Transportation Research Part C 84 (2017) 196–218

Contents lists available at ScienceDirect

Transportation Research Part C

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

Train routing and timetabling problem for heterogeneous train traffic with switchable scheduling rules

Yan Xu [a,b], Bin Jia [a,*], Amir Ghiasi [b], Xiaopeng Li [b]

[a] State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, 100044, China
[b] Department of Civil and Environmental Engineering, University of South Florida, 33620, USA

ARTICLE INFO

Article history:
Received 28 December 2016
Received in revised form 8 August 2017
Accepted 15 August 2017
Available online 9 September 2017

Keywords:
Train timetabling
Routing
Heterogeneous train traffic
Switchable scheduling rule

ABSTRACT

This paper proposes a mathematical model for the train routing and timetabling problem that allows a train to occasionally switch to the opposite track when it is not occupied, which we define it as switchable scheduling rule. The layouts of stations are taken into account in the proposed mathematical model to avoid head-on and rear-end collisions in stations. In this paper, train timetable could be scheduled by three different scheduling rules, i.e., no switchable scheduling rule (No-SSR) which allows trains switching track neither at stations and segments, incomplete switchable scheduling rule (In-SSR) which allows trains switching track at stations but not at segments, and complete switchable scheduling rule (Co-SSR) which allows trains switching track both at stations and segments. Numerical experiments are carried out on a small-scale railway corridor and a large-scale railway corridor based on Beijing–Shanghai high-speed railway (HSR) corridor respectively. The results of case studies indicate that Co-SSR outperforms the other two scheduling rules. It is also found that the proposed model can improve train operational efficiency.

1. Introduction

While rail transportation provides economic mobility for both passengers and freight across the world, its efficiency faces an unprecedented challenge because its limited capacity receives increasingly growing passenger and freight transportation demands (Xu et al., 2015). Urban railway transit system also faces such issues (He et al., 2016). This highlights the need for planning efficient train timetables and routes that can best utilize the limited railway capacity throughout the time while guarantying the safety operation. One outstanding challenge to these operations is scheduling heterogeneous trains (e.g., fast trains and slow trains) on high-speed railway with aims to meet different customers' transportation demands. However, such arrangements likely cause excessive travel delay, particularly when a fast train follows a slow one. To address this challenge, Mu and Dessouky (2013) firstly proposed a switchable dispatching strategy (SDS) for a double-track railway link, which essentially enables a fast train to overtake its slow front train by using the opposite directional track when it is in vacancy. Their results showed that the SDS can be able to reduce a fast train's knock-on delay by as much as 30% compared with the dedicated dispatching policy. Inspired by this, we propose a mathematical model to obtain the optimal scheduling solutions for heterogeneous trains on HSR corridors with consideration of switchable scheduling rules based on the SDS in Mu and Dessouky (2013), aiming to produce efficient timetables without infrastructure extensions. With this model, it can

* Corresponding author.
E-mail address: bjia@bjtu.edu.cn (B. Jia).

http://dx.doi.org/10.1016/j.trc.2017.08.010

find the optimal scheduling solutions while the slow and fast trains have optional chance to be scheduled on the opposite track, which is different from previous work where the faster trains were always dispatched on opposite track with SDS (Mu and Dessouky, 2013).

1.1. Literature review

Generally, the goal of the train timetable problem (TTP) is to minimize railroad system operational costs (often measured by the total train delay) by optimally scheduling these arrival, departure times as well as the orders inside and outside stations without causing collision risks or violating certain side constraints (Hansen and Pachl, 2014). TTP is an important issue for train operation and is also known as a hard problem due to large problem scales and complex problem structure. Numerous research has been conducted to solve TTP with mathematical programming techniques and heuristic algorithms in the past few decades (Bešinović et al., 2016; Cacchiani et al., 2016; Zhou and Zhong, 2005, 2007). Cacchiani and Toth (2012) presented an overview of the main works on nominal and robust train timetabling problems. Recently, Caimi et al. (2017) surveyed practical applications and the combinatorial optimization models for railway timetable problem.

While, the train routing problem (TRP) is to select a sequence of tracks for a train from its origin to destination, with the objective of minimizing the knock-on/secondary delay and/or increasing the capacity of railway networks (Mu and Dessouky, 2013). To obtain the optimal solutions for train operations, TTP and TRP are often considered simultaneously as one joint optimization problem, i.e., train timetabling and routing problem (TTRP) (Caimi et al., 2011; Lamorgese et al., 2016; Lee and Chen, 2009; Meng and Zhou, 2014; Samà et al., 2017; Zhou and Teng, 2016). Lee and Chen (2009) proposed a fast heuristic algorithm for the TTRP, where the operation time of trains depends on the assigned track, and the minimum headway between the trains depends on the trains' relative status. Caimi et al. (2011) considered sequentially timetabling and routing problem, at the macroscopic level, the timetable was produced by a periodic event scheduling problem model with continuous variables rather than the event times, then adopted Resource Tree Conflict Graph model to resolute the conflict at stations at the microscopic level. Similarly, Lamorgese et al. (2016) proposed an exact micro–macro approach to cyclic and non-cyclic train timetabling, where the routing problem involved the track choices at stations in the decomposition problem and the master problem was the line problem that finds the arrival/departure time at each station. For scheduling the freight trains, Mu and Dessouky (2011) considered fixed and flexible paths to enhance utilization rates of infrastructures. For real-time train traffic management, Törnquist and Persson (2007) first present a mathematical formulation based on a discrete-event theory for train rescheduling and rerouting problem, then they developed it in an optimization-based computational re-scheduling support for railway traffic networks (Törnquist Krasemann, 2015). On the basis of discrete-event theory, Qi et al. (2016) present a track choice-based bi-level formulation for integrating multi-track station layout design and train scheduling models problem. Meng and Zhou (2014) reformulated the train rerouting and rescheduling on an N-track network by considering a vector of cumulative flow variables and provided an efficient decomposition solution algorithm based on Lagrangian relaxation. Zhou and Teng (2016) furtherly built an integer linear programming for the simultaneous passenger train routing and timetabling problem in a space-time discretized network. A Lagrangian relaxation decomposition framework together with a heuristic method was proposed to solve it. Samà et al. (2017) designed a variable neighborhood search algorithm for fast train scheduling and routing during disturbed rail traffic situations. Fang et al. (2015) surveyed the recent models and methods on train rescheduling in railway networks.

Over the last decade, simulation techniques emerged in TTRP area. Dorfman and Medanic (2004) proposed a discrete-event model (DEM) to schedule the two-way train traffic on a single-track railway line. In this work, train movements were controlled by local feedback-based travel advance strategies. Their work was further improved by Li et al. (2008) into a global feedback-based travel advance strategy for train timetabling problem. Later, an optimal scheduling model with DEM was proposed to find the optimal velocity for each train running on the railway line (Xu et al. (2014). They further applied the DEM to scheduling heterogeneous traffic on high-speed railway corridor considering train switchable scheduling rules (Xu et al., 2015). Besides, Xu et al. (2016) proposed an improved discrete-time model for heterogeneous high-speed passenger train movements. All these studies focus on numerical heuristic solutions by assigning tracks to trains according to local information. Despite the breakthroughs of these studies, they do not ensure solution optimality.

1.2. Main focuses of this study

Although most existing studies consider arrival/departure times and routes along train trips, the switchable scheduling rules (SSRs) in the train timetable design process are rarely investigated. Besides, most existing literature addressed the train timetabling problem neglected the layouts of stations (Qi et al., 2016; Törnquist and Persson, 2007; Törnquist Krasemann, 2015), which may cause infeasible train timetables. Therefore, this study intends to provide the following contributions to the train routing and timetabling problem optimization methods.

(i). Based on the train switchable dispatching rules in Mu and Dessouky (2013, 2014), we first propose an integer linear programming formulation for the train routing and timetabling problem with switchable scheduling rule (SSR). Different from the methodology of queue theory in Mu and Dessouky (2013) that could be applied to a small-scale network, and the heuristic simulation method in Xu et al. (2015) that uses the local information-based switchable dispatching rules to identify feasible solutions, the proposed mathematical models in this paper can obtain a feasible exact optimal solution or a near-optimum solution of the train routing and timetabling problem. In this model, we assume that a train traveling on an N-track section has two choices: (a) traveling on the designated track, and (b) using the vacant reverse direction track.

(ii). As we consider the routing and timetabling problem for heterogeneous train traffic at a macroscopic level, the maximum speeds of trains over the dedicated and opposite tracks are assumed to be equal. In the previous works, Caimi et al. (2011) considered the timetabling problem at the macroscopic level and routing problem in stations at the microscopic level sequentially, while Lamorgese et al. (2016) decomposed the TTRP and considered the routing problem in stations. Different from their work, we consider the train routing and timetabling problem simultaneously. Since the routing problem involves the track choices both at stations and segments in this paper, the layouts of stations should be considered carefully to avoid train collisions. Thus, we define one pair of binary matrixes to depict the layout of each station, to formulate the safety operation constraints at stations. We refine the train safety operation constraints at stations and segments that further helps to improve computational efficiency. Furthermore, the objective is modified to minimize the deviations from the ideal timetable while reducing the track switch times.

(iii). To demonstrate the effectiveness of the SSR, we present two other scheduling rules, i.e., No-SSR and In-SSR, for the TTRP. For each scheduling rule, we provide mathematical details of the constraints and objective function.

(iv). Two sets of numerical experiments with a seven-station railway corridor and a large-scale railway corridor based on the Beijing–Shanghai high-speed railway corridor in China are performed to illustrate the applications and to evaluate the performance of the proposed methods. We employ the CPLEX solver to solve the proposed models. The results of the small-scale case studies demonstrate the correctness of the proposed models. Further, in comparison to the other scheduling rules in large-scale case studies, the Co-SRR outperforms the other two scheduling rules.

Note that it is worth investigating ways of building such an exact model for the following two reasons. On the academic side, an exact model can provide a better understanding of TTRP with SSR structure and can reveal theoretical insights into how much improvement SSR can make at maximum for various problem instances. On the practical side, there might still exist significant room for improvement from current heuristic solutions, and any improvement in rail operation practices, even small in percentage, is worth pursuing, because it may translate into substantial cost savings and social benefits due to the enormous mass of this industry. Although the switchable scheduling rule is not implemented in real railway operation, it is worth researching the efficiency and feasibility of this emerging technology in advance. With the emerge of advanced infrastructure and efficient signaling systems, it is possible to implement these scheduling rules in practice, especially in the cases where the numbers of trains between two directions have a larger difference.

The rest of this paper is organized as follows. Section 2 describes problem setting, problem statement, and the notations. Section 3 presents the simultaneous train routing and timetabling mathematical formulation for three different scheduling rules. Section 4 provides two sets of numerical examples, in which a seven-station railway corridor and a large-scale railway corridor based on Beijing–Shanghai high-speed railway corridor in China are considered to evaluate the performance of the proposed mathematical model. Finally, concluding remarks and future research directions are given in Section 5.

2. Problem description

2.1. Problem setting

2.1.1. Infrastructure

We consider a single railway corridor as illustrated in Fig. 1. This railway corridor consists of some sections indexed by $j \in J := \{1, 2, \ldots, J\}$. Each section is either a station where trains load and unload passengers, or a segment that connects two consecutive stations. We denote the type of section j as s_j, where $s_j = 0$ if section j is a station, or $s_j = 1$ if section j is a segment between two consecutive stations. Further, let P_j be the set of tracks in section j and t_{jn} be the n^{th} track in section j. We assume that all tracks in $P_j, \forall j \in J$ are bi-directional so that they can be occupied by trains running in both directions. Without loss of generality, we call the direction from left to right as inbound and the corresponding tracks as inbound (i.e., track 1 at segments and tracks 1 and 2 at stations in Fig. 1). We define the other direction as outbound and the corresponding tracks as outbound (i.e., track 2 at segments and tracks 3 and 4 at stations in Fig. 1). Further, for the inbound trains, we set the inbound tracks (e.g., track 1 at segments in Fig. 1) as the dedicated tracks, while the outbound tracks (e.g., track 2 at segments in Fig. 1) are opposite tracks. Similarly, the dedicated and opposite tracks are defined for the outbound trains.

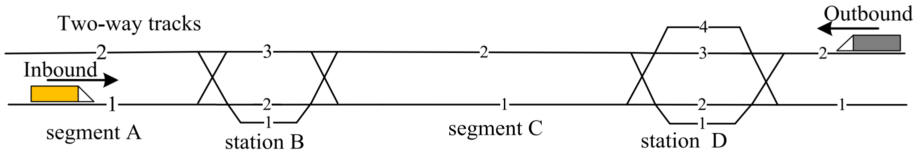

Fig. 1. A rail network with two stations and two segments.

2.1.2. Trains

We consider a set of trains moving along the railway corridor in both directions. We index these trains by $i \in I := \{1, 2, \ldots, I\}$. We assume that these trains are categorized into two types: slow and fast trains, and these two types of train sets are denoted by I_f and I_s, respectively. With the definition of discrete events in Törnquist and Persson (2007), each train $i \in I$ passing through each section $j \in J$ is considered as an independent event denoted by $k \in K := \{1, 2, \ldots, K\}$. Thus, each event connects one train and one section, and the traveling process of trains from the origin to the destination stations are composed of a number of discrete events. Fig. 2(a) illustrates the time-space diagram for two trains moving in the opposite directions, and Fig. 2(b) shows the corresponding event diagram of these two trains.

Let K_i be the event set for train i and L_j be the event set for segment j. As it is shown in Fig. 2(b), the event set for train 1 moving from segments A to D consists of independent events {1, 2, 3, 4} in K_1, and the corresponding event set for train 2 moving from segments D to A in the opposite direction is {5, 6, 7, 8} in K_2, while the event set related to segment A is {1, 8} in L_A. As illustrated in Fig. 2(b), event 3 is related to train 1, station C, and the train utilizes track 1 of station C, while event 6 is related to train 2, station C and it occupies track 2 in station C. Apparently, there is no conflict between events 3 and 6 since they occupy different tracks at the same station. Thus each independent event $k \in K$ corresponds to one train and one section.

The decision variables are defined as follows. Integer variables x_k^{begin} and x_k^{end} are defined as the entering and exiting times of t rain i to and from section j, respectively. Binary decision variables q_{kn} indicate the track choice, i.e., $q_{kn} = 1$ if train i chooses track t_{jn} to traverse section j, or 0 otherwise.

2.1.3. Switchable scheduling rule (SSR)

In the timetable design process, the traditional track assignment within the railway corridor is based on the dedicated scheduling rule, in which trains in one direction can only use the track assigned for the same direction and are forbidden from using the track in the opposite direction, even when the opposite track is in vacant. Different from the dedicated scheduling rule, switchable scheduling rule (SSR) is developed from the switchable dispatching rules that were proposed to reschedule trains in real-time train traffic management problems (Mu and Dessouky, 2013, 2014; Törnquist and Persson, 2007; Törnquist Krasemann, 2015). SSR allows trains to travel on the opposite vacant track to complete overtaking. Fig. 3 gives an example of SSR in a simple railway network. As illustrated in this figure, slow train u and fast train u^+ travel in the outbound direction and train i travels in the inbound direction. With the dedicated scheduling rule, train i can only use track 1 in segments A and C, while train u and train u^+ can only use outbound track 2 in segments C and A. At the time t_1, train u travels through segment C followed by a fast train u^+ (see Fig. 3(a)). If train u^+ is about to catch up with train u in segment C, train u^+ would be delayed as it should reduce its speed while traveling in segment C or dwell more time at station D to avoid rear-end collision with its front train u. According to SSR, when train u is entering track 2 in segment C and train i is moving on inbound track 1 in segment A, and if the entering time of train i to track 1 in segment C is later than the exiting time of train u^+ from the same track, the outbound fast train u^+ can be scheduled on inbound track 1 in segment C (see Fig. 3(b)), then return to outbound track 3 at station B without any delay (see Fig. 3(c)). This example shows that SSR can improve the performance of timetable.

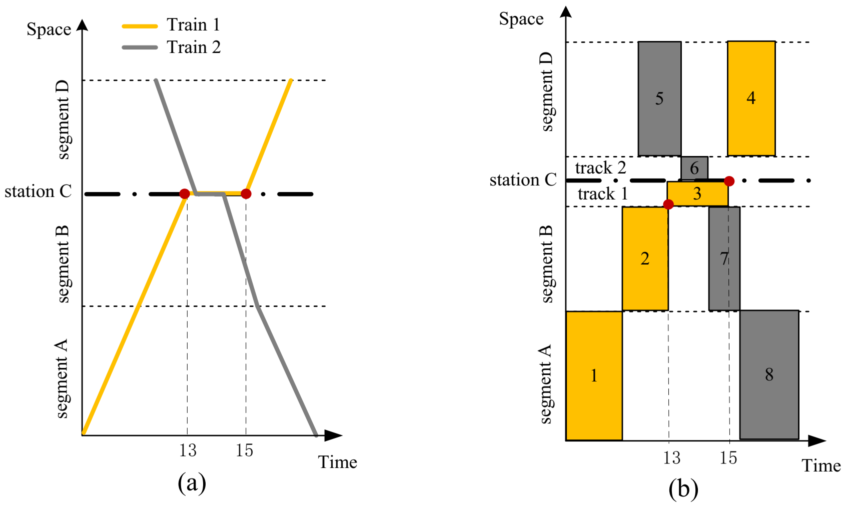

Fig. 2. (a) An illustrative time-space diagram for a railway corridor and its train traffic; (b) the corresponding event diagram, which shows how resources (i.e., segments A-D) are allocated to the trains.

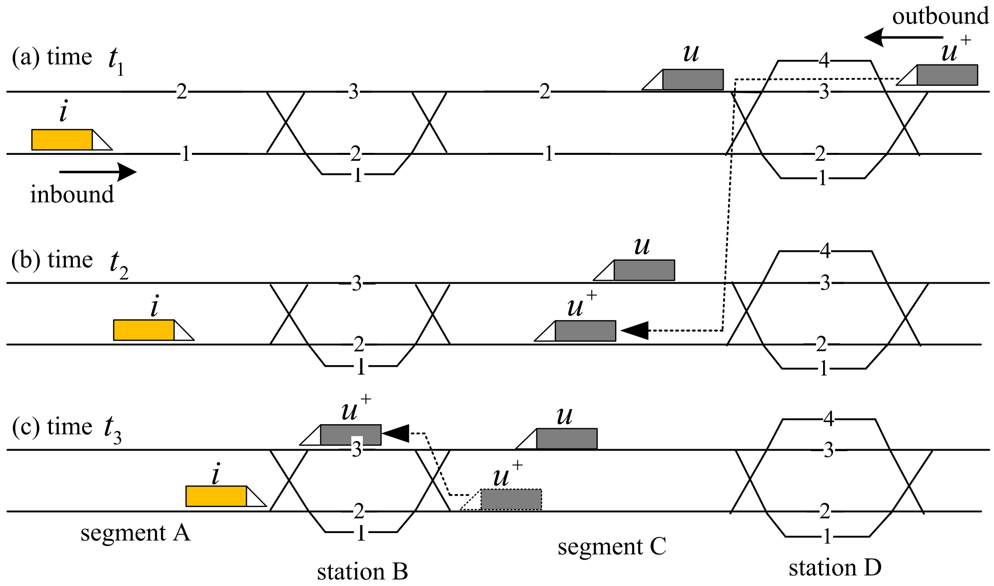

Fig. 3. An illustrative example of SSR.

2.1.4. Routes

A train route is defined as the sequence of tracks along its visited segments from the origin to the destination stations. The route of train i is composed of decision variables q_{kn}, where $k \in K_i \cap L_j, n \in P_j, i \in I, j \in J$. Fig. 4 shows an illustration of alternative routes in a railway corridor of Fig. 1. As illustrated in Fig. 4, there are four candidate routes for a train starting from track 1 in segment A to track 1 in station D. As an example, route 1 consists of track $\{t_{A1}, t_{B3}, t_{C2}, t_{D1}\}$. Apparently, without the switchable scheduling rule, routes 1 and 2 are not available for the inbound trains that depart from segment A to segment D.

2.2. Problem statement

In the previous works that adopted the discrete event mathematical model, e.g. Qi et al. (2016), Törnquist and Persson (2007), and Törnquist Krasemann (2015), most of the authors have considered the train conflicts when the two trains occupied the same track at a certain segment, as shown in Fig. 5(a). In Fig. 5(a), T1 and T2 selected the same track to dwell at station B, in this situation, the two trains have a move sequence, i.e., T1 first leaves then T2 arrives later (see Fig. 5(b)) or T2 first arrives then T1 leaves later (see Fig. 5(c)). However, to thoroughly dealing with all the collisions at stations, it is not enough to only consider the situations with two consecutive trains visiting the same track. For example, as shown in Fig. 6(a), when train T2 from the outbound direction is backing to the outbound track of segment A from the southernmost track of station B, it may have a head-on collision with inbound train T1, which is entering into station B from the inbound track of segment A. In this condition, the two involved trains should have a move sequence and respect a time separation, i.e., T1 first arrives then T2 leaves later (see Fig. 6(b)) or T2 first leaves then T1 arrives later (see Fig. 6(c)). Similarly, trains traveling in the same direction may also have a potential rear-end collision despite they occupying different track on the same segment. As shown in Fig. 7, where T3, T4 travel in the inbound direction, and when T4 is entering track 2 in segment E and T3 is entering track 1 in segment E, they may have a rear-end collision at the end of station D, so they should be separated by a minimum time interval to avoid a collision. To our knowledge, few studies have considered this type of problem.

To resolve these potential collisions between two consecutive trains in situations of they visiting different tracks at stations (e.g. in Figa. 6(a) and 7), we should consider the layouts of stations carefully. Generally, the joint of two tracks in stations can be classified into two types. One type is turnout where allows trains to run over one track or another, and the other

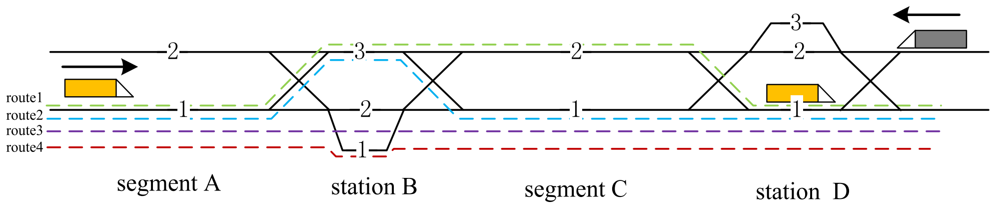

Fig. 4. An example of a railway corridor with four routes.

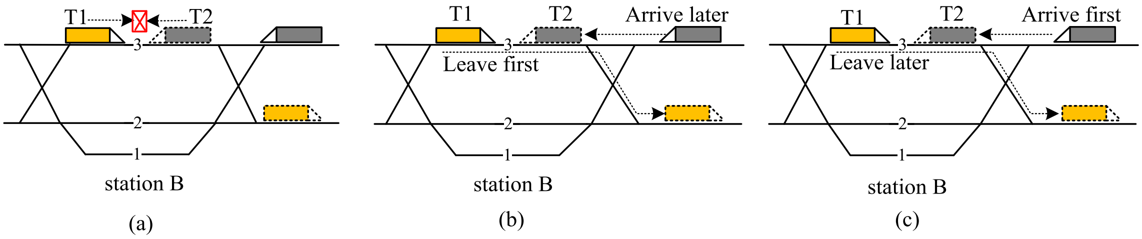

Fig. 5. Conflict resolution when (a) two consecutive trains visiting the same track at stations, where (b) is T1 leave first, T2 arrive later; (c) is T2 arrive first, T1 leave later.

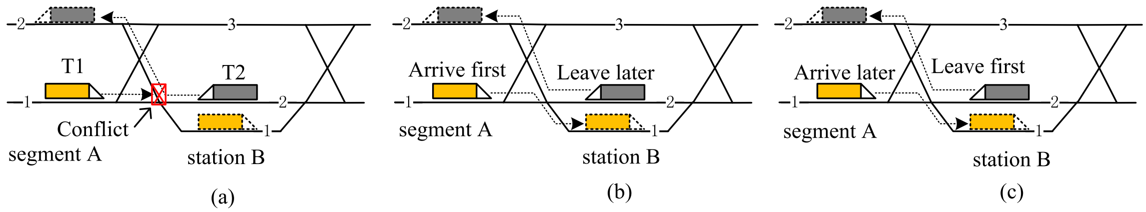

Fig. 6. Conflict resolution when (a) two consecutive trains visiting different track at stations, where (b) is T1 arrive first, T2 leave later; (c) is T2 leave first, T1 arrive later.

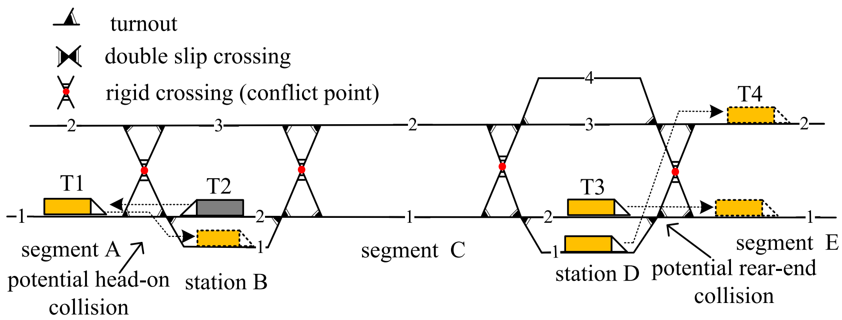

Fig. 7. Examples of crossings in different types of station.

type is crossing that effects two tracks to cross at grade (Pachl, 2014). Further, the crossings could be a slip or rigid crossing depending on whether allows train to switch track, i.e., slip crossing is equipped with additional points providing a slip connection to permit movements from one track to another but rigid crossing not (turnouts and slip crossing are usually referred to as switches in North American railway). A crossing with a slip connection at one side is called a 'single slip', and a crossing with slip connections at both sides is called a 'double slip'. In this work, the slip crossings are assumed as double slip crossing. For trains through these turnouts and crossings, there are two different situations. One is that both of two consecutive trains intend to pass the turnouts or slip crossings but neither of them passes the rigid crossings, i.e., the referred two trains would occupied the same track on the following segment which links with the turnouts or slip crossings. In this situation, the time intervals between the two trains can be guaranteed by restricting the headway constraints on the following segment (as shown in Fig. 7, where both T1 and T2 occupy inbound track 1 to through segment A). The other is that one (or both) of the two consecutive trains intend to the rigid crossing (as shown in Fig. 6 (a), where T2 is backing to its outbound track at segment A but T1 is from its inbound track at segment A). In the latter situation, the two trains should be separated by a minimum time intervals at the end station B to avoid a head-on collision.

Based on above analysis, we know that rigid crossing may cause train head-on or rear-end collision, which cannot be avoided by the safety constraints on the next segment, so we define the rigid crossings as conflict points in this paper. For depicting these conflict points in stations, we define one pair of incidence matrix, i.e., the first one is the incidence matrix of the tracks from segment j-1 to station j, denoted by $c^{nm}_{j,l-s}$, $n \in P_{j-1}, m \in P_j, s_j = 0, j \in J \setminus \{1\}$, which shows the layout of the conflict point at the upstream of station j, and the other one is the incidence matrix of the tracks from station j to segment j+1, denoted by $c^{mn}_{j,s-l}$, $n \in P_{j+1}, m \in P_j, s_j = 0, j \in J \setminus \{|J|\}$, which shows the layout of conflict point at the downstream of station j. In the incidence matrix of $c^{nm}_{j,l-s}$, the rows represent the track index at segment $j-1$, and the columns represent the track index in station j. $c^{nm}_{j,l-s} = 1$ if the conflict point is on the way of trains from track n at segment $j-1$ to track m in station j, and

0 otherwise. Similarly, at the downstream of station j, $c^{mn}_{j,s-l} = 1$ if the conflict point is on the way of trains from track m in station j to track n at segment $j + 1$, and 0 otherwise. For instance, the following binary matrices show the layout of conflict points in stations B and D in Fig. 7, respectively.

Table 1
Indices, input parameters in this problem.

Notations	Definition
I	Set of trains
I_f, I_s	Set of fast trains and slow trains, respectively
i, u	Index of trains
J	Set of sections
j	Index of sections
K	Set of events
$k, \hat{k}$	Indices of discrete events
b_i, l_i	The first and last event of train i, respectively
K_i, L_j	Set of events related to train i and section j, respectively
P_j	Set of tracks in section j
P^1_j, P^2_j	Set of inbound and outbound tracks in section j
m, n	Indices of track
t_{jn}	The n^{th} track in section j
$c^{nm}_{j,l-s}$, $c^{mn}_{j,s-l}$	The incidence matrix of the tracks from segment j-1 to station j, and that of the tracks from station j to segment $j + 1$ for the inbound trains, respectively
$c^{nm}_{j,l-s,1}$, $c^{mn}_{j,s-l,1}$	The incidence matrix of the tracks from segment $j + 1$ to station j, and that of the tracks from station j to segment $j-1$ for the outbound trains, respectively
H	Departure time interval
d_k	Minimum duration time of event k, i.e., the minimum traveling time at segment or dwell time at station
s_j	Section j attribute indicator ($s_j = 1$ if section j is a segment between stations, and 0 otherwise)
α, β	Times corresponding to acceleration and deceleration of trains, respectively
ε_f, ε_s	Weights of one-time switching cost for the fast and slow trains, respectively
$b^{departure}_i$	Planned departure time from origin station for train i
o_k	Travel direction indicator of event k ($o_k = 1$ if event k is an outbound event, and 0 otherwise)
Δ^{meet}_j	Meet headway time, which is the time interval between two involved trains when the two trains visiting the same track at section j or the two trains visiting different track but have potential collisions at stations
$\Delta^{dd}_j, \Delta^{aa}_j$	Departure headway time, and arrival headway time between two successive trains running in the same directions when the two trains visiting the same track in section j, respectively
$a^f_{l_i}, a^o_{l_i}$	Arrival time for train i at its destination in ideal timetable (without consideration of any conflict with other trains and trains travel in free-travel situations) and in the feasible timetable (with consideration of train interact and trains travel in obstacle situations), respectively
$a^f_{L,Des}, a^o_{L,Des}$	Arrival time for the last train at its destination in the ideal and feasible timetable, respectively

Table 2
Decision variables in this problem.

Notations	Definition
x^{begin}_k, x^{end}_k	Integer variables, indicating the beginning and ending times of event k, respectively, and these decision variables constitute train timetable
q_{kn}	Binary variables, indicating event k utilizes track n or not ($q_{kn} = 1$ if event k chose track t_{jn} to through section j, $k \in L_j, n \in P_j, j \in J$, and 0 otherwise), and these decision variables constitute train routes
θ_k	Binary variables, indicating event k attribute indicator variable ($\theta_k = 1$ if event k is a stopped event, and 0 otherwise), and they are used to estimate whether it should consider the acceleration and deceleration time
p_k	Binary variables, indicating event k switching track or not ($p_k = 1$ if event k is scheduled on opposite track and the switching cost should be considered, and 0 otherwise), and they are added to objective function to control the track switch times
$\gamma_{k\hat{k}}, \lambda_{k\hat{k}}$	Binary variables, indicating the event k and $\hat{k}$ move sequence (when event k and $\hat{k}$ visiting the same track, we have $\gamma_{k\hat{k}} + \lambda_{k\hat{k}} = 1$, i.e., $\gamma_{k\hat{k}} = 1$ and $\lambda_{k\hat{k}} = 0$ if event k happens before event $\hat{k}$, $k, \hat{k} \in L_j, j \in J$, or $\gamma_{k\hat{k}} = 0$ and $\lambda_{k\hat{k}} = 1$ if event k happens after event $\hat{k}$, $k, \hat{k} \in L_j, j \in J$; when event k and $\hat{k}$ visiting different track, we have $\gamma_{k\hat{k}} + \lambda_{k\hat{k}} = 0$), and they are used to activate the headway constraints
$C_{k,u}, C_{k,d}$	Binary variables, indicating event k across conflict point or not in stations ($C_{k,u} = 1$ if event k goes across the active conflict point at the upstream of station j, $k \in L_j, j \in J$, and 0 otherwise; $C_{k,d} = 1$ if event k goes across the active conflict point at the downstream of station j, $k \in L_j, j \in J$, and 0 otherwise.) and they are used to estimate whether the two referred trains should satisfy the time intervals
$h_{k\hat{k},u}, f_{k\hat{k},u}$	Binary variables, indicating the event k and $\hat{k}$ move sequence at the upstream of stations (when at least one of event k and $\hat{k}$ across the conflict point at upstream of station, we have $h_{k\hat{k},u} + f_{k\hat{k},u} = 1$, i.e., $h_{k\hat{k},u} = 1$ and $f_{k\hat{k},u} = 0$ if event k happens before event $\hat{k}$, or $h_{k\hat{k},u} = 0$ and $f_{k\hat{k},u} = 1$ if event k happens after event $\hat{k}$; otherwise, we have $h_{k\hat{k},u} + f_{k\hat{k},u} = 0$) and they are used to guarantee the time intervals at the upstream of stations
$h_{k\hat{k},d}, f_{k\hat{k},d}$	Binary variables, indicating the event k and $\hat{k}$ move sequence at the downstream of stations (when at least one of event k and $\hat{k}$ across the conflict point at downstream of station, we have $h_{k\hat{k},d} + f_{k\hat{k},d} = 1$, i.e., $h_{k\hat{k},d} = 1$ and $f_{k\hat{k},d} = 0$ if event k happens before event $\hat{k}$, or $h_{k\hat{k},d} = 0$ and $f_{k\hat{k},d} = 1$ if event k happens after event $\hat{k}$; otherwise, we have $h_{k\hat{k},d} + f_{k\hat{k},d} = 0$) and they are used to guarantee the time intervals at the downstream of stations

$$c_{B,l-s} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, c_{B,s-l} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{1}$$

$$c_{D,l-s} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}, c_{D,s-l} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}, \tag{2}$$

In this example, the value of the first row and the third column is 1 in the binary matrix of $c_{D,l-s}$, which implies that trains from track 1 at segment C to track 3 in station D must go across the conflict point at the upstream of station D.

Since the outbound trains are always from segment $j+1$ to station j, then to segment $j-1$. So, for the outbound trains, we define the incidence matrix of the tracks from segment $j+1$ to station j as $c^{nm}_{j,l-s,1}$ and the incidence matrix of the tracks from station j to segment $j-1$ as $c^{mn}_{j,s-l,1}$. Since the tracks in the considered corridor are bidirectional, we can use the method of matrix transpose to depict the layouts of stations in the outbound direction. For instance, the incidence matrix of tracks from segment $j+1$ to station j is $c^{nm}_{j,l-s,1} = c^{mn\ T}_{j,s-l}$ and that from station j to segment $j-1$ is $c^{mn}_{j,s-l,1} = c^{nm\ T}_{j,l-s}$

2.3. Notations

For the reader's convenience, the notation system including indices, input parameters, and decision variables is described in Tables 1 and 2, respectively.

3. Simultaneous train routing and timetabling model

This section presents the simultaneous train routing and timetabling mathematical model. Firstly, the assumptions are listed as follows:

Assumption 1: All tracks in railway network are bi-directional and can be occupied by trains in both directions at different timestamps, i.e., a train can use opposite track for running.
Assumption 2: A train's travel times on a segment along a track is independent of whether this train switches the track before or after this segment.
Assumption 3: Each train has a pre-determined stop plan along with its trip, and it must stop as required for loading or unloading passengers.
Assumption 4: Each train has a pre-determined earliest departure time and cannot depart from its origin station earlier than the pre-determined departure time.

3.1. Model formulation for train routing and timetabling problem with switchable scheduling rule

- ***Objective function***

Practical train scheduling and routing problems may be associated with different objectives depending on applications. Specifically, planning applications are concerned with generating an efficient train timetable, e.g., minimizing the costs or the total passenger travel time, while real-time rescheduling applications aim to adjust the daily or hourly train operation schedules to recover the delay as soon as possible. In this paper, we mainly focus on exploring the effectiveness of switchable scheduling rule to improve train operational efficiency in the planning stage. Therefore, we define the following objective function to minimize the train deviations (measured in minutes) between the trains' arrival time at the destination in the feasible and ideal timetables.

$$\min z = \sum_{i \in I} (x^{end}_{l_i} - a^{f}_{l_i}). \tag{3}$$

In the objective function (3), $x^{end}_{l_i}$ is the arrival time for train i, for all $i \in I$, at its destination in the feasible timetable after scheduling, and $a^{f}_{l_i}$ is the arrival time for train i, for all $i \in I$, at its destination station in ideal train timetable. In ideal train timetable, the arrival/departure times for trains at each station they visited are equal to the pre-determined departure times simply plus the sum of minimum running time at each visited section, the dwell time at each station and the acceleration/deceleration time, here conflicts among trains are not considered. The ideal timetables are only used as the reference timetables to evaluate the performance of train timetable scheduled by different scheduling rules, and are not feasible as they are obtained by relaxing the safety operation constraints as well as the capacity limitation constraints.

Further, we define the track switch times to count the number of times that a train switch to its opposite track. In other words, if a train selects its opposite track to traverse a section, this can be seen as one track switch time. In practical railway operation, fewer track switch times on segments are desirable as it reduces the operating difficulties for train operators. This inspires us to find a way to control track switch times with Co-SSR and In-SSR. Therefore, a switching cost can be added to the objective function to control the track switch times as follows

$$\min z = \sum_{i \in I} (x_{l_i}^{end} - a_{l_i}^{f}) + \sum_{k \in K_i, i \in I} p_k, \tag{4}$$

where p_k is a binary variable that is used to express whether event k is scheduled on its opposite track or not. If event k should be scheduled on its opposite track of the segment (i.e., $p_k = 1$), a switching cost will be added to the objective function. The value of p_k can be calculated with the following equations

$$\begin{cases} \sum_{n \in P_j^2} q_{kn} \leqslant M \cdot p_k, k \in L_j, j \in J, o_k = 1, \\ \sum_{n \in P_j^1} q_{kn} \leqslant M \cdot p_k, k \in L_j, j \in J, o_k = -1. \end{cases} \tag{5}$$

In constraints (5), we can see that if event k is related to an inbound train (i.e., $o_k = 1$), and the train travels on any one of its opposite tracks (i.e., the outbound direction track), then $\sum_{n \in P_j^2} q_{kn} = 1$, and we obtain $p_k = 1$. Whereas if the train travels on its designated track (i.e., the inbound direction track in this example), then $\sum_{n \in P_j^2} q_{kn} = 0$, and the penalty index p_k can be set to 0 or 1. Since this problem is a minimization problem, we obtain $p_k = 0$, and no switching cost for event k is added into the objective.

The fast trains were always scheduled on the opposite track if they were able to catch their front slow trains (Mu and Dessouky, 2013). However, we can change the penalty coefficient weight for different types of trains in the objective function to schedule the preferred trains to the opposite vacant tracks when potential conflict may occur. Thus we reformulate the objective as

$$\min z = \sum_{i \in I} (x_{l_i}^{end} - a_{l_i}^{f}) + \sum_{k \in K_i, i \in I_f} \varepsilon_f \cdot p_k + \sum_{k \in K_i, i \in I_s} \varepsilon_s \cdot p_k. \tag{6}$$

Based on the weight of the switching cost in Eq. (6), we are able to schedule the preferred trains to the opposite vacant tracks. For instance, if $\varepsilon_f = 1$ and $\varepsilon_s = 0.5$ (i.e., $\varepsilon_f > \varepsilon_s$), the optimal solution is to schedule the slow train to the vacant track in its opposite direction when scheduling train on the opposite track is inevitable. Namely, through the penalty parameters, we can flexibly control the track switch times and schedule the preferred trains to the opposite tracks. On the other hand, switching train track may influence the robustness of the timetable so that greater penalty parameters may be better in some conditions. However, the focus of this paper is to investigate the effectiveness of the switchable scheduling rule, thus for simplicity, we set the same parameter values for the two types of trains in all the following case studies, i.e., $\varepsilon_f = \varepsilon_s = 0.5$.

- ***Constraints***

In the feasible train timetable design process, we should consider (i) train operational constraints, (ii) limited resource constraints, (iii) train operational safety constraints at segments, (iv) train operational safety constraints at stations and (v) decision variable feasible range constraints. More detailed formulation of each set of constraints are provided as follow.

(i) Train operational restrictions

Since we consider trains traversing a station or a segment as independent events, the movement of each train from the origin to its final station consists of a series of events and each event $k, k \in K$ has the properties of beginning time x_k^{begin}, ending time x_k^{end}, minimum duration time d_k, and the track choice q_{kn}. We formulate the train operational restrictions constraints as follows.

$$x_{k+1}^{begin} = x_k^{end}, k \in K_i, i \in I : k \neq l_i, \tag{7}$$

$$x_{b_i}^{begin} \geq b_i^{departure}, i \in I : b_i^{departure} \geqslant 0 \tag{8}$$

$$x_k^{end} \geq x_k^{begin} + d_k, k \in L_j, s_j = 0, j \in J, \tag{9}$$

$$x_k^{end} \geq x_k^{begin} + d_k + \alpha \cdot \theta_{k-1} + \beta \cdot \theta_{k+1}, k \in L_j, s_j = 1, j \in J, \tag{10}$$

$$x_k^{end} - x_k^{begin} \leq M \cdot \theta_k, k \in K_i, k \in L_j, s_j = 0, j \in J. \tag{11}$$

Constraint (7) ensures the coherence of trains traveling from the origins to the destinations. The beginning time of next event k + 1 should be equal to the ending time of current event k. Constraint (8) guarantees that a train cannot depart from its origin station earlier than its planning departure time $b_i^{departure}$. Constraint (9) ensures that trains dwell enough time as required at a station. Taking acceleration and deceleration time into account, the ending time of event k traveling in section j should be greater or equal to the beginning time of event k plus the acceleration or deceleration time in Constraint (10). For instance, if an upstream event of event k is a stopped event (i.e., $\theta_{k-1} = 1$), the ending time of event k should be equal to the beginning time of event k plus the acceleration time. However, since its downstream event *k + 1* is not a stopped event (i.e., $\theta_{k+1} = 0$), we obtain the ending time of event k as $x_k^{end} = x_k^{begin} + d_k + \alpha$. The beginning and ending time of events occurring in the station are already calculated through Constraint (9) so that whether event k is a stopped event or not (i.e., pass the station without stop) can be estimated by Constraint (11), where M is a large enough constant (which is set to 2000 in these experiments). For instance, if $x_k^{end} - x_k^{begin} > 0$, then train i is arranged to dwell at station j (i.e., $\theta_k = 1$). If $x_k^{end} - x_k^{begin} = 0$, then it represents that train i passes station j without stop, i.e., θ_k can be 0 or 1. However, since the model objective is to minimize the trains' total travel time we obtain $\theta_k = 0$. For example, if θ_k is set to 1, we obtain a greater x_{k+1}^{end} value (i.e., $x_{k+1}^{end} = x_{k+1}^{begin} + d_{k+1} + \alpha \cdot \theta_k + \beta \cdot \theta_{k+2}$) than if we set $\theta_k = 0$ ($x_{k+1}^{end} = x_{k+1}^{begin} + d_{k+1} + \beta \cdot \theta_{k+2}$). Apparently, setting $\theta_k = 0$ yields lower deviation from the free-travel timetable. In particular, we have $\theta_k = 1$ when event k has related a train at its origin and destination station.

(ii) The limited capacity restrictions

This constraint set considers the limited infrastructure capacity. When two consecutive trains plan to occupy the same track through a section (see Fig. 5(a)), they should satisfy a headway times (i.e., one train occupies this track a specific headway time later than the other one). So, we adopt binary variables $\gamma_{k\hat{k}}$ and $\lambda_{k\hat{k}}$ to imply the move sequence between the two trains, where k and $\hat{k}$ are the indices of events related to the two considered trains. $\gamma_{k\hat{k}}$ indicates that event k happens before $\hat{k}$ while $\lambda_{k\hat{k}}$ indicates event k happens after $\hat{k}$. For instance, we have $\gamma_{k\hat{k}} = 1$, $\lambda_{k\hat{k}} = 0$ in Fig. 5(b), and $\gamma_{k\hat{k}} = 0$, $\lambda_{k\hat{k}} = 1$ in Fig. 5(c), where event k is related to T1 and event $\hat{k}$ is related to T2. If the two trains occupy different tracks through this segment, we have $\gamma_{k\hat{k}} = 0$, $\lambda_{k\hat{k}} = 0$, it means the two trains should not satisfy a headway time in this situation. Thus, we formulate the following constraints to describe trains' track choices and the move sequence between the trains.

$$\sum_{n \in P_j} q_{kn} = 1, k \in L_j, j \in J, \tag{12}$$

$$\lambda_{k\hat{k}} + \gamma_{k\hat{k}} \leq 1, k, \hat{k} \in L_j, j \in J : k < \hat{k}. \tag{13}$$

Constraint (12) guarantees that each event k must utilize one and only one track to traverse section j, and event k can occupy any track of section j, regardless of the attribute of track direction. Constraints (13) indicates that $\gamma_{k\hat{k}}$ and $\lambda_{k\hat{k}}$ cannot be equal to 1 simultaneously. The following constraints formulate the relationship between the trains based on train track choice.

$$q_{kn} + q_{\hat{k}n} - 1 \leq \lambda_{k\hat{k}} + \gamma_{k\hat{k}}, k, \hat{k} \in L_j, n \in P_j, j \in J : k < \hat{k}. \tag{14}$$

Constraint (14) ensures that if the two trains choose the same track, we obtain $1 \leqslant \lambda_{k\hat{k}} + \gamma_{k\hat{k}}$. Combined with (13), it yields $\lambda_{k\hat{k}} + \gamma_{k\hat{k}} = 1$, which implies that either $\gamma_{k\hat{k}}$ or $\lambda_{k\hat{k}}$ should be equal to 1 in this case.

(iii) Train operational safety constraints at segments

On the railway corridor, safety constraints are modeled by the minimum headway times, including the departure and arrival headway times, to guarantee the safety operation of trains traveling in the same direction when they occupied the same track at segments. These constraints are listed as follows.

$$x_{\hat{k}}^{begin} - x_k^{begin} \geq \Delta_j^{dd} - M \cdot (1 - \gamma_{k\hat{k}}), k, \hat{k} \in L_j, o_k = o_{\hat{k}}, s_j = 1, j \in J : k < \hat{k}, \tag{15}$$

$$x_{\hat{k}}^{end} - x_k^{end} \geq \Delta_j^{aa} - M \cdot (1 - \gamma_{k\hat{k}}), k, \hat{k} \in L_j, o_k = o_{\hat{k}}, s_j = 1, j \in J : k < \hat{k}, \tag{16}$$

$$x_k^{begin} - x_{\hat{k}}^{begin} \geq \Delta_j^{dd} - M \cdot (1 - \lambda_{k\hat{k}}), k, \hat{k} \in L_j, o_k = o_{\hat{k}}, s_j = 1, j \in J : k < \hat{k}, \tag{17}$$

$$x_k^{end} - x_{\hat{k}}^{end} \geq \Delta_j^{aa} - M \cdot (1 - \lambda_{k\hat{k}}), k, \hat{k} \in L_j, o_k = o_{\hat{k}}, s_j = 1, j \in J : k < \hat{k}. \tag{18}$$

When the two trains are traveling in different directions and occupying the same track at segments, they should satisfy the following headway time safety constraints.

$$x_{\hat{k}}^{begin} - x_k^{end} \geq \Delta_j^{meet} - M \cdot (1 - \gamma_{k\hat{k}}), k, \hat{k} \in L_j, o_k \neq o_{\hat{k}}, s_j = 1, j \in J : k < \hat{k}, \tag{19}$$

$$x_k^{begin} - x_{\hat{k}}^{end} \geq \Delta_j^{meet} - M \cdot (1 - \lambda_{k\hat{k}}), k, \hat{k} \in L_j, o_k \neq o_{\hat{k}}, s_j = 1, j \in J : k < \hat{k}, \tag{20}$$

Constraints (15)–(18) guarantee the departure and arrival safety headway time when events k and $\hat{k}$ occur in the same direction (i.e., $o_k = o_{\hat{k}}$). Additionally, these constraints can guarantee that events k and $\hat{k}$ cannot change the running order within the same section j (i.e., first depart, first arrive). As the logical relationship between two successive trains depicted in constraints (13), $\lambda_{k\hat{k}}$ and $\gamma_{k\hat{k}}$ could not be equal to 1 simultaneously. Therefore, when events k and $\hat{k}$ are at the same track on the same segment, constraints (15) and (16) are active while constraints (17) and (18) are inactive and vise versa. Constraints (19) and (20) guarantee the meet headway time separation when events k and $\hat{k}$ occupy the same track through a segment or they intend to move through the same crossing in stations.

(iv) Train operational safety constraints at stations

At stations, trains may occupy the same track due to the limited capacity of stations. In these conditions, they should satisfy the following meet headway time constraints no matter the two consecutive trains traveling in the same direction or not.

$$x_{\hat{k}}^{begin} - x_k^{end} \geq \Delta_j^{meet} - M \cdot (1 - \gamma_{k\hat{k}}), k, \hat{k} \in L_j, s_j = 0, j \in J : k < \hat{k}, \tag{21}$$

$$x_k^{begin} - x_{\hat{k}}^{end} \geq \Delta_j^{meet} - M \cdot (1 - \lambda_{k\hat{k}}), k, \hat{k} \in L_j, s_j = 0, j \in J : k < \hat{k}, \tag{22}$$

In the two-way rail traffic flow, head-on collisions may happen when the two trains traveling in the opposite direction (see Fig. 6(a)). Moreover, a rear-end collision may occur when the two trains traveling in the same direction but entering different directional track at next segment (see Fig. 7(b)). Base on the analysis in Section 2.2, if one (or both) of the two trains are claiming to go across the conflict points, they must be separated by minimum time intervals to avoid a collision. Therefore, we define binary variables $C_{k,u}$ and $C_{k,d}$ to represent whether the trains go across the conflict point at the upstream and downstream of the stations, respectively. It is worth mentioning that the outbound trains always pass the downstream of the stations firstly and then enter the stations. Thus, $C_{k,u}$ and $C_{k,d}$ for both directional trains are considered at the following constraints.

$$q_{(k-1)n} + q_{km} + c_{j,l-s}^{nm} - 2 \leq M \cdot C_{k,u}, k \in L_j, o_k = 1, n \in P_{j-1}, m \in P_j, s_j = 0, j \in J \setminus \{1\}, \tag{23}$$

$$q_{km} + q_{(k+1)n} + c_{j,s-l}^{mn} - 2 \leq M \cdot C_{k,d}, k \in L_j, o_k = 1, m \in P_j, n \in P_{j+1}, s_j = 0, j \in J \setminus \{|J|\}, \tag{24}$$

$$q_{(k-1)n} + q_{km} + c_{j,l-s,1}^{nm} - 2 \leq M \cdot C_{k,d}, k \in L_j, o_k = -1, m \in P_j, n \in P_{j+1}, s_j = 0, j \in J \setminus \{|J|\}, \tag{25}$$

$$q_{km} + q_{(k+1)n} + c_{j,s-l,1}^{mn} - 2 \geq M \cdot C_{k,u}, k \in L_j, o_k = -1, n \in P_{j-1}, m \in P_j, s_j = 0, j \in J \setminus \{1\}. \tag{26}$$

Further, we define binary variables $h_{k\hat{k},u}, f_{k\hat{k},u}, h_{k\hat{k},d}, f_{k\hat{k},d}$ to represent the relationship between the trains at the same stations. With the binary variables of trains going (or not going) across the conflict points, i.e., $C_{k,u}$ and $C_{k,d}$, the relationship of the considered trains at the upstream and downstream of stations can be estimated by following constraints.

$$C_{k,u} + C_{\hat{k},u} \leq M \cdot (h_{k\hat{k},u} + f_{k\hat{k},u}), k, \hat{k} \in L_j, s_j = 0, j \in J : k < \hat{k}, \tag{27}$$

$$C_{k,d} + C_{\hat{k},d} \leq M \cdot (h_{k\hat{k},d} + f_{k\hat{k},d}), k, \hat{k} \in L_j, s_j = 0, j \in J : k < \hat{k}, \tag{28}$$

$$h_{k\hat{k},u} + f_{k\hat{k},u} \leqslant 1, k, \hat{k} \in L_j, s_j = 0, j \in J : k < \hat{k}, \tag{29}$$

$$h_{k\hat{k},d} + f_{k\hat{k},d} \leq 1, k, \hat{k} \in L_j, s_j = 0, j \in J : k < \hat{k}. \tag{30}$$

Fig. 8 illustrates the time intervals between two consecutive trains at different track of stations but they are at the risk of collisions. In these conditions, the involved two trains should be separated by a time intervals. For simplicity, we set the minimum time intervals between the two consecutive trains as the same as the meet headway time, i.e., 1 min.

From the illustration of the time interval between two consecutive trains traveling in the inbound direction in Fig. 8(a), we obtain the following constraints to avoid train collisions.

$$x_{\hat{k}}^{begin} - x_k^{begin} \geq \Delta_j^{meet} - M \cdot (1 - h_{k\hat{k},u}), k, \hat{k} \in L_j, o_k = o_{\hat{k}} = 1, s_j = 0, j \in J : k < \hat{k}, \tag{31}$$

$$x_k^{begin} - x_{\hat{k}}^{begin} \geq \Delta_j^{meet} - M \cdot (1 - f_{k\hat{k},u}), k, \hat{k} \in L_j, o_k = o_{\hat{k}} = 1, s_j = 0, j \in J : k < \hat{k}, \tag{32}$$

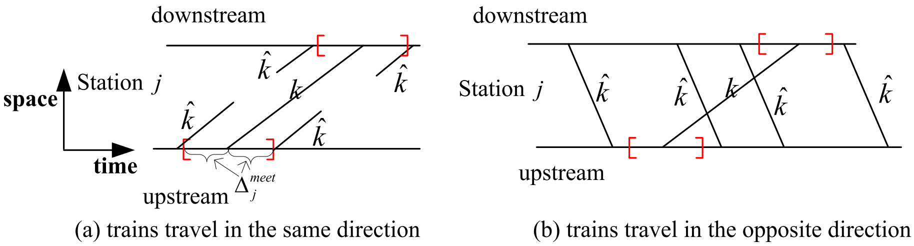

Fig. 8. Time intervals between two consecutive trains at stations when they are at the risk of collisions.

$$x_{\hat{k}}^{end} - x_k^{end} \geq \Delta_j^{meet} - M \cdot (1 - h_{k\hat{k},d}), k, \hat{k} \in L_j, o_k = o_{\hat{k}} = 1, s_j = 0, j \in J : k < \hat{k}, \tag{33}$$

$$x_k^{end} - x_{\hat{k}}^{end} \geq \Delta_j^{meet} - M \cdot (1 - f_{k\hat{k},d}), k, \hat{k} \in L_j, o_k = o_{\hat{k}} = 1, s_j = 0, j \in J : k < \hat{k}. \tag{34}$$

As the trips for outbound trains are opposite to inbound trains, the outbound trains first pass the downstream of stations and then enter the stations. Therefore, we have the following constraints to avoid train rear-end collisions in the outbound direction.

$$x_{\hat{k}}^{begin} - x_k^{begin} \geq \Delta_j^{meet} - M \cdot (1 - h_{k\hat{k},d}), k, \hat{k} \in L_j, o_k = o_{\hat{k}} = -1, s_j = 0, j \in J : k < \hat{k}, \tag{35}$$

$$x_k^{begin} - x_{\hat{k}}^{begin} \geq \Delta_j^{meet} - M \cdot (1 - f_{k\hat{k},d}), k, \hat{k} \in L_j, o_k = o_{\hat{k}} = -1, s_j = 0, j \in J : k < \hat{k}, \tag{36}$$

$$x_{\hat{k}}^{end} - x_k^{end} \geq \Delta_j^{meet} - M \cdot (1 - h_{k\hat{k},u}), k, \hat{k} \in L_j, o_k = o_{\hat{k}} = -1, s_j = 0, j \in J : k < \hat{k}, \tag{37}$$

$$x_k^{end} - x_{\hat{k}}^{end} \geq \Delta_j^{meet} - M \cdot (1 - f_{k\hat{k},u}), k, \hat{k} \in L_j, o_k = o_{\hat{k}} = -1, s_j = 0, j \in J : k < \hat{k}. \tag{38}$$

Similarly, when the two trains travel in the opposite directions and visit different tracks (see Fig. 8(b)), they should satisfy the following constraints.

$$x_{\hat{k}}^{end} - x_k^{begin} \geq \Delta_j^{meet} - M \cdot (1 - h_{k\hat{k},u}), k, \hat{k} \in L_j, o_k \neq o_{\hat{k}}, s_j = 0, j \in J : k < \hat{k}, \tag{39}$$

$$x_k^{begin} - x_{\hat{k}}^{end} \geq \Delta_j^{meet} - M \cdot (1 - f_{k\hat{k},u}), k, \hat{k} \in L_j, o_k \neq o_{\hat{k}}, s_j = 0, j \in J : k < \hat{k}, \tag{40}$$

$$x_{\hat{k}}^{begin} - x_k^{end} \geq \Delta_j^{meet} - M \cdot (1 - h_{k\hat{k},d}), k, \hat{k} \in L_j, o_k \neq o_{\hat{k}}, s_j = 0, j \in J : k < \hat{k}, \tag{41}$$

$$x_k^{end} - x_{\hat{k}}^{begin} \geq \Delta_j^{meet} - M \cdot (1 - f_{k\hat{k},d}), k, \hat{k} \in L_j, o_k \neq o_{\hat{k}}, s_j = 0, j \in J : k < \hat{k}. \tag{42}$$

(v) Variable feasible ranges

The following constraints indicate the feasible ranges of the variables.

$$x_k^{begin}, x_k^{end} \in Z, k \in L_j, j \in J. \tag{43}$$

$$\gamma_{k\hat{k}}, \lambda_{k\hat{k}}, \theta_k \in \{0, 1\}, k, \hat{k} \in L_j, j \in J : k < \hat{k}, \tag{44}$$

$$q_{kn} \in \{0, 1\}, k \in L_j, n \in P_j, j \in J, \tag{45}$$

$$C_{k,u}, C_{k,d} \in \{0, 1\}, k \in L_j, j \in J, \tag{46}$$

$$h_{k\hat{k},u}, f_{k\hat{k},u}, h_{k\hat{k},d}, f_{k\hat{k},d} \in \{0, 1\}, k, \hat{k} \in L_j, j \in J : k < \hat{k}, \tag{47}$$

where integer decision variables x_k^{begin}, x_k^{end} indicate the beginning and ending time of each dependent event, respectively, binary variables $\gamma_{k\hat{k}}, \lambda_{k\hat{k}}$ state the logical relationship between two successive trains on the same section, binary variable θ_k indicates whether event k is a stop event at station, binary decision variable q_{kn} denotes whether event k selects n^{th} track to pass section j, for all $j \in J$, binary variables $C_{k,u}, C_{k,d}$ imply whether event k goes across the conflict point at the upstream and downstream of stations, respectively, and binary variables $h_{k\hat{k},u}, f_{k\hat{k},u}$ denote the logical relationship between two involved trains at the upstream of stations, and $h_{k\hat{k},d}, f_{k\hat{k},d}$ denote the logical relationship between two involved trains at the downstream of stations.

As trains are allowed switching track at stations and segments in above scheduling rule, we call it as complete switchable scheduling rules, i.e., Co-SRR. Herein, the train routing and timetabling problem for heterogeneous train traffic with Co-SRR can be formulated as the following model (M1), which is essentially an integer linear programming model (ILPM).

$$\begin{cases} \min z = \sum_{i \in I} (x_{l_i}^{end} - a_{l_i}^{f}) + \varepsilon \cdot \sum_{k \in K_i, i \in I} p_k, \\ s.t. \text{ constraints } (5), (7)-(47). \end{cases} \tag{48}$$

3.2. Formulation of other scheduling rules

To evaluate the proposed switchable scheduling rule model, we define two other different scheduling rules: incomplete switchable scheduling rule (In-SSR), and no switchable scheduling rule (No-SSR) (i.e., dedicated scheduling rule). These scheduling rules are described below.

(i) **In-SSR:** With this scheduling rule, trains are allowed to occupy the reverse direction tracks in stations but not at segments. More specifically, trains traveling in the inbound direction can only occupy inbound track 1 on segments and trains traveling in the outbound direction can only occupy outbound track 2 on segments. Nevertheless, trains in both directions can occupy any track in stations. In this case, we define the following constraints to depict track choice of trains at segments and stations, respectively.

$$\begin{cases} q_{k1} = 1, k \in L_j, s_j = 1, j \in J, o_k = 1, \\ q_{k2} = 1, k \in L_j, s_j = 1, j \in J, o_k = -1. \end{cases} \tag{49}$$

$$\sum_{n \in P_j} q_{kn} = 1, k \in L_j, s_j = 0, j \in J. \tag{50}$$

As trains can switch tracks at stations but not at segments, the move sequence based on the train track choice should meet the following constraints.

$$q_{kn} + q_{\hat{k}n} - 1 \le \lambda_{k\hat{k}} + \gamma_{k\hat{k}} \leqslant 1, k, \hat{k} \in L_j, n \in P_j, s_j = 0, j \in J : k < \hat{k}. \tag{51}$$

$$\lambda_{k\hat{k}} + \gamma_{k\hat{k}} = 1, k, \hat{k} \in L_j, o_k = o_{\hat{k}}, s_j = 1, j \in J : k < \hat{k}. \tag{52}$$

Meanwhile, the departure headway and arrival headway constraints (15)–(18) for trains traveling in the same direction should be adopted to guarantee safe operation at segments.

Since trains can switch tracks at stations, both head-on collisions for the trains traveling in different directions and rear-end collisions for trains traveling in the same direction at stations should be addressed. Further, since trains could not switch tracks at segments, the constraints (23)–(26) that are used to estimate whether trains cross the conflict points should be simplified for the In-SRR as follows.

$$q_{(k-1)1} + q_{km} + c_{j,l-s}^{1m} - 2 \le M \cdot C_{k,u}, k \in L_j, o_k = 1, m \in P_j, s_j = 0, j \in J \setminus \{1\}, \tag{53}$$

$$q_{km} + q_{(k+1)1} + c_{j,s-l}^{m1} - 2 \le M \cdot C_{k,d}, k \in L_j, o_k = 1, m \in P_j, s_j = 0, j \in J \setminus \{|J|\}, \tag{54}$$

$$q_{(k-1)2} + q_{km} + c_{j,l-s,1}^{2m} - 2 \le M \cdot C_{k,d}, k \in L_j, o_k = -1, m \in P_j, s_j = 0, j \in J \setminus \{|J|\}, \tag{55}$$

$$q_{km} + q_{(k+1)2} + c_{j,s-l,1}^{m2} - 2 \le M \cdot C_{k,u}, k \in L_j, o_k = -1, m \in P_j, s_j = 0, j \in J \setminus \{1\}. \tag{56}$$

Moreover, the operational safety constraints at stations should be satisfied, i.e., constraints (21) and (22), (27)–(42). Besides, constraints (5) are included to estimate whether an event of trains switches tracks at stations.

Therefore, the train routing and timetabling problem for heterogeneous train traffic with In-SSR can be formulated as the following model (M2), which is subjected to train operational restrictions (7)–(11), decision variables feasible ranges constraints (43)–(47), train track choice constraints (49), (50), constraints of train move sequence at sections (51), (52), operational safety constraints for trains running in the same direction at segments (15)–(18), as well as operational safety constraints at stations (21) and (22), (27)–(42), and (53)–(56). Thus M2 is formulated as

$$\begin{cases} \min z = \sum_{i \in I} (x_{l_i}^{end} - a_{l_i}^{f}) + \varepsilon \cdot \sum_{k \in K_i, i \in I} p_k, \\ s.t. \text{ constraints } (5), (7)-(11), (15)-(18), (21)-(22), (27)-(47), (49)-(56). \end{cases} \tag{57}$$

(ii) **No-SSR:** With this scheduling rule, trains can only be scheduled on the designated tracks at stations and segments. For example, in Fig. 1, trains traveling in the inbound direction can only occupy inbound track 1 in segments as well as the inbound tracks in stations throughout the whole trip, and trains traveling in the outbound direction can only occupy inbound track 2 in segments as well as inbound tracks at stations throughout the trip. To account for this restrictions, track choice constraints for trains at segments and stations should be modified as follows.

$$\begin{cases} \sum_{n \in P_j^1} q_{kn} = 1, k \in L_j, j \in J, o_k = 1, \\ \sum_{n \in P_j^2} q_{kn} = 1, k \in L_j, j \in J, o_k = -1. \end{cases} \tag{58}$$

Since trains can switch tracks at neither segments nor stations, we should only consider the rear-end collisions between two consecutive trains running in the same direction. Further, the potential collisions at stations can be avoided by considering the departure and arrival headway time at segments. Thus, we only should consider the safety operations constraints (15)–(18) for trains running in the same direction at segments as well as trains move sequence constraints at segments (52).

In brief, the train routing and timetabling problem for heterogeneous train traffic with No-SSR can be formulated as the following model (M3), which is subjected to train operational restrictions (7)–(11), decision variables feasible ranges (43)–(45), trains track choice constraints (58), trains move sequence constraints based on track choice (52), as well as safety operational constraints (15)–(18) for trains running in the same direction at segments.

$$\begin{cases} \min z = \sum_{i \in I} (x_{l_i}^{end} - a_{l_i}^f), \\ s.t. \text{ constraints } (7)-(11), (15)-(18), (43)-(45), (52), (58). \end{cases} \tag{59}$$

3.3. Solution methodology

Since the proposed objective function and constraints are all linear except for binary integer constraints, the presented problem is an integer linear programming problem that can be solved by existing commercial solvers. We used solver CPLEX 12.6 to solve the presented mathematical model with the OPL language. The numerical tests are performed on a PC with Windows 7 platform, Intel(R) Core(TM) i7-2130 with 2.7 GHz CPU and 8.00 GB memory.

To evaluate the performance of the proposed train routing and timetabling model, the three evaluation criteria mentioned in Dorfman and Medanic (2004) are adopted, namely (1) the time to clear the line, (2) the deviations of all trains, (3) the maximal deviations, and (4) time-efficiency ratio. The detail of these criteria is as follows.

(1) The time to clear the line is defined as $J_1 = a_{L,Des}^o - t_0$, where $a_{L,Des}^o$ is the arrival time for the last train at its destination in the obstacle situation (with train delay), and t_0 is the departure time of the first train leaving its origin station.
(2) The total deviations of all trains can be calculated as $J_2 = \sum_{i \in I}(a_{l_i}^o - a_{l_i}^f)$ where $a_{l_i}^f$ and $a_{l_i}^o$ are the arrival time for train i at its pre-determined destination in the ideal and feasible timetables (i.e., $a_{l_i}^o = x_{l_i}^{end}$), respectively.
(3) The maximal deviations of trains arriving at their destination stations is formulated as $J_3 = \max_{i \in I}(a_{l_i}^o - a_{l_i}^f)$.
(4) Time-efficiency ratio is formulated as $\eta = \frac{a_{L,Des}^f - t_0}{a_{L,Des}^o - t_0}$, where $a_{L,Des}^f$ is the arrival time for the last train at its destination in the free-travel situation. Note that $0 < \eta < 1$ and a higher value of η indicates a higher time-efficiency.

4. Numerical experiments

4.1. A small-scale case study

In this case, we consider a two-way double-track railway corridor with 7 stations and 6 segments to validate the proposed mathematical models. As shown in Fig. 9, the stations and segments are numbered consecutively with indices $S1, S2, \ldots, S13$ along the inbound direction. A total number of 18 trains, which are categorized into slow and fast types, are operated on this line. We assume that trains T1, T2, T3, T4, T5, T6, T7, T8, T9 are in the inbound direction (i.e., from S1 to S13), and the remain-

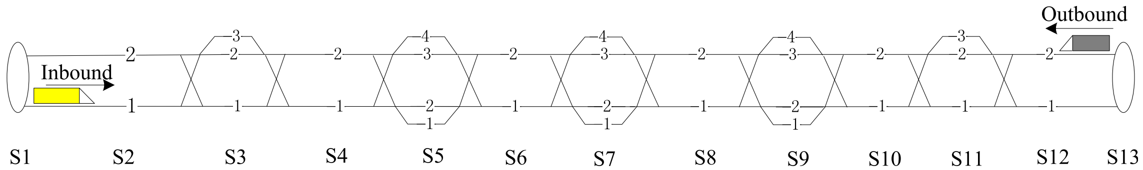

Fig. 9. A small-scale railway corridor.

ing trains are in the outbound direction (i.e., from S13 to S1). The ratio of the slow train to fast train are set as 2:1 and the first two departures are slow trains, thus T1, T2, T4, T5, T7, T8 are slow and T3, T6, T9 are fast trains in the outbound direction. We set the same setting in the outbound direction. Further, the expected departure interval time is set as 10 min, thus in the inbound (outbound) direction, T1 (T10) departs from origin station S1 (S13) at time 0, then T2 (T11) departs from origin station S1 (S13) at time 10 and so on. For simplicity, we assume the same value for the three types of safety headway time at any segment j for all $j \in J$, i.e., for all $j \in J$, i.e., $\Delta_j^{meet} = 2$ min, $\Delta_j^{dd} = \Delta_j^{aa} = 2$ min. Moreover, we set the acceleration time $\alpha = 2$, and deceleration time $\beta = 1$ in the following experiments. The minimum interstation travel time and stop plan for the fast and slow trains are presented in Table 3.

According to the definition of train discrete event, each train experiences 13 events in this experiment. All events associated with a train are labeled from inbound to outbound in the sequence according to the train's number. For instance, the event set for T1 in the inbound direction is $\{1, 2, \ldots, 13\}$, the one for T2 is $\{14, 15, \ldots, 26\}$, and so on.

Fig. 10 shows the timetables for trains in the small-case scheduled by Co-SSR, where (a) is the trains in the inbound direction, (b) is the trains in the outbound direction.

Table 4 shows the detail results of timetables scheduled by all aforementioned scheduling rules.

As shown in Table 4, the total deviations is 60 min and 54 min scheduled by the No-SSR and the In-SSR respectively, while it is 40 min with 6 track switch times by the Co-SSR. More specifically, the main deviations occur on the fast trains in the inbound direction of timetable scheduled by No-SSR, e.g., T3, T6, T9 suffer 6 min deviations during their trips, which might be caused by their front slow trains. On the other side, the total deviations of T3 and T9 reduce from 6 to 0 min, while both T2 and T8 suffer 3 more minutes in the timetable scheduled by the In-SSR. That is because trains are allowed switching track at stations in the In-SSR, thus T3 chooses the outbound track 2 in S3 to complete overtaking its front slow train T2, and resulting in T2 to dwell more 5 min at S3. Similarly, with In-SSR, T8 dwells at outbound track 2 of S3 to wait for being overtaken by fast train T9. These results show that In-SSR outperforms No-SSR. With the Co-SSR, T3 is properly scheduled on outbound

Table 3
The minimum interstation travel time and stop plan for the fast and slow trains.

Section index	Inbound track index	Outbound track index	Distance (km)	Fast train		Slow train	
				Minimum travel time (min)	Stop plan	Minimum travel time (min)	Stop plan
1	1,2	3,4	0	0	1	0	1
2	1	2	45	9	0	12	0
3	1	2,3	0	0	0	2	1
4	1	2	50	10	0	15	0
5	1,2	3,4	5	1	1	2	1
6	1	2	25	5	0	7	0
7	1,2	3,4	0	0	0	2	1
8	1	2	55	11	0	14	0
9	1,2	3,4	5	1	1	2	1
10	1	2	30	6	0	8	0
11	1	2,3	0	0	0	2	1
12	1	2	40	8	0	10	0
13	1,2	3,4	0	0	1	0	1

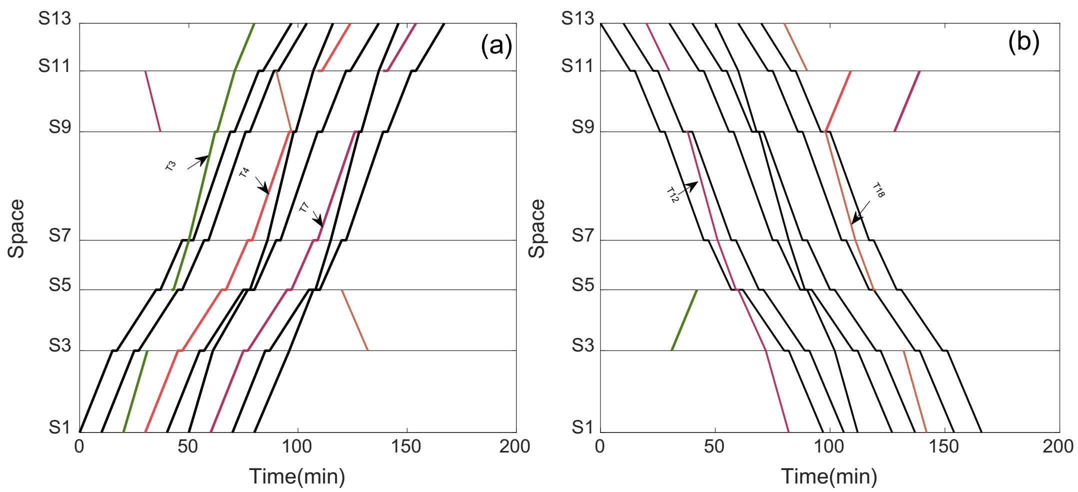

Fig. 10. Timetables for the trains in the small-case scheduled by Co-SSR: (a) inbound direction, (b) outbound direction.

Table 4
Train timetable performance with different scheduling rules.

Train index	No-SSR		In-SSR			Co-SSR		
	J_2 (min)	Deviation location	J_2 (min)	Deviation location	Switching location	J_2 (min)	Deviation location	Switching location
T1	3	S7	3	S7	–	3	S7	–
T2	3	S5	6	S3	–	–	–	–
T3	5 + 1	S4 + S8	–	–	$S_{3,2}$	–	–	$S_{4,2}$
T4	3	S9	3	S9	–	–	–	$S_{10,2}$
T5	3	S5	3	S5	–	3	S5	–
T6	5 + 1	S4 + S8	5 + 1	S4 + S8	–	5 + 1	S4 + S6	–
T7	3	S9	3	S7	–	–		$S_{10,2}$
T8	3	S5	6	S3	$S_{3,2}$	3	S5	–
T9	5 + 1	S4 + S8	–	–	–	5 + 1	S2 + S8	–
T10	3	S5	3	S5	–	3	S5	–
T11	3	S9	3	S9	–	2	S9	–
T12	1 + 1	S10 + S6	1 + 1	S10 + S6	–	2	S6	$S_{10,1}$
T13	3	S5	3	S5	–	3	S5	–
T14	3	S9	3	S9	–	3	S9	–
T15	1 + 1	S10 + S6	1 + 1	S10 + S6	–	1 + 1	S10 + S6	–
T16	3	S5	3	S5	–	–	–	–
T17	3	S9	3	S9	–	2	S9	–
T18	1 + 1	S10 + S6	1 + 1	S10 + S6	–	2	S6	$S_{10,1}$; $S_{4,1}$
Total deviations	60		54			40		
Track switch times					2			6

track 2 at S4, thus T2 does not need to wait at S3 for being overtaken that results in no deviations for T2 and T3. The deviations are distributed on the fast and slow trains with no preference because we assume uniform deviations for both fast and slow trains. However, it is easy to transfer the deviations from the fast trains to the slow ones by setting the proper deviation weights in the objective function.

Further, as shown in Table 4, most deviations for slow trains occur at stations while that for fast trains are at segments. That is because the running time for the fast trains at segments is shorter than that of slow trains, but they are held up by their front slow trains. Whereas the slow trains may be arranged to dwell more time at stations to wait for being overtaken by their following fast trains.

Fig. 11 gives an example to illustrate the conflict resolution at conflict point.

As shown in Fig. 11, T4 and T18, which are traveling in different directions, are intending to go across the conflict point at the downstream of S9, and they should be separated by the minimum time interval at the downstream of S9. Thus, T18 first goes across the downstream conflict point and arrives at S9 at time 97, then T4 leave S9 later and departs from S9 at time 98, which is 1 min later and satisfy the meet headway time separation. Meanwhile, T16 and T18 are traveling in the same direction, and T18 is claiming to go across the conflict point at upstream of S5 while T16 is entering the outbound track 2 at S4. In this case, they should meet the headway constraints. As expected, T16 leaves S5 1 min earlier than T18 (i.e., T16 leaves S5 at time 119 and T18 leaves at time 120). Furthermore, as an example, Fig. 11 shows the optimal route for T18 from S13 to S1 that is {$S_{13,3}$, $S_{12,2}$, $S_{11,3}$, $\mathbf{S_{10,1}}$, $S_{9,3}$ $S_{8,2}$, $S_{7,3}$ $S_{6,2}$, $S_{5,4}$, $\mathbf{S_{4,1}}$, $S_{3,3}$, $S_{2,2}$ $S_{1,3}$}.

4.2. Large-scale case experiments

To test the effectiveness and efficiency of our proposed routing and timetabling approach, this section applies the proposed model to Beijing–Shanghai high-speed railway (HSR) corridor in China, which consists of 23 stations and 22 segments (i.e., 45 railway sections in total) with a total length 1032 km, as shown in Fig. 12. Assume that there are 4 tracks within each station and 2 tracks on each segment. The types of all stations are assumed to be the same as station D in Fig. 7. Further, the direction from Shanghai toward Beijing is defined as the inbound direction. Hence, Shanghai Hongqiao is set as station 1, the segment between Shanghai Hongqiao and Kunshan S is set as segment 2, Kunshan S station is set as station 3, and so on all

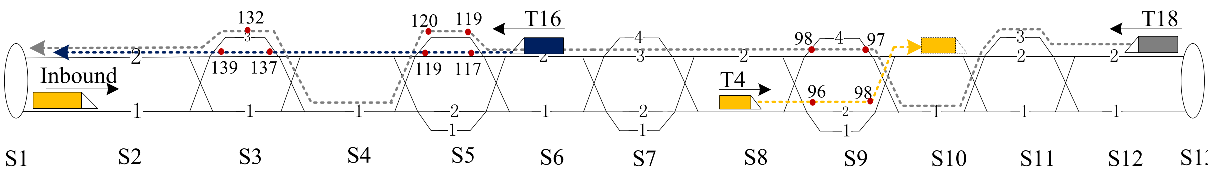

Fig. 11. Illustration of the time and route for two trains in different direction scheduled by Co-SRR.

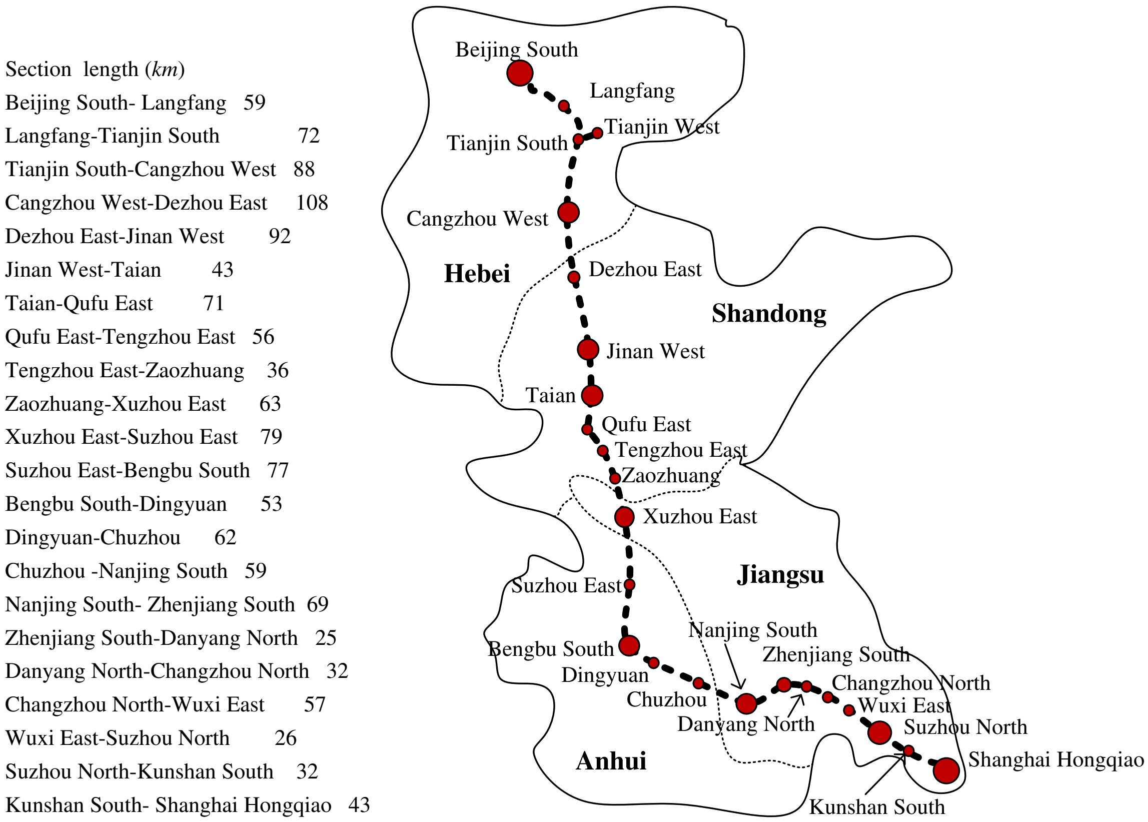

Fig. 12. Map of Beijing–Shanghai high-speed railway corridor.

the way to the Beijing South station. Further, more information about the problem parameters including the minimum travel time on each segment and the dwell time requirement on each station for the fast and slow trains are listed in Table 5, in which the maximum travel speed of fast and slow trains are set as 300 km/h and 250 km/h, respectively.

Note that it is practically difficult to collect all realistic information related to Beijing–Shanghai railway corridor. Hence, to test the application of the proposed routing and timetabling with SSR, inaccessible parameters (e.g., stop plan) are reasonably assumed.

4.2.1. Case studies with three scheduling rules

In this example, we investigate the validity and effectiveness of the proposed SSRs approach. As benchmarks, we apply the aforementioned three scheduling rules to schedule heterogeneous trains on Beijing–Shanghai HSR corridor. In this example, 20 inbound and 20 outbound trains are considered in total. In each experiment, we assume that the ratio of fast to slow trains is 2:1, and the departure time interval is assumed to be 30 min. Further, the inbound and outbound trains are numbered from T1 to T20, and T21 to T40, respectively. Since the timetables scheduled by No-SSR are similar to that scheduled by In-SSR, we only present the train timetables scheduled by In-SSR and Co-SSR. These timetables are shown in Figs. 13 and 14, respectively.

Table 5
The basic information for fast and slow trains on Beijing-Shanghai railway corridor (units: min).

Section	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$t^{fast}_{\min}$	0	9	0	6	0	5	2	11	0	6	0	5	0	14	2
$t^{slow}_{\min}$	0	10	2	8	0	6	0	14	1	8	3	6	2	17	0
Section	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
$t^{fast}_{\min}$	12	0	12	0	11	2	15	0	16	3	13	3	7	0	11
$t^{slow}_{\min}$	14	4	15	4	13	0	18	5	19	0	15	0	9	4	13
Section	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
$t^{fast}_{\min}$	0	14	0	9	0	18	3	22	0	18	0	14	0	12	0
$t^{slow}_{\min}$	0	17	5	10	3	22	0	26	3	21	0	17	3	14	0

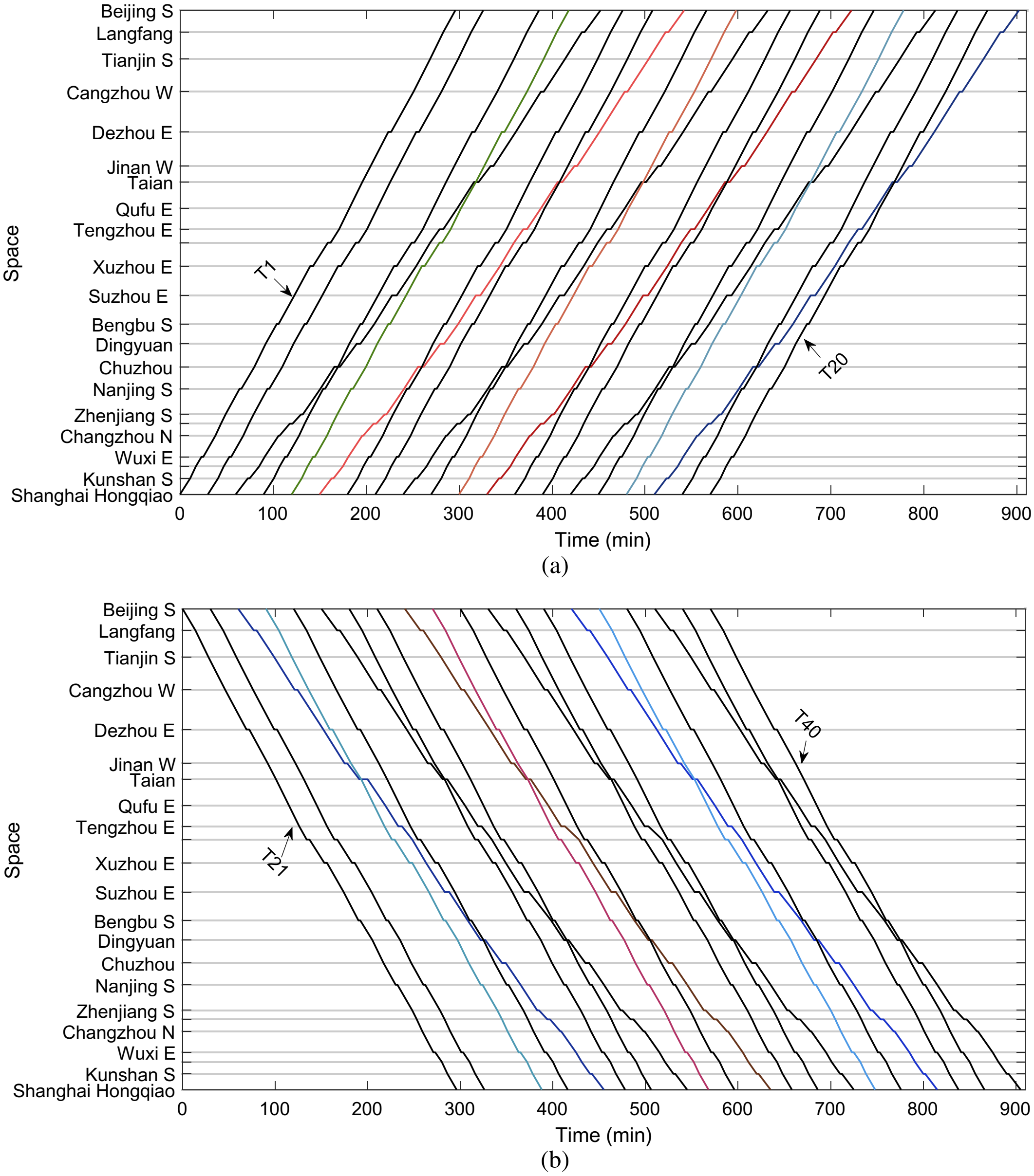

Fig. 13. Train timetables scheduled by In-SSR: (a) inbound direction, (b) outbound direction.

Fig. 15 illustrates an example of overtaking scheduled by In-SSR and Co-SSR, where T4 can catch up with T3 on segment 18 (Chuzhou to Dingyuan) in the inbound direction. In In-SSR, T3 is arranged to wait at station Chuzhou to be overtaken by fast train T4 that results in 2 min delay for T3, i.e., T3 dwells at Chuzhou S for 6 min versus the required dwell time was 4 min. Since trains can move on the vacant tracks in reverse directions in Co-SSR, T3 is taken to the vacant track 4 in its reverse direction at station Chuzhou and then enters track 2 at segment 18, while its following T4 is traveling on its designated track to traverse segment 18 without any delays, i.e., $q_{108,2} = 1, q_{153,1} = 1$. As the two successive trains choose different tracks to traverse segment 18 and neither of them goes across the conflict point, it is not necessary to consider the safety departure headway time and meet headway time, thus T3 and T4 could enter segment 18 simultaneously, i.e., $x_{108}^{begin} = x_{153}^{begin} = 170$. On the other hand, due to the fact that trains cannot use their reverse direction track in In-SSR, the inbound trains T3 and T4 could only utilize the same track to traverse segment 18, i.e., $q_{108,1} = 1$, $q_{153,1} = 1$. Therefore, they should satisfy the departure safety headway constraint to guarantee the operational safety. As a result, T3 and T4 enter segment 18 at time 170 and 172, respectively. Similarly, as shown in Fig. 14, the inbound trains T7, T10, T12, T16, T19 are properly arranged on vacant outbound track 2 at segment 18 (Chuzhou to Dingyuan), while the outbound trains T23, T28, T29, T32, T37, T38 are allocated to vacant inbound track 1 at segment 34 (Jinan West to Tainan). Moreover, trains T23, T27, T30,

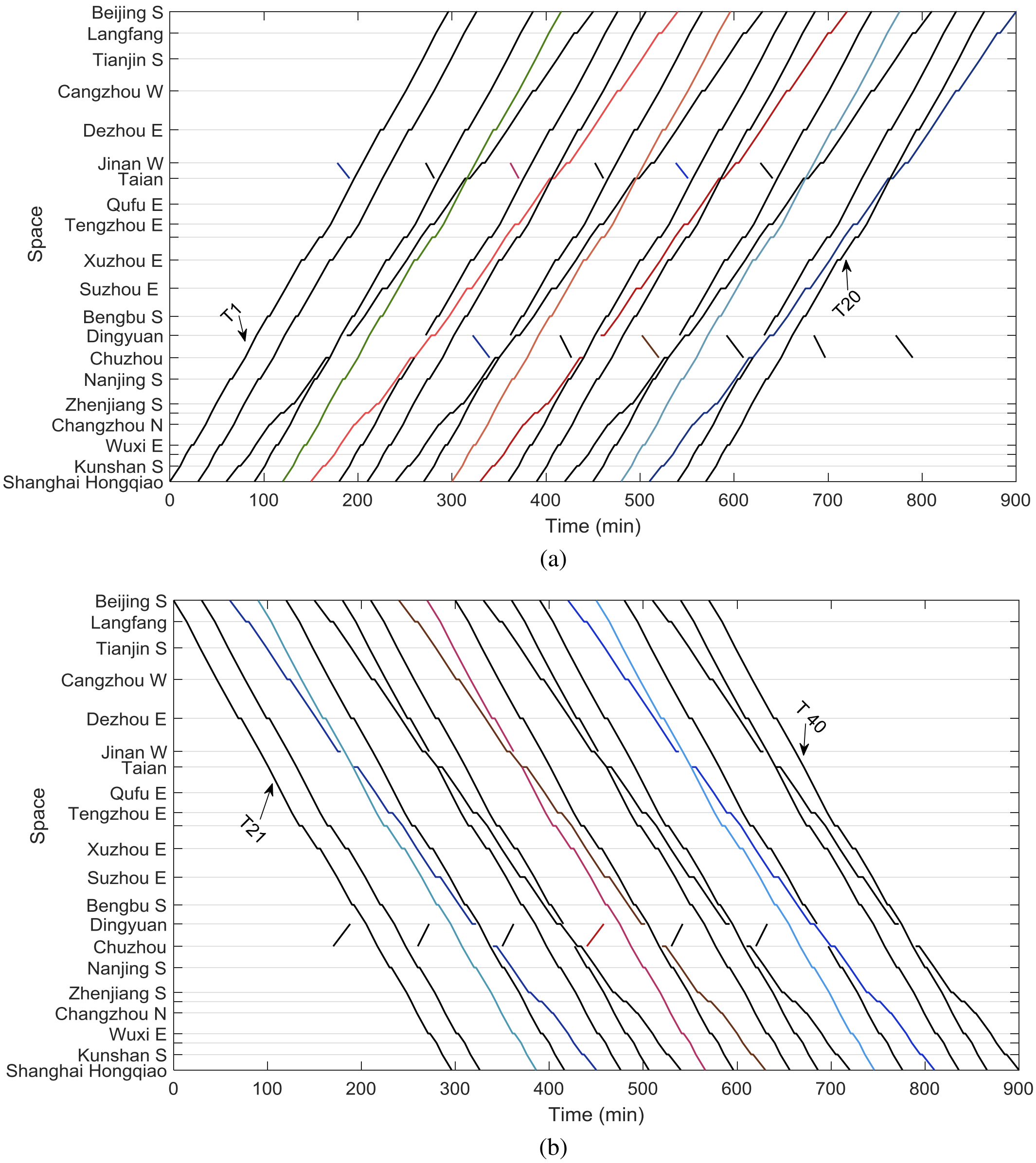

Fig. 14. Train timetables scheduled by Co-SSR: (a) inbound direction, (b) outbound direction.

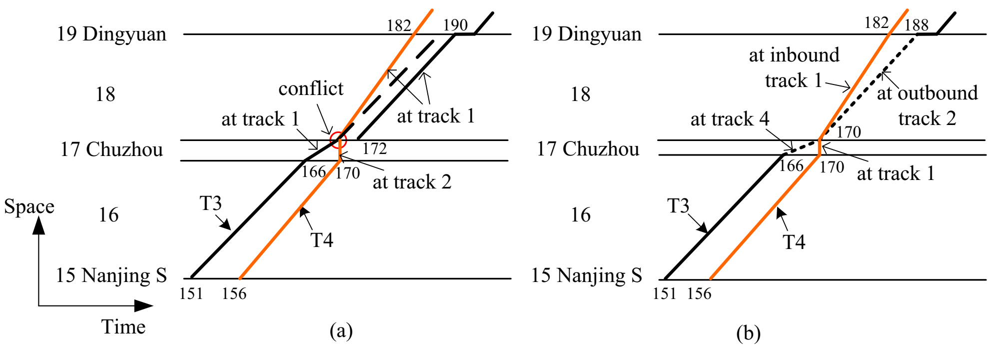

Fig. 15. An illustration of an overtaking: (a) scheduled by In-SSR, (b) scheduled by Co-SSR.

T33, T35, T38 are scheduled on vacant inbound track 1 at the segment of 18 that forms similar opposite-directional trajectories.

Table 6 shows the comparison results for all three scheduling rules. The presented criteria are time to clear line J_1, total deviations for all trains J_2, maximum deviations J_3, time-efficiency ratio η, and the deviant/switched trains.

As shown in this table, Co-SSR outperforms No-SSR and In-SSR. In fact, the time to clear line (i.e., J_1) is largely determined by the velocities of the last several trains, thus its values are almost the same among all three scheduling rules. With Co-SSR, the total deviations (i.e., J_2) is reduced from 66 to 0, with 18 track switch times. Further, in No-SSR and In-SSR, the total number of deviated trains are 24, most of which are slow trains, while there is no deviated train in Co-SSR results. As discussed in Section 2, with Co-SSR, if fast trains can catch up with the preceding slow trains, they are allowed to travel on the reverse tracks. The results show that a total of 6 inbound and 12 outbound trains are scheduled on the reverse tracks. As a result, the utilization rate of railway tracks are improved and train deviations from the ideal timetable are reduced. Co-SSR could also reduce the maximum deviations (i.e., J_3) from 5 to 0. Note that the corresponding values for No-SSR and In-SSR are identical, which implies that the capacity of stations with 4 tracks are sufficient, thus concerning this criterion, In-SSR is not better than No-SSR. Based on the presented quantitative results, we conclude that Co-SSR outperforms the other two common approaches for the train routing and timetabling problem.

4.2.2. Timetables' performance evaluations concerning different parameters settings

4.2.2.1. Influence of the number of trains. In this section, we use the three aforementioned train scheduling rules to test the influence of train numbers on the performance of the proposed model. In this example, the number of total trains varies between 6 and 60 and trains are scheduled with the departure interval time of 30 min. Moreover, the ratio of fast to slow trains is set to be 5:1, and the first departure train is assumed to be a fast one. The results of J_2 (which is the most important considered criterion) and the solution times are shown in Fig. 16. In this figure, the solution time 1, 2, 3 represent the solution time for No-SSR, In-SSR, and Co-SSR, respectively.

Table 6
Comparison results for different scheduling rules.

Scheduling rule	J_1 (min)	J_2 (min)	J_3 (min)	Track switch times	η	Deviated trains (time: min)/Switched trains
No-SSR	902	66	5	0	0.9978	T3(2), T5(2), T6(2), T8(2), T9(2), T11(2), T12(2), T14(2), T15 (2), T17(2), T18(2), T20(2), T23(5), T24 (2), T26(5), T27(2), T29 (5), T30(2), T32(5), T33(2), T35(5), T36(2), T38(5), T39(2)
In-SSR	902	66	5	0	0.9978	T3(2), T5(2), T6(2), T8(2), T9(2), T11(2), T12(2), T14(2), T15 (2), T17(2), T18(2), T20(2), T23(5), T24 (2), T26(5), T27(2), T29 (5), T30(2), T32(5), T33(2), T35(5), T36(2), T38(5), T39(2)
Co-SSR	900	0	0	24	1	At stations: T3, T10, T12 switch to $S_{17,4}$; T7, T16, T19 switch to $S_{17,3}$; At segments: T3, T10, T12, T7, T16, T19 switch to $S_{18,2}$; T23, T28, T29, T32, T37, T38 switch to $S_{18,1}$; T23, T27, T30, T33, T35, T38 switch to $S_{34,1}$.

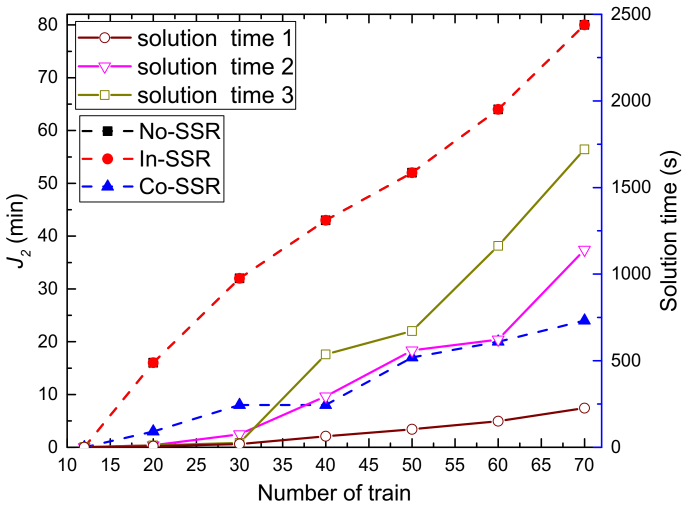

Fig. 16. The influence of different numbers of trains on J_2 and the solution times.

As shown in Fig. 16, J_2 and the solution times increases with the number of trains, which is as we expected because as the number of train increases, there may be more conflicts between different trains. It is also worth mentioning that J_2 results in Co-SSR is much better than that of the other two strategies. For instance, when the number of the train is 40, the values of J_2 in No-SSR and In-SSR are both 43 min and that is 8 min in Co-SSR, with 9 track switch times at segments and 1 track switch times at stations. However, the solution times are 63.61, 294.38 and 536.06 s, respectively. Since trains are allowed to switch tracks at stations, it may take more time to find an optimal solution in In-SSR compared with No-SSR, but the final results are identical to No-SSR. That implies that although the capacity of stations with 4 tracks are sufficient, the throughput capacity of segments is not enough. In other words, limited throughput capacity of sections is the main factor affecting the trains' operation efficiencies in Jing-Hu HSR corridor. Moreover, the maximum deviations in No-SSR and In-SSR are always 5 min when the number of trains increases from 10 to 70, but it is only 2 or 3 min in Co-SSR, which further confirms the efficiency of Co-SSR. Generally, the solution time in Co-SSR is longer than that in No-SSR. However, since the studied problem is meant for planning with a long time horizon, the Co-SSR solution times are still reasonable.

4.2.2.2. Influence of departure interval. In this experiment, we investigate the influence of departure time interval H on timetables' performance measures considering different scheduling rules. In this example, the departure time interval H varies from 20 to 60 min with 10 min increments. The ratio of fast to slow trains are set as 2:1 and the first departing train at the origin is a fast one. The results of J_2 are presented in Table 7. In this table, symbol '-' denotes that the solution time exceeds the time limit of 2 h without any solution.

As shown in Table 7, in No-SSR and In-SSR, total deviations mainly increase as the departure time intervals decrease. That may be because there is more mutual interference between trains on higher train densities. Moreover, with the same departure time interval, more number of trains on railway may cause more deviations because mutual interferences between trains are more intensive. Besides, there are no switching events in In-SSR since 4 tracks in stations already meet the operation of the trains on this corridor and the limitation is the throughput capacity of segments as mentioned in Section 4.2.2.1, which resulting in no difference between No-SSR and In-SSR in these cases. On the contrary, the train timetables scheduled by Co-SSR generally outperform those from the other two strategies. Especially when $H = 30$ and $H = 60$ min, the total Co-SSR deviations is 0 in all experiments, which means that all fast trains are scheduled on the reverse tracks to avoid deviation. However, for $H = 30$ the track switch times are a little more, which may be caused by more intensive mutual interference between trains on the HSR corridor. In practice, considering the minimum total travel time on this HSR corridor (i.e., nearly 5 h), the service time window of high-speed trains at each origin station is roughly between 6:00 am and 6:00 pm. Thus the train operators only can schedule 24 trains if the departure time interval is 60 min, which apparently could not meet the passenger demand. On the other hand, when $H = 30$ min, almost 48 trains can be scheduled on the HSR corridor. Therefore, to decide an appropriate departure time interval, the operators should consider timetable efficiency, transportation demands, and more other factors. Further, as it is shown in Table 7, when the departure intervals are relatively small (i.e., 20 min) and there is a high density of trains on the railway corridor, the model cannot find an optimal or a near-optimum solution. This implies the need for an efficient algorithm for the proposed models.

Table 7
Comparisons for different departure interval.

Departure intervals	Number of trains	No-SRR	In-SRR	Co-SRR		
		J_2 (min)	J_2 (min)	J_2 (min)	Track switch times at segments	Track switch times at stations
$H = 20$	10	12	12	8	2	0
	20	49	49	20	10	0
	40	125	125	–	–	–
	60	194	–	–	–	–
$H = 30$	10	11	11	0	3	1
	20	26	26	0	8	3
	40	66	66	0	18	6
	60	99	99	0	27	9
$H = 40$	10	14	14	4	4	0
	20	37	37	13	9	1
	40	84	84	38	14	2
	60	126	126	60	20	1
$H = 50$	10	12	12	4	3	0
	20	36	36	9	10	0
	40	72	72	24	18	0
	60	108	108	36	27	0
$H = 60$	10	5	5	0	1	0
	20	15	15	0	3	0
	40	30	30	0	6	0
	60	45	45	0	8	0

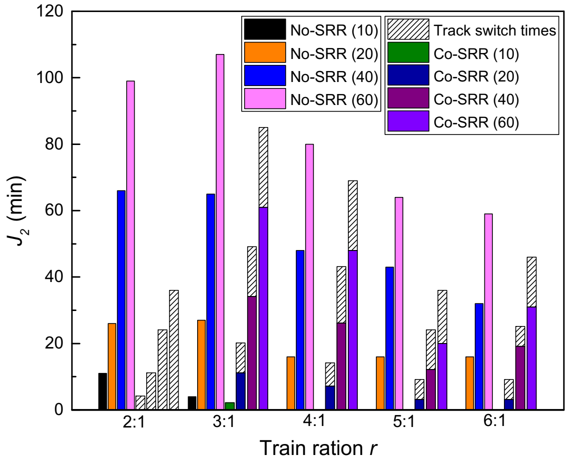

Fig. 17. The values of criteria J_2 with different r values (H = 30 min).

4.2.2.3. Influence of ratios of train types. This experiment focuses on investigating the performance of timetables with different ratios (denoted by r) of fast to slow trains. As the analyses in s 4.2.2.1 and 4.2.2.2 indicate, the train timetables scheduled by In-SSR and No-SSR are almost identical. Therefore, this section only compares No-SSR and Co-SSR results, where r varies from 2:1 to 6:1. In this experiment, the number of trains increases from 10 to 60, and departure interval H is set to 30 min. The results of J_2 are shown in Fig. 17.

Fig. 17 shows that J_2 increases with the number of involved trains. However, the total deviations are much less with Co-SSR than that with No-SSR in a certain situation, which again demonstrates the effectiveness of switchable scheduling rules.

Greater train ratios of fast to slow trains result in less total train deviations, excluding when the ratio is 2:1. For instance, when the total number of trains on the HSR corridor is 60, and if $r = 3:1$, we obtain $|I_f| = 45, |I_s| = 15$, the total deviations $J_2^{3:1} = 107$ with No-SRR, and that is 61 with 34 track switch times with Co-SRR. Moreover, if $r = 4:1$, we obtain $|I_f| = 48, |I_s| = 12, J_2^{4:1} = 80$ in No-SRR, and that is 48 with 21 track switch times in Co-SRR. Further, if $r = 5:1$, it yields $|I_f| = 50, |I_s| = 10$, and the last departure trains are two slow trains, $J_2^{5:1} = 64$ in No-SRR, and that is 20 with 16 track switch times in Co-SRR. Whereas if $r = 6:1$, we obtain $|I_f| = 50, |I_s| = 10, J_2^{6:1} = 59$ in No-SRR, and that is 31 with 15 track switch times in Co-SRR which is greater than that in $r = 5:1$. That is because the last departure trains are 4 fast trains that may overtake the previous slow trains and result in deviations on these slow trains. Particularly, the values of $J_2^{2:1}$ are all 0, no matter how many trains are involved. Combined with the analysis in Section 4.2.1, we know that a slow train is overtaken no more than twice before arriving at its destination when the departure time interval is 30 min and $r = 2:1$. Moreover, one of the two close trains has been scheduled on their reverse vacant tracks when a conflict occurs, eventually forming similar groups of trajectory. These experiments show that if the departure time interval is 30 min, the best operational ratio is 2:1 followed by 5:1 with Co-SSR.

From the above analysis, the ratio of fast to slow trains and departure time interval have significant impacts on the efficiency of train operations. Therefore, in the real train operation, it needs to do multi-party verification to obtain the optimal operational scheme.

5. Conclusion

This paper proposes a new mathematical formulation for the train routing and timetabling problem with a switchable scheduling rule to train operational efficiency on the double-track railway corridor. Three integer linear programming models are built to investigate different scheduling rules, including (1) No-SSR that requires trains to utilize only their designated tracks, (2) In-SSR where trains could utilize their reverse vacant tracks only at stations, and (3) Co-SSR where trains can be scheduled on their reverse vacant tracks both on segments and at stations if necessary.

A number of numerical experiments are performed to assess these three scheduling rules. We employ the CPLEX solver to find the optimal scheduling solutions for trains. The results indicate that Co-SSR outperforms the other two scheduling rules. It is also found that In-SSR performs better than No-SSR when the capacity of stations are insufficient. The results also imply that the main factor affecting the efficiency of train operations on the double-track railway corridor is the capacity of sections rather than that of stations. Moreover, with No-SSR and In-SSR, the total train deviations increase as the departure time intervals decrease because higher train densities on HSR corridor increase the interference between trains. Besides, greater ratios of fast to slow trains result in less total train deviations from the ideal timetable.

The proposed mathematical formulation obtains arrival and departure times, and the resource allocations for all trains. Although the presented problem can achieve the optimal scheduling solutions for trains, the problem solution times may not be very efficient for real-time application. Therefore, one possible future direction is to develop an efficient algorithm for larger real-time operation problems. Besides, it is interesting to investigate the difference time of trains along different directional tracks among a segment. For instance, if one track in the opposite direction is in maintenance task, the train should slow down (i.e., extend the minimum travel time for the train on this segment) to guarantee the safety. Further, flexible departure time (i.e., allow trains departing earlier than their pre-determined departure time at origin station) or skipping some stop along their trips could find better scheduling solutions, so that is also an interesting research direction.

Acknowledgment

The research is supported in part by the National Natural Science Foundation of China (Nos. 71631002, 71621001, 71471012, 71390332), the Research Foundation of State Key Laboratory of Rail Traffic Control and Safety (No. RCS2016ZT003), and the U.S. National Science Foundation through Grants CMMI#1453949.

References

Bešinović, N., Goverde, R.M.P., Quaglietta, E., Roberti, R., 2016. An integrated micro–macro approach to robust railway timetabling. Transp. Res. Part B: Methodol. 87, 14–32.
Cacchiani, V., Furini, F., Kidd, M.P., 2016. Approaches to a real-world train timetabling problem in a railway node. Omega 58, 97–110.
Cacchiani, V., Toth, P., 2012. Nominal and robust train timetabling problems. Eur. J. Oper. Res. 219, 727–737.
Caimi, G., Fuchsberger, M., Laumanns, M., Schüpbach, K., 2011. A multi-level framework for generating train schedules in highly utilised networks. Public Transp. 3 (1), 3–24.
Caimi, G., Kroon, L., Liebchen, C., 2017. Models for railway timetable optimization: applicability and applications in practice. J. Rail Transp. Plann. Manage. 6, 285–312.
Dorfman, M.J., Medanic, J., 2004. Scheduling trains on a railway network using a discrete event model of railway traffic. Transp. Res. Part B: Methodol. 38, 81–98.
Fang, W., Yang, S., Yao, X., 2015. A survey on problem models and solution approaches to rescheduling in railway networks. IEEE Trans. Intell. Transp. Syst. 16, 2997–3016.
Hansen, I.A., Pachl, J. (Eds.), 2014. Railway Timetabling & Operations: Analysis, Modelling, Optimisation, Simulation, Performance Evaluation. DVV Media Group GmBH.
He, L., Liang, Q., Fang, S., 2016. Challenges and innovative solutions in urban rail transit network operations and management: China's Guangzhou metro experience. Urban Rail Transit 2 (1), 33–45.
Lamorgese, L., Mannino, C., Natvig, E., 2016. An exact micro–macro approach to cyclic and non-cyclic train timetabling. Omega.
Lee, Y., Chen, C.Y., 2009. A heuristic for the train pathing and timetabling problem. Transp. Res. Part B: Methodol. 43, 837–851.
Li, F., Gao, Z.Y., Li, K.P., Yang, L.X., 2008. Efficient scheduling of railway traffic based on global information of train. Transp. Res. Part B: Methodol. 42, 1008–1030.
Meng, L.Y., Zhou, X.S., 2014. Simultaneous train rerouting and rescheduling on an N-track network: a model reformulation with network-based cumulative flow variables. Transp. Res. Part B: Methodol. 67, 208–234.
Mu, S., Dessouky, M., 2011. Scheduling freight trains traveling on complex networks. Transp. Res. Part B: Methodol. 45, 1103–1123.
Mu, S., Dessouky, M., 2013. Efficient dispatching rules on double tracks with heterogeneous train traffic. Transp. Res. Part B: Methodol. 51, 45–64.
Mu, S., Dessouky, M.M., 2014. Control rules for dispatching trains on general networks with multiple train speeds. In: Conference Proceedings Transport Research Arena.
Pachl, J., 2014. Railway Operation and Control. VTD Rail Publishing, Mountlake Terrace (USA).
Qi, J., Yang, L., Gao, Y., Li, S., Gao, Z., 2016. Integrated multi-track station layout design and train scheduling models on railway corridors. Transp. Res. Part C: Emerg. Technol. 69, 91–119.
Samà, M., D'Ariano, A., Corman, F., Pacciarelli, D., 2017. A variable neighbourhood search for fast train scheduling and routing during disturbed railway traffic situations. Comput. Oper. Res. 78, 480–499.
Törnquist, J., Persson, J.A., 2007. N-tracked railway traffic re-scheduling during disturbances. Transp. Res. Part B: Methodol. 41, 342–362.
Törnquist Krasemann, J., 2015. Computational decision-support for railway traffic management and associated configuration challenges: an experimental study. J. Rail Transp. Plann. Manage. 5, 95–109.
Xu, X.M., Li, K.P., Yang, L.X., 2015. Scheduling heterogeneous train traffic on double tracks with efficient dispatching rules. Transp. Res. Part B: Methodol. 78, 364–384.
Xu, X.M., Li, K.P., Yang, L.X., Ye, J.J., 2014. Balanced train timetabling on a single-line railway with optimized velocity. Appl. Math. Model. 38, 894–909.
Xu, Y., Jia, B., Li, M.H., Li, X.G., 2016. An improved discrete-time model for heterogeneous high-speed train traffic flow. Commun. Theor. Phys. 65 (3), 381.
Zhou, W., Teng, H., 2016. Simultaneous passenger train routing and timetabling using an efficient train-based Lagrangian relaxation decomposition. Transp. Res. Part B: Methodol. 94, 409–439.
Zhou, X., Zhong, M., 2005. Bicriteria train scheduling for high-speed passenger railroad planning applications. Eur. J. Oper. Res. 167, 752–771.
Zhou, X., Zhong, M., 2007. Single-track train timetabling with guaranteed optimality: branch-and-bound algorithms with enhanced lower bounds. Transp. Res. Part B: Methodol. 41, 320–341.