Article:

Li, F, Gao, Z, Wang, D et al. (2017) A subjective capacity evaluation model for single-track railway system with δ-balanced traffic and λ-tolerance level. Transportation Research Part B: Methodological, 105. pp. 43-66. ISSN: 0191-2615

https://doi.org/10.1016/j.trb.2017.08.020

Please cite the paper as:

Li F, Gao Z, Wang D, Liu, R., Tang, T., Wu, J. and Yang, L. (2017) A subjective capacity evaluation model for single-track railway system with δ-balanced traffic and λ-tolerance level. Transportation Research Part B: Methodological 105: pp. 43-66. https://publications.leeds.ac.uk/viewobject.html?cid=1&id=437179

A subjective capacity evaluation model for single-track railway system with δ-balanced traffic and λ-tolerance level

Feng Li
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China, fengli0925@gmail.com

Ziyou Gao
Institute of Transportation System Science and Engineering, Beijing Jiaotong University, Beijing 100044, China, zygao@bjtu.edu.cn

David Z.W. Wang
School of Civil and Environmental Engineering, Nanyang Technological University, Singapore 639798, Singapore, WangZhiwei@ntu.edu.sg

Ronghui Liu
Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, United Kingdom, R.Liu@its.leeds.ac.uk

Tao Tang, Jianjun Wu, Lixing Yang
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China, tangtao@bjtu.edu.cn; jjwu1@bjtu.edu.cn; lxyang@bjtu.edu.cn

Abstract:
In this paper, we propose a method to measure the capacity of single-track railway corridors subject to a given degree of balance between the two directional traffic loads and a permitted overall delay level. We introduce the concepts of δ-balance degree and λ-tolerance level to reflect the subjective measures of the railway administrator for capacity evaluation. A train balance scheduling problem with initial departure time choice of trains is embedded into the measure of railway capacity. The combined scheduling and capacity evaluation method is formulated as a 0-1 mixed integer programming model, and solved using a simple dichotomization-based heuristic method. A highly efficient heuristic procedure based on the concept of compaction pattern is developed to solve the train balance scheduling problem, and the numerical results demonstrate that the method yields high-quality solutions close to the optimal ones using the CPLEX solver. The two-way traffic loading capacity of a single-track railway corridor is analyzed in detail under different tolerance levels and balance degrees. The transition regions of traffic loading capacity are identified, and provide a useful decision support tool for the railway administrators in dealing with train rescheduling requests under disturbance or disruption scenarios.

Key words: railway capacity; tolerance level; balance condition; compaction pattern; departure time choice.

1. Introduction

The capacity of a railway system is a key measure and is of significant importance to the railway industry. Whether it is to add more trains in an existing system (Burdett and Kozan, 2009) or to build new rail infrastructure (Burdett, 2016), it is crucial to know where the spare capacity lies or where the new capacity needs are. Krueger

(1999) defined the railway capacity as "a measure of the ability to move a specific amount of traffic over a defined rail line with a given set of resources under a specific service plan". A more generally adopted definition is the maximum number of trains that can traverse the entire railway line in a given period of time (Burdett and Kozan, 2006; Mussone and Calvo, 2013). Whilst these definitions seem to be self-explanatory, their quantification is not straight forward because it depends not only on the assortment of railway layouts, but also the proportions of different train types as well as the dispatching rules of trains in the railway system.

Most of the existing studies focus on the capacity of double-tracks or multi-tracks railway system (Prinz, 2005; Alex Landex et al, 2006; Wahlborg, 2005; Melody and Preston, 2010; Lindner, 2011). However, single-track railroads still have important transportation roles to play in many countries. For example, single-track railroad in USA accounts for approximately 80% of the entire railway network (CS-I, 2007; Tolliver, 2010). Freight transport is usually undertaken along single-track railway corridors in some countries of Northern Europe, such as Sweden, Denmark and Norway (Landex, 2008). The famous Qing-Zang railway corridor, which links 89 stations and traverses the whole of the southwest of China at a length of 1956km, is single-track all the way.

The distinct characteristic of the single-track railroad is that it carries two-way traffic, i.e., the segment between stations can be occupied by trains travelling in both directions. The meeting-crossing and overtaking among trains make single-track railroad more complicated to plan and manage than other railway system. As a consequence, the transport capacity of a single-track railroad is rarely able to achieve what is expected by the railway administrators. Part of the reason for that is the complication associated with assessing the actual capacity of the single-track system, and more specifically the lack of a clear definition that reflects explicitly the two-way traffic characteristic of single-track railway. Compare with double- and multi-track railway system, two-way traffic in the single-track railway system results in more conflicts between train flows in different directions. It is insufficient to only focus on the line or station capacity. Additionally, an accurate capacity evaluation is closely related to how the trains are scheduled to run in the railway system, which is often unknown at the stage of exploring the capacity.

There has been a rich literature on railway capacity (Frank, 1966; Petersen, 1974; Assad, 1980; Yokota, 1980; Petersen and Taylor, 1982; Welch and Gussow, 1986; De Kort et al, 2003; Kozan and Burdett, 2005; Lai and Barkan, 2009; Bevrani et al, 2015; Burdett, 2015a, 2015b, 2016). Most of them however are focused on capacity of segments or stations, and these capacity analyses emphasize the influence of railway infrastructure layout only. Due to two-way traffic characteristics and the strong dependence between segments and stations in the single-track railroad, it is essential to consider the single-track railway as a whole system. In addition to the needs to consider train types and schedule plans, it would also be interesting to evaluate capacity from the viewpoint of railway administrators, to take into account the constraints or flexibility they wish to put into the system.

In this paper, we analyze single-track railway system capacity from the viewpoint of railway administrators: giving a set of objectives the administrators wish to achieve, what the railway capacity would be. More specifically, we set out to explore: ***if the average delay of trains is confined to a given range, what is the maximal number of trains that can be loaded onto the single-track railway system***? Clearly, with increasing train numbers, more delays would be expected in order to accommodate the increased number of meet-crossings. Being able to accurately quantify the railway capacity under different delay tolerance levels provides decision support for the administrators to balance the trade-off between the demand and the service levels. In addition to delay considerations, the administrators usually aim to keep the balance between train flows in both directions. The relative balance of in- and out-bound train flows has a significant impact on the delays of trains and capacity of the single-tracks railway. So a new question can be proposed as: ***if the average delay of trains is confined to a certain range and a relative balance between the in- and out-bound train flows is maintained, what is the maximal number of trains that can be loaded onto the single-track railway system***?

To the best of our knowledge, the delay tolerance level and relative balance have not been jointly considered previously in the analysis of railway capacity of single-track system. In this paper, we set out to derive a two-way balanced traffic loading capacity for the single-track railway system subject to a given delay tolerance level. We present an analytical formulation of the model and develop a highly efficient algorithm to derive the solutions. The outcomes of our results provide a useful decision support tool for the administrators.

The major contributions of this paper are listed as follows. Firstly, the concept of a two-way balanced traffic loading capacity is explicitly expressed, in which a λ-tolerance level is introduced to describe the control of the administrators on train delays, and a δ-balance degree is defined to reflect the expectation of the administrators for the relative balance of in- and out-bound train flows. Secondly, a 0-1 mixed integer programming model is formulated to quantify this. The objective of the model peruses the maximal allowed number of ***train-pairs*** based on λ-tolerance level of administrators in the single-track railway corridor. The deviation between the average travel times of in- and out-bound train flow is subject to δ-balance condition. An important characteristic of the model is that the departure times of trains from their original stations can vary within a given hard time-window. Our third contribution is a simple dichotomization-based method proposed to solve the above model. But a key issue is how to solve efficiently train δ-balance scheduling problem with initial departure choice. A heuristic procedure based on ***compaction pattern*** of time-distances is designed to search the optimal departure times of trains from their original stations. The optimal solution satisfying δ-balance condition is identified during the search process.

The outcomes include not only a method to evaluate capacity from the tactical level, but also a decision support tool for the railway administrators at the operation level. Since the train scheduling problem with departure choice is embedded into the capacity evaluation model, the proposed model and solution method can capture the optimal departure time of trains from the original stations. Additionally, the model and method proposed in this paper can be readily extended to double-tracks/multi-tracks railway system. Another important extension is to apply the proposed method to different disruption scenarios, and identify quantitatively the capacity loss from the viewpoint of railway administrators.

This paper is organized as follows. Section 2 reviews the related literature. The definition of two-way δ-balance traffic loading capacity in the single-track railway system is presented in Section 3. A 0-1 mixed integer programming model is represented in Section 4. The proposed solution method is introduced in Section 5, and experimental results are analyzed in Section 6. Finally, conclusions are presented in Section 7.

2. Literature review

Traditionally, railway capacity is defined as the maximal number of trains that can safely traverse the entire railway line in a given period of time. In Abril et al. (2008), the railway capacity can be classified as *theoretical capacity*, *practical capacity*, *used capacity* and *available capacity* depending on different objectives, and the capacity evaluation can be generalized into three methods: analytical, optimization and simulation. The analytical approach adopts mathematical equations or algebraic expressions to quantify railway capacity, and is often used to calculate *theoretical capacity* of railway lines. The earliest analytical model was developed by Frank (1966) for a single-track railway line. The number of possible trains on a given segment was estimated based on trains travelling at an average speed between two consecutive sidings. Extending on Frank's method, Petersen (1974) considered trains with three different velocities run at a segment to reflect the influence of heterogonous trains on the capacity. In these earlier studies, the departure times of trains are uniformly distributed over a given time period. De Kort et al (2003) adopted a probabilistic (max, +) approach to evaluate theoretical capacity of a high-speed railway corridor under uncertainty in different demand levels. Burdett and Kozan (2006, 2009) analyzed the influence of mixed traffic, signal locations and dwell times of trains on theoretical capacity of a railway corridor. They developed analytical techniques based on the critical section and train proportions. An improved railway capacity analysis method (Burdett, 2015a) was devoted to schedule trains with return paths in the railway system. The proposed approach allowed planners to identify how many train paths are achievable and how many return paths are possible. Burdett (2015b) formulated and solved a comprehensive set of multi-objective models that perform a trade-off analysis of theoretical capacity. In particular, those models determined theoretical capacity as the most equitable solution, and also provided a set of non-dominated solutions for later analysis and comparison.

An enhanced parametric capacity evaluation was proposed by Lai and Barkan (2009) to assist railroad companies in capacity expansion projects. Based on an estimated future demand and available budget, the proposed model can generate possible expansion alternatives, and compute line capacity and investment costs. In Bevrani et al. (2015), an optimization approach was applied to a case study of the Iran national railway in order to identify its current theoretical capacity and to optimally expand it given a variety of technical conditions. It tentatively demonstrated how an analytical approach for capacity expansion is more efficient than a manual process. Burdett (2016) considered two capacity expansion possibilities, i.e., track duplications and section subdivisions. The case study showed that section subdivision is the best and cheapest option as the cost of track duplications is proportional to its length, whereas subdivision is a static cost.

Most analytical models in the literature address the calculation of theoretical capacity, and are usually used to identify the bottlenecks of the railway lines. However, the analytical approaches ignored the effects of variations in traffic and operations that occur in reality. In practice, the actual railway capacity was far lower than the value obtained by the analytical approaches (Abril et al., 2008).

Optimization methods for capacity evaluation are linked closely to the determination of saturated timetables. The UIC 406 (2004) is one such method, which is developed by the International Union of Railways in Europe to calculate the saturated capacity and is widely adopted in many Europe countries (Robert, 2005; Alex Landex et al, 2006; Wahlborg, 2005; Melody and Preston, 2010; Lindner, 2011). The UIC 406 modifies a pre-determined timetable and reschedules the trains as close as possible to each other (Abril et al., 2008). If the compression indicates free capacity, more trains can be added to the railway system. Landex et al (2006) described in detail the application of UIC 406 in Denmark, while Lindner (2011) applied UIC 406 to evaluate the corridor and station capacity. However, Mussone and Calvo (2013) pointed out that UIC 406 was inadequate for capacity evaluation of railway junctions and station tracks. Additionally, the timetable compression method was designed primarily to analyze capacity of double- and multi-tracks railway system.

Simulation techniques have often been used to model the movement of trains across a railway network. They allow a real world railway environment to be mimicked in great detail. It has already applied into train scheduling problem (Li et al., 2008, 2014; Xu et al., 2015; Mu and Dessouky, 2011, 2013; Liu et al., 2014). Because of its flexibility and high-efficiency, simulation can be used to evaluate practical capacity of railway system by combining with other optimization methods.

Petersen (1974), Petersen and Taylor (1982) considered the combination technique for a single-track rail line, in which the dynamic programming and the branch-and-bound were embedded into the simulation process. Welch and Gussow (1986) developed two “what-if” simulation models to evaluate the relative effect of many factors influencing main line capacity in Canada. Kaas (1991) presented a general simulation model to evaluate railway network capacity at different levels. Dessouky and Leachman (1995) used a simulation framework to analyze the relationship between track capacity and train delay. Their simulation model considered important physical parameters such as train length, speed limits and train headways.

Previous research has focused upon focused on factors affecting railway capacity, such as railway infrastructure layout, mixed traffic and operational parameters. Very few previous research works have ever considered the capacity analysis of single-track railway system from the viewpoint of the administrators. Furthermore, it is very difficult to evaluate the capacity of the single-track railroad as a whole system due to the strong coupling relationship between rail segments and stations. The following two sections discuss in detail the characteristics of railway capacity under the viewpoint of railway administrators, and present a 0-1 mixed integer programming model for capacity analysis of single-track railway corridor.

3. Capacity of a single-track railway corridor with two-way balanced traffic

3.1 Two-way traffic characteristic of the single-track railway

A single-track railway corridor is made up by a series of single-track segments that link stations and sidings.

Frank (1966) was the first to characterize the distinct characteristics of two-way traffic in single-track railway systems, where a segment between stations can be used by the trains in different directions (though of course, only trains travelling in the same direction can occupy the segment at the same time). Here, we name the two travel directions as out- and in-bound. The number of outbound and inbound trains is set to be equal so as to impose quantity balance in two directions. We couple one outbound with one inbound train to form a ***train-pair***. The capacity of a single-track railway corridor is defined as the maximal number of ***train-pairs*** that can travel along the corridor during a fixed time period.

3.2 Average travel time of all trains: a λ-tolerance factor

The more train-pairs in a single-track system, the more interactions among trains (on track and segment occupancy by trains in different directions) there will be and hence longer travel time of trains. More meeting-crossings between trains result also in more waiting time of trains at stations. An interesting problem discussed in this paper is to investigate railway capacity under a certain delay tolerance range. The acceptable maximal delay of trains is considered as an input parameter of the proposed model. However, due to unknown timetable, the value of the maximal delay is unbounded and cannot be estimated. And hence, the value of the free travel time of train is adopted as a benchmark of evaluating the acceptable delay. The question on single-track railway capacity can be better expressed as: *what is the maximal number of train-pairs that can be loaded onto the single-track railway corridor if the average travel time of trains does not exceed a given level?*

We introduce a parameter λ to describe the acceptable level of the administrators. Assume that the number of train-pairs to be loaded is *N* and the loaded train types are denoted as $J=\{1,2,...,j,...,|J|\}$, where $|J|$ is the number of train types. The average free travel time of trains is $\overline{f}=\frac{1}{2N}\sum_{u\in V}\sum_{j\in J}\beta_{u,j}\cdot f_j$. Here, the binary parameter $\beta_{u,j}$ identifies whether train u is of type j, while f_j is the free travel time of *j*-type train, which denotes the time required by train passing through railway system without unnecessary waiting.

Administrators are interested in whether the average travel time of these loaded trains does not exceed $\lambda\cdot\overline{f}$, or what is the maximal number of loaded train-pairs when the average travel time of trains is within the acceptable tolerance level $\lambda\cdot\overline{f}$. Here, the parameter λ is a real number ($\lambda>1$), and we term it "the acceptable travel time factor (the **λ-tolerance** factor)".

3.3 Travel times of trains in different directions: a δ-balance factor

The meeting and crossing of trains from different directions is a key feature of single track railway system. It must be carefully managed. When it happens, trains from one direction have to wait at stations to let the trains in the other direction pass. As well as to minimize total travel time of all trains, the administrators usually also hope that large deviation in travel times between train flows in different directions can be avoided as possible.

The concept of relative balance is to represent the deviation between out- and in-bound travel time, and it reflects the subjective non-preference of the administrators. Let $\overline{f}_{out}$ and $\overline{f}_{in}$ denotes the average travel time of the out- and in-bound train flows, respectively. The relative balance is described as follows:

$$|\overline{f}_{out}-\overline{f}_{in}|\ \leq\delta\cdot\mathcal{D}_N^{\max} \tag{1}$$

where, parameter δ is called the balance degree and is a real number ($0<\delta\leq 1$). Eq. (1) is called "**δ-balance**

condition”. The other important parameter $\mathcal{D}_N^{\max}$ in Eq. (1) is the maximal deviation between the outbound and inbound train flows. It should be emphasized that $\mathcal{D}_N^{\max}$ is not an input parameter of the model; rather its value is dependent on the number of loaded train-pairs. The value of $\mathcal{D}_N^{\max}$ is determined by the solution of a specific train scheduling process. Assume that the number of the loaded train-pairs is N and all outbound trains travel freely in the single-track railway corridor. When the meet-crossing between trains appears, all inbound trains must dwell on the stations to avoid the outbound trains. According to the waiting time of all inbound trains at stations, the maximal deviation can be obtained. It should be pointed out that, for the specific train scheduling process, the waiting time of all inbound trains is required to be minimal because of the flexibility of departure time of trains from the original station. In Appendix II-B, this specific train scheduling process is described in detail, and a pre-processing procedure is presented to obtain the value of $\mathcal{D}_N^{\max}$.

3.4 Initial departure-time choice of trains

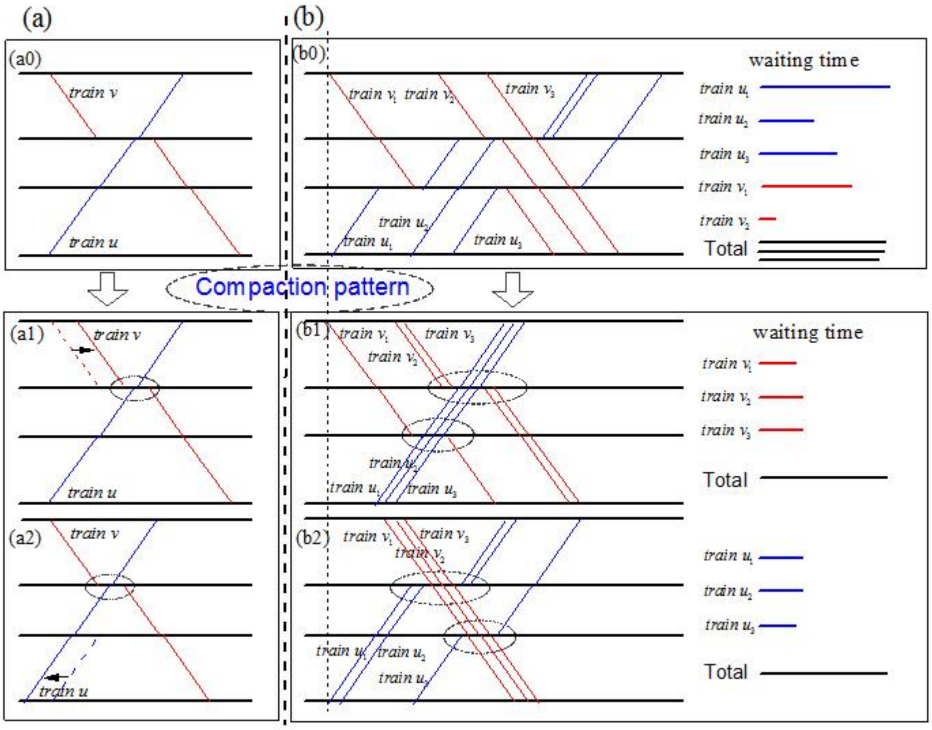

Figure 1 two examples for different departure time of trains: (a) one pair of trains (b) three pairs of trains

The capacity analysis proposed in this paper takes into account different travel tolerance levels set by administrators. Minimizing the total travel times of trains is the basis of accurate capacity evaluation. In a train scheduling problem, the appropriate initial departure times of trains can reduce the travel times of trains in the railway system. Figure 1 shows that selecting the appropriate departure time can significantly reduce the unnecessary waiting time of trains at stations. Hence, the initial departure times of trains from their original stations should be regarded as the decision variables rather than the input parameters. It is emphasized that train scheduling problem with initial departure choice is an important element in the capacity evaluation model proposed in this paper.

4. Model formulation: a 0-1 mixed integer programming

This section presents a 0-1 mixed-integer programming formulation for the two-way δ-balance traffic loading capacity problem in a single-track railway corridor. A summary of the notations adopted in the model is presented in Appendix I.

Model:

Maximize N (2)

Subject to:

(a) flow conservation constraints:

$$n_j = \lfloor N \cdot \gamma_j \rfloor \qquad j = 1,2,\ldots,|J|-1 \tag{3}$$

$$\sum_{j \in J} n_j = N \tag{4}$$

(b) train proportion conservation constraints:

$$n_j = \sum_{u \in V^O} \beta_{u,j} \qquad j = 1,2,\ldots,|J| \tag{5a}$$

$$n_j = \sum_{u \in V^I} \beta_{u,j} \qquad j = 1,2,\ldots,|J| \tag{5b}$$

(c) Travel tolerance level constraints:

$$\sum_{u \in V^O \cup V^I} (t^a_{u,\bar{r}_u} - t^d_{u,\tilde{r}_u}) \le \lambda \cdot [\sum_{u \in V^O \cup V^I} \sum_{j \in J} \beta_{u,j} \cdot f_j] \tag{6}$$

(d) δ-balance constraints:

$$|\sum_{u \in V^O} (t^a_{u,\bar{r}_u} - t^d_{u,\tilde{r}_u}) - \sum_{v \in V^I} (t^a_{v,\bar{r}_v} - t^d_{v,\tilde{r}_v})| \le \delta \cdot \mathfrak{D}^{\max}_N \cdot N \tag{7}$$

(e) Departure time choice constraints:

$$0 \le t^d_{u,\tilde{r}_u} < T \qquad \forall \ u \in V^O \cup V^I \tag{8}$$

(g) Constraints II-4~II-11 in Appendix II-A. (9)

The objective of the model is to maximize the number N of train-pairs that can be loaded into the single-track railway corridor. The input parameter of the model is the proportion of different types of trains, which is indicated by symbol γ_j, and $\sum_{j \in J} \gamma_j = 1$. According to the proportion coefficient γ_j, the number (n_j) of different types of trains in the out- and in-bound directions is deduced by the number of train-pairs (see constraints (3) and (4)). Because train number is always an integer, symbol "$\lfloor - \rfloor$" denotes the integer part of $N \cdot \gamma_j$. Clearly, the number n_j is related to the decision variable N. Constraints (5) ensure that the loaded trains in the out- and in-bound directions satisfy the proportion of different types of trains. These trains is recorded in set V^O and V^I.

As we described in Section 3.2, this study focuses on the maximal number of train-pairs when the average travel time of trains is confined to a given level. Constraint (6) ensures that the total travel time of the loaded trains does not exceed the expected value ($\lambda \cdot [\sum_{u \in V^O \cup V^I} \sum_{j \in J} \beta_{u,j} \cdot f_j]$), which is corresponding to λ-tolerance level. Variable $t^d_{u,\tilde{r}_u}$ is the departure time of train u from its original station $\tilde{r}_u$, and $t^a_{u,\bar{r}_u}$ is the arrival time of train u at its destination station $\bar{r}_u$. Clearly, the capacity evaluation investigated in this study is closely related to a train scheduling process. Different from the standard train scheduling problem, the specific scheduling process emphasizes the relative balance in travel times between train flows in different directions. Constraint (7), which is called as "*δ-balance condition*", ensures that the travel deviation between out- and in-bound train flows is confined to an expected range of railway administrators.

Constraints (8) ensure that all loaded trains must depart from their original stations within the time window $[0, T]$, where T is the minimum free-flow travel time of all loaded trains, i.e., $T = \min(f_j \mid j \in J)$. The time window ensures that no train can leave the system before all trains have been loaded onto the railway corridor.

Similar to the standard train scheduling problem, certain additional constraints are necessary to reflect the travelling characteristic of trains in the single-track railway system, which include ***headway constraints***, ***meeting-crossing constraints***, ***station capacity constraints***, ***segment running time constraints*** *and* ***stopping/non-stopping constraints***. These constraints have already been discussed in detail in our previous works (Li et al, 2014). And hence, we list these constraints in Appendix II.A (constraints (II-3) ~ (II-11)).

5 Solution algorithm

5.1 Model analysis and heuristic framework

The model proposed above yields a 0-1 mixed integer programming formulation for the evaluation of two-way δ-balance traffic loading capacity in the single-track railway corridor. Constraints (6) ~ (9) mean that capacity evaluation is related closely to a train schedule plan. Constraint (6) is an evaluation criterion, which identifies whether there is a feasible train schedule plan that satisfies the accepted tolerance level. If the maximal number of train-pairs is N, it is concluded that no feasible schedule plan can satisfy constraint (6) when the number of train-pairs is $N+1$. In other words, even the schedule plan with the minimal total travel time also exceeds the acceptable tolerance level set by the administrators. While constraints (7) and constraints (II-3 ~ II-11) in Appendix II.A reflect the travel process of trains loaded onto the single-track railway system.

Assume that the number of trains loaded into the railway corridor is known. We formulate a specific train scheduling problem with initial departure choice, which is subject to the relative balance of train flows in different directions, and minimize the total travel times of all trains loaded onto the single-track railway corridor. This model is noted by symbol $\mathfrak{M}(N)$, and is presented in Appendix II.A. From the solution of model $\mathfrak{M}(N)$, it is identified whether tolerance level constraint (6) is satisfied.

A simple dichotomizing-based method is adopted to explore the maximal number of train-pairs in the single-track railway corridor. Firstly, we set the initial lower bound $\tilde{n}_{lb}$ and upper bound $\tilde{n}_{ub}$ of the number of train-pairs. The initial lower bound may be set to 1, and the initial upper bound is set to $\lfloor T / h^{dd} \rfloor$; the latter is the possible maximal number of train-pairs in fixed time window $[0, T]$. Here, parameter h^{dd} is the safety headway between two trains departing from the original station. Moreover, we analyze whether the solution of model $\mathfrak{M}(\tilde{n}_{lb} + \lfloor (\tilde{n}_{ub} - \tilde{n}_{lb}) / 2 \rfloor)$ satisfies travel tolerance condition (constraint (6)). If it is, $\tilde{n}_{lb} + \lfloor (\tilde{n}_{ub} - \tilde{n}_{lb}) / 2 \rfloor$ is set to new lower bound; otherwise, it is regarded as the value of upper bound. Table 1 presents a detailed heuristic procedure.

Table 1: Dichotomizing-based heuristic search

Set initial values for $\tilde{n}_{lb}$ and $\tilde{n}_{ub}$ ($\tilde{n}_{lb} = 1$, $\tilde{n}_{ub} = \lfloor T / h^{dd} \rfloor$);

While $\tilde{n}_{lb} < \tilde{n}_{ub}$ do

Repeat

Set $N = \tilde{n}_{lb} + \lfloor (\tilde{n}_{ub} - \tilde{n}_{lb}) / 2 \rfloor$

Solve the train scheduling problem $\mathfrak{M}(N)$;

Update train-pair numbers:

If $\frac{1}{2N}\sum_{u}(t^{a}_{u,\bar{r}_u} - t^{d}_{u,\tilde{r}_u}) \le \lambda \cdot \overline{f}$, then $\tilde{n}_{lb} = N$

If $\frac{1}{2N}\sum_{u}(t^{a}_{u,\bar{r}_u} - t^{d}_{u,\tilde{r}_u}) > \lambda \cdot \overline{f}$ or no feasible solution is found, then $\tilde{n}_{ub} = N$

End While

Output the value of N

The above dichotomizing-based heuristic is straightforward. However, a pivotal issue is how to solve model $\mathfrak{M}(N)$ efficiently. The solution of model $\mathfrak{M}(N)$ includes: the initial departure time of each train from their original stations, and their arrival and departure times at other stations. This can be expressed as $\mathcal{S}=\{\mathcal{T}(V), \mathbb{S}(V)\}$, where $\mathcal{T}(V)$ records the departure times of trains from their original stations, i.e., $\mathcal{T}(V) = \{t^{d}_{u,\tilde{r}_u} \mid u \in V\}$, and $\mathbb{S}(V)$ records the arrival and departure times of trains at stations, i.e., $\mathbb{S}(V) = \{(t^{a}_{u,r}, t^{d}_{u,r}) \mid u \in V, r \in R_u\}$. Here, $t^{a}_{u,r}$ and $t^{d}_{u,r}$ are the arrival and departure time of train u at station r, respectively.

It is well-known that the branch-and-bound algorithm is a precise method to solve the 0-1 mixed-integer programming problem. However, as a non-polynomial method, the branch-and-bound may be unable to obtain the optimal solution. For a large-scale problem, even a feasible solution can hardly be obtained within finite computational time. If the departure times of trains from their original stations are relaxed, solving train scheduling problem becomes even more difficult. Compared with train scheduling problem with expected initial departure time, the choice of train departure time and order in $\mathfrak{M}(N)$ will result in a larger feasible region.

We adopt symbol $\mathfrak{M}(N \mid \mathcal{T}(V))$ to denote train schedule problem with expected/fixed departure times. There are many excellent methods for train scheduling in the literature (e.g. Carey, 1994; Higgins et al., 1996, 1997; Cai et al., 1998; Zhou and Zhong, 2007; Burdett and Kozan, 2009a, 2009b, 2014a, 2014b)). In our previous works (Li et al, 2014), a Confliction-Distribution-Prediction method (CDP) was developed to solve $\mathfrak{M}(N \mid \mathcal{T}(V))$ efficiently. However, the CDP focused on train scheduling problem with expected departure times. Figure 1 provides two simple examples to demonstrate that proper initial departure times of trains can largely reduce unnecessary waiting times of trains at stations. In Figure 1 (a1), the waiting time of train v at station is reduced only by changing the departure time of train u or v from the original station. The proper departure time of trains in Figure 1 (b1 and b2) make the waiting times of all trains at stations reduce three times approximately. Hence, how to determine the proper initial departure time for each train is the key issue to solve $\mathfrak{M}(N)$. Based on the comparison between two optimal schedule plans, we develop an initial departure choice procedure based on "compaction pattern" to determine the optimal or suboptimal initial departure time of trains.

5.2 Determine the initial departure time of trains at the original stations

The initial departure choice of train is influenced by many factors, such as crew and rolling stock. However, in this paper, the evaluation of two-way traffic loading capacity is based on the minimal total travel times of all trains. Hence, we only focus on how to determine the initial departure times of trains so as to minimize the total travel time of trains.

5.2.1 The definition of compaction pattern

Determining the optimal initial departure time of train is very difficult due to the unknown schedule plan. Different initial departure times of trains will result in the schedule plans with different structures. For instance, Figure 1 (b0) presents an optimal schedule plan, in which the initial departure times of trains are given in advance; while Figure 1 (b1) and (b2) show two schedule plans with optimal initial departure times of trains. In Figure 1 (b1 and b2), a ***compaction pattern*** is developed where trains wait at the station for meet-crossing between trains. ***Compaction pattern*** denotes that the arrival or departure interval between trains at stations reach the minimal headway. In other words, the waiting times of trains in compaction pattern cannot be compressed any further.

Compaction pattern provides a novel idea to seek the optimal or near-optimal departure times of trains. Assume we can obtain quickly a train schedule plan based on a given initial departure times of trains. According to the arrival and departure time distribution of trains at stations, the compressible time-distances among trains can be analyzed. By adjusting the initial departure times of trains, time-points distribution is gradually converged towards compaction pattern. We call the algorithm proposed for the optimal initial departure of trains as “the initial departure choice based on compaction pattern”, or simply “IDC_CP”.

5.2.2 Descriptions of compaction pattern at station

Let set $\mathfrak{D}$ denote the travel information of trains at stations given by a schedule plan, and it can be expressed as $\mathfrak{D}=\{\mathcal{D}_u \mid u \in V\}$, where $\mathcal{D}_u$ records the travel information of train u at each station, i.e., $\mathcal{D}_u=\{\mathcal{D}_u^r \mid r \in R_u\}$. The information unit $\mathcal{D}_u^r$ contains three elements, and is expressed as $\mathcal{D}_u^r=((t_{u,r}^a, t_{u,r}^d), \mathcal{X}_u^r(\mathcal{V}_u^r), \mathcal{F}_u^r(\mathcal{V}_u^r))$, where the first part denotes the time interval between $t_{u,r}^a$ and $t_{u,r}^d$, $\mathcal{X}_u^r(\mathcal{V}_u^r)$ and $\mathcal{F}_u^r(\mathcal{V}_u^r)$ identify the arrival or departure characteristic and time-point distribution of trains in region $(t_{u,r}^a, t_{u,r}^d)$. Here, set $\mathcal{V}_u^r$ records the trains which have arrived and/or departed during time interval $(t_{u,r}^a, t_{u,r}^d)$, i.e., $\mathcal{V}_u^r=\{\lambda_v \mid v \in V, \text{ and } t_{u,r}^a < t_{v,r}^a < t_{u,r}^d \text{ or } t_{u,r}^a < t_{v,r}^d < t_{u,r}^d\}$, and λ_v is the ID of train v. $\mathcal{X}_u^r(\mathcal{V}_u^r)$ and $\mathcal{F}_u^r(\mathcal{V}_u^r)$ can be expressed as $\mathcal{X}_u^r(\mathcal{V}_u^r)=\{\mathcal{X}_v^r \mid v \in \mathcal{V}_u^r\}$ and $\mathcal{F}_u^r(\mathcal{V}_u^r)=\{\mathrm{p}_v^r \mid v \in \mathcal{V}_u^r\}$, respectively. Their definitions are listed as follows:

$$\mathcal{X}_v^r=\begin{cases}1 & \text{if time point indicates train } v \text{ arrives at station } r, \text{ and } v \in \mathcal{V}_u^r \\ 0 & \text{if time point indicates train } v \text{ departes from station } r, \text{ and } v \in \mathcal{V}_u^r\end{cases} \tag{10}$$

$$\mathrm{p}_v^r=\begin{cases} t_{v,r}^a & \text{if } \mathcal{X}_v^r=1, \text{ and } v \in \mathcal{V}_u^r \\ t_{v,r}^d & \text{if } \mathcal{X}_v^r=0, \text{ and } v \in \mathcal{V}_u^r\end{cases} \tag{11}$$

According to $\mathcal{F}_u^r(\mathcal{V}_u^r)$, the time-points distribution is expressed as $\Gamma_u^r=\{t_{u,r}^a, t_{u,r}^d\} \cup \mathcal{F}_u^r(\mathcal{V}_u^r)$. Figure 2 (a) presents a simple example to explain intuitively the definitions of the above symbols. In the region $(t_{u,r}^a, t_{u,r}^d)$, three trains (v_1, v_2, v_3) arrive at or depart from station r. It is concluded that, $\mathcal{V}_u^r=\{\lambda_{v_1},\lambda_{v_3},\lambda_{v_1},\lambda_{v_2},\lambda_{v_3},\lambda_{v_2}\}$, $\mathcal{X}_u^r=\{1,\ 1,\ 0,\ 1,\ 0,\ 0\}$, $\mathcal{F}_u^r=\{t_{v_1,r}^a,t_{v_3,r}^a,t_{v_1,r}^d,t_{v_2,r}^a,t_{v_3,r}^d,t_{v_2,r}^d\}$ and $\Gamma_u^r=\{t_{u,r}^a,t_{v_1,r}^a,t_{v_3,r}^a,t_{v_1,r}^d,\ t_{v_2,r}^a,t_{v_3,r}^d,\ t_{v_2,r}^d,t_{u,r}^d\}$.

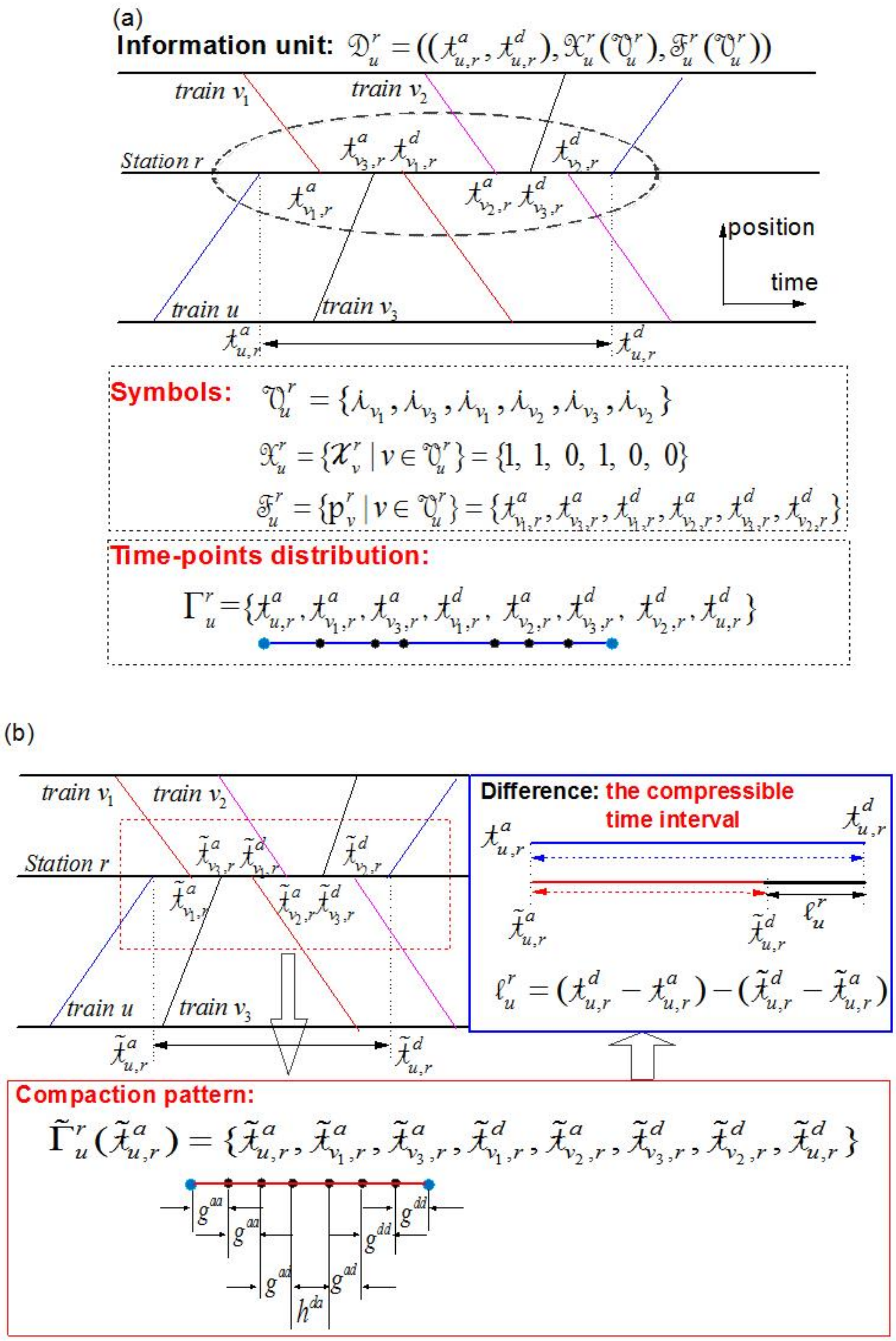

Figure 2 Graphic descriptions of symbols adopted in compaction pattern (a) time-points distribution Γ_u^r and (b) compaction pattern $\tilde{\Gamma}_u^r$

- **The compaction pattern formulation $\tilde{\Gamma}_u^r$**

The aim of initial departure time choice of trains is to make Γ_u^r closely to its compaction pattern $\tilde{\Gamma}_u^r$ as possible, and reduce the unnecessary waiting time of trains. Assume that a new arrival time of train u at station r is $\tilde{\mathfrak{t}}_{u,r}^{a}$ after initial departure times of trains are adjusted. Based on $\tilde{\mathfrak{t}}_{u,r}^{a}$, arrival or departure characteristic $\mathfrak{X}_u^r(\mathfrak{V}_u^r)$ in set $\mathfrak{V}_u^r$, and minimum headways between trains, an ideal compressed time-points distribution $\tilde{\mathfrak{F}}_u^r(\mathfrak{V}_u^r)$ can be reformulated by a mapping function $\mathfrak{W}(\tilde{\mathfrak{t}}_{u,r}^{a}, \mathfrak{X}_u^r(\mathfrak{V}_u^r))$, i.e., $\tilde{\mathfrak{F}}_u^r(\mathfrak{V}_u^r) = \mathfrak{W}(\tilde{\mathfrak{t}}_{u,r}^{a}, \mathfrak{X}_u^r(\mathfrak{V}_u^r))$. The mapping rule of $\mathfrak{W}(\tilde{\mathfrak{t}}_{u,r}^{a}, \mathfrak{X}_u^r(\mathfrak{V}_u^r))$ is presented in Table 2. We adopt the first time-point of the distribution presented in Figure 2 (a) to explain the mapping rule in Table 2. The first time-point is the arrival time of train v_1 at station r. Because the direction of train v_1 is opposite to train u, the arrival-arrival safety headway (g^{aa}) is considered as critical time interval in compaction pattern. And hence, the first time-point in compaction pattern can be written to $\tilde{\mathfrak{t}}_{u,r}^{a} + g^{aa}$.

Table 2: The mapping rule in $\mathfrak{W}(\tilde{\mathfrak{t}}_{u,r}^{a}, \mathfrak{X}_u^r(\mathfrak{V}_u^r))$ for $\tilde{\mathfrak{F}}_u^r(\mathfrak{V}_u^r)$

The characteristic of train	conditions	$\tilde{\mathrm{p}}_v^r$
$v = \bar{v}, v \in \mathfrak{V}_u^{r\,*}$	$\mathcal{X}_v^r = 1;\ u, v \in V^O$ or $u, v \in V^I$	$\tilde{\mathfrak{t}}_{u,r}^{a} + h^{aa}$
	$\mathcal{X}_v^r = 0;\ u, v \in V^O$ or $u, v \in V^I$	$\tilde{\mathfrak{t}}_{u,r}^{a} + h^{ad}$
	$\mathcal{X}_v^r = 1;\ u \in V^O, v \in V^I$ or $u \in V^O, v \in V^I$	$\tilde{\mathfrak{t}}_{u,r}^{a} + g^{aa}$
	$\mathcal{X}_v^r = 0;\ u \in V^O, v \in V^I$ or $u \in V^O, v \in V^I$	$\tilde{\mathfrak{t}}_{u,r}^{a} + g^{ad}$
$v \neq \bar{v}, v \in \mathfrak{V}_u^r$	$\lambda_v = \lambda_{v^-}$	$\tilde{\mathrm{p}}_{v^-}^r + \tau_b$
	$\lambda_v \neq \lambda_{\bar{v}}; \mathcal{X}_{v^-}^r = 1, \mathcal{X}_v^r = 1; v, v^- \in V^O$ or $v, v^- \in V^I$	$\tilde{\mathrm{p}}_{v^-}^r + h^{aa}$
	$\lambda_v \neq \lambda_{\bar{v}}; \mathcal{X}_{v^-}^r = 1, \mathcal{X}_v^r = 0; v, v^- \in V^O$ or $v, v^- \in V^I$	$\tilde{\mathrm{p}}_{v^-}^r + h^{ad}$
	$\lambda_v \neq \lambda_{\bar{v}}; \mathcal{X}_{v^-}^r = 0, \mathcal{X}_v^r = 0; v, v^- \in V^O$ or $v, v^- \in V^I$	$\tilde{\mathrm{p}}_{v^-}^r + h^{dd}$
	$\lambda_v \neq \lambda_{\bar{v}}; \mathcal{X}_{v^-}^r = 0, \mathcal{X}_v^r = 1; v, v^- \in V^O$ or $v, v^- \in V^I$	$\tilde{\mathrm{p}}_{v^-}^r + h^{da}$
	$\lambda_v \neq \lambda_{\bar{v}}; \mathcal{X}_{v^-}^r = 1, \mathcal{X}_v^r = 1; v \in V^O, v^- \in V^I$ or $v \in V^I, v^- \in V^O$	$\tilde{\mathrm{p}}_{v^-}^r + g^{aa}$
	$\lambda_v \neq \lambda_{\bar{v}}; \mathcal{X}_{v^-}^r = 1, \mathcal{X}_v^r = 0; v \in V^O, v^- \in V^I$ or $v \in V^I, v^- \in V^O$	$\tilde{\mathrm{p}}_{v^-}^r + g^{ad}$
	$\lambda_v \neq \lambda_{\bar{v}}, \mathcal{X}_{v^-}^r = 0, \mathcal{X}_v^r = 0; v \in V^O, v^- \in V^I$ or $v \in V^I, v^- \in V^O$	$\tilde{\mathrm{p}}_{v^-}^r + g^{dd}$

*Note: $\bar{v}$ is the first train in set $\mathcal{V}_u^r$; v^- is the ahead train of train v in set $\mathcal{V}_u^r$.

Combined ($\tilde{t}_{u,r}^a$, $\tilde{t}_{u,r}^d$) with $\tilde{\mathfrak{F}}_u^r(\mathcal{V}_u^r)$, the ideal compaction pattern $\tilde{\Gamma}_u^r(\tilde{t}_{u,r}^a)$ is expressed as $\tilde{\Gamma}_u^r(\tilde{t}_{u,r}^a)=\{\tilde{t}_{u,r}^a,\tilde{t}_{u,r}^d\}\bigcup\tilde{\mathfrak{F}}_u^r(\mathcal{V}_u^r)$. Figure 2 (b) shows the compaction pattern of the example in Figure 2(a), which is expressed as $\tilde{\Gamma}_u^r(\tilde{t}_{u,r}^a)=\{\tilde{t}_{u,r}^a,\tilde{t}_{v_1,r}^a,\tilde{t}_{v_3,r}^a,\tilde{t}_{v_1,r}^d,\tilde{t}_{v_2,r}^a,\tilde{t}_{v_3,r}^d,\tilde{t}_{v_2,r}^d,\tilde{t}_{u,r}^d\}$, and $\tilde{t}_{v_1,r}^a=\tilde{t}_{u,r}^a+g^{aa}$; $\tilde{t}_{v_3,r}^a=\tilde{t}_{v_1,r}^a+g^{aa}$; $\tilde{t}_{v_1,r}^d=\tilde{t}_{v_3,r}^a+g^{ad}$; $\tilde{t}_{v_2,r}^a=\tilde{t}_{v_1,r}^d+h^{da}$; $\tilde{t}_{v_3,r}^d=\tilde{t}_{v_2,r}^a+g^{ad}$; $\tilde{t}_{v_2,r}^d=\tilde{t}_{v_3,r}^d+g^{dd}$; $\tilde{t}_{u,r}^d=\tilde{t}_{v_2,r}^d+g^{dd}$. Clearly, in the compaction pattern, the time-distance between the neighboring time points is compressed to a critical value.

- **The compressible time interval between Γ_u^r and $\tilde{\Gamma}_u^r$**

The difference between Γ_u^r and $\tilde{\Gamma}_u^r$ can be measured by the compressible time interval ℓ_u^r , i.e., $\ell_u^r=(t_{u,r}^d-t_{u,r}^a)-(\tilde{t}_{u,r}^d-\tilde{t}_{u,r}^a)$, which is an important evaluation criterion for designing the departure choice procedure of trains.

5.2.3 The characteristic descriptions of compaction pattern at segment

The behaviors of trains on a segment can also be included into the compaction pattern. The travel information of trains at segments are recorded in set $\mathfrak{L}=\{\mathfrak{L}_u\,|\,u\in V\}$. Here, $\mathfrak{L}_u$ can be expressed as $\mathfrak{L}_u=\{\delta_u^{r,r^+}\,|\,r,r^+\in R_u\}$, where δ_u^{r,r^+} is the delay time of train u at the segment between station r and r^+, i.e., $\delta_u^{r,r^+}=(t_{u,r^+}^a-t_{u,r}^d)-(p_u^{r,r^+}+\vartheta_u^r\cdot\tau_{u,e}^a+\vartheta_u^{r^+}\cdot\tau_{u,e}^d)$. If $\delta_u^{r,r^+}=0$, the travel of train u at the segment between station r and r^+ subjects to compaction pattern of segment.

5.2.4 The initial departure choice of trains

- **The departure adjustment of trains based on compaction pattern**

As been depicted by Figure 1, the appropriate initial departure time can efficiently reduce unnecessary waiting time of trains at stations for the meeting and crossing between trains, and make the arrival and departure times of trains at stations closely to compaction pattern as possible. And hence, the aim of the initial departure choice is to reduce the difference between Γ_u^r and $\tilde{\Gamma}_u^r$.

Let the departure times of trains in an initial schedule to be $\mathcal{T}=\{t_{u,\tilde{r}_u}^d\,|\,u\in V\}$. The mapping function $\mathfrak{H}_r(\mathcal{T},\Gamma_u^r,\tilde{\Gamma}_u^r)$ is formulated to determine the new initial departure time of trains, i.e., $\mathcal{T}'=\mathfrak{H}_r(\mathcal{T},\Gamma_u^r,\tilde{\Gamma}_u^r)$. The mapping rule of $\mathfrak{H}_r(\cdot)$ is defined as follows:

$$t^{d}_{v,\tilde{r}_u}{}' = \begin{cases} t^{d}_{v,\tilde{r}_u} + (\tilde{t}^{a}_{v,r} - t^{a}_{v,r}) & \text{if } v = u \\ t^{d}_{v,\tilde{r}_u} + (\tilde{p}^{r}_{v} - p^{r}_{v}) & \text{if } v \in \mathcal{V}^{r}_{u} \\ t^{d}_{v,\tilde{r}_u} & \text{others} \end{cases} \tag{12}$$

Note that $\mathcal{H}_r\,(\mathcal{T},\Gamma^r_u,\tilde{\Gamma}^r_u)$ not only focuses on the departure time choice of train u, but also emphasizes the departure adjustment of trains in set $\mathcal{V}^r_u$. The mapping rule reflects a strong coupling relation among trains in the single-track railway system.

A simple rule that applies delays to the initial departure time of trains is used to reduce train segment delay. Assume that the delay of train u is b^{r,r^+}_u at the segment between station r and r^+ in an initial schedule plan. The new departure time of trains, $\mathcal{T}' = \mathcal{H}_s\,(\mathcal{T}, b^{r,r^+}_u)$, can be formulated by function $\mathcal{H}_s\,(\mathcal{T}, b^{r,r^+}_u)$ as follows.

$$t^{d}_{v,\tilde{r}_u}{}' = \begin{cases} t^{d}_{v,\tilde{r}_u} + b^{r,r^+}_u & \text{if } v = u \\ t^{d}_{v,\tilde{r}_u} & \text{others} \end{cases} \tag{13}$$

- **The feasibility analysis of train initial departure**

Constraints (9) ensure that the initial departure times of trains are restricted to time window $[0,T]$. The departure-departure time headway must also be satisfied for the adjusted initial departure times. Hence, a feasible analysis is necessary for the new train departure sequence generated by $\mathcal{H}_r(\text{--})$ and $\mathcal{H}_s(\text{--})$.

Consider a situation where train u is scheduled to depart before train v, but their initial departure time interval does not satisfy the Departure-Departure headway. We examine the earlier extensible time space of train u and the later extensible time space of train v. The train with more extensible space is selected, and its initial departure time is moved till the Departure-Departure headway is satisfied. Once no extensible time space is found, the examining range is extended to other trains before train u and after train v. The bound analysis of time window is also similar. When the initial departure time of train is left or right bound of time window, the extensible space is set to zero.

5.3 The uniformity apportionment mechanism for balance constraints

According to the above initial departure choice and the CDP method (Li et.al, 2014), a schedule plan can be quickly obtained. However, the balance constraints are not considered in the CDP method. Hence, it is necessary to modify the CDP so that the balance constraints are satisfied. A specific characteristic in the CDP is the travel optimization mechanism, that the travel strategies of trains are analyzed based on the confliction distribution prediction achieved by the greedy method. We adopt a uniformity apportionment mechanism to ensure that the subsequent schedule plan obtained by the greedy method satisfies relative balance condition.

Note that the hard time windows [0, T) in the proposed model can ensure that no train can leave before all trains have been loaded into the railway corridor. When a train travels at its last segment, all meeting-crossings between it and the trains in opposite direction have occurred. It is concluded that all trains travel freely at their last segment. And hence, the uniformity apportionment mechanism is to adjust the travel time of out- or in-bound train flows on their last segment of travel.

In the schedule plan obtained by the greedy mechanism (Li et.al, 2014), total travel time of outbound and inbound train flow are presented as $T^O = \sum_{u\in V^O} (t^{a}_{u,\bar{r}_u} - t^{d}_{u,\tilde{r}_u})$ and $T^I = \sum_{u\in V^I} (t^{a}_{u,\bar{r}_u} - t^{d}_{u,\tilde{r}_u})$, respectively. If

$|T^O - T^I| > \delta \cdot \mathfrak{D}_N^{\max} \cdot N$, the balance constraint cannot be satisfied. Assume that $T^O > T^I$, and the compensated difference between outbound and inbound train flows is $T^O - T^I - \delta \cdot \mathfrak{D}_N^{\max} \cdot N$. The uniformity apportionment mechanism ensures that the compensated difference is assigned equally to all inbound trains. The travel times of all inbound trains at their last segment are delayed till the balance condition is satisfied. The uniformity apportionment mechanism is described as follows.

$$t_{u,\bar{r}_u}^a = \begin{cases} t_{u,\bar{r}_u}^a + (T^O - T^I)/N - \delta \cdot \mathfrak{D}_N^{\max}, & if\ T^O - T^I > \delta \cdot \mathfrak{D}_N^{\max} \cdot N, u \in U^I \\ t_{u,\bar{r}_u}^a + (T^I - T^O)/N - \delta \cdot \mathfrak{D}_N^{\max}, & if\ T^I - T^O > \delta \cdot \mathfrak{D}_N^{\max} \cdot N, u \in U^O \\ t_{u,\bar{r}_u}^a & if\ |T^O - T^I| \le \delta \cdot \mathfrak{D}_N^{\max} \cdot N \end{cases} \tag{14}$$

Based on the integration of uniformity apportionment and greedy mechanism, the modified optimization mechanism in the CDP can identify the satisfactory travel strategies of trains, and ensure that the obtained schedule plan satisfy the relative balance between outbound and inbound train flows.

5.4 The algorithm procedure for solving $\mathfrak{M}(N)$

Table 3: Algorithm IDC_CP (The initial departure choice based on the compaction pattern)

Initialization: Generate $\mathcal{T}^{(0)}(V)$ randomly, and solve $\mathfrak{M}(N|\mathcal{T}^{(0)}(V))$. And then obtain the solution $\mathcal{S}^* = \{\mathcal{T}^*(V), \mathfrak{S}^*(V)\}$, i.e., $z^* \Leftarrow z^{initial}, \mathcal{T}^*(V) = \mathcal{T}^{(0)}(V)$ and $\mathfrak{S}^*(V) = \{(t_{u,r}^{a\ *}, t_{u,r}^{d\ *}) | u \in V, r \in R_u\}$; set up the initial set $\mathfrak{D}$ and $\mathfrak{L}$, i.e., $\mathfrak{D} = \{\mathfrak{D}_u | u \in V\}$ and $\mathfrak{L} = \{\mathfrak{L}_u | u \in V\}$.

While $u < |V|$ (initial $u = 1$) do

Repeat

Detect new initial departure of trains and schedule plan based on ***sub-procedure1(*** $\mathfrak{D}_u$ ***) and sub-procedure2(*** $\mathfrak{L}_u$ ***)***;

If a better solution is found, update z^*, $\mathcal{S}^*$, $\mathfrak{D}$ and $\mathfrak{L}$. Reset $u = 1$;

Otherwise, $u \Leftarrow u + 1$.

End While

Output the value of z^*, and $\mathcal{S}^*(V)$.

Sub-procedure 1 ($\mathfrak{D}_u$**):** the detecting procedure based on $\mathfrak{D}_u = \{\mathfrak{D}_u^r | r \in R_u\}$

While $r \le |R_u|$ (initial $r = 1$) do

Repeat

Based on $\mathfrak{D}_u^r = ((t_{u,r}^a, t_{u,r}^d), \mathfrak{F}_u^r(\mathfrak{V}_u^r), \mathfrak{X}_u^r(\mathfrak{V}_u^r))$, analyze the compressible time-distance ℓ_u^r.

While $k_1 \cdot t_{step} < \ell_u^r$ (initial $k_1 = 0$) do

Repeat

Set $\tilde{\mathcal{t}}^{a}_{u,r} = \mathcal{t}^{a}_{u,r} + k_1 \cdot t_{step}$, and formulate $\tilde{\Gamma}^{r}_{u}$;

Determine the attempted departure initial departure of trains based on $\mathcal{T}' = \mathfrak{H}_r(\mathcal{T}, \Gamma^{r}_{u}, \tilde{\Gamma}^{r}_{u})$, and feasible analysis for $\mathcal{T}'$;

Solve $\mathfrak{M}(N \,|\, \mathcal{T}'(V))$ and analyze the results:

If the better solution is found, then update $\mathcal{S}^*$.

$k_1 \Leftarrow k_1 + 1$;

End while

$r \Leftarrow r + 1$;

End while

Sub-procedure 2 ($\mathfrak{L}_u$): the detection algorithm for $\mathfrak{L}_u$ ($\mathfrak{L}_u = \{ b_u^{r,r^+} \,|\, r, r^+ \in R_u \}$)

While $r \le |R_u| - 1$ (initial $r = 1$) do

Repeat

Set $\mathcal{T}' = \mathfrak{H}_s(\mathcal{T}, b_u^{r,r^+})$, and feasible analysis for $\mathcal{T}'$;

Solve $\mathfrak{M}(N \,|\, \mathcal{T}'(V))$ and analyze the results:

If the better solution is found, then update $\mathcal{S}^*$.

$r \Leftarrow r + 1$;

End while

Algorithm IDC_CP presented in Table 3 starts from an initial schedule plan obtained using the CDP method (Li et.al, 2014). Based on the travel information of each train at station and segment, i.e., $\mathfrak{D}$ and $\mathfrak{L}$, the departure choice procedure is executed for the compaction pattern. If a better solution is found, the information in set $\mathfrak{D}$ and $\mathfrak{L}$ is reset.

6. Numerical experiments

Two important features are investigated through a series of numerical experiments: (1) the quality and computational efficiency of the proposed IDC_CP, and (2) the two-way traffic loading capacity characteristics under different tolerance levels and balance degrees. The algorithms proposed in Section 5 is implemented in C++ language and executed on a PC with Windows 7 operating system, equipped with an Intel E5-4620 2.2 GHz processor and 8G RAM.

We consider a five-station and four-segment single track railway corridor. We randomly generate ten scenarios with small-scale variations in total length of the corridor and the lengths of the four segments. Table 4 lists the instances generated.

Table 4: The list of fourteen examples generated randomly (unit: km)

Inst.	Total length	Segment 1	Segment 2	Segment 3	Segment 4	Inst.	Total length	Segment 1	Segment 2	Segment 3	Segment 4

1	156	32	46	38	40	6	158	35	40	45	38
2	154	30	40	48	36	7	160	38	43	47	32
3	160	40	36	46	38	8	158	32	41	38	47
4	152	35	42	39	36	9	154	36	47	31	40
5	158	42	36	43	37	10	160	33	42	47	38

The segment lengths in each sample are uniformly distributed values between 30 and 50. The number of sidings at each station is set to 3. Small scale examples are adopted to evaluate the difference between the solution obtained by the IDC_CP method and the optimal solution. The optimal solutions are obtained by the branch-and-bound method, which is realized by the standard CPLEX MIP algorithm (version 12.6).

6.1 The performance of IDC_CP for train balance scheduling with departure choice

The initial departure choice of trains and the balance degree are two distinct characteristics of model $\mathfrak{M}(N)$. In the following experiments, we focus on these two characteristics of the model $\mathfrak{M}(N)$ and the performance of algorithm IDC_CP.

6.1.1 The importance analysis of train initial departure choice

Firstly, we identify the influence of flexible initial departure on the performance of train scheduling problem. Table 5 presents the results of model $\mathfrak{M}(N)$ and $\mathfrak{M}(N|\mathcal{T}(V))$ obtained by the branch-and-bound and the proposed IDC_CP. The number of train-pairs is set to 4. In model $\mathfrak{M}(N|\mathcal{T}(V))$, the interval between initial departure times of trains is set to 20 min. The balance constraint is relaxed in the results presented in Table 5.

The results show that the total travel time of all trains in $\mathfrak{M}(N)$ is reduced by 0.1504 compared to that in $\mathfrak{M}(N|\mathcal{T}(V))$. Three indictors, i.e., the number of the compressible interval (n^c), the total compressible time-distance ($\sum \ell_u^r$) and the maximal compressible interval ($\max\{\ell_u^r\}$), are indicated to identify the difference between the solutions of $\mathfrak{M}(N)$ and $\mathfrak{M}(N|\mathcal{T}(V))$. These indictors reflect unnecessary waiting or delay times of trains at stations and segments. The average values of three indictors ($\bar{n}^c$, $\overline{\sum \ell_u^r}$, $\overline{\max\{\ell_u^r\}}$) in ten examples for $\mathfrak{M}(N)$ and $\mathfrak{M}(N|\mathcal{T}(V))$ are (11.6, 65.5, 21.6) and (3.0, 3.3, 1.3), respectively. Clearly, it is proved that the rational initial departure times of trains can efficiently avoid the unnecessary delay of trains, and make the arrival or departure time distribution of trains at stations closely to the compaction pattern.

However, when the branch-and-bound is applied, the average computational time for $\mathfrak{M}(N|\mathcal{T}(V))$ and $\mathfrak{M}(N)$ is about 0.142h and 1.363h respectively, i.e. significantly higher computation time for $\mathfrak{M}(N)$ with branch-and-bound. The flexibility of initial departure time makes model $\mathfrak{M}(N)$ more complexity than $\mathfrak{M}(N|\mathcal{T}(V))$. Even with homogenous trains, the binary variables ($\xi_{u,v}^{r,r^+}$, $\xi_{u,v,r}^{AD}$, $\xi_{u,v,r}^{DA}$ and $\varsigma_{u,r}^{i}$ [1]), which reflect the priority of trains with same direction at station and segment, still need be identified because of the unknown initial departure times of trains.

With our proposed algorithm IDC_CP, however, we can see in Table 5 that the computation time is reduced by over a thousand times (from an average of 1.363 hours down to 3.653 sec). The quality of the solutions is compared to the optimal solutions, with an average optimality gap $\bar{\varepsilon}$ of only 0.0018. The average value of three indicators (n^c , $\sum \ell_u^r$ and $\max\{\ell_u^r\}$) is 3.0, 3.0 and 1.0, respectively. It indicates that the solutions obtained by the IDC_CP have similar structure as the optimums, and proves the effectiveness of compaction-distribution based in IDC_CP.

[1] See constraints (II-4) ~ (II-8) in Appendix II.A.

6.1.2 The influence of balance constraints

As been shown in constraints (7), another important characteristic of model $\mathfrak{M}(N)$ is to keep the relative balance between train flows in different directions. Table 6 presents the results of $\mathfrak{M}(N)$ under different balance-degrees δ: 0.2, 0.4, … , 0.8. The computational time of CPLEX MIP algorithm is restricted within 24 hours.

When the balance constraints are added, a distinct difference compared to those in Table 5 (without balance constraints) is that the computational time to reach the optimal solution is much higher. For instance, for the case of δ=0.2, the optimal solution in six examples is not obtained within 24h, and the average computational time for other four examples also reaches 20.54h (see Table 6). Though the added balance constraints reduces the feasible region of model $\mathfrak{M}(N)$, it results in large difficulty of pruning and bounding, and increases the computational complexity of the decision tree.

Algorithm IDC_CP still has good performance when balance constraint is considered in model $\mathfrak{M}(N)$. The results in Table 6 show the solutions obtained by algorithm IDC_CP are very close to the best solutions obtained by the branch-and-bound. For instance, for the cases of δ=0.2, the optimality gap ε between the IDC_CP and the branch-and-bound is about 0.0203. When δ=0.8, the optimality gap is only 0.0035. With the gradual relaxation of balance constraints, algorithm IDC_CP can obtain the solution with better quality. In terms with computational efficiency, the average computational time is only about 9.498s when the IDC_CP is adopted. Obviously, compared with the branch-and-bound, algorithm IDC_CP can be applied to large scale cases in the real world. Algorithm IDC_CP is tested in the part of the Qing-Zang single-track corridor, which has the length of 830km and links 13 stations. The numerical results (Table 7) show that, the feasible solution by the branch-and-bound is not obtained when the number of train pairs exceeds five. The computational time required by the IDC_CP is between 12.89s and 94.87s; while the optimal gap is between 0.0063 and 0.0118.

Table 5: Results of train scheduling problem with fixed and flexible departure time

Inst.	Fixed departure time					Flexible departure time											
	Branch-and-bound					Branch-and-bound						IDC_CP					
	obj	CPU /h	n^c	$\sum \ell_u^r$	$\max\{\ell_u^r\}$	obj	CPU /h	n^c	$\sum \ell_u^r$	$\max\{\ell_u^r\}$	deviation	obj	CPU /s	n^c	$\sum \ell_u^r$	$\max\{\ell_u^r\}$	gap
1	827	0.209	12	63.0	19.5	723	0.391	3	3.0	1.0	0.1438	723	4.072	3	3.0	1.0	0.0000
2	831	0.049	12	81.0	29.5	715	1.839	3	3.0	1.0	0.1622	715	8.444	3	3.0	1.0	0.0000
3	847	0.079	12	58.0	17.5	737	0.593	3	5.0	3.0	0.1493	739	2.218	3	3.0	1.0	0.0027
4	815	0.048	12	60.0	20.5	707	1.594	3	3.0	1.0	0.1528	707	1.299	3	3.0	1.0	0.0000
5	833	0.481	12	61.0	17.5	727	1.067	3	3.0	1.0	0.1458	731	1.932	3	3.0	1.0	0.0055
6	844	0.111	12	67.0	26.5	731	0.269	3	3.0	1.0	0.1546	731	2.023	3	3.0	1.0	0.0000
7	851	0.307	10	64.0	24.5	739	0.471	3	3.0	1.0	0.1516	739	4.013	3	3.0	1.0	0.0000
8	828	0.090	11	68.0	19.5	727	0.425	3	3.0	1.0	0.1389	730.8	9.462	3	3.0	1.0	0.0052
9	817	0.009	11	55.0	12.5	712	5.370	3	4.0	2.0	0.1475	715	0.717	3	3.0	1.0	0.0042
10	855	0.040	12	78.0	28.5	739	1.613	3	3.0	1.0	0.1570	739	2.350	3	3.0	1.0	0.0000
Average		0.142h	11.6	65.5	21.6		1.363h	3	3.3	1.3	0.1504		3.653s	3.0	3.0	1.0	0.0018

Table 6: Performance results of train balance scheduling under different balance degrees.

Balance Degree	δ=0.2						δ=0.4						δ=0.6						δ=0.8					
Inst.	Branch-and-bound			IDC_CP			Branch-and-bound			IDC_CP			Branch-and-bound			IDC_CP			Branch-and-bound			IDC_CP		
	Obj^{up}	gap	CPU/h	Obj	CPU/s	$\mathcal{E}$	Obj^{up}	gap	CPU/h	Obj	CPU/s	$\mathcal{E}$	Obj^{up}	gap	CPU/h	Obj	CPU/s	$\mathcal{E}$	Obj^{up}	gap	CPU/h	Obj	CPU/s	$\mathcal{E}$
1	750.8	0.000	4.05	766.0	4.36	0.020	741.6	0.000	9.76	750.0	8.23	0.011	735.0	0.000	5.198	735.0	7.65	0.000	730.4	0.000	5.38	732.0	5.72	0.002
2	742.8	0.000	14.15	758.0	11.52	0.020	733.6	0.000	11.35	742.0	11.78	0.011	727.0	0.120	24.0	727.0	14.28	0.000	722.4	0.000	6.69	726.0	3.34	0.005
3	766.0	0.133	24.0	773.7	61.14	0.010	754.4	0.000	6.48	757.7	69.01	0.004	743.0	0.000	5.02	747.0	30.98	0.005	743	0.000	5.02	747.0	36.87	0.005
4	737.0	0.000	22.5	764.2	1.17	0.037	734.0	0.000	7.78	748.2	1.24	0.019	726.0	0.114	24.0	733.4	1.29	0.010	714.4	0.000	11.20	717.5	1.23	0.004
5	761.2	0.000	20.7	778.0	0.92	0.022	750.0	0.000	15.9	762.0	0.99	0.016	744.0	0.098	24.0	746.0	0.94	0.003	731.5	0.000	5.36	731.5	0.82	0.000
6	762.0	0.037	24.0	776.8	0.31	0.019	753.2	0.000	12.3	760.7	0.28	0.010	750.0	0.096	24.0	757.4	0.32	0.010	738.4	0.000	4.18	741.4	0.30	0.004
7	770.0	0.132	24.0	797.2	11.89	0.035	763.6	0.123	24.0	781.2	3.87	0.023	758.0	0.096	24.0	764.4	4.23	0.008	746.4	0.000	7.39	748.4	1.85	0.003
8	761.0	0.106	24.0	767.4	1.40	0.008	758.0	0.135	24.0	758.0	1.42	0.000	744.0	0.071	24.0	745.0	0.67	0.001	731.5	0.000	5.31	740.0	0.67	0.012
9	742.0	0.137	24.0	750.0	0.69	0.011	734.0	0.000	5.90	734.0	0.68	0.000	729.0	0.114	24.0	731.0	0.68	0.003	717.5	0.108	24.0	717.5	0.068	0.000

10	770.0	0.138	24.0	786.0	1.58	0.021	761.2	0.000	2.01	770.0	1.57	0.012	758.0	0.000	16.3	758.0	1.66	0.000	746.4	0.000	13.96	746.4	1.56	0.000
Aver.		0.068	20.54h		9.498s	0.0203		0.026	11.95h		9.907s	0.0106		0.0709	19.45h		6.27s	0.0040		0.0108	8.85h		5.243s	0.0035

* $\varepsilon = (Obj^{IDC\_CD} - Obj^{up}) / Obj^{up}$;

Table 7: Numerical results when the IDC_CP is applied into the part of Qing-Zang single-track railway corridor

Balance Degree	δ=0.2					δ=0.4					δ=0.6					δ=0.8				
Train pairs.	Branch-and-bound		IDC_CP		ε	Branch-and-bound		IDC_CP		ε	Branch-and-bound		IDC_CP		ε	Branch-and-bound		IDC_CP		ε
	Obj/min	CPU/h	Obj/min	CPU/s		Obj/min	CPU/h	Obj/min	CPU/s		Obj/min	CPU/h	Obj/min	CPU/s		Obj/min	CPU/h	Obj/min	CPU/s	
3	2540.0	24	2576.0	12.89	0.0142	2537.5	24	2560.6	28.82	0.0091	2527.4	24	2545.2	25.47	0.0071	2513.8	24	2529.8	27.49	0.0064
4	3431.1	24	3457.1	32.76	0.0076	3410.8	24	3432.3	42.86	0.0063	3389.5	24	3407.5	53.65	0.0053	3366.7	24	3382.7	60.86	0.0048
5	4314.2	24	4373.7	74.63	0.0138	4294.9	24	4341.3	82.58	0.0108	4266.1	24	4301.8	86.68	0.0084	4235.6	24	4267.6	94.76	0.0076
Aver.					0.0118					0.0087					0.0069					0.0063

6.2 Two-way balance traffic loading capacity evaluation

The two-way balance traffic loading capacity proposed in this paper depends not only on the topological structure of single-track railway corridor, but also on the different tolerance levels and balance degrees. Intuitively, the set of tolerance level and balance degree restrain the allowed maximal number of train-pairs passing through the single-track railway system.

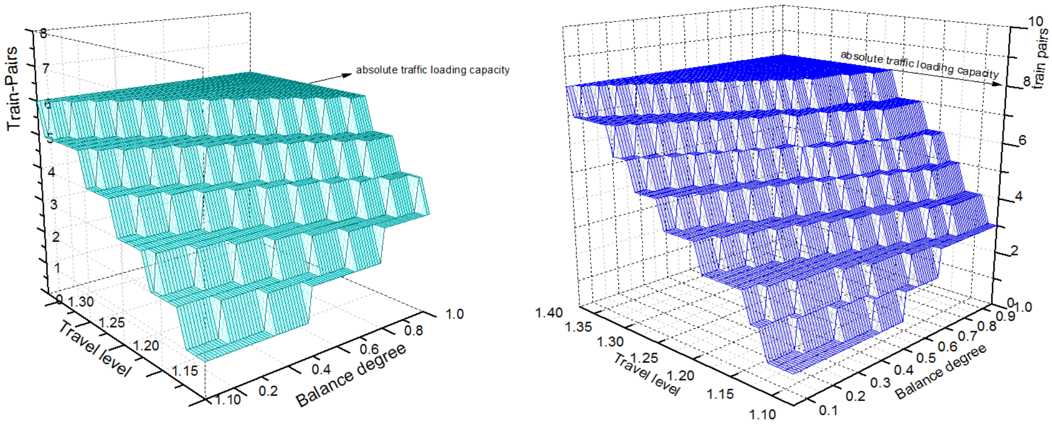

Figure 3 Three-dimension graphical depictions of two-way balance traffic loading capacity under different travels and balance degrees (a) track number at stations is 3 (b) track number at stations is 4

We take the first randomly generated instance in Table 4 to illustrate the influence of tolerance levels and balance degrees on the two-way balance traffic loading capacity. Figure 3 presents a three-dimensional depiction of the achieved traffic loading capacity under different tolerance levels and balance degrees. The two horizontal axes denote the tolerance level and balance degree, respectively, and the vertical axis is the maximal number of train-pairs that can be scheduled to travel in the system. With increasing tolerance level and balance degree, the top of the two-way traffic loading capacity keeps at 6 train-pairs for the 3-track case (Figure 3 (a)). This top value is decided by the topology structure, i.e., the absolute two-way traffic loading capacity. It is influenced by the number of tracks (or sidings) of stations, and does not depend on the tolerance levels and balance degrees. For example, when track number in stations is set to 4, the absolute two-way traffic loading capacity increases to 8 train-pairs (Figure 3 (b)).

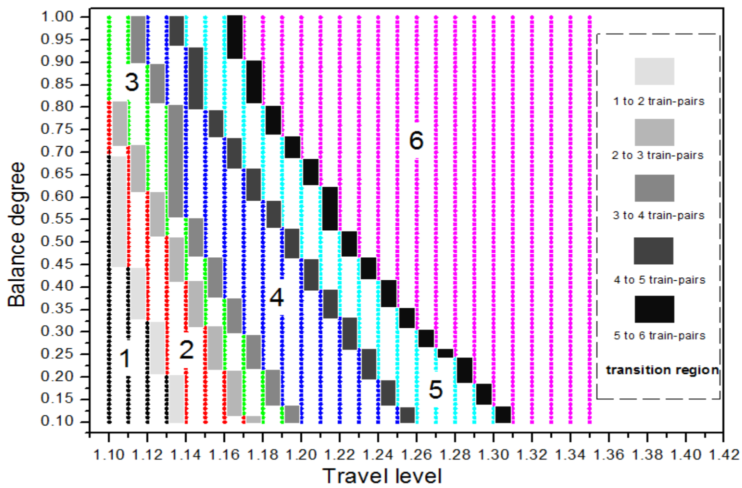

Figure 4 Transition description of two-way balance traffic loading capacity under different delay tolerance levels and balance degrees

Figure 4 presents the cross-section of three-dimension graph in Figure 3 (a). The results are divided into six regions, and the Arabic numerals denote the number of train-pairs in each region which satisfy the tolerance level and balance degree constraints. The results show that with more relaxed tolerance levels and less balanced train flows in both directions, the more train-pairs can be scheduled to the system and greater system capacity.

The results also show that, the capacity is restrained when the tolerance level is lower than 1.17. However, when tolerance level exceeds 1.30, the two-way loading capacity is not influenced by balance degree and tolerance level and reaches the absolute top value. Figure 4 also presents the transition regions (marked in different shades of grey) in capacity gains. For instance, when balance degree is kept at 0.10, the transition region of tolerance level is between 1.13 and 1.14, in which the loading capacity varies from 1 train-pair to 2 train-pairs. Other transition regions are also distributed at (1.17, 1.18), (1.19, 1.20), (1.25, 1.26), and (1.30, 1.31). These results can explore the relation between travel delay of train and capacity loss, and provide decision support for railway administrator dealing with train rescheduling under disturbance or disruption scenarios.

Figure 5 further presents the average travel time of each train under different tolerance levels and balance degrees. The black grid surface represents the travel time front which is the allowed average travel time of train under the different tolerance levels, and the complicate zigzag structure below the front surface indicates the actual average travel time of train. It can be visually found that, with increasing the tolerance level and balance degree, the average travel time of trains gradually reduce. The complicated zigzag structures are developed with the variation of the tolerance level and balance degree.

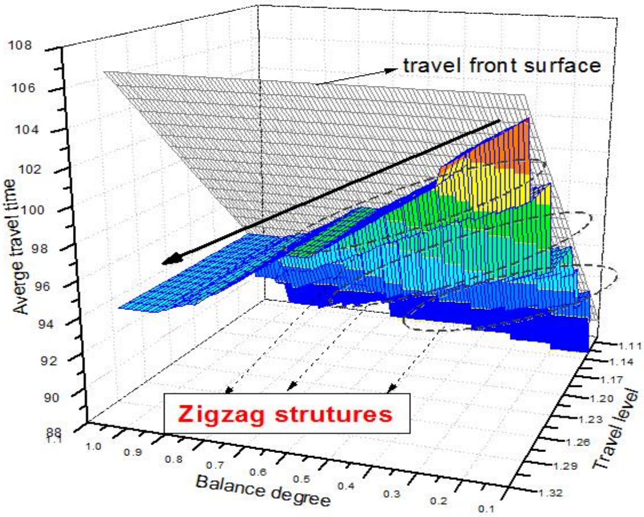

Figure 5 the average travel time of train under different travels and balance degrees

The zigzag structures in the actual travel time is further depicted and explained by the results in Figure 6. Two black dashed lines are travel fronts corresponding with two tolerance levels $\lambda = 1.12$ and $\lambda = 1.29$. There are four phases are emerging for the case of $\lambda = 1.12$. Only one train-pair is allowed to run when balance degree is between 0.1 and 0.32. With the relaxation of balance degree, the average travel time of trains is gradually reduced. The transition occurs when balance degree loads the region between 0.32 and 0.33, in which the allowed number of train-pairs increases from one to two. Near the transition region, the average travel time of train is close to the travel front. Thus, the zigzag profiles are developed with a further relaxation of balance degree. However, for the case of $\lambda = 1.29$, the absolute capacity is reached in the second phase. And hence, only a zigzag structure is developed.

The information presented in Figure 5 and Figure 6 can be used to identify explicitly the difference between the actual travel time and travel front, and they provide an intuitive decision support for railway administrator to consider the trade-off between travel time of trains and relative balance of outbound and inbound train flows.

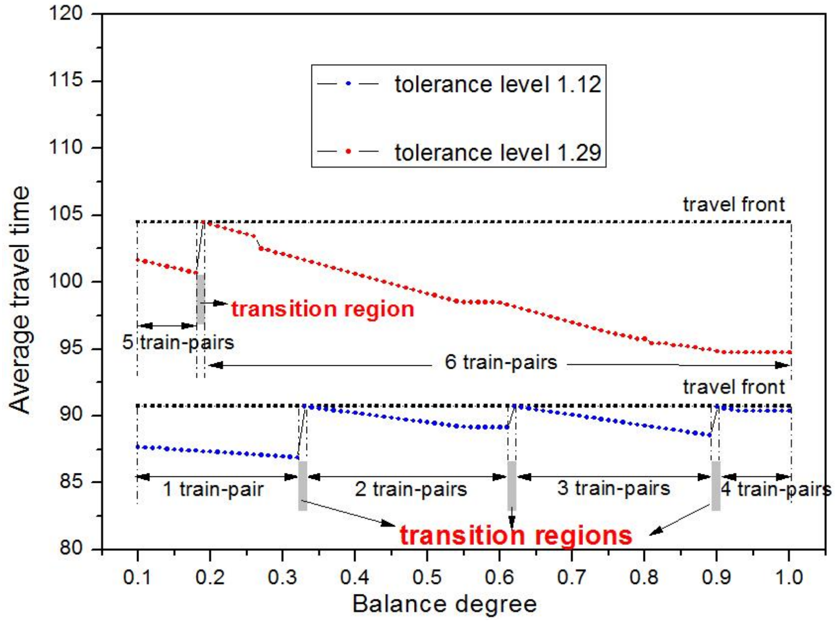

Figure 6 the average travel time of trains under different balance degrees

7. Conclusions

This paper addresses the issues of capacity evaluation of single-track railway corridor from the perspective of the railway administrators. A sophisticated 0-1 mixed-integer programming is formulated to obtain the maximum number of trains which can be scheduled along a single-track railway corridor subject to two constraints the administrators regularly face: the travel tolerance level and the relative balance between the two-way traffic loads. The initial departure times of the scheduled trains are allowed to vary within a specific time window to ensure the two constraints are met. A dichotomization based solution framework is proposed, which iteratively solve the initial departure time of the scheduled trains and adjust the number of trains that can be scheduled.

The proposed solution framework relies upon solving a train scheduling problem with initial departure time decisions. A method based on the concept of compact distribution (IDC_CP) is developed to solve the optimal departure times of trains from original stations. We show that the solutions based on the IDC_CP method are comparable (with an optimality gap within 2%) to those based on traditional branch-and-bound method and solved using the standard CPLEX solver. Most significantly, however, our proposed IDC-CP solver is more efficient: a problem for case of δ=0.6 taking 19.45 hours to solve using the traditional method is solved by ICD_CP method in just 6.27 seconds, with an optimality gap of 0.4%. The efficiency of the ICD-CP solver allows our proposed capacity evaluation method to be applied not only as a planning tool, but also during operations to maximize a single-track system capacity.

We apply the proposed method to investigate the two-way traffic loading capacity of single-track railway corridor under the different travel tolerance levels and different balance degrees. We show that, with increasing tolerance level and balance degree, the two-way capacity tends to a top value (the absolute capacity), which is decided by the topology structure of railway system. We can identify explicitly the transition regions of traffic loading capacity, and average travel time of trains under different tolerance levels and balance degrees. These results can explore the relation between travel delay of train and capacity loss.

The proposed method provides an efficient and subjective framework for capacity evaluation and initial departure-time rescheduling of a single-track railway system. We have assumed so far that all scheduled trains traverse along the corridor without interruptions. An important and natural extension of our research is to consider disruption (planned or un-planned), so as to provide a practical tool to the railway administrators to identify quantitatively the loss of capacity in the event of disruption.

Acknowledgement

This work was financially supported by the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (No. 71621001), National Natural Science Foundation of China (No. 71571018 and 71401008), the Fundamental Research Funds for the Central Universities (No. 2016JBZ007 and 2017JBM074) and the State key Laboratory of Rail Traffic Control and Safety (No. RCS2017ZT001), and the UK Rail Safety and Standard Board Project RSSB-T1071 (DITTO).

Reference

Abril, M., Barber, F., Ingolotti, L., Salido, M.A., Tormos, P., Lova, A., 2008. An assessment of railway capacity. Transportation Research Part E **44** (5), 774-806.

Assad, A. A., 1980. Models for rail transportation. Transportation Research Part A **14** (3), 205-220.

Bevrani, B., Burdett, R.L., Yarlagadda, P., 2015. A case study of the Iranian national railway network and its absolute capacity expansion using analytical models. Transport, 1-17, doi:10.3846/16484142.2015.1099053.

Burdett, R.L., 2015a. Incorporating complex train paths in an analysis of absolute capacity. International Journal of Railway Technology **4** (4), 73-83.

Burdett, R.L., 2015b. Multi-objective models and techniques for railway capacity analysis. European Journal of Operational Research **245**, 489-505.

Burdett, R.L., 2016. Optimization models for expanding a railways theoretical capacity. European Journal of Operational Research **251** (3), 783-797.

Burdett, R.L., Kozan, E., 2006. Techniques for absolute capacity determination in railways. Transportation Research Part B **40** (8), 616-632.

Burdett, R.L., Kozan, E., 2009. Techniques for inserting additional trains into existing timetables. Transportation Research B **43** (8), 821-836.

Burdett, R.L., Kozan, E., 2009a. Scheduling trains on parallel lines with crossover points. Journal of Intelligent Transportation Systems: Technology, Planning, and Operations **13** (5), 171-187.

Burdett, R.L., Kozan, E., 2009b. Techniques for restricting multiple overtaking conflicts and performing compound moves when train scheduling. Mathematical and Computer Modelling **50** (1-2), 314-328.

Burdett, R.L., Kozan, E., 2014a. Performance profiling for predictive train schedules. Journal of Rail Transport Planning and Management **4**, 98-114.

Burdett, R.L., Kozan, E., 2014b. Determining operations affected by delay in predictive train timetables. Computers and Operations Research **41**, 150-166.

Cai, X., Goh, C. J., Mees, A. I., 1998. Greedy heuristics for rapid scheduling of trains on a single track. IIE Trans. **30** (5), 481-493.

CS-I, 2007. National rail freight infrastructure capacity and investment study. Association of American Railroads (AAR), Cambridge systematics.

Carey, M., 1994. Extending a train pathing model from one-way to two-way track. Transportation Research Part B **28** (5), 395-400.

Caprara, A., Fischetti, M., Toth, P., 2002. Modeling and solving the train timetabling problem. Operations Research **50** (5), 851-861.

De Kort, A.F., Heidergott, B., Ayhan, H., 2003. A probabilistic approach for determining railway infrastructure capacity. European Journal of Operational Research **148** (3), 644-661.

Dessouky, M., Leachman, R.C., 1995. A simulation modeling methodology for analyzing large complex rail networks. Simulation **65** (2),131-142.

ERRAC, 2004. A comparison of member state public research programmes with the ERRAC SRRA 2020. European Rail Research Advisory Council (ERRAC), Brussels.

Frank, O., 1966. Two-way traffic in a single line of railway. Operations Research **14** (5), 801-811.

Higgins, A., Kozan, E., Ferreira, L., 1996. Optimal scheduling of trains on a single line track. Transportation Research Part B **30** (2), 147-161.
Higgins, A., Kozan, E., Ferreira, L., 1997. Heuristic techniques for single line train scheduling. Journal of Heuristics **3** (1), 43-62.
Kaas, A.H., 1991. Strategic Capacity Analysis of Networks: Developing and Practical use of Capacity Model for Railway Networks, ScanRail Consult, Technical University of Denmark.
Kozan, E., Burdett, R.L., 2005. A railway capacity determination model and rail access charging methodologies. Transportation Planning and Technology **28** (1), 27-45.
Krueger, H., 1999. Parametric modeling in rail capacity planning. Proceedings of the 1999 Winter simulation conference, 1194-2000.
Lai, Y-C., Barkan, C., 2009. Enhanced parametric railway capacity evaluation tool. Journal of Transportation Research Board 2117, 33–40.
Landex, A., 2008. Methods to estimate railway capacity and passenger delays. Ph. D. thesis, Technical University of Denmark Department of Transport, Denmark.
Landex, A., Kaas, A.H., Schittenhelm, B., Schneider-Tilli, J., 2006. Practical use of the UIC 406 capacity leaflet by including timetable tools in the investigations. Computers in Railways X **88**, 643-652.
Li, F., Gao, Z.Y., Li, K.P., Yang, L.X., 2008. Efficient scheduling of railway traffic based on global information of train. Transportation Research Part B **42** (10), 1008-1030.
Li, F., Sheu J-B., Gao, Z.Y., 2014. Deadlock analysis, prevention and train optimal travel mechanism in single-track railway system. Transportation Research Part B **68**, 385-414.
Lindner, T., 2011. Applicability of the analytical UIC Code 406 compression method for evaluating line and station capacity. Journal of Rail Transport Plan and Management **1**(1), 49-57.
Liu, R., Whiteing, A.E., Koh, A., 2013. Challenging established rules for train control through a fault tolerance approach: Applications at a Classic Railway Junction, Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit **227** (6), 685-692.
Melody K., Preston, J.M., 2010. Revising the UIC 406 method: revenue generating capacity. In: Joint Rail Conference (JRC) 2010, Urbana, IL, USA.
Mu, S., Dessouky, M., 2011. Scheduling freight trains traveling on complex networks, Transportation Research Part B **45** (7), 1103-1123.
Mu, S., Dessouky, M., 2013. Efficient dispatching rules on double tracks with heterogeneous train traffic. Transportation Research Part B **51**, 45-64.
Mussone, L., Calvo, R.W., 2013. An analytical approach to calculate the capacity of a railway system. European Journal of Operational Research **228** (1): 11-23.
Petersen, E.R., 1974. Over-the-road transit time for a single track railway. Transportation Science **8** (1), 65-74.
Petersen, E.R., Taylor, A.J., 1982. A structured model for rail line simulation and optimization. Transportation Science **16** (2), 192-206.
Prinz, J.H., 2005. Implementation of UIC 406 capacity calculation method at Austrian Railway (OBB). Austrian Railway (OBB), Vienna.
Tolliver, D., 2010. Railroad planning & design. CE 456/656 Lecture Notes, North Dakota State University.
UIC, 2004. UIC CODE 406 R-capacity, 1st edn. International Union of Railways (UIC), Paris.
Wahlborg, M., 2005. Application of the UIC capacity leaflet at Banverket. In: 1st International Seminar on Railway Operations Modelling and Analysis, I.A. Hansen, et al., Editors. Delft, Netherlands.
Welch, N., Gussow, J., 1986. Expansion of Canadian National Railway's line capacity. Interfaces **16** (1), 51-64.
Xu, X.M., Li, K.P., Yang, L.X., 2015. Scheduling heterogeneous train traffic on double tracks with efficient dispatching rules. Transportation Research Part B **78**, 364-384.
Yokota, H., 1980. Performance analysis of passing track and planning principles for its layout on single-track line,

Quarterly Reports **21** (3), 105-114.
Zhou, X.S., Zhong, M., 2007. Single-track train timetabling with guaranteed optimality: Branch-and-bound algorithms with enhanced lower bounds. Transportation Research Part B **41** (2), 320-341.

Appendix I: Symbol Descriptions

1. Symbols description in the model:

1.1 Index and Set

j: Train type index.

J: The set of train types, and $J = \{1, 2, ..., |J|\}$, $|J|$ is the number of train types.

u, v: Train index.

V^O: Set of all outbound trains, and $|V^O|$ is the number of outbound trains.

V^I: Set of all inbound trains, and $|V^I|$ is the number of inbound trains.

V: Set of all trains, where $V = V^O \bigcup V^I$.

r, r^+: Station index.

$\tilde{r}_u$: The origin (i.e. station) of train u.

$\overline{r}_u$: The destination (i.e. station) of train u.

R_u: The stations visited by train u.

$I_{u,r}$: Feasible tracks set of train u at station r ($r \in R_u$).

1.2 Parameter

δ: The balance degree.

λ: The tolerance level.

γ_j: The proportion of *j*-type train in the loaded train set.

$\mathfrak{D}_N^{\max}$: The maximal average deviation between the outbound and inbound train flows.

f_j: The free travel time of j type train in the single-track railway corridor.

T: The time window, where $T = \min(f_j \mid j \in J)$.

$\beta_{u,j}$: 0-1 parameter, if train *u* is of type *j*, then it is 1, otherwise 0.

p_u^{r,r^+}: The free running time for train u on the segment between station r and its next station r^+ (r、$r^+ \in R_u$).

h^{aa}, h^{dd}, h^{da} and h^{ad}: The time headway between two trains at a station travelling in the same direction, where the superscripts represents the status of the trains as respectively: arrival-arrival, departure-departure, departure-arrival, and arrival-departure.

g^{aa}, g^{dd}, g^{da} and g^{ad}: The time headway between two trains at a station travelling in opposite directions; the superscripts represent the same as above.

τ_b: The traversing time of train at station.

$\tau^a_{u,e}$ ($\tau^d_{u,e}$): The time required by train u when acceleration from a station (or deceleration to stop at a station).

M: A large number.

1.3 Decision Variable

N: The number of train-pairs loading the single-track railway corridor.

n_j: The number of *j*-type train loading the single-track railway corridor.

$t^a_{u,r} / t^d_{u,r}$: The arrival/departure time of train u at station r.

$\bar{f}_{out}$: The average travel time of outbound train flows.

$\bar{f}_{in}$: The average travel time of inbound train flows.

$\xi^{r,r^+}_{u,v}$: 0-1 binary variable. If train u has prior to occupy the segment between station r and station r^+ than train v,, then $\xi^{r,r^+}_{u,v} = 1$, otherwise $\xi^{r,r^+}_{u,v} = 0$.

$\xi^{AD}_{u,v,r}$ ($\xi^{DA}_{u,v,r}$): 0-1 binary variable. If train u arrives at (departs from) station r before train v departs from (arrives at) station r, then $\xi^{AD}_{u,v,r} = 1$ ($\xi^{DA}_{u,v,r} = 1$), otherwise $\xi^{AD}_{u,v,r} = 0$ ($\xi^{DA}_{u,v,r} = 0$).

$\zeta^{r,r^+}_{u,v}$: same as $\xi^{r,r^+}_{u,v}$, but for trains travelling in opposite direction.

$\zeta^{AA}_{u,v,r}$ ($\zeta^{DD}_{u,v,r}$): 0-1 binary variable. If train u arrives (departs) earlier at station r than train v, then $\xi^{AA}_{u,v,r} = 1$ ($\xi^{DD}_{u,v,r} = 1$), otherwise $\xi^{AA}_{u,v,r} = 0$ ($\xi^{DD}_{u,v,r} = 0$).

$\zeta^{DA}_{u,v,r}$ ($\zeta^{AD}_{u,v,r}$): 0-1 binary variable. If train u departs from (arrives at) station r before train v arrives at (departs from) station r, then $\zeta^{DA}_{u,v,r} = 1$ ($\zeta^{AD}_{u,v,r} = 1$), otherwise $\zeta^{DA}_{u,v,r} = 0$ ($\zeta^{AD}_{u,v,r} = 0$).

$\varsigma^i_{u,r}$: 0-1 binary variable. If train u occupies track i at station r, then $\varsigma^i_{u,r} = 1$, otherwise $\varsigma^i_{u,r} = 0$.

ϑ^r_u: 0-1 binary variable. If train u stops at station r, then $\vartheta^r_u = 1$, otherwise $\vartheta^r_u = 0$.

μ_{OI}: 0-1 binary variable. If all outbound trains travel through the railway system without any delay, then

$\mu_{OI} = 1$; otherwise, if all inbound trains are not delayed, then $\mu_{OI} = 0$.

2. Symbols description in the algorithm:

$\tilde{n}_{lb}$: The low bound of the number of train-pairs passing through the single-track railway corridor.

$\tilde{n}_{ub}$: The up bound of the number of train-pairs passing through the single-track railway corridor.

$t^{a}_{u,r}$: The arrival time of train u at station r in a given schedule plan.

$t^{d}_{u,r}$: The departure time of train u from station r in a given schedule plan.

$\mathbb{S}(V)$: The set that records the arrival and departure times of each train at stations, i.e., $\mathbb{S}(V) = \{(t^{a}_{u,r}, t^{d}_{u,r}) \mid u \in V, r \in R_u\}$.

$t^{d}_{v,\tilde{r}_u}$: The initial departure time of train u at its original station $\tilde{r}_u$ in a known schedule plan.

$\mathcal{T}(V)$: The set that records the initial departure time of each train, i.e., $\mathcal{T}(V) = \{t^{d}_{u,\tilde{r}_u} \mid u \in V\}$.

$\mathcal{S}$: The solution of model $\mathfrak{M}(N_i)$, which can be expressed by $\mathcal{S}=\{\mathcal{T}(V), \mathbb{S}(V)\}$.

$\mathfrak{D}$: The set that records the travel information of trains at stations in a known schedule plan, and is expressed by $\mathfrak{D} = \{\mathcal{D}_u \mid u \in V\}$, and $\mathcal{D}_u = \{\mathcal{D}^r_u \mid r \in R_u\}$.

$\mathcal{D}^r_u$: The information set which includes the arrival and departure time of train u at station r, and the meet-crossing or overtaking between train u and other trains. And it is expressed by $\mathcal{D}^r_u = ((t^{a}_{u,r}, t^{d}_{u,r}), \mathfrak{X}^r_u(\mathfrak{V}^r_u), \mathfrak{F}^r_u(\mathfrak{V}^r_u))$.

$\mathfrak{V}^r_u$: The set that records the ID of trains that meet train u at station r, i.e., $\mathfrak{V}^r_u = \{\lambda_v \mid v \in V, \text{ and } t^{a}_{u,r} < t^{a}_{v,r} < t^{d}_{u,r} \text{ or } t^{a}_{u,r} < t^{d}_{v,r} < t^{d}_{u,r}\}$. And λ_v is the ID of train v.

$\mathfrak{X}^r_u(\mathfrak{V}^r_u)$: The set that records the arrival and departure characteristic of trains in set $\mathfrak{V}^r_u$, and is expressed by $\mathfrak{X}^r_u(\mathfrak{V}^r_u) = \{\mathcal{X}^r_v \mid v \in \mathfrak{V}^r_u\}$. If train v ($v \in \mathfrak{V}^r_u$) is an arrival train at station r, then $\mathcal{X}^r_v = 1$; otherwise if it is a departure train, then $\mathcal{X}^r_v = 0$.

$\mathfrak{F}^r_u(\mathfrak{V}^r_u)$: The set that records the time-points distribution of trains in set $\mathfrak{V}^r_u$ at station r, and is expressed by $\mathfrak{F}^r_u(\mathfrak{V}^r_u) = \{p^r_v \mid v \in \mathfrak{V}^r_u\}$. If $\mathcal{X}^r_v = 1$ ($v \in \mathfrak{V}^r_u$), then $p^r_v = t^{a}_{v,r}$; otherwise if $\mathcal{X}^r_v = 0$ ($v \in \mathfrak{V}^r_u$), then $p^r_v = t^{d}_{v,r}$.

Γ_u^r: The time-points distribution resulted by the arrival and departure of train u at station r, and it is expressed as $\Gamma_u^r=\{t^a_{u,r},t^d_{u,r}\}\bigcup \mathfrak{F}_u^r(\mathcal{V}_u^r)$.

$\tilde{\Gamma}_u^r$: The compaction pattern corresponding with Γ_u^r.

ℓ_u^r: The compressible time distance, which reflects the difference between the original time-points distribution and its compaction pattern. And it can be deduced by $\ell_u^r=(t^d_{u,r}-t^a_{u,r})-(\tilde{t}^d_{u,r}-\tilde{t}^a_{u,r})$.

$\mathfrak{L}$: The set that records the travel information of trains at segments in a obtained schedule plan, and is expressed by $\mathfrak{L}=\{\mathfrak{L}_u \mid u\in V\}$, and $\mathfrak{L}_u=\{\delta_u^{r,r^+}\mid r,r^+\in R_u\}$; δ_u^{r,r^+} is the delay time of train u at the segment between station r and r^+, i.e., $\delta_u^{r,r^+}=(t^a_{u,r^+}-t^d_{u,r})-(p_u^{r,r^+}+\vartheta_u^r\cdot\tau^a_{u,e}+\vartheta_u^{r^+}\cdot\tau^d_{u,e})$.

$\mathfrak{K}_r(\mathcal{T},\Gamma_u^r,\tilde{\Gamma}_u^r)$: A mapping function, which determine new initial departure of trains based on Γ_u^r and $\tilde{\Gamma}_u^r$, i.e., $\mathcal{T}'=\mathfrak{K}_r(\mathcal{T},\Gamma_u^r,\tilde{\Gamma}_u^r)$.

$\mathfrak{K}_s(\mathcal{T},\delta_u^{r,r^+})$: A mapping function, which determine new initial departure of trains based on the delay of train u at the segment between station r and r^+, i.e., $\mathcal{T}'=\mathfrak{K}_s(\mathcal{T},\delta_u^{r,r^+})$.

Appendix II:

II.A The formulation of train balance scheduling problem with initial departure choice ($\mathfrak{M}(N)$)

$$\text{Minimize} \sum_{u|u\in V}(t^a_{u,\bar{r}_u}-t^d_{u,\tilde{r}_u}) \tag{II-1}$$

Subject to:

δ*-balance constraints:*

$$\Big|\sum_{u\in V^O}(t^a_{u,\bar{r}_u}-t^d_{u,\tilde{r}_u})-\sum_{v\in V^I}(t^a_{v,\bar{r}_v}-t^d_{v,\tilde{r}_v})\Big| \le \delta\cdot\mathfrak{D}_N^{\max}\cdot N \tag{II-2}$$

Departure time choice constraints:

$$0\le t^d_{u,\tilde{r}_u}\le T \qquad \forall\ u\in V^O\bigcup V^I \tag{II-3}$$

Departure-Departure and Arrival-Arrival headway constraints between the trains with same direction:

$$t^d_{u,r}+h^{dd}\le t^d_{v,r}+(1-\xi^{r,r^+}_{u,v})\cdot M \qquad \forall\ u,v\in V^O \text{ or } u,v\in V^I,\ u\neq v,\ r,r^+\in R_u\bigcap R_v \tag{II-4a}$$

$$t^a_{u,r^+}+h^{aa}\le t^a_{v,r^+}+(1-\xi^{r,r^+}_{u,v})\cdot M \qquad \forall\ u,v\in V^O \text{ or } u,v\in V^I,\ u\neq v,\ r,r^+\in R_u\bigcap R_v \tag{II-4b}$$

$$\xi^{r,r^+}_{u,v}+\xi^{r,r^+}_{v,u}=1 \qquad \forall\ u,v\in V^O \text{ or } u,v\in V^I,\ u\neq v,\ r,r^+\in R_u\bigcap R_v \tag{II-4c}$$

Arrival-Departure and Departure-Arrival headway constraints between the trains with same direction:

$$t^a_{u,r} + h^{ad} \le t^d_{v,r} + (1-\xi^{AD}_{u,v,r})\cdot M \qquad \forall\ u,v \in V^O \text{ or } u,v \in V^I;\ u \ne v;\ r \in R_u \cap R_v;\ r \ne \tilde{r}_u, \tilde{r}_v, \bar{r}_u, \bar{r}_v \qquad \text{(II-5a)}$$

$$\xi^{AD}_{u,v,r} + \xi^{AD}_{v,u,r} = 1 \qquad \forall\ u,v \in V^O \text{ or } u,v \in V^I;\ u \ne v;\ r \in R_u \cap R_v;\ r \ne \tilde{r}_u, \tilde{r}_v, \bar{r}_u, \bar{r}_v \qquad \text{(II-5b)}$$

$$t^d_{u,r} + h^{da} \le t^a_{v,r} + (1-\xi^{DA}_{u,v,r})\cdot M \qquad \forall\ u,v \in V^O \text{ or } u,v \in V^I;\ u \ne v;\ r \in R_u \cap R_v;\ r \ne \tilde{r}_u, \tilde{r}_v, \bar{r}_u, \bar{r}_v \qquad \text{(II-5c)}$$

$$\xi^{DA}_{u,v,r} + \xi^{DA}_{v,u,r} = 1 \qquad \forall\ u,v \in V^O \text{ or } u,v \in V^I;\ u \ne v;\ r \in R_u \cap R_v;\ r \ne \tilde{r}_u, \tilde{r}_v, \bar{r}_u, \bar{r}_v \qquad \text{(II-5d)}$$

Meeting-crossing constraints between trains with opposite direction:

$$t^a_{u,r^+} + g^{ad} \le t^d_{v,r^+} + (1-\zeta^{r,r^+}_{u,v})M \qquad \forall\ u \in V^O, v \in V^I \text{ or } u \in V^I, v \in V^O;\ r, r^+ \in R_u \cap R_v;\ r \ne \bar{r}_u \qquad \text{(II-6a)}$$

$$\zeta^{r,r^+}_{u,v} + \zeta^{r,r^+}_{v,u} = 1 \qquad \forall\ u \in V^O, v \in V^I \text{ or } u \in V^I, v \in V^O;\ r, r^+ \in R_u \cap R_v;\ r \ne \bar{r}_u \qquad \text{(II-6b)}$$

Departure-Departure and Arrival-Arrival headway constraints between the trains with opposite direction:

$$t^a_{u,r} + g^{aa} \le t^a_{v,r} + (1-\zeta^{AA}_{u,v,r})\cdot M \qquad \forall\ u \in V^O, v \in V^I \text{ or } u \in V^O, v \in V^I;\ r \in R_u \cap R_v,\ r \ne \tilde{r}_u, \tilde{r}_v, \bar{r}_u, \bar{r}_v \qquad \text{(II-7a)}$$

$$\zeta^{AA}_{u,v,r} + \zeta^{AA}_{v,u,r} = 1 \qquad \forall\ u \in V^O, v \in V^I \text{ or } u \in V^I, v \in V^O;\ r \in R_u \cap R_v;\ r \ne \tilde{r}_u, \tilde{r}_v, \bar{r}_u, \bar{r}_v \qquad \text{(II-7b)}$$

$$t^d_{u,r} + g^{dd} \le t^d_{v,r} + (1-\zeta^{DD}_{u,v,r})\cdot M \qquad \forall\ u \in V^O, v \in V^I \text{ or } u \in V^I, v \in V^O;\ r \in R_u \cap R_v;\ r \ne \tilde{r}_u, \tilde{r}_v, \bar{r}_u, \bar{r}_v \qquad \text{(II-7c)}$$

$$\zeta^{DD}_{u,v,r} + \zeta^{DD}_{v,u,r} = 1 \qquad \forall\ u \in V^O, v \in V^I \text{ or } u \in V^I, v \in V^O;\ r \in R_u \cap R_v;\ r \ne \tilde{r}_u, \tilde{r}_v, \bar{r}_u, \bar{r}_v \qquad \text{(II-7d)}$$

Station capacity constraints:

$$\sum_{i|i \in I_{ur}} \varsigma^i_{u,r} = 1 \qquad \forall\ u \in V;\ r \in R_u \qquad \text{(II-8a)}$$

$$t^d_{u,r} + h^{da} \le t^a_{v,r} + (1-\xi^{DA}_{u,v,r})M + (1-\varsigma^i_{u,r})M + (1-\varsigma^i_{v,r})M$$

$$\forall\ u,v \in V^O \text{ or } u,v \in V^I;\ u \ne v;\ r \in R_u \cap R_v;\ i \in I_{ur} \cap I_{vr} \qquad \text{(II-8b)}$$

$$t^d_{u,r} + g^{da} \le t^a_{v,r} + (1-\zeta^{DA}_{u,v,r})M + (1-\varsigma^i_{u,r})M + (1-\varsigma^i_{v,r})M$$

$$\forall\ u \in V^O, v \in V^I \text{ or } u \in V^O, v \in V^I;\ r \in R_u \cap R_v;\ i \in I_{ur} \cap I_{vr} \qquad \text{(II-8c)}$$

Segment running time constraints：

$$t^d_{u,r} + p_u^{r,r^+} + \vartheta_u^r \cdot \tau^a_{u,e} + \vartheta_u^{r^+} \cdot \tau^d_{u,e} \le t^a_{u,r^+} \qquad \forall\ u \in V;\ r, r^+ \in R_u \qquad \text{(II-9)}$$

Stopping/non-stopping constraints:

$$t^a_{u,r} + \tau_b - t^d_{u,r} \le M \cdot \vartheta_u^r \qquad \forall\ u \in V;\ r \in R_u \qquad \text{(II-10a)}$$

$$t^d_{u,r} - t^a_{u,r} - \tau_b \le M \cdot \vartheta_u^r \qquad \forall\ u \in V;\ r \in R_u \qquad \text{(II-10b)}$$

$$t^a_{u,r} + \tau_b - t^d_{u,r} < M \cdot (1-\vartheta_u^r) \qquad \forall\ u \in V;\ r \in R_u \qquad \text{(II-10c)}$$

Binary variables:

$$\xi_{u,v}^{r,r^+}, \xi_{u,v,r}^{AD}, \xi_{u,v,r}^{DA}, \zeta_{u,v}^{r,r^+}, \zeta_{u,v,r}^{AA}, \zeta_{u,v,r}^{DD}, \varsigma_{u,r}^{i}, \vartheta_{u}^{r} \in \{0,1\} \tag{II-11}$$

The model' purpose is to minimize the total travel times of all trains loaded in the single-track railway corridor. Constraint (II-2) denotes that the travel average deviation between out- and in-bound train flows is confined to a certain range ($\delta \cdot \mathcal{D}_N^{\max}$).Constraints (II-3) ensure that all trains must depart from their original stations at a given time windows, and their initial departure time is free.

Moving block signal system has been widely discussed in the railway operation. The block is defined in real time by computers as safe zones around each train. Moving block allows trains to run closer together, while maintaining required safety margins. Constraints (II-4) emphasize the *Departure-Departure headway* h^{dd} *and Arrival-Arrival headway* h^{aa} when two trains with same direction depart from and arrive at the same station. The binary variable $\xi_{u,v}^{r,r^+}$ describes the priority of train u and v depart from station r and arrive at station r^+, which also is the priority of train u and v occupy the segment between station r and r^+. Specially, if $r = \tilde{r}_u$ and $r = \tilde{r}_v$, constraints (II-4a) also reflect the departure order of two trains from the same original station.

In the single-track railway system, when a train is entering into the station and the other train with the same direction is ready to depart from the same station, a safety time interval must be guaranteed so that station dispatchers have enough time to switch signals to arrange routes for different trains. Constraints (II-5a) and (II-5b) ensure that the *Arrival-Departure headway* h^{ad} is satisfied between the arrival and departure trains with the same direction. The binary variable $\xi_{u,v,r}^{AD}$ presents the arrival and departure priority of train u and v at station r. Similarly, the *Departure-Arrival* headway h^{da} is ensured by constraints (II-5c) and (II-5d). It should be pointed out, theses headways are not considered at the original and destination stations. In this paper, the original and destination stations are assumed to be the yard stations. Different to the intermediate stations, the yard stations have sufficient track number and signal equipment, and may pull in and out trains at the same time. When a train arrives at a destination station, it is moved from railway system immediately. A train may departure from the original station when its departure time is satisfied and no trains with opposite direction travel on its next segment.

Constraints (II-6) specify the meet-crossing behavior between two trains in opposite directions, which is a distinct characteristic of single-track railway system. If two trains in opposite directions need to occupy the same segment at the same time, one train must wait at station so that the other train can meet and cross. The binary variable $\zeta_{u,v}^{r,r^+}$ is introduced to describe the priority of train u and v for the segment between station r and r^+. Similar to constraints (II-4), constraints (II-7) ensure the safety headway when two trains with opposite directions arrive at and depart from the same station. The binary variable $\zeta_{u,v,r}^{AA}$ and $\zeta_{u,v,r}^{DD}$ describe the arrival and departure priority of train u and v at station r, respectively. And parameters g^{aa} and g^{dd} denote the Arrival-Arrival and Departure-Departure headway between the trains in opposite directions, respectively.

Constraints (II-8) focus on the finite track number in the stations. Typically, the station capacity is related to the number of tracks or platforms at station. In this paper, it is assumed that one track (or one siding) in a station only provides service for at most one train. And hence, at any time, the number of trains dwelling on the station cannot exceed the number of tracks. We adopt the track choices of trains at stations to reflect the finite station capacity. Binary variable $\varsigma_{u,r}^{i}$ represents whether train u select the track i in station r ($r \in R_u$). If it is true, then $\varsigma_{u,r}^{i} = 1$, otherwise $\varsigma_{u,r}^{i} = 0$. Constraints (II-8a) state that one train can only hold one track in a station. If two trains select the same track in a station, one train can only arrive at a station after the other train has departed from the station, and the Departure-Arrival headway between them is guaranteed. Clearly, constraints (II-8b) and (II-8c) ensure that one track in station can only provide service for at most one train at a time, and moreover guarantee that the number of trains at station does not exceed the station capacity at any moment.

Additionally, constraints (II-9) link the entering and leaving times of each train on a segment. Parameter p_u^{r,r^+} is the free running time of train u at the segment between station r and r^+. If the train stops at station r or r^+, two extra time loss $\tau_{u,e}^{a}$ and $\tau_{u,e}^{d}$ are taken into account due to the acceleration of train departing from station and deceleration of train arriving at station, respectively. Here, the binary variable ϑ_u^r is introduced to reflect whether train u stop at station r, and its identification is presented by constraints (II-10). Obviously, if $\vartheta_u^r = 0$, constraints (10) ensure $t_{u,r}^{a} + \tau_b = t_{u,r}^{d}$; otherwise, $t_{u,r}^{a} + \tau_b < t_{u,r}^{d}$. Note that parameter τ_b is the basic running time of train at station. Finally, constraints (II-11) model the binary characteristic of the variables.

II.B The model formulation and solution method for identifying the maximal average deviation parameter ($\mathfrak{D}_N^{\max}$) between in- and out-bound train flows

Model formulation:

The model for identifying parameter $\mathfrak{D}_N^{\max}$ is described as follows. Firstly, the objective of the model is to minimize the total travel time of the loaded trains in the single-track railway system (Eq. (II-1)). Constraints (II-3)-(II-11) are included to ensure that the travel paths of trains satisfy the characteristic of single-track railway system. A class of specific constraints, which are expressed by (II-12a) and (II-12b), are required for ensuring that either outbound or inbound trains is free flow. The binary variable μ_{OI} is introduced to identify whether outbound or inbound train flow is free. If $\mu_{OI} = 1$, constraints (II-11a) indicate that the travels of all outbound trains are free; while constraints (II-12b) are redundant. If $\mu_{OI} = 0$, constraints (II-12b) ensure inbound trains are free flow.

$$\mu_{OI} \cdot \sum_{j \in J} (t_{u,\bar{r}_u}^{a} - t_{u,\tilde{r}_u}^{d}) \cdot \beta_{u,j} \le f_j \qquad \forall \ u \in V^O \qquad \text{(II-12a)}$$

$$(1 - \mu_{OI}) \cdot \sum_{j \in J} (t_{u,\bar{r}_u}^{a} - t_{u,\tilde{r}_u}^{d}) \cdot \beta_{u,j} \le f_j \qquad \forall \ u \in V^I \qquad \text{(II-12b)}$$

$$\mu_{OI} \in \{0,1\} \qquad \text{(II-13)}$$

Based on the departure and arrival time of in- and out- bound trains at their original and destination stations, the value of the maximal average deviation between in- and out- bound train flows is easily calculated by Eq. (II-14).

$$\mathcal{D}_N^{\max} = \left| \sum_{u\in V^O} (t^a_{u,\bar{r}_u} - t^d_{u,\tilde{r}_u}) - \sum_{u\in V^I} (t^a_{u,\bar{r}_u} - t^d_{u,\tilde{r}_u}) \right| / N \qquad \text{(II-14)}$$

Solution method:

The above model indicates that the value of parameter $\mathcal{D}_N^{\max}$ is related to the number of loaded train-pairs and dispatch rule of in- and out- bound trains. We adopts a simple scheduling rule to estimate the value of $\mathcal{D}_N^{\max}$. Assume the number of the loaded train-pairs is N, and the out-bound train flow is free. The simple rule is described in Table AII-1.

Table AII-1: A simple rule for calculating the value of $\mathcal{D}_N^{\max}$

Step 1:	Select a random time, and adopt the successive departure pattern to schedule the free outbound train flow.
Step 2:	According to the arrival time of outbound train flow at their first station, determine the initial time of the first inbound train v_1, which is regarded as the left bound of time windows. Moreover, the right bound of time windows is also decided, i.e., $t^d_{v_1,\tilde{r}_{v_1}} + \min(f_j \mid j \in J)$.
Step 3:	Based on the track number at the intermediate stations, schedule gradually all inbound trains.
Step 4:	According to the obtained schedule plan, the value of $\mathcal{D}_N^{\max}$ is calculated.

The "successive departure pattern" in Step 1 is that all outbound trains or inbound trains depart sequentially from the same original station, and their departure time interval from the origin is the Departure-Departure headway (h^{dd}). For the case of heterogonous trains, the train with higher speed has priority to depart from the original station for avoiding the delay of trains resulted by the overtaking behavior. In step 3, the number of inbound trains allowed to successive depart is decided by the track number in the intermediate station. Additionally, the departure times of inbound trains are also constrained by time windows.

We adopt a simple example to illustrate the above method for calculating $\mathcal{D}_N^{\max}$, which is depicted in Figure AII-1. The track number of the intermediate stations is set to 3, and the number of the loaded train-pairs is 4.

Firstly, outbound trains (u_1, u_2, u_3 and u_4) are freely scheduled in the single-track railway system based on the successive departure pattern. According to the arrival time of train u_1 at its first station (R_1) and idle track number of the station, the departure time of the first inbound train (v_1) can be deduced. The initial departure time of train v_1 is set to the left bound of time windows, and moreover the whole time windows ([0, T]) can be developed. All inbound trains must depart from the original station in this time window.

According to the idle track number of station R_1 (p.s., a track of the station has been occupied by outbound train flows), the trajectories of two inbound trains with successive departure pattern can be determined based on the arrival-arrival and arrival-departure headway (h^{aa} and g^{ad}). Similarly, other inbound trains can be scheduled in the single-track railway system based on the idle track number of next station R_2 .

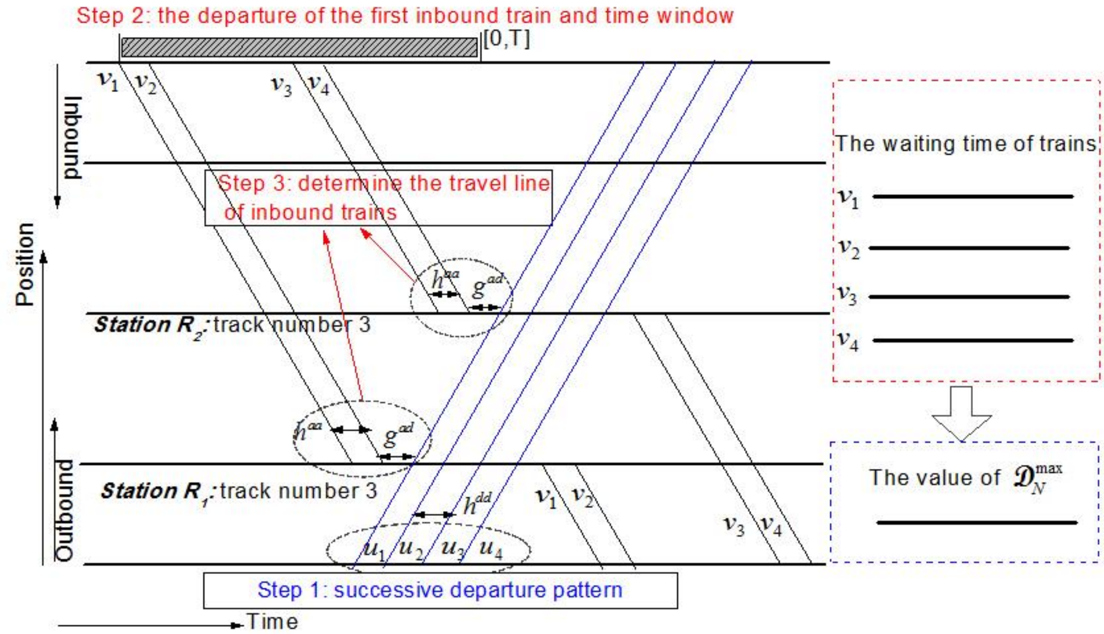

Figure AII-1 An sample for calculating the value of $\mathcal{D}_N^{\max}$