover and mutation operations. The algorithm will not stop to generate new generation until the creation of a qualified generation satisfying the convergence criterion or other terminal conditions. The detailed techniques in GA will be listed in the following. Firstly, we shall introduce some relevant notations in the procedure:

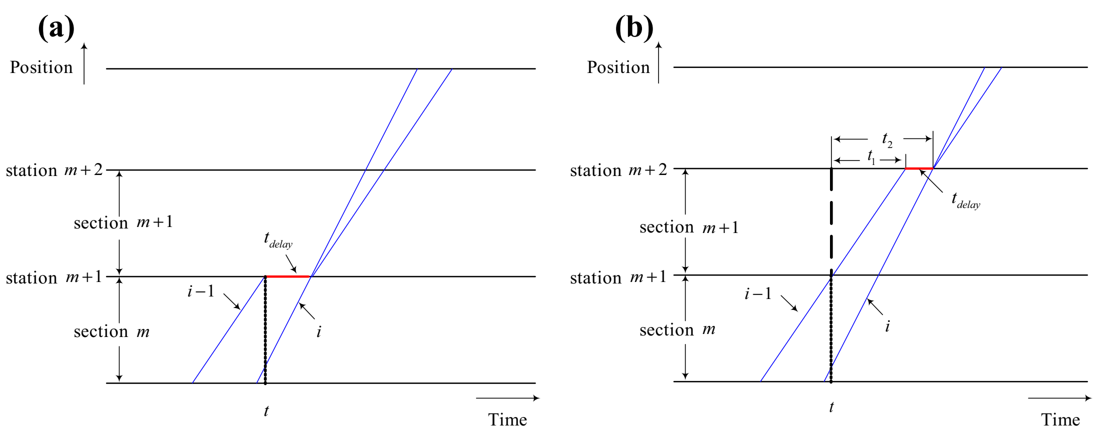

Fig. 5. Comparison of two different rules: (a) rule in Ref. [10]; (b) improved rule in this paper.

pop_size: size of the populations in one generation;
gen_size: size of the genes in one chromosome;
P_c: probability of crossover operation;
P_m: probability of mutation operation.

3.2.2.1. Solution representation. For simplicity, we take the trains' velocities as genes in the chromosome, which is represented by $C = V = [v_1, v_2, \ldots, v_N]$. In GA, we need to firstly initialize *pop_size* feasible chromosomes. Note that a chromosome is feasible if all genes in the chromosome satisfy the velocity constraints (1).

3.2.2.2. Selection operation. For selecting the chromosomes for the next generation, the roulette-wheel method is employed here for selection operation. Firstly, the chromosomes in the population will be ranked in an increasing sequence according to their fitness, marked by $V_1, V_2, \ldots, V_{pop\_size}$.

Then, an evaluation function is used to measure the likelihood of reproduction for each chromosome. Specifically, we use the function $Eval(V_i) = \alpha(1-\alpha)^{i-1}$ to firstly give the fitness of each chromosome, where $\alpha \in [0, 1]$ is a pre-specified parameter. Moreover, an accumulative sequence $\{W_k\}$ is presented for characterizing the roulette wheel operation, i.e., $W_0 = 0$, $W_k = \sum_{i=1}^{k} Eval(V_i)$, $k = 1, 2, \ldots, pop\_size$.

3.2.2.3. Crossover operation. In this operation, for each chromosome i, we randomly generate a real number p_i $(0 < p_i < 1)$, if $p_i < P_c$, we select chromosome i as a parent chromosome to breed children chromosomes. Crossover operations can be performed as follows. Assume that $V_i = (v_1^i, v_2^i, \ldots, v_N^i)$ and $V_j = (v_1^j, v_2^j, \ldots, v_N^j)$ are two parent chromosomes. Then, randomly generate an integer number $t (1 \leqslant t \leqslant N)$, with the one-point crossover operation, we produce $X = (v_1^i, v_2^i, \ldots, v_t^j, \ldots, v_N^j)$ and $Y = (v_1^j, v_2^j, \ldots, v_t^i, \ldots, v_N^i)$. This progress is shown in Fig. 6. Otherwise, chromosome i will be automatically transferred to the next generation.

3.2.2.4. Mutation operation. In this operation, for each chromosome i, we randomly generate a real number p_i $(0 < p_i < 1)$. If $p_i < P_m$, we randomly generate an integer number $t(1 \leqslant t \leqslant gen\_size)$, then randomly generate a feasible gene which satisfies the constraint (1) to replace the tth gene of chromosome i (see Fig. 7). Otherwise, chromosome i will be automatically transferred to the next generation.

3.2.3. The GA-ITAS algorithm

Based on the aforementioned detailed analysis, GA-ITAS approach can be designed as the following procedure:

Step 1. Initialize the population which is used as the input information of the train system.
Step 2. Based on the current system information, apply ITAS approach to scheduling trains for each velocity vector and calculate the corresponding delay-ratios.
Step 3. If the current status does not meet the terminal condition, use the selection, crossover and mutation operations to update the chromosomes, and go to step 2; otherwise, go to step 4.
Step 4. Output the best solution, i.e., the optimal combination of the trains' velocities; and then produce the corresponding train schedule.

4. Numerical experiments

To show the efficiency and effectiveness of the proposed approach, this section will apply GA-ITAS to scheduling trains on a railway line similar to that in Ref. [16], illustrated in Fig. 8. In this experimental railway, a total number of 16 single-track sections, which can be used by the trains with both directions, are taken into consideration. The total length of the railway is set as 275 km and the section lengths vary between 5 km and 30 km. Station A and station B represent the terminal stations for trains originating and terminating, and the other stations denote "intermediate" stations for trains passing and overtak-

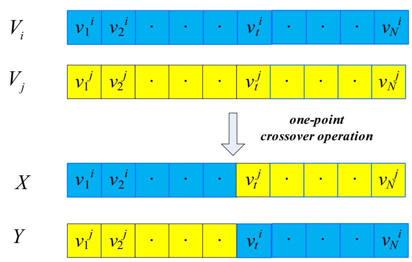

Fig. 6. Crossover operation in genetic algorithm.

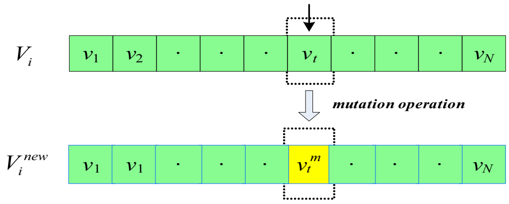

Fig. 7. Mutation operation in the genetic algorithm.

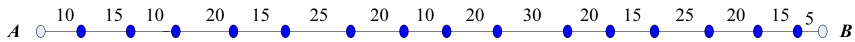

Fig. 8. A simple railway line example.

ing. Two classes of trains are considered in this set of experiments, where one class departs from station A and terminates at station B, and the other one runs in the opposite direction. The departure interval time is assumed to be one hour. Moreover, in GA-ITAS, P_c is set as 0.6, P_m is set as 0.5, the *pop_size* is set as 20 and the number of generations is set as 150.

4.1. Computational results

4.1.1. The case of homogeneous trains

In this case, we use GA-ITAS method to schedule 18 homogeneous trains, in which velocity variables are relaxed from a pre-specified fixed value 20 m/s to the potential integers varying between 18 m/s and 22 m/s. The ITAS method is also employed to schedule 18 homogeneous trains with the fixed velocity 20 m/s for comparative convenience. In the experimental results, the optimal combination of trains velocities in GA-ITAS turns out to be V^G = {20, 19, 22, 19, 18, 20, 18, 22, 22, 18, 19, 22, 21, 22, 19, 21, 18, 18} (unit: m/s), which leads to the total travel time 4512.23 min and produces 100.27 min reduction compared with 4612.5 min in the ITAS. It is noteworthy that in comparison to that in the ITAS, the trajectories of some trains in the result of GA-ITAS can be improved to a great extent. For instance, the eighth up stream train (direction: from station A to station B) is scheduled to stop at the stations for five times in the ITAS, and its total waiting time is 2950 s. However, in the GA-ITAS, the same train is only planned to stop at the stations twice due to the generated optimal velocities, and its total waiting time is reduced to 1770 s. The timetables obtained by ITAS and GA-ITAS are given in Fig. 9a and b, respectively.

In order to further compare the performance of the generated schedules, we give the results of three criteria and two ratios corresponding to solutions produced by the ITAS and GA-ITAS in Table 2. The relative errors ε ($\varepsilon = (Z_{GA-ITAS} - Z_{ITAS})/Z_{ITAS}$) are also reported for different items. In the experimental results, it is obvious that the total delay time J_2 in GA-ITAS is notably less than that in ITAS, which indicates that relaxing the constraints of train velocity essentially contributes to the reduction of the total delay time in the process of scheduling the trains. Likewise, the maximal delay time (criteria J_3) in the GA-ITAS reduce almost 995 s compared to that produced by ITAS. Additionally, although the time-efficient ratio η by the ITAS takes a high value, it is still less than that produced by the GA-ITAS, which also states the time-efficiency of the schedule obtained by the GA-ITAS is better than that via the TAS. As for the objective function, the optimal objective function by GA-ITAS algorithm is only 0.0834 in our model, while the optimal objective produce by ITAS turns out to be 0.1182. This result essentially verifies that the GA-ITAS can keep the delay time of the trains in a more balanced state. However, we need to point out that since the average velocity (18 m/s) of the last two trains in the GA-ITAS is lower than the fixed velocity (20 m/s) in the ITAS, the time to clear the line, i.e., J_1 in the former algorithm, is slightly longer than that in the latter, though there is only one stop of the two trains in the GA-ITAS.

On the basis of the analysis above, we finally conclude that the train timetable obtained by GA-ITAS is apparently more suitable than that obtained via ITAS in the case of homogeneous trains, above all, in the aspects of the total delay time and the time-efficiency.

4.1.2. The case of heterogeneous trains

In this set of experiments, two classes of trains, i.e., slow train and fast train, will be taken into consideration to test the proposed approaches. Assumptions for slow trains are the same to that in the case of homogeneous trains, while the velocity of the fast train is assumed to be an integer and varies between 28 m/s and 32 m/s. According to the service requirements, slow train and fast train are scheduled to depart from their origin in turn, and the departure time interval is set as one hour. In the experiment results, the total travel time of trains are 4006.11 min in the TAS, 3900.28 min in the ITAS, and 3848.91 min in the GA-ITAS, respectively. The relative errors $\mu = (Z_{ITAS} - Z_{TAS})/Z_{ITAS}$, for each item are given in Table 3 where we can see almost all of the criteria in ITAS is better than those in TAS. That is, the proposed ITAS is more effective than TAS.

In addition, the optimal conbination of the velocities obtained by the GA-ITAS is V^G = {19, 22, 31, 29, 18, 21, 30, 30, 18, 18, 30, 32, 18, 19, 30, 31, 21, 18} (unit: m/s). In Table 3, results of the case of hetergogeneous trains are given, and the relative errors ε ($\varepsilon = (Z_{GA-ITAS} - Z_{ITAS})/Z_{ITAS}$) for each item are also calculated.

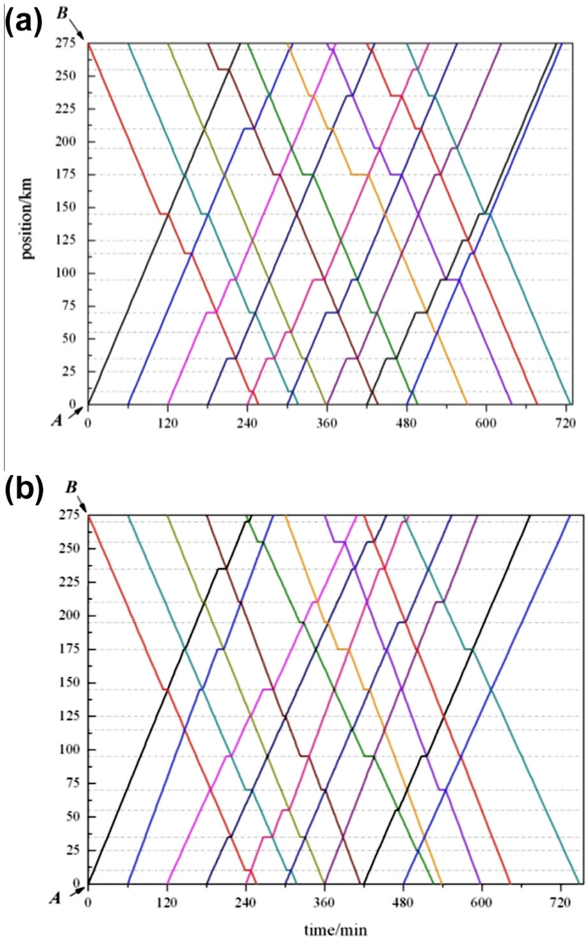

Fig. 9. Schedule for the case of homogeneous trains in two models: (a) ITAS and (b) GA-ITAS ($P_c = 0.6$, $P_m = 0.5$).

Table 2
Performance criteria and ratios for the case of homogeneous trains ($P_c = 0.6$, $P_m = 0.5$).

	J_1	J_2	J_3	η	*delay-ratio*
ITAS	43600 s	29250 s	3400 s	0.9760	0.1182
GA-ITAS	44911 s	20800 s	2405 s	0.9814	0.0834
ε	0.0301	−0.2889	−0.2926	0.0055	−0.2944

Table 3
Performance criteria and ratios for the case of heterogeneous trains ($P_c = 0.6$, $P_m = 0.5$).

	J_1	J_2	J_3	η	*delay-ratio*
TAS	44183 s	29533 s	4650 s	0.9630	0.1401
ITAS	43783 s	28183 s	4550 s	0.9718	0.1099
GA-ITAS	44652 s	14423 s	3567 s	0.9871	0.0666
μ	−0.0091	−0.0457	−0.0215	0.0091	−0.2156
ε	0.0198	−0.4882	−0.216	0.0157	−0.394

Clearly, we can see that the statistical characteristics of the data in Tables 2 and 3 are almost consistent to each other. Thus, the results of the case of homogeneous trains can be equivalently applicable to the heterogeneous case, i.e., the GA-ITAS method in the train scheduling problem works more efficiently than the ITAS method. However, some differences still exist between these two set of experiments. Specifically, most of the relative errors ε in this case are roughly equal to or better than that in the homogeneous case, except for J_1 which is greatly influenced by the velocities of the last two trains. Compared with the ITAS, the GA-ITAS greatly reduces the total delay time J_2 by 48.82% in this case, while only 28.89% in

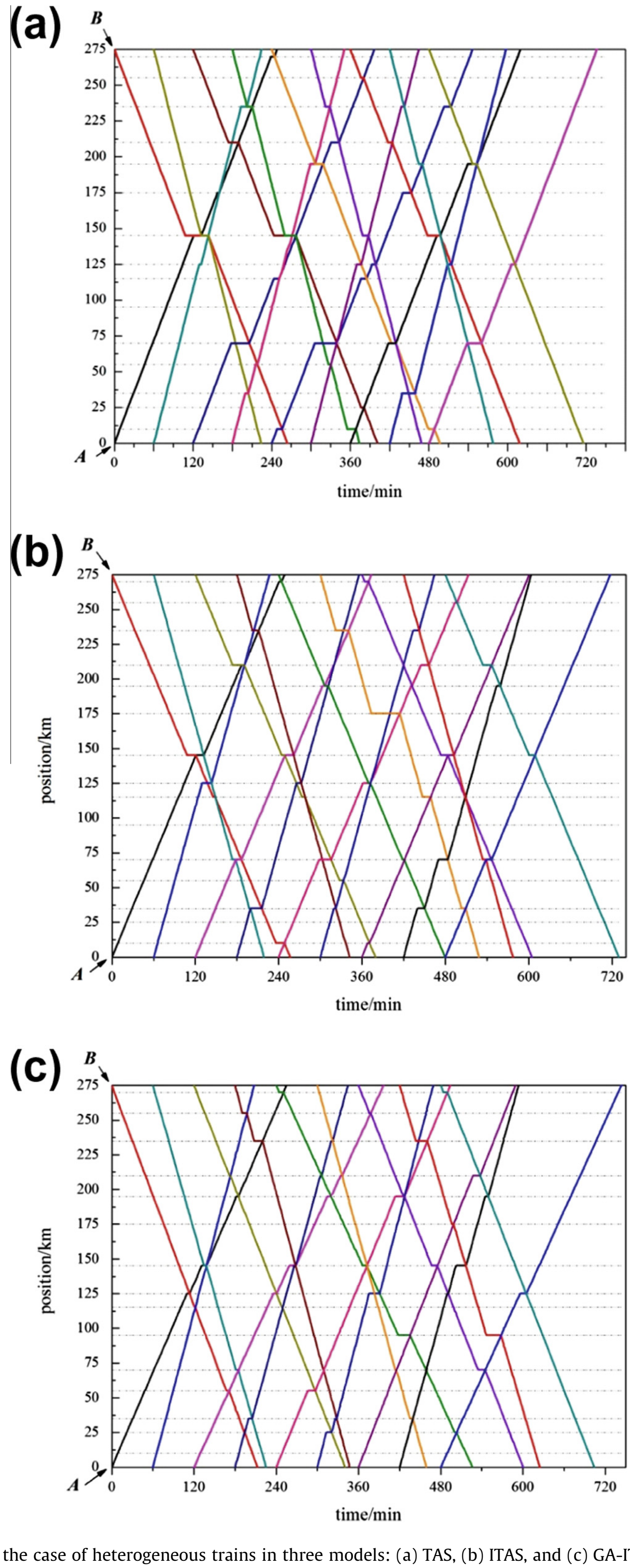

Fig. 10. Schedule for the case of heterogeneous trains in three models: (a) TAS, (b) ITAS, and (c) GA-ITAS (P_c = 0.6, P_m = 0.5).

Table 4
Means and standard deviations of different performance criteria.

	J_1	J_2	J_3	η	*delay-ratio*
$P_c = 0.3$, $P_m = 0.3$	44652 s	14423 s	3567 s	0.9871	0.0666
$P_c = 0.3$, $P_m = 0.8$	44490 s	14412 s	3567 s	0.9907	0.0667
$P_c = 0.6$, $P_m = 0.5$	44652 s	14423 s	3567 s	0.9871	0.0666
$P_c = 0.8$, $P_m = 0.3$	44228 s	14603 s	3450 s	0.9784	0.0677
$P_c = 0.8$, $P_m = 0.8$	44490 s	14412 s	3567 s	0.9907	0.0667
Mean	44502 s	14455 s	3544 s	0.9868	0.0669
Standard deviation	69.4	33.2	21.0	0.0020	0.0002

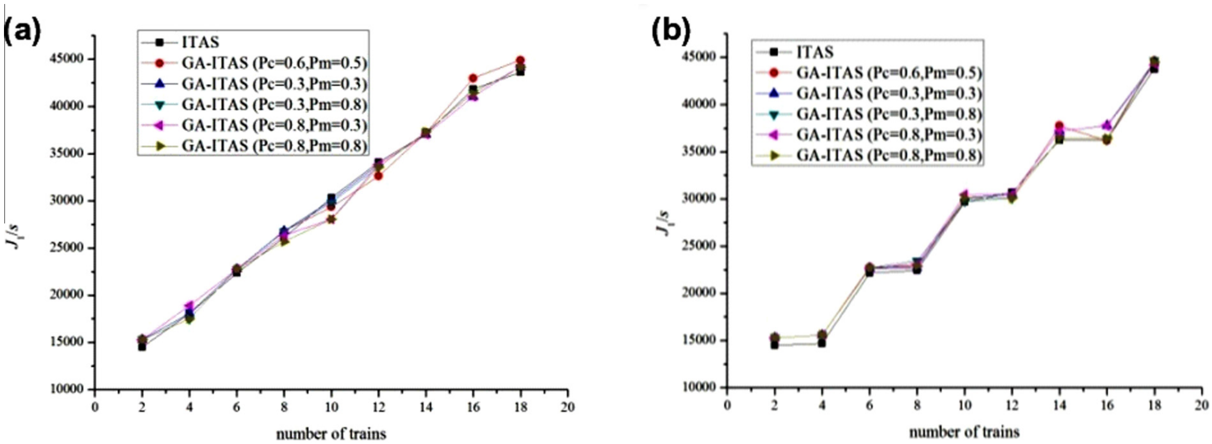

Fig. 11. The value of J_1 ((a) homogeneous trains; (b) heterogeneous trains).

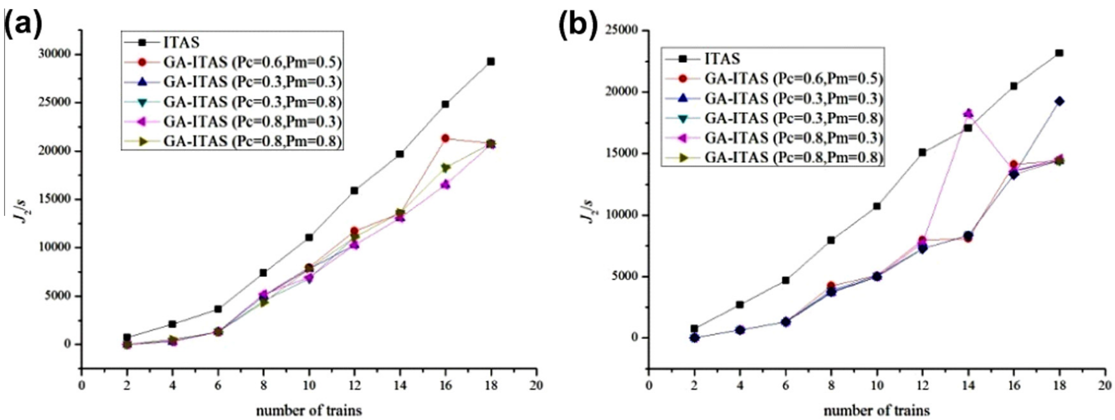

Fig. 12. The value of J_2 ((a) homogeneous trains; (b) heterogeneous trains).

the case of homogeneous trains. This is to say, the larger differences among the trains' velocities are, the more coupling effect among the trains will be produced. Therefore, when the constraint of train velocity is relaxed, the limited railway resource might be employed with a much high utilization rate. More simulations and results will be presented in Section 4.1.3, which essentially strength the illustration of the above analysis. In Fig. 10, the scheduling diagrams in the case of heterogeneous trains in the TAS, ITAS and GA-ITAS are given.

To make a more meaningful comparison between the ITAS and GA-ITAS methods, we run the GA-ITAS method many times and report the mean and standard deviation of the evaluation values. Detailed results can be referred to Table 4.

4.1.3. The performance of the GA-ITAS

In this section, for different numbers of the trains scheduled on the rail line, the evaluation criteria and ratios mentioned above are computed to demonstrate the characteristics of the generated schedules. In these results, we can analyze the

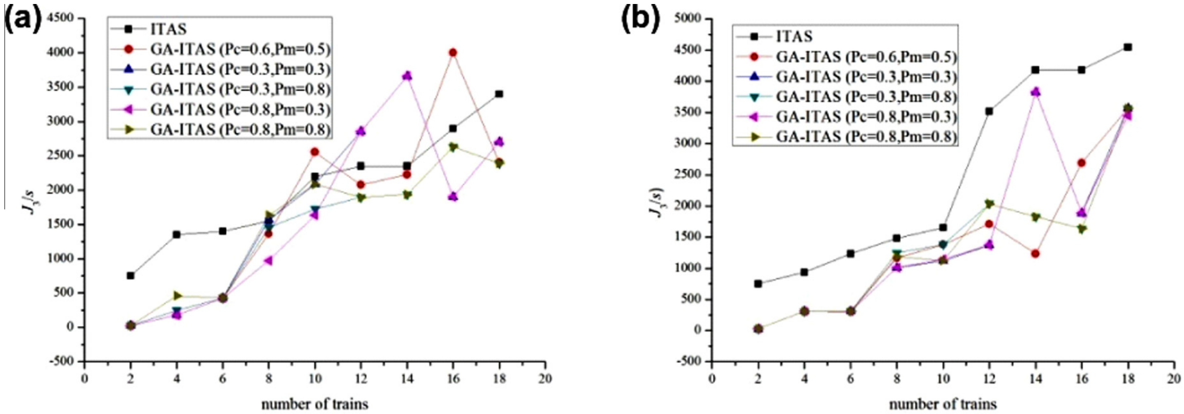

Fig. 13. The value of J_3 ((a) homogeneous trains; (b) heterogeneous trains).

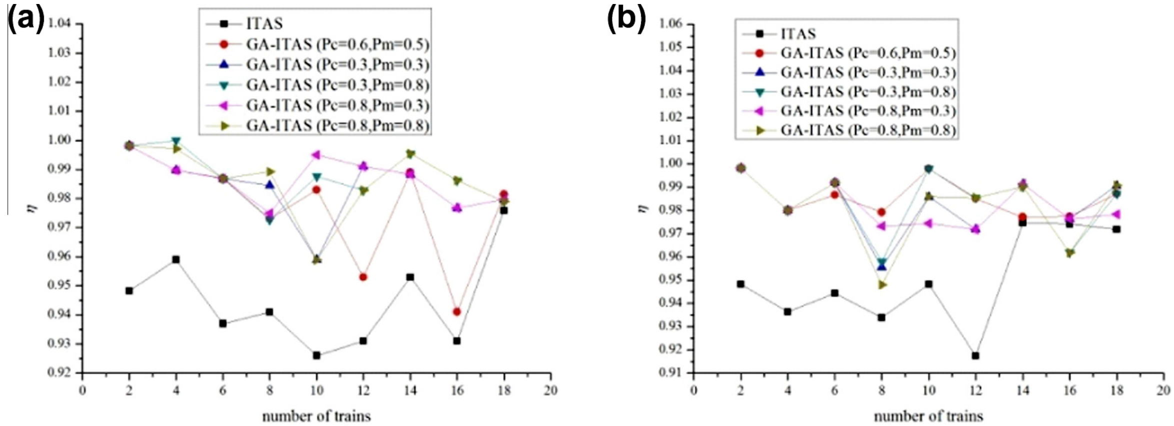

Fig. 14. The value of η ((a) homogeneous trains; (b) heterogeneous trains).

performance of the GA-ITAS method from the statistical perspective compared with the benchmarks provided by ITAS. In the experiments, the relevant assumptions for trains are also the same to that listed in Sections 4.1.2 and 4.1.3. Fig. 11 shows the values of J_1 produced by ITAS and GA-ITAS, respectively, with different numbers of the trains and different parameters P_c and P_m. It follows from this figure that in the cases of homogeneous trains, J_1 produced by GA-ITAS are roughly consistent to that in ITAS, but the former is generally bigger than the latter in the case of heterogeneous trains.

As addressed above, J_2 is used to represent the total delay time in the schedule of the involved trains. In Fig. 12, it is clear that J_2 in ITAS is greater than that in GA-ITAS for two sets of experiments, and what's more, the gap is fairly large for the case of heterogeneous trains. As J_2 is a crucial criterion to reflect the performance of the train schedule, the GA-ITAS method shows its efficiency and effectiveness in producing a high-quality schedule for both cases of the homogeneous and heterogeneous trains when compared with the ITAS approach.In Fig. 13, the maximal delays J_3 produced by ITAS and GA-ITAS, respectively, have not clear tendencies with the increase of the number of trains. Nevertheless, in the case of heterogeneous trains, for each schedule with respect to different number of trains, J_3 in GA-ITAS is always less than that in ITAS.

"The high value of η shows that the railway system indeed works efficiently" is a conclusion in Ref. [14]. With this result, according to the values of η in Fig. 14, both the GA-ITAS method and the ITAS method can make the railway system works efficiently, and as expected, the former is significantly better than the latter.

Fig. 15 displays the values of the delay-ratio $F(V)$. It is easy to see that, compared with the values of $F(V)$ in TAS, the ones in GA-ITAS decreases remarkably. On one hand, this phenomenon verifies the validity of the GA-ITAS algorithm in generating a balanced train schedule. On the other hand, it proves the operability of our core idea. That is, it is possible for us to find an optimal combination of trains' velocities via optimizing the relative delay level ratio $F(V)$, and then obtain a schedule with the least unnecessary delay time, i.e., a balanced schedule.

As a conclusion, in most situations, during the process of searching for an optimal schedule with the least delay-ratio, the other criteria and ratios can be also optimized simultaneously, especially for J_2. Interestingly, in both homogeneous and heterogeneous cases, it is easy to see that the gap between the total delay time (i.e., J_2) produced by ITAS and GA-ITAS might be enlarged with the increasing of the numbers of trains. This result also states that the proposed GA-ITAS is very effective in generating a high-quality schedule for the large-scale problem. However, from Figs. 11 to 15, there are also few points are aberrant. This is a norm phenomenon because that GA is essentially a probabilistic search method.

4.2. The convergence of the GA-ITAS algorithm

To further investigate the convergence of the GA-ITAS approach, the iterative processes of this algorithm in case of heterogeneous trains are given in Fig. 16. In this set of experiments, different numbers of trains, i.e., 6, 10, 14 and 18, respectively, are considered on the railway link listed in Fig. 8. It is easy to see from Fig. 16 that the proposed GA-ITAS algorithm can be quickly convergent for each group of trains within 80 iterations, where 20 chromosomes are generated for each population in the algorithms. This result also indicates the steadiness and robustness of genetic algorithm in seeking an optimal scheduling plan.

Since the GA is a probabilistic search technique, different random seeds will produce different answers for a same problem. Take the case of 18 heterogeneous trains as an example. In Fig. 17, we report the optimization processes produced with different values of P_c and P_m, in which although the finally generated approximately optimal objectives are slightly different, their convergent tendencies are still similar to each other.

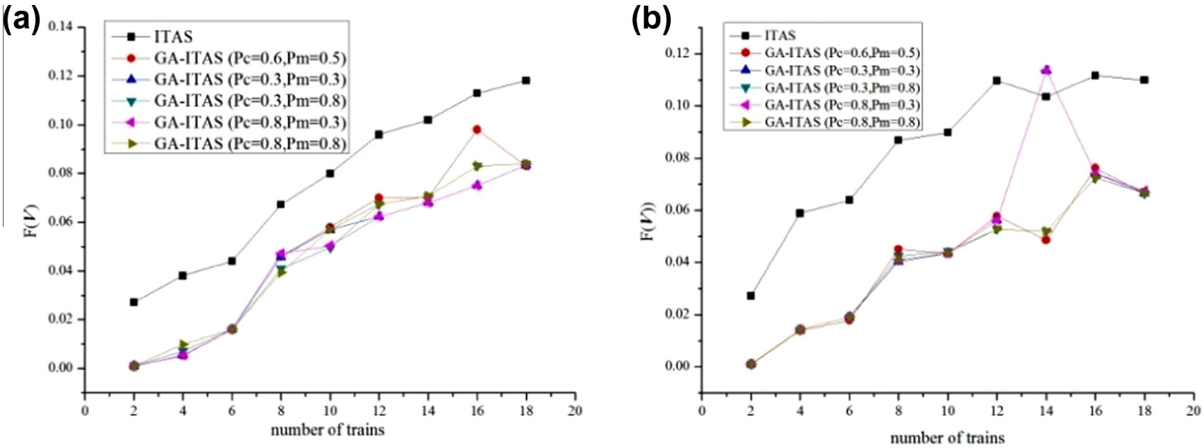

Fig. 15. The value of $F(V)$ ((a) homogeneous trains; (b) heterogeneous trains).

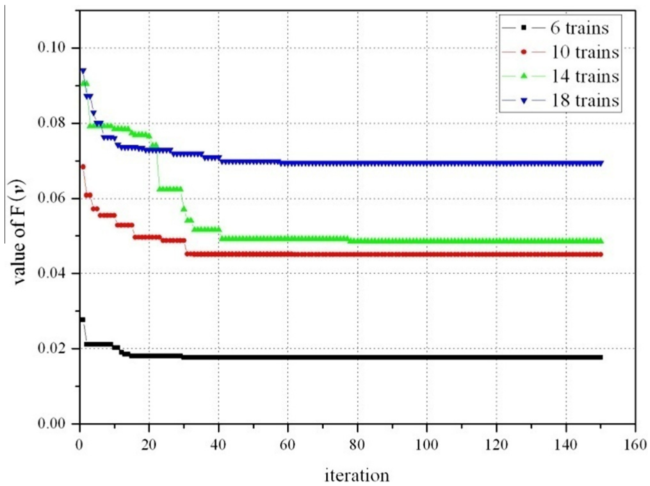

Fig. 16. Optimization processes of GA-ITAS ($P_c = 0.6$, $P_m = 0.5$).

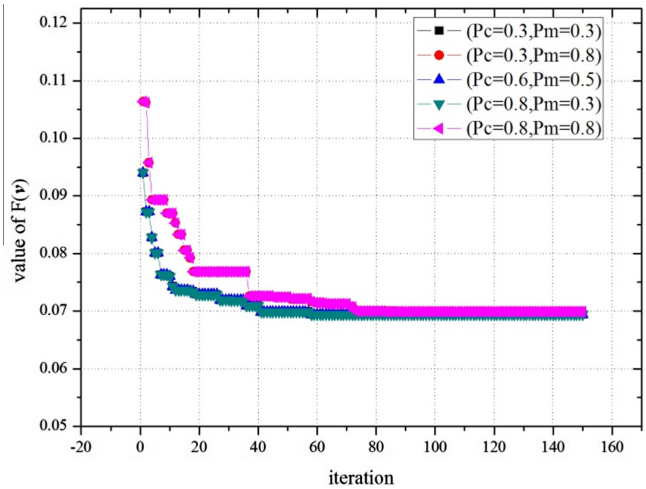

Fig. 17. Optimization processes of GA-ITAS with different P_c and P_m.

5. Conclusions

Combining the ITAS method and genetic algorithm, this research proposed an improved simulation model for scheduling trains, called GA-ITAS method, where the ITAS method is the improved method developed based on the TAS method proposed by Dorfman and Medanic [14]. Different from the researches in literature, the velocity optimization was particularly considered as a decision variable to reduce the coupling effects caused by the velocity difference, and then generated an optimal schedule with the balanced characteristics. Several indices, including the time to clear the line, the delay of all trains, and the maximal delay, were employed in this research to further evaluate the performance of the generated schedule. The numerical experiments were implemented on a well-connected rail line with 17 stations for the cases of homogenous and heterogeneous trains. The computational results demonstrate that compared with TAS, the proposed approaches in this research can find a more high-quality train schedule under the evaluation of different indices.

Additionally, it is worth pointing out that the proposed approaches can be easily generalized to a more complex case, i.e., considering the velocity decision variables on each link rather than on the entire journey. Besides, the micro operations and energy consumption are not considered in the scheduling process. These aspects can be treated as the new topics in future researches.

Acknowledgements

This research was supported by the National High Technology Research and Development Program of China (No. 2011AA110502), the National Natural Science Foundation of China (Nos. 71271022, 71271020), the Research Foundation of State Key Laboratory of Railway Traffic Control and Safety, Beijing Jiaotong University (Nos. RCS2012ZZ001, RCS2011ZZ003) and Program for New Century Excellent Talents in University under Grant no. NCET-10-0218.

Reference

[1] I. Amit, D. Goldfarb, The timetable problem for railways, Develop. Oper. Res. 2 (1971) 379–387.
[2] A. Higgins, E. Kozan, L. Ferreira, Optimal scheduling of trains on a single line track, Transp. Res. B 30 (2) (1996) 147–161.
[3] I. Sahin, Railway traffic control and train scheduling based on inter-train conflict management, Transp. Res. B 33 (7) (1999) 511–534.
[4] B. Szpigel, Optimal train scheduling on a single track railway, Oper. Res. 7 (1973) 344–351.
[5] L.X. Yang, K.P. Li, Z.Y. Gao, Train timetable problem on a single-line railway with fuzzy passenger demand, IEEE Trans. Fuzzy Syst. 17 (3) (2009) 617–629.
[6] X.S. Zhou, M. Zhong, Single-track train timetabling with guaranteed optimality: branch-and-bound algorithms with enhanced lower bounds, Transp. Res. Part B 41 (2007) 320–341.
[7] M. Carey, A model and strategy for train pathing with choice of lines, platforms and routes, Transp. Res. B 28 (5) (1994) 333–353.
[8] M. Carey, D. Lockwood, A model, algorithms and strategy for train pathing, J. Oper. Res. Soc. 46 (8) (1995) 988–1005.
[9] K. Ghoseiri, F. Szidarovszky, M.J. Asgharpour, A multi-objective train scheduling model and solution, Transp. Res. Part B 38 (2004) 927–952.
[10] L.X. Yang, K.P. Li, Z.Y. Gao, X. Li, Optimizing trains movement on a railway network, Omega Int. J. Manage. Sci. 40 (5) (2012) 619–633.
[11] M. Carey, Extending a train pathing model from one-way to two-way track, Transp. Res. Part B 28 (5) (1994) 395–400.
[12] U. Brannlund, P.O. Lindberg, A. Nou, J.E. Nilsson, Railway timetabling using Lagrangian relaxation, Transport. Sci. 32 (1998) 358–369.
[13] X. Cai, C.J. Goh, A fast heuristic for the train scheduling problem, Comput. Oper. Res. 21 (5) (1994) 499–510.
[14] M.J. Dorfman, J. Medanic, Scheduling trains on a railway network using a discrete event model of railway traffic, Transp. Res. Part B 38 (1) (2004) 81–98.
[15] J. Medanic, M.J. Dorfman, Efficient scheduling of traffic on a railway line, J. Optim. Theory Appl. 115 (3) (2002) 587–602.
[16] F. Li, Z.Y. Gao, K.P. Li, L.X. Yang, Efficient scheduling of railway traffic based on global information of train, Transp. Res. Part B 42 (10) (2008) 1008–1030.

[17] R.G.J. Mills, S.E. Perkins, P.J. Pudney, Dynamic rescheduling of long-haul trains for improved timekeeping and energy conservation, Asia-Pacific J. Oper. Res. 8 (2) (1991) 146–165.
[18] O. Frank, Two-way traffic on a single line of railway, Oper. Res. 14 (5) (1965) 801–811.
[19] Peat, Marwick, Mitchell & Co., 1975. Train dispatching simulation model: capabilities and description USA. Report No. DOT-FR-4-5014-1, March 1975, Prepared for Federal Railroad Administration, Department of Transportation, Washington, DC.
[20] D.A. Rudd, A.J. Storry, Single track railway simulation, new models and old, Rail Int. 7 (6) (1976) 335–342.
[21] E.R. Petersen, A.J. Taylor, C.D. Martland, An introduction to computer-assisted train dispatch, J. Adv. Transport. 20 (1) (1986) 63–72.
[22] R.L. Sauder, W.M. Westerman, Computer aided train dispatching: decision support through optimization, Interfaces 13 (6) (1983) 24–37.
[23] D. Jovanovic, P.T. Harker, Tactical scheduling of rail operations: the SCAN I system, Transport. Sci. 25 (1) (1991) 46–64.
[24] P. Pudney, A. Wardrop, Generating train plans with problem space search, in: M. Hickman, P. Mirchandani, S. Voß (Eds.), Computer-aided Systems in Public Transport, Lecture Notes in Economics and Mathematical Systems, Springer, 2008.
[25] J.H. Holland, Adaptation in Natural and Artificial Systems, University of Michigan Press, Ann Arbor, 1975.
[26] S.S. Chaudhry, W. Luo, Application of genetic algorithms in production and operations management: a review, Int. J. Prod. Res. 43 (19) (2005) 4083–4101.
[27] D.E. Goldgerg, Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Esley, 1989.
[28] C.H. Martin, A hybrid genetic algorithm/mathematical programming approach to the multi-family flowshop scheduling problem with lot streaming, Omega Int. J. Manage. Sci. 37 (2009) 126–137.
[29] A. Udhayakumar, V. Charles, M. Kumar, Stochastic simulation based genetic algorithm for chance constrained data envelopment analysis problems, Omega Int. J. Manage. Sci. 39 (2011) 387–397.
[30] A.E. Carter, C.T. Ragsdale, Scheduling pre-printed newspaper advertising inserts using genetic algorithms, Omega Int. J. Manage. Sci. 30 (2002) 415–421.
[31] J.W. Chung, S.M. Oh, I.C. Choi, A hybrid genetic algorithm for train sequencing in the Korean railway, Omega Int. J. Manage. Sci. 37 (2009) 555–565.
[32] S.M. Lee, A.A. Asllani, Job scheduling with dual criteria and sequence-dependent setups: mathematical versus genetic programing, Omega Int. J. Manage. Sci. 32 (2004) 145–453.
[33] K. Nachtigallt, S. Voget, A genetic algorithm approach to periodic railway synchronization, Comput. Oper. Res. 23 (5) (1996) 453–463.
[34] V. Salima, X. Cai, A genetic algorithm for railway scheduling with environmental considerations, Environ. Modell. Software 12 (4) (1997) 301–309.
[35] X. Yang, X. Li, Z.Y. Gao, H.W. Wang, T. Tang, A cooperative scheduling model for timetable optimization in subway systems, IEEE Trans. Intell. Transp. Syst. 14 (1) (2013) 438–447.