



**UNIVERSITY OF LEEDS**

This is a repository copy of *Integrated Capacity Allocation and Timetable Coordination for Multimodal Railway Networks*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/212894/>

Version: Accepted Version

---

**Article:**

Chai, S., Yin, J., Tang, T. et al. (3 more authors) (2024) Integrated Capacity Allocation and Timetable Coordination for Multimodal Railway Networks. *Transportation Research Part C: Emerging Technologies*, 165. 104681. ISSN 0968-090X

<https://doi.org/10.1016/j.trc.2024.104681>

---

© 2024, Elsevier. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>. This is an author produced version of an item published in *Transportation Research Part C: Emerging Technologies*. Uploaded in accordance with the publisher's self-archiving policy.

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# Integrated Capacity Allocation and Timetable Coordination for Multimodal Railway Networks

Simin Chai<sup>1</sup>, Jiateng Yin<sup>1,2\*</sup>, Tao Tang<sup>1</sup>, Lixing Yang<sup>2</sup>, Ronghui Liu<sup>3</sup>, Qin Luo<sup>4</sup>

1. *State Key Laboratory of Advanced Rail Autonomous Operation, Beijing Jiaotong University, Beijing, China*

2. *School of Systems Science, Beijing Jiaotong University, Beijing, China*

3. *Institute for Transport Studies, University of Leeds, Leeds, U.K.*

4. *College of Urban Transportation and Logistics, Shenzhen Technology University, Shenzhen, China*

## Abstract

Modern transportation systems are moving towards shared mobility with diverse demands, and resource aggregation and coordination among different transport modes have become more and more significant. In this study, we investigate the integrated train capacity allocation and timetable coordination for multimodal railway networks to release the congestion of transfer hubs between metro and mainline rail networks. We consider that a part of vehicles can be dynamically allocated to some metro trains to increase their capacity to take more passengers at transfer hubs alighting from mainline rail trains. We formulate this problem into a mixed-integer linear programming (MILP) model, which simultaneously generates the coordinated timetables of both metro and mainline rail trains, as well as the train capacity allocation strategy in the network. The objectives are to minimize the passenger travel time, passenger transfer time at the hubs and the operational costs for rail managers. To tackle computational challenges in real-world instances, we develop an exact branch-and-cut solution algorithm to generate (near-)optimal solutions more efficiently. In our algorithm, we propose five sets of valid inequalities that are dynamically added to the model to strengthen the linear relaxation bounds at each node. We also design a customized branching strategy in the search tree by imposing branching on the key decision variables regarding the train sequences at transfer stations. Real-world case studies based on the operational data of a realistic multimodal railway network in Beijing are conducted to verify the effectiveness of our approach. The results demonstrate that our branch-and-cut-based approach outperforms commercial solvers regarding solution quality and computational efficiency. Compared to the current non-coordinated train timetable in practice, our approach by flexibly allocation train capacities can reduce the passenger transfer waiting time by over 40%.

**Keywords:** Multimodal railway network, Capacity allocation, Train timetable coordination, Passenger demand

## 1 Introduction

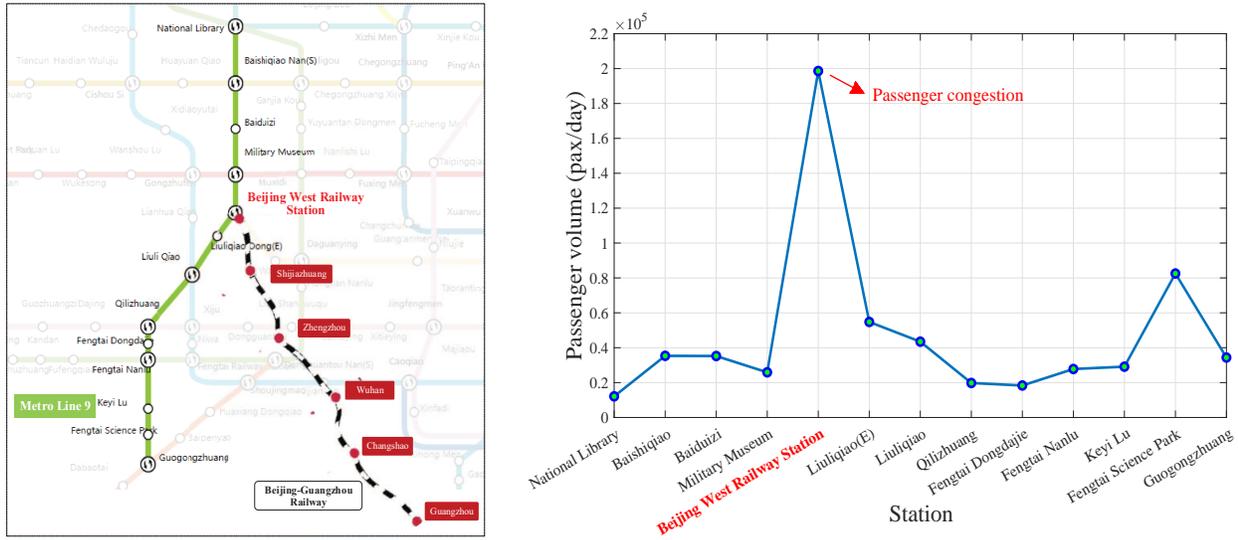
In an era of globalization, characterized by a surge in the demands for both human and freight mobility, swift and efficient transportation service is increasingly crucial. Among the various evolving transportation services, multimodal transportation, which stands out as an integrated solution catering to diverse needs, has been promoted throughout the world and attracted significant attention in recent years (OECD, 2020; Wang et al., 2023c). For

---

\*E-mail address: jtyin@bjtu.edu.cn (J. Yin)

example, the Shift2Rail Joint Undertaking launched Shift2MaaS to create an EU-wide multimodal travel experience for all European citizens by connecting rail with other transport modes (Shift2Rail, 2020). Siemens Mobility, a division of Siemens, has recognized multimodal transportation as a core business imperative for developing seamless travel experiences for passengers in the next two decades (Siemens, 2023).

In public transportation systems, metro stations, as comprehensive passenger transport hubs of multiple transportation modes (such as railways, buses, and aviation), play a prominent role in fostering regional economic development, tourism, and cultural exchanges. With the increase in passenger demand, these metro stations always suffer from significant transportation pressures due to the substantial influx of transfer passengers from external transportation systems and the daily commuting needs of nearby residents. Each day, tens of thousands of passengers gather, disperse, or transfer at these hubs for social and economic activities. For example, over two hundred thousand passengers utilize trains and metros at the prominent transportation hub, Beijing West Railway Station for their leisure and business needs, as illustrated in Figure 1. To enhance the operational efficiency of the urban and regional transportation systems, there is an urgent requirement for well-integrated transport stations that facilitate convenient travel and smooth transfer. A highly effective method to realize this objective involves the collaborative optimization of schedules across various transportation modes, especially at these stations.



(a) Layout of Line No.9 and Beijing-Guangzhou Railway Line

(b) Passenger volume at stations in Line No.9 of Beijing metro

Figure 1: Illustration of passenger congestion at transfer hubs

Moreover, the excessive passenger demand at metro stations takes challenges as a large number of passengers often cannot board trains promptly due to the insufficient existing train loading capacity. This results in severe congestion and operational risks at these stations. To alleviate the congestion, the concept of virtual coupling (VC), which enables multiple vehicles to be virtually coupled into a platoon with very short following distance, has been promoted globally and garnered significant interest in recent years (Quaglietta et al., 2020; Chai et al., 2023a). In the context of virtual coupling, multiple vehicles can be rapidly coupled into a platoon with very short following distances, while a platoon comprising multiple vehicles can also be decoupled to separated vehicles. Different compositions of train platoons offer the flexibility to modify the maximum number of passengers to which a train can cater. In this sense, virtual coupling can potentially mitigate station congestion to improve service quality for passengers while saving operational costs from the rail managers' perspective in metro systems. Therefore, virtual coupling

has become a hot research topic recently. Nevertheless, we notice that most of these studies are devoted to the conceptual frameworks (Di Meo et al., 2019; Quaglietta et al., 2020) or control methodologies for virtual coupled trains (Liu et al., 2021b). Meanwhile, the existing literature hardly addresses the scheduling of trains with different compositions within multimodal networks at the tactical planning level.

Therefore, this paper focuses on optimizing the scheduling of trains with different loading capacities in a multimodal railway network, including mainline railways and metros. In the network, mainline railways cater to long-distance passengers traveling between cities, while metros not only bridge regional areas with external railway systems but also serve the daily traveling needs of nearby residents, as shown in Figure 2. We aim to design a coordinated timetable for railway and metro trains to minimize passenger traveling times and enhance transfer efficiency. Meanwhile, with the high flexibility of virtual coupled trains, we consider employing trains with multiple vehicles or carriages with larger loading capacities for busy metro stations to alleviate congestion while scheduling trains with only a single group of carriages along less crowded metro lines to reduce operational costs. To this end, we formulate a rigorous mathematical model with the objective of minimizing passenger travel time and operational costs to optimize the coordinated timetable and train capacity allocation strategies simultaneously. By analyzing the mathematical properties of the proposed model, we develop an exact branch-and-cut algorithm, consisting of several sets of valid inequalities and a series of customized branching rules, to obtain high-quality solutions efficiently. We also conduct real-world instances to quantitatively reveal the benefits of future multimodal rail mobility and cast marginal insights for rail managers.

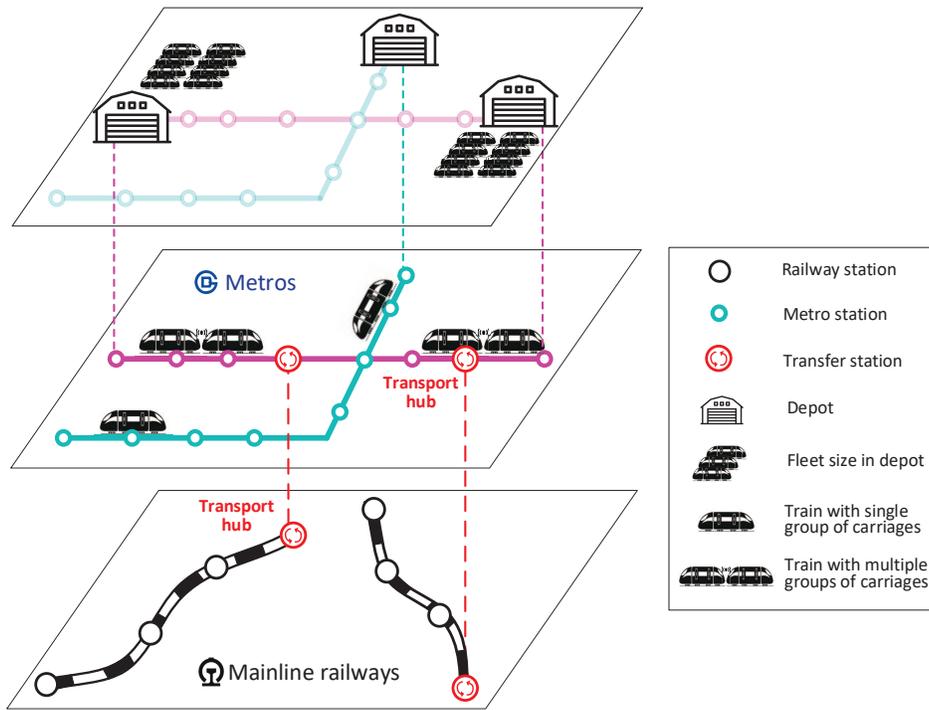


Figure 2: Illustration of VC trains for multimodal railway transportation with transport hubs

## 1.1 Literature review

The train scheduling problem, which aims to generate an optimized timetable that is to be carried out by a limited train fleet, is a very active and significant research field in railway traffic management over the past years.

In particular, most recent studies in this field focus on the demand-oriented train schedule optimization problems, for both single-line and network.

In single-line cases, the passenger demand-oriented train scheduling problem aims to optimize the service frequency and arrival/departure times, in order to adapt to the time-varying demand (Wang et al., 2018; Lu et al., 2023). For example, Niu and Zhou (2013) derived a time-variant demand-oriented timetabling approach and formulated a nonlinear 0-1 integer model, solved by a genetic algorithm to minimize the total number of waiting passengers and weighted remaining passengers. Barrena et al. (2014) addressed non-periodic timetable design with time-variant passenger demands in a metro corridor and developed two nonlinear mathematical formulations to minimize the passenger waiting time at the stations. Due to the computational intensity of these two models, they proposed a fast adaptive large neighborhood search (ALNS) algorithm to obtain a good solution.

In the case of a network with multiple lines, the demand-oriented train scheduling problem becomes much more complex due to the large volume of transfer passengers. There has been a growing focus on train timetable coordination in urban railway networks to enhance transfer efficiency and provide seamless service for passengers. For instance, Wu et al. (2015) developed a timetable synchronization optimization model to minimize the transfer waiting time in a metro network during peak hours. Kang et al. (2016) explored the first train scheduling synchronization problems, aiming to reduce the passenger’s long waiting time when transferring from one metro line to another. Guo et al. (2020) proposed an MIP model to optimize the last train schedule to maximize the successful transfer events and shorten the big difference between the last departure times. *Considering a metro network where different lines share physical tracks and/or platforms, Liu et al. (2023b) integrated train timetables, passenger flows, and train speed profiles to dynamically adjust train schedules online, aiming to optimize the passenger satisfaction and operational costs. They developed a bi-level model predictive control approach by considering rolling stock availability and passenger demands at the higher level, while designing timetables and train speed profiles at the lower level. In their follow-up work Liu et al. (2023a), they further proposed a scenario-based distributed model predictive control approach to handle uncertain passenger flows in urban rail transit networks considering transfer passengers between different lines. Note that the above literature only focuses on passenger-oriented train scheduling within a single railway transportation mode, without considering the arrival patterns of transfer passengers between multimodal railway networks.*

Different from the above studies on specific operation periods (e.g., coordination of first or last trains), we observe a growing number of studies exploring network-level train coordination problems to improve transfer conveniences through the whole operation horizon (Wang et al., 2020; Yin et al., 2023). For example, Yin et al. (2021) developed an MILP model for the coordination of train timetables among different metro lines to minimize the crowdedness of transfer stations. A decomposition-based ALNS algorithm is proposed to enhance the computational efficiency of large-scale instances. Yuan et al. (2023) integrated the train timetable coordination and skip-stop plan optimization for metro networks to improve service quality. A decomposition and approximate dynamic programming approach is designed to convert the original network-level problem into a series of small-scale subproblems for each metro line.

Meanwhile, a few studies also addressed the coordination of trains in a multi-mode transportation framework, and enhancing multimodal transfer efficiency has become a significant area of interest. Liu et al. (2021a) and Gkiotsalitis et al. (2023a) reviewed the existing studies on public transport transfer coordination and synchronization in recent years at the tactical planning phase and the real-time control phase respectively. In their follow-up work (Gkiotsalitis et al., 2023b), a mixed integer quadratic program with a convex objective function was proposed to coordinate the arrival times of trains on different urban rail lines that serve stations along a joint corridor, in order to maintain an even headway among trains and reduce passenger transfer times. Long et al. (2020) formulated a bi-objective mixed-integer linear programming model for the scheduling coordination of the last metro train and high-speed railway

trains. [Huang et al. \(2021\)](#) proposed three integer programming models progressively to optimize the coordination of the last metro train schedule with high-speed railway and aviation system. The estimation of transfer time between different modes is the most important factor affecting the coordination quality. [Wang et al. \(2023c\)](#) focused on an urban passenger transport hub and utilizing multimodal passenger simulation. They proposed a bi-objective integer nonlinear programming model (INLP) incorporating train schedules and flexible routing plans in metro systems to minimize passenger waiting time and operating costs simultaneously. [Ning et al. \(2023\)](#) investigate the integrated optimization of last train timetabling in urban rail transit lines and bridging service (taxis and buses) design with consideration of passenger path choices. Using pre-constructed path sets, they proposed a bi-objective mixed-integer nonlinear programming (MINLP) model to minimize total passenger travel time and total passenger travel cost. [Ke et al. \(2024\)](#) formulated a multi-commodity flow model in a time-space network to address the intermodal timetabling problem for trains and flight services considering two different transfer modes: rapid walking transfer and shuttle bus transfers. An Alternating Direction Method of Multipliers method is developed to decompose the problem into four routing subproblems (i.e., train routing, flight routing, shuttle routing, and passenger routing) to enhance the computation efficiency. As they considered unlimited train capacity, the effects of transfer passenger demands were not explicitly captured in this work.

As the emerging technology VC allows for more flexible train services through rapid coupling and decoupling, recent studies have emphasized that the advantage of VC is to dynamically adjust train capacity to meet fluctuating passenger demand. For example, considering vehicle capacity can be dynamically adjusted according to the arrival passengers at each station, [Shi et al. \(2020\)](#) and [Shi and Li \(2021\)](#) proposed an MILP model to optimize the timetable and vehicle capacities in an urban transit line, with the objective of minimizing passenger waiting time and vehicle operating costs. [Zhou et al. \(2022\)](#) developed an integrated approach to jointly optimize the train timetable and rolling stock circulation plan in a metro line. They addressed non-equilibrium and time-dependent passenger demand in this bidirectional line by considering a flexible train composition mode, which allows rolling stocks to change their compositions through (un)coupling operations according to the passenger demand. Furthermore, [Guo et al. \(2022\)](#) and [Shi et al. \(2022\)](#) presented a train capacity allocation strategy by flexibly reserving and releasing carriages in an overcrowded metro line to minimize the total passenger waiting time and passenger accumulation risk at all involved stations. Note that the above literature only considers a single urban transit line to address the train scheduling problem, and the train timetables among multiple transportation modes are ignored. Considering multiple connected lines in an urban rail transit network, [Chai et al. \(2023a\)](#) proposed an MILP formulation to optimize the schedule of flexibly coupled trains. However, the model was solved by the commercial solver CPLEX, which limits its implementation in real-world scale instances. In addition, the formulation only allows the train composition of a maximum of two vehicles.

## 1.2 Paper contributions

From the relevant literature in this field, we see that there has been a growing trend to investigate the management and operation of urban rail transit networks. Nevertheless, we notice that little research has been devoted to multimodal railway networks, e.g., metro, mainline railway networks, etc. Different from the existing studies, this paper aims to propose an integrated modeling approach for the capacity allocation and train coordination of multiple rail networks with the help of emerging VC strategies. A detailed comparison of our study with relevant literature is presented in Table 1. Our unique contributions are summarized as follows.

(1) We first propose the integrated capacity allocation and timetable coordination for multimodal railway networks, in which a part of vehicles can be dynamically allocated to different metro trains to alleviate the congestion of transfer hubs between metro and mainline rail networks. We develop a novel MILP formulation for this problem

in order to collaboratively optimize the timetables of both metro and rail networks as well as the train capacity allocation strategy. **The formulation integrates two types of network-dependent modeling approaches, that is the space-time network for metro systems and the event-activity network for railway systems, to tackle train scheduling problems in multimodal railway networks.** Our objective function is constructed by considering multiple factors, including passenger travel time, passenger transfer time at the hubs, and the operational costs for rail managers. Different from the state-of-art, our approach not only incorporates internal passenger demand in railway and metro systems but also addresses the coordination of these two modes by considering transfer passenger demand at transfer stations in transport hubs.

(2) As our models contain a few sets of binary variables to indicate the synchronization of multiple networks, which leads to computational difficulties with MIP solvers, we analyze the characteristics of our formulation and then develop an exact branch-and-cut-based solution algorithm to generate high-quality solutions more efficiently. In our approach, we propose and prove five sets of valid inequalities, which are dynamically added to the model to enhance linear programming (LP) bound at each branching node. We also design a tailored branching strategy by directing branching on critical binary variables (i.e., those determine the coordinated timetable of trains) within the search tree to obtain better feasible solutions more quickly.

(3) To verify the effectiveness of our approach, we conduct two sets of experiments involving a set of small-scale instances and a set of real-world large-scale instances based on the historical data of the railway-metro network in Beijing. We compare the computational results obtained by our approach with two benchmarks, i.e., the commercial MIP solver CPLEX and the practical timetable in Beijing. The experiments demonstrate that our branch-and-cut-based solution approach evidently outperforms CPLEX. Our experiments also find that our approach can reduce passenger transfer waiting time by over 40% compared with the current train timetable. Our results also reveal several meaningful insights for rail managers in practice.

The rest of this paper is organized as follows. Section 2 presents a detailed description of the studied problem. Section 3 formulates the problem into an MILP model. Section 4 proposes an exact branch-and-cut-based solution approach. Section 5 reports two case studies, i.e., a small case and a real-world case based on the detected data in a realistic railway-metro network. Section 6 gives our conclusions and future research directions.

## 2 Problem Statement

### 2.1 Description of multimodal railway networks

Our study considers a multimodal railway transportation network consisting of multiple interconnecting railway lines  $\mathcal{I}$  and metro lines  $\mathcal{J}$ . In railway systems, each line  $i \in \mathcal{I}$  is associated with a set of railway stations  $\mathcal{U}_i$ , numbered as  $\{1, 2, \dots, U_i\}$ . Passengers in railway systems typically book their tickets in advance using a seat reservation system to plan their journeys. Thus, in our problem, passengers with the same journey (i.e., sharing the same origin and destination) are consolidated into passenger group  $p \in \mathcal{P}_i$  on line  $i$ , and each passenger group must be served by a train. We denote the number of passengers in group  $p$  as  $N_p$ , and their planned origin and destination stations are denoted as  $o_p$  and  $d_p$ , respectively. In metro systems, each metro line  $j \in \mathcal{J}$  is associated with a set of metro stations  $\mathcal{S}_j$  numbered as  $\{1, 2, \dots, S_j\}$ , where stations at both ends of each metro line  $j$  is connected with a depot. The metro network involves a set of depots  $\mathcal{D}$  for rolling stock allocation. The maximum capacity of each depot  $d \in \mathcal{D}$  to allocate rolling stock is  $Z_d^{\max}$ . In addition, each depot  $d$  is connected to one or more stations, i.e., trains from depot  $d$  can serve multiple lines simultaneously. We define  $\mathcal{J}_d \subset \mathcal{J}$  as the set of lines that can be served by the rolling stock from depot  $d$ . Our problems consider to schedule a group of metro trains  $\mathcal{K}_j$  from a depot, passing through a sequence of stations  $\{1, 2, \dots, S_j\}$  and back to another depot on each metro

Table 1: Summary of recent relevant studies on train timetable coordination and train capacity allocation in comparison with our work

Publication	Infrastructure	Capacity allocation	Research problem	Objective	Model structure	Solution algorithm
Kang et al. (2016)	Subway network	No	The first train timetabling	Train arrival time differences and the number of missed trains	MIP	Local search
Yin et al. (2021)	Urban rail network	No	Timetable coordination	The congestion level of stations	MILP	ALNS
Wang et al. (2023b)	Metro network	No	The last train timetabling	The latest time for passengers	MILP	Improved GA
Yuan et al. (2023)	Metro network	No	Timetable coordination with skip-stop plan	The total passenger waiting time and station crowding	MINLP	Decomposition and ADP
Yin et al. (2023)	Urban rail network	No	Timetable synchronization	The total waiting time of passengers and synchronization quality indicator	MIP	ALNS
Ning et al. (2023)	Multimodal urban network	No	The last train timetabling and bridging service design	The total passenger travel time and total passenger travel cost	MINLP	Adaptive iterative algorithm
Zhou et al. (2022)	Single metro line	Yes	Train timetabling and capacity allocation	The operational costs and passenger waiting time	MILP	VNS
Pan et al. (2023)	Single metro line	Yes	Train timetabling and capacity allocation	The total involved costs	MILP	CG based heuristics
This paper	Multimodal rail network	Yes	Capacity allocation and timetable coordination and	The passenger travel time and operational costs	MILP	Branch-and-cut

<sup>1</sup> Symbols description in Table 1: mixed integer programming (MIP); genetic algorithm (GA); approximate dynamic programming (ADP); column generation (CG); Adaptive Large Neighborhood Search (ALNS); variable neighbourhood search (VNS)

line  $j$ . In this process, we need to consider that the rolling stock on each line  $j \in \mathcal{J}_d$  is required to cover the trains  $\mathcal{K}_j$  with limited resources in depot  $d$ .

In the multimodal railway network, a set of transfer stations  $\mathcal{V}_{ij}$  connects railway line  $i$  with metro line  $j$ , where passengers can transfer between railway and metro trains. In our problem, we consider a passenger from lines  $i$  being able to transfer from a railway train to a metro train on line  $j$  if the departure time of the metro train  $k \in \mathcal{K}_j$  is neither earlier than nor too late with respect to the passenger’s arrival time at the transfer station  $s \in \mathcal{V}_{ij}$ . This transfer indicates walking from one railway station to a metro station within the same transfer hub. When such a transfer is feasible, we say that the timetables of two trains coordinate. Specifically, two trains on different lines coordinate if the departure time of metro train  $k$  falls within a suitable time window with respect to the arrival time of railway train  $l$ . We refer to this time window as *coordination time window*, denoted as  $[\underline{L}_s^{ij}, \bar{L}_s^{ij}]$ , where  $\underline{L}_s^{ij}$  and  $\bar{L}_s^{ij}$  represent the minimum and maximum coordinated times at station  $s$ , respectively. **Note that the minimum coordinated time  $\underline{L}_s^{ij}$ , which denotes the minimum difference for a feasible transfer between the departure time of metro trains and the arrival time of railway trains, ensures that passengers can walk from one railway station on line  $i$  to a metro station on line  $j$ , implicitly representing the transfer walk time for passengers at transfer station  $s$  in our problem.** As shown in Figure 4, trains 1 and 4 on line  $i$  fail to coordinate with train  $k$  because train 1 arrives too early and train 4 arrives too late at the transfer station  $s$ . Conversely, railway trains 2 and 3 do coordinate with metro train  $k$  at station  $s$ . Our study aims to design train timetables in the multimodal rail network that facilitate as many smooth transfers as possible. To this end, we design train timetables for both railway and metro systems by slightly adjusting the departure and arrival times of trains in the network.

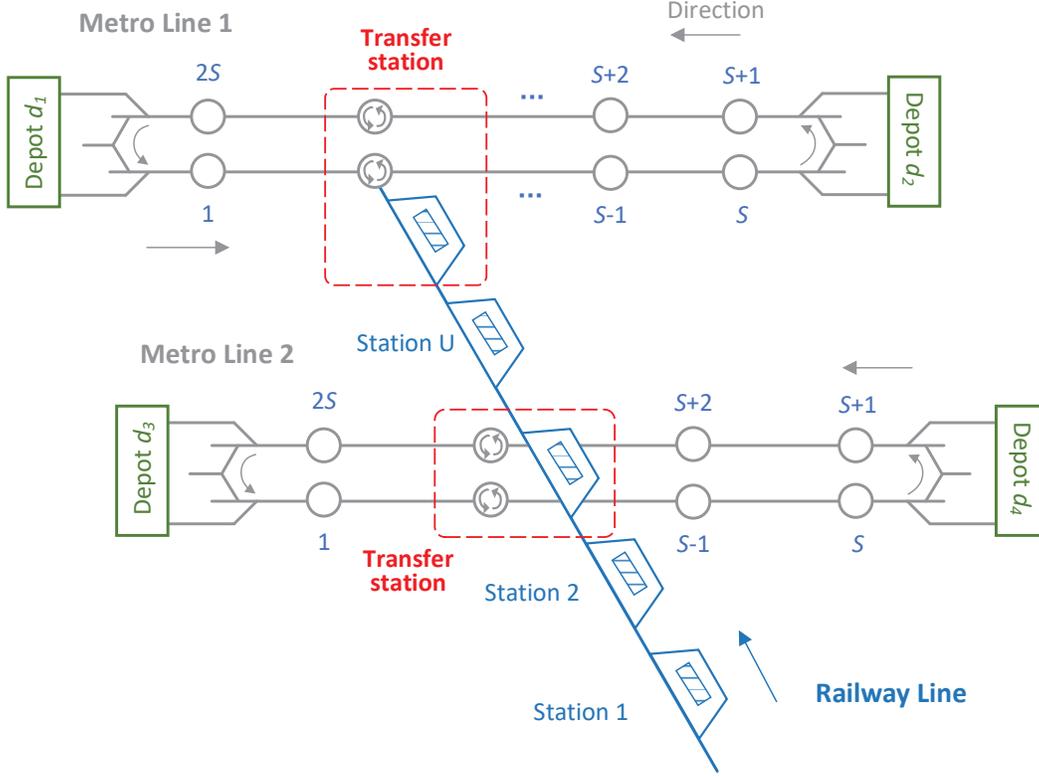


Figure 3: Illustration of network layout in our problem

In addition to considering transfer passengers from railway trains at transfer hubs, we also need to account for the arrival of passengers within the metro systems. In metro systems, it is widely recognized that the passenger demand is time-dependent with varying arrival rates at different time periods. The dynamic passenger demand is typically captured by using the passenger arrival rate for every station over time, which is a continuous function of time. To model this time-dependent passenger demand, we discretize the planning time horizon (e.g., from 7:00 am to 11:00 pm) into a set of timestamps  $\mathcal{T} = \{1, 2, \dots, T\}$ , where  $\tau$  is the index of timestamps. We assume that the passenger demand remains constant within each timestamp, similar to the treatment in [Huang et al. \(2020\)](#). Thus, we denote  $n_{s,\tau,j}^{\text{arr}}$  as the number of arrival passengers at station  $s$  and timestamp  $\tau$  on line  $j$ .

**Remark 2.1** In principle, all passenger demand for rail transportation networks is essentially origin-destination (OD) dependent. Nevertheless, the dynamics of passenger flow on the mainline railway are less evident than that of the metro system. In the mainline railways, passengers book their tickets and make their travel plans in advance according to the timetable proposed by the railway operators, so the number of passengers is almost predetermined by the sold tickets based on the published schedule. That is to say, the specific OD-based passenger demand can be captured accurately and treated as static constants (see e.g., [Yang et al. \(2016\)](#); [Zhang et al. \(2020\)](#)). In comparison, the passenger flow in metro systems has different characteristics, since passengers usually do not care about the train timetables before their trips, leading to the dynamic (time-dependent) features due to the randomness of demands. In this sense, the OD matrices are seldom available as it's uncertain where passengers will alight when they arrive at the origin station. Instead, the arrival rates at each station can be easily collected in real time. Thus, a common method to address dynamic time-dependent passenger demand in metro systems is to characterize it through the passenger arrival rate at each station over time. For example, [Wang et al. \(2018\)](#), [Huang et al. \(2020\)](#),

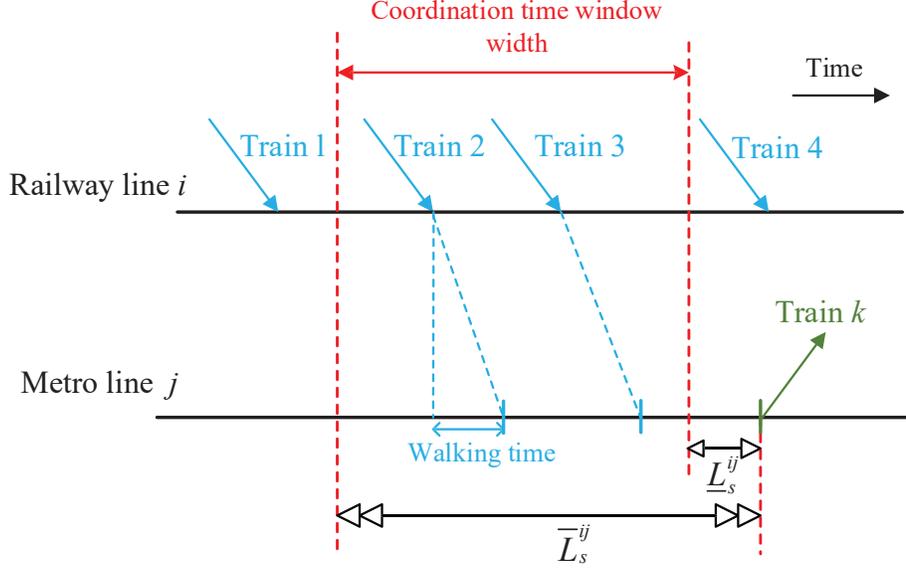


Figure 4: Illustration of coordination time window for lines  $i$  and  $j$  at transfer station  $s$

and Zhou et al. (2022) optimized train schedules by considering time-dependent passenger arrival rates and verified the effectiveness of their models on Beijing metro systems.

As there is a sudden influx of transfer passengers from railway systems, along with the internal arrival of passenger flows, extreme congestion often occurs at busy transfer stations in metro lines. To alleviate such congestion, we can flexibly allocate additional capacity for trains on the overcrowded lines. One effective strategy involves employing trains with multiple carriage groups on congested metro lines, while utilizing trains with a single carriage group for other lines. In our study, the train composition, i.e., involving the number of carriage groups, can be changed according to passenger demand at the depots  $\mathcal{D}$ , located at both ends of each metro line. Note that in our problem, train composition is only allowed to change at the depots, while remaining unchanged at stations during their operations, similar to the assumptions in Zhou et al. (2022) and Pan et al. (2023). We note that since train compositions remain unchanged at stations during their operations, but are changed before departure at the depots located at both ends of each metro line, the dwell times of trains at each station are not affected by the flexible capacity allocation. Thus, in our study, we consider the dwell times of trains at stations as given input parameters. For each metro line  $j$ , we denote set  $\mathcal{C}_j = \{c_j^{\min}, c_j^{\min} + 1, \dots, c_j^{\max}\}$  as the number of carriage groups involved in a train, where  $c_j^{\min}$  and  $c_j^{\max}$  respectively represent the minimum and maximum numbers of carriage groups that can be included in a train. Taking Figure 5 as an example, the capacity of each group of carriages is set as 100, resulting in train capacities of 100, 200, and 300 with one, two, and three groups of carriages, respectively. Figure 5(a) shows the number of cumulative passengers at each station without train capacity allocation, which is limited by the fixed loading capacity of trains with two groups of carriages. Consequently, 100 passengers cannot board train 1 at stations A and B, reducing service quality and increasing safety risks. In contrast, Figure 5(b) shows the flexible capacity allocation strategy. Train 2 consists of only one group of carriages to save the operational costs, while train 3 is executed by three groups of carriages. Obviously, train 3 with a larger loading capacity can transport more passengers than the train without capacity allocation in Figure 5(a). As a result, the number of cumulative passengers at stations is significantly reduced (e.g., from 100 to 0 at stations A and B), improving passenger service quality.

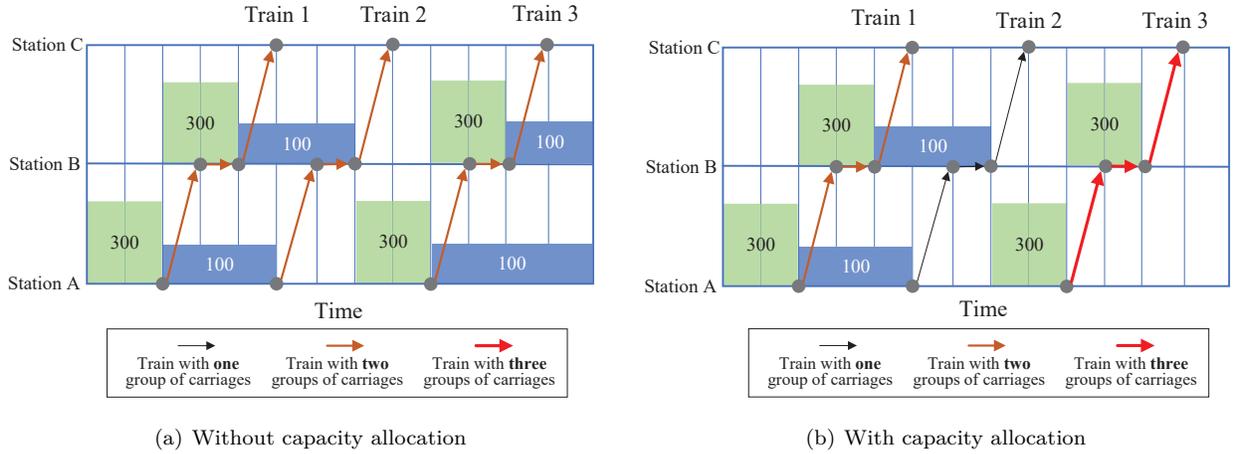


Figure 5: A simple example of train capacity allocation

## 2.2 Objective and Assumptions

The rail managers aim to provide high-quality services to passengers while also considering operational costs. In the context of multimodal synchronization between railway and metro systems, the efficiency of passenger transfers at transport hubs and the overall travel experience in each transportation mode are crucial for enhancing service quality. Hence, our consideration involves the passenger travel and transfer time within railway and metro systems. Furthermore, in congested metro networks such as the Beijing metro, passenger demand may exceed the total train capacity, particularly during peak hours or in densely populated areas. This can result in a significant accumulation of passengers at all involved stations, causing inconvenience and dissatisfaction. To avoid passengers being stranded at stations, we implement a flexible train capacity allocation strategy in metro systems, ensuring that all waiting passengers can board the first arriving train. In practice, the train capacity allocation strategy will certainly influence the cost of train operations. Arising from the strategy, there is a trade-off relationship between the passenger travel times and operational costs. Specifically, more trains involving more carriages with larger loading capacity can reduce passenger travel and waiting times, while undertaking more carriages in service will inevitably increase the operation cost for rail companies (Shi et al., 2022). Thus, our study aims to minimize the total travel time of all considered passengers in multimodal transportation systems and the operational costs of train capacity allocation.

To formulate the described problem without loss of generality, we adopt the following assumptions throughout this paper.

**Assumption 1.** Our study only considers tackling congestion at metro stations caused by the substantial influx of transfer passengers from railway systems, along with the daily commuting needs of nearby residents.

**Assumption 2.** In our study, the flexible capacity allocation strategy can better match the passenger demand by employing trains with multiple carriage groups on congested metro lines, while utilizing trains with a single carriage group for other lines. This strategy allows the train composition, i.e., involving carriage groups, to be changed according to passenger demand (i.e., the number of waiting passengers). Thus, we assume that all waiting passengers can board the first coming train with the flexible train capacity allocation strategy in metro systems.

**Assumption 3.** We assume that each railway station can accommodate all the trains that stop at the station.

**Assumption 4.** In our study, we assume that the section running time and station dwell time of each metro train are predetermined, in order to optimize the train headway, departure times, and train formulation plans (Chai et al., 2023b).

The basic idea that Assumption 1 is valid in our study is as follows: In multimodal railway networks, train timetable coordination is crucial to improve service quality for two-way transfer passengers—those moving between metro and railway lines. In practice, the rail managers are more concerned with the large volume of passengers alighting a mainline train (over 1,000 passengers for each mainline train), since a large part of these passengers taking metro trains will lead to overcrowded transfer stations. Thus, we do not explicitly consider passengers transferring from the metro to the railway. Assumption 3 is derived from the premise that, in principle, the capacity of railway stations is generally sufficient to accommodate train stops except in cases of disruptions and disturbances. In this sense, the capacity of railway stations in our problem has little effect on train scheduling, so we do not consider the capacity of railway stations to simplify the model formulation.

### 3 Mathematical Formulations

This section presents the construction of mathematical formulations for the studied problem. Specifically, section 3.1 introduces our mathematical formulation. Here, we formulate the objective function and derive five sets of constraints related to metro and railway train scheduling, the flow of passengers within these systems, and the coordination of timetables. The linearization of our formulation is presented in Section 3.2. The overall MILP optimization model is provided in Section 3.3.

#### 3.1 Model Formulation

In this section, we construct the mathematical formulation for the investigated problem. First, we list all the sets, parameters and decision variables used in the model formulation in Table 8 and Table 9 in Appendix A, respectively. Then we present the derived constraints related to the integrated optimization problem. Finally, we introduce the objective function in our formulation.

##### 3.1.1 Constraints regarding metro train timetable

In metro systems, the train timetable contains information such as the departure and arrival times of each train at each station, which provides important guidance for train operations. In our problem, the running time and the dwell time of metro trains are given as input parameters, thus we only need to formulate the departure times of each train to determine the metro train timetable. To this end, we define integer decision variables  $w_{k,s}^j$  to represent the departure time of train  $k$  at station  $s$  on metro line  $j$ , and the arrival times of each train can be obtained according to  $w_{k,s}^j$ . The constraints related to the metro train timetable are formulated in constraints (1)-(6):

$$w_{k,s+1}^j - w_{k,s}^j = T_{k,s,j}^{\text{run}} + T_{k,s+1,j}^{\text{dwell}}, \quad \forall k \in \mathcal{K}_j, s, s+1 \in \mathcal{S}_j, j \in \mathcal{J} \quad (1)$$

$$w_{k+1,s}^j - w_{k,s}^j \geq h_j^{\min}, \quad \forall k, k+1 \in \mathcal{K}_j, s \in \mathcal{S}_j, j \in \mathcal{J} \quad (2)$$

$$w_{k+1,s}^j - w_{k,s}^j \leq h_j^{\max}, \quad \forall k, k+1 \in \mathcal{K}_j, s \in \mathcal{S}_j, j \in \mathcal{J} \quad (3)$$

$$b_{k,s,\tau}^j = \begin{cases} 1 - d_{k,s,\tau}^j, & \forall k = 1, s \in \mathcal{S}_j, j \in \mathcal{J}, \tau \in \mathcal{T} \\ d_{k-1,s,\tau}^j - d_{k,s,\tau}^j, & \forall k \in \mathcal{K}_j \setminus \{1\}, s \in \mathcal{S}_j, j \in \mathcal{J}, \tau \in \mathcal{T} \end{cases} \quad (4)$$

$$w_{k,s}^j = d_{k,s,1}^j + \sum_{\tau \in \mathcal{T} \setminus \{1\}} \tau \cdot (d_{k,s,\tau}^j - d_{k,s,\tau-1}^j), \quad \forall k \in \mathcal{K}_j, s \in \mathcal{S}_j, j \in \mathcal{J} \quad (5)$$

$$d_{k,s,\tau-1}^j - d_{k,s,\tau}^j \leq 0. \quad \forall \tau, \tau-1 \in \mathcal{T}, k \in \mathcal{K}_j, s \in \mathcal{S}_j, j \in \mathcal{J} \quad (6)$$

Constraints (1) are formulated to calculate the departure time of each train between two stations, where  $T_{k,s,j}^{\text{run}}$  represents the running time of train  $k$  from station  $s$  to station  $s+1$  and  $T_{k,s,j}^{\text{dwell}}$  represents the dwell time of train  $k$

at station  $s$  on line  $j$ . In general, the parameters  $T_{k,s,j}^{\text{run}}$  and  $T_{k,s,j}^{\text{dwell}}$  are pre-given by the dispatchers. To ensure the safety of train operations, constraints (2) specify that the time difference between the departure of two successive trains should be larger than the minimum headway time. Furthermore, constraints (3) impose a maximum headway time between successive trains to guarantee a certain level of service in metro systems. To formulate train departures with a time-indexed approach, an auxiliary binary variable  $d_{k,s,\tau}^j$  is introduced to represent train departures. Here,  $d_{k,s,\tau}^j = 1$  indicates that train  $k$  departs from station  $s$  before or at timestamp  $\tau$ , and  $d_{k,s,\tau}^j = 0$  otherwise. Another auxiliary binary variable  $b_{k,s,\tau}^j$  is introduced to determine if the timestamp  $\tau$  falls between trains  $k-1$  and  $k$  at station  $s$  on line  $j$ . The value of variables  $b_{k,s,\tau}^j$  can be derived from variables  $d_{k,s,\tau}^j$ , formulated as  $b_{k,s,\tau}^j = d_{k-1,s,\tau}^j - d_{k,s,\tau}^j$  in constraints (4). Specifically, if and only if train  $k-1$  departs before or at  $\tau$  and train  $k$  departs after  $\tau$ , resulting in  $d_{k-1,s,\tau}^j = 1$  and  $d_{k,s,\tau}^j = 0$ , then we have  $b_{k,s,\tau}^j = d_{k-1,s,\tau}^j - d_{k,s,\tau}^j = 1$ ; otherwise  $b_{k,s,\tau}^j = 0$ . Constraints (5) describe the relationship between  $w_{k,s}^j$  and  $d_{k,s,\tau}^j$ . Specifically, when train  $k$  departs from station  $s$  at the first timestamp (i.e.,  $\tau = 1$ ),  $w_{k,s}^j = d_{k,s,1}^j$ ; otherwise, the departure time  $w_{k,s}^j$  of train  $k$  at station  $s$  is equal to the earliest timestamp  $\tau$  at which  $d_{k,s,\tau}^j$  equals 1, i.e.,  $\sum_{\tau \in \mathcal{T} \setminus \{1\}} \tau \cdot (d_{k,s,\tau}^j - d_{k,s,\tau-1}^j)$ . Constraints (6) ensures that variable  $d_{k,s,\tau}^j$  is non-decreasing with respect to the timestamp  $\tau$ .

We use Figure 6 as an example to illustrate the departures of two successive trains  $k-1$  and  $k$  from the same station  $s$  on line  $j$ . Trains  $k-1$  and  $k$  depart from station  $s$  at timestamps 2 and 5, respectively, as indicated by the green dotted rectangle. For any timestamp  $\tau \geq 2$ , we have  $d_{k-1,s,\tau}^j = 1$ , and for any  $\tau \geq 5$ ,  $d_{k,s,\tau}^j = 1$ . Thus, we can derive that  $b_{k,s,\tau}^j = d_{k-1,s,\tau}^j - d_{k,s,\tau}^j = 1$ , for  $\tau = 2, 3, 4$ , which represents the time interval between the departures of these two successive trains from station  $s$ . Some previous studies in the literature used this approach to formulate linear constraints for passenger boarding operations (e.g., Niu and Zhou (2013), Zhou et al. (2022)). In our study, we use the representation of  $b_{k,s,\tau}^j$  to model waiting passengers in metro systems, as detailed in the next section.

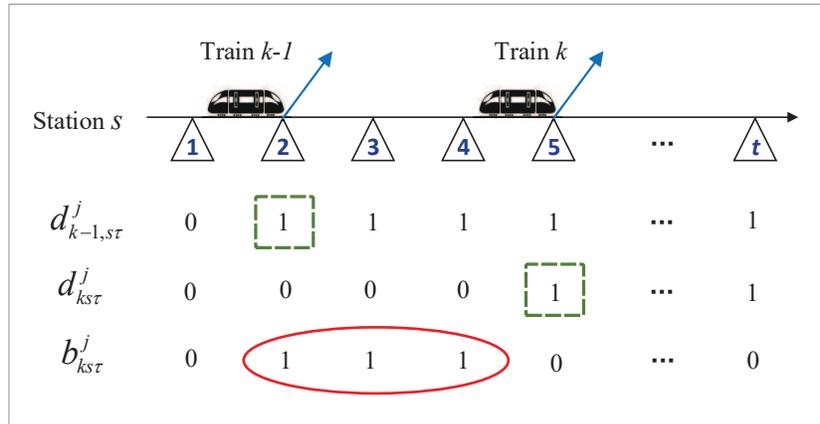


Figure 6: Illustration of two train departures with the time-indexed formulation

One of the most significant differences in train capacity allocation is its ability for trains to change compositions flexibly, involving groups of carriages and improving the utilization of rolling stock resources. This flexibility enhances the utilization of rolling stock resources. Thus, an essential consideration for the flexible allocation of train capacity is the efficient utilization of resources within the constraints of a limited fleet size. To this end, we define a set of integer variables  $v_d^j$ ,  $j \in \mathcal{J}_d$ ,  $d \in \mathcal{D}$  to indicate the fleet size of rolling stock assigned to line  $j$  from depot  $d$ , and integer variables  $c_k^j$  to represent the number of carriages involved by train  $k$  on line  $j$ . As the use of carriages does not exceed the designated fleet size, we derive the following set of constraints regarding limited

rolling stock resources:

$$\sum_{j \in \mathcal{J}_d} v_d^j \leq Z_d^{\max}, \quad \forall d \in \mathcal{D} \quad (7)$$

$$\sum_{k \in \mathcal{K}_j} c_k^j \leq v_d^j, \quad \forall j \in \mathcal{J}_d, d \in \mathcal{D} \quad (8)$$

$$c_j^{\min} \leq c_k^j \leq c_j^{\max}, \quad \forall k \in \mathcal{K}_j, j \in \mathcal{J} \quad (9)$$

Constraints (7) guarantee that the fleet size of rolling stock from depot  $d$  to the connected lines  $\mathcal{J}_d$  should be less than the maximum fleet size  $Z_d^{\max}$  of available rolling stock in depot  $d$  due to the limited budget of rail managers. Constraints (8) ensure that the number of carriages utilized by trains operating on line  $j$  should be less than the allocated fleet size of rolling stock for that line from depot  $d$ . Constraints (9) set the threshold  $[c_j^{\min}, c_j^{\max}]$  for the number of carriages involved by each train.

### 3.1.2 Constraints for modeling passenger flow in metro systems

To model the passenger flows on each metro line, we define the following five sets of decision variables for line  $j$ :  $nb_{k,s}^j$  denotes the number of passengers boarding train  $k$  at station  $s$ ;  $nw_{k,s}^j$  represents the count of passengers waiting for train  $k$  at station  $s$ ;  $nc_{k,s}^j$  indicates the residual capacity of train  $k$  at station  $s$ ;  $na_{k,s}^j$  represents the number of passengers alighting from train  $k$  at station  $s$ ; and  $nr_{k,s}^j$  represents the number of passengers remaining in train  $k$  after departing from station  $s$ . The constraints related to the modeling of passenger flows are formulated in constraints (10)-(15):

$$nb_{k,s}^j = nw_{k,s}^j, \quad \forall k \in \mathcal{K}_j, s \in \mathcal{S}_j, j \in \mathcal{J} \quad (10)$$

$$nb_{k,s}^j \leq nc_{k,s}^j, \quad \forall k \in \mathcal{K}_j, s \in \mathcal{S}_j, j \in \mathcal{J} \quad (11)$$

$$nw_{k,s}^j = \begin{cases} \sum_{\tau \in \mathcal{T}} b_{k,s,\tau}^j \cdot n_{s,\tau,j}^{\text{arr}} + \sum_{e \in E_{is}^{\text{arr}}} \sum_{p \in \mathcal{P}_i} N_p \cdot r_{p,e}^i \cdot \varphi_{k,e,s}^{ij} \cdot R_s^j, & \forall s \in \mathcal{V}_{ij}, k \in \mathcal{K}_j, i \in \mathcal{I}, j \in \mathcal{J} \\ \sum_{\tau \in \mathcal{T}} b_{k,s,\tau}^j \cdot n_{s,\tau,j}^{\text{arr}}, & \forall s \in \mathcal{S}_j, k \in \mathcal{K}_j, j \in \mathcal{J} \end{cases} \quad (12)$$

$$nc_{k,s}^j = \begin{cases} c_k^j \cdot D^{\text{met}}, & \forall s = 1, k \in \mathcal{K}_j, j \in \mathcal{J} \\ c_k^j \cdot D^{\text{met}} - nr_{k,s-1}^j + na_{k,s}^j, & \forall s \in \mathcal{S}_j \setminus \{1\}, k \in \mathcal{K}_j, j \in \mathcal{J} \end{cases} \quad (13)$$

$$na_{k,s}^j = \begin{cases} 0, & \forall s = 1, k \in \mathcal{K}_j, j \in \mathcal{J} \\ Q_s^j \cdot nr_{k,s-1}^j, & \forall s \in \mathcal{S}_j \setminus \{1\}, k \in \mathcal{K}_j, j \in \mathcal{J} \end{cases} \quad (14)$$

$$nr_{k,s}^j = \begin{cases} nb_{k,s}^j, & \forall s = 1, k \in \mathcal{K}_j, j \in \mathcal{J} \\ nr_{k,s-1}^j + nb_{k,s}^j - na_{k,s}^j, & \forall k \in \mathcal{K}_j, s \in \mathcal{S}_j \setminus \{1\}, j \in \mathcal{J} \end{cases} \quad (15)$$

Constraints (10) and (11) indicate the number of boarding passengers  $nb_{k,s}^j$  is associated with the number of waiting passengers  $nw_{k,s}^j$  and the residual capacity  $nc_{k,s}^j$  of train  $k$  at station  $s$ . In our problem, as the flexible train capacity allocation strategy can provide a larger loading passenger capacity with more carriages into service, all waiting passengers can board the first coming train, which is formulated by constraints (10). In addition, constraints (11) guarantee that the number of boarding passengers  $nb_{k,s}^j$  at station  $s$  does not exceed the residual capacity  $nc_{k,s}^j$  of train  $k$ . **In fact, constraints (10) and (11) indicate that there are no stranded passengers (who are unable to board the first coming train due to the limited residual capacity) at stations in metro systems addressed in our problem.** Constraints (12) describe that the passengers  $nw_{k,s}^j$  waiting for train  $k$  consist of those arriving at station  $s$  after the departure of train  $k-1$  and those transferring from railway train connected by train  $k$  at transfer station

$s \in \mathcal{V}_{ij}$ . Constraints (13) indicate the residual capacity  $nc_{k,s}^j$  of train  $k$  at station  $s$  (when the passengers have alighted at the station  $s$  but have not boarded) is related to the number of in-vehicle passengers (i.e., the difference between the number of onboard passengers  $nr_{k,s-1}^j$  from previous stations and the number of alighting passengers  $na_{k,s}^j$  at current station) and the number of carriages  $c_k^j$  involved by the train at current station. Constraints (14) indicate the number of alighting passengers at station  $s$  is proportional  $Q_s^j$  to the number of passengers on board  $nr_{k,s-1}^j$  at the previous station. We note that there are no alighting passengers at the origin station  $s = 1$  for each line, which is the starting point of the passenger loading dynamics. Constraints (15) model the number of passengers remaining onboard when train  $k$  departs from station  $s$ . These constraints indicate that when a train dwells at a station, some in-vehicle passengers whose destination is the current station need to alight, and then passengers waiting on the platform are allowed to board this train if there exists available capacity.

**Remark 3.1** In our model, binary variables  $b_{k,s,\tau}^j$  and  $d_{k,s,\tau}^j$  are mainly used to establish the relationship between traffic flow and passenger flow within metro networks. Specifically, by incorporating variables  $b_{k,s,\tau}^j$  and  $d_{k,s,\tau}^j$ , we derive constraints (12) that describe the number of waiting passengers based on time-independent passenger arrivals and train departure times. In this sense, variables  $b_{k,s,\tau}^j$  and  $d_{k,s,\tau}^j$  are actually auxiliary decision variables in our model to model the demand-driven timetable problem by interconnecting with passenger flow demand and traffic flow supply. We note that this type of binary variables  $b_{k,s,\tau}^j$  and  $d_{k,s,\tau}^j$  is associated with a very huge number, which makes the model difficult to be solved by commercial solvers effectively. In our problem, given the fixed train running times and dwell times, once the departure time of each train at the first station is determined, i.e., for any train  $k$  on line  $j$ , when we have the value of  $d_{k,s,\tau}^j$  for  $s = 1$  and  $\tau \in \mathcal{T}$ , the values of all remaining variables  $d_{k,s',\tau}^j$  for  $s' \in \mathcal{S}_j \setminus \{1\}$  and  $\tau \in \mathcal{T}$  can be subsequently obtained. Leveraging this property, we can derive the following set of equalities regarding variables  $d_{k,s,\tau}^j$ :

$$d_{k,s',\tau'}^j - d_{k,s,\tau}^j = 0, \quad \forall \tau, \tau' \in \mathcal{T} : \tau' = \tau + \sum_{s''=1}^{s'-1} (T_{k,s'',j}^{\text{run}} + T_{k,s'',j}^{\text{dwell}}), \quad s = 1, s' \in \mathcal{S}_j \setminus \{1\}, k \in \mathcal{K}_j, j \in \mathcal{J}$$

It is worth noting that although this set of equalities is quite extensive, it implicitly simplifies the complexity involved in determining the variables  $d_{k,s,\tau}^j$ , to improve the computation efficiency.

### 3.1.3 Constraints regarding railway train scheduling

In our problem, the scheduling of railway trains can be formulated using an event-activity network, a widely employed method for traffic management in railway systems (Zhu and Goverde, 2019). Within railway networks, the train scheduling for each line  $i \in \mathcal{I}$  can be represented by an event-activity network. This network is denoted as a directed graph  $G_i = (E_i, A_i)$ , where  $E_i$  is the set of events, and  $A_i$  is the set of activities for line  $i$ . For a railway line  $i$ , an event  $e \in E_i$  signifies either the arrival or departure of a train at a station, while an activity  $a \in A_i$  connects two such events. Let  $E_{iu}^{\text{dep}} \subset E$  represent the subset of events corresponding to departures from station  $u$ , and  $E_{iu}^{\text{arr}} \subset E$  denote the subset of events corresponding to arrivals at station  $u$ . We use  $l_e$  to represent the train associated with event  $e$ . Activities  $A_i$  are categorized into train activities and headway activities. A train activity  $a \in A_i^{\text{train}}$  can be either a running activity  $a \in A_i^{\text{run}}$  between departure and arrival events at adjacent stations or a dwell activity  $a \in A_i^{\text{dwell}}$  within the same station. A headway activity represents the headway time between two trains running on the same track in a segment, denoted as  $a \in A_i^{\text{track}}$ . Figure 7 illustrates the event-activity network  $G_i = (E_i, A_i)$  for two trains running on a railway line  $i$  with three stations.

To ensure the safety of trains in railway lines, we define the binary decision variable  $\lambda_a^i$ , which determines whether event  $e$  takes place before or after event  $f$ ,  $a = (e, f)$ , and variable  $t_e^i$ , which represents the beginning time of event

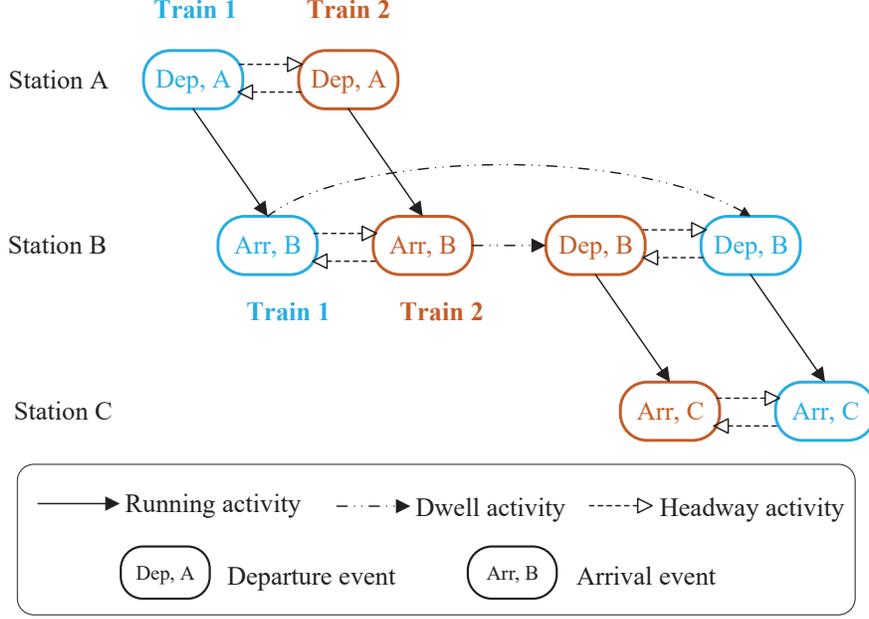


Figure 7: Illustration of an event-activity network formulation for two trains and three stations

$e$  in line  $i$ . Based on the above decision variables, we next present the following constraints to formulate railway train timetable:

$$t_e^i - q_e^i \geq 0, \quad \forall e \in E_i^{cr}, i \in \mathcal{I} \quad (16)$$

$$t_e^i - q_e^i = 0, \quad \forall e \in E_i \setminus E_i^{cr}, i \in \mathcal{I} \quad (17)$$

$$t_f^i - t_e^i = T_a^{\text{run}}, \quad \forall a = (e, f) \in A_i^{\text{run}}, i \in \mathcal{I} \quad (18)$$

$$t_f^i - t_e^i \geq T_a^{\text{dwe}}, \quad \forall a = (e, f) \in A_i^{\text{dw}}, i \in \mathcal{I} \quad (19)$$

$$\lambda_a^i + \lambda_{a'}^i = 1, \quad \forall a = (e, f) \in A_i^{\text{track}} \wedge a' = (f, e) \in A_i^{\text{track}}, i \in \mathcal{I} \quad (20)$$

$$t_f^i - t_e^i + M_1(1 - \lambda_a^i) \geq h_a^{\text{min}}, \quad \forall a = (e, f) \in A_i^{\text{track}}, i \in \mathcal{I} \quad (21)$$

$$\lambda_a^i - \lambda_{a'}^i = 0. \quad \forall (a, a') \in B_i, i \in \mathcal{I} \quad (22)$$

In our study, we focus on the coordination between mainline railway and metro rail systems, instead of optimizing the train schedules for the whole mainline railway network. Therefore, our model focuses on optimizing the schedules of trains that stop at the transfer stations. To avoid the case that our optimized train schedules affect the whole mainline railway network, constraints (16) and (17) specify that the beginning time of events  $E_i^{cr}$  can be adjusted slightly based on the planned timetable to improve the coordination of train timetables in multimodal networks. Constraints (18) indicate the running time  $T_a^{\text{run}}$  of a train in each segment, where activity  $a$  corresponds to the running activity in line  $i$ . Constraints (19) indicate the dwell time  $T_a^{\text{dwe}}$  of a train in each railway station, where activity  $a$  corresponds to the dwell activity in line  $i$ . Constraints (20) indicate the sequence of two departure/arrival events  $e$  and  $f$ , i.e., event  $e$  takes place before or after event  $f$ . If event  $e$  takes place before event  $f$ ,  $\lambda_a = 1$ ; otherwise,  $\lambda_a = 0$ . Constraints (21) guarantee that if event  $e \in E$  takes place after event  $f \in E$ , represented by  $t_f^i - t_e^i < 0$ , then  $\lambda_a^i$  must be equal to 0. Furthermore, if event  $e \in E$  takes place before event  $f \in E$ , then  $\lambda_a^i = 1$ , and hence the minimum headway time  $h_a^{\text{min}}$  between events  $e$  and  $f$  is respected. Note that the parameter  $M$  is a large positive number (e.g., the considered time horizon). Constraints (22) ensure that no overtaking takes place with a segment in railway lines. Specifically, both  $a = (e, f)$  and  $a' = (e', f')$  are segment headway activities (i.e.,  $a, a' \in A_i^{\text{track}}$ ),

where  $e, f \in E_{iu}^{\text{dep}}$ ,  $e', f' \in E_{iu+1}^{\text{arr}}$ ,  $l_e = l_{e'}$  and  $l_f = l_{f'}$ . In other words,  $a$  corresponds to two departures and  $a'$  corresponds to two arrivals of the same trains. The activity pair  $(a, a')$  is called an *order activity pair* and the set  $B_i$  is defined as the set of all order activity pairs on line  $i$ . Thus, constraints (22) ensure that trains  $l_e$  and  $l_f$  arrive in station  $u + 1$  in the same order as how they departed from station  $u$ , i.e.,  $\lambda_a = \lambda_{a'}$ .

### 3.1.4 Constraints for modeling passenger flow in railway systems

In railway systems, passengers who travel on the railway network book tickets in advance. Therefore, they use a seat reservation system to plan the trains to take during their journey (Zhan et al., 2021). In the context of seat reservation systems, the passenger demand is given and fixed, which can be split into different groups  $\mathcal{P}$  according to their journey. Typically, passengers make the train choice decisions to maximize their utility, i.e., minimize their traveling time between their origin station  $o_p$  and their destination  $d_p$ . To formulate the train choice decisions, we define the binary decision variables  $r_{p,e}^i$  to represent if passengers in group  $p$  choose train  $l_e$  to travel. The constraints related to the modeling of passenger flows in railway systems are formulated in constraints (23)–(28):

$$\sum_{e \in E_{io_p}^{\text{dep}}} r_{p,e}^i = 1, \quad \forall p \in \mathcal{P}_i, i \in \mathcal{I} \quad (23)$$

$$\sum_{e \in E_{id_p}^{\text{arr}}} r_{p,e}^i = 1, \quad \forall p \in \mathcal{P}_i, i \in \mathcal{I} \quad (24)$$

$$r_{p,e}^i = r_{p,f}^i, \quad \forall p \in \mathcal{P}_i, (e, f) \in A_i^{\text{train}}, i \in \mathcal{I} \quad (25)$$

$$r_{p,e}^i \leq \max\{0, \min\{t_e^i - T_p + 1, T_a^{\text{dwe}}\}\}, \quad \forall p \in \mathcal{P}_i, a = (e, f) \in A_i^{\text{dw}} : e \in E_{io_p}^{\text{dep}}, i \in \mathcal{I} \quad (26)$$

$$r_{p,e}^i \leq T_a^{\text{dwe}}, \quad \forall p \in \mathcal{P}_i, a = (e, f) \in A_i^{\text{dw}} : e \in E_{id_p}^{\text{arr}}, i \in \mathcal{I} \quad (27)$$

$$\sum_{p \in \mathcal{P}_i} r_{p,e}^i \cdot N_p \leq D^{\text{rail}}. \quad \forall e \in E_i, i \in \mathcal{I} \quad (28)$$

Constraints (23), (24), and (25) ensure the transition of passengers from their origin stations, to their destination, and via some intermediate stations respectively. Constraints (26) indicate that passengers can get on a train at a station if and only if the train stops there (i.e.,  $T_a^{\text{dwe}} > 0$ ,  $a \in A_i^{\text{dwell}}$ ) and departs no earlier than their intended departure time (i.e.,  $t_e^i \geq T_p$ ,  $e \in E_{io_p}^{\text{dep}}$ ). Note that when any term of the  $\min\{\cdot\}$  operator is smaller than 1 (i.e.,  $\min\{\cdot\} \leq 0$ , since  $t_e^i$ ,  $T_p$ , and  $T_a^{\text{dwe}}$  are all integers), constraints (26) will reduce to  $r_{p,e}^i \leq 0$ . To ensure the first term of the  $\min\cdot$  operator is not less than 1 when the condition  $t_e^i \geq T_p$  holds, we express it as  $t_e^i - T_p + 1$ . Similarly, constraints (27) ensure that passengers can arrive at their destination only by taking trains that stop at the destination. Constraints (28) guarantee that the number of boarding passengers in train  $l_e$  does not exceed its loading passenger capacity.

### 3.1.5 Constraints for the coordination of train timetable

On the basis of the above formulation, the railway train schedule determines the arrival times  $t_e^i$  of each event  $e \in E_{iu}^{\text{arr}}$  at station  $u$  on each railway line  $i$ , and the metro timetable determines the departure times  $w_{k,s}^j$  of each metro train  $k$  at station  $s$  on each line  $j \in \mathcal{J}$ . The timetable of railway train  $l_e$  of line  $i$  and metro train  $k$  of line  $j$  will then be coordinated at transfer station  $s \in \mathcal{V}_{ij}$  if  $\underline{L}_s^{ij} \leq w_{k,s}^j - t_e^i \leq \overline{L}_s^{ij}$  is satisfied and metro train  $k$  is identified as the first connecting train for train  $l_e$ . To formulate the coordination relationship between railway and metro trains, we introduce four sets of binary decision variables as follows. The first set of variables  $x_{k,e,s}^{ij}$  determines if the difference between the departure time of train  $k$  on line  $j$  and the arrival time of train  $l_e$  on line  $i$  is greater than or equal to  $\underline{L}_s^{ij}$  at transfer station  $s \in \mathcal{V}_{ij}$ . The second  $y_{k,e,s}^{ij}$  determines if train  $k$  on line  $j$  is the first available connecting train for the arrival railway train  $l_e$  on line  $i$  at transfer station  $s$ . The third set of

variables  $z_{k,e,s}^{ij}$  determines whether the difference between the departure time of train  $k$  on line  $j$  and the arrival time of train  $l_e$  on line  $i$  is less than or equal to  $\bar{L}_s^{ij}$  at transfer station  $s$ . The last set of variables  $\varphi_{k,e,s}^{ij}$  indicate whether the timetables of metro train  $k$  on line  $j$  and railway train  $l_e$  on line  $i$  are coordinated at transfer station  $s \in \mathcal{V}_{ij}$ . Given these defined variables, we can denote the following constraints to formulate the coordination of railway train timetable and metro train timetable in the network:

$$w_{k,s}^j - t_e^i \geq \underline{L}_s^{ij} - M_2(1 - x_{k,e,s}^{ij}), \quad \forall e \in E_{is}^{\text{arr}}, k \in \mathcal{K}_j, s \in \mathcal{V}_{ij}, i \in \mathcal{I}, j \in \mathcal{J} \quad (29)$$

$$w_{k,s}^j - t_e^i < \underline{L}_s^{ij} + M_3 x_{k,e,s}^{ij}, \quad \forall e \in E_{is}^{\text{arr}}, k \in \mathcal{K}_j, s \in \mathcal{V}_{ij}, i \in \mathcal{I}, j \in \mathcal{J} \quad (30)$$

$$w_{k,s}^j - t_e^i \leq \bar{L}_s^{ij} + M_4(1 - z_{k,e,s}^{ij}), \quad \forall e \in E_{is}^{\text{arr}}, k \in \mathcal{K}_j, s \in \mathcal{V}_{ij}, i \in \mathcal{I}, j \in \mathcal{J} \quad (31)$$

$$w_{k,s}^j - t_e^i > \bar{L}_s^{ij} - M_5 z_{k,e,s}^{ij}, \quad \forall e \in E_{is}^{\text{arr}}, k \in \mathcal{K}_j, s \in \mathcal{V}_{ij}, i \in \mathcal{I}, j \in \mathcal{J} \quad (32)$$

$$y_{k,e,s}^{ij} = x_{k,e,s}^{ij} - x_{k-1,e,s}^{ij}, \quad \forall e \in E_{is}^{\text{arr}}, k, k-1 \in \mathcal{K}_j, s \in \mathcal{V}_{ij}, i \in \mathcal{I}, j \in \mathcal{J} \quad (33)$$

$$\varphi_{k,e,s}^{ij} = y_{k,e,s}^{ij} \cdot z_{k,e,s}^{ij}, \quad \forall e \in E_{is}^{\text{arr}}, k \in \mathcal{K}_j, s \in \mathcal{V}_{ij}, i \in \mathcal{I}, j \in \mathcal{J} \quad (34)$$

$$\sum_{k \in \mathcal{K}_j} \varphi_{k,e,s}^{ij} = 1, \quad \forall e \in E_{is}^{\text{arr}}, s \in \mathcal{V}_{ij}, i \in \mathcal{I}, j \in \mathcal{J} \quad (35)$$

Constraints (29) and (30) with a big- $M$  structure indicate if the time difference between the departure of metro train  $k$  on line  $j$  and the arrival of train  $l_e$  on line  $i$  is larger than or equal to  $\underline{L}_s^{ij}$ . Specifically, when  $x_{k,e,s}^{ij} = 1$ , constraint (29) becomes  $w_{k,s}^j - t_e^i \geq \underline{L}_s^{ij}$ , which means that train  $k$  departs from station  $s$  at least  $\underline{L}_s^{ij}$  time units later than the arrival time of train  $l_e$ . Meanwhile, constraint (30) becomes  $w_{k,s}^j - t_e^i < \underline{L}_s^{ij} + M_3$ , and this constraint will always be satisfied due to the big- $M$  value. If  $x_{k,e,s}^{ij} = 0$ , constraint (30) becomes  $w_{k,s}^j - t_e^i < \underline{L}_s^{ij}$  and constraint (29) will always be satisfied. Similarly, constraints (31) and (32) indicate if the time difference between the departure of metro train  $k$  on line  $j$  and the arrival of train  $l_e$  on line  $i$  is less than or equal to  $\bar{L}_s^{ij}$ . Constraints (33) indicate that the values of variables  $y_{k,e,s}^{ij}$  can be determined by  $x_{k,e,s}^{ij}$  and  $x_{k-1,e,s}^{ij}$ . Specifically, variable  $y_{k,e,s}^{ij}$  is equal to 1 if and only if  $x_{k,e,s}^{ij} = 1$  and  $x_{k-1,e,s}^{ij} = 0$ , which means train  $k$  is the first connecting train for railway train  $l_e$ . Constraints (34) ensure that the timetables of metro and railway trains are coordinated only when the time difference between the departure of metro train  $k$  on line  $j$  and the arrival of train  $l_e$  on line  $i$  at transfer station  $s$  falls within  $[\underline{L}_s^{ij}, \bar{L}_s^{ij}]$ . Specifically, if  $y_{k,e,s}^{ij} = 1$  and  $z_{k,e,s}^{ij} = 1$ ,  $\varphi_{k,e,s}^{ij}$  is set to 1; otherwise  $\varphi_{k,e,s}^{ij}$  equals 0. Finally, constraints (35) guarantee each railway train must be coordinated by a metro train in the network.

We take Figure 8 as an example to illustrate the timetable coordination at transfer station  $s$  of a railway train  $l_e$  on line  $i$  and metro train  $k$  on line  $j$ . Because the departure times of trains  $k$ ,  $k+1$ , and  $k+2$  are later than the arrival time of train  $l_e$  at station  $s$  plus the minimum coordinated time  $\underline{L}_s^{ij}$ , we can easily obtain  $x_{k-1,e,s}^{ij} = 0$ ,  $x_{k,e,s}^{ij} = 1$ ,  $x_{k+1,e,s}^{ij} = 1$  and  $x_{k+2,e,s}^{ij} = 1$ . Meanwhile, as train  $k$  is the first connecting train for the arrival train  $l_e$ ,  $y_{k,e,s}^{ij} = 1$ . In addition, since trains  $k-1$  and  $k$  depart from the station earlier than the arrival time of train  $l_e$  at station  $s$  plus the maximum coordinated time  $\bar{L}_s^{ij}$ , we have  $z_{k-1,e,s}^{ij} = 1$  and  $z_{k,e,s}^{ij} = 1$ . As indicated in constraints (34), we thus obtain  $\varphi_{k,e,s}^{ij} = 1$ , which represents train  $l_e$  is coordinated by train  $k$  at station  $s$ .

### 3.1.6 Objective Function

In this study, both the economic and service objectives are considered from the perspectives of operators and passengers. In reality, the operational costs of the rolling stock (carriage in our problem) undertaking train services are affected by many complex factors, e.g., energy consumption, maintenance costs, running kilometers, etc. In our study, for convenience, we introduce parameter  $g$  to represent the operational cost of employing a single carriage in a train running from the original station to the terminal station (i.e., implementing a service), as motivated by

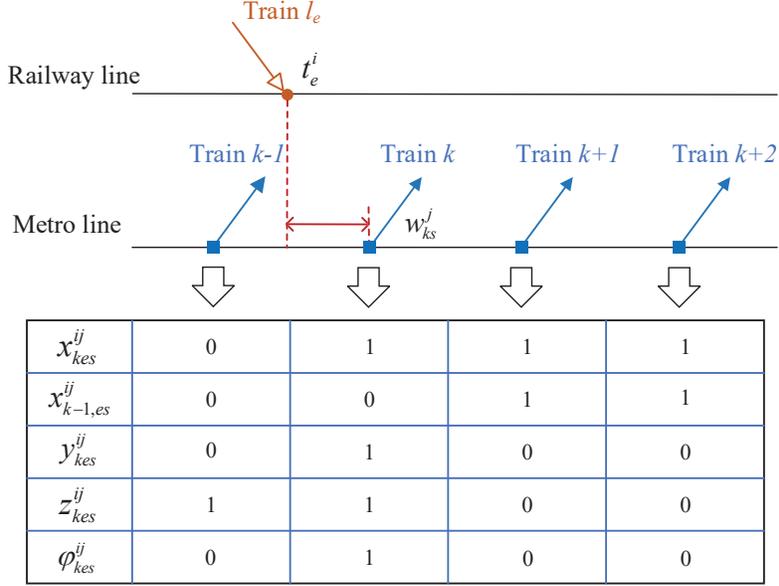


Figure 8: Illustration of coordination events at transfer station

Wang et al. (2018), Shi et al. (2022) and Zhou et al. (2022). Thus, the operational costs can be calculated as follows.

$$M^{\text{cost}} = \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}_j} g \cdot c_k^j \quad (36)$$

From the perspective of passengers, the total passenger travel time, including waiting time and onboard travel time, will be considered in the objective function to improve service quality and transfer efficiency. In our problem, the total passenger travel time includes the passenger travel time  $F^{\text{rail}}$  in railway systems and the passenger travel time  $F^{\text{metro}}$  in metro systems, and the transfer time  $F^{\text{tra}}$  for passengers at the transfer stations. The computation of total passenger travel time and waiting time in metro systems is indicated by equation (37). The total passenger waiting time is derived as the sum of the waiting time for newly arriving passengers between two successive trains and the waiting time for the detained passengers, the latter equal to 0 according to constraints (10).

$$F^{\text{metro}} = \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}_j} \sum_{s \in \mathcal{S}_j} \sum_{\tau \in \mathcal{T}} \sum_{t \geq \tau} b_{k,s,t}^j \cdot n_{s,t,j}^{\text{arr}} \cdot T^{\text{unit}} + \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}_j} \sum_{s \in \mathcal{S}_j} T_{k,s,j}^{\text{run}} \cdot nr_{k,s}^j \quad (37)$$

The passenger travel time in railway systems is calculated as:

$$F^{\text{rail}} = \sum_{i \in \mathcal{I}} \sum_{p \in \mathcal{P}_i} \sum_{e \in E_{id_p}^{\text{arr}}} (t_e^i \cdot r_{p,e}^i - T_p) \cdot N_p \quad (38)$$

The transfer time for passengers at transfer stations is derived as follows.

$$F^{\text{tra}} = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}_j} \sum_{s \in \mathcal{V}_{ij}} \sum_{p \in \mathcal{P}_i} \sum_{e \in E_{is}^{\text{arr}}} N_p \cdot r_{p,e}^i \cdot \varphi_{k,e,s}^{ij} \cdot (w_{k,s}^j - t_e^i) \cdot R_s^{ij} \quad (39)$$

Finally, by introducing four weight coefficients, i.e.,  $\theta_1$ ,  $\theta_2$ ,  $\theta_3$ , and  $\theta_4$  to balance the above four parts, the objective function in our study can be written as follows:

$$\min \theta_1 \cdot M^{\text{cost}} + \theta_2 \cdot F^{\text{rail}} + \theta_3 \cdot F^{\text{metro}} + \theta_4 \cdot F^{\text{tra}}$$

Naturally, decision-makers can strike a balance among different perspectives by adjusting the values of these four coefficients in the objective function. For example, if the decision-makers prefer higher passenger satisfaction with

less passenger traveling time, coefficient  $\theta_2, \theta_3, \theta_4$  can be set much larger than  $\theta_1$ ; otherwise, a large value of  $\theta_1$  can be employed for fewer operational costs.

In our proposed model, some constraints have nonlinear forms, e.g., constraints (12) and constraints (34), which make the model hard to be solved by some commercial solvers directly (e.g., CPLEX and GUROBI). **Thus, we linearize these nonlinear constraints by applying transformation properties proposed by Williams (2013) for solution convenience. Interested readers may refer to this reference for more details. For the sake of compactness, the details of these linearization techniques are detailed in Appendix B.**

### 3.2 Overall optimization model

Since the involved nonlinear constraints (12), (34), (38) and (39) are reformulated as the equivalent linear forms, we construct the following optimization model for the studied problem:

$$\begin{aligned} \min \quad & \Omega = \theta_1 \cdot M^{\text{cost}} + \theta_2 \cdot F^{\text{rail}} + \theta_3 \cdot F^{\text{metro}} + \theta_4 \cdot F^{\text{tra}} \\ \text{s.t.} \quad & \text{Constraints (1)-(11), (13)-(33), (35)-(37), (49)-(56)} \end{aligned} \quad (40)$$

Notice that model (40) is an MILP formulation, where the decision variables can be classified into four categories. The first type of variables, i.e.,  $x_{k,e,s}^{ij}, y_{k,e,s}^{ij}, z_{k,e,s}^{ij}, \varphi_{k,e,s}^{ij}$ , is associated with the binary decision variables that determine the timetable coordination of railway and metro trains. The second refers to integer variables, i.e., the timing decision variables  $w_{k,s}^j$  and  $t_e^i$ , which are employed to formulate the times of metro and railway trains, respectively. The third type indicates the binary variable  $r_{p,e}^i$  to represent the train choice for passenger groups in railway systems. The last refers to integer variables  $c_k^j$ , representing the number of carriages involved by metro trains. The total number of variables and constraints is listed in Table 2.

Table 2: Number of involved variables and constraints in the optimization model (40)

Variables or constraints	Total number
<b>Variables</b>	
Binary variables $x_{k,e,s}^{ij}, y_{k,e,s}^{ij}, z_{k,e,s}^{ij}, \varphi_{k,e,s}^{ij}$	$4 \cdot  \mathcal{I}  \cdot  \mathcal{J}  \cdot  \mathcal{K}  \cdot  E  \cdot  \mathcal{V}_{ij} $
Binary variables $r_{p,e}^j$	$ \mathcal{J}  \cdot  \mathcal{P}  \cdot  E $
Integer variables $w_{k,s}^j, c_k^j$	$2 \cdot  \mathcal{J}  \cdot  \mathcal{K}  \cdot  \mathcal{S} $
Integer variables $t_e^i$	$ \mathcal{I}  \cdot  E $
Continuous variables $nw_{k,s}^j, nc_{k,s}^j, na_{k,s}^j, nb_{k,s}^j$	$4 \cdot  \mathcal{J}  \cdot  \mathcal{K}  \cdot  \mathcal{S} $
<b>Constraints regarding metro systems</b>	
Constraints (1)-(15)	$15 \cdot  \mathcal{K}_j  \cdot  \mathcal{S}_j  \cdot  \mathcal{J} $
<b>Constraints regarding railway train scheduling</b>	
Constraints (18)-(19)	$3 \cdot  A_i^{\text{train}}  \cdot  \mathcal{I} $
Constraints (20)-(21)	$3 \cdot  A_i^{\text{track}}  \cdot  \mathcal{I} $
<b>Constraints regarding passenger flow in railway train systems</b>	
Constraints (23)-(24)	$2 \cdot  \mathcal{P}_i  \cdot  \mathcal{I} $
Constraints (25)-(28)	$3 \cdot  \mathcal{P}_i  \cdot  \mathcal{I}  \cdot  A_i^{\text{train}} $
<b>Constraints for the coordination of timetable</b>	
Constraints (29)-(35)	$7 \cdot  E_{is}^{\text{arr}}  \cdot  \mathcal{K}_j  \cdot  \mathcal{V}_{ij}  \cdot  \mathcal{I}  \cdot  \mathcal{J} $

To obtain a better root relaxation gap in the branch-and-bound tree, we introduce different values (that are as small as possible) to replace the big- $M$  parameters while still keeping the correctness of the formulation. In our

model (40), a total of 6 groups of big- $M$  are employed in constraints (21), (29-32), (53) and (55-56). We next present the determined values of the big- $M$  used in the model. The big- $M$  in constraints (21) needs to be determined as the maximum value of  $|\mathcal{T}| + h_a^{\min}$  to guarantee that the left-hand side of constraints (21) is always larger than  $h_a^{\min}$ ,  $a \in A_i^{\text{track}}$ . The big- $M$  in constraints (29) needs to ensure the left-hand side of these constraints is always larger than or equal to the right-hand side when  $x_{k,e,s}^{ij} = 0$ . Thus, the  $M_2$  is taken as the maximum value of  $|\mathcal{T}| + \underline{L}_s^{ij}$ . The big- $M$  in constraints (30) needs to ensure the left-hand side of the one is always rigorously less than the right-hand side when  $x_{k,e,s}^{ij} = 1$ . Thus, the  $M_3$  is taken as the maximum value of  $|\mathcal{T}| - \underline{L}_s^{ij}$ . With the analogous idea as the  $M_2$  and  $M_3$ ,  $M_4$  and  $M_5$  in constraints (31) and (32) are determined as  $|\mathcal{T}| - \bar{L}_s^{ij}$  and  $|\mathcal{T}| - \bar{L}_s^{ij}$ , respectively. The big- $M$  in constraints (53) and (55-56) are taken as the maximum value of  $|\mathcal{T}|$  to ensure that the auxiliary variables must be less than the considered time horizon.

**Remark 3.2** The space-time network and event-activity network are two widely recognized modeling techniques in train scheduling formulation, each offering distinct characteristics and benefits. In our model formulation, we adopt the event-activity network to formulate train scheduling in railway systems, while leveraging space-time network-dependent approaches to tackle the metro train scheduling problem, to take advantage of these two modeling techniques. The event-activity network method denotes an event as the arrival or departure of a train at a station, while an activity bridges two events, representing trains running between stations, dwelling at stations, departure intervals, and other related activities. This method effectively characterizes the departure, arrival, dwell, and movement of each train through events and activities, thereby facilitating constraints such as train overtaking and running orders. Thus, it is frequently utilized in modeling train scheduling in railway networks comprising multiple tracks and routes, as studied by Veelenturf et al. (2016) and Zhu and Goverde (2019). On the other hand, the space-time network modeling method excels in capturing flow-based problems across high-dimensional times and spaces by comprehensively representing all temporal and spatial points. Given the continuous influx of passengers and the high frequency of train departures in metro systems, the space-time network method, with high-density time and space division, is often employed to describe the passenger flow, traffic flow, and their interrelationships, as demonstrated in Yin et al. (2017) and Wang et al. (2023a).

## 4 Branch-and-Cut Algorithm

Due to the large number of binary auxiliary and decision variables, the size of the model (40) escalates rapidly. Consequently, solving real-life instances to optimality using MIP solvers may become impractical due to the excessive computational time required. To address this challenge, we next present a branch-and-cut (BaC) approach. This method involves identifying valid inequalities and designing tailored branching strategies based on the theoretical properties of the model.

### 4.1 Valid inequalities

The above model (40) is sufficient to solve the addressed problem, while it contains numerous linear programming (LP) relaxation feasible spaces. It is possible to add additional inequalities to reduce the feasible space without removing integer solutions. In this section, we introduce five sets of valid inequalities to refine the formulation of the proposed MILP model (40). Theoretically, an inequality is deemed *valid* for an MILP if every integral solution of the MILP satisfies it. In other words, valid inequalities are constraints that approximate the feasible space to the integer convex hull. As a result, a valid inequality known as a “cutting plane”, can be incorporated into an MILP to enhance its LP bounds and formulation (Fischetti et al., 2017; Wolsey, 2020) that will be more tractable

to solve.

In our problem, one main reason for a weak LP relaxation stems from the binary variables  $x_{k,e,s}^{ij}$ ,  $y_{k,e,s}^{ij}$ ,  $z_{k,e,s}^{ij}$  and  $\varphi_{k,e,s}^{ij}$ , because these four sets of variables determine the coordination of train timetables in multimodal railway networks. They are involved in the discretization constraints modeled with the help of big- $M$  parameters. Thus, we introduce some valid inequalities to strengthen the linear relaxation and speed up the solution process. **We prove these valid inequalities are valid for the proposed model. This means that in any integer feasible solution, the valid inequalities are not violated. On the other hand, when solving a LP relaxation, they can further constrain the solution space and thus cut off fractional solutions.**

Our first set of valid inequalities arises from the binary variables  $y_{k,e,s}^{ij}$  that represents whether metro train  $k$  on line  $j$  is the first available connecting train for arrival railway train  $l_e$  on line  $i$  at the transfer station. Since each railway train can only be coordinated by a metro train in the considered network as indicated in constraints (35), the first connecting train  $k$  is the one that coordinates with each train  $l_e$  in railway systems. Using this property, we introduce the following inequalities related to  $y_{k,e,s}^{ij}$  to strengthen our MILP formulation:

$$\sum_{k \in \mathcal{K}_j} y_{k,e,s}^{ij} = 1, \quad \forall e \in E_{is}^{\text{arr}}, \quad s \in \mathcal{V}_{ij}, \quad i \in \mathcal{I}, \quad j \in \mathcal{J} \quad (41)$$

**Proposition 4.1** *Inequities (41) are valid for the MILP model (40).*

**Proof 4.1** For any railway train  $l_e$ ,  $e \in E_{is}^{\text{arr}}$ ,  $i \in \mathcal{I}$ , let us suppose that there exist feasible solutions  $x_{k,e,s}^{ij}$  representing the time relationship between railway train  $l_e$  and metro trains  $\mathcal{K}_j$  on line  $j$ , such that for trains  $k'$  departing earlier than train  $k$  (i.e.,  $k' < k$ ), we have  $x_{k',e,s}^{ij} = 0$ , and for trains  $k$  and  $k''$  departing after  $k$ ,  $x_{k,e,s}^{ij} = x_{k'',e,s}^{ij} = 1$ . According to constraints (33), we can calculate  $y_{k,e,s}^{ij} = 1$  for train  $k$ , while for the remaining trains  $k' \in \mathcal{K}_j \setminus \{k\}$  on line  $j$ , we have  $y_{k',e,s}^{ij} = 0$ . This indicates that only train  $k$  is the first available connecting train of arrival railway train  $l_e$  at the transfer station  $s \in \mathcal{V}_{ij}$ . Therefore, the proof is complete.  $\square$

Based on the description of the first set of valid inequalities above, once metro train  $k$  is the first connecting train for railway train  $l_e$ , i.e.,  $y_{k,e,s}^{ij} = 1$ , the train  $k$  must be able to coordinate with the railway train at station  $s$  ( $\varphi_{k,e,s}^{ij} = 1$ ). Thus, we can obtain  $z_{k,e,s}^{ij} = 1$  according to constraints (34). In addition, if a metro train  $k$  fails to coordinate with a railway train  $l_e$  due to their time difference larger than the maximum coordinated time  $\bar{L}_s^{ij}$ , i.e.,  $z_{k,e,s}^{ij} = 0$ , then the railway train must not be coordinated by the metro train  $k$ . In other words, train  $k$  must not be the first connecting train for train  $l_e$ . Thus, we have  $y_{k,e,s}^{ij} = 0$ . Using this property, we can derive the following class of inequalities:

$$z_{k,e,s}^{ij} - y_{k,e,s}^{ij} \geq 0, \quad \forall e \in E_{is}^{\text{arr}}, \quad k \in \mathcal{K}_j, \quad s \in \mathcal{V}_{ij}, \quad i \in \mathcal{I}, \quad j \in \mathcal{J} \quad (42)$$

**Proposition 4.2** *Inequities (42) are valid for the MILP model (40).*

**Proof 4.2** For a railway train  $l_e$ ,  $e \in E_{is}^{\text{arr}}$ ,  $i \in \mathcal{I}$ , let us suppose that there exist feasible solutions  $y_{k,e,s}^{ij}$  and  $z_{k,e,s}^{ij}$  that represent the coordination relationship between train  $k$  and  $l_e$ , such that constraints (42) is violated, i.e.,  $y_{k,e,s}^{ij} = 1$  and  $z_{k,e,s}^{ij} = 0$ . This implies that train  $k$  does not coordinate with train  $l_e$ . According to constraints (29)-(33), we thus derive the following set of equalities:

$$\begin{aligned} x_{k',e,s}^{ij} &= 1, \quad k' \in \{k, k+1, \dots, K_j\} \\ x_{k',e,s}^{ij} &= 0, \quad k' \in \{1, 2, \dots, k-1\} \\ z_{k',e,s}^{ij} &= 0, \quad k' \in \{k, k+1, \dots, K_j\} \end{aligned}$$

Combining the above three sets of equalities, we can explicitly obtain that for the train  $l_e$ , the variables  $\varphi_{k,e,s}^{ij} = 0$  for all trains  $k \in \mathcal{K}_j$  on metro line  $j$ . This violates constraints (35) that each railway train should be coordinated

by a metro train in the network. In other words, there is no feasible solution in which inequalities (42) are violated. Therefore, the proof is complete.  $\square$

Next, we consider another type of valid inequalities arising from the coordination time window. Recall that the minimum and maximum coordinated times exist for each transfer station  $s \in \mathcal{V}_{ij}$  connecting lines  $i$  and  $j$  in the network. As specified in constraints (29)-(32), binary variables  $x_{k,e,s}^{ij}$  impose the time difference between the train  $k$  departure and train  $l_e$  arrival at station  $s$  to satisfy the minimum coordinated time, while variables  $z_{k,e,s}^{ij}$  impose their time difference to meet the maximum coordinated time. Thus, for any railway train  $l_e$ ,  $e \in E_{is}^{\text{arr}}$ ,  $i \in \mathcal{I}$ , the departure time of each metro train  $k$  on line  $j$  is either larger than the arrival time of the train  $l_e$  plus  $\underline{L}_s^{ij}$  or less than the arrival time of the train  $l_e$  plus  $\bar{L}_s^{ij}$ . This condition can be expressed symbolically as  $x_{k,e,s}^{ij} = 1$  or  $z_{k,e,s}^{ij} = 1$  or both  $x_{k,e,s}^{ij}$  and  $z_{k,e,s}^{ij}$  equal to 1 for each train  $k$ . Using this property, we introduce the following two sets of inequalities related to  $x_{k,e,s}^{ij}$  and  $z_{k,e,s}^{ij}$  to strengthen our MILP formulation:

$$\sum_{k \in \mathcal{K}_j} z_{k,e,s}^{ij} + \sum_{k \in \mathcal{K}_j} x_{k,e,s}^{ij} \geq |\mathcal{K}_j| + 1, \quad \forall e \in E_{is}^{\text{arr}}, s \in \mathcal{V}_{ij}, i \in \mathcal{I}, j \in \mathcal{J} \quad (43)$$

$$\sum_{k \in \mathcal{K}_j} z_{k,e,s}^{ij} + \sum_{k \in \mathcal{K}_j} x_{k,e,s}^{ij} \leq \frac{\bar{L}_s^{ij} - \underline{L}_s^{ij}}{h_j^{\min}} + |\mathcal{K}_j| + 1, \quad \forall e \in E_{is}^{\text{arr}}, s \in \mathcal{V}_{ij}, i \in \mathcal{I}, j \in \mathcal{J} \quad (44)$$

**Proposition 4.3** *Inequalities (43) and (44) are valid for the MILP model (40).*

**Proof 4.3** We first prove that inequalities (43) are valid. For railway train  $l_e$ ,  $e \in E_{is}^{\text{arr}}$ ,  $i \in \mathcal{I}$ , we assume that a total of  $|\mathcal{K}_j|$  metro trains operate on line  $j$ . These trains depart from the transfer station  $s \in \mathcal{V}_{ij}$  connected with railway line  $i$ , and are potentially coordinated with train  $l_e$ . Let us suppose that the time difference between the departures of two trains  $k$  and  $k+1$  and the arrival of train  $l_e$  falls within the coordination time window. According to constraints (29)-(32), we can obtain the value of variables  $x_{k,e,s}^{ij}$  and  $z_{k,e,s}^{ij}$  for each metro train  $k$  on the line  $j$ , as shown in Figure 9. Furthermore, we can derive the following set of equalities:

$$\begin{aligned} x_{k',e,s}^{ij} &= 0, \quad k' \in \{1, 2, \dots, k-1\} \\ x_{k',e,s}^{ij} &= 1, \quad k' \in \{k, k+1, \dots, K_j\} \\ z_{k',e,s}^{ij} &= 1, \quad k' \in \{1, 2, \dots, k+1\} \\ z_{k',e,s}^{ij} &= 0, \quad k' \in \{k+2, k+3, \dots, K_j\} \end{aligned}$$

Combining the above four sets of equalities, we can explicitly derive the following equalities for railway train  $l_e$ ,  $e \in E_{is}^{\text{arr}}$ ,  $i \in \mathcal{I}$ :

$$\begin{aligned} \sum_{k \in \mathcal{K}_j} x_{k,e,s}^{ij} &= |\mathcal{K}_j| - k + 1, \\ \sum_{k \in \mathcal{K}_j} z_{k,e,s}^{ij} &= k + 1, \end{aligned}$$

Then, we can obtain that

$$\sum_{k \in \mathcal{K}_j} z_{k,e,s}^{ij} + \sum_{k \in \mathcal{K}_j} x_{k,e,s}^{ij} = |\mathcal{K}_j| + 2 \geq |\mathcal{K}_j| + 1,$$

Therefore, the proof of the validity of inequalities (43) is complete.

Next, we prove that inequalities (44) are valid. Using the same example illustrated in Figure 9, let us assume  $h_j^{\min} \leq \bar{L}_s^{ij} - \underline{L}_s^{ij} \leq 2h_j^{\min}$ . Thus we have  $|\mathcal{K}_j| + 2 \leq \frac{\bar{L}_s^{ij} - \underline{L}_s^{ij}}{h_j^{\min}} + |\mathcal{K}_j| + 1 \leq |\mathcal{K}_j| + 3$ , and we can derive the following

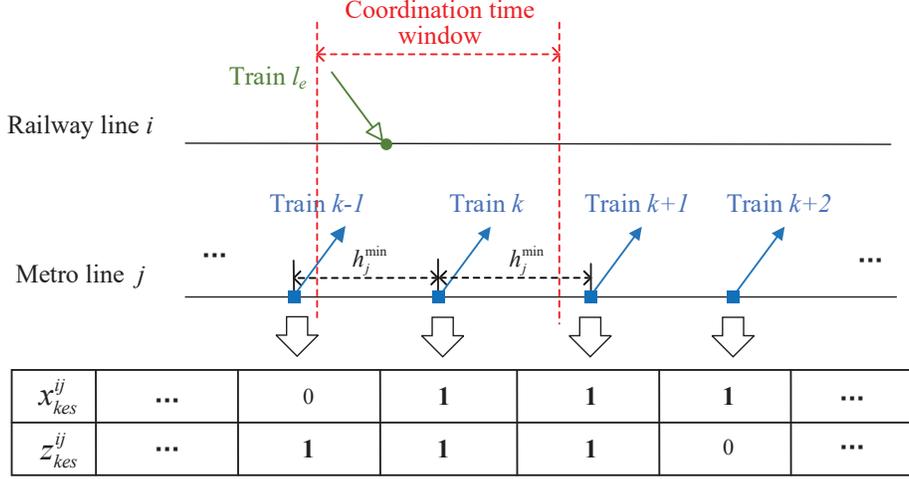


Figure 9: An illustrative example of valid inequalities (43) and (44)

inequality:

$$\sum_{k \in \mathcal{K}_j} z_{k,e,s}^{ij} + \sum_{k \in \mathcal{K}_j} x_{k,e,s}^{ij} = |\mathcal{K}_j| + 2 \leq \frac{\bar{L}_s^{ij} - \underline{L}_s^{ij}}{h_j^{\min}} + |\mathcal{K}_j| + 1$$

It is essential to note that this inequality remains valid for other cases involving different values of the minimum headway time  $h_j^{\min}$  and the coordination time window  $[\underline{L}_s^{ij}, \bar{L}_s^{ij}]$ . For instance, if  $2h_j^{\min} \leq \bar{L}_s^{ij} - \underline{L}_s^{ij} \leq 3h_j^{\min}$ , then train  $k+2$  may coordinate with train  $l_e$ . In this case, we have

$$\sum_{k \in \mathcal{K}_j} z_{k,e,s}^{ij} + \sum_{k \in \mathcal{K}_j} x_{k,e,s}^{ij} = |\mathcal{K}_j| + 3 \leq \frac{\bar{L}_s^{ij} - \underline{L}_s^{ij}}{h_j^{\min}} + |\mathcal{K}_j| + 1$$

Hence, the proof is complete.  $\square$

We can take advantage of the railway headway time parameters to define the last family of valid inequalities. Let us consider the train timetables on two lines  $i$  and  $j$  to be coordinated at a given transfer station  $s$  such that the minimum headway time  $h_a^{\min}$  for railway trains is greater than the length of the coordination time window, i.e.,  $h_a^{\min} > \bar{L}_s^{ij} - \underline{L}_s^{ij}$ . If a railway train  $l_e$  coordinates with a metro train  $k$ , the coordination of  $k$  with railway trains  $l_{e1}$  or  $l_{e2}$  is impossible since there would not be enough time units to ensure solution feasibility, where  $l_{e1}$  and  $l_{e2}$  represent the train departing before and after train  $l_e$ , respectively. Generalizing the previous idea, we obtain the following result by extending the *headway inequalities* proposed in [Fouilhoux et al. \(2016\)](#):

$$\sum_{e \in E_{is}^{\text{arr}}} \varphi_{k,e,s}^{ij} \leq \frac{\bar{L}_s^{ij} - \underline{L}_s^{ij}}{h_a^{\min}} + 1, \quad \forall k \in \mathcal{K}_j, s \in \mathcal{V}_{ij}, i \in \mathcal{I}, j \in \mathcal{J} \quad (45)$$

**Proposition 4.4** *Inequities (45) are valid for the MILP model (40).*

**Proof 4.4** We consider that there is a group of railway trains arriving at transfer station  $s$  with the minimum headway time  $h_a^{\min}$ . Let us assume  $h_a^{\min} \leq \bar{L}_s^{ij} - \underline{L}_s^{ij} \leq 2h_a^{\min}$  such that train  $k$  is coordinated with trains  $l_{e2}$  and  $l_{e3}$  at the station  $s$ , i.e.,  $\varphi_{k,e2,s}^{ij} = \varphi_{k,e3,s}^{ij} = 1$ , as shown in Figure 10. So, we can derive the following inequality:

$$\sum_{e \in E_{is}^{\text{arr}}} \varphi_{k,e,s}^{ij} = 2 \leq \frac{\bar{L}_s^{ij} - \underline{L}_s^{ij}}{h_a^{\min}} + 1$$

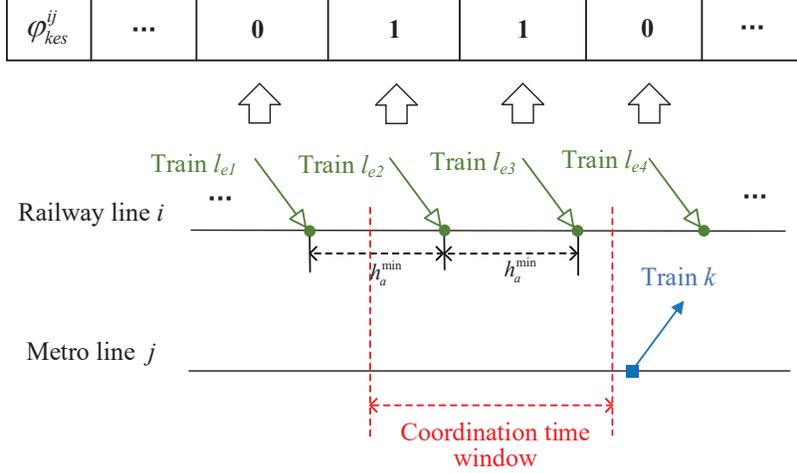


Figure 10: An illustrative example of valid inequalities (45)

We note that this inequality remains valid for other cases involving different values of the minimum headway time  $h_a^{\min}$  and the coordination time window  $[\underline{L}_s^{ij}, \bar{L}_s^{ij}]$ . For instance, if  $2h_a^{\min} \leq \bar{L}_s^{ij} - \underline{L}_s^{ij} \leq 3h_a^{\min}$ , then train  $l_{e4}$  runs on line  $i$  with the minimum headway time may be coordinated by train  $k$  in this example. In this case, we have

$$\sum_{e \in E_{is}^{\text{arr}}} \varphi_{k,e,s}^{ij} = 3 \leq \frac{\bar{L}_s^{ij} - \underline{L}_s^{ij}}{h_j^{\min}} + 1$$

Thus, the proof is complete.  $\square$

## 4.2 Branching strategy

In the BaC framework, the branching strategy emerges as a pivotal step, given its profound influence on algorithmic performance through the generation of child nodes. In the context of MILPs, the branching strategy typically unfolds in two phases: first, identifying a candidate variable or a set of them for branching, and then creating child subproblems by imposing bounds on these variables to compel them away from fractional values.

In this study, we introduce a branching strategy (BS) according to the properties of the model (40) to decide the selection of node children. The fundamental insight driving our branching strategy is that the values of variables  $\varphi_{k,e,s}^{ij}$  determine the ranges of railway train arrival times and metro train departure times. Therefore, branching on these variables at the beginning of a branch-and-bound search can potentially produce a rough schedule of involved trains. This initial branching may direct the search and expedite the branching process. Specifically, we set priority to branch on the coordinated variable  $\varphi_{k,e,s}^{ij}$ , while other variables are branched according to the default strategies provided by CPLEX, i.e., the node with the lowest bound is chosen for branching, following the common practice in the field of operations research (see e.g., Zhang et al. (2021); Chai et al. (2024)). The priority given to branching on variable  $\varphi_{k,e,s}^{ij}$  is adjusted based on the number of nodes explored within the branch-and-bound search tree. Specifically, if the number of nodes explored falls within the range  $[O^\sigma, O^\sigma + m]$  with  $m$  denoting the size of the range and  $\sigma$  representing the frequency of our branching strategy execution, we apply our branching strategy to create two subproblems (nodes) by branching on the variable  $\varphi_{k,e,s}^{ij}$ . Otherwise, if the number of explored nodes lies outside this range, the node with the lowest bound is chosen for branching on other variables. After executing the aforementioned branching strategy at  $\sigma$  branching, it is necessary to update the range of nodes from  $[O^\sigma, O^\sigma + m]$

to  $[O^{\sigma+1}, O^{\sigma+1} + m]$  for the subsequent  $\sigma + 1$  branching. Specifically, we calculate  $O^{\sigma+1}$  using the formula:

$$O^{\sigma+1} = O^{\sigma} + \kappa \cdot m,$$

where  $\kappa$  represents the updated step-size. To ensure the branching captures the search procedure by updating the nodes range, we establish a link between  $\kappa$  and the objective function value (VOF) of the LP relaxation at the root node, as well as the current best lower bound (LB). Thus, we heuristically determine the step-size  $\kappa$  as follows:

$$\kappa = \ln(\text{LB}/\text{VOF}) + 1.$$

This equation indicates that  $\kappa$  is equal to 1 at the root node due to  $\text{LB}=\text{VOF}$ . Subsequently,  $\kappa$  is marginally larger than 1 at the start of the solving process since LB is close to VOF. Nevertheless, as the solution progresses, the difference between LB and VOF is getting bigger, leading to a much larger step-size  $\kappa$ . In summary, our branching strategy BS is frequently executed with a small step-size (i.e., marginally above 1) on the frontier (near the root node) of the search tree. As the search proceeds, our branching strategy is utilized less frequently, and the branching node is automatically chosen instead (i.e., choosing the node with the lowest bound for branching).

Next, we introduce our branching strategy on the coordinated variables  $\varphi_{k,e,s}^{ij}$ . In the solution process, if an optimal solution to the coordination of train timetables is not an integer, then there must exist a railway train  $l_e$  such that  $0 < \sum_{k \in \mathcal{K}_j} \varphi_{k,e,s}^{ij} < 1$ . Then branching is applied on a fractional  $\sum_{k \in \mathcal{K}_j} \varphi_{k,e,s}^{ij}$ , where we select the *special ordered set* strategy for branching. In the realm of discrete optimization, the special ordered set strategy typically pertains to cases where at most one variable can take a non-zero value for a set of variables, with all others remaining at 0 (Wolsey, 2020). This implies a selection of at most one from a set of possibilities. The fundamental concept of the special ordered set strategy involves prioritizing sets of binary variables by establishing an order among variables rather than individually addressing each variable in existing literature (see e.g., Coniglio et al. (2021); Adelgren and Gupte (2022)).

In our algorithm, the idea of the strategy involves prioritizing certain sets of binary variables by defining an order among variables. In our problem, recalling constraints (35), we have

$$\sum_{k \in \mathcal{K}_j} \tilde{\varphi}_{k,e,s}^{ij} = 1, \quad \forall e \in E_{is}^{\text{arr}}, \quad s \in \mathcal{V}_{ij}, \quad i \in \mathcal{I}, \quad j \in \mathcal{J} \quad (46)$$

with  $\tilde{\varphi}_{k,e,s}^{ij} \in \{0, 1\}$  for any train  $l_e$ ,  $e \in E_{is}^{\text{arr}}$  at station  $s \in \mathcal{V}_{ij}$ . It is evident that the aforementioned constraints adhere to the implementation case of the special ordered set strategy; each feasible solution for  $e \in E_{is}^{\text{arr}}$  and  $s \in \mathcal{V}_{ij}$  ensures that at most one variable in the set  $\mathcal{K}_j = \{1, 2, \dots, K_j\}$  can be set to one. Thus, at each branching node, we can effectively construct a new set  $\Upsilon_j$  from  $\mathcal{K}_j$  according to the value of  $\tilde{\varphi}_{k,e,s}^{ij}$ , where some variables may have fractional values at the current node, and then explore the nodes in  $\Upsilon_j$  at first.

Next, we introduce the observation behind the selection of  $\Upsilon_j \subseteq \mathcal{K}_j$ : for each arrival event  $e \in E_{is}^{\text{arr}}$  and station  $s \in \mathcal{V}_{ij}$ , the variables  $\tilde{\varphi}_{k,e,s}^{ij}$  in constraints (46) can be sequenced according to the order of metro trains, i.e., the index of train  $k \in \mathcal{K}_j$ . The priority assigned to each variable is inversely proportional to the train index; the earlier the train departs, the higher the priority. This is because a lower index implies an earlier train departure, which subsequently leads to reduced passenger transfer time within the coordination time window in our problem. Consequently, we establish the following principles to determine the set  $\Upsilon_j$  at each branching node:

$$\Upsilon_j := \{k \in \mathcal{K}_j | k \leq \eta\} \quad \text{where } \eta := \sum_{k \in \mathcal{K}_j} k \cdot \tilde{\varphi}_{k,e,s}^{ij} \quad (47)$$

where  $\eta$  is a parameter indicating the branching value of the branching strategy. Within this special ordered set strategy, we denote a penalty to the value of  $\eta$  with the train  $k$ , favoring the selection of variables associated with

earlier trains  $k \in \Upsilon_j$ . According to the constructed set  $\Upsilon_j$ , the search direction within each branching node of the searching tree changes based on the following cases:

$$\sum_{k \in \Upsilon_j} \varphi_{k,e,s}^{ij} = 1 \quad \text{or} \quad \sum_{k \in \Upsilon_j} \varphi_{k,e,s}^{ij} = 0 \quad (48)$$

Using the LP relaxation solution  $\tilde{\varphi}_{k,e,s}^{ij} = \{0.1, 0.7, 0, 0.2\}$  for  $k \in \mathcal{K}_j = \{1, 2, 3, 4\}$  as an illustrative example, we calculate the branching value  $\eta$  as follows:  $1.1 + 2 \times 0.7 + 4 \times 0.2 = 2.3$ . Thus, we get the set  $\Upsilon_j = \{k \in \mathcal{K}_j | k \leq 2.3\}$ . Consequently, we split the fractional variables  $\tilde{\varphi}_{k,e,s}^{ij}$  into two new subproblems as  $\varphi_{1,e,s}^{ij} + \varphi_{2,e,s}^{ij} = 1$  and  $\varphi_{3,e,s}^{ij} + \varphi_{4,e,s}^{ij} = 0$ , i.e.,  $\varphi_{3,e,s}^{ij} = \varphi_{4,e,s}^{ij} = 0$  due to  $\varphi_{k,e,s}^{ij} \in \{0, 1\}$ . In this case, the branching scheme proves significantly more effective compared to the traditional approach of branching on a single variable at each step, leading to a considerable reduction in the number of nodes required in the branch-and-bound tree.

### 4.3 Branch-and-cut implementation

In this section, we describe the whole framework of our BaC algorithm (as shown in Figure 11) and provide detailed of its implementation mainly from the following four phases:

- **Phase 1: Initialize.** During the initialization phase, only the constraints involved in the model (40) are included in the MILP. To get a tighter formulation for the model (40), five sets of valid inequalities (41-45) proposed in Section 4.1 can be added to the model, either during the initialization phase or dynamically based on the LP relaxation solutions, i.e., they are added later in the algorithm when the LP relaxation solutions violate these inequalities.

- **Phase 2: Select and branch on nodes.** Following the initialization phase, the branching tree comprises only a single root node. This node is automatically chosen to initiate the branch-and-bound process. By solving the LP relaxation of the root node and branching on variables randomly, two (or more) new nodes are created, which then constitute the list of unexplored nodes. The determination of the nodes to be explored next in the solving process is essential. As demonstrated in Section 4.2, if the number of explored nodes falls within the range  $[O^\sigma, O^\sigma + m]$ , the special ordered set branching strategy is employed to branch on variables  $\varphi_{k,e,s}^{ij}$ , generating two subproblems (nodes). Otherwise, the node with the best bound is chosen for branching on other integer and binary variables, following the default branching decision in the CPLEX solver. After executing the branching strategy at  $\sigma$  branching, it is necessary to update the range of nodes from  $[O^\sigma, O^\sigma + m]$  to  $[O^{\sigma+1}, O^{\sigma+1} + m]$  for the next  $\sigma + 1$  branching.

- **Phase 3: Solve LP and update bounds.** Solving the LP relaxation of one resulting subproblem at the generated node aims to derive potentially new bounds. The other node(s) are added to the list of unexplored nodes. The upper bound (UB) and LB of the node preceding the current node in the branching tree are updated with the newly found bounds. If a set of new integer solutions is identified, it updates the UB accordingly. If the LB of this node surpasses the best-known UB, i.e.,  $LB \geq UB$ , there is no need for further investigation of this node, and it is pruned, thereby being not explored anymore. After pruning, another node in the list is chosen for solving. If its bound is lower than the best-known upper bound, there is still room for improvement, prompting the algorithm to continue branching on this node eventually.

- **Phase 4: Find and add cuts.** At this phase, the algorithm searches for violated inequalities, i.e., finding cuts. In principle, there are two potential scenarios for finding cuts: Firstly, if the LP solutions of a node fail to satisfy one of the valid inequalities (41-45), the unsatisfied inequality is added to cut fractional solutions of the LP relaxation, after which the LP relaxation at the node is re-solved. Secondly, if the LP solutions of a node satisfy all proposed valid inequalities, the algorithm assesses whether the number of explored nodes falls within the range

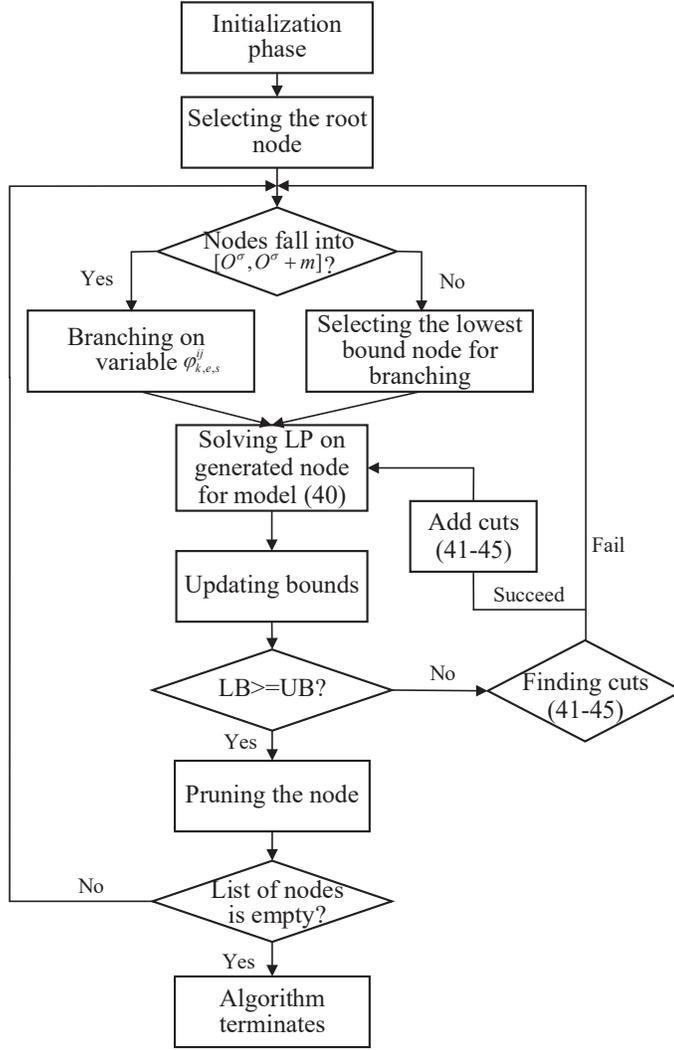


Figure 11: The framework of the branch-and-cut algorithm

$[O^\sigma, O^\sigma + m]$  and selects a branching node in the list of unexplored nodes to execute the next iteration. It is worth noting that if the list of unexplored nodes is empty, the current best upper bound represents the optimal solution.

## 5 Numerical Experiments

In this section, we present the computational results on two sets of numerical experiments. The first set of experiments, conducted on a small-scale network, aims to demonstrate the effectiveness of different valid inequalities and branching strategies. The second set is implemented on a real-world instance of a railway-metro network in Beijing, to assess potential benefits for railway operators through the application of train capacity allocation and the coordination of train timetables. All models and algorithms in this section are coded in C++ on a workstation equipped with 32 Intel Xeon 6140 processors and 32GB RAM. In addition, we utilize the IBM ILOG CPLEX Optimization Studio 12.10.0 solver on the same platform.

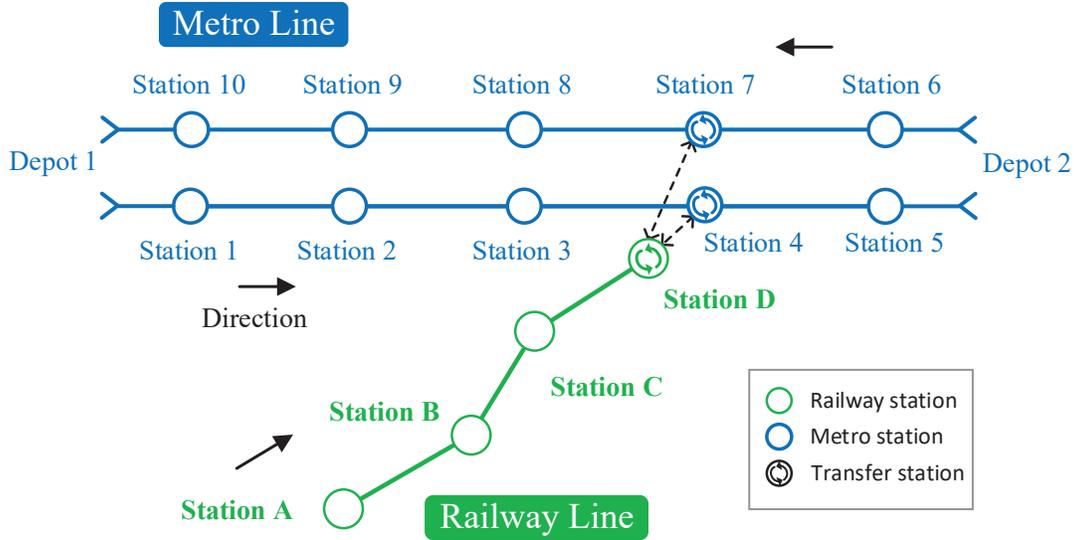


Figure 12: Network layout in the small-scale case study

## 5.1 Small-scale case study

### 5.1.1 Network layout and parameter settings

As shown in Figure 12, the small-scale case study considers a network with a bi-directional metro line and a unidirectional railway line. The network involves four railway stations (denoted as stations A to D), ten metro stations (numbered from 1 to 10), and two depots located at the ends of the metro line. In this network, there is a transfer hub involving railway station D and metro stations 4 and 7, where passengers can alight from railway trains at station D and transfer onto metro trains at either station 4 or station 7 according to their needs. The planning time horizon is set as two hours, and we divide it into 120 time units (i.e.,  $|\mathcal{T}| = 120$ ).

To simulate passenger flow dynamics during peak and off-peak hours in metro systems, we generate time-dependent passenger demand using the following strategy: For the initial 60 time units (i.e., one hour), we utilize a roughly monotone increasing value to represent the number of arrival passengers  $n_{s,\tau,j}^{\text{arr}}$ ,  $\tau \in \{1, 2, \dots, 60\}$ ; in the subsequent 60 time units, i.e.,  $\tau \in \{61, 62, \dots, 120\}$ , we generate a roughly decreasing passenger arrival numbers. **The passenger demand in metro systems is depicted in Figure 13.** Concerning passenger flows in railway systems, we consider 20 passenger groups, each containing 300 passengers. We randomly generate their intended trips including the origin station  $o_p$ , destination station  $d_p$ , and the intended departure time  $T_p$ . **The details of the generated passenger groups for railway systems are provided in Table 3.** The other parameters associated with the small-scale case study are summarized in Table 4.

Using the network described above, we construct 6 instances by changing the number of scheduled trains and the coordination time window in our experiments. Specifically, we fix the number of railway trains to 10, while the number of scheduled trains on each metro line is either 10 or 20, and the coordination time window is set as [1,5], [3,6] or [5,10]. Combining these input parameters, we construct a total of 6 instances. For example, instance 1 involves 10 metro trains with a time window of [1,5], while instance 2 involves 10 metro trains with a time window of [3,6]. To assess the effectiveness of different valid inequalities (41-45) and branching strategies, we adopt the following methods to tackle these 6 instances. **“None” signifies that the instance is directly solved by MIP solver CPLEX, utilizing the branch and cut algorithm which involves solving a sequence of LP subproblems, branching, adding cuts, and leveraging heuristics, all with the default parameter settings;** “Constraints” indicates that the proposed valid

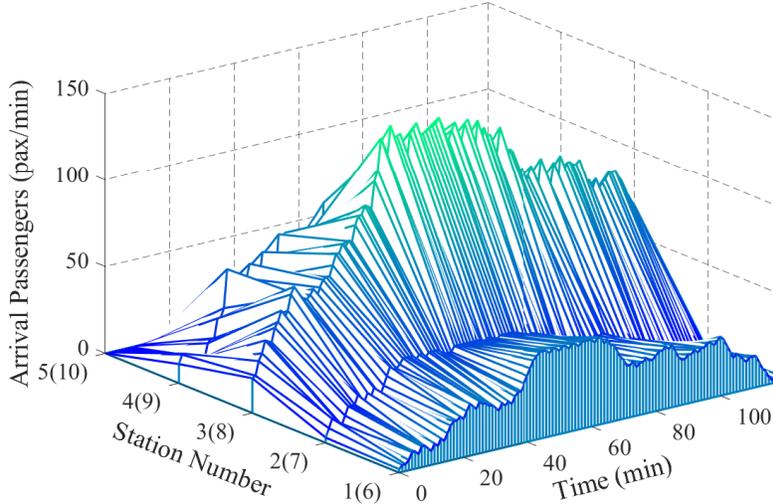


Figure 13: Passenger demand variations in metro systems in the small case study

inequalities are directly added into the model right from the beginning; “Cuts” denotes that the valid inequalities are added to the model dynamically during the BaC process, facilitated by setting generic callbacks; “BS” represents the utilization of the proposed branching strategy without employing valid inequalities; Finally, “BaC” combines the branching strategy with dynamically added cuts. In these experiments, the computational time limit is set as 600s, i.e., we terminate the algorithm and return the current best solution when reaching ten minutes.

### 5.1.2 Computational results with different cuts and branching strategies

Table 5 reports the computational results for six small-scale instances. In Table 5, we report the number of branch-and-bound nodes, the computational times, and optimality gaps. In the final column of Table 5, we respectively report the LP relaxation lower bound values (at the root node in the branch-and-bound tree) for “Constraints” strategies, to assess the tightness of each set of valid inequalities, the number of added user cuts for “Cuts” strategies, and the number of created branches for BS, to evaluate the effectiveness of our branching strategy. For our BaC method, we list the number of added user cuts and created branches within parentheses.

From the experimental results of Table 5, we observe that all the methods can obtain the (near-)optimal solutions within the computational time limit for the majority of instances due to their relatively small scales. Particularly, for instances 1 and 2, the default CPLEX solver (“None” in Table 5) efficiently solves the model in 90 seconds. However, its performance rapidly decreases as the number of metro trains increases and the coordination time window expands. Regarding the “Constraints” strategies, we notice that the reduction in computational times is not obvious compared to default CPLEX, despite slightly larger objective function values at the root node (i.e., LP bound) by incorporating constraints into the model. Furthermore, for instances 4-6, adding all constraints (i.e., “All” in Table 5) shows less computationally effective than the default CPLEX solver. This is attributed to the substantial number of constraints required in the model initially, leading to a time-consuming solution process. The computational results on these six instances also reveal that (45) is the strongest set of cuts, as the strategy exclusively adding constraints (45) yields the most positive impact on reducing the computational time and the number of explored nodes in the search tree.

In terms of the “Cuts” strategies, it is obvious that these strategies obtain better solutions compared to both

Table 3: The information of passenger groups in the small-scale case study

Passenger group $p$	Origin station $o_p$	Destination station $d_p$	Intended departure time $T_p$ (min)	Volume $N_p$
1	Station A	Station D	1	300
2	Station B	Station D	9	300
3	Station A	Station C	5	300
4	Station A	Station D	15	300
5	Station A	Station B	20	300
6	Station A	Station D	20	300
7	Station B	Station D	10	300
8	Station A	Station C	15	300
9	Station A	Station D	10	300
10	Station A	Station B	13	300
11	Station A	Station D	1	300
12	Station B	Station C	35	300
13	Station A	Station C	5	300
14	Station C	Station D	20	300
15	Station A	Station B	10	300
16	Station A	Station D	31	300
17	Station B	Station D	19	300
18	Station A	Station C	35	300
19	Station A	Station D	20	300
20	Station A	Station D	20	300

“Constraints” strategies and the default CPLEX solver. For instance, in instance 5, we can get a feasible solution with an optimality gap of 2.78% by dynamically adding inequalities (45). Similarly, in instance 6, the optimality gap is further reduced to 4.01% by adding the valid inequalities dynamically, which evidently outperforms the other methods and reveals the great benefits of these valid inequalities. From the results in Table 5, we observe that when dynamically adding inequalities (41), the number of cuts remains zero for all instances 1-6. This suggests that inequalities (41) may be a set of redundant constraints for our MILP model. In other words, the original constraints of the model (40) already imply the imposition of inequality (41), resulting in inequality (41) failure to cut any fractional solution in the domain of LP in the model. Another interesting phenomenon is that, dynamically adding cuts (42) can cut some fractional solutions in instances 2 and 3, i.e., 3 and 8 cuts are added in the model for instances 2 and 3, respectively. Nevertheless, in instances 4-6, no user cuts (42) are added in the solution process. This is probably because for different instances, the different input parameters (e.g., the minimum headway time and the coordination time window) affect the value of the fractional solution and the constant term in the inequalities (42), which further affect the action of the cuts. The computational results on these six instances also reveal that adding all inequalities or all but (41) is the best strategy for obtaining a high-quality solution for the MILP model, especially for large instances.

An important insight from the table is that both BS and BaC can still obtain (near-)optimal solutions across all six instances. Even in the challenging instance 6, both strategies produce solutions with less than a 4% optimality gap. The reason is that our BS branching strategy facilitates early branching on the coordination decision variables (i.e.,  $\varphi_{k,e,s}^{ij}$ ) to guide the search direction effectively. Therefore, BS can help find high-quality and feasible solutions much faster than the default CPLEX solver, which employs random branching. These results reveal the significance

Table 4: The input parameters in the small case study

Parameters	Notations	Values	Unit
Loading passenger capacity of a railway train	$D^{\text{rail}}$	1200	person
Loading passenger capacity of a metro train	$D^{\text{met}}$	800	person
Alighting rate	$Q_s^j$	0.5	
Passenger transfer rate	$R_s^{ij}$	0.3, 0.5	
Metro train dwell time	$T_{k,s,j}^{\text{dwell}}$	1	minutes
Metro train running time	$T_{k,s,j}^{\text{run}}$	3	minutes
Coordination time window	$[\underline{L}_s^{ij}, \bar{L}_s^{ij}]$	[1,5], [3,6], [5,10]	minutes
Metro time headway	$h_j^{\min}, h_j^{\max}$	3,15	minutes
Railway time headway	$h_a^{\min}$	3	minutes
The minimum/maximum number of carriages in metro trains	$c_j^{\min}, c_j^{\max}$	1,3	
Cost of using one carriage	$g$	200	
The maximum rolling stock allocated in depot	$Z_d^{\max}$	30	

of tailored branching strategies. In particular, we also notice that the BaC method outperforms BS for instances 5 and 6, illustrating the effectiveness of combining valid inequalities with our customized branching strategies. We report the computational results of the obtained train timetables, the capacity allocation and the coordination strategies for instance 1 in the small case study in Appendix C.

## 5.2 Real-world case study

In this section, we conduct real-world case studies on the realistic railway-metro network in Beijing. We use historical passenger demand data in the network and compare our approach with the non-coordinated train timetables. Furthermore, we investigate the benefits of flexible capacity allocation in alleviating congestion. We also implement a series of sensitivity analyses on the key parameters, such as the coordination time window, to provide rail managers with valuable managerial insights.

### 5.2.1 Description of field data and parameter settings

As illustrated in Figure 14, our real-world experiments are conducted on a network involving a metro line, i.e., Line No.4 in Beijing metro, and two railway lines, i.e., Beijing-Zhangjiakou high-speed railway line and Beijing-Tianjin intercity railway line.

These lines are selected due to their significance as the primary commuter and travel corridors in both urban and regional areas of Beijing. Specifically, Line No.4 serves as the main north-south corridor for commuters in Beijing, and carries transfer passengers from Beijing South Railway Station and Beijing North Railway Station. Both the Beijing-Zhangjiakou high-speed railway and the Beijing-Tianjin intercity railway are major components of the “Eight Vertical and Eight Horizontal High-Speed Railway Network” in China, with a large number of passengers traveling in and out of Beijing on these two railway lines every day. In particular, as a pivotal transportation project for the 2022 Winter Olympics, the Beijing-Zhangjiakou high-speed Railway caters to a substantial volume of athletes traveling between the two cities.

In the network, Line No.4, including 34 physical stations, has bidirectional tracks with two operation directions for train services. In this paper, the operation direction from TGY station to AHQB station is the upstream direction, while the opposite direction is the downstream direction. Due to the independent nature of passenger

Table 5: Computational results in the small-scale case study

Selection	Number of nodes						Computation time (s)						Gap (%)						LP bound/Number of cuts/Branches						
	Instance	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
<b>Constraints<sup>a</sup></b>																				LP Bound					
None	4629	8781	13455	890	2922	4412	19	89	600	600	600	600	0	0	3.11	1.7	3.59	5.01	200163	200453	200875	200057	200348	200882	
Only (41)	4722	9032	13211	1566	2567	4456	13	90	600	600	600	600	0	0	3	1.97	3.61	4.44	200163	200453	200875	200057	200348	200882	
Only (42)	4231	5287	12169	2110	2761	4501	26	45	600	600	600	600	0	0	2.58	2.22	3.57	5.33	200163	200503	200905	200057	200348	200882	
Only (43)	4482	5198	9256	1362	2453	4182	29	52	600	600	600	600	0	0	2.97	2.71	3.87	4.35	201212	200809	200889	200060	200361	200900	
Only (44)	4539	4286	10162	2408	2216	2025	22	44	600	600	600	600	0	0	2.69	1.22	3.63	5.5	200163	200453	200879	200073	200348	200882	
Only (45)	<b>251</b>	1082	9437	<b>731</b>	1205	4499	<b>10</b>	43	600	600	600	600	0	0	<b>2.43</b>	1.41	<b>2.89</b>	4.66	202316	202968	200875	200064	200355	200882	
All but (41)	406	924	10013	592	299	3388	10	45	600	600	600	600	0	0	2.38	1.83	3.56	5.23	202316	202971	200905	200074	200360	200901	
All but (42)	296	1391	9308	807	1312	4501	9	36	600	600	600	600	0	0	2.4	1.13	3.71	4.88	202316	202971	200875	200074	200355	200889	
All but (43)	3211	2172	9788	4212	1514	3222	10	40	600	600	600	600	0	0	2.4	1.88	3.98	4.44	202316	202968	200905	200074	200355	200882	
All but (44)	165	1480	8999	1483	2211	3434	10	31	600	600	600	600	0	0	2.52	1.13	3.65	4.89	202316	202968	200905	200064	200355	200889	
All but (45)	4861	4421	11164	2803	2324	1988	20	42	600	600	600	600	0	0	3.12	1.88	3.62	5.01	200163	200453	200905	200074	200348	200901	
All	406	910	10233	634	454	3509	15	26	600	600	600	600	0	0	2.43	1.83	3.62	5.23	202316	202971	200905	200074	200360	200901	
<b>Cuts<sup>a</sup></b>																			User cuts						
Only (41)	4623	8781	8913	886	2770	4412	22	92	600	600	600	600	0	0	3.02	1.7	3.59	4.56	0	0	0	0	0	0	
Only (42)	4623	8962	9340	890	2810	4454	23	55	600	600	600	600	0	0	2.44	1.7	3.59	4.65	0	3	8	0	0	0	
Only (43)	5110	5598	10733	744	470	3279	24	54	600	600	600	600	0	0	3.34	1.78	3.68	4.56	4	6	19	14	14	19	
Only (44)	4623	8781	11009	1119	350	761	23	45	600	600	600	600	0	0	2.37	1.14	3.54	5.22	0	0	22	16	14	16	
Only (45)	2157	4383	8913	658	1697	3772	12	39	600	600	600	600	0	0	3.68	1.15	<b>2.78</b>	4.64	5	2	11	4	10	8	
All but (41)	1965	4321	9676	357	1823	3437	26	58	600	600	600	600	0	0	<b>2.4</b>	1.1	2.84	<b>4.01</b>	7	5	52	27	44	41	
All but (42)	1965	4321	11576	354	2126	3256	17	34	600	600	600	600	0	0	<b>2.22</b>	1.1	2.84	4.73	7	5	51	27	44	42	
All but (43)	1857	4383	10212	1981	488	1883	18	36	600	600	600	600	0	0	2.89	<b>0.92</b>	3.99	4.44	6	2	32	17	20	24	
All but (44)	1965	4321	12266	2489	874	3812	18	35	600	600	600	600	0	0	2.23	1	3.18	4.24	7	5	32	22	23	25	
All but (45)	5233	5598	11964	766	4534	3862	26	40	600	600	600	600	0	0	2.33	1.31	3.96	4.76	2	6	42	25	38	36	
All	1965	4321	10102	354	2107	3456	26	58	600	600	600	600	0	0	<b>2.4</b>	1.1	2.84	<b>4.01</b>	7	5	52	27	44	41	
<b>Branch</b>																			Branches						
BS	2061	5657	10045	367	670	1035	26	59	600	600	600	600	0	0	2.4	1.2	<b>2.67</b>	3.51	1	3	4	11	16	13	
BaC	1860	4451	8805	402	1919	1562	26	61	600	600	600	600	0	0	<b>2.39</b>	1.25	<b>2.21</b>	<b>3.04</b>	7(1 <sup>b</sup> )	6(2 <sup>b</sup> )	38(6 <sup>b</sup> )	30(4 <sup>b</sup> )	37(7 <sup>b</sup> )	45(9 <sup>b</sup> )	

<sup>a</sup> Since we propose three groups of valid inequalities (41)-(45), we respectively test these cuts in our experiments.

For example, “All but (41)” means that we add inequalities (42-45) in the corresponding experiment.

<sup>b</sup> The values in parentheses represent the number of created branches.

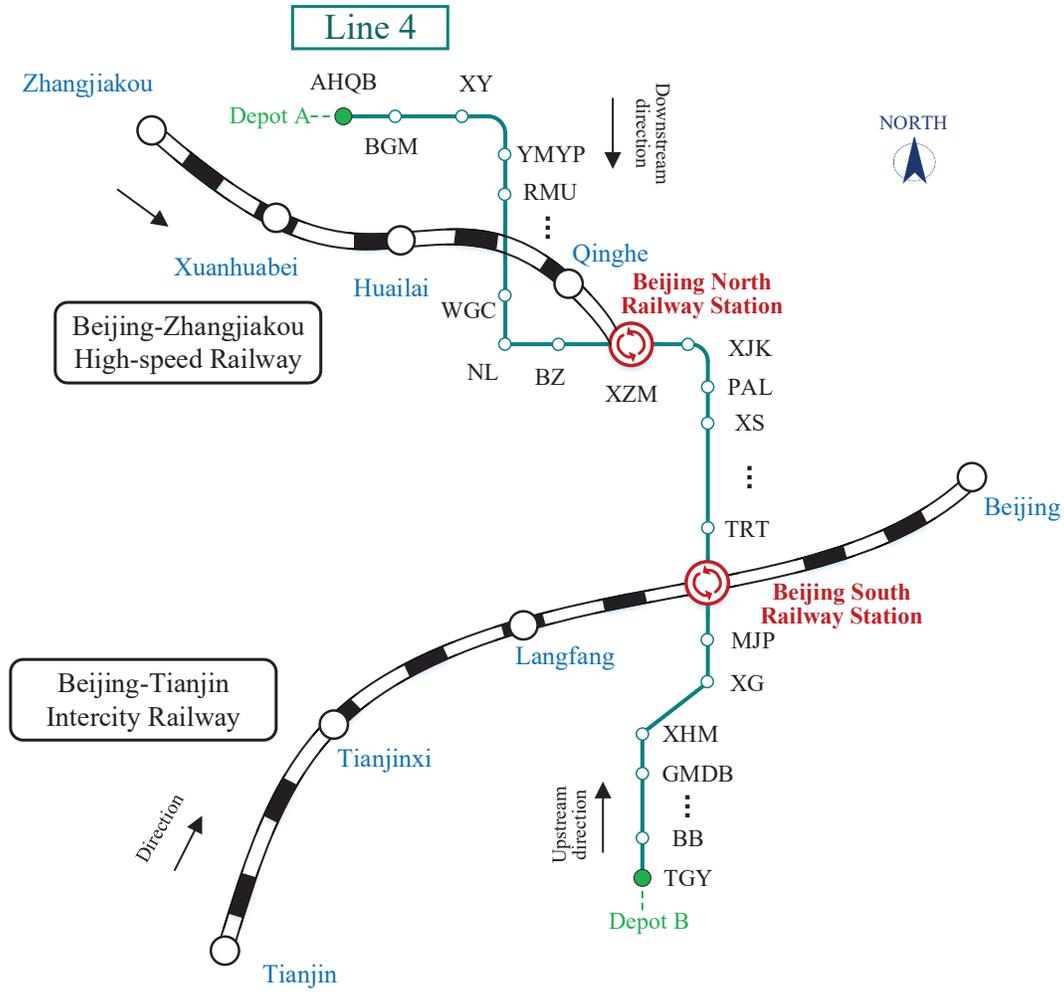


Figure 14: Layout of a railway-metro network in Beijing

flow in the upstream and downstream directions, these are considered two separate metro lines. Besides, there are two depots connected with the origin and destination stations of these two lines. Specifically, depots A and B are connected to station AHQB and station TGY, respectively, where trains can perform turnaround operations, as indicated by green dots in Figure 14. For the Beijing-Zhangjiakou high-speed railway line, we focus on five large stations, i.e., Zhangjiakou, Xuanhuabei, Huailai, Qinghe, and Beijing North Railway station, that could be used for train original departure and final arrival. For Beijing-Tianjin intercity railway line, we consider five stations including Tianjin, Tianjinxi, Langfang, Beijing South, and Beijing railway station for train departures and arrivals in our experiments.

These three lines are interconnected in the network, resulting in two transfer hubs, i.e., Beijing North Railway Station and Beijing South Railway Station. Beijing North Railway station and metro station XZM are located in the same transport hub, which connects with the Beijing-Zhangjiakou high-speed railway line and metro Line No.4, while Beijing South Railway station connects the Beijing-Tianjin intercity railway line and the metro line. Therefore, in addition to passengers traveling on the railway lines and the metro lines, we also account for passengers alighting from railway trains and transferring onto metro trains at these transfer stations in the network. For example, the passengers can travel from Zhangjiakou Railway station, alight at Beijing North Railway station, and then walk to metro station XZM to board metro trains on line No.4.

In the real-world experiments study, we construct a total of 10 testing instances with different numbers of planned metro trains during the period 7:00 am to 10:00 am. As the railway train timetable is fixed in a cycle time (e.g., three months) in daily operation, we fix the planned number of trains on both railway lines as eight during this period in our case study. In railway systems, the passenger information of the railway line is based on real data from the ticket booking system, but we have revised some of this because we only considered part of the data, and the actual passenger data is confidential. We generate 20 passenger groups, each comprising passengers according to their intended trip. The details of the passenger groups are reported in Table 13 in Appendix D. For the metro systems, we use the real-world passenger demand data collected on a weekday in January 2020 in the Beijing metro in this set of numerical experiments. The weight coefficients  $\theta_1$ ,  $\theta_2$ ,  $\theta_3$  and  $\theta_4$  in the objective function are set as 0.3, 0.07, 0.003, 0.6, respectively. The input parameters such as the coordination time window for each transfer station, the cost of using one carriage, and the minimum and maximum groups of carriages involved by each metro train are detailed in Appendix D.

As the computational results in the previous section have shown that the BaC algorithm performs the best in solving instances, we next only use our BaC algorithm to solve the constructed ten instances. In the following experiments, the computational time limit is set as 9000s.

### 5.2.2 Comparison with the non-coordinated train timetable in practice

In this subsection, we compare the coordinated train timetable with the non-coordinated timetable currently used in practice. The non-coordinated timetable is obtained by solving two independent models for railway systems (constraints (16-28)) and metro systems (constraints (1-15)), respectively, without considering the passenger transfers. Based on the solutions obtained from the train timetables of railway and metro trains, we can calculate the passenger transfer time  $F^{\text{tra}}$  analytically according to the passenger flows.

Table 6 reports the computational results for the two train timetables on the ten instances. Specifically, we present the objective function value (denoted as “Obj”), passenger transfer time at the transfer stations (i.e.,  $F^{\text{tra}}$ ), operational costs (i.e.,  $M^{\text{cost}}$ ), the computation time (denotes as “Com”), and the optimality gap (denoted as “Gap”) until the time limit. The value of Gap is the percentage gap between the best upper bound (UB) and the best lower bound (LB) computed by our BaC algorithm, i.e.,  $\text{Gap} = (\text{UB} - \text{LB})/\text{UB} \times 100\%$ . Furthermore, in the last two columns of Table 6, we respectively report the performance improvements “OPI” for Obj and “TPI” for  $F^{\text{tra}}$  with the coordinated timetable compared to the non-coordinated timetable. Here, we note that CPLEX takes about 10 minutes to obtain the solution for the non-coordinated timetables of railway and metro trains. Thus, we do not report the computation time and gap of this strategy in Table 6 for brevity.

From the results in Table 6, we see that with the increase of metro trains, the value of  $F^{\text{tra}}$  decreases, while the operational cost  $M^{\text{cost}}$  increases for the coordinated timetable. This is because higher train service frequency, which reduces the departure interval of metro trains at transfer stations, can reduce the passenger transfer time. Naturally, rail managers need to employ more carriages to perform more train services. In addition, we see that the coordinated timetable obtains much better Obj values than the non-coordinated timetable for all instances. For example, in instance 3, the coordinate timetable can reduce the Obj by 12.70% (from 477,005 to 416,445) compared to the non-coordinated timetable. In terms of the operational costs  $M^{\text{cost}}$ , we notice that the coordinated train timetable results in lower costs, i.e., fewer carriages in service, compared to the strategy without coordination (except for instances 2 and 4). For example, in instance 8, the operational cost  $M^{\text{cost}}$  decreases from 522,000 to 512,000 with the coordinated timetable, indicating cost savings for rail managers. This result further demonstrates that since the flexible capacity allocation strategy allows trains to change the carriages involved according to the specific passenger demand, the coordinated train timetable with fewer waiting passengers at transfer stations can reduce

the usage of carriages in service. Compared with the non-coordinated timetable, the most outstanding advantage of the coordinated timetable is the reduction of the passenger transfer time (i.e.,  $F^{\text{tra}}$ ). For example, in instance 4, the passenger transfer time decreases from 207,920 to 102,640, indicating that the waiting time for passengers who alight from railway trains to transfer to metro trains can be evidently reduced. From the last two columns of Table 6, we see that on average, the OPI and TPI with coordination compared with the non-coordinated timetable reach 9.05% and 40.17%, respectively, in the case study. These values demonstrate that a well-coordinated train timetable enables metro trains to depart from transfer stations to connect railway trains within a coordination time window, thereby reducing operational costs for rail managers and simultaneously enhancing service quality for passengers. Here, an interesting phenomenon is that, instead of increasing the number of planned metro trains, an appropriate number of planned metro trains can significantly reduce the passenger transfer time with the coordinated timetable compared to the non-coordinated timetable. For example, in instances 9 and 10, the value of TPI is approximately 27%, while the TPI exceeds 65% for instances 5 and 6. This observation suggests that an appropriate number of metro train services (e.g., 40 trains in our experiments) can yield more significant benefits for improving service quality for transfer passengers with the coordinated timetable, particularly in terms of operating cost savings.

Table 6: Comparison between the coordinated and non-coordinated train timetables on the real-world instances

Instance index	Metro trains	Coordinated timetable					Non-Coordinated timetable			OPI (%)	TPI (%)
		Obj	$F^{\text{tra}}$	$M^{\text{cost}}$	Com(s)	Gap(%)	Obj	$F^{\text{tra}}$	$M^{\text{cost}}$		
1	16	501953	274400	322000	4500	0	538528	412480	326000	6.79	33.48
2	20	466932	222880	342000	6120	0	482459	238400	342000	3.22	6.51
3	26	416445	113520	366000	8510	0	477005	300960	368000	12.70	62.28
4	30	407686	102640	382000	9000	3.83	436729	207920	382000	6.65	50.63
5	36	436661	113440	436000	9000	7.23	546079	357280	448000	20.04	68.25
6	40	417758	95600	420000	9000	7.97	544711	366960	424000	23.31	73.95
7	46	401882	92100	478000	9000	9.12	413405	123200	480000	2.79	25.24
8	50	415359	92920	512000	9000	9.78	430973	128560	522000	3.62	27.72
9	56	410845	87500	503000	9000	11.23	446787	120560	504000	8.04	27.42
10	60	405153	86960	502000	9000	12.56	419207	117920	504000	3.35	26.26
Average	-	428067	128196	426300	8213	6.17	473588	237424	430000	9.05	<b>40.17</b>

### 5.2.3 Experiments with different capacity allocation strategies

In this subsection, we compare the performance of three strategies for computing train capacity allocation plans. The first strategy involves no train capacity allocation (denoted as “No TCA”) in the metro systems, and the second strategy is to randomly generate a set of trains in which we allocate carriages flexibly (denoted as “Random TCA”). The third strategy is to use our optimization-based approach to allocate carriages for each train, which is denoted as “TCA-based approach”. Regarding the no TCA and random TCA strategies, we first generate a set of train capacity allocations (i.e., the number of carriages involved by each train) using these strategies. We then use each train capacity allocation plan as input to compute the performances, including the average carriage crowding degree, the average number of waiting passengers at each station, the average number of stranded passengers unable to board the first train, the maximum number of stranded passengers at all stations, and the operational costs.

Table 7 presents these performances for three strategies in the case study. The testing instances 1-5 are the same as those in Section 5.2.2. We see that for all performances, apart from the operational costs, the no TCA strategy

is the worst among the three strategies. This is expected because of the large difference between the over-saturated passenger demand and the limited train capacity. With the randomly generated train capacity allocation plan, the numbers of waiting and stranded passengers are both reduced compared to the strategy with no TCA. For example, in instance 1, the random TCA reduces the average number of stranded passengers by 13.6% (from 8,743 to 7,696) compared to no TCA. Nevertheless, the random allocation strategy requires a large number of carriages, resulting in the largest operational costs for all testing instances. From the results in Table 7, we can see that our TCA-based approach evidently outperforms other strategies in all testing instances. Although the TCA-based approach needs to employ more carriages to perform train services, the carriage crowding degree can be reduced by an average of 0.64 compared to no TCA, which can significantly enhance the passenger travel experience. Furthermore, we also notice that the TCA-based approach notably mitigates the stranded passengers to 0, i.e., all waiting passengers can board the first train at stations, highlighting the importance of an optimized train capacity allocation plan for reducing passenger travel time and operational risks. In summary, we conclude that the overall performance can be noticeably improved with our optimization-based approach.

Table 7: Comparison among the different capacity allocation strategies on the real-world instances

Performance	Strategy	Instance				
		1	2	3	4	5
Carriage crowding degree	No TCA	0.95	0.95	0.92	0.92	0.91
	Random TCA	0.82	0.82	0.81	0.79	0.8
	TCA-based approach	0.34	0.32	0.29	0.27	0.24
Average waiting passengers	No TCA	13224	9779	6599	5186	3655
	Random TCA	12426	9062	6432	4982	3234
	TCA-based approach	1798	1639	1491	1426	1355
Average stranded passengers	No TCA	8743	5792	3255	2129	1204
	Random TCA	7696	5144	2802	1313	870
	TCA-based approach	0	0	0	0	0
Maximum stranded passengers	No TCA	25489	17279	11234	8519	5543
	Random TCA	24089	15035	10781	4532	2234
	TCA-based approach	0	0	0	0	0
Operational costs	No TCA	260000	280000	302000	314000	340000
	Random TCA	336000	384000	418000	392000	452000
	TCA-based approach	322000	342000	366000	382000	436000

#### 5.2.4 Sensitivity analysis with different coordination time window and weight coefficients

In this subsection, we investigate the impact of the coordination time window and the weight coefficients in the objective function on the integrated capacity allocation and timetable coordination problem. Two sets of experiments are conducted: In the first set of experiments, we vary the coordination time window by changing the minimum and the maximum coordinated times. In the second set of experiments, we vary the weight coefficients to analyze the trade-off between the operational costs and passenger transfer time in the objective function. The other input parameter settings remain consistent with those in instance 2 of Section 5.2.2.

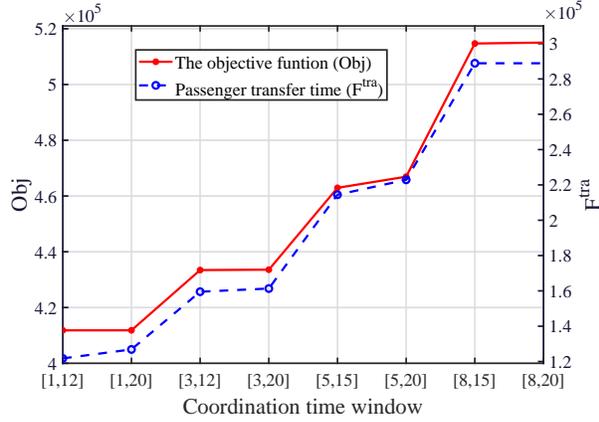
First, we investigate the impact of the coordination time window on our model. To this end, we incrementally set the coordination time window  $[\underline{L}_s^{ij}, \bar{L}_s^{ij}]$  as  $[1, 15]$ ,  $[1, 20]$ ,  $[3, 15]$ ,  $[3, 20]$ ,  $[5, 15]$ ,  $[5, 20]$ ,  $[8, 15]$  and  $[8, 20]$ . The cor-

responding results are presented in Figure 15(a). Analyzing these results, we notice that both the Obj and the  $F^{\text{tra}}$  exhibit a rising trend as the minimum and maximum coordinated times increase. For instance, when the coordination time window is  $[8, 20]$ , Obj is approximately 515,000, while the passenger transfer time  $F^{\text{tra}}$  reaches 289,000. In addition, we also notice that when only the maximum coordinated time is changed while keeping the minimum coordinated time constant, the values of Obj and  $F^{\text{tra}}$  show slight variation. For example, when the coordination time window shifts from  $[3, 12]$  to  $[3, 20]$ , the Obj remains approximately 433,500, and the  $F^{\text{tra}}$  remains unchanged at 160,000. Conversely, Obj and  $F^{\text{tra}}$  evidently increase if the minimum coordinated time is altered while keeping the maximum time constant. This demonstrates that the obtained coordinated train timetable is more dependent on the minimum coordinated time and less affected by the maximum coordinated time, as the objective function favors shorter passenger transfer times. An interesting feature of Figure 15(a) is that when the coordination time window changes, the Obj and  $F^{\text{tra}}$  exhibit similar and closely aligned trends in this set of experiments. This indicates that changes in the coordination time window primarily impact passenger transfer time, with minimal effect on the operational efficiency of individual transportation systems in multimodal transportation. This insight is very useful for rail managers to design a well-coordinated train timetable within a reasonable coordination window in the future.

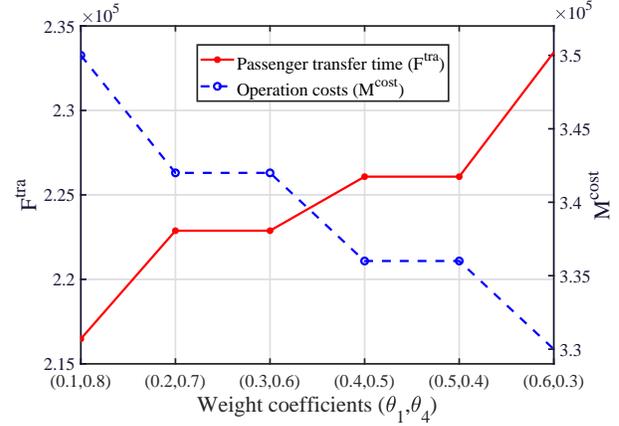
Next, we analyze the sensitivity with different weight coefficients in the objective function. In our formulation, four weight coefficients ( $\theta_1$ ,  $\theta_2$ ,  $\theta_3$ , and  $\theta_4$ ) are introduced to integrate the passenger travel time and operational costs into a single objective. Since the passenger travel time in railway and metro systems has less effect on the objective function compared to the passage transfer and operational costs according to the results in Section 5.2.2, we only vary the weight coefficients of  $\theta_1$  and  $\theta_4$  to explore the trade-off between operational costs  $M^{\text{cost}}$  and passenger transfer time  $F^{\text{tra}}$ . In this set of experiments, we sequentially set these two weight coefficients as  $\theta_1 = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6$  and  $\theta_2 = 0.8, 0.7, 0.6, 0.5, 0.4, 0.3$ . From the results in Figure 15(b), we can observe that as the weight  $\theta_4$  involving  $F^{\text{tra}}$  in the objective function decreases, passenger transfer time  $F^{\text{tra}}$  exhibits an upward trend. On the other hand, when we increase the value of  $\theta_1$ , the operational costs take a decreasing tendency in this set of experiments. These results demonstrate that when rail managers pay more attention to service quality for transfer passengers, more carriages involved in trains will be employed in services, increasing operational costs. In addition, an interesting observation is that when the weight coefficients change only slightly, the values of  $F^{\text{tra}}$  and  $M^{\text{cost}}$  remain almost constant. For example, when the weights change from  $(0.2, 0.7)$  to  $(0.3, 0.6)$ , the transfer time  $F^{\text{tra}}$  is approximately 222,880, and the operational costs are both 342,000. This is probably because small changes in weight coefficient do not evidently impact the optimal solution, i.e., the number of carriages involved in trains and the departure/arrival times of all considered trains in our experiments. In summary, the trends in the values of  $F^{\text{tra}}$  and  $M^{\text{cost}}$  suggest that trains with more carriages can enhance service quality for passengers, but employing more carriages in service inevitably increases operational costs for rail managers.

## 6 Conclusions and future research

In this paper, we studied the integrated train capacity allocation and timetable coordination in multimodal railway networks, where a part of vehicles can be dynamically allocated to different metro trains to alleviate the congestion of transfer hubs. We proposed a new mathematical model that optimizes the timetables of both metro and railway networks and the train capacity allocation strategy. The objective is to minimize the passenger travel time, passenger transfer time and operational costs for rail managers. Different from the state-of-the-art, our formulation not only incorporates internal passenger demand in railway and metro systems but also addresses the coordination of these two modes by considering transfer passenger demand at transfer stations. To tackle the computational difficulties



(a) Obj and  $F^{\text{tra}}$  with the different coordination time window



(b)  $F^{\text{tra}}$  and  $M^{\text{cost}}$  with different weight coefficients

Figure 15: Sensitivity analysis with different coordination time window and weight coefficients

arising from the binary variables related to the synchronization of multiple networks, we developed an exact branch-and-cut-based solution algorithm. The algorithm involves five sets of valid inequities and a customized branching strategy based on the model properties, to obtain high-quality solutions within a reasonable computation time. Finally, we conducted experiments on two sets of instances, including small-scale instances and real-world large-scale instances based on the field data of the Beijing railway-metro network. Our experiment results demonstrate that our proposed branch-and-cut-based approach computes optimal or near-optimal solutions and outperforms the state-of-the-art MIP solvers. Meanwhile, we also find that our approach by flexibly allocation train capacities can reduce the passenger transfer waiting time by over 40%, compared to the current non-coordinated train timetable. Additionally, several sets of experiments on sensitivity analysis were conducted to further explore the impact of key parameters in the model.

Our future work will focus on the following two key aspects. First, this study only takes a first step of integrating capacity allocation and timetable coordination for multimodal railway networks. In future research, we aim to extend this integration to more transportation modes and broader transport hubs to provide door-to-door services for passengers. Another meaningful extension is to consider a more specific rolling stock circulation plan in the multimodal railway network, to explore more practical scenarios of virtual coupling patterns in metro and railway systems.

## Acknowledgments

This research was supported by National Natural Science Foundation of China (Nos. 72288101, 71825004), by the Beijing Municipal Natural Science Foundation, China (No. 4222051), by the State Key Laboratory of Advanced Rail Autonomous Operation (No. RCS2023K008), by a grant from Department of Education of Guangdong Province (No. 2022KCXTD027), and by the Guangdong Key Construction Discipline Research Ability Enhancement Project (2021ZDJS108).

## References

Adelgren, N., Gupte, A., 2022. Branch-and-bound for biobjective mixed-integer linear programming. *INFORMS Journal on Computing* 34, 909–933.

- Barrena, E., Canca, D., Coelho, L.C., Laporte, G., 2014. Single-line rail rapid transit timetabling under dynamic passenger demand. *Transportation Research Part B: Methodological* 70, 134–150.
- Chai, S., Yin, J., D’Ariano, A., Liu, R., Yang, L., Tang, T., 2024. A branch-and-cut algorithm for scheduling train platoons in urban rail networks. *Transportation Research Part B: Methodological* 181, 102891.
- Chai, S., Yin, J., D’Ariano, A., Samà, M., Tang, T., 2023a. Scheduling of coupled train platoons for metro networks: A passenger demand-oriented approach. *Transportation Research Record* 2677, 1671–1689.
- Chai, S., Yin, J., D’Ariano, A., Samà, M., Tang, T., 2023b. Train schedule optimization for commuter-metro networks. *Transportation Research Part C: Emerging Technologies* 155, 104278.
- Coniglio, S., Furini, F., San Segundo, P., 2021. A new combinatorial branch-and-bound algorithm for the knapsack problem with conflicts. *European Journal of Operational Research* 289, 435–455.
- Di Meo, C., Di Vaio, M., Flammini, F., Nardone, R., Santini, S., Vittorini, V., 2019. Ertms/etcs virtual coupling: proof of concept and numerical analysis. *IEEE Transactions on Intelligent Transportation Systems* 21, 2545–2556.
- Fischetti, M., Ljubić, I., Monaci, M., Sinnl, M., 2017. A new general-purpose algorithm for mixed-integer bilevel linear programs. *Operations Research* 65, 1615–1637.
- Fouilhox, P., Ibarra-Rojas, O.J., Kedad-Sidhoum, S., Rios-Solis, Y.A., 2016. Valid inequalities for the synchronization bus timetabling problem. *European Journal of Operational Research* 251, 442–450.
- Gkiotsalitis, K., Cats, O., Liu, T., 2023a. A review of public transport transfer synchronisation at the real-time control phase. *Transport Reviews* 43, 88–107.
- Gkiotsalitis, K., Cats, O., Liu, T., Bult, J., 2023b. An exact optimization method for coordinating the arrival times of urban rail lines at a common corridor. *Transportation Research Part E: Logistics and Transportation Review* 178, 103265.
- Guo, J., Xiao, X., Shi, J., 2022. Dynamic carriage reserving for an over-crowded metro junction station. *Transportation Letters* 14, 464–477.
- Guo, X., Wu, J., Sun, H., Yang, X., Jin, J.G., Wang, D.Z., 2020. Scheduling synchronization in urban rail transit networks: Trade-offs between transfer passenger and last train operation. *Transportation Research Part A: Policy and Practice* 138, 463–490.
- Huang, K., Wu, J., Liao, F., Sun, H., He, F., Gao, Z., 2021. Incorporating multimodal coordination into timetabling optimization of the last trains in an urban railway network. *Transportation Research Part C: Emerging Technologies* 124, 102889.
- Huang, Y., Mannino, C., Yang, L., Tang, T., 2020. Coupling time-indexed and big-m formulations for real-time train scheduling during metro service disruptions. *Transportation Research Part B: Methodological* 133, 38–61.
- Kang, L., Zhu, X., Sun, H., Puchinger, J., Ruthmair, M., Hu, B., 2016. Modeling the first train timetabling problem with minimal missed trains and synchronization time differences in subway networks. *Transportation Research Part B: Methodological* 93, 17–36.
- Ke, Y., Wu, X., Nie, L., Yao, Z., Chen, Y., 2024. Synchronizing train, aircraft, shuttle, and passenger flows in intermodal timetabling: A time–space network-based formulation and a decomposition algorithm using alternating direction method of multipliers. *Transportation Research Part C: Emerging Technologies* 159, 104464.

- Liu, T., Cats, O., Gkiotsalitis, K., 2021a. A review of public transport transfer coordination at the tactical planning phase. *Transportation Research Part C: Emerging Technologies* 133, 103450.
- Liu, X., Dabiri, A., Wang, Y., De Schutter, B., 2023a. Real-time train scheduling with uncertain passenger flows: A scenario-based distributed model predictive control approach. *IEEE Transactions on Intelligent Transportation Systems* .
- Liu, X., Dabiri, A., Xun, J., De Schutter, B., 2023b. Bi-level model predictive control for metro networks: Integration of timetables, passenger flows, and train speed profiles. *Transportation Research Part E: Logistics and Transportation Review* 180, 103339.
- Liu, Y., Zhou, Y., Su, S., Xun, J., Tang, T., 2021b. An analytical optimal control approach for virtually coupled high-speed trains with local and string stability. *Transportation Research Part C: Emerging Technologies* 125, 102886.
- Long, S., Meng, L., Miao, J., Hong, X., Corman, F., 2020. Synchronizing last trains of urban rail transit system to better serve passengers from late night trains of high-speed railway lines. *Networks and Spatial Economics* 20, 599–633.
- Lu, Y., Yang, L., Yang, H., Zhou, H., Gao, Z., 2023. Robust collaborative passenger flow control on a congested metro line: A joint optimization with train timetabling. *Transportation Research Part B: Methodological* 168, 27–55.
- Ning, J., Peng, Q., Zhu, Y., Xing, X., Nielsen, O.A., 2023. Bi-objective optimization of last-train timetabling with multimodal coordination in urban transportation. *Transportation Research Part C: Emerging Technologies* 154, 104260.
- Niu, H., Zhou, X., 2013. Optimizing urban rail timetable under time-dependent demand and oversaturated conditions. *Transportation Research Part C: Emerging Technologies* 36, 212–230.
- OECD, 2020. Safe and seamless travel and improved traveller experience URL: <https://www.oecd-ilibrary.org/content/paper/d717f6ea-en>, doi:<https://doi.org/https://doi.org/10.1787/d717f6ea-en>.
- Pan, H., Yang, L., Liang, Z., 2023. Demand-oriented integration optimization of train timetabling and rolling stock circulation planning with flexible train compositions: A column-generation-based approach. *European Journal of Operational Research* 305, 184–206.
- Quaglietta, E., Wang, M., Goverde, R.M., 2020. A multi-state train-following model for the analysis of virtual coupling railway operations. *Journal of Rail Transport Planning & Management* 15, 100195.
- Shi, J., Qin, T., Yang, L., Xiao, X., Guo, J., Shen, Y., Zhou, H., 2022. Flexible train capacity allocation for an overcrowded metro line: A new passenger flow control approach. *Transportation Research Part C: Emerging Technologies* 140, 103676.
- Shi, X., Chen, Z., Pei, M., Li, X., 2020. Variable-capacity operations with modular transits for shared-use corridors. *Transportation Research Record* 2674, 230–244.
- Shi, X., Li, X., 2021. Operations design of modular vehicles on an oversaturated corridor with first-in, first-out passenger queueing. *Transportation Science* 55, 1187–1205.

- Shift2Rail, 2020. Shift2rail joint undertaking multi-annual action plan. <https://shift2rail.org/wp-content/uploads/2019/05/Draft-Shift2Rail-Multi-Annual-Action-Plan-Part-B-20.5.2019.pdf>.
- Siemens, 2023. Siemens mobility 2023 URL: <https://www.mobility.siemens.com/global/en/portfolio/digital-solutions-software/mobility-as-a-service.html>.
- Veelenturf, L.P., Kidd, M.P., Cacchiani, V., Kroon, L.G., Toth, P., 2016. A railway timetable rescheduling approach for handling large-scale disruptions. *Transportation Science* 50, 841–862.
- Wang, E., Yang, L., Li, P., Zhang, C., Gao, Z., 2023a. Joint optimization of train scheduling and routing in a coupled multi-resolution space–time railway network. *Transportation Research Part C: Emerging Technologies* 147, 103994.
- Wang, F., Xu, R., Song, X., Wang, P., 2023b. Collaborative optimization of last-train timetables for metro network to increase service time for passengers. *Computers & Operations Research* 151, 106091.
- Wang, X., Lv, Y., Sun, H., Xu, G., Qu, Y., Wu, J., 2023c. A simulation-based metro train scheduling optimization incorporating multimodal coordination and flexible routing plans. *Transportation Research Part C: Emerging Technologies* 146, 103964.
- Wang, Y., D’Ariano, A., Yin, J., Meng, L., Tang, T., Ning, B., 2018. Passenger demand oriented train scheduling and rolling stock circulation planning for an urban rail transit line. *Transportation Research Part B: Methodological* 118, 193–227.
- Wang, Y., Li, D., Cao, Z., 2020. Integrated timetable synchronization optimization with capacity constraint under time-dependent demand for a rail transit network. *Computers & Industrial Engineering* 142, 106374.
- Williams, H.P., 2013. *Model building in mathematical programming*. John Wiley & Sons.
- Wolsey, L.A., 2020. *Integer programming*. John Wiley & Sons.
- Wu, J., Liu, M., Sun, H., Li, T., Gao, Z., Wang, D.Z., 2015. Equity-based timetable synchronization optimization in urban subway network. *Transportation Research Part C: Emerging Technologies* 51, 1–18.
- Yang, L., Qi, J., Li, S., Gao, Y., 2016. Collaborative optimization for train scheduling and train stop planning on high-speed railways. *Omega* 64, 57–76.
- Yin, J., D’Ariano, A., Wang, Y., Yang, L., Tang, T., 2021. Timetable coordination in a rail transit network with time-dependent passenger demand. *European Journal of Operational Research* 295, 183–202.
- Yin, J., Wang, M., D’Ariano, A., Zhang, J., Yang, L., 2023. Synchronization of train timetables in an urban rail network: A bi-objective optimization approach. *Transportation Research Part E: Logistics and Transportation Review* 174, 103142.
- Yin, J., Yang, L., Tang, T., Gao, Z., Ran, B., 2017. Dynamic passenger demand oriented metro train scheduling with energy-efficiency and waiting time minimization: Mixed-integer linear programming approaches. *Transportation Research Part B: Methodological* 97, 182–213.
- Yuan, Y., Li, S., Liu, R., Yang, L., Gao, Z., 2023. Decomposition and approximate dynamic programming approach to optimization of train timetable and skip-stop plan for metro networks. *Transportation Research Part C: Emerging Technologies* 157, 104393.

- Zhan, S., Wong, S., Shang, P., Peng, Q., Xie, J., Lo, S., 2021. Integrated railway timetable rescheduling and dynamic passenger routing during a complete blockage. *Transportation Research Part B: Methodological* 143, 86–123.
- Zhang, C., Gao, Y., Yang, L., Gao, Z., Qi, J., 2020. Joint optimization of train scheduling and maintenance planning in a railway network: A heuristic algorithm using lagrangian relaxation. *Transportation Research Part B: Methodological* 134, 64–92.
- Zhang, L., Wang, S., Qu, X., 2021. Optimal electric bus fleet scheduling considering battery degradation and non-linear charging profile. *Transportation Research Part E: Logistics and Transportation Review* 154, 102445.
- Zhou, H., Qi, J., Yang, L., Shi, J., Pan, H., Gao, Y., 2022. Joint optimization of train timetabling and rolling stock circulation planning: A novel flexible train composition mode. *Transportation Research Part B: Methodological* 162, 352–385.
- Zhu, Y., Goverde, R.M., 2019. Railway timetable rescheduling with flexible stopping and flexible short-turning during disruptions. *Transportation Research Part B: Methodological* 123, 149–181.

# Appendix A. Sets, parameters and variables for the model formulation

Table 8 lists all the sets, parameters, and subscripts used in the model formulation. Table 9 lists all the decision variables defined in the model formulation.

Table 8: Sets, parameters and subscripts for the model formulation

Notations	Detailed definition
<b>Sets</b>	
$\mathcal{I} = \{1, 2, \dots, I\}$	Set of railway lines
$\mathcal{J} = \{1, 2, \dots, J\}$	Set of metro lines
$\mathcal{D} = \{1, 2, \dots, D\}$	Set of depots in metro systems
$\mathcal{T} = \{1, 2, \dots, T\}$	Set of timestamps in the discretized time horizon, indexed by $\tau$
$\mathcal{K}_j = \{1, 2, \dots, K_j\}$	Set of trains on line $j$ in metro systems
$\mathcal{S}_j$	Set of stations of line $j$ in metro systems
$\mathcal{U}_i$	Set of stations of line $i$ in railway systems
$\mathcal{J}_d$	Set of metro lines served by the rolling stock from depot $d$
$\mathcal{P}_i$	Set of passenger groups on railway line $i$
$\mathcal{V}_{i,j}$	Set of transfer stations connected with railway line $i$ and metro line $j$
$E$	Set of events
$A$	Set of activities
$E_{iu}^{\text{dep}}/E_{iu}^{\text{arr}}$	Subset of events corresponding to the departures/arrivals from/at station $u$
$E_i^{\text{cr}}$	Set of events that can be slightly adjusted on railway line $i$
$A_j^{\text{train}}$	Set of train activities including running activities and dwell activities on railway line $j$
$A_i^{\text{run}}/A_i^{\text{dw}}/A_i^{\text{track}}$	Set of train running/dwell/track headway activities on railway line $i$
$B_i$	Set of all order activity pairs on railway line $i$ , $(a, a') \in \mathcal{B}_i$
<b>Input parameters</b>	
$T_{k,s,j}^{\text{run}}$	The running time of train $k$ from station $s$ to $s+1$ on line $j$
$T_{k,s,j}^{\text{dwell}}$	The dwelling time of train $k$ at station $s$ on line $j$
$T_a^{\text{run}}/T_a^{\text{dwe}}$	The running/dwell time for running activity $a$
$N_p$	The number of passengers in group $p$
$R_s^{i,j}$	Transferring rate of passengers from railway line $i$ to metro line $j$ at station $s$ , $R_s^{i,j} \in [0, 1]$
$D^{\text{met}}$	The loading passenger capacity of a carriage in metro systems
$D^{\text{rail}}$	The loading passenger capacity of a railway train
$Z_d^{\text{max}}$	The maximum capacity of depot $d$ to allocate rolling stock
$Q_s^j$	Alighting rate of onboard remaining passengers at station $s$ on metro line $j$ , $Q_s^j \in [0, 1]$
$T_p$	The intended departure time of passengers in group $p$
$\underline{L}_s^{i,j}/\bar{L}_s^{i,j}$	The minimum/maximum coordinated time at station $s$ between railway line $i$ and metro line $j$
$T^{\text{unit}}$	The unit length of discretized time intervals
$h_j^{\text{min}}/h_j^{\text{max}}$	The minimum/maximum headway time between two successive trains from stations in metro line $j$
$h_a^{\text{min}}$	The minimum headway time for track headway activity $a$
$c_j^{\text{min}}/c_j^{\text{max}}$	The minimum/maximum number of carriages in service train on metro line $j$
$n_{s,\tau,j}^{\text{arr}}$	The number of arrival passengers at station $s$ and timestamp $\tau$ in metro line $i$
$o_p/d_p$	Origin station and destination station of passenger group $p$
$q_e^i$	The planned beginning time of event $e$ on railway line $i$
$g$	The cost of one carriage involved by metro trains

## Appendix B. Linearization of nonlinear formulations

Table 9: The decision variables for the model formulation

Notations	Detailed definition
<b>Auxiliary variables</b>	
$b_{k,s,\tau}^j$	Determines if the timestamp $\tau$ is between trains $k - 1$ and $k$ at station $s$ on line $j$
$d_{k,s,\tau}^j$	Determines if train $k$ departs from station $s$ before or at timestamp $\tau$ on line $j$
$nw_{k,s}^j$	The number of passengers waiting for train $k$ at station $s$ in line $j$
$nb_{k,s}^j$	The number of passengers boarding on train $k$ at station $s$ in line $j$
$nc_{k,s}^j$	The residual loading capacity of train $k$ at station $s$ in line $j$
$na_{k,s}^j$	The number of passengers alighting from train $k$ at station $s$ in line $j$
$nr_{k,s}^j$	The number of passengers remaining in train $k$ after departing from station $s$ in line $j$
<b>Main variables</b>	
$w_{k,s}^j$	The departure time of train $k$ at station $s$ in metro line $j$
$c_k^j$	The number of carriages involved by train $k$ on metro line $j$
$t_e^i$	<b>The beginning time of event <math>e</math> on railway line <math>i</math></b>
$\lambda_a^i$	Determines if event $e$ takes place before event $f$ in railway line $i$ , where $a = (e, f)$
$r_{p,e}^i$	Determines if passengers in group $p$ choose train $l_e$ in railway line $i$
$v_d^j$	The fleet size of rolling stock assigned to line $j$ from depot $d$
$x_{k,e,s}^{ij}$	Determines if the time difference between departure of train $k$ and event $e$ at station $s$ is larger than $\underline{L}_s^{ij}$
$y_{k,e,s}^{ij}$	Determines if train $k$ is the first available connecting train for arrival train $l_e$ at station $s$
$z_{k,e,s}^{ij}$	Determines if the time difference between departure of train $k$ and event $e$ at station $s$ is less than $\bar{L}_s^{ij}$
$\varphi_{k,e,s}^{ij}$	Determines if train $l_e$ on line $i$ and train $k$ on line $j$ are coordinated at transfer station $s$

**Lemma 3.1.** For constraints (12), the nonlinear term  $r_{p,e}^i \cdot \varphi_{k,e,s}^{ij}$  is a multiplication of two binary variables. By introducing an auxiliary variable  $\beta_{peks}^{ij}$ , constraints (12) can be rewritten as the following linear form:

$$nw_{k,s}^j = \begin{cases} \sum_{\tau \in \mathcal{T}} b_{k,s,\tau}^j \cdot n_{s,\tau,j}^{\text{arr}} + \sum_{e \in E_{is}^{\text{arr}}} \sum_{p \in \mathcal{P}_i} N_p \cdot \beta_{peks}^{ij} \cdot R_s^{ij}, & \forall s \in \mathcal{V}_{ij}, k \in \mathcal{K}_j, i \in \mathcal{I}, j \in \mathcal{J} \\ \sum_{\tau \in \mathcal{T}} b_{k,s,\tau}^j \cdot n_{s,\tau,j}^{\text{arr}}, & \forall s \in \mathcal{S}_j, k \in \mathcal{K}_j, j \in \mathcal{J} \end{cases} \quad (49)$$

$$\begin{cases} \beta_{peks}^{ij} \leq r_{p,e}^i, \\ \beta_{peks}^{ij} \leq \varphi_{k,e,s}^{ij}, & \forall p \in \mathcal{P}_i, e \in E_{is}^{\text{arr}}, s \in \mathcal{V}_{ij}, k \in \mathcal{K}_j, i \in \mathcal{I}, j \in \mathcal{J} \\ \beta_{peks}^{ij} \geq r_{p,e}^i + \varphi_{k,e,s}^{ij} - 1, \\ \beta_{peks}^{ij} \in \{0, 1\}. \end{cases} \quad (50)$$

**Lemma 3.2.** For constraints (34), the nonlinear expression is a multiplication of two binary variables. We can replace these nonlinear constraints using the following linear form:

$$\begin{cases} \varphi_{k,e,s}^{ij} \leq y_{k,e,s}^{ij}, & \forall e \in E_{is}^{\text{arr}}, k \in \mathcal{K}_j, s \in \mathcal{V}_{ij}, i \in \mathcal{I}, j \in \mathcal{J} \\ \varphi_{k,e,s}^{ij} \leq z_{k,e,s}^{ij}, & \forall e \in E_{is}^{\text{arr}}, k \in \mathcal{K}_j, s \in \mathcal{V}_{ij}, i \in \mathcal{I}, j \in \mathcal{J} \\ \varphi_{k,e,s}^{ij} \geq y_{k,e,s}^{ij} + z_{k,e,s}^{ij} - 1. & \forall e \in E_{is}^{\text{arr}}, k \in \mathcal{K}_j, s \in \mathcal{V}_{ij}, i \in \mathcal{I}, j \in \mathcal{J} \end{cases} \quad (51)$$

**Lemma 3.3.** For constraints (38), the nonlinear term  $t_e^i \cdot r_{p,e}^i$  is a integer variable multiplied by a binary variable.

By introducing an auxiliary variable  $\gamma_{pe}^i = a_e^i \cdot r_{p,e}^i$ , constraints (36) can be rewritten as the following linear form:

$$F^{\text{rail}} = \sum_{i \in \mathcal{I}} \sum_{p \in \mathcal{P}_i} \sum_{e \in E_{id_p}^{\text{arr}}} (\gamma_{pe}^i - T_p) \cdot N_p \quad (52)$$

$$\begin{cases} \gamma_{pe}^i \leq M_6 \cdot r_{p,e}^i, \\ \gamma_{pe}^i \leq t_e^i, \\ \gamma_{pe}^i \geq t_e^i - M_6(1 - r_{p,e}^i), \\ \gamma_{pe}^i \geq 0. \end{cases} \quad \forall p \in \mathcal{P}_i, e \in E_{id_p}^{\text{arr}}, i \in \mathcal{I} \quad (53)$$

**Lemma 3.4.** Based on the linearization process in constraints (50), we can replace the nonlinear term  $r_{p,e}^i \cdot \varphi_{k,e,s}^{ij} \cdot (w_{k,s}^j - t_e^i)$  in constraints (39) by introducing two auxiliary variables  $\hat{\mu}_{peks}^{ij} = \beta_{peks}^{ij} \cdot z_{ks}^j$  and  $\check{\mu}_{peks}^{ij} = \beta_{peks}^{ij} \cdot t_e^i$ . Thus, constraints (39) can be rewritten as the following linear form:

$$F^{\text{tra}} = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}_j} \sum_{s \in \mathcal{V}_{ij}} \sum_{p \in \mathcal{P}_i} \sum_{e \in E_{is}^{\text{arr}}} N_p \cdot (\hat{\mu}_{peks}^{ij} - \check{\mu}_{peks}^{ij}) \cdot R_s^{ij} \quad (54)$$

$$\begin{cases} \hat{\mu}_{peks}^{ij} \leq M_6 \beta_{peks}^{ij}, \\ \hat{\mu}_{peks}^{ij} \leq w_{k,s}^j, \\ \hat{\mu}_{peks}^{ij} \geq w_{k,s}^j - M_6(1 - \beta_{peks}^{ij}), \\ \hat{\mu}_{peks}^{ij} \geq 0. \end{cases} \quad \forall p \in \mathcal{P}_i, e \in E_{is}^{\text{arr}}, s \in \mathcal{V}_{ij}, k \in \mathcal{K}, i \in \mathcal{I}, j \in \mathcal{J} \quad (55)$$

$$\begin{cases} \check{\mu}_{peks}^{ij} \leq M_6 \beta_{peks}^{ij}, \\ \check{\mu}_{peks}^{ij} \leq t_e^i, \\ \check{\mu}_{peks}^{ij} \geq t_e^i - M_6(1 - \beta_{peks}^{ij}), \\ \check{\mu}_{peks}^{ij} \geq 0. \end{cases} \quad \forall p \in \mathcal{P}_i, e \in E_{is}^{\text{arr}}, s \in \mathcal{V}_{ij}, k \in \mathcal{K}, i \in \mathcal{I}, j \in \mathcal{J} \quad (56)$$

## Appendix C. The computational results for instance 1 in the small-scale case study

Table 10 reports the computational results of the obtained train timetables in railway systems and the coordination strategies, i.e., the connected metro trains, for instance 1 in the small case study. Table 11 reports the computational results of the obtained train timetables in metro systems and the capacity allocation strategies, i.e., the number of carriages involved, for instance 1 in the small case study.

## Appendix D. The input parameters in the real-world case study

Table 12 reports the input parameters associated with the real-world case study. In the real-world case study, the coordination time window between two railway lines and two metro lines at two transfer stations (i.e., Beijing North Railway Station and Beijing South Railway Station) is the same and equal to [5,20]. Besides, we set the minimum and maximum numbers of carriages involved by each metro train as 6 and 10, respectively. **In our study, we do not consider the maximum time headway between trains in railway systems. This is because the objective function for railway systems favors less passenger travel time. As a result, trains typically tend to be scheduled to**

Table 10: The computational results for railway train timetables and coordination strategies for instance 1 in the small-scale case study

Railway train $l_e$	Departure times $t_e^i, e \in E_{iu}^{dep}$ (min)			Arrival times $t_e^i, e \in E_{iu}^{arr}$ (min)			Loading Passenger group $p$	Coordinated metro train $k$	
	Station A	Station B	Station C	Station A	Station B	Station C		Upstream	Downstream
1	0	13	26	10	23	36	2,7,14	2	2
2	5	18	31	15	28	41	1,3,11,13	3	3
3	10	23	36	20	33	46	9,15	4	4
4	15	28	41	25	38	51	4,8,10	5	5
5	20	33	46	30	43	56	5,6,17	6	6
6	25	38	72	35	48	82	12	9	9
7	30	43	68	40	53	78	19	9	9
8	35	48	65	45	58	75	16	8	8
9	40	53	87	50	63	97	20	10	10
10	45	58	83	55	68	93	18	10	10

Table 11: The computational results for metro train timetables and capacity allocation strategies for instance 1 in the small-scale case study

Metro train $k$	Departure times $w_{k,s}^j$ (min)					The number of carriages $c_k^j$	
	Station 1(6)	Station 2(7)	Station 3(8)	Station 4(9)	Station 5(10)		
1	11	15	19	23	27	2	
2	25	29	33	37	41	2	
3	30	34	38	42	46	1	
4	35	39	43	47	51	1	
Upstream direction	5	40	44	48	52	56	1
	6	45	49	53	57	61	1
	7	53	57	61	65	69	1
	8	64	68	72	76	80	2
	9	71	75	79	83	87	2
	10	86	90	94	98	102	3
.....							
	1	22	26	30	34	38	3
	2	33	37	41	45	49	2
	3	38	42	46	50	54	1
	4	43	47	51	55	59	1
Downstream direction	5	48	52	56	60	64	1
	6	53	57	61	65	69	1
	7	60	64	68	72	76	1
	8	72	76	80	84	88	1
	9	79	83	87	91	95	2
	10	94	98	102	106	110	3

depart from each station at the minimum time headway, aiming to minimize overall passenger travel time. Table 13 provides the detailed information of passenger groups in railway systems for the real-world case study.

Table 12: The input parameters in the real-world case study

Parameters	Notations	Values	Unit
Loading passenger capacity of a railway train	$D^{\text{rail}}$	2400	person
Loading passenger capacity of a metro train	$D^{\text{met}}$	400	person
Alighting rate	$Q_s^j$	0.2 0.6	
Passenger transfer rate	$R_s^{ij}$	{0.3, 0.5, 0.5, 0.3}	
Metro train dwell time	$T_{k,s,j}^{\text{dwell}}$	0.5	minutes
Metro train running time	$T_{k,s,j}^{\text{run}}$	2	minutes
Coordination time window	$[\underline{L}_s^{ij}, \bar{L}_s^{ij}]$	[5,20]	minutes
Metro time headway	$h_j^{\min}, h_j^{\max}$	3,20	minutes
Railway time headway	$h_a^{\min}$	3	minutes
The minimum/maximum number of carriages	$c_j^{\min}, c_j^{\max}$	6,10	
Cost of using one carriage	$g$	2000	
The maximum rolling stock allocated in depot	$Z_d^{\max}$	150	

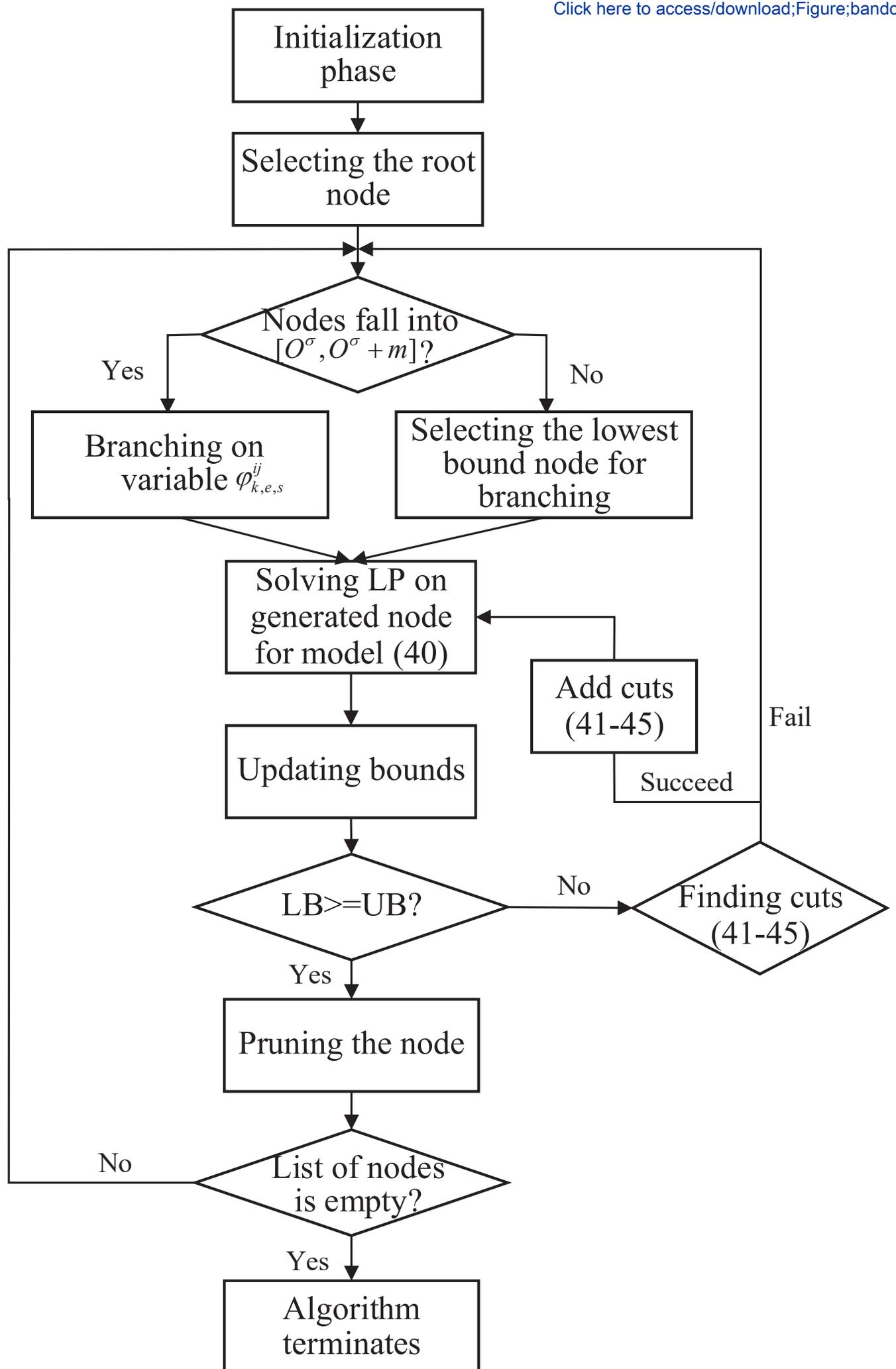
Table 13: The information of passenger groups in the real-world case study

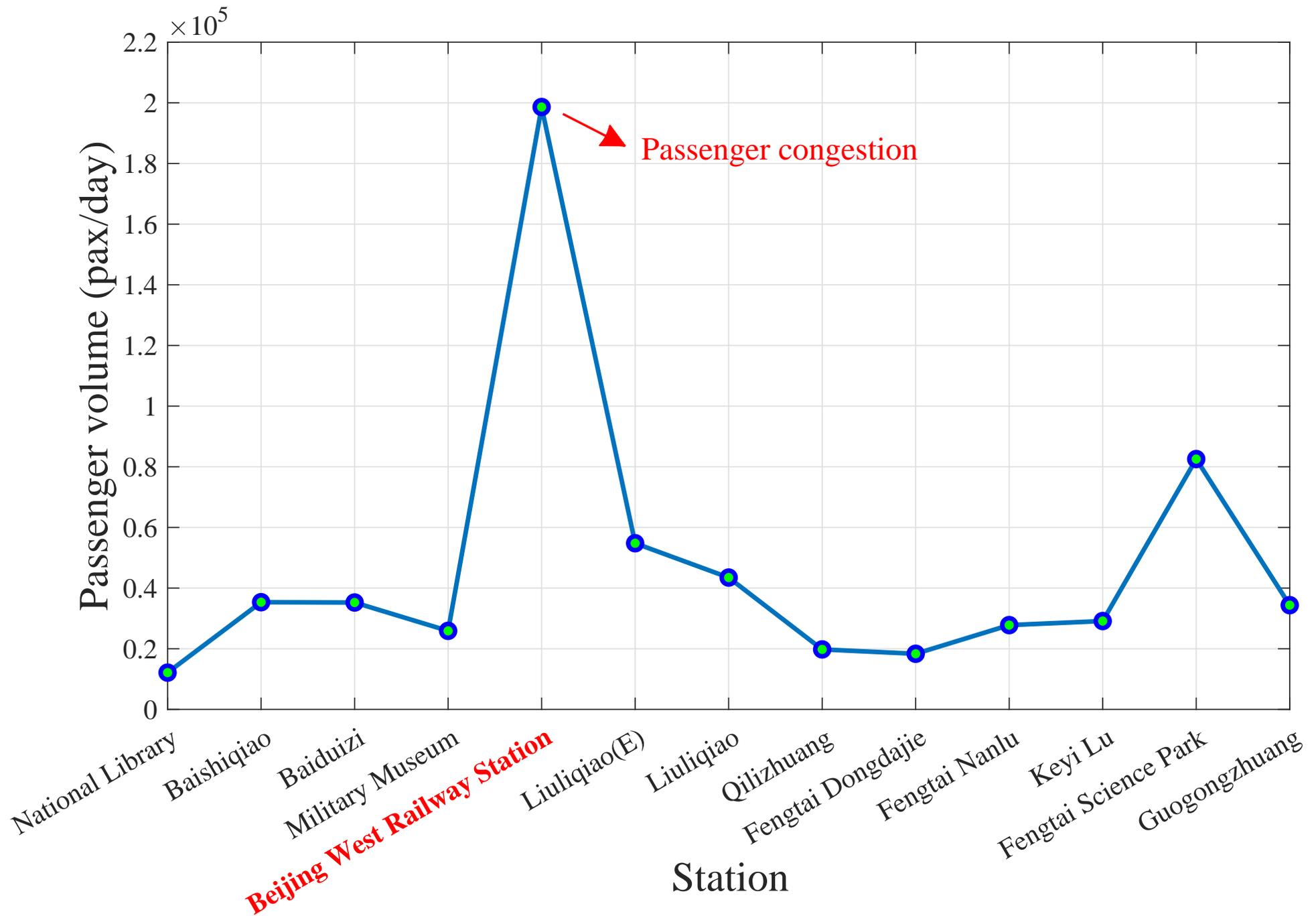
Passenger group $p$	Origin station $o_p$	Destination station $d_p$	Intended departure time $T_p$ (min)	Volume $N_p$
<b>Beijing-Zhangjiakou High-speed Railway</b>				
1	ZJK	HL	0	900
2	ZJK	QH	5	800
3	ZJK	QH	8	1000
4	ZJK	QH	10	800
5	ZJK	QH	13	800
6	ZJK	BJN	9	600
7	ZJK	BJN	12	600
8	ZJK	BJN	16	800
9	ZJK	BJN	23	800
10	ZJK	BJN	28	1200
11	ZJK	BJN	31	400
12	ZJK	BJN	34	800
13	XHB	QH	10	600
14	XHB	QH	18	600
15	XHB	QH	25	800
16	XHB	BJN	29	700
17	XHB	BJN	36	700
18	XHB	BJN	40	800
19	HL	QH	26	800
20	HL	BJN	46	900
<b>Beijing-Tianjin Intercity Railway</b>				
1	TJ	LF	0	800
2	TJ	LF	5	800
3	TJ	BJS	8	800
4	TJ	BJS	10	1000
5	TJ	BJS	13	800
6	TJ	BJS	20	400
7	TJ	BJS	32	800
8	TJ	BJ	9	600
9	TJ	BJ	17	800
10	TJX	LF	10	600
11	TJX	BJS	17	400
12	TJX	BJS	23	800
13	TJX	BJS	27	1400
14	TJX	BJS	36	800
15	TJX	BJS	39	1000
16	TJX	BJ	46	800
17	TJX	BJ	48	800
18	LF	BJS	44	800
19	LF	BJS	49	1000
20	LF	BJ	56	800

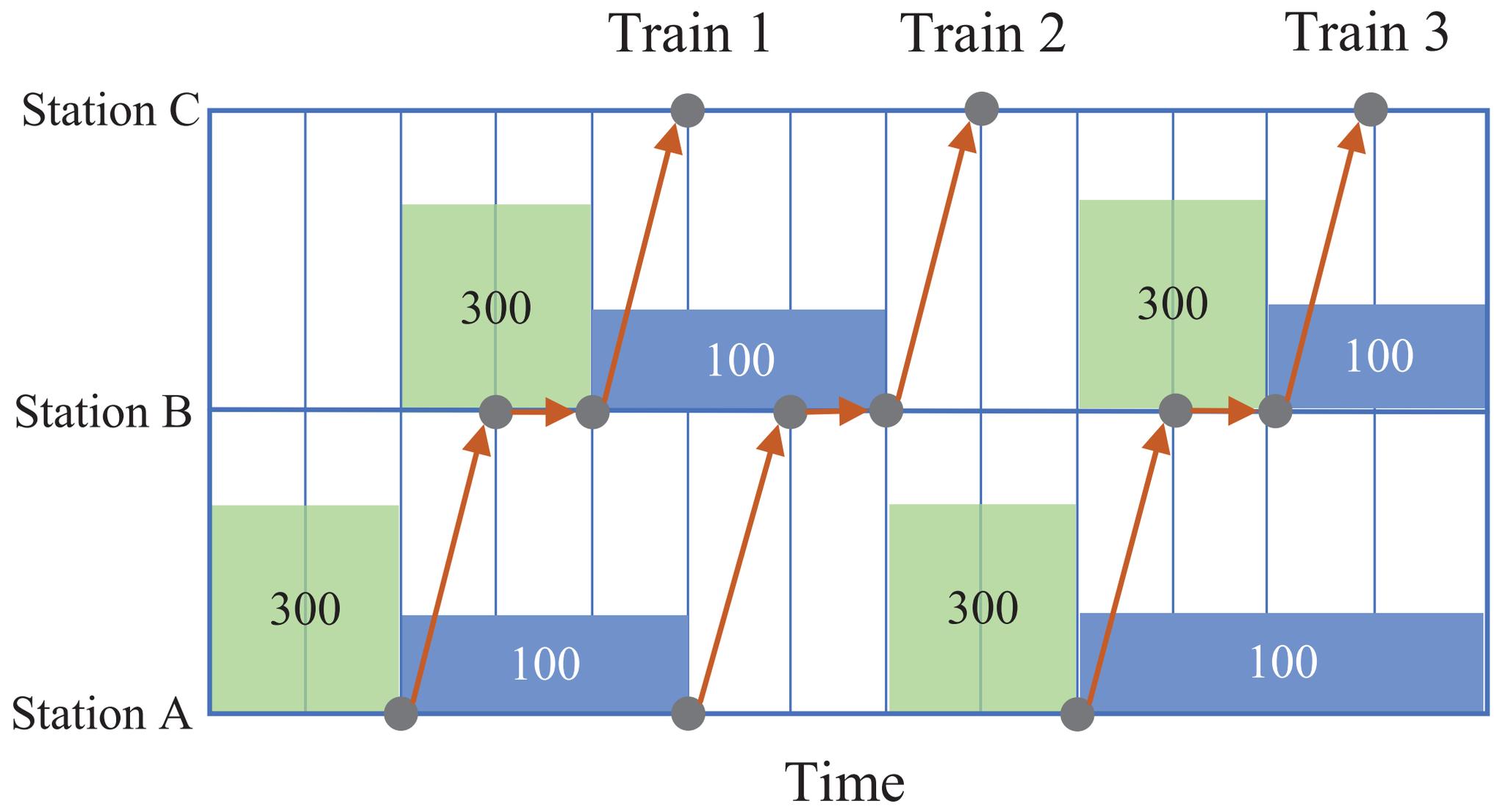


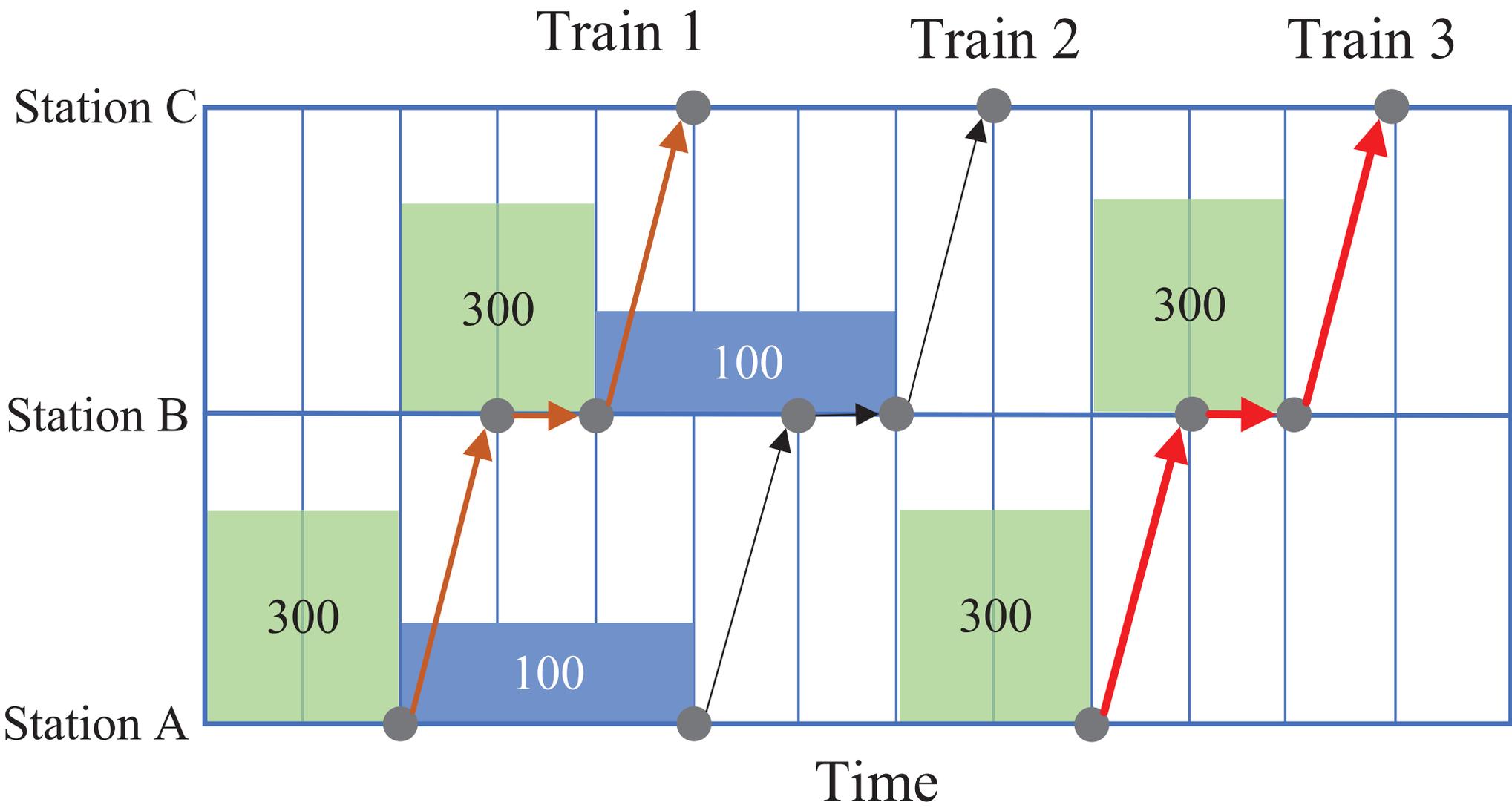
[Click here to access/download](#)

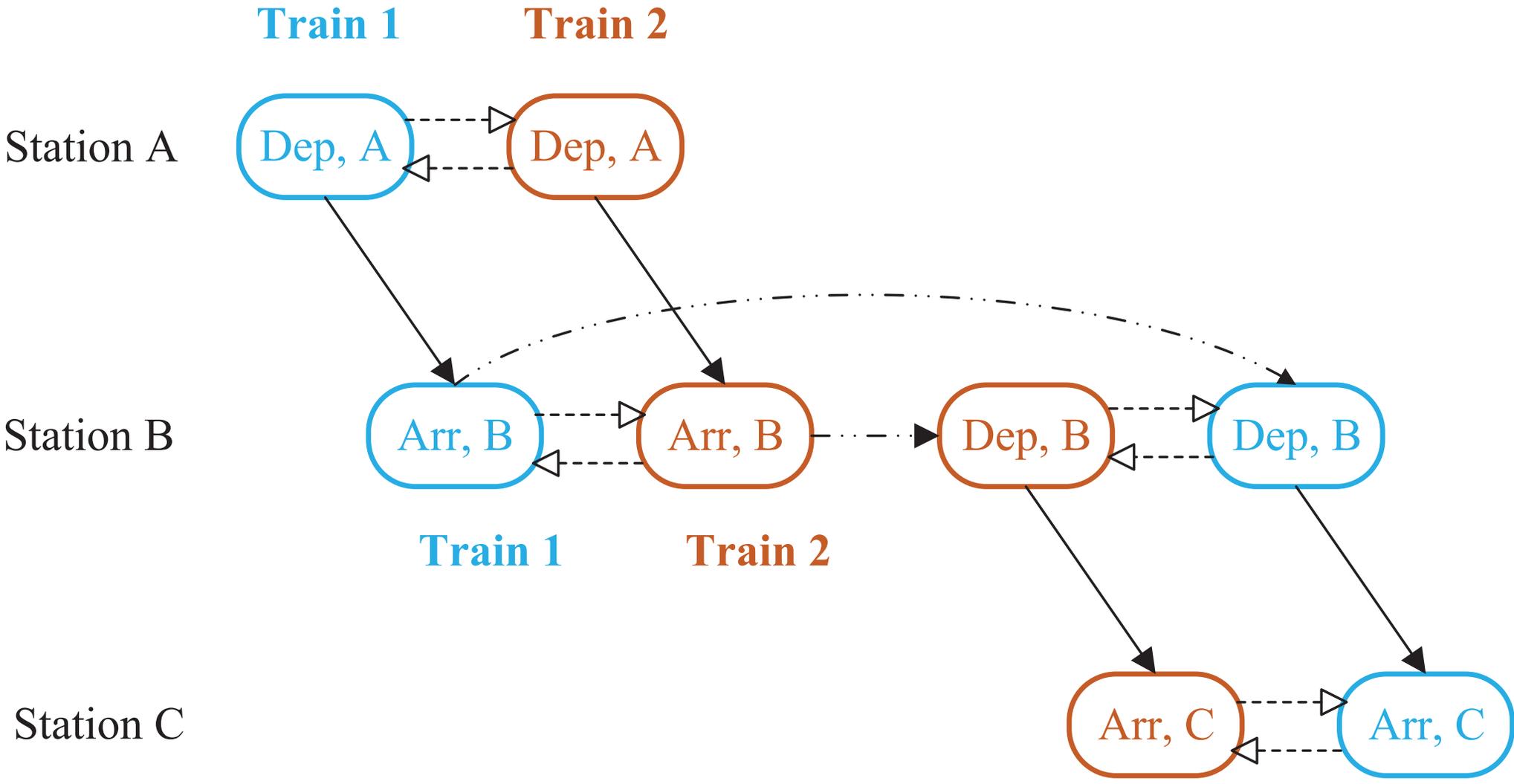
**LaTeX Source File**  
TS-template.tex





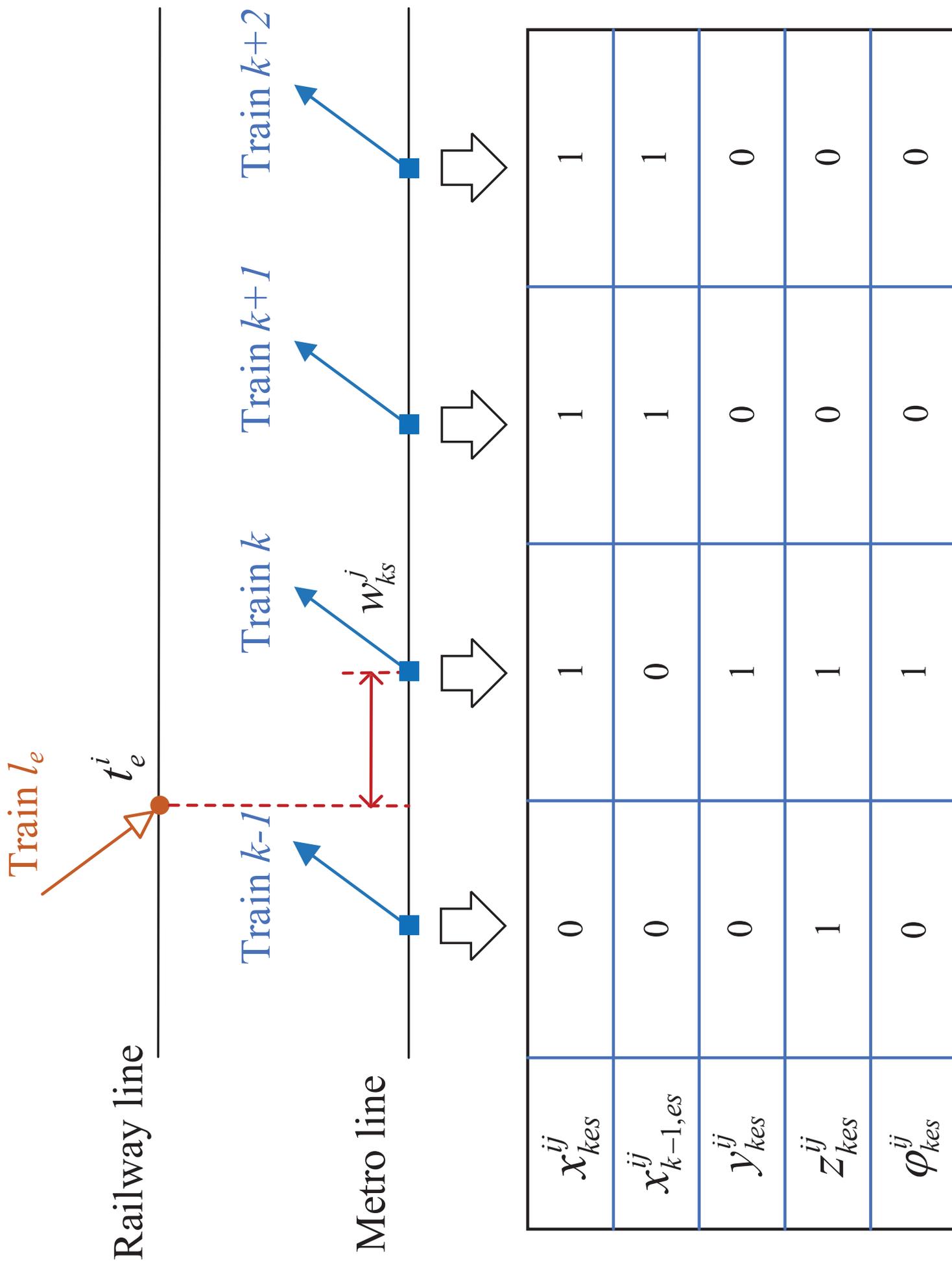






→ Running activity    - · - · → Dwell activity    - - - - ▷ Headway activity

**Dep, A** Departure event      **Arr, B** Arrival event



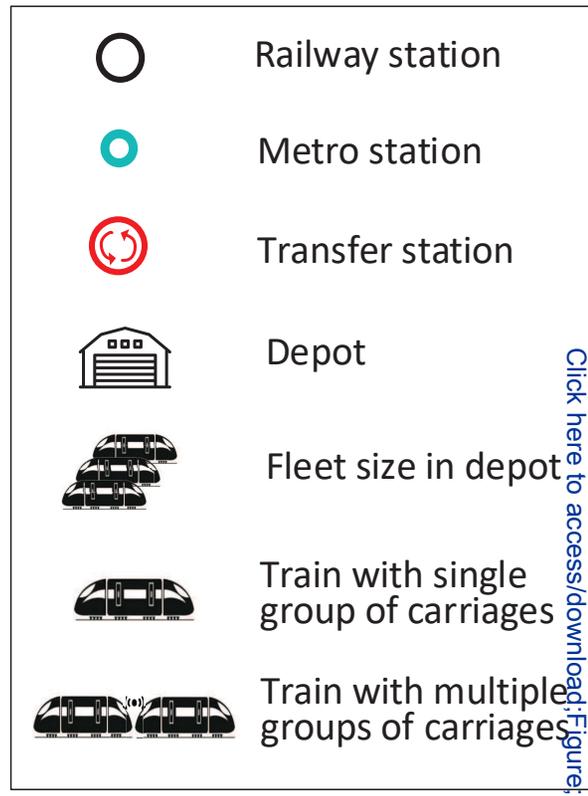
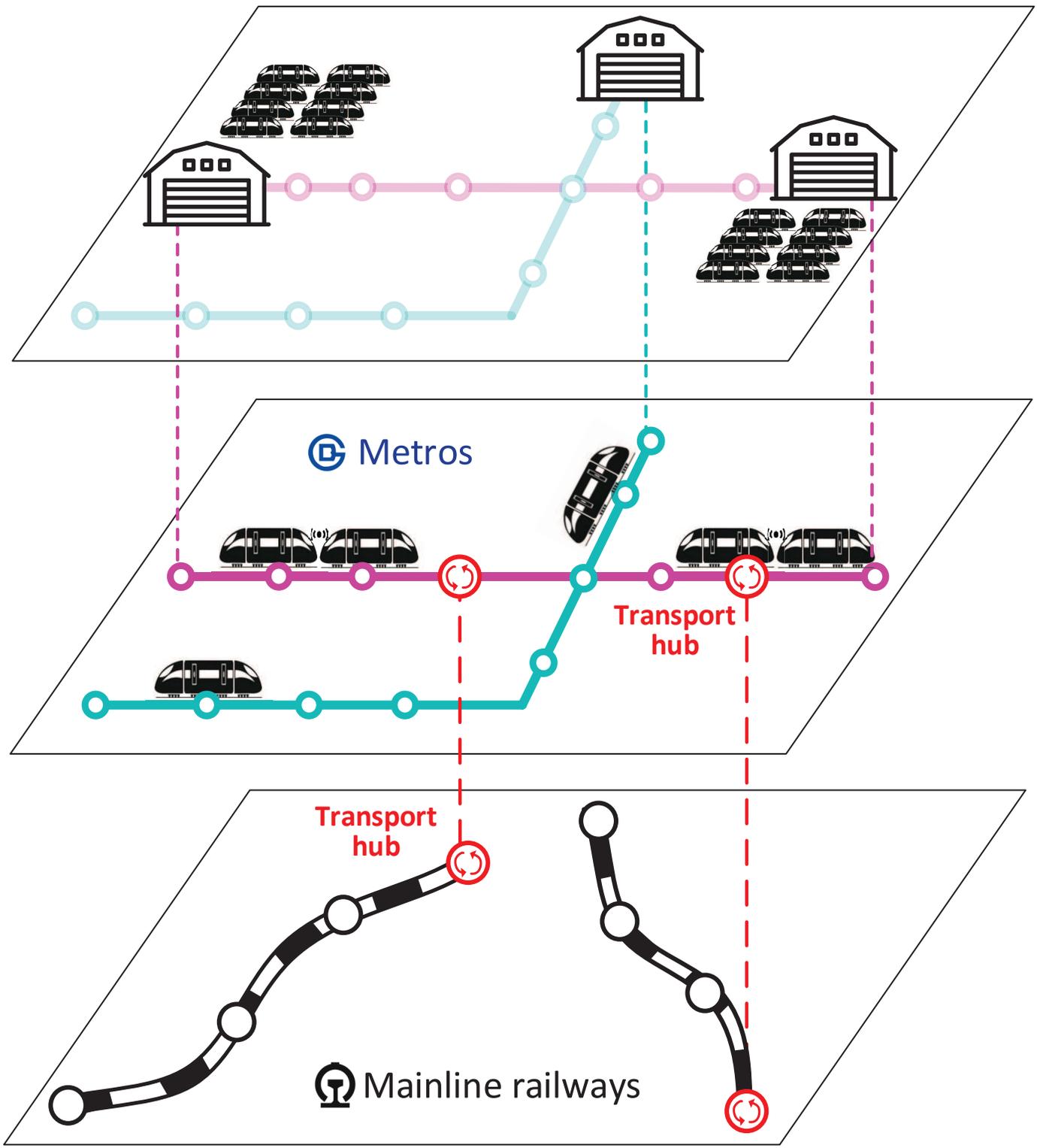
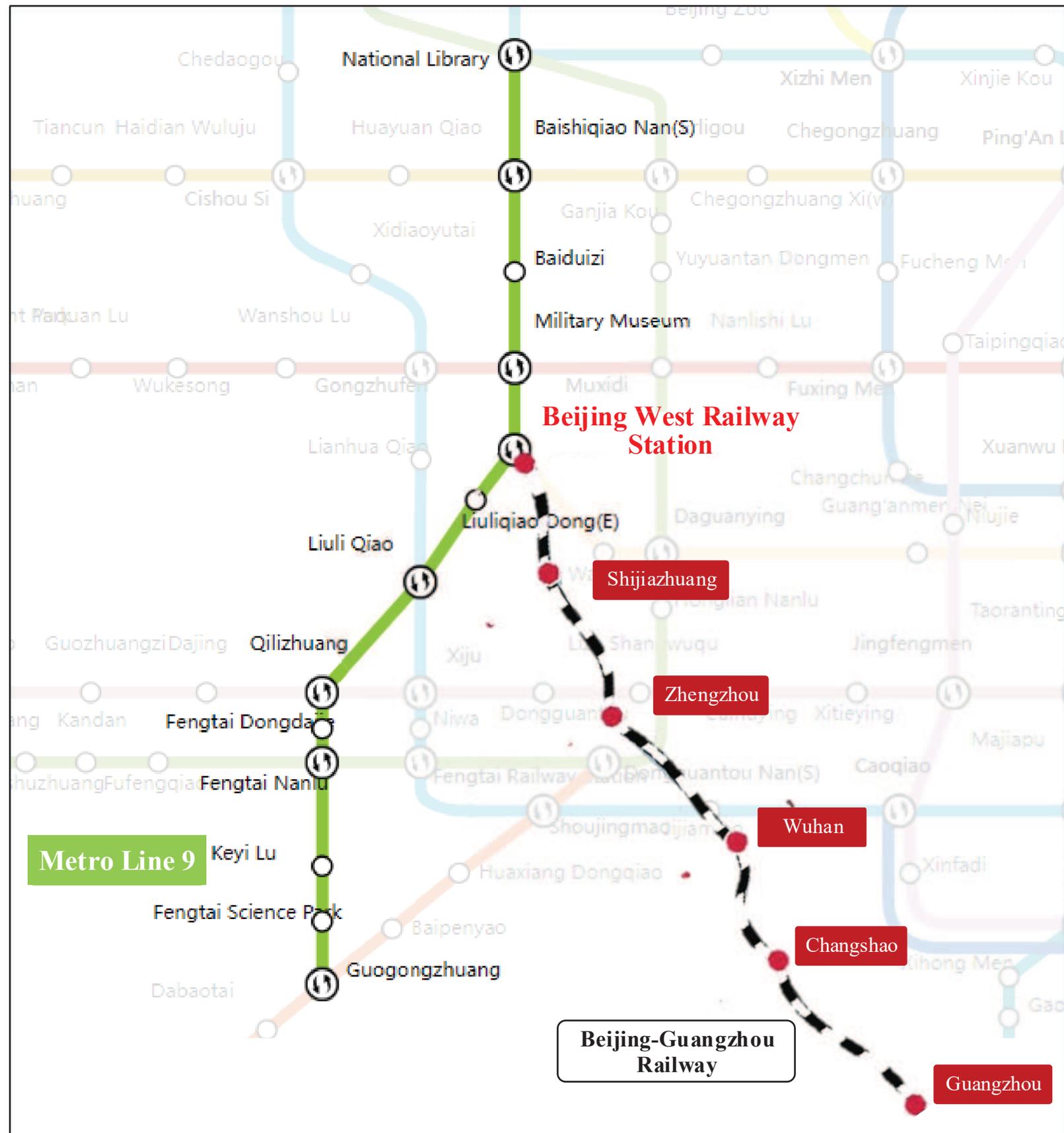
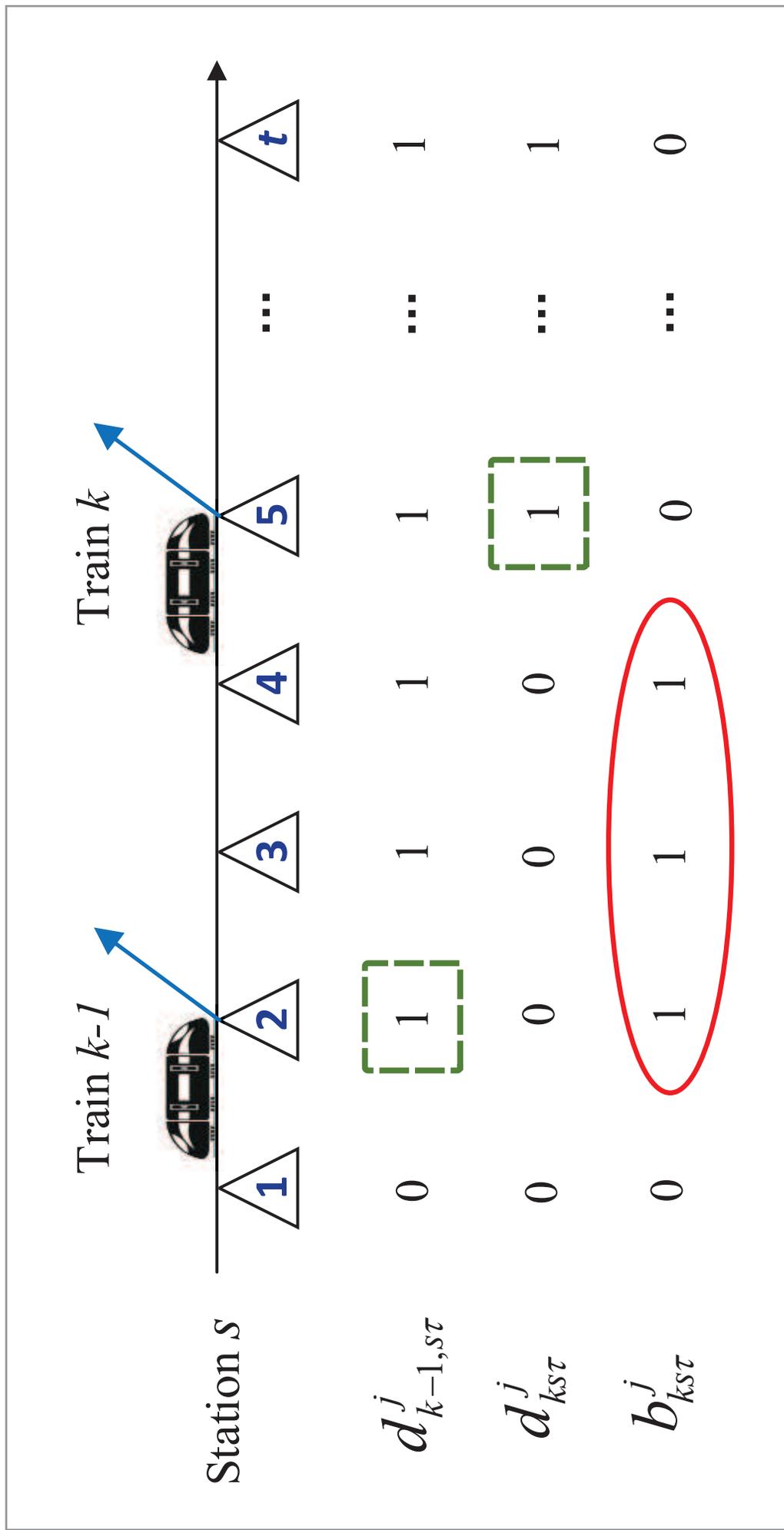
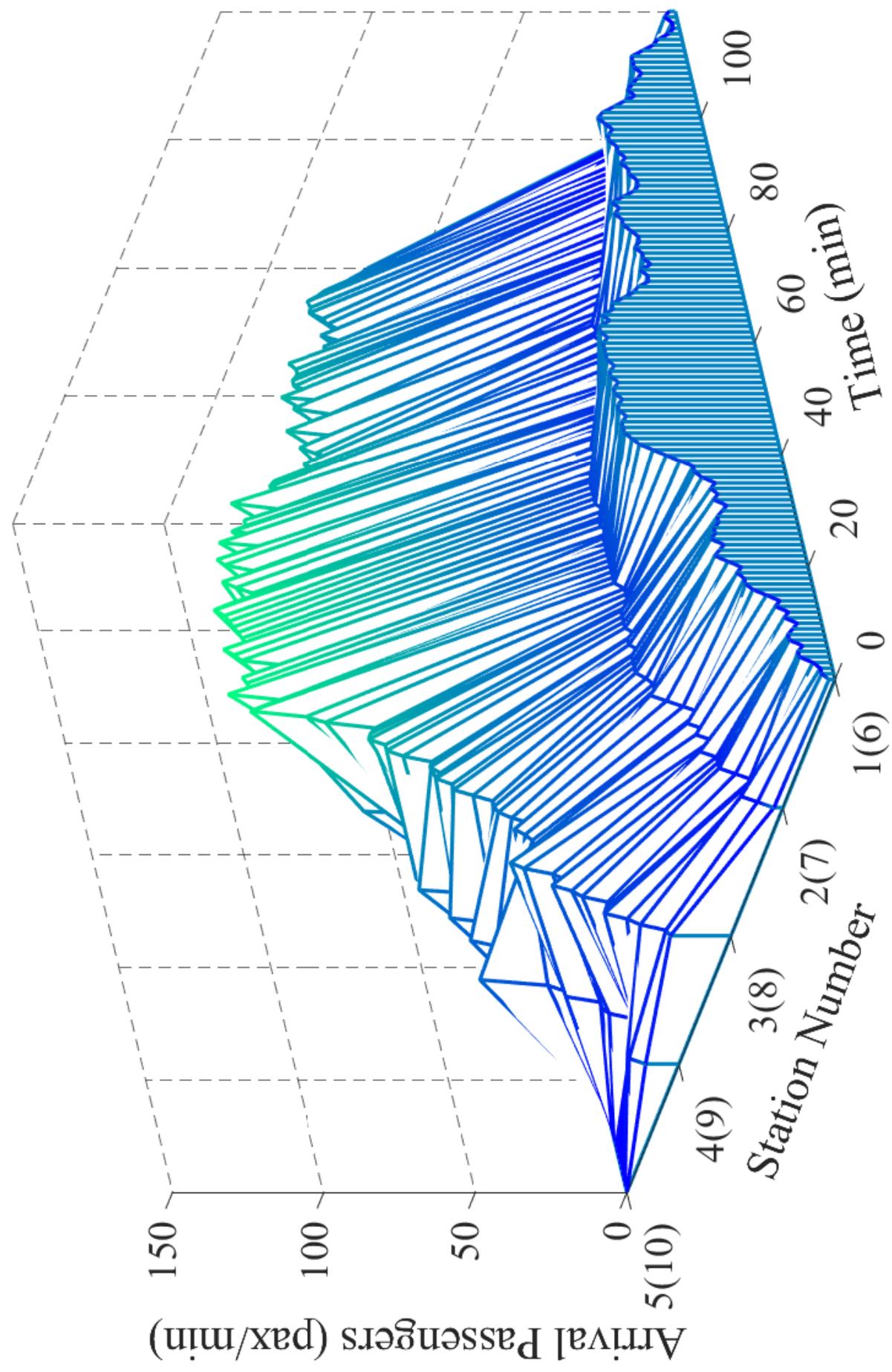
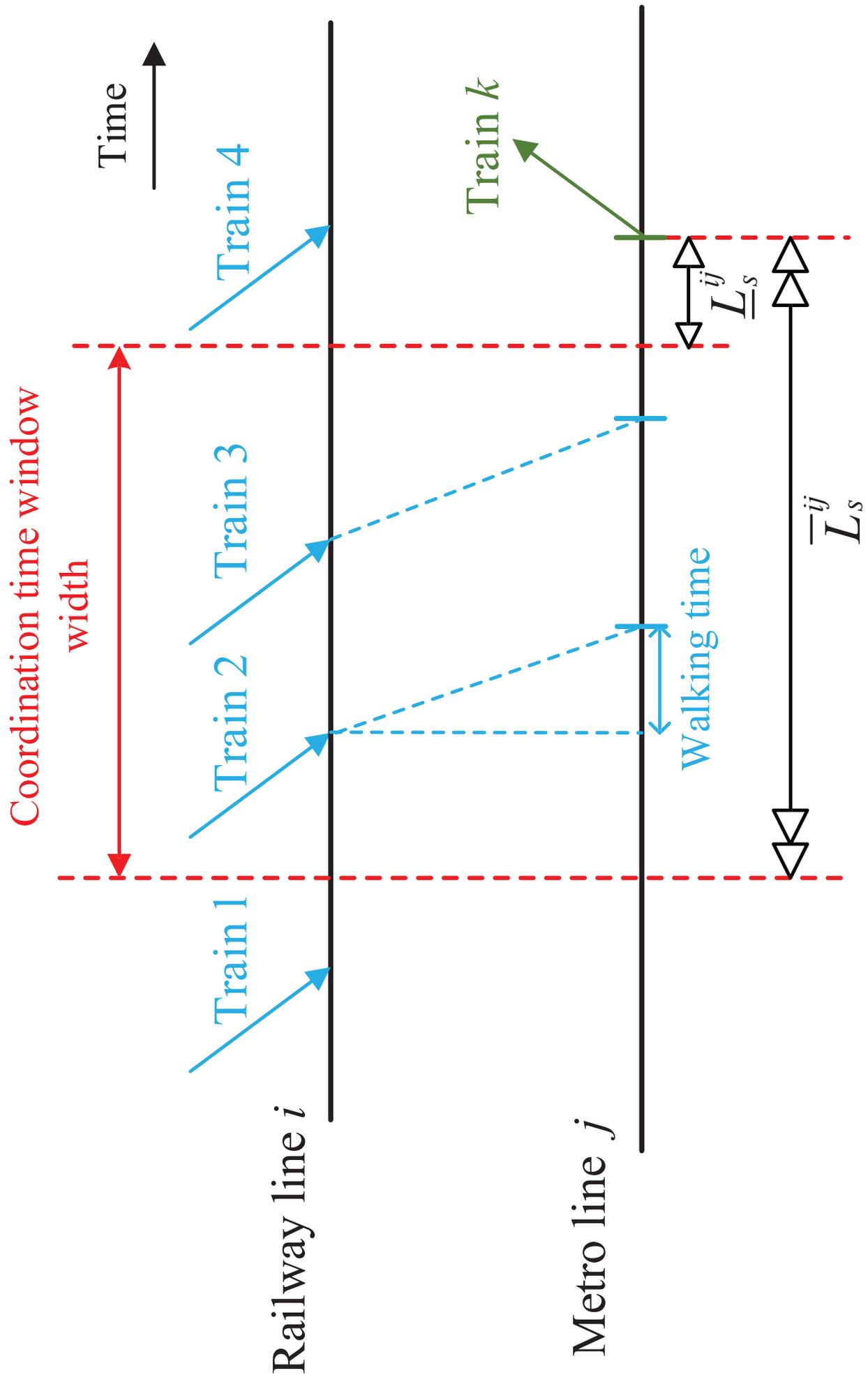


Figure [Click here to access/download: Figure/introduction.eps](#)

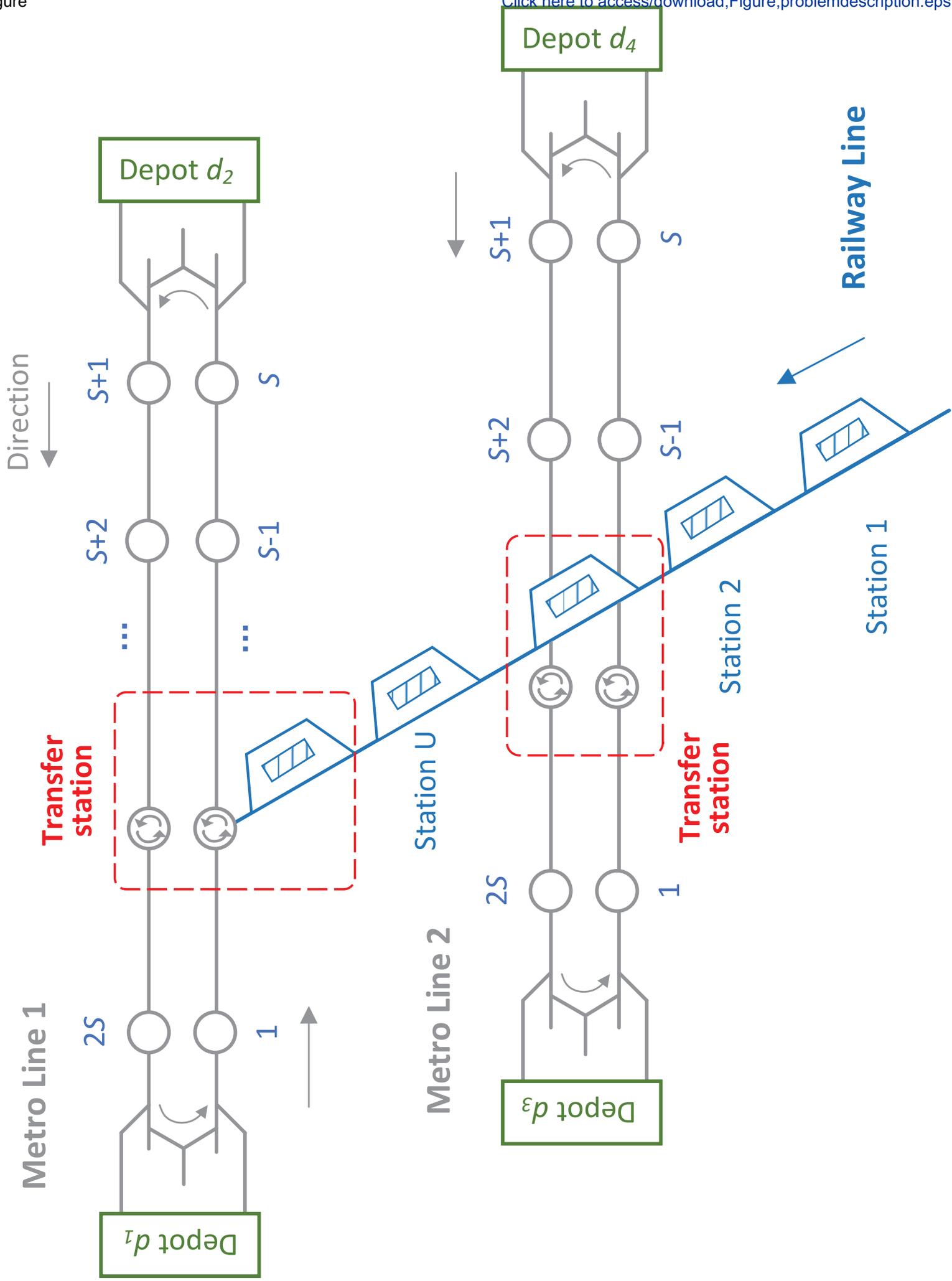


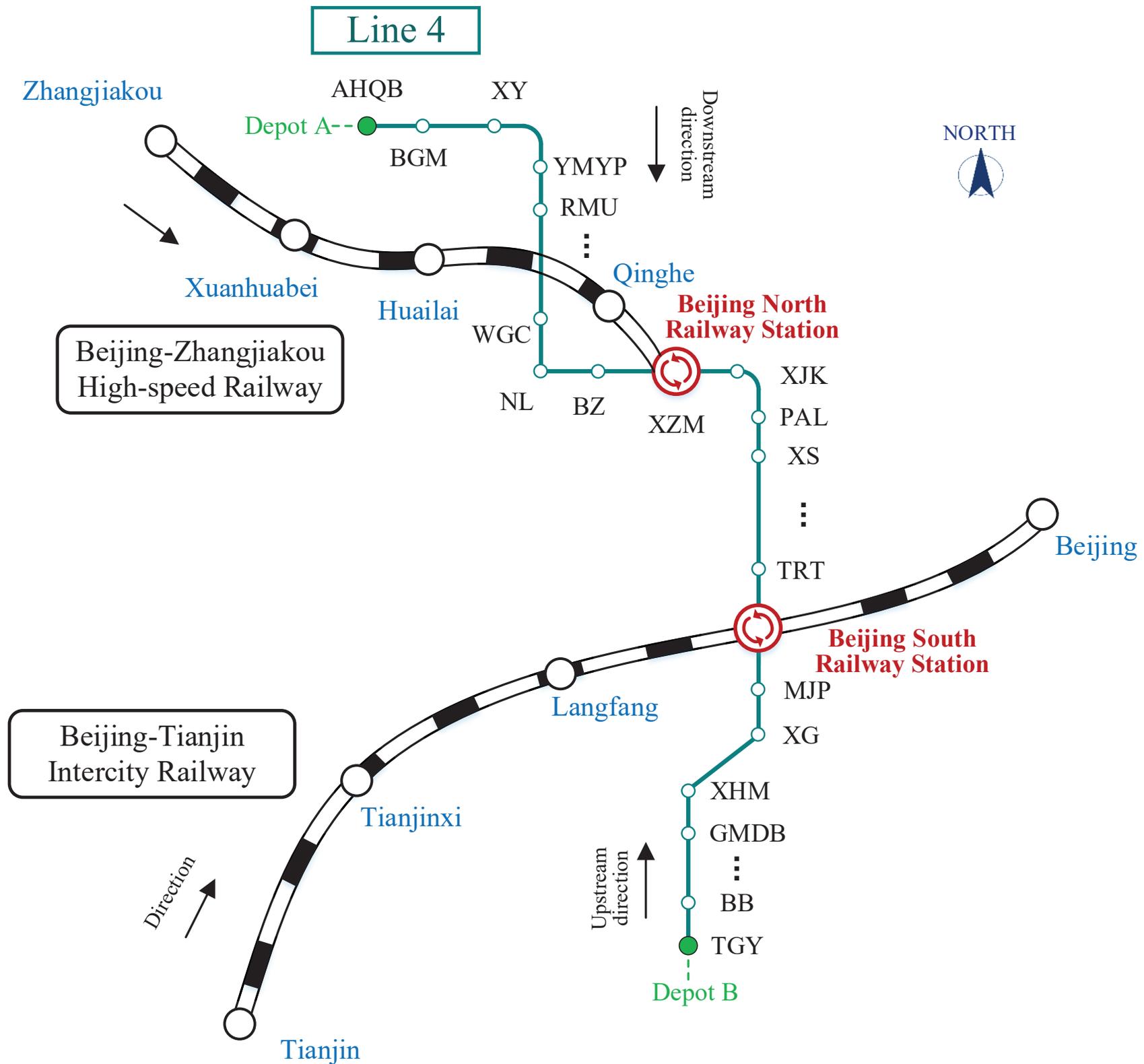


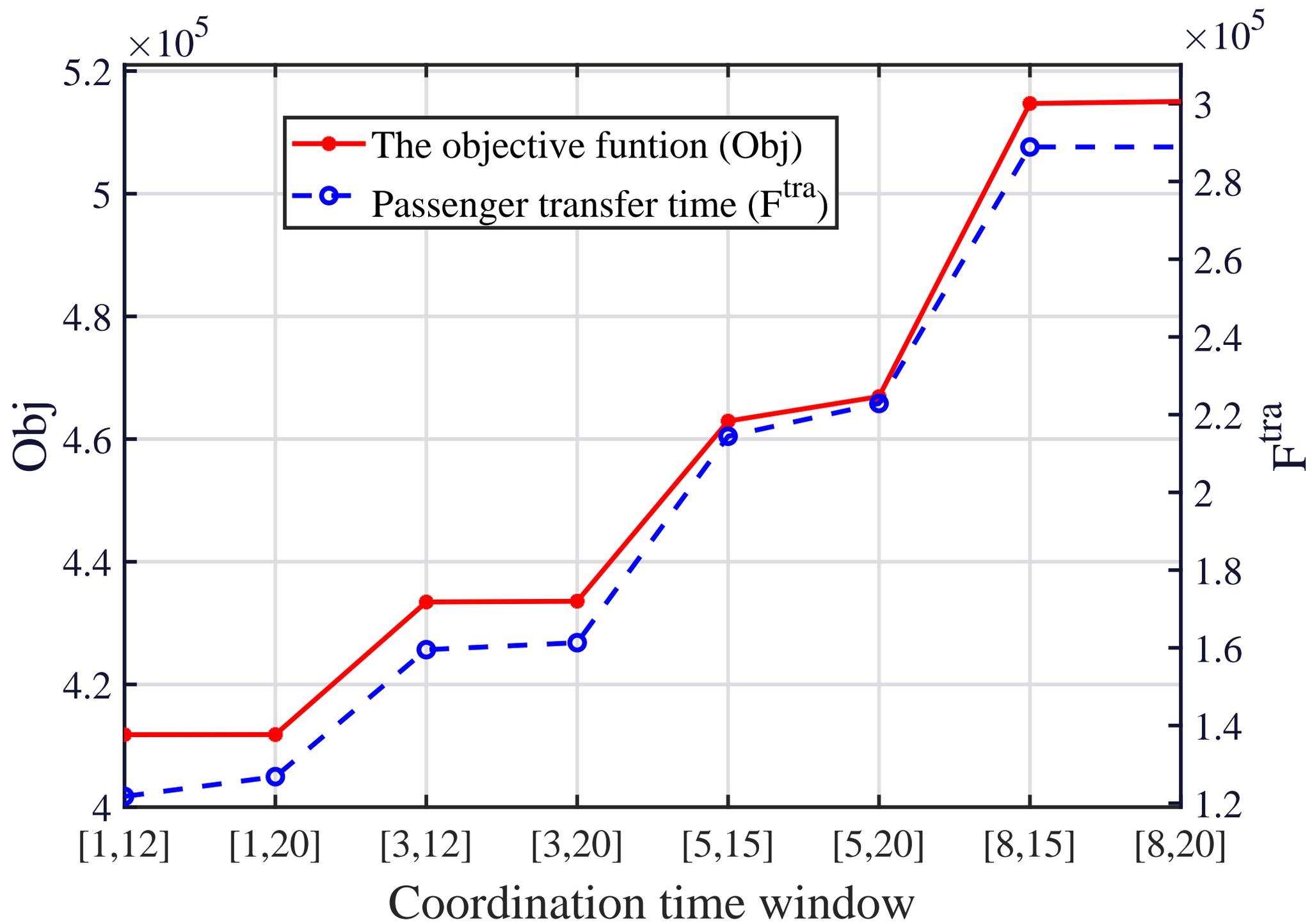


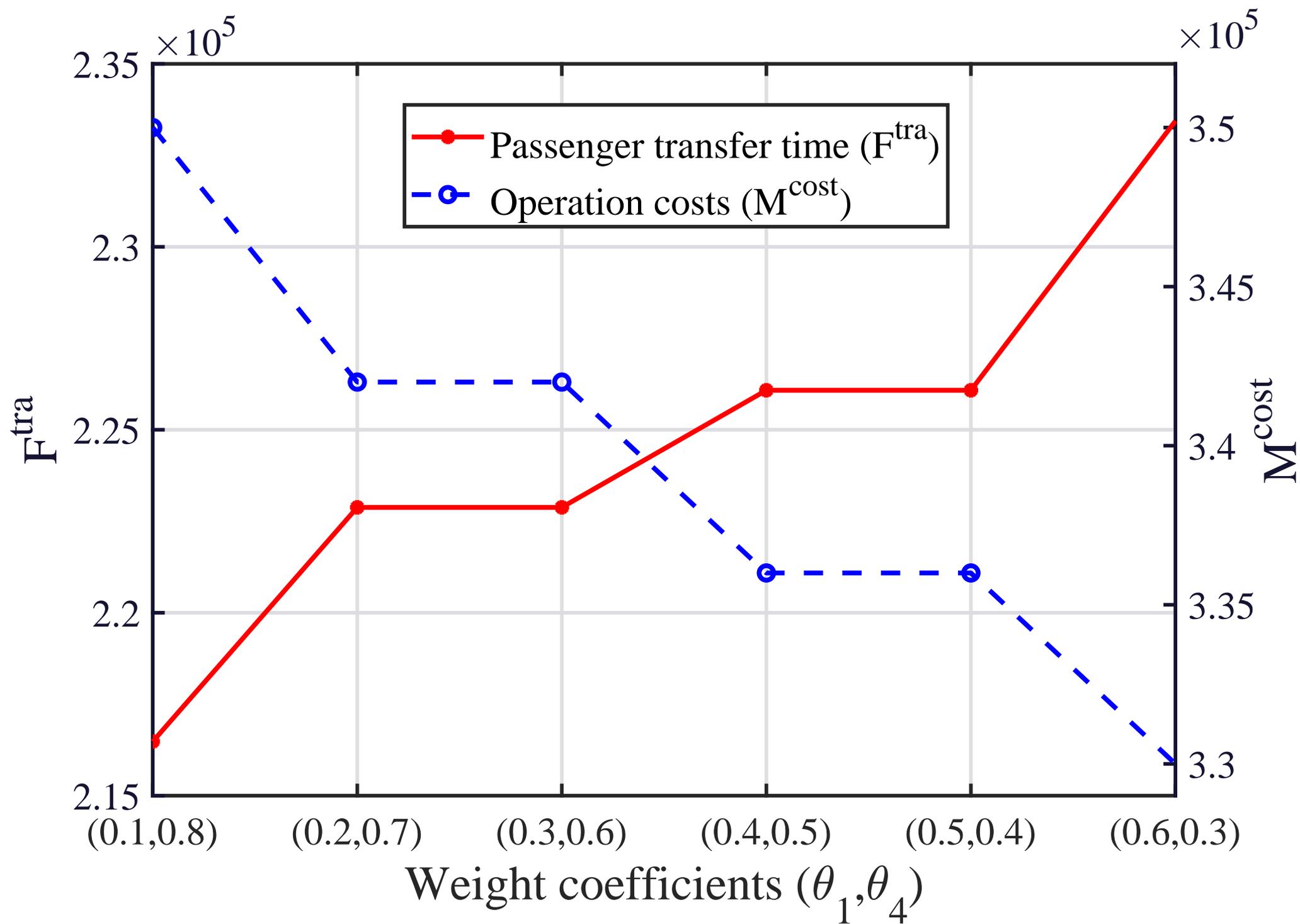


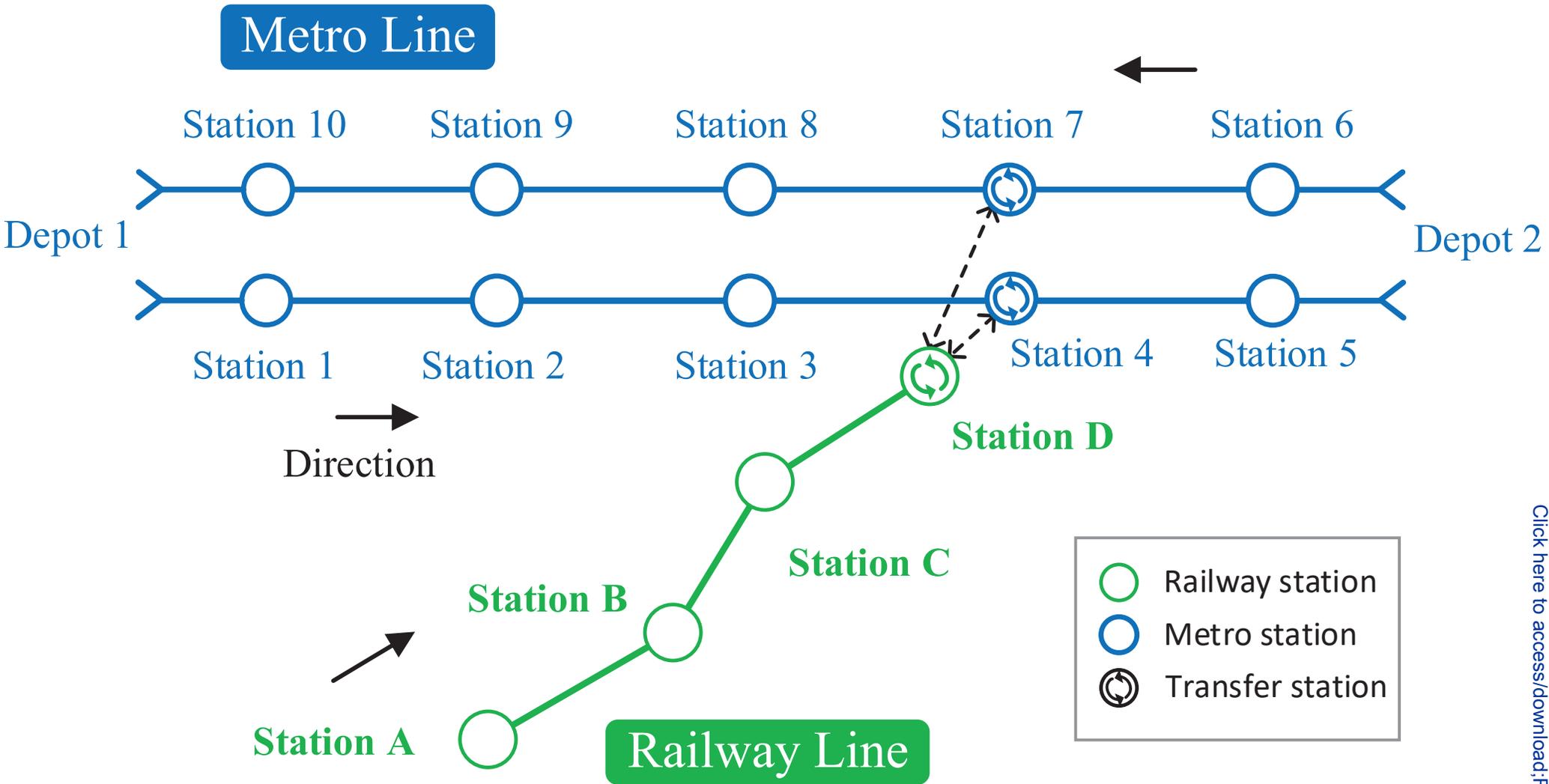
Figure

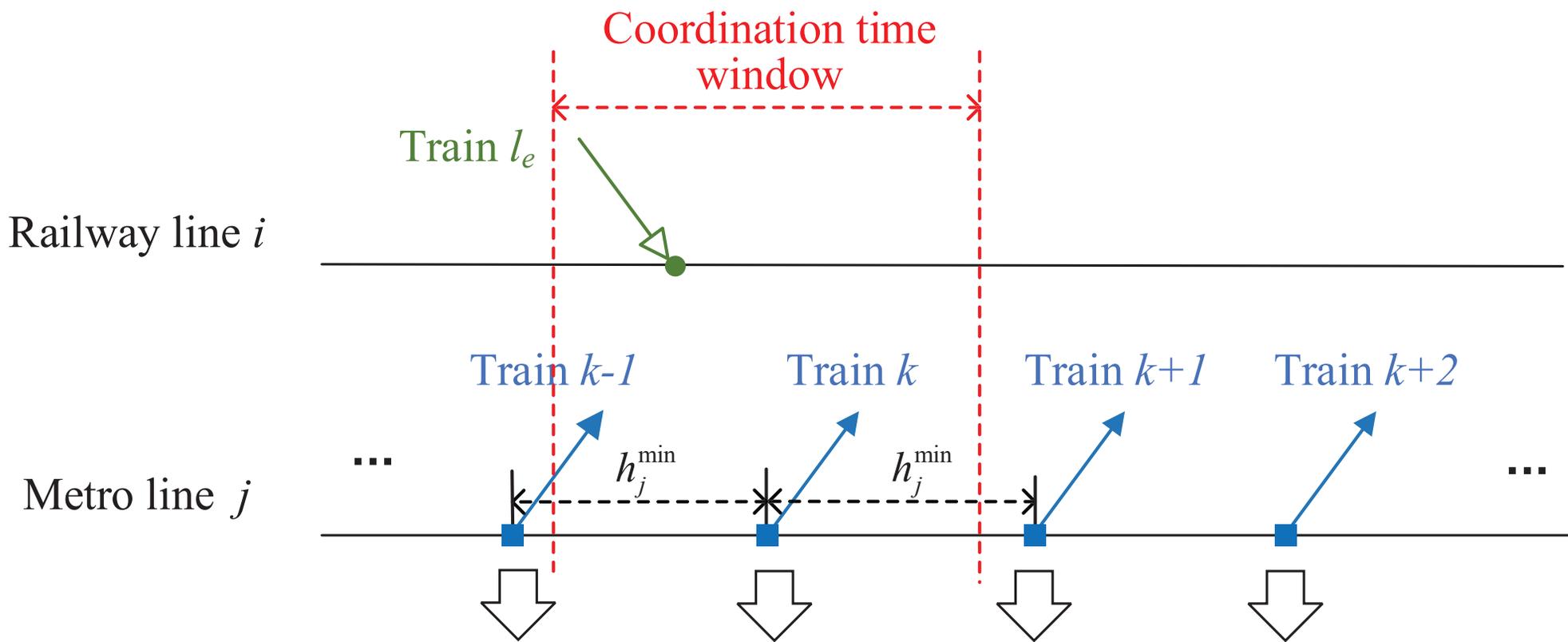




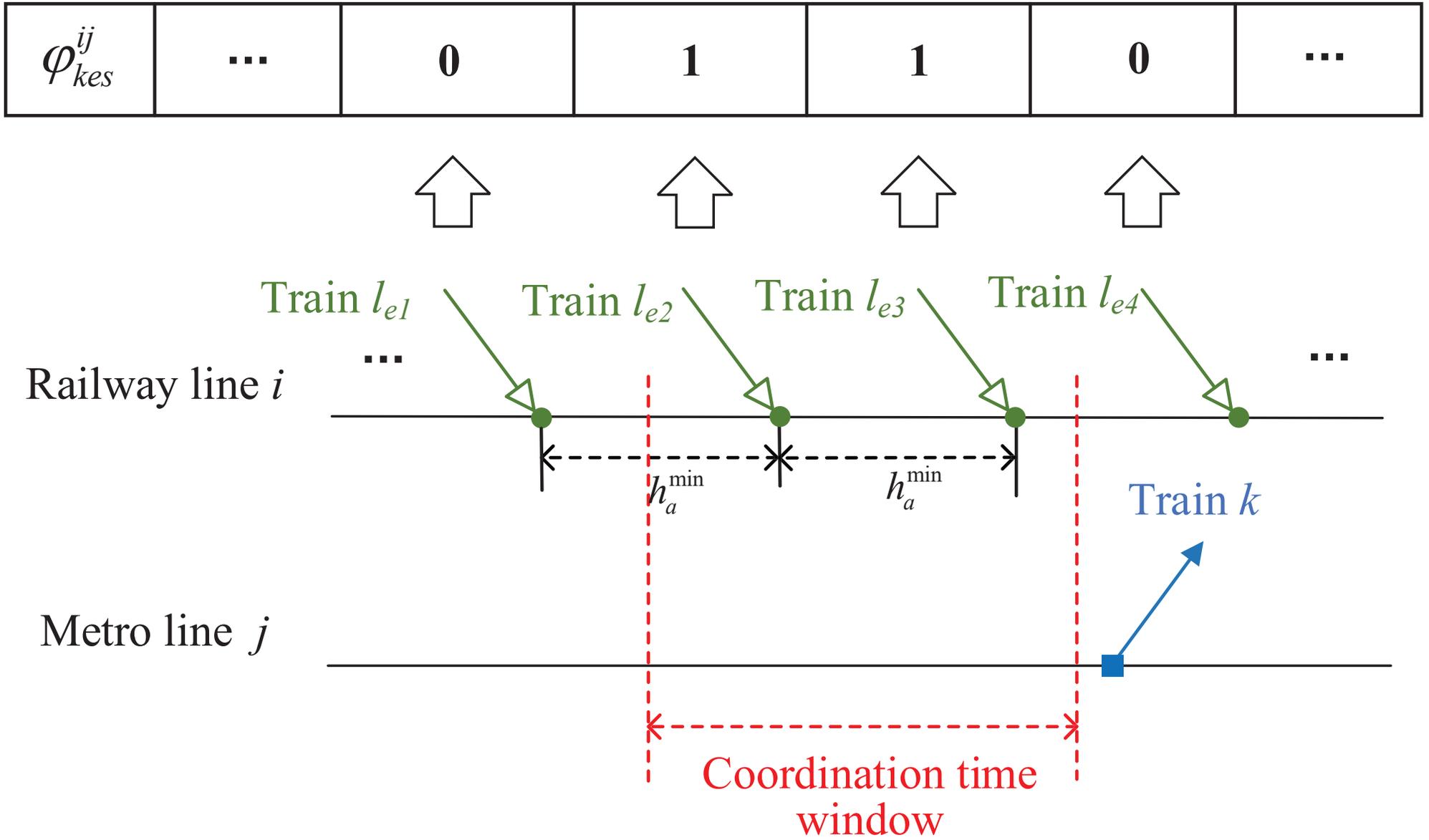








$x_{kes}^{ij}$	...	0	1	1	1	...
$z_{kes}^{ij}$	...	1	1	1	0	...



## Author Statement

The authors confirm contribution to the paper as follows:

**Simin Chai:** study conception and design, data collection, analysis and interpretation of results, methodology, draft manuscript preparation; **Jiateng Yin:** study conception and design, supervision, methodology, analysis and interpretation of results, draft manuscript preparation; **Tao Tang:** study conception and design, supervision, methodology; **Lixing Yang:** study conception and design, data collection; **Ronghui Liu:** study conception and design, data collection; **Qin Luo:** study conception and design, supervision.