University of Southampton

Faculty of Engineering and Physics Science

School of Engineering

Research on High-speed Railway Capacity Utilisation and Optimisation

by

JIAXI LI

ORCID ID: 0000-0001-6298-5747

Thesis for the degree of Doctor of Doctor of Philosophy

May 2025

University of Southampton

Abstract

Faculty of Engineering and Physics Science

School of Engineering

Doctor of Philosophy

Research on High-speed Railway Network Capacity Utilisation and Optimisation

by

JIAXI LI

Fast, frequent, direct and favourable time railway travels are the common goal of passenger railways, which has been summarised as three challenges: balancing direct services and transfers, managing the trade-off between coverage and short travel times, and addressing scheduling complexities across different national railway contexts after analysing worldwide cases and related railway planning stages. A research focus on line planning, timetable generation, and platforming is therefore established.

This thesis proposes a review framework for capacity in passenger railways, covering concepts, influencing factors, measurement, and passenger evaluation. Both train paths and timespan have their own suitable scenarios. Open tracks and platform numbers are the most critical factors for mesoscopic modelling. Measuring capacity through timetabling aligns with the research goal and includes key capacity-related passenger experience aspects such as availability and travel time. Following this, this thesis proposes a capacity-related service scheduling process for China's High-Speed Railway (HSR) by systematically analysing network management approaches, passenger characteristics, and operational factors for regional and long-distance HSR. Centralising congested HSR sections is regarded as the approach of determining train densities across other parts of the network. Origin and Destination (OD) groups are discussed as the suitable dimension for describing passenger behaviours and demand features. Operational factors are classified into attributes and timetabling details, and dwell time and overtaking are recognised as crucial for a short timetable timespan. Offering transfers for cross-line trains moving in the opposing direction at busy nodes could potentially benefit passengers.

As service scheduling problem containing line planning, timetable generation and platforming. It first divides the line planning into two stages. The first stage is a selective work which minimises interchange locations using a weighted maximum spanning tree approach. The second stage optimises service frequency for OD pairs through a weighted set covering problem. The objective is to minimise both the number of intermediate stops and operational costs, while meeting service requirements for different OD. The analysis shows that offering flexibility in train services (fast, semi-fast, and local trains) allows each OD pair to benefit from shorter travel times. This thesis then introduces an integrated timetable generation and platforming mesoscopic model that uses train path flow balance, time-mapping, and precedence constraints. The advantage of the model's structure is verified against macroscopic models. A heuristic algorithm is developed to generate feasible timetables for large-scale networks efficiently. The analysis reveals that the number of platforms plays a more significant role in reducing timespan than overtaking rules. Additionally, adding extra stops on fast trains brings certain levels of extra passenger benefit without extending the timespan. This thesis is the first to propose classifying cross-line train movements into following and opposing types, and the trade-off is defined as the potential affected train path number.

In short, this thesis provides a rich theoretical framework for railway capacity in academic research and an integrated capacity management solution that encompasses a wide range of factors, detailed operational levels, and feasible methods for real-world railway operations.

Table of Contents

Table of Tables

Table of Tables

Table of Figures

Research Thesis: Declaration of Authorship

Print name: JIAXI LI

Title of thesis: Research on High-speed Railway Network Capacity Utilisation and Optimisation

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

   Article 1: **Li, J.,** Preston, J., Armstrong, J., & Yuan, W., (2023) Evaluating Timetable's Capacity Utilisation with An Extended Event-activity Network Method. *Rail Belgrade 2023,* Belgrade, Serbia. https://railbelgrade2023.sf.bg.ac.rs/download/Book_of_abstracts_RailBelgrade_2023.pdf

   Article 2: Hu, H., Yue, Y., Fu, H., & **Li, J.** (2023). Improving service qualities of cyclic line plans considering heterogeneous passenger origin-destination (OD) flow groups. *Transportation Planning and Technology*, *46*(7), 864–887. https://doi.org/10.1080/03081060.2023.2226117

   Article 3: Hu, H., Yue, Y., Fu, H., & **Li, J.** (2024). How to optimize train lines for diverse passenger demands: A line planning approach providing matched train services for each O-D market. *Transportation Research Part a Policy and Practice*, *186*, 104154. https://doi.org/10.1016/j.tra.2024.104154

In Article 2 and Article 3, I led the following aspects: (1) the structure and drafting of the literature review, (2) proofreading of critical academic definitions and concepts, (3) logical verification and rationale development for model constraints, and (4) drafting of the discussion section. To be more specific：

- For Article 2: My contributions primarily focused on the research motivation and rationale, explanation of constraints, and the design of policy recommendations.
- For Article 3: My contributions were concentrated on explaining the research pathway, which was directly applied to Figure 4.4 in this thesis, designing the experimental approach, and determining the presentation format of the computational results.

Alternatively, the Credit authorship contribution statement is:

- For Article 2: Investigation, Writing-review & editing, Refinement.
- For Article 3: Formal analysis, Investigation, Writing – review & editing, Refinement.

Regarding foundational principles directly referenced in my thesis:

- For Article 2, I adopted the concept of grouping OD pairs at the line planning stage.
- For Article 3, I incorporated the idea that all passengers should have the opportunity to access "fast" train services, which is interpreted as a reduced number of intermediate stops in this thesis.

Signature: Date: May 2025

Acknowledgement

John Preston, John Armstrong, Melaine Forest, and Simon Blainey: My two Johns have been a constant source of support throughout my PhD. You always make the efforts to understand and respect my academic views and combine them with your thoughtful insights, and this has been invaluable across various aspects of my PhD: academic work, analytical thinking, and publishing papers. I would also like to thank Melaine for taking care of all administrative details; her great care allowed me to focus fully on my research without worrying about anything else. A special thanks to Simon for the rigorous standards in academic writing—your guidance was the turning point that led me to critically rethink and refine my approach to academic research.

Novak Djokovic, Grigor Dimitrov, Jack Draper, and other tennis players and friends： **Nole,** you've truly transformed my life. Without you, I might never have chosen to study abroad, met so many amazing friends through tennis, or developed the healthy habits that now already shaped my daily routine, personal values, and the drive to become better. You are my lifetime role model, and your influence has become as essential to me as **air and water**. **Grisho & Jack**, your kindness has made it possible for me to consider both you as my close friends, and you guys are the reason I visit Queens Cup the grass court season every year. **Grigor,** we know each other since 2016, I never think this connection could last this long, and I could not be honoured that you are truly happy every time you see me. **Jack,** the first time we met in 2023 was truly unforgettable, although I have started supporting you long time ago. You are very pleasant person to get along with and very hard-working in the court. I feel lucky to support you in a long future. Sincerely wish many many years to come.

Liao Zhengwen, Yuan Wuyang, Cheng Yan, Li Tang, Zhang Yongxiang, and Liu Yigu: You were the "strategical, tactical, and practical tutor" throughout my PhD research. You shared your academic experiences, insights, encouragement, and practical to-do lists with me generously and without reservation. I thoroughly enjoyed each discussion we had, and your ideas formed the "backbone" of my academic foundation, while your encouragement served as a beacon guiding me through challenging times. I hope to continue growing, absorbing all of your strengths, and becoming someone who can also support and guide others in their academic paths.

Xu Jingxiao & Ji Hezi, Andrej Trajkovski, and Li Peng: You were my closest companions during my time in Southampton. Throughout my PhD, I gradually became less social, but it wasn't until each of you left after graduation that I truly realized how much your presence meant to me. I'll always cherish those moments we spent watching tennis and playing mahjong, especially as we witnessed Djokovic rise to become the GOAT.

Nie Lei, Fu Huiling, He Zhenhuan, Jing Yun, Li Dewei, Miao Jianrui, He Bisheng, Jiang Xinguo, Zhao Lan, Taku Fujiyama, and Lin Youfang: As I ventured into the world of higher education, each of you became a role model in different ways. Some of you extended a helping hand when I was at my most confused and challenged, offering me repeated opportunities to reflect and grow. Others provided encouraging words during conferences or casual conversations, helping me rediscover optimism amidst the self-doubt and over-analysis that marked these four years. Some of you were always there, responding instantly on WeChat, offering invaluable advice and guidance on my academic concerns and challenges. For all of this, I extend my deepest thanks to each of you.

Xu Minghao, Li Yueyi & Zhang Zeyu, Qi Xinjia, Zhuge Yixiu & ONG PEI BOON, Han Peiran, Hu Huaibin, Li Haoyu, Zhao Chen, Yu Liuqing, Li Chen & Gong Yantao, Zhang Peng & Guo Enze, Wang Shuling, Huang Ying & Wu Fei, Wang Jiachun, Chen Teng & Gao Jinxiang, Zhao Tianyin, Xu Zhengjia & An Mingyue, Yu Chao, Yu Zongze, Liang Ke, Wang Tairan, and Sui Xupeng.: I want to thank each of you for being there during my moments of frustration and doubt. It's through our shared support and encouragement that I've managed to bring this dissertation to completion. I wish each of us all the best in whatever lies ahead. You are the guys I turn to first whenever I feel lost. Your unwavering trust and encouragement have been instrumental in helping me navigate this journey, and I couldn't have reached this point without you guys.

Zhao Tianyu, Ye Zi (Exland), Li Zonghuan, Qin Guowei, Liu Sipei (Wingspan), Li Kunling, Shu Yihui, Li Zhenxing, Xu Lei, Song Shiping, Gui Chenhuan, Li Bolin, Zhao Jingyi, Li Jiawei, Zhen Xiaosheng, Liu Yang & Wen Xin, Guo Hao, Wu Tong, Ma Eryu, Yang Jun, Zhang Xiaoliang, Wu You, Mao Zijin, Xiao Tinghan, Ma Deming, Coco & Dawan, Hao Xuan, Luo Xiaoxi, Jian Mengzhen, Karim Biaddang & Shen Kang: you have been my greatest outlet beyond tennis, always there to listen to my rants and frustrations. Your constant encouragement has driven my progress, and I look forward to the day when I can confidently show what I have become.

Li Weizhen, Wang Lei, Zhou Chiming, Shen Yifei, Yao Yu & Yao Hanlin, Hao Qi (Hocky), Yang Xu, Yuki, Bao Fucheng, Chen Xin, Gao Yan, Zhou Jiahuan, Wang Zi'an, Zhu Wenhan, Lin Yiyi, Hu Yawen, Xiao Yuexin, Liu Zheng, Yu Yun, and Huang Shuo: Novak Djokovic gave me more than just the love for tennis—he brought all of you into my life. Tennis connected us, and through it, we became friends who could talk about anything. I cherish the memories of our time together at Lin Cuiqiao in Beijing, Qizhong Tennis Centre in Shanghai, and SW19. I'm certain we'll find ourselves back on the tennis courts together someday.

The gym Portswood, Puregym Southampton Central, Liu Dun, Tian Rui, Wang Chengpeng, Liam Day, James Hicks, Hubert Poletek, Sun Shuang, Zhang Xiangyu, Zheng Weichao, Bryan Yong, Will Lewis, Nick Finch, Adam Mikietinski, Shentu Zeyu: I always thinking whether I am too late for the gym, and the workout is a great part of my life and I am honour to make so many friends in this journey, wish I could confidently be topless in a very near future.

Excellent youtubers and up-stars from Blibli: you guys form as another university where I could always find answers and personal inspirations. At times, I've wondered how my life might have unfolded in parallel universes. In this age of content, I've had the privilege of seeing so many of you excel in your fields. I hope that I can continue to grow and shine in the future just like you.

Long Jin, Li Yina, He Haitao, Zhang Peng, Deng Toulian, Zhou Xiaolian, Huang Xiao, Xiong Zeliang, and Wang Lingjin: Before university, my idea of education was solely focused on getting into a good secondary school in Changsha, Hunan. The six years I spent there, in a top academic environment, instilled in me the resilience to face challenges and refuse to surrender to fate. I wasn't a confident child then, but it was the unwavering trust and support of my teachers that helped me grow into the person I am today, able to give back through my own academic journey. I extend my heartfelt thanks to my high school teachers.

Mom and Dad, my Coady family, Li Yitian, Wang Zhicheng, Li Jianxiong & Luo Huasheng, Li Ye, and other family members: To my parents, thank you for your unwavering love and care throughout all these years. It hasn't always been easy, especially when you had to push past societal expectations to support and respect every decision I made. To the Coady family, you have been my home in Southampton and my greatest source of strength in the UK.

My Motherland: finally, I would like to express my gratitude to my country. Coming from a modest background, it was your financial support that made my studies in the UK possible. I hope that in the long future, I will continue to grow and evolve, becoming someone who can truly contribute to my country and the world with expertise in high-speed rail development.

Abbreviation

CR ....China Railway

CUI ....Capacity Utilisation Index

DB ....Deutsche Bahn

EAN ....Event-Activity Network.

GJT....................................Generalised Journey Time

HK.....................................Hong Kong

HSR ....High-Speed Railway

ITF.....................................International Transport Forum

JR......................................Japan Railway

LPPA ....Line Planning based on Passenger Assignment

OD ....Origin and Destination (passenger route)

OECD...............................Organisation for Economic Co-operation and Development

OT.....................................Origin and Terminus (train starting and ending)

ORR ....Office of Rail and Road

PEI ....Passenger Experience Indicator

RSI ....Railway Supply Indicator

RP ....Revealed Preference

SNCF ....Société nationale des chemins de fer français

SP ....Stated Preference

SILMST.............................Selecting Interchange Location based on Weighted Spanning Tree

STN ....Space-time network

UIC ....International Union of Railways

UK.....................................United Kingdom

V&V...................................Verify and validate

WMST ....Weighted Maximum Spanning Tree

WSTP ....Weighted Set Covering Problem

Chapter 1 Introduction

1.1 General motivation

Railways are a fast, efficient, and green public transport mode (International Union of Railway (UIC), 2018). The overall economic, social, and environmental benefits lead to more countries proposing new ‘fast’ passenger railway or high-speed railway (HSR) projects to provide links between regions and cities. Taking HSR as an example, there were 56,000 km of HSR in operation worldwide in 2021, and this mileage is expected to double by 2050 (UIC, 2022).

The railway network is typically formed and interconnected through the construction of railway lines in phases, and maintaining fast, frequent, and direct services for every journey, namely every passenger origin-destination (OD) pair, is increasingly challenging as the network develops. If railway operators aim to cater for more passenger requirements and improving service quality in the passenger railway planning process, it will cause additional challenges in terms of the capacity utilisation of the network (Yan et al., 2019). The rationale behind this is that the available track occupation time on a daily basis and infrastructure amounts are limited, against many possible service combinations which can be scheduled into the network. Not all possible passenger journeys, can be arranged by 'attractive' train services which have:

- preferred arrival and departure times aligned with OD characteristics,
- minimised in-vehicle time, and
- the provision of multiple daily train options.

A direct solution is building more railways and purchasing more rolling stocks so that the network railway capacity will have an obvious increase, but the HSR or the fast railway is one of the most expensive transport systems. Before proposing any railway infrastructure or fleet expansions, railway companies prefer modifying operations first, represented by the timetables, to see whether they can:

- improve the capacity utilisation: adding more trains while retaining a similar capacity utilisation level or using a lower capacity utilisation to deliver a similar level of service.
- improve the service quality of the timetable: the timetable works better towards a series of goals (e.g., fast travel and dense train traffic) to make an optimal capacity utilisation of the infrastructure.

This essentially requires railway operators to process a systematic knowledge of the capacity in passenger railway scenarios, enabling an optimal railway capacity utilisation. When managing a fully-connected railway network, whether it is HSR or otherwise, the consideration arises that what kind of train services should be scheduled into the network? What kind of timetable strategy should be adopted? Is it possible to balance service quality and railway capacity?

1.2 Research scope analysis

The capacity of a railway network is influenced by many stages of railway planning process (UIC, 2013). A service scheduling problem typically covers network planning, line planning, timetable generation, platforming and routing (track allocation problem), and possibly other tactical sub-schedules (e.g., rolling stock circulation, crew scheduling). A representative passenger railway planning process can be seen in Figure 1.1. Consequently, exploring an optimal capacity utilisation problem can be conducted based on these planning stages with various perspectives and modelling details.

Besides, considering the funding part of this PhD research, this thesis mainly takes China HSR as case study background, and it is reasonable to give a research scope analysis, introduce some key facts of China HSR, identify the commonality of the problem regardless of railway contexts, and set up several primary research boundaries.

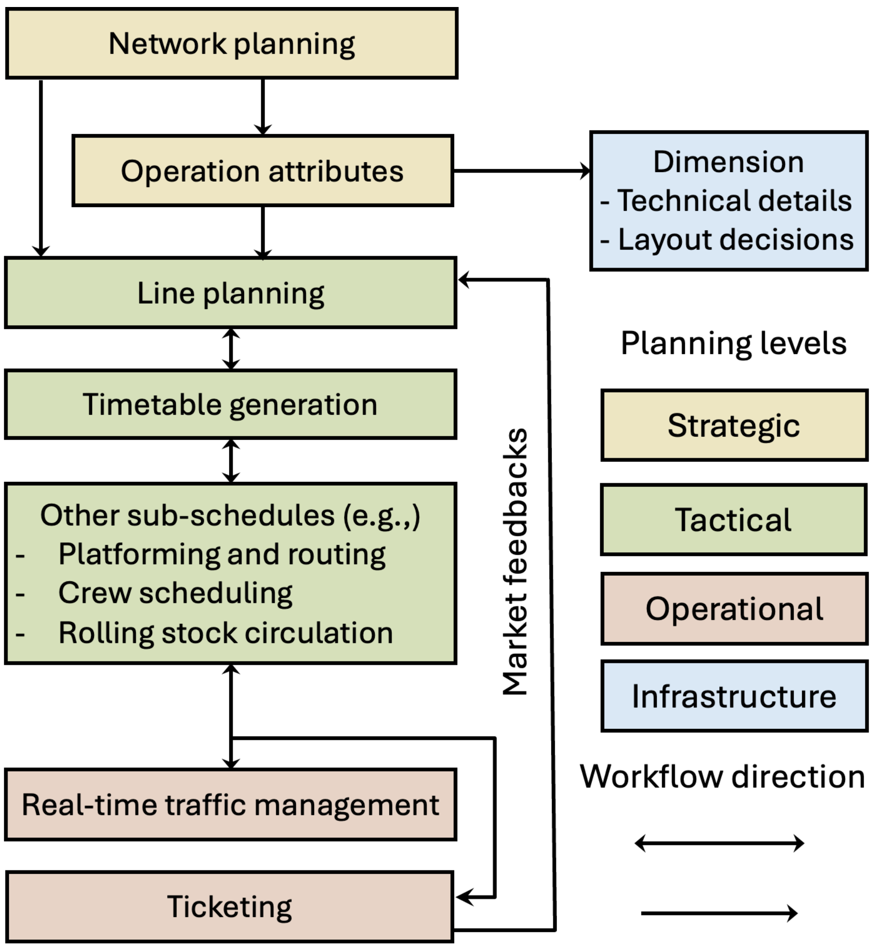


Figure 1.1 The passenger railway planning process (Liao, 2021; UIC, 2013)

1.2.1 Railway capacity at railway planning stages

1. Network planning.

The working content of network planning is determining the nodes and links (Canca et al., 2019) to construct railways and is deeply affected by different countries' planning and operating traditions. Another important part of network planning is to decide the operation attributes of a railway project. Deciding operation attributes is positioning some key transport characteristics (Liao, 2021) of a railway project, reflecting the demand features cross the network, consisting of:

- network interdependency (i.e. whether network can be divided into independent parts)
- service type (e.g., direct services or requiring interchange),
- train mix type (e.g., passenger trains only network with the mix of stopping patterns),
- top train speed levels (e.g., 350km/h, 300km/h, 250km/h, 160km/h),
- the periodicity of the timetable.

The dimension of a railway project can therefore be processed into the next phase, consisting of technical details and layout decisions (UIC, 2013), such as selecting the suitable rolling stock type, matching the signalling system, deciding the minimum distance between two stations. UIC (2018) lists three primary passenger railway types with their distinctive operation attributes:

- long-distance intercity railways,
- regional (short-distance intercity) & suburban railways, and
- urban railways.

However, it should be noted that the definitions of these three railway types entirely depends on railway organisations in countries. For example, NetworkRail (2013) in the UK defines long-distance rail travel as exceeding 80.5 km (50 miles) or between large cities (towns) separated by over 48km (30 miles). Deutsche Bahn (DB) in Germany (DB, 2023) defines regional railways to be within travel distance shorter than 50km or with travel time under one hour. SNCF (ARAFER, 2016) in France categorises regional railways as trains operating within the same region. The Office of Rail and Road (Office of Rail and Road (ORR), 2023) in the UK, on the other hand, classifies regional railways by jointly considering region and travel distance: ScotRail (by region), TransPennine Express (by region and distance), and some CrossCountry services (by distance).

Interestingly, some rail networks can have multiple railway types or combined features, such as rail transit networks in London, Paris, and Tokyo. For example, the service coverage of Thameslink indicates that it can act as urban railway serving central London, regional railway serving satellite towns around North and South London, and intercity railway serving long-distance destinations, such as Peterborough and Brighton. Meanwhile, Transilien SNCF, exhibits both features of regional and urban railways serving central and suburban Paris.

In short, the network planning frames the overall capacity utilisation features of a railway network. Accordingly, studies that examine whether planned or proposed expansion of infrastructure can meet transport needs are classified as capacity planning or expansion in this thesis. Representative works include Lai & Barkan (2011) and Rosell & Codina (2020).

2. Line planning.

The working content of line planning is to determine the operation elements for each train, namely a train line: the stop plan, the rolling stock composition, and frequency (Fu et al., 2015; Schöbel, 2012). To be more detailed, the stop plan contains originating (O) and terminus (T) stations and intermediate stopping stations of a train line. When a line plan is proposed, most service qualities can be analysed, such as expected In-vehicle time for each passenger OD, service frequency, direct travel, or a transfer. Railway capacity at this stage can be considered in various ways. One approach is to treat it as a constraint by defining upper limits on the number of train (lines) that can pass through each railway section (Burggraeve et al., 2017; Hu et al., 2023). Alternatively, capacity can be incorporated into the objective function, such as minimising operational costs, which often leads to models that conclude with a reasonably small number of train (lines) (Zhang et al., 2020). In some cases, however, railway capacity is omitted entirely from the model (Fuchs & Corman, 2019). A major issue arises that a line plan might not guarantee a feasible timetable at the timetable generation stage (Bešinović et al., 2021; Fuchs et al., 2022).

In summary, line planning determines most of the service content of a passenger railway network. As railway operators might wish to schedule a line plan which satisfies most service qualities for all passenger ODs, it is possible that the proposed line plan is infeasible as a result of ignoring the railway capacity constraints.

3. Timetable generation and platforming.

In most cases, line plans act as inputs of the timetable generation which forms the train traffic by assigning conflict-free time slots to train lines. At the timetable generation stage, infrastructure constraints (e.g., platforming) and operational rules (e.g., dwell time, headways) need to be carefully considered (Zhang, 2022; Zhang, 2019). The issue of infeasibility remains in the sequential steps: timetabling, platforming and routing. This is because the results from timetable generation are usually at a macroscopic level, so conflicts or constraints at the platforming and routing stage can still render the overall schedule plan infeasible (Goverde et al., 2016). When observing the train traffic, it is natural to consider buffer time (Landex, 2008) or slack time of critical train paths (Goverde, 2007) to schedule a robust timetable, see Bešinovic (2017) as an example, so that the timetable can be 'reliable' and 'stable' at the real-time traffic management stage.

In summary, the timetable generation stage finalises the remaining service contents of a railway network: arrival and departure times of trains and interchange opportunities (Guo, 2015; Li, 2018) for some passenger OD pairs. Compared to other passenger railway planning stages, the timetable is regarded as the core denotation of capacity in passenger railways (Zhang & Nie, 2016) because it represents the capacity consumption of a service proposal for a given railway infrastructure. Accordingly, studies that modify service networks with capacity constraints, or deliver capacity-related objectives, are classified as capacity analysis or optimisation in this thesis.

4. Other tactical planning stages.

The working content of rolling stock circulation is to match the train paths on the timetable while satisfying their maintenance requirements (Li, 2019). Moreover, Liao (2021) experiments with the relationship between the railway capacity and rolling stock amount via a hybrid space-time network formulation. The result shows that the marginal increase in railway capacity decreases with the growth of rolling stock, and potentially affecting the real-time performance of the timetable, such as causing delays (Armstrong & Preston, 2017). A similar conclusion has also been reached by Yin et al. (2023).

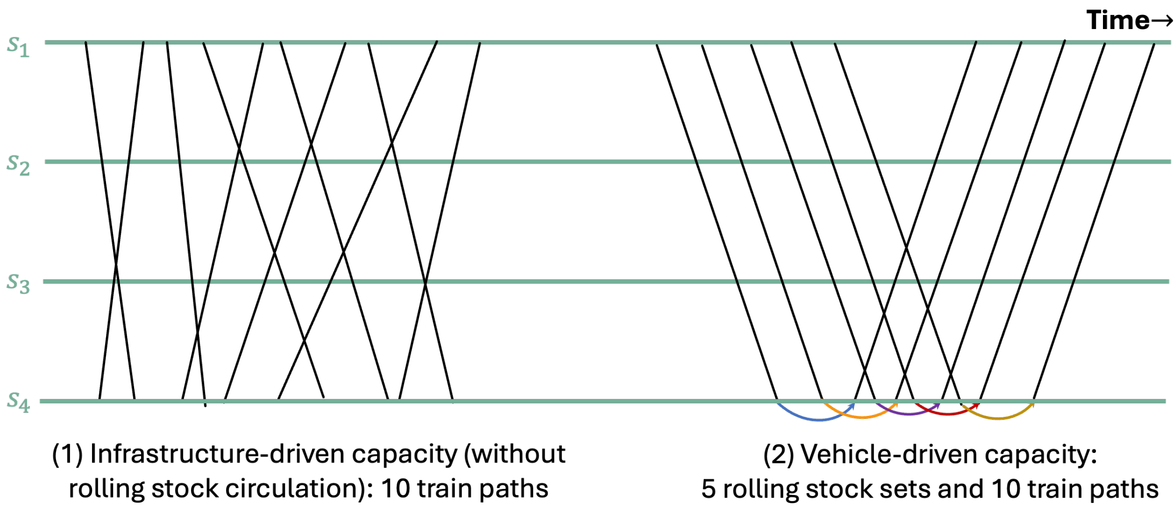


Figure 1.2 Different types of railway capacity shown in time-space graph (Liao et al., 2021)

5. Summary

To schedule a service plan in a dense and busy railway network, it requires concluding all service contents and assessing critical service qualities. Meanwhile, this thesis focuses on a known railway context with established infrastructure layouts, omitting network planning from its scope.

While acknowledging the important impact of crew scheduling and rolling stock circulation on railway capacity, they are regarded as subsequent steps following the timetable generation in this thesis. Therefore, this thesis is defined as a service scheduling problem primarily covering the line planning and timetable generation stages, and carefully considering the influence of platforming (Figure 1.3). The term of timetabling is introduced in the following paragraph to cover these three stages of passenger railway planning. A service plan shall contain the line plan, the feasible timetable (or time-space graph), and the platform arrangements.

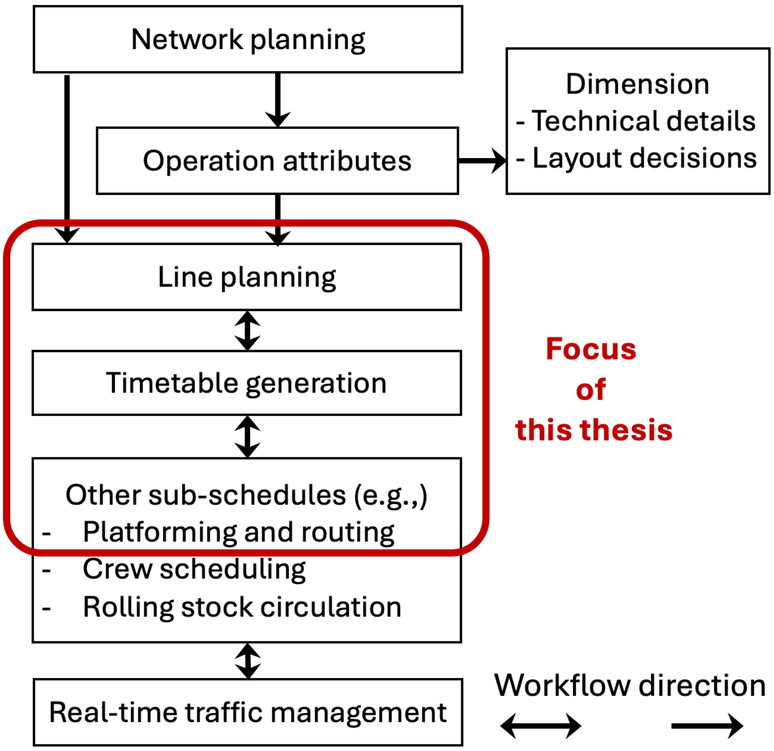


Figure 1.3 Railway planning stages covered in this thesis

1.2.2 Typical features of China high-speed railway

China's HSR network has been planned and constructed in phases. At the end of 2023, China had built 45,000-kilometres of HSR for passenger-only transport, finishing 80% of the '8+8' grid-based HSR master plan (Figure 1.4), which is updated from a '4+4' grid-based HSR master plan in 2016. China Railway comprises 18 railway bureaus where each bureau serves several provinces and/or municipalities: e.g. Beijing, Tianjin, Shanghai, and Chongqing. As a latecomer to operating HSR, the operation of China HSR shows some mixed infrastructural and operational features between Japanese Shinkansen, railways in Western Europe, and China conventional railways. Regarding rolling stock holding, China Railway possesses 3,541 sets of HSR multiple units, equivalent to 4,073 standard sets by regarding each 8-car multiple units as a standard set (Lu, 2023).

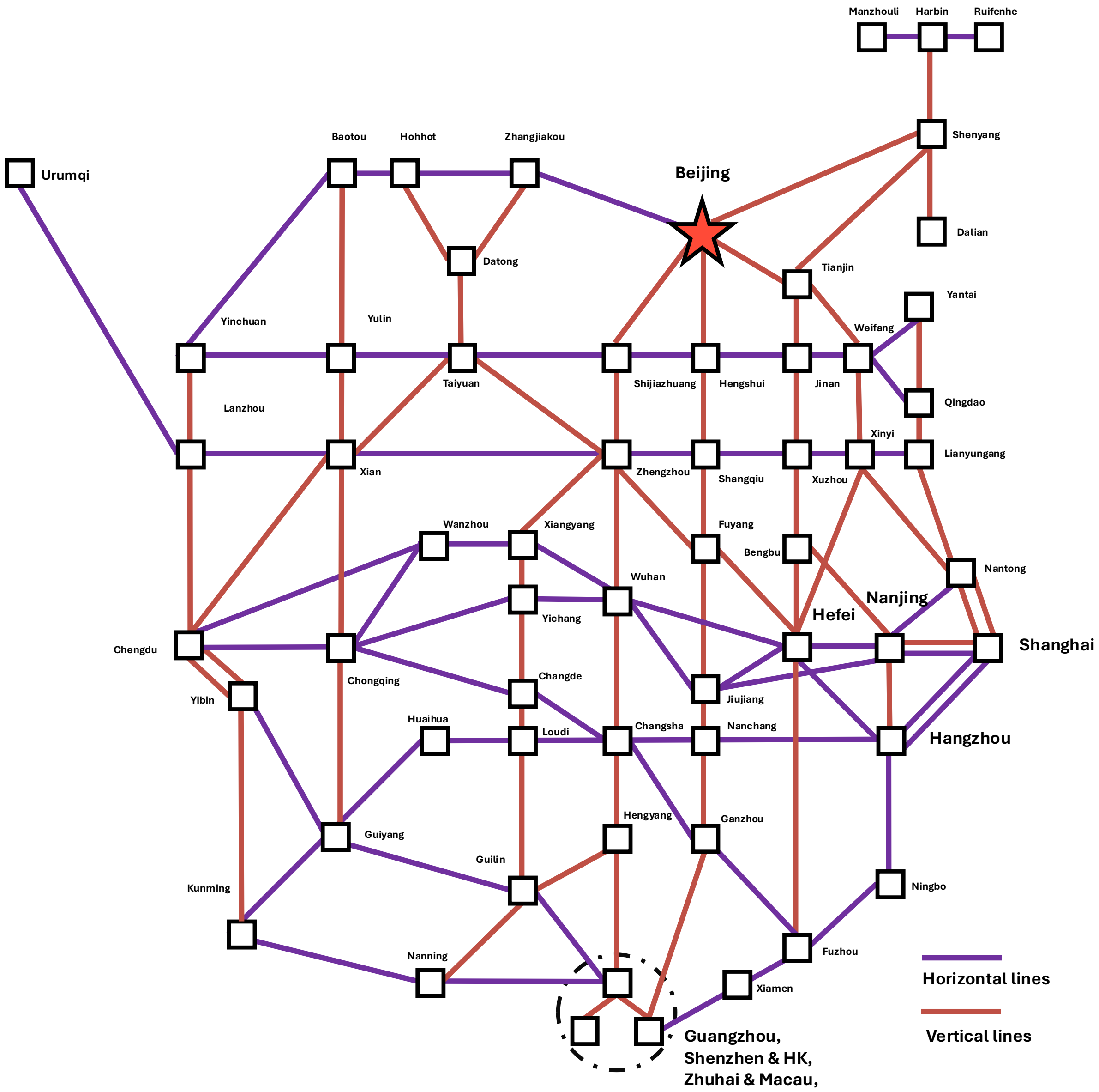


Figure 1.4 China mainland '8+8' gird-based HSR master plan.

1.2.2.1 Infrastructure and layouts

1. Basic layout: station and HSR lines

For intermediate stations, China HSR sets dedicated platform tracks for stopping trains, and the open tracks are separate for passing trains (see Layout A in Figure 1.5). This basic station layout is the same with most stations in Japanese Shinkansen. From a perspective of capacity utilisation, HSR stations like this can host over-taking behaviours between trains, and trains can run with a shorter headway. In many intermediate stations in Great Britian and Netherland, for example, open tracks also act as platform tracks. In that case, the next train on the last scenario needs a longer headway to ensure the operation safety.

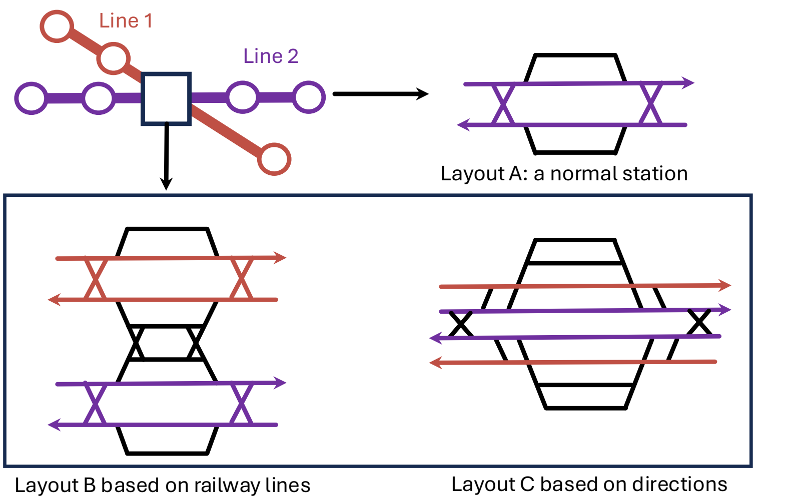


Figure 1.5 China HSR stations and layouts (Zhang, 2022)

For node stations, they connect different railway lines of the network. The Layout B and C in Figure 1.5 show typical HSR node station layouts in China. In layout B, line one and line two are arranged in parallel, while open tracks of line two are placed at each side of line one in layout C. From a capacity utilisation perspective, if there is not a train service that needs to cross these two lines, trains only running at these two lines are separated. Compared to layout B, a stopping train from line two in layout C has to cross the open track of line one, possibly affecting passing trains from line one.

2. HSR major cities with multiple stations.

HSR stations in most capital cities of provinces, provincial-level cities (Dalian, Shenzhen, Qingdao, Ningbo, Xiamen), and municipalities usually connect multiple HSR lines, significantly increasing the route conflicts if attempting to arrange direct trains linking every direction. Evidence has shown that stations linking multiple HSR lines can easily become the capacity bottlenecks of the network (Nie et al., 2021; Zhang, 2022). Figure 1.6 shows the direction roles of Nanjing South HSR station.

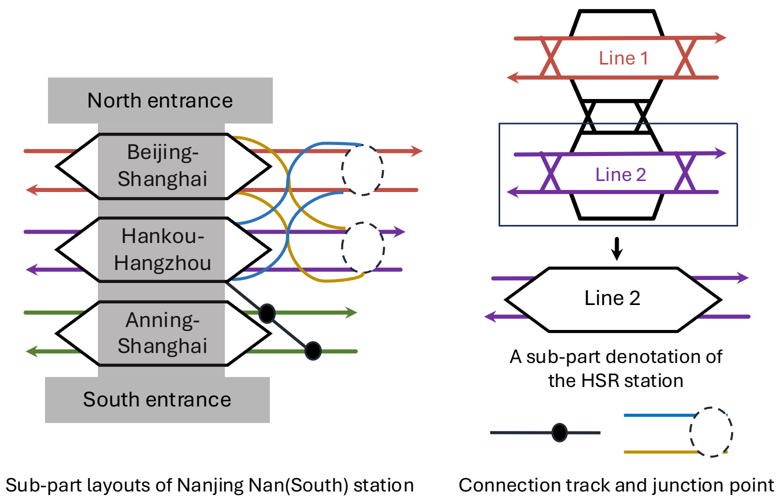
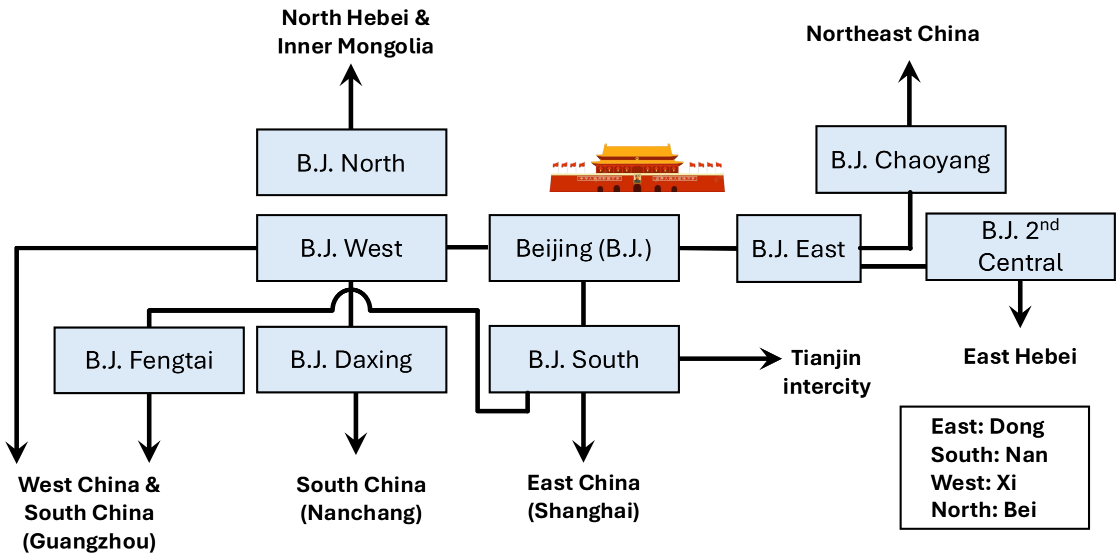
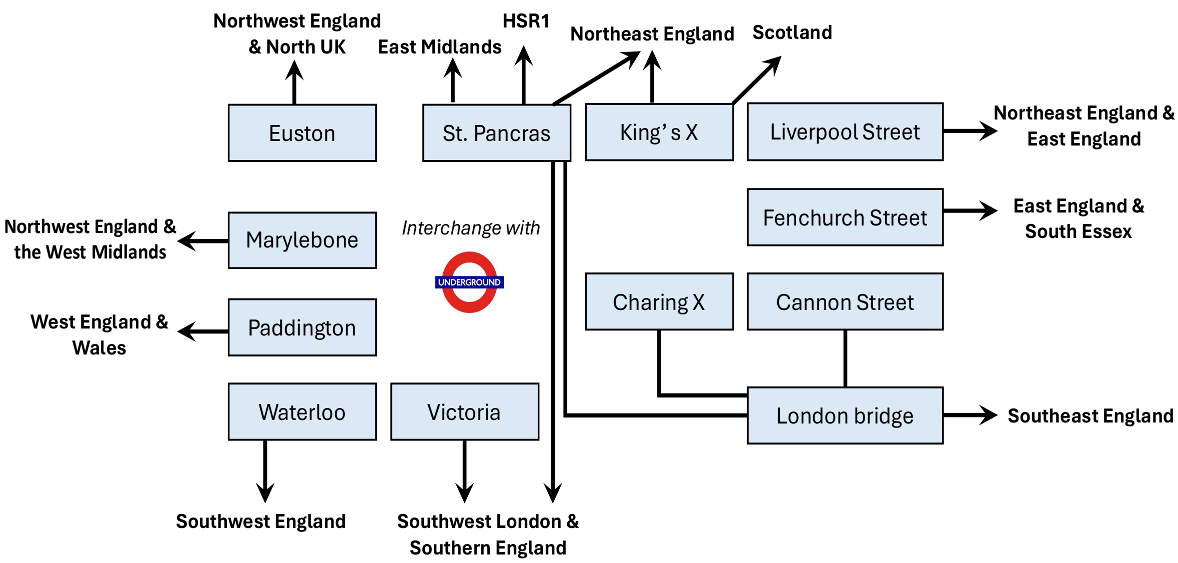


Figure 1.6 Direction roles of Nanjing south HSR station.

Previous literature has testified and proved that the terminus (T) and originating (O) stations easily suffer from insufficient capacity (Liao, 2021; Wang et al., 2023; Zhang, 2022). However, this situation is expected to change because the network planning of China HSR (China Government, 2020) plans that those cities will build more than one HSR station if that is not already the case. Each station accommodates specific directional roles in serving trains. In this thesis, these cities are collectively named as 'China HSR major cities'. Accordingly, China HSR major cities are equipped with depots and maintenance bases, so that stations can either be the terminus or origin of train services. From a capacity utilisation perspective, multiple stations can increase the total hosted train volumes, and affiliated track layouts can help avoid route conflicts for all directions. Figure 1.7 shows directional functions of HSR stations in Beijing and London.



(1) HSR stations in Beijing, China



(2) London mainland railway terminals, the UK

Figure 1.7 Directional roles of Beijing HSR stations and London terminals

3. Regional HSR:

Within the China HSR context, there is no rigid definition of regional HSR. The conceptualisation of China regional HSR is similar to Transport Express Régional of SNCF: regional HSR lines are translated as 'inter-city line' (China Railway, 2020), in which regional HSR trains are dispatched by a single railway bureau, train codes begin with 'C', and have the following features:

- Usually run within a single province or a metropolitan area.
- Might run between HSR stations and airports.
- Originating or terminating at stations in China major HSR cities.

Inevitably, metropolitan areas often boast connectivity through multiple HSR lines in China. Taking Nanjing and Shanghai as examples (Figure 1.8), these cities are linked by six HSR lines, offering passengers in both cities a choice of three stations. Similar connectivity is observed in other city pairs or groups such as Beijing and Tianjin, Chengdu and Chongqing, Guangdong Province. From a capacity perspective, regional HSR significantly enhances the service coverage of the HSR network. Multiple HSR lines, in turn, allocate more capacity for long-distance HSR trains. Concurrently, long-distance HSR trains can extend their service range into regional HSR lines, enabling passengers from smaller cities to access direct trains to major cities. Similar operational models can be seen with HSR in France and Italy, where HSR trains run on conventional railways to extend their service range into more cities.

To summarise, the main topology feature of China HSR network consists of 'artery' lines linking all HSR major cities and 'capillary' lines as regional HSR. Overall, China HSR plays similar roles of long-distance and regional railways classified by UIC (2018).

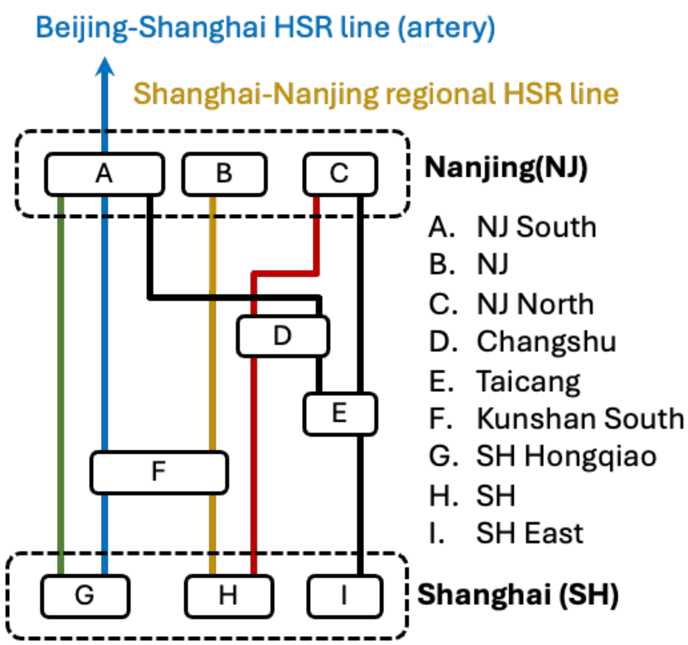


Figure 1.8 HSR links between Nanjing and Shanghai

1.2.2.2 China HSR trains and service outlook

1. General train types of China HSR

There are two ways of categorising HSR trains in China. The first way is determined according to whether a train path needs to traverse multiple railway lines: original-line and cross-line trains. Original-line trains run solely on a single railway line, while cross-line trains run on multiple lines. The full China HSR railway lines list can be read in UIC (2022). The second way, neither based on travel distance nor travel time, is summarised according to whether a train path needs to traverse multiple provinces. Contrasting with regional HSR trains described above, long-distance HSR trains are collectively defined in this thesis as train paths that need to traverse multiple provinces, excluding Xinjiang and Gansu, and Table 1.1 gives a conclusive summary.

Table 1.1 the China HSR train category

<table>
<tr><th rowspan="2">Dispatching unit</th><th rowspan="2">Province level</th><th colspan="2">Railway line</th></tr>
<tr><td>Original-line train</td><td>Cross-line train</td></tr>
<tr><td rowspan="2">Single bureau</td><td>Single province & metropolitan areas</td><td colspan="2">Regional HSR trains</td></tr>
<tr><td rowspan="2">Cross province</td><td colspan="2" rowspan="2">Long-distance HSR trains</td></tr>
<tr><td>Cross bureau</td></tr>
</table>

2. Priority of cross-line and long-distance trains

The timetable generation for China's High-Speed Rail (HSR) follows the same principles as China conventional railway timetabling. Train paths are graphed into the timetable by the following priority (Zhang, 2019):

- The first rank: long-distance and cross-line trains,
- The second rank: long-distance and original-line trains,
- The third rank: other trains dispatched by a single railway bureau,
- Adjusting lower ranked train paths first if conflicting with higher ranked trains,
- Within the same rank, cross bureau train paths hold a higher priority than single bureau train paths.
- Rolling stock circulation is determined after the timetable generation.

Presently, two top 'artery' HSR lines, which are Beijing-Shanghai line and Beijing-Guangzhou-Shenzhen line are occupied by a large number of cross-line trains, both exceeding 70% in proportion. The rationale behind this priority arrangement has two reasons (Li, 2022; Li, 2020; Xu, 2020; Zhang, 2019): the first is that the long running time of long-distance HSR trains only allows limited time windows either for reasonable arrivals or departures. Arranging their train paths first can prevent graphing infeasibility. The second is that the China HSR put significant efforts to maintain direct transports to/among China HSR major cities. From a capacity utilisation perspective, fixed positions of high ranked train paths limit the potential for capacity increase or any timetable adjustments.

Figure 1.9 shows a train density comparison between China HSR and Japan Shinkansen collected on 13th January 2024 (modified from a GIS work by Li & Pi (2024) (Other details can see Appendix A). Beijing-Shanghai line and Beijing-Guangzhou-Shenzhen line has more HSR line connections, while Japan Tokaido and Sanyo Shinkansen does not have HSR line connection with each other. The Tokyo-Osaka HSR train density of Japan Shinkansen is much higher than these two China HSR arterial lines, suggesting the negative capacity impacts of running excessive cross-line trains in China.

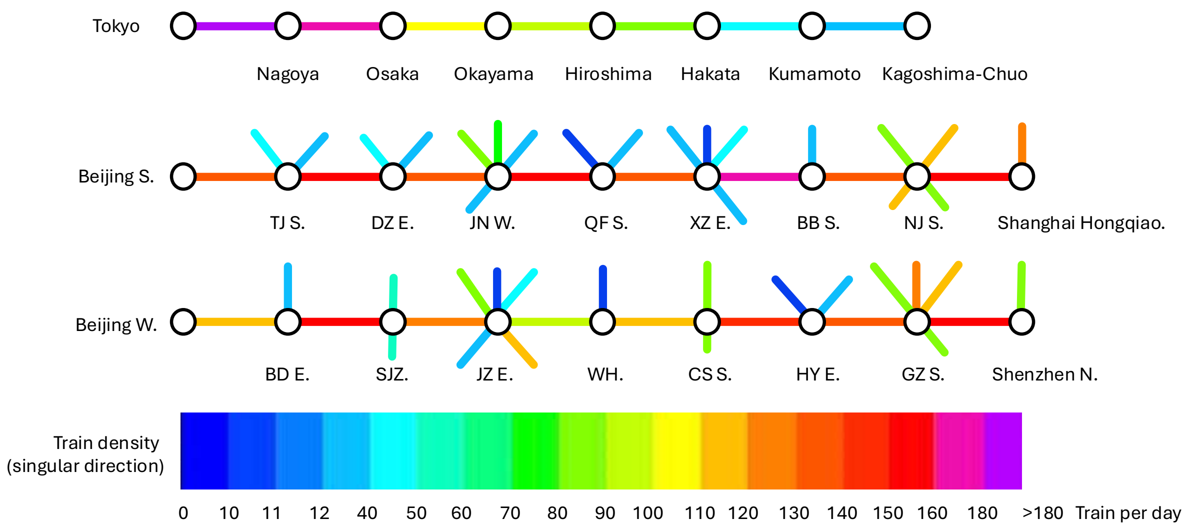


Figure 1.9 HSR train density comparison between China and Japan (Li & Pi, 2024).

3. The stop plan and timetable periodicity

China HSR shows similar stop plan features when comparing trains in the UK, Western Europe, and Japan Shinkansen. A summarised figure (Figure 1.10) is given based on Zhang et al (2020).

Starting with trains in in Western Europe and Japan, one of the common scheduling styles is based on a hierarchical station and stop setting. Stations are classified as big, medium, and small stations, and trains are generally divided into fast, semi-fast, and local trains. Fast trains (F-trains) stop at big stations, semi-fast (S-trains) trains stop at big and a subset of medium stations, and local trains (L-trains) stop nearly at every station. However, Japan Shinkansen set medium and small stations as selective stops, so that the overall service frequency can match the varying passenger OD. Notice that different countries adopt distinctive names for this three-level train setting, and a summary is given in Table 1.2.

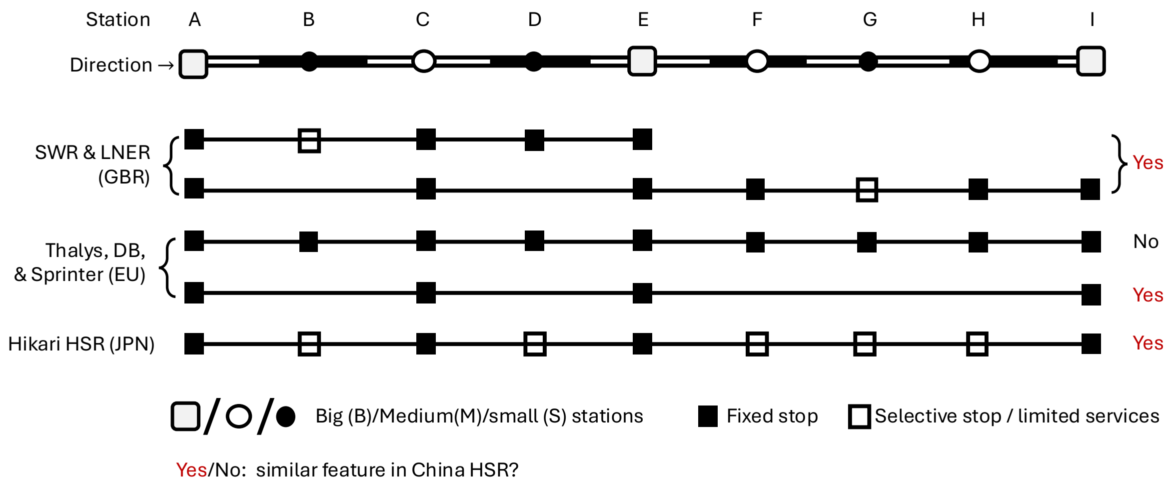


Figure 1.10 Examples of stop plans based on distance and station levels.

Moving forward, the terminus stations and stop setting of Southwestern Railway (SWR) and London North Eastern Railway (LNER) services in the UK are influenced by two key factors: the outbound distance from the origin station and the hierarchical intermediate station levels. In Figure 1.10, Station A represents London terminals, while Stations E and I serve as terminus stations located at varying travel distances. For instance, in the case of LNER, the train service between Stations A and E corresponds to the route from London King's Cross to York. Similarly, the train service marked from Station A to Station I represents longer-distance services from London King's Cross to northern destinations, such as Edinburgh Waverley and Glasgow Central. In the case of SWR 'mainline service', according to their own definition, Station A is London Waterloo. Station E and I refers to Southampton Central and Poole (or Weymouth), respectively.

As a summary, China's HSR stop planning incorporates both running distance and station hierarchy, but its stop-setting patterns are less distinct compared to the more structured approaches observed in Western Europe and Japan.

Stop plan features shape the service pattern of a railway network. Scholars in China (Nie et al., 2021) believe 'well- characterised' stop plans have a correlation with the periodic timetable. In a periodic timetable, trains with similar or the same stop plan repeat their services at set time intervals (e.g., every 15, 30, 45, or 60 minutes) in an operating day. By introducing periodic timetables, the full-service contents of a railway network can be easily memorable for the public.

As for aperiodic timetables, trains with similar or the same stop plans do not need to repeat on the same operating day. Therefore, it is possible to schedule multiple stop setting combinations to form the service network. Currently, China HSR adopts the aperiodic timetable as the main norm, but periodic services can be seen on some regional HSR and partial F-trains.

Table 1.2 Examples of three-level-based train setting

Country	F-train (level 1)	S-train (level 2)	L-train (level 3)
Netherlands	Intercity	Inter Region	Intra Region
Germany	Intercity Express	Intercity, Europe City	S-bahn, Regional-Bahn Regional Express
Switzerland	Intercity (Neigezug)	Inter Region, Reseau Express	S-bahn, Reseau Express Regional
Spain	AVE, Larga Distancia	Media Distancia	Avant
Italy	Frecciarossa (360km/h)	Frecciargento (250km/h)	Frecciabianca (200km/h)
Japan	Nozomi	Hikari	Kodama

1.2.2.3 Worldwide HSR comparison

Despite the unique topology and service outlook of China HSR, some similarities between China HSR and other railway contexts can still be observed. Table 1.3 gives a topology and timetable summary of major HSR countries, and Table 1.4 lists representative HSR lines as detailed layout comparison. Table 1.3 is an extended work based on Drábek et al. (2021) and railway websites.

Table 1.3 HSR in representative countries

Country	Topology	Service range	Service scheduling	Periodicity
China	Fully-connected N.	HSR only	Network-based	D, E
Germany	Fully-connected N.	Mixture	Network-based	B
Japan	Tokyo-based H.	HSR only	Line-based	B
France	Paris-based H.	Extending	Line-based	C
Spain	Madrid-based H.	Mixture	Line-based	D
Italy	Bologna-based H.	Extending	Network-based	D

The explanation (Drábek et al., 2021) of Table 1.3 is as follows:

1. **Topology**

- Fully-connected network (N.): multiple railway lines are connected as a network.

- Single-hub-spoke (H.): several railway lines are built radially around an obvious centre.

2. **Service range:**
   - HSR only: HSR trains only run on the newly built or upgraded HSR tracks.
   - Extending: HSR trains can extend their services on conventional railway networks while HSR tracks only allow HSR trains running.
   - Mixture: the conventional and HSR tracks are connected as a network, and HSR trains and other trains use the same infrastructure.
3. **Service scheduling:**
   - Network-based: scheduling timetables at network levels.
   - Line-based: scheduling timetables line by line and allow some branch services.
4. **Timetable periodicity:**
   - Fully periodic (A) and almost fully periodic: only a minor irregularity (B)
   - Periodic services in peak hours only (C)
   - Partially periodic: only a certain extent of periodic services in some areas (D)
   - Aperiodic: no periodic services (E)

Table 1.4 Selected HSR line and station information

Country	Railway line	Length	Number of stations	Line connections
China	Beijing-Guangzhou-HK	2430 km	47	12
China	Shanghai-Kunming	2266 km	55	13
China	Beijing-Shanghai	1318 km	23	12
Japan	Tokyo-Osaka-Hakata	1068 km	35	1
Spain	Madrid-Barcelona-Figueres	752 km	9	3
France	Paris-Lyon-Marseille	718 km	10	10
Germany	Berlin-Hannover-Wurzburg	606 km	8	＞ 10
Italy	Turin-Milan-Roma-Salerno	591 km	9	9

To determine the line connection amount, two conditions must be satisfied:

- There exists track connection linking two HSR lines.
- The HSR train services use the track connection to extend or cross-line.

Regarding timetable periodicity, most countries introduce various levels of periodic services. It can be argued that timetable level (A) might be entirely notional/theoretical. From a practical perspective, timetable level (B) is more reasonable by allowing varying service frequency during in/off peak hours to match time-dependent passenger demands. In terms of HSR timetable periodicity in China, only partial networks, such as trains running between Shanghai and Nanjing, Beijing and Tianjin, Guangzhou and Shenzhen, have periodic inter-city train services.

Regarding network topology, only China and Germany have fully-connected networks, and both share a similar level of railway line connections, but China HSR has much longer line length and more station numbers in general. Japan can only reach similar levels of station number and line length after combining Tokaido and Sanyo shinkansen, however, there is no 'cross-line' train running from Tokyo to Hakata (Japan Railpass, 2023). The only line connection is an extension, not a tree branch, from Hakata to Kagoshima-Chuo. Additionally, the operating distance of China HSR trains are much longer than any other trains in the world. Tian (2018) calculated the total operation distance of China's HSR trains railway timetable in 2017 shown as the Table 1.5.

Regarding service range, France and Italy can extend HSR services on conventional railways, and Germany adopts a mixture between Intercity Express (HSR name) with other trains, making the service scheduling seemingly challenging. But the network statement of EU (Rail Net Europe, 2021) show that HSR trains in those countries have higher priority in the timetable generation, which is similar with cross-line and long-distance HSR in China. Although Japan, France, Spain, and Italy HSR are all shaped as one-centre trees, all HSR lines and trains terminates at Tokyo station, most HSR trains terminates at stations in Paris, while HSR lines and trains go through the Madrid and Bologna station, leading to different distances of service range.

Table 1.5 Operating distance distribution of China HSR trains in 2018.

Distance	**HSR train number**	**Percentage**
0-200km	210	9.78%
200-500km	613	28.54%
500-1000km	622	28.96%
1000-1500km	466	21.69%
1500-2000km	143	6.66%
2000km	94	4.36%

1.2.2.4 Summary

To summarise, China HSR adopts the similar infrastructure layout (double-open-track with platform tracks) as the majority of Japan Shinkansen, maintaining a certain level of independence in operation tasks such as train running directions and platforming. Most China major HSR cities have multiple HSR stations to accommodate trains from/to certain directions, increasing the total railway capacity and potentially preventing route conflicts. The railway line length and number of stations on China HSR are much longer and greater than any other railway systems. Based on this, China HSR still has complex line connections, which is similar to Germany. China aims to operate HSR services as the positions of long-distance and regional railways classified by UIC (2018). Notably, a clear timetable generation priority is given to cross-line and long-distance China HSR trains. Other service plan features such as hierarchical stop or train settings are less obvious, and aperiodic timetables are typical of China HSR.

1.2.3 Identified common challenges

1.2.3.1 Trade-off between direct transport versus interchange

Beginning with a sample railway network with six cities shown in Figure 1.11 (1), if a railway operator aims to ensure direct transport for every passenger OD, it requires at least six train lines for each direction (i.e. 12 types of train service) in Figure 1.11 (2). Supposing track layouts near the station B and C can support trains for all possible directions, it can be estimated that there would be a significant amounts of routing conflicts among these 12 types of service.

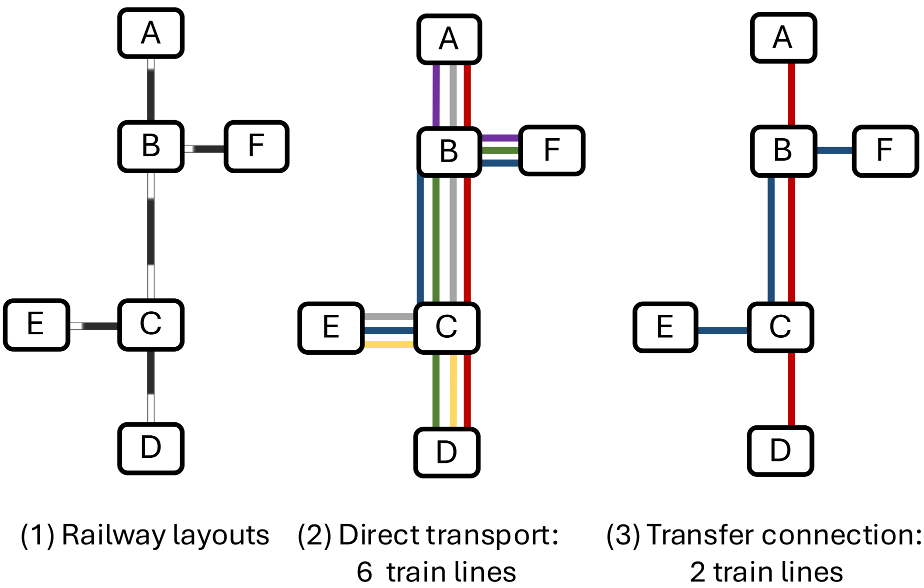


Figure 1.11 Railway links and two service scheduling styles.

Comparatively, if allowing passenger transfer at station B and C, only two train lines (four types of service) are needed in Figure 1.11 (3), and route conflict types are less than Figure 1.11 (2). Furthermore, if station B and C adopts layout B in Figure 1.5, there is no route conflict among all trains, both train lines could arrange a high frequency of services. In fact, Figure 1.11 (3) is how cities operate underground (metro) or some urban railways where only limited train line types are provided (e.g., Docklands Light Railway in London).

For a more complex railway network or those stations linking multiple railway lines, it triggers the discussion regarding the trade-off between direct transport and interchange. If track layouts cannot support direct transport between all ODs, is it possible to arrange successful transfer connections? if track layouts can support trains for all directions, what would be the suitable underlying capacity utilisation strategies and rationales? What is the boundary between requiring passengers to transfer or arranging a direct train that causes extra route conflicts? Combining with China HSR contexts, what are the potential capacity benefits of reducing the priority of cross-line trains?

1.2.3.2 Increasing coverage versus challenging in-vehicle time & capacity

Stop setting of a train line inherently leads to two conflicting aspects. The first conflict lies between railway operators and passengers. Any additional stops at the middle of a passenger journey (i.e. OD) will extend the in-vehicle time (Hu et al., 2023). On the other hand, adding one extra stop will boost more direct links among passenger ODs. In long distance travels, the additional stops and the extended in-vehicle time for passengers will be much longer compared to fast trains.

Table 1.6 Travel time comparison between F-trains and L-trains in the UK and China HSR

From	to	In-vehicle time	Total stops	Max. speed	Distance
Southampton Airport	London Waterloo	1h 11min	4	160km/h	112km
		1h 46min	14		
Beijing Nan (South)	Shanghai Hongqiao	4h 29min	3	350km/h	1318km
		6h 34min	14		

Table 1.6 presents a comparison of travel times between two sets of passenger trains with the highest and lowest number of stops. From Southampton Airport to London Waterloo, adding 10 more stops only extends the In-vehicle times by 35 minutes on this 112km journey. Conversely, due to the greater acceleration and deceleration penalties characteristic of HSR rolling stock, adding 11 more stops on a 1318 km China HSR line results in an increase of 125 minutes.

Although the HSR train with 12 stops can serve 66 passenger ODs, there is uncertainty regarding the types and numbers of passenger ODs attracted by this train.

Reducing extra stops of each train line and increasing the combination of train lines with fewer stops seems to be a good approach to reduce the In-vehicle time and enhance the attractiveness. However, this strategy gives rise to the second conflict between the capacity and diverse combinations of train lines, which has partially explained in Figure 1.10. Consequently, it is impossible to arrange non-stop direct trains for every passenger OD.

In short, there seems to be an 'impossible trinity' between coverage (if providing direct services to all passenger ODs), frequency, and In-vehicle time due to the limited railway capacity, requiring a systematic analysis and review in this thesis, and corresponding solutions.

1.2.3.3 Scheduling styles versus capacity utilisation features

The capacity utilisation not only reflects the passenger demand features across different countries but also incorporates subjective preferences when scheduling HSR services, such as partially shown in Table 1.2 and Figure 1.10. Other scheduling preferences can be based on total passenger demand volume, turnover (passenger-kilometre), or equality among passenger ODs. Higher demand volumes can be seen on 'short' distance railway travels, but 'long' distance railway travels contribute to higher turnovers. such as in China HSR (Li, 2022; Li, 2020), France (Autorite De Regulation des Transports, 2019), the UK (ORR, 2023), and Germany (Bundesnetzaguntur, 2022). Interestingly, both short or long-distance railway travels are associated with large and major cities in these countries, triggering the capacity challenges on busy lines or around the major stations over which types of travel should be prioritised.

A common service scheduling style is that F-trains are designed for fast travels with less intermediate stops across major cities, whether in the UK, Europe, or East Asia. In Tokaido Shinkansen, the frequency of F-trains, which is 12 trains per hour (Japan Railpass, 2023), is much higher than S-trains and L-trains (2-3 trains in total per hour), while in China HSR, the total frequency of F-trains are far less than other trains with more stops, and there is less obvious frequency differences between F-trains, S-trains, and L-trains in Netherland (Luijt et al., 2017) or major services in the UK (Rabot, 2023). The service pattern difference between China and Japan could be seen as the results from city population scales (i.e. potential travel demand difference) along the HSR lines. Trains in Netherland and the UK, however, offer more equal service frequency for more passenger OD compared to Japan Shinkansen. On the other hand, retaining direct transport connections, represented by cross-line trains across the China HSR network, can also be prioritised as a means of scheduling an equal and fair service network.

Whatever the circumstances, the determination of a service plan, containing any service pattern or scheduling style, should satisfy the demand features of the corresponding railway context and try to deliver a balance between service quality and railway capacity.

1.2.3.4 Summary

An optimal capacity utilisation of a passenger railway network must sufficiently consider service quality of a schedule plan for most ODs. Conversely, railway capacity serves as an effective assessment method to verify and validate (V&V) the deliverability and feasibility of a service plan proposal. As stated in Section 1.1, the optimal capacity utilisation of a passenger railway network should make full use of the existing infrastructure where the trains are expected to provide direct, fast, and frequent services to the most passenger ODs. This essentially requires this thesis to establish the theorical relationship between railway capacity and service qualities of a passenger railway and provide a comprehensive method to compare the capacity and passenger experience trade-offs between different service plan proposals.

1.3 The research aim and objectives

To schedule a service plan that contains fast, frequent, and ideally direct transport service for most ODs across the network, the resulting shared challenges motivate this thesis to formulate the main research aim:

To design a systematic timetabling solution which could improve both capacity utilisation and timetable service quality.

The objectives of this research are:

1. To identify the passenger and railway operator factors affecting railway capacity.
2. To assess the key capacity-related service qualities in passenger railways.
3. For a given railway network, develop a line planning method which can test various service scheduling styles.
4. For a given railway network and a proposed train line plan, develop a method to measure the capacity utilisation of the feasible timetable.
5. To determine an optimal timetable that has either: a) reduced capacity utilisation (i.e. timespan) whilst maintaining similar service qualities, b) improved service qualities while retaining a similar capacity utilisation level, or c) reduced capacity utilisation and improved service at the same time.

1.4 Research significance

From a perspective of academic knowledge, this thesis will make the following contributions: it will summary the methodological equivalence of timetable scheduling and capacity assessment and analysis. By refining the passenger travel experience into a service indicator system, it will establish a theoretical relationship between service quality indicators and capacity utilisation indicators. The thesis will identify a collaborative optimisation method suitable for enhancing both service quality and capacity by reviewing previous studies and analysing the key affecting factors of railway capacity from the aspects of passenger and railway operation. It will also propose a comparative framework for service quality indicators and capacity utilisation indicators, enabling a trade-off analysis between passenger convenience and operational efficiency for railway operators.

From a perspective of practical operation, this thesis will help railway operators or planners integrate passenger experience into the service planning process, represented by line planning and timetable generation. It will allow the scheduling of high-quality timetables that fully utilise the capacity under varying railway contexts, represented by infrastructure and facility layouts, demand features, planning traditions. Additionally, this thesis will systematically outline the representation form of railway capacity within major railway planning stages and summary the infrastructural and operational affecting factors of railway capacity from macro, meso, and micro aspects. This can help railway operators and planners to identify the key capacity-affecting factors relevant to their current operation status and to adopt improvement measures.

1.5 Thesis structure

Chapter 1 identifies three common challenges through an analysis of railway planning stages, operational features, and capacity utilisation across various countries. These challenges, which are balancing direct services and transfers, managing the trade-off between coverage and short travel times, and exploring scheduling complexities for a specific railway context, are distilled into the research aim, with specific objectives set to achieve this aim.

Chapter 2 reviews the definitions, influencing factors, measurement methods, and passenger considerations by categorising key research elements. It constructs a comprehensive knowledge framework and identifies key aspects for capacity planning in passenger railways.

Chapter 3 establishes well-founded passenger assumptions based on OD groups. Through a macro-to-micro analysis, it develops a general approach to management, highlighting key points in operational attributes and timetable details. A workflow for capacity optimisation is formed, defining a staged solution for line planning and timetable generation.

Chapter 4 first introduces a method based on the maximum spanning tree to locate interchange points without manual setting based HSR connecting major cities. Line planning is modelled using the set covering problem, aiming to minimise intermediate stations, ensuring equitable access to fast services for all OD pairs.

Chapter 5 builds a mesoscopic-level integrated capacity measurement& timetable generation model based on discussions in Chapter 2 and validate the effectiveness with the macroscopic model. The model is able to generate a feasible timetable with a minimum timespan. As selective workflows, four potential capacity improvement methods are proposed, tested, and discussed.

Chapter 6 summarises the findings of the previous chapters, present the overall conclusions and contributions of this thesis, identify the limitation and future work directions based on this thesis.

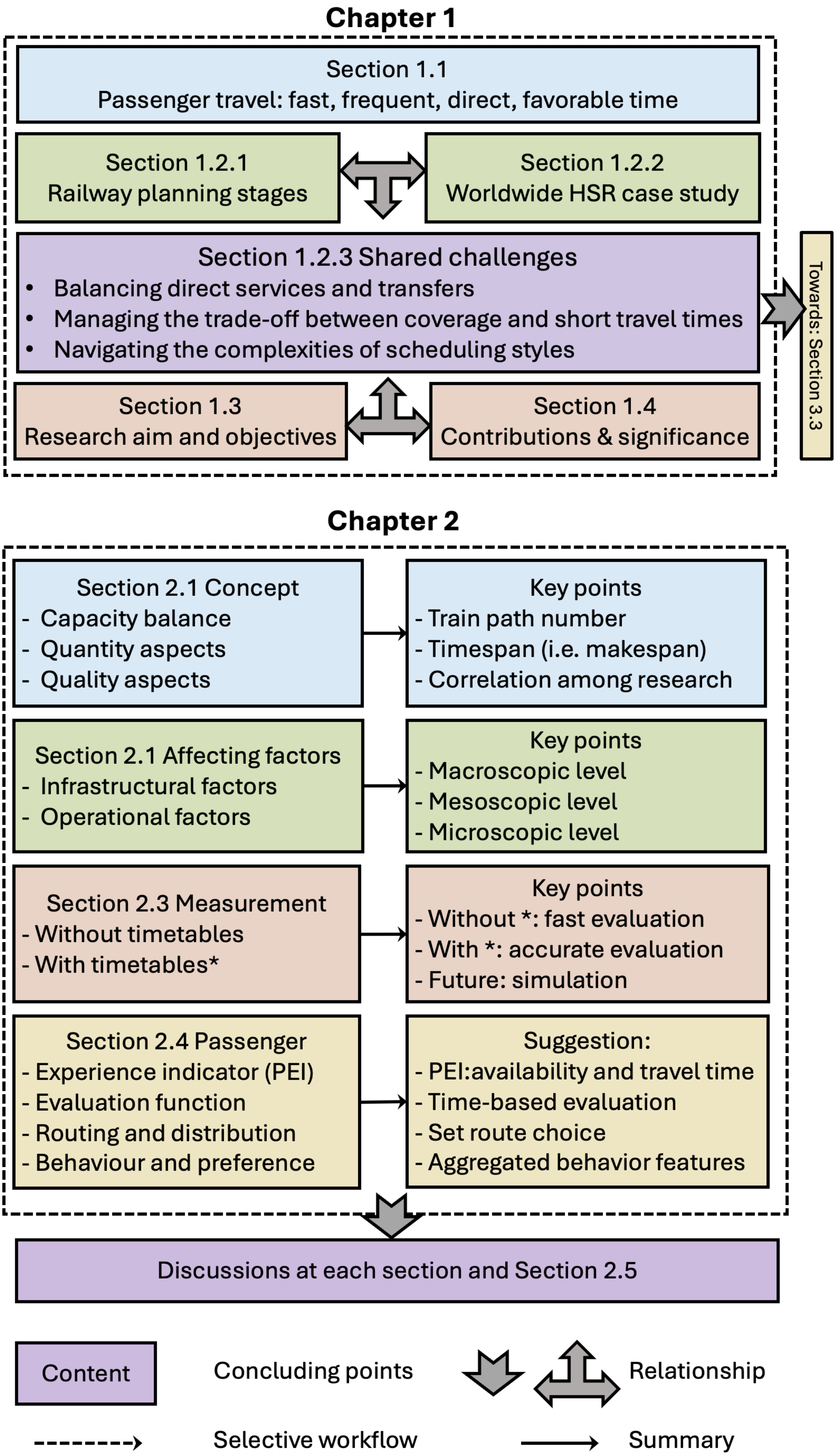


Figure 1.12 The top half of the thesis structure

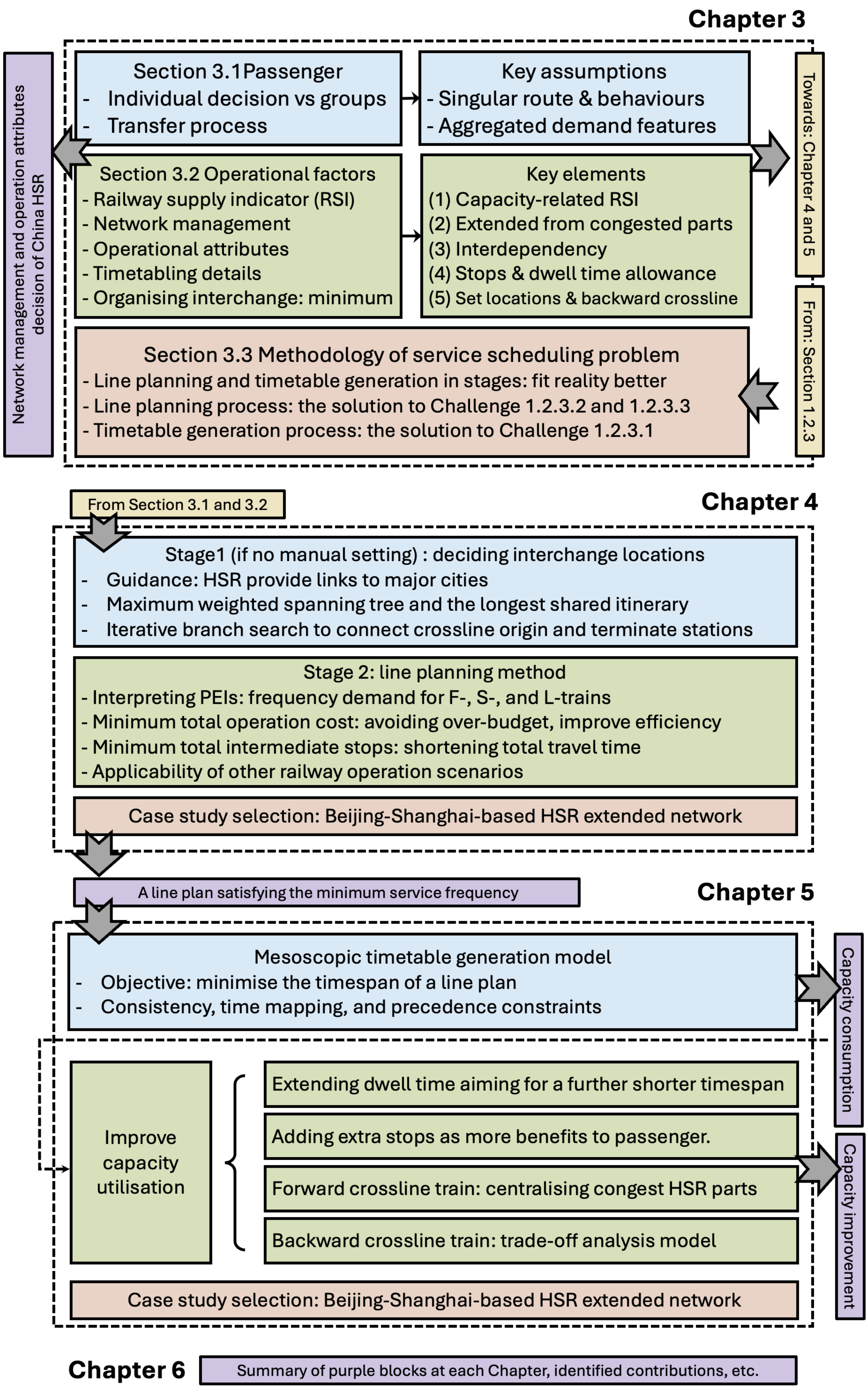


Figure 1.13 The bottom half of the thesis structure

Chapter 2 Literature review

The capacity utilisation of a passenger railway network is jointly affected by the infrastructure layout and the structure of the service plan (i.e. the line plan and its produced timetable) in detailed. At the same time, scheduling a service plan with high service qualities are the goal of improving the capacity utilisation of the network and passenger experience. Based on this view, the literature review of this thesis needs to answer the following questions, and the following paragraphs are titled and structured accordingly:

- What is the concept framework of railway capacity? (Section 2.1)
- What are the affecting factors of railway capacity? (Section 2.2)
- How to measure the railway capacity? (Section 2.3)
- What is the mathematical linkage between railway capacity and the timetabling process? (Section 2.3.2)
- How previous research considers demand features and evaluates passenger experience (section 2.4)?
- How demand features and passenger experience evaluation affect the railway capacity and other supplies by railway operators (Section 2.4.1)?
- What assumptions and conclusions can be collected? (Section 2.4.2-2.4.5)
- What are the academic gaps among previous research works? (Section 2.5)

2.1 The concept framework of railway capacity

The debate on the definitions of railway capacity stems from different emphasis on ‘volume (quantity)’ and ‘value (quality)’ aspects of capacity (Armstrong & Preston, 2017). Despite no unified definition achieved over the years, a common pattern between different railway capacity definitions is that the capacity is the maximum 'qualified' rail transport products by subjecting to limited railway resources and operational constraints. In the former sentence, the ‘maximum’ represents the volume, while defining a ‘qualified’ rail transport product depends on different stakeholders (i.e., railway operators, planners, passengers, infrastructure managers).

2.1.1 Capacity balance factors from UIC406

Train speed, train number, stability, and heterogeneity are four inter-related factors and are identified together as capacity balance factors in UIC (2004), which is one of first standards indicating the relationship between quantity (i.e. train number) and quality aspects (i.e. stability) of railway capacity. More derived concepts of quantity and quality aspects of railway capacity will be introduced in the following sections (Section 2.1.2 and 2.1.3).

These four balance factors are represented as four axes, with the capacity features of a railway context forming a quadrilateral defined by the chords from the four quadrants in a qualitative manner. The capacity balance remains constant for a given level of infrastructure provision; therefore, any adjustment to one of the four contributing factors must be offset by corresponding changes in one or more of the others. Heterogeneity here is the diversity of train traffics caused by the train line types (Khadem Sameni, 2012). Other heterogeneity definitions can be seen in Lusby et al. (2018). The stability here represents the margins and buffers that have to be added to the running time of trains and between trains to ensure that minor delays are absorbed (UIC, 2004).

Figure 2.1 shows capacity utilisation scenarios for different HSR railways and a metro system, attributed to each background's own infrastructure features and layouts, operational attributes and service scheduling styles, and diverse transport needs.

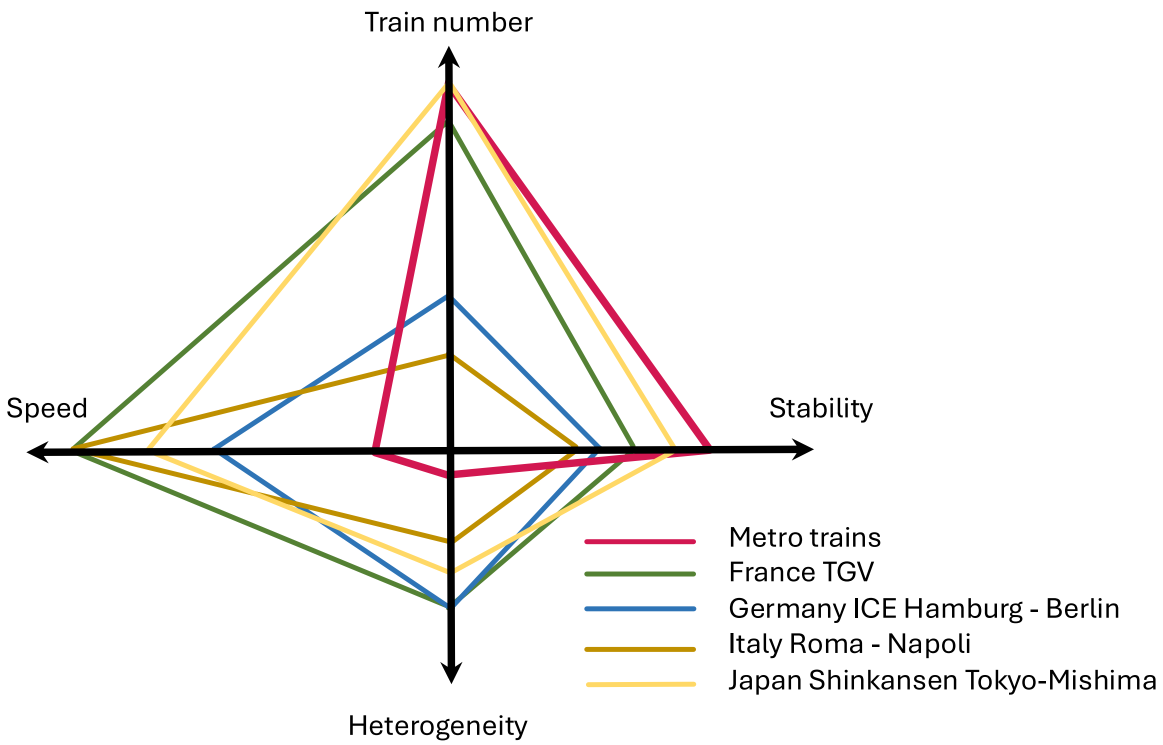


Figure 2.1 Capacity balance factor features of different railway systems (UIC, 2015)

2.1.2 Quantity aspects of capacity in passenger railways

There are three familiar capacity concepts in the literature which are theoretical capacity, practical capacity, and capacity occupation rate (Krueger, 1999). Theoretical capacity is accumulated from the homogeneous train traffic subject to a strict minimum headway and removing the maintenance time and other allowances. Theoretical capacity is the upper limit of a railway system. Practical capacity, on the other hand, needs to take account of infrastructure constraints, train traffic mix, and a certain level of stability. The capacity occupation time rate (UIC, 2013) is the ratio of occupancy time or used time by train paths to a defined time period in a given railway network, also commonly known as capacity utilisation or consumption (level). Other similar capacity definitions can be obtained from Khadem Sameni (2012) and Borndörfer & Lamorgese (2018), and a partial summary is given as Table 2.1.

Table 2.1 Partial terminology of railway capacity (Besinovic & Goverde, 2015)

Term	**Parts of similar terms**
Theoretical Capacity	Design Capacity (Ryes et al., 2013), Absolute Capacity (Kozan & Burdett, 2005)
Practical Capacity	Achievable Capacity (Ryes et al., 2013), Effective Capacity (Goverde & Hansen, 2013)
Capacity Occupation	Infrastructure Occupation (UIC, 2004), Occupancy Time (UIC, 2013), Capacity Utilisation (e.g., (Goverde, 2007) , (Petering et al., 2016)) Used Capacity (Abril et al., 2008)
Capacity Occupation Rate	Utilisation Rate (Landex, 2008) Capacity Consumption (UIC, 2013)

Depending on the research objectives or infrastructure details required, railway capacity can also be divided into the station capacity, the line capacity, the corridor capacity, the junction capacity, the network capacity, etc. (Mussone & Wolfler Calvo, 2013; UIC, 2013). The railway station is a special case because major stations are the key nodes of a railway network, subject to long dwell times and various routing conflicts. They can easily suffer insufficient capacity resulting in train congestion around stations, particularly if train services are growing dramatically, as for example in the cases of London Gatwick Airport and East Croydon in the UK (Network Rail, 2020). To describe this dynamic traffic congestion at railway stations, an inventory (rate), which was originally used in freight yard scenarios, can be introduced to represent the number of current dwell trains in a station by jointly calculating the arrival and departure patterns of trains (Dick, 2023).

When denoting the railway capacity of a timetable, it is commonly accepted (Besinovic & Goverde, 2015) that railway capacity is defined as the maximum number of completed train paths that can be run in each railway infrastructure within a given period, subject to a set of operational rules. The dual description is defined in this paper as capacity occupation: the minimum timespan (i.e. makespan) of running and competing a certain number of train paths occupying a given railway infrastructure, subject to a series of operational rules, which accounts starting from the departure time of the first train to the arrival time of the last train in a timetable (Petering et al., 2016; Sparing & Goverde, 2017; Zhang & Nie, 2016). The difference is drafted as Figure 2.2.

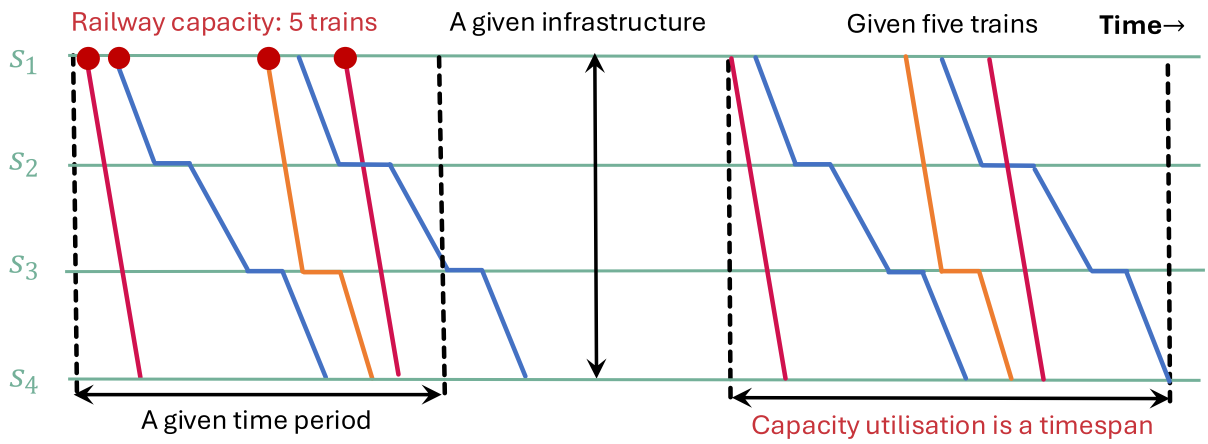


Figure 2.2 Different railway capacity denotations

Liao et al (2021) further divided the capacity of the timetable into infrastructure-driven capacity and vehicle-driven capacity respectively. Vehicle-driven capacity is the produced timetable after train paths have matched all rolling stock circulations, while the infrastructure-driven capacity does not necessarily take into account rolling stock circulations. Liao (2021) pointed out that the delivery process, driver preparation time, and maintenance requirements of rolling stock made vehicle-driven capacity value usually stay below the infrastructure-driven capacity, especially in rolling stock shortage situations.

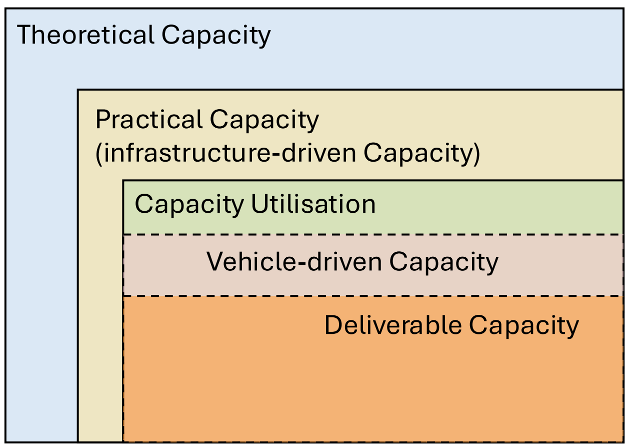


Figure 2.3 Relationship among railway capacity concepts

Moreover, errors, delays, disruptions, and other unexpected factors might happen when a timetable is put into practice. Therefore, the timetable will not perform exactly as planned. Researchers mentioned at Bešinović et al. (2021) and Robenek et al. (2016) regarded how timetables actually work as one kind of railway capacity as well. To avoid confusion, the real-time timetable performance in each railway network is defined as the deliverable capacity. The relationships among different definitions can see in Figure 2.3. The dotted line in the Figure means that some exceptions can happen. For example, if a timetable is strictly scheduled by following the rolling stock circulation (Liao et al., 2020), the capacity utilisation value equals the vehicle-driven capacity. Similarly, if no significant changes and cancellations happen, the deliverable capacity value equals the capacity utilisation value.

2.1.3 Quality aspects of capacity in passenger railways

A concern that comes from only viewing quantity aspects of railway capacity is that this single dimension cannot represent the full nature of rail transport products. A simple example is that the same quantity capacity value might be measured from two different railway networks and/or under different operation situations. Considering the different transport needs, operation styles & objectives, railway system performance etc, this thesis defines quality aspects of railway capacity as efficiency and performance indicators.

The indicators of railway efficiency depend on stakeholders' viewpoints (ITF, 2019). The basic capacity-related efficiency indicators in passenger railways are mainly railway usage data based on distance and products, such as train/person/seat kilometres in a defined time, as reported by the UK (ORR, 2023). In a reserved-seat environment, Li (2020) and Zhang (2012) believe that the capacity provision of a timetable should effectively meet the demand features across the network. The capacity-related efficiency in this case is based on how the supplies of a timetable meet the demands. Comparatively, in an unreserved-seat environment, the capacity-related efficiency is regarded as crowdedness level. If a train is too crowded, it suggests demands for additional train services (ORR, 2023). Similarly, In urban railways or other railway contexts whether service pattern is simple and periodic, passenger flow during peak/off-peak hours represents the intensity of transport demands (Cheng, 2018). Next, scholars like Burdett (2015) believe that every railway line has different 'prices' because of profitability or potential economic contributions to regional development, and railway capacity should not simply be the sum of train paths. Therefore, it is a natural thought to combine those usage data with economic terms that are based on revenue or cost data (ITF, 2019).

Regarding performance indicators, delays, disruptions, and other unexpected errors can happen during the real-time traffic management stage of a timetable. According to Goverde & Hansen (2013), performance indicators have four levels in the timetabling process: feasibility, stability, robustness, and resilience:

- *Timetable feasibility is the ability of all trains to adhere to their scheduled train paths with conflict-free condition.*

- *Timetable stability is the ability of a timetable to absorb initial and primary delays so that delayed trains return to their scheduled train paths. Hence, a timetable is stable if any train delay can be absorbed by the time allowances in the timetable without active dispatching (i.e., re-routeing, re-scheduling and/or cancellation of trains).*

- *Timetable robustness is the ability of a timetable to withstand design errors, parameter variations, and changing operational conditions.*

- *Timetable resilience is the flexibility of a timetable to prevent or reduce secondary delays using dispatching (and) can be viewed as the complement of robustness.*

More recently, Bešinović (2020) has detailed the inter-relationships between robustness and resilience. The whole resilience of the railway system consists of the robustness with which timetables are maintaining original performance levels when disruptions happen, the vulnerability which represents how much timetable performances drop given disruptions, the survivability which represents how long the timetable performance stays below the original level, the response which represents the period that the system starts to take actions and the timetable performance stops dropping, and the recovery which is when the timetable performance comes back to normal. The relationship among these concepts is drafted as Figure 2.4.

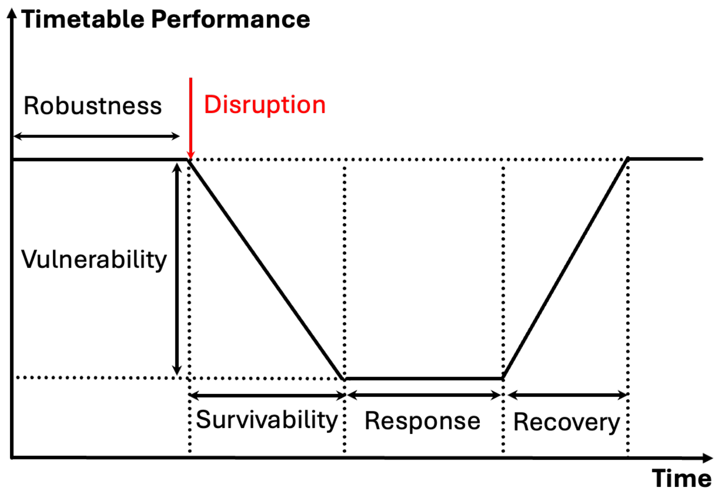


Figure 2.4 The resilience of a railway system

2.1.4 Discussion

The capacity balance factor graph shown as Figure 2.1 can effectively compare capacity utilisations of different railway systems from a general level, primarily caused by the network planning of railway systems. The dimensionality unit of measuring railway capacity is either the number of train paths or the timespan of running train paths, and they both belong to practical capacity category. There is a slight application scenario difference between adopting train paths and timespan. Train paths are usually represented as a possible 'provided' capacity from of a defined railway infrastructure, depending on the variety of train paths. The timespan, however, is more likely to be regarded as a capacity consumption of a proposed line plan in a defined railway infrastructure.

The provision of a 'good' capacity cannot only view total train paths or capacity occupation rate. To illustrate this with an example based on Figure 2.2, suppose an underground line has a minimum headway of two minutes, allowing for 30 trains to be dispatched in an hour. These 30 trains could all be non-stop services from the origin to the terminus station, or they could all be stop-by-stop services. Both line plans would make full use of the railway infrastructure with high occupation rates. However, it is evident that arranging a stop-by-stop line plan requires a much longer timespan and provides much greater service coverage compared to the non-stop line plan.

Therefore, it is necessary to introduce efficiency indicators of railway capacity, which are mainly developed from statistical usage data and evaluated with economic terms. Performance indicators of delivered capacity need to be analysed via a timetable. It can be concluded a timetable shall have sufficient tolerances against initial real time errors. Thus, the capacity utilisation level of a railway system cannot be 100% of the practical maximum.

2.2 The affecting factor of railway capacity.

From the perspective of infrastructure manager and railway operators, most studies (e.g., Landex, 2008; Abril et al., 2008; Khadem Sameni, 2012; Borndörfer & Lamorgese, 2018; Liao, 2021) commonly agree that the influencing factors can be divided into infrastructural factors, technical factors, and operational factors. Technical factors mainly contain track conditions which determines the speed limits, signalling system and block section length which are the vital components determining the headway among trains, station distributions across the network (network topology), track layouts (platforms and switch areas), and rolling stock features etc. Liao (2021) systematically summarised and testified the capacity increase measures regarding infrastructural factors, and the summary is given as Table 2.2.

Table 2.2 Key capacity influence factors at different levels of railway infrastructural detail

Detail levels	Key influence factor	Capacity increase measures	Capacity denotation
Macroscopic	Station positions (terminus station) Depot	Increasing station number in major cities Increasing depot number	Train path

Detail levels	Key influence factor	Capacity increase measures	Capacity denotation
Mesoscopic	Open track number Platform number	Increasing open tracks Isolating platform tracks	Headway Timespan
Microscopic	Track layout Track circuit	Conflict-free routes Route aggregation	Blocking time Occupation rate Utilisation Gannt chart

The amount of open track, platform tracks, and track layout are the vital factors that distinguish the railway infrastructure layout of different countries. It affects how overtaking behaviours and the headway among trains. Gao (2022) gives some examples of track layouts on a railway line (Figure 2.5) when operating a mix of fast (i.e. fewer stops) and slow trains (i.e. more stops) in the network.

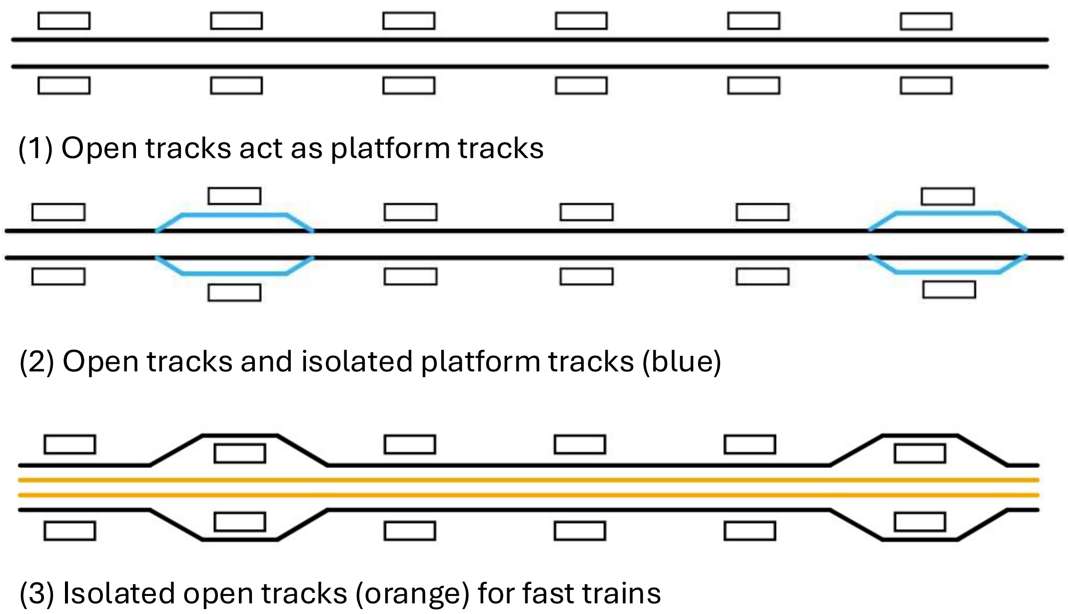


Figure 2.5 Different track layouts of a railway line

Operational factors, which will be explained in detail in Chapter 3, mainly contain network interdependency (Parbo et al., 2016; Zhang, 2019), traffic mixture (heterogeneity) of timetables, routing and platforming etc. It should be noticed that the heterogeneity of the timetable is also affect by the line plan whose major task is to decide the stop plan of trains.

2.3 Measuring railway capacity.

Most literature still follows a well-accepted but conventional standard based on Abril et al. (2008), and recently supplemented by Khadem Sameni & Moradi (2022), where capacity studies are simply divided into analysis (mathematic formular-driven method), simulation (supported by business software), and optimisation (models and algorithms). However, there is a clear trend that the methodology of capacity studies is becoming more integrated and comprehensive. For example, Mussone & Wolfler Calvo (2013) created a capacity optimisation model with constraints based on analytical formulas that calculate the junction capacity, line capacity, and interference between trains. Similarly, Zhou (2022) calculated unused time-space block areas by using the arrival and departure times of successive trains. Then, each used time-space block area is further abstracted as weighted vertices, and capacity measurement is therefore defined as a traveling salesman problem (i.e. optimisation). In short, the increasingly various research processes make the existing capacity classification method reluctantly applicable.

The nature of thesis indicates that the timetable is not only the transport product for passengers, but also a denotation of the railway capacity. Based on whether a research process requires a timetable, this thesis uses a new railway capacity study classification, and the following title is listed accordingly.

2.3.1 Measuring capacity without a timetable

2.3.1.1 Formula-based methods

1. Headway formula

The UIC in code 405 (Khadem Sameni, 2012; Liao, 2021) gives the following formular as the basic principle for measuring the railway capacity:

$$N = \frac{T}{t_h + t_b + t_e} \tag{2-1}$$

- N is the railway capacity, counted as competed train path number.
- T is the defined time period, which can be hour (s), a day, etc.
- t_h is the average minimum headway between trains.
- t_b is the necessary buffer time added to the headway to against initial delays.
- t_e is the extra time parameter that compensates for the diminishing returns in travel time reduction when the number of block sections increases.

The research focus of later studies based on headway formula is how to reasonably value t_h by jointly considering track condition, signalling systems, rolling stock features. Examples can see previous works (Abril et al., 2008; Kozan & Burdett, 2005; Mussone & Wolfler Calvo, 2013).

Considering the variety of train stops, leading to the heterogeneity, another branch of development based on headway formulae is the coefficient method (Lai et al., 2015; Hao Yang et al., 2011). The basic principle is first introducing a particular stopping pattern as the ‘standard train', regarded as the 'basic train equivalent.' Other trains with stopping patterns that differ from the standard train are categorised as 'other types of trains.' Due to the variations in stopping patterns, these trains may produce different headways at each station, resulting in a corresponding coefficient (2-4). Railway capacity (2-2) is then calculated as the total number of standard trains (2-3) minus the equivalent number of standard trains represented by the coefficients of the other types of trains.

$$N = n_s - \sum(\varepsilon_i - 1)\, n_i \tag{2-2}$$

$$n_s = \frac{T}{t_{h'} + t_b + t_e} \tag{2-3}$$

$$\varepsilon_i = \frac{h_a + h_d}{t_h} \tag{2-4}$$

- N is the railway capacity, counted as train number.
- n_s is the 'standard' train number in the timetable.
- n_i is other type of train number in the timetable.
- ε_i is the coefficient of train n_i.
- T is the defined time period, which can be hour (s), a day, etc.
- $t_{h'}$ is the average minimum headway by assuming standard trains are the only train types in the timetable.
- t_b and t_e are the same meaning with those in formula (2-1).
- h_a is the arrival headway time between the standard train and train n_i.
- h_d is the departure headway time between the standard train and train n_i.

The 'average' headway essentially indicates a fact that the train traffic among trains are usually heterogenous, which is caused by the variation in speed, stop patterns and headway times. Landex (2008) proposed the heterogeneity formula of train traffic as following, suggesting that if the calculation result approaches to 0, train traffic is homogenous, and vice versa.

$$Heterogeneity = 1 - \frac{\sum_{i=1}^{P}[RDeparture \times RArrival]}{P - 1} \tag{2-5}$$

$$RDeparture = min\{\frac{h_{d,i}}{h_{d,i+1}}; \frac{h_{d,i+1}}{h_{d,i}}\} \tag{2-6}$$

$$RArrival = min\{\frac{h_{a,i}}{h_{a,i+1}}; \frac{h_{a,i+1}}{h_{a,i}}\} \tag{2-7}$$

- P is the number of observed headways, indexed by i.
- i is the index of headway.
- $h_{d,i}$ is the ith departure headway time between a pair of successive trains.

- $h_{a,i}$ is the ith arrival headway time between a pair of successive trains.

2. Queuing theory

The queuing theory regards facilities inside a railway network as servers and trains as customers. The 'arrival- dwell or other operation tasks - departure' process is assumed as stomatic possibility. By setting the Kendall’s notation first, several performance indexes could be evaluated, such as queue length, queuing time, delays. For example, Weik et al. (2016) employed a single-server queueing system to calculate the average delay time, analysing the relationship between the maximum available train number under a time period (i.e., capacity) and the level of service (i.e., delays). Huisman et al. (2002) utilised train occupancy times and a semi-Markov queueing model to predict train waiting times, thereby estimating the maximum available train number under a time period (i.e. capacity) within a specified period. Yuan & Hansen (2007) analysed the relationship among train number, timetable reliability, and punctuality via a queuing model. Further interest on queuing theory application on railway can read the review given by Weik et al. (2016).

2.3.1.2 Graph-based method

Railway capacity can be measured by simulating the train path on time-space graph, which can be divided into simulation methods and compression methods.

1. Compression methods

There are two compression methods: UIC 406 compressions (UIC, 2013) and Capacity Utilisation Index (CUI) (Armstrong et al., 2011). The former takes the theoretical/technical headway as a compression unit, while CUI takes the ‘planning’ (i.e., theoretical plus supplement) time intervals at a railway station or a railway line as a compression unit. Despite the unit difference, the principles of these two methods are the same: moving every train path in the same direction without changing the train precedence and train trajectory until all train paths are closely-arranged with each other. More details can read Landex (2008), Armstrong et al. (2011), and Zhang (2015). Moreover, Max-plus algebra (Goverde (1998, 2007)) can be understood as a 'compression’ method: the route arrangement of trains that follows a pre-set order and blocking time can be denoted as a matrix form, thus the capacity can be measured.

2. Simulation methods

Simulation methods typically analyse proposed performance indexes of timetables before implementation, providing adjustment suggestions, and allowing railway capacity to be calculated and evaluated. Typical business software includes OpenTrack (Dicembre & Ricci, 2011) from Switzerland, RailSys (Lindfeldt, 2015) from Germany, Rail Traffic Control (Shih et al.,

2015) from the US, Capacity Management System from the UK, DNOS from Netherland, PETER from TU Delft, RAILCAP from Belgium, DEMIURGE from France (Khadem Sameni, 2012), and Treno (Trenolab, 2022) from Italy.

Besides business software, other techniques can simulate the railway network. For instance, Chen (2009) scheduled regular timetables on a high-speed railway via Petri Network. Iliasov et al. (2013) developed a signalling platform called SafeCap, which identifies bottlenecks in the railway network. Ma (2015) estimated the capacity influence of train combinations with different speeds. Vieira et al. (2018) developed a discrete events system to analyse railway efficiency, defined as the rate of total train paths generated by a heuristic algorithm and used capacity.

2.3.1.3 Discussion

Formula-based methods can give a quick and general capacity estimation of railway system and some possible system performances. However, both headway formula and queuing theory have significant limitations when applied to large and dense railway networks. The headway formula relies heavily on the accurate determination of parameter values, which requires extensive empirical data. Additionally, the coefficient method is limited to scenarios with a restricted variety of train types or affected by a dominant standard train type (Gao, 2022).

Regarding queuing theories, one of the fast evaluation applications is that analysing the 'queuing length' and 'delays' suggest the potential capacity bottlenecks and a need to expand the infrastructure. Assuming an average arrival headway is a minutes from n directions and the average dwell time is c minutes, the minimum platform number required is $p = \left\lceil \frac{nc}{a} \right\rceil$, where $\lceil \quad \rceil$,is the ceiling function symbol (Zhao et al., 2023). Accordingly, the switch area, which handles all departure operations of trains, can be simulated as a single server. Because all trains must merge into the open track, both the dwell time and platform number can affect the station inventory significantly.

Compared to formula-based methods, graph-based methods can conclude more precise railway capacity measurement results. In most cases, both compression and simulation methods are preceded by a timetable generation step, thereby forming as a research cycle of timetable generation, railway capacity measurement and assessment, and feedback. Examples can see Abril et al, (2008); Gao (2022) and Goverde et al. (2016). The compression method allows for the easy assessment of the occupation intensity, represented as capacity occupation rate, of different parts of the railway infrastructure. This helps identify network bottlenecks and determine which parts of the network have the potential to accommodate more trains. The accompanying drawback is that the compression methods exclusively emphasise on capacity occupation rate while offering limited explanatory potential regarding service quality of the timetable. For example, some trains show strong attractiveness and competitiveness due to their timeliness, giving them priority in the timetabling process. However, they may have a negative impact on the overall capacity occupation rate.

The simulation method provides the highest precision in capacity measurement by simulating train operations to verify the feasibility of timetable proposals. It also should be noticed that the development of advanced signalling systems (i.e., moving block signalling) (Landex & Wittrup, 2023) will possibly change the fundamental principles of railway capacity measurement in future, such as in forming microscopic train trajectories (Wang & Goverde, 2016), describing train relationships (e.g., virtual coupling) (Quaglietta et al., 2020), and trains can even behave like a road traffic congestion: queuing and then forming a 'shockwave' (Diaz de Rivera & Dick, 2021). With the growth of computing power and developments of artificial intelligence, simulation methods have a promising future in railway capacity study. However, due to the extensive initial scenario setup and parameter preparation required, the simulation method demonstrates its advantages more effectively in research motivations such as testing the influence of random disturbances (e.g., severe weather, equipment failures), microscopic operational details (e.g., switch operations), and nonlinear operations (e.g., train dynamics, passenger boarding and alighting), compared to the nature of this doctoral thesis.

2.3.2 Measuring capacity with a timetable

When a train operates within the network, it utilises the time-space resource (Liao, 2021) and is represented as capacity consumption (UIC, 2013): the signalling system divides a railway network into individual block sections, the total occupation time, which consists of signal preparation, reaction, approaching, actual occupation, and clearing time, is also called the blocking time. Comparatively, the headway is the time interval between a pair of successive trains occupying a block section. Capacity measurement establishes a mathematical link with the fundamental safety rules in timetabling: headway or blocking time.

Measuring capacity with a timetable in this section essentially means generating a feasible timetable with optimisation methods while integrating capacity constraints or objectives. The great extensibility of optimisation models can deeply combine railway operation rules, goals, and interests of different parties (i.e. railway operators, passengers) and railway planning stages, thus forming integrated problems. Representative reviews can see Ball et al., (2007); Bešinović et al., (2021); Cacchiani et al. (2015); Caimi et al. (2017); Harrod (2012); Niu (2021); Zhang (2020) , Li (2022) and Zhang (2019).

2.3.2.1 Principles of scheduling theory: job (flow)-shops models

Scheduling theory was first applied in railway planning by Szpigel (1973) because the modelling principle is similar to the occupation of block sections: a job (i.e. train) can only be processed (i.e. an operation, a movement) by one machine (i.e. station and block section) at a time, and each job follows a specific operation precedence to complete the train route. The scheduling theory in railway timetable scheduling is to find the job (i.e. trains) schedule which has the minimum complement time (i.e. makespan) or tardiness (Pinedo, 2007).

Depending on interpretation difference, scheduling theory on railway timetable scheduling can be modelled as either flow-shop scheduling problem or job-shop scheduling problem. All jobs (i.e. trains) at flow-shop scheduling follow the same occupation precedence of machine (i.e. stations and block sections) while each job at job-shop scheduling problem could have their own precedence. It appears that the flow-shop scheduling model could show more advantage on scheduling timetables for singular direction of a railway line, but the literature evidence shows a preference on job-shop scheduling, see the review given by Lange & Werner, (2018).

Depending on perspective difference, the solution process of the job-shop scheduling problem needs to convert time-dependent graphs into static forms such as alternative graphs (Corman et al., 2010) or disjunctive graphs (Burdett & Kozan, 2006). Alternative graphs show alternative job processing precedences at each machine, while disjunctive graphs represent alternative in machines (i.e. stations, or platforms of a station) of each operation. A natural concern rises when dealing with stations which contains multiple platforms, Lange & Werner (2018) therefore summaries three types of parallel-machine (units) problems (Figure 2.6):

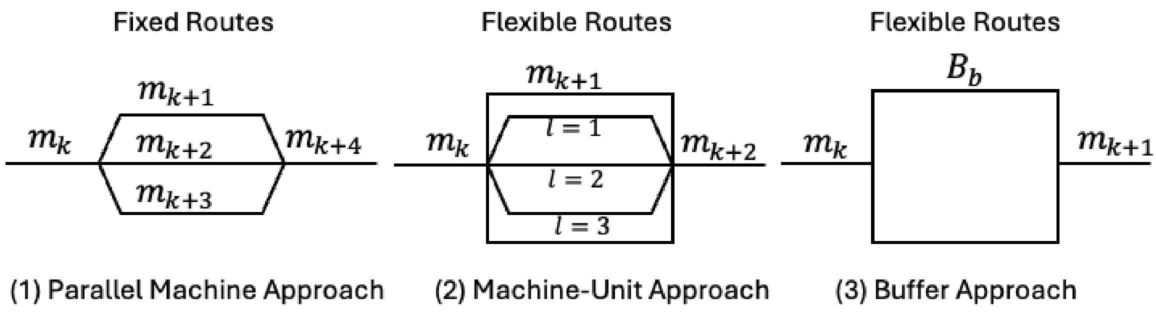


Figure 2.6 Different approaches of parallel machine (units)

- The parallel-machine approach (D’Ariano et al., 2007): the train routes are fixed, and the station layout is known.
- The machine-unit approach (Liu & Kozan, 2009): the train routes are flexible, and the station layout is known.

- The buffer approach (Burdett & Kozan, 2010): the train routes are flexible, and the station layout is unknown.

A machine-unit job-shop scheduling model example (Liu & Kozan, 2009) is shown below.

Table 2.3 Symbols and the explanation of a job-scheduling problem

Symbols	**Meaning**
n	numbers of jobs (i.e. trains) where the index is $i = 1,2,3,..i,..j...n$
m	numbers of machine (i.e. stations or block section) where the index is $k = 1,...,m$
h	numbers of units of a machine (i.e. platforms) where the index is $l = 1,...,h$
o	index of operation precedence of a job where the index is $o = 1,...,m$
s_{ilk}	Starting time of job i at $l\ th$ unit of machine m_k. $s_{ilk} \geq 0$.
p_{ilk}	Processing time of job i at $l\ th$ unit of machine m_k. $p_{ilk} \geq 0$
x_{ilk}	$= 1$ if job i is assigned $l\ th$ unit of machine m_k. $x_{ilk} = 0$, otherwise.
y_{ijlk}	$= 1$ if both job i and job j are assigned $l\ th$ unit of machine m_k, and job i is proceed before job j. $y_{ijlk} = 0$, otherwise.
r_{iolk}	$= 1$ if $o\ th$ operation of job i requires $l\ th$ unit of machine m_k. $r_{iolk} = 0$, otherwise.
C	The job competition time.
M	A very large positive number.

The model structure is formed as:

$$Minimise\ C$$

$$\sum_{l=1}^{h} \sum_{k=1}^{m} r_{iolk} (s_{ilk} + p_{ilk}) \leq \sum_{l=1}^{h} \sum_{k=1}^{m} r_{io+1lk}\, s_{ilk} \tag{2-8}$$

$$\sum_{l=1}^{h} \sum_{k=1}^{m} r_{iolk} (s_{ilk} + p_{ilk}) \leq C \tag{2-9}$$

$$s_{jlk} + p_{jlk} + M(y_{ijlk} - 1) \leq s_{ilk} \tag{2-10}$$

$$s_{ilk} + p_{ilk} + M(y_{jilk} - 1) \leq s_{jlk} \tag{2-11}$$

$$\sum_{l=1}^{l_k} \sum_{k=1}^{m} x_{ilk} = 1 \tag{2-12}$$

$$y_{ijlk} + y_{jilk} \leq 1 \tag{2-13}$$

$$x_{ilk} + x_{jlk} - 1 \leq y_{ijlk} + y_{jilk} \tag{2-14}$$

Constraint (2-8) restricts that the starting time of $(o+1)th$ operation of train i should not be earlier than the complement time of $o\ th$ operation of train i. Constraint (2-9) restricts that the complement time of each operation of each job at the machine m (i.e. the last) is no earlier than the makespan. $i+1$ cannot enter the block section unless train i leave the block section. Constraint (2-10) - Constraint (2-14) indicates the coupling relationships among decision variables, guarantying the there is only one operation of one job is processed at one unit of a machine.

2.3.2.2 Principles of space-time network modelling

The difference between Job-shop models and space-time network (STN) modelling is that STN sets a binary variable representing the occupation of discrete time element while time element in job-shop models is continuous (Lange & Werner, 2018). Brännlund et al., (1998) divided time period of STN and tracks into one-minute time slots and block sections, and the binary decision variable $x_{nbt}=1$ means train n occupies block section b at time instant t. Some basic concepts are list as following: the graph is $G=\ (V,A),\ v=(i,t)\in V$ represents the time instant of a vertex, and $a=(i,j,t,t')\in A$ represents a space-time arc. $i,j,i',j'$are the space index, and t,t',τ,τ' are the time index. The objective function is written as $min\ Z=\sum_n\sum_{a\in A}c_{n,a}x_{n,a}$, and $c_{n,a}$ represents as a cost.

The model is mainly structured via the multi-commodity flow and safety constraints. The multi-commodity flow is written as $\sum_{a\in A^+}x_{n,a}=\sum_{a\in A^-}x_{n,a}$ where A^+ represents entering direction, and A^-represents leaving direction. Regarding safety constraints, there are three methods which are incompatible arc set (Caprara et al., 2002, 2006; Zhang et al., 2020; Zhang et al., 2019), cumulative flow variables (Meng & Zhou, 2014, 2019), and recourse variables (Wang et al., 2023):

- The incompatible arc set $\Psi(a)$ collect all possible arcs representing the overlapping (i.e. incompatible and conflicted) between a pair of successive trains regarding their arrivals, departures, and running between stations. The corresponding constraints are written analogously as $\sum_n\sum_{a\in\Psi(a)}x_{n,a}\le 1$. A similar modelling principle is hypergraph (Harrod, 2012) which introduce a second binary decision variable to represent this afflicted occupation of block sections spatio-temporally.
- The cumulative flow variable essentially establishes the logic link between the actual time and decision binary variables. The actual arrival time and departure time of train n is defined as $\overline{a_{n,i,j}}$ and $\overline{d_{n,i,j}}$ respectively, and supportive binary variable $a_{n,i,j,t}$ and $d_{n,i,j,t}$ represent the train n arrive or depart from i to j respectively. The actual arrival and departure time are calculated as $\overline{a_{n,i,j}}=\sum_{t\in T}t\times(a_{n,i,j,t}-a_{n,i,j,t-1})$ and $\overline{d_{n,i,j}}=\sum_{t\in T}t\times(d_{n,i,j,t}-d_{n,i,j,t-1})$, so that the headway can be written with big-M constraints.

- The resource variable essentially defines that occupation of a block section is a period $h_{n,i,j,\pi} = 1; \pi \in [t, t + headway]$, simulating the headway constraints as $\sum_n \sum_\pi h \leq 1$. By setting $h_{n,i,j,\pi} = x_{n,a}$, the decision variable is therefore coupling with resource variable.

STN demonstrates excellent extensibility in timetabling modelling. It is effective for both macro and micro timetable modelling and is applicable to both periodic and non-periodic timetables. For instance, Zhang et al. (2020) researched routing and platforming problems based on the macro-level timetable based on the incompatible arc sets from (Caprara et al., 2002, 2006). Wang et al. (2023) employed a resource variable modelling system and a branch-and-price strategy to study the issue of overlapping conflicting routes in routing and platforming problem. Meng & Zhou (2014, 2019) used cumulative flow variables to investigate micro-level robust timetable modelling. Zhang et al. (2019) adopted a replication variable strategy to replicate a single-period timetable into a periodic timetable. Li (2022) set the concept of actual time in cumulative flow variables as a cost, making the STN objective function representative of the capacity consumed by the timetable. Liao et al. (2021) combined decision variable with the occupancy of micro-block sections to design a hybrid STN modelling method for capacity evaluation. Yao et al. (2023) embedded passenger routing problems into STN-based timetable modelling, enabling the assessment of timetable service quality. For further detailed reviews regarding STN, refer to Wang et al. (2022)'s work.

2.3.2.3 Principles of event-activity network modelling

Given a set of events (vertexes): $i, j \in N$, activity (arc): $[i, j] \in A$ (i.e. representing from event i to event j), and the time instants of events v_j, v_i, and a time window (l_{ij}, u_{ij}) , the model of event-activity network (EAN) is to find a solution which satisfies all $l_{ij} \leq v_j - v_i \leq u_{ij}$. Any activity of trains, such as running, dwell, headway, can all be written analogously.

Some major extensions of $l_{ij} \leq v_j - v_i \leq u_{ij}$ are as follows. If the timetable is operated in a cyclic nature, the model introduces a cycle time T and a module binary variable p_{ij} which equals to 1 only if $v_j \leq v_i$, and the aim is to find solution which satisfies all $l_{ij} \leq v_j - v_i + Tp_{ij} \leq u_{ij}$, named as periodic event scheduling problem (PESP). Define $x_a = v_j - v_i + Tp_a$ and C^+, C^- as two directions of an undirected graph G, the PESP can be re-written as cycle periodicity formulation (CPF). Supporting proofs and detailed formulations can be found in Odijk (1996) and Peeters (2003). Furthermore, if set v_i has the flexibility of a time window$(\underline{v_i}, \overline{v_i})$, and $v_j - v_i + Tp_{ij} \leq u_{ij}$ is changed into $l_{ij} + \left(\overline{v_i} - \underline{v_i}\right) \leq \underline{v_j} - \underline{v_i} + Tp_{ij} \leq u_{ij} - (\overline{v_j} - \underline{v_j})$. Caimi et al. (2011) regarded this as a method to deal with microscopic timetable infeasibility after achieving a feasible macroscopic timetable from PESP. Further interests on the extension and integration of EAN can read a review given by Liebchen & Möhring (2008) and Zhang (2020).

The evidence indicates that EAN has conducted extensive research on capacity measurement. Sparing & Goverde (2013) established a capacity assessment model with the objective function of minimizing the cycle time. In their subsequent work (Sparing & Goverde, 2017), they analysed the relationship between flexible dwell times, running times, and the stability of the timetable. Similarly, Heydar et al. (2013) and Petering et al. (2016) developed a capacity assessment model for a single-track railway in one direction, also with the objective function of minimizing cycle time. This model operates on a micro-scale (i.e., including platform selection), but did not allow overtaking behaviours among trains. Zhang & Nie (2016) then designed constraints for modelling overtaking behaviour, analysing the impact of factors such as speed differentials and stopping patterns on capacity. In their subsequent studies (Zhang et al., 2019; Zhang & Nie, 2017), they attempted to introduce the concept of time window to analyse the impact of cross-line trains on capacity. Chen (2017) focused on minimising the latest arrival time to enhance the bottleneck segment throughput capacity of a railway line. Yan & Goverde (2019), based on their own previous work (Yan & Goverde, 2017), designed four optimization objectives—minimising travel time, timetable robustness, among others—to explore the trade-offs between these objectives. Li et al. (2023) introduce an extended EAN to access the capacity of railway network where the model contains platform choices and the capacity impact of cross-line trains. The results are very similar to the work by Li (2020) that if cross-line trains account for over 40%, the overall capacity of the network will decrease. Apart from capacity assessment, Lamorgese et al. (2017) designed a feedback mechanism between macro and micro timetables based on the EAN to assess whether infrastructure expansion is needed to achieve a capacity increase.

2.3.2.4 Considering robustness in the timetable

As a common feature, robustness refers to how a system reacts in the presence of arbitrary errors. Depending on data availability, how to convey errors into model language mainly divides studies into recoverable robustness optimisation (Anderegg et al., 2009) and stochastic programming (Khan & Zhou, 2010). Recoverable robustness optimisation inevitably generates overly conversative solutions because every uncertainty needs to be considered, leading to a low-capacity occupation level. Stochastic programming usually gives specific solutions after errors and initial plans become known. However, stochastic errors programming still proved to be very challenging considering the complex model structure and the large problem scale at the timetable generation stage (Bešinović et al., 2021; Goverde & Hansen, 2013).

Besides, robustness perspectives contribute to various measurements and definitions of robustness: buffer time (Landex, 2008), slack time of critical paths (Goverde, 2007), realised passenger travel time (Dewilde & Sels, 2011), and metric-based assessment (Andersson et al., 2013). Lusby et al. (2018) give an exhaustive summary of robustness at each railway planning stage, similar terminologies and methodologies.

2.3.2.5 Discussion

1. Mathematical features of modelling methods

As general, most timetabling model principles belong to mixed integer programming where the train operation rules and relationships among trains are either written as constraints or linked by binary variables. Although the amount of literature on timetable scheduling using job-shop models is significantly less than that on space-time networks (STN) and event-activity networks (EAN), the modelling principles, such as big-M and precedence constraints, have profoundly influenced the development of timetabling models. These constraints effectively describe the sequencing of trains and numerous if-then scenarios in train operations.

From the summary of safety constraints in Section 2.3.2.2, it is clear that blocking time and headway are fundamentally equivalent (Liao, 2021; Meng & Zhou, 2019), differing only in their perspective on capacity (Figure 2.7). Blocking time describes the occupation of time-space resources, whereas headway refers to the time interval between successive events, such as trains' behaviours: arrival, dwell, and departure.

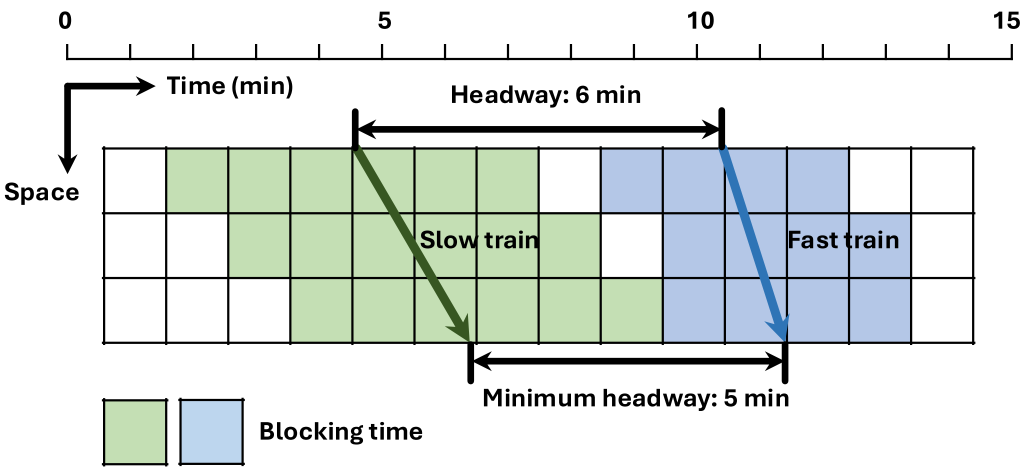


Figure 2.7 Blocking time and headway denotations on a time-space graph

In terms of representing railway capacity, both STN and EAN demonstrate distinct advantages at the application level. If train numbers are used as the capacity form within a defined railway infrastructure, the STN, where decision variables primarily represent the formation of an integrated train path, offers greater structural benefits. In contrast, when assessing the capacity required by a proposed line plan, the EAN where the main model structure represent the time instants and cost could easily represent the timespan as a 'capacity consumption'.

2. The debate of model detail levels

Depending on data availability or research motivation required, it is common to classify timetables into macroscopic and microscopic levels. A macroscopic level timetable regards railway stations as 'grey boxes' while a microscopic level timetable usually regards every block section as a modelling unit. Microscopic level timetabling allows regard every track circuit as a modelling unit, therefore the routing conflicts can be considered in detail. However, the drawback of microscopic timetabling is requiring significantly larger data input compared to macroscopic level (Bešinović et al., 2021). A recent trend has started to consider railway infrastructure at the mesoscopic level, see Goverde et al. (2016) and Zhang et al. (2020, 2022) as examples. Mesoscopic level modelling simplifies the routing conflicts at the switch areas. One strong supporting reason is that the routing at the switch area can be 'partially packed' as an aggregation (Corman et al., 2009; Lu et al., 2022; Schlechte et al., 2011), and an empirical headway verification has been tested by Li (2019). Figure 2.8 shows detail levels of railway infrastructure and routing aggregations. Nevertheless, routing and platforming is a natural integrated step with timetable generation and determining the route aggregation level would be key to achieving the balance of controlling model complexity and considering more operational rules. Further interests on platforming and routing can see the review by Schlechte (2012).

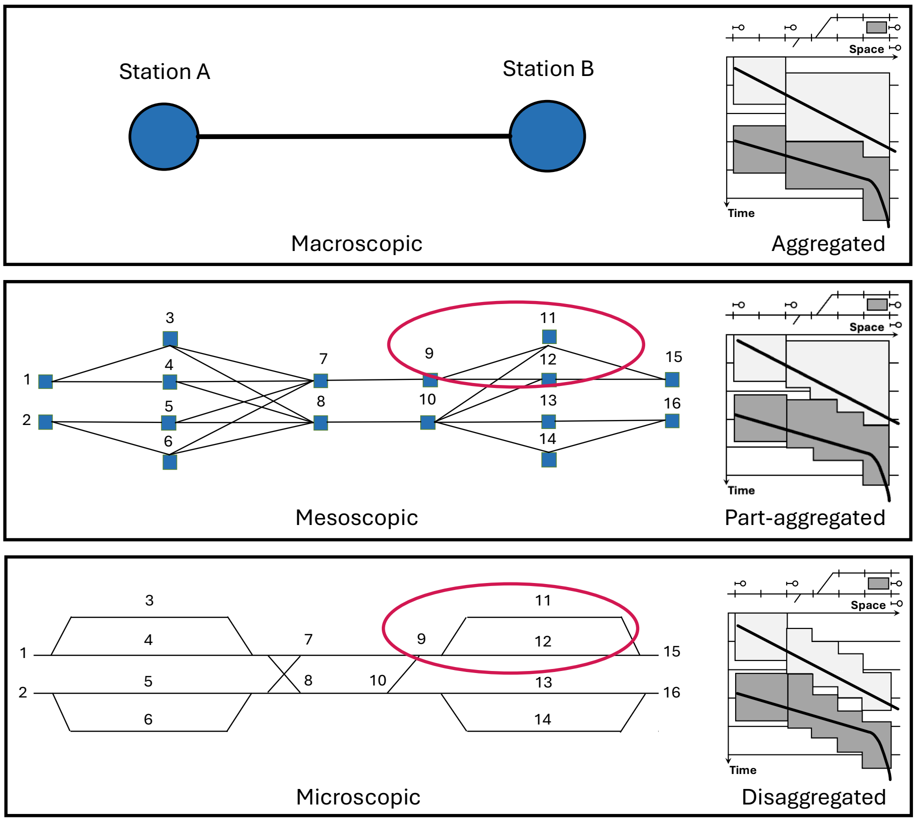


Figure 2.8 Levels of railway infrastructures and route aggregation

3. The conflict between integration trend and decomposition algorithm development

The integrated problems contain several isolated passenger railway planning stages as a whole, which is increasingly popular in the academic field (Lestari et al., 2024). Most claimed that integrated problems own the advantage of avoiding the repetitive loop between feedback and adjustments (Shown in Figure 1.1) with higher solution efficiency. However, by comparing at their solution processes, decomposition techniques are their dominating methods (Leutwiler & Corman, 2023). Decomposition techniques divide the problem into either parallel smaller problem, where the search for the satisfying gaps between upper and lower bounds, or into a master problem and sub-problems. In the latter approach, each sub-problem begins its solution process based on a base solution from the master problem, followed by the iterations and stopped by reaching predetermined gaps. From the algorithm perspective, the feedback and adjustment process in the integrated problem merely transfers human intervention of passenger railway planning in stages to the programming. Additionally, to maintain solution efficiency, the integrated problem model either oversimplifies model constraints, operation rules, or railway scenarios. This simplification leads to weaker interpretability of the integrated problem concerning railway operation reality. Based on this observation, it is questionable whether defining the service scheduling problem in this thesis as an integrated problem is a more advantageous choice.

4. The uniqueness for simulation-based method

This thesis would agree that simulation methods mentioned in section 2.3.1 can exert their full potential in researching resilience and robustness in the timetabling, because the basic modelling unit from simulation represents a probability of change. Therefore, they can simulate even more severe situations such as disruptions or closures. Simulation methods follow the general features of discrete event dynamic systems (Vieira et al., 2018): defining transitions or status updating (changing) rules.

Additionally, focusing on capacity of the network does not mean ignoring initial robustness or stability of the timetable. The determination process of the minimum headway applied in a railway network shall consider certain levels of buffer time (Ning et al., 2009; UIC, 2013).

2.4 Evaluating passenger experience.

Passenger experience, sometimes referred to in other literature as passenger-oriented (Hu et al., 2024) or passenger-centric (Parbo et al., 2016) approaches, is a significant academic branch within timetabling problem. However, research within this field can present diverse definitions, interpretations, methodologies, and research pathways depending on the attributes of the railway project. The recent work by Hartleb et al. (2019) has highlighted that any timetabling concerning passenger experience should consider the following aspects: what elements constitute passenger experience (section 2.4.1)? How is it expressed (section 2.4.2)? How are passengers' origin-destination (OD) routing handled (section 2.4.3)? Additionally, how to reasonably hypothesise about passenger behaviours and intentions (2.4.4)? This section accordingly uses these questions as the following section headings.

2.4.1 Passenger experience indicator (PEI) list

Passenger experience covers not only the In-vehicle stage, but also before and after boarding stages (OECD/ITF, 2014). Academic and empirical evidence across European, Asian, and North American countries have commonly summarised the passenger experience into indicators within the following aspects (OECD/ITF, 2014; Transportation Research Board of the National Academies, 2013): availability, travel time, reliability, accessibility, care, comfort, safety, and information. Among these aspects, availability, travel time, and reliability are closely related to railway capacity because they essentially exert requirements for timetabling process. The remaining five aspects help improve passengers' well-being' during their travels. Those indicators are named as railway capacity-related passenger experience indicators (PEI) in Table 2.4.

Table 2.4 Railway capacity-related passenger experience indicators

Aspect	Passenger experience indicator (PEI)
Availability	(1) Coverage
	(2) Direct travel
	(3) Transfer
	(4) Service hours
	(5) Daily/weekly frequency
	(6) Hourly frequency
Travel time	(7) In-vehicle time
	(8) Transfer time
	(9) Schedule delays/adoption time
Reliability	(10) Punctuality
	(11) Cancellation

Availability refers to train services in terms of geography, service time range, and frequency. It is common that not all travels can be served by a direct train, and not all railway operations follow a cyclic nature (i.e., the same train service repeats with a periodicity); travel mode is divided into direct travel and transfer, and the frequency is divided into daily (or weekly) and hourly frequency accordingly. Daily (or weekly) frequency means there is at least one desired train in a day or in a week but does not necessarily follow a cyclic nature. Comparatively, hourly frequency usually refers to periodic timetables (e.g., timetables in the UK, the Netherlands, and Germany). Hourly frequency can increase in peak hours and decrease in off-peak hours to meet demand changes.

The travel time has In-vehicle and out-of-vehicle parts in general. Only the schedule delays are not wholly a travel time cost PEI. Schedule delays are the difference, either being earlier or later, between passenger desired arrival/departure time and actual times given by the timetable (van Wee et al., 2008). Desired arrival/departure time can vary from person to person, evidence (Lestari et al., 2024; Wheat & Wardman, 2017) shows that increasing the frequency of the train services can decrease the schedule delays statistically and negative mental feelings of passengers and extending the operating hours can satisfy more special travel needs. (e.g., the potential to catch the last metro after a late-night arrival) (Hu et al, 2023).

The reliability of a railway system in passengers' view is essentially its punctuality (Parbo et al., 2016) and the extent to which trains run as planned. The concern of cancellations (ORR, 2023) is also valued by passengers as well.

2.4.2 Passenger experience evaluation function

There are mainly two types of passenger evaluation functions: travel time-based and utility-based function. For travel time, it further divided into absolute travel time and perceived travel time. The perceived travel time is also known as generalised journey time (GJT) in the UK (Railway Delivery Group, 2018). Absolute travel time is the actual time cost (i.e. station-to-station time) for every passengers, jointly optimising the absolute travel time with either a timetable or a line plan is one of the objectives in research, see examples like Dewilde & Sels, (2011); Wang et al., (2015); Yan et al. (2019); Zhu (2019) and Hu et al. (2023). Perceived travel time is a weighted travel time which covers the whole travel process (Ortúzar & Willumsen, 2011). In urban railway scenarios where trains are usually operated with a cyclic nature, how to quantitatively express transfer 'cost' is a key focus (Cheng, 2018). In intercity railway scenarios, the UK adopts the GJT consisting of the following components (Railway Delivery Group, 2018): $GJT = T + H + I$, where T is the station-to-station travel time, including any transfer time, H is service interval penalty, and I is the sum of all interchange penalties. In a series of works by Wardman (2012, 2022) and Wheat & Wardman (2017), the GJT is rewritten as $GJT = T + \alpha H + \beta I$, where α, β are newly added coefficients. The motivation behind this was to challenge whether the current formulation of GJT used the appropriate weighting and adopted the historical ticket data to determine the elasticity (i.e. how easy the demand can be affected). The findings indicate that greater weight should be assigned to interchange. In short, the investigation result shows very limited evidence integrating perceived travel time at the timetabling stage. Comparatively, introducing perceived travel time in delay management is more common, such as Goverde (2005); Högdahl & Bohlin (2023) and Isaai & Singh, (2001), because perceived travel time can represent the priority among passengers (groups): who go first.

On the other hand, utility-based functions are commonly shown on passenger routing (i.e. OD assignment or distribution) or choice modelling (read section 2.4.3), rather than in timetabling issues (Hartleb et al., 2019). Cheng (2018) further summarises that studying utility-based function essentially is the quantitative research of influence factors or key stages of passenger travel process, such as waiting time (Gong et al., 2021; Niu & Zhou, 2013), adoption time (i.e. weighted waiting time) proposed by Polinder et al. (2020), and transfer time (Li, 2018; Wong et al., 2008). Typical function structure types are additive functions (Li, 2020), multiplicative functions (Liu, 2012), natural exponential functions (OECD/ITF, 2014). The investigation of this thesis shows that (Mei et al., 2021) is one of rare studies who integrates passenger utility and timetabling where the passenger utility is written as continuous approximation form. Further interest regarding utility-based function can see the work by Ortúzar & Willumsen (2011).

2.4.3 Passenger routing or distribution

Choice modelling is assigning a utility to each route, and the percentage or the possibility of passengers selecting a certain route depending on the utility of the route or alternative routes. Multiple available routes generate the following step called passenger OD distribution (i.e. assignment), which is seen more frequent on road traffic. Comparatively, singular route is much more common for railway travels, especially for long-distance railway travellers, and predetermining the passenger route choice is therefore an acceptable assumption. Following those natures, representative examples of the timetabling problem considering passenger routing or distributions can classified into Table 2.5. Further interests regarding passenger routing could see the review part by Hartleb & Schmidt (2022) and Yao et al. (2023).

Regarding choice modelling structure, logit-based models are dominant players in this field, see evidence proved by Cheng (2018) and Li (2020). Because of the non-linear nature of logit-based models, the effectiveness of logit-based model linearisation is proved by Haase & Müller (2014) and Hartleb & Schmidt (2022), and the following works (Hartleb et al., 2023; Hu et al., 2024) can integrate choice modelling better with their overall optimisation models in timetabling.

Alternatively, the choice modelling can be denoted as a linear additive form consisting of deterministic parts and an error term. Deterministic part represents the 'known and concluded' utility, such as station-to-station travel time, and error terms represent the 'unknown and hidden part', such as passenger preferences or statistical errors (Train, 2002).

Table 2.5 Research pathways of passenger routing and distribution in timetabling

	Predetermined route	**Integrated choice modelling**
Singular route	Pätzold & Schöbel (2016) Liebchen (2018) **This thesis**	Yao et al. (2023) Gattermann et al. (2016) Goerigk & Schmidt (2017) Löbel et al. (2020)
Distribution	Robenek et al. (2016) Yao et al. (2024) Sels et al. (2015)	Hartleb & Schmidt (2022) Hu et al. (2024)

2.4.4 Passenger preferences

Passenger investigation data is the input of choice models, provides vital reference for reasonable assumptions, and verifies service quality of timetables. There are two types of investigations: stated preference (SP) and revealed preference (RP). SP investigation is presented with hypothetical scenarios and collects passenger preferences, representing passenger expectations on service quality. RP investigation presents the observed and collected behaviour, using data from ticketing and any other historical record. The insights of RP suggest the passenger decision preferences and real-world behaviour patterns, indicating the compromise between current railway supply and passenger expectations (Lee, 2014). RP conclusions not only reflect the gap between current service levels and passenger expectation, but also validate the model results based on SP data.

Structures of SP or RP questionnaires are highly flexible depending on the investigation goals (Guo, 2015; Li, 2022; Li, 2015; Li, 2020; Wardman, 2022; Wheat & Wardman, 2017; Zhang et al., 2016). Despite the differences in railway contexts, some characteristics of long-distance travellers are commonly shared:

- mostly planned travels,
- less elasticity (i.e. stable demands),
- showing stronger concern regarding transfer and interchange, and
- valuing more on arrival time rather than departure time.

Regarding transport modes, China HSR has a strong competition with civil aviation (Wang et al., 2020; Yang et al., 2018), and evidence shows that flights only show its advantage when the travel time is later than 19.00 or earlier than 10.00, and for passengers from/to Northwest China.

2.4.5 Discussion

It could be concluded that the line planning decides critical contents of a service plan, thus delivering most aspects in PEIs. Time-based and utility-based functions are two typical passenger evaluation functions reflecting passenger experience. Utility functions further require choice modelling and input from SP or RP investigations, or any other empirical data, which can evaluate the full passenger travel process. If trying to integrate utility functions into optimisation models, a core challenge is the linearisation of the utility function. In terms of valuation of the coefficients, perceived travel time, or GJT in UK, requires empirical data as a basis for the component values.

Considering passenger experience in timetabling stages can help deliver more precise and satisfactory services to different passengers (or passenger groups), specially based on key insights from SP and RP investigation results. However, Wheat & Wardman (2017) challenges the validity of SP investigation: some questions may mislead passengers due to their lack of professional knowledge. For instance, a transfer time of approximately 15 minutes was found to be preferred by most passengers in China HSR during a SP investigation (Guo, 2015; Li, 2020; Wheat & Wardman, 2017), however, this preference is often unrealistic, as passengers frequently overlook many factors, such as the anxiety caused by crowdedness, unfamiliarity, facilities (Wheat & Wardman, 2017), in the transfer process that prolong their actual transfer time.

There is a clear trend that research related to utility function is mostly applied to urban railways, regional railways or other periodic services with short travel distances (Lestari et al., 2024). These railway scenarios also typically do not require seat reservation, and the service patterns (i.e. the stop plans of services) are simple. Consequently, passengers can alter their travel behaviours, train service at other time, train service with different routes, or other transport modes at a relatively low cost. Demand characteristics are therefore highly time-dependent (Gong et al., 2021; Meng & Zhou, 2019; Niu & Zhou, 2013). This observation aligns with Wheat & Wardman, (2017), who noted that short-distance travellers might exhibit higher price and service elasticity, indicating that demands of short-distance travellers are more easily influenced by the varying price regarding during/off-peak trains, crowdedness levels and availability of trains. Therefore, utility functions can investigate the key influencing factors and are used for enhancing the attractiveness of railway services, especially for travellers who are not commuters.

Comparatively, it is worth debating whether to adopt the utility function in railway contexts dominated by long-distance passengers. Firstly, from the perspective of choice modelling, alternative services that occupy different routes for long-distance passengers are in fact very limited. For those OD pairs with less demand, provided services only have one route with lower frequency, passengers have to either make compromise to those service or change into other transport modes (Hu et al., 2024). For those popular OD pairs, they are often served by high-frequency fast trains on a singular route. Additionally, Lu (2021) has testified the alternative routes of all ODs within China HSR network and concluded that 'impedance' of alternative routes is very high in such a large network scale, represented by either no available HSR services or excessively additional travel distance. This indicates that the alternative routes, even provided with HSR services, are very unattractiveness. Secondly, from a perspective of service time difference, services with high 'utility' values due to time-related preferences, thus more likely attracting more passengers, will be eliminated by dynamic ticketing in subsequent stages to adjust actual passenger demand (Yuan et al., 2018). Thirdly, from the perspective of utility function form, the utility functions usually contain factors that the timetabling cannot cover, such as comfort, safety, accessibility, and information, which is far less relevant with railway capacity. In fact, Wheat & Wardman (2017) verified that the GJT function can better fit ticket sales data to reflect passenger preferences compared to the utility functions, and they argued that travel time represents the most vital aspect of railway travels.

2.5 Summary and gaps

This literature review systematically explains the concept framework and measurement methods of railway capacity, briefly introduces the influencing factors of railway capacity, and summarises methods of evaluating passenger experience. Each section mainly sorts the principles of methods, highlights the difference among research pathways and natures, analyse the research motivations behind, and ends up with a corresponding discussion. At the same time, supplementary review recommendations are listed for further interests.

Regarding a conceptual framework of railway capacity, an extended railway capacity concept framework is built based on the capacity balance factors graph from UIC406, concluding as quantity and quality aspects and covering the whole railway planning process. Regarding influencing factors of railway capacity, the key infrastructure factor is the track layout including open track number and isolated platform tracks. Regarding railway capacity measurement, the development history of optimisation models shows that the majority of models involves mixed integer programming, and the mode complexity depends on detail levels of simulating the infrastructures, describing the operation rules, and integration levels of models. Besides, it can be expected that software-based simulation methods will have broader integration with operational research due to advances in chip technology and artificial intelligence. Another conclusion is that timetabling models with capacity constraints or objectives fit better with the research nature of this thesis. Regarding evaluating passenger experience, availability, travel time and reliability are first concluded as capacity-related PEI aspects, which can be addressed by timetabling goals. Although more precise passenger behaviours reflected by choice models, applying utility-based evaluation function in long-distance railway travels has been questioned by strong reasoning, and time-based evaluation function is therefore more applicable.

By comparing the insights derived from this literature review and the latest statistical review given by Lestari et al. (2024) who clustered 333 effective timetabling papers from 944 samples, this thesis confidently identifies the following academic gaps:

- Very few studies have explored the reasons behind the controversial and diverse definitions and interpretations of railway capacity. As explained and answered in Section 1 and 2.1, this can be attributed to specific focused railway planning stages, railway contexts influenced by distinctive styles of network planning, varying transport needs and demand features (i.e. passengers in urban railways, regional railways, or long-distance railways), and the variety of research motivations.
- Although multiple objectives have been mentioned in some papers, few studies have been able to combine passenger experience indicators (PEI) and capacity in a deep and comprehensive way. As a result, the trend of capacity-related papers predominantly focuses on the increased capacity rather than improved capacity. There is a significant need to develop railway capacity methodologies that better meet passenger demand features and make more effective utilisation of infrastructure.
- From a perspective of addressing the challenges proposed in section 1.2.3, existing academic results lack convincing theories or methods, primarily due to the adaptability of existing methods to different application scenarios. For examples, alternative routes and frequent transfers are commonly accepted by urban railway passengers but not by long-distance passengers. Arranging 'stop-by-stop' trains does not increase travel time significantly for short-distance travel. In other words, it is necessary to divide cluster ODs in a reasonable manner, rather than uniformly setting all ODs to be completely equal in terms of meeting PEIs. Managing network-level services require more coordination and constraints within network parts. There is a need to divide country-based macroscopic operation features (i.e. services in Japan, Germany, Netherland, the UK, China) into operation factors (Chapter 3) to conclude effective solution principles.

In summary, this literature review establishes extensive academic linkages with existing research and serves as a guidebook connecting related topics beyond just railway capacity. It provides a reference for the research pathway in the subsequent chapters and highlights the key points of modelling and assumption aspects.

Chapter 3 Capacity-related service scheduling process

The introductory chapters have concluded that line planning, timetable generation, and platforming are crucial stages in passenger railway planning that affect capacity utilisation of a railway network. Furthermore, improving capacity utilisation necessitates a comprehensive consideration of the service quality across various OD pairs, which has been categorised in the Passenger Experience Indicators (PEI) table. This chapter aims to develop a capacity-related service scheduling process based on a detailed analysis of the passenger and operational factors influencing railway capacity.

3.1 Passenger factors

3.1.1 Understanding passenger decision and clustered groups

The ticket purchasing process for an individual passenger is influenced by personal attributes such as occupation, gender, age, financial affordability. These factors contribute to the passenger's preferences on certain service qualities. After comparing against train booking information and dynamic prices, passenger will make the decision that suits them best (Yuan, 2020). In other words, the decision to purchase the same train ticket can stem from entirely different passenger attributes, and for the same OD travel demand, different purchase outcomes may arise, such as variations in In-vehicle time, departure or arrival times, and seat class.

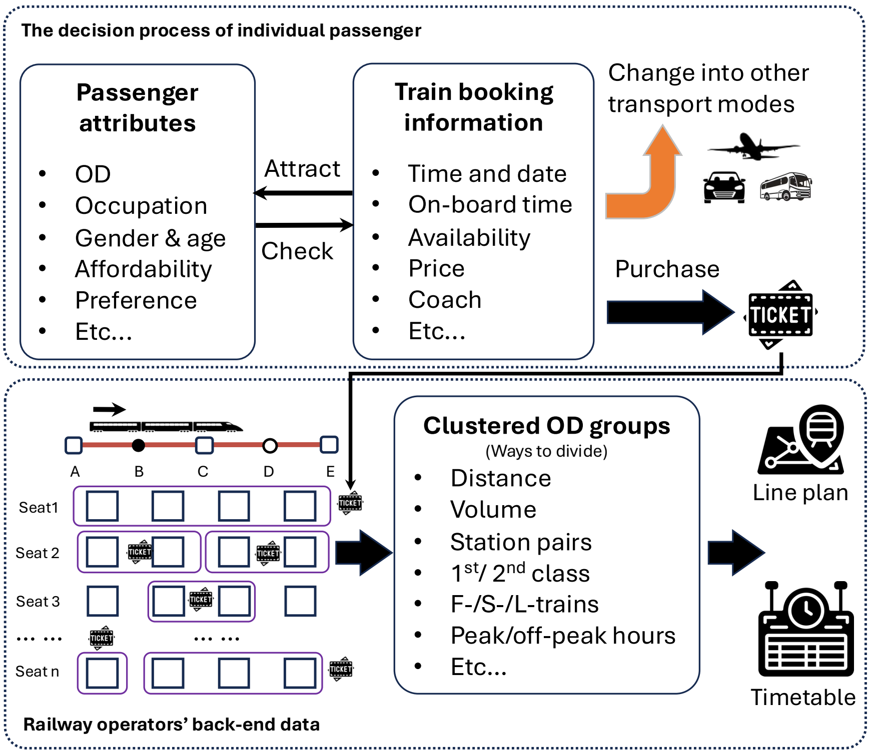


Figure 3.1 Formation of clustered OD groups

Correspondingly, from the railway operator's perspective (Figure 3.1), these individual ticket decisions accumulate as OD flows and reveal several objective attributes, such as the OD, date, seat class, and train code (i.e. indicating the travel time), distance, demand volume. Railway operators cluster into specific OD groups based on these objective attributes. Thus, when interpreting the demand characteristics of ODs, various dimensions emerge, such as long-distance versus short-distance ODs, ODs between different city tiers (i.e. station pairs), and distinctions between 1st and 2nd class passengers, which links to various travel purposes such as business, commuting, tourism etc. All ODs together constitute the diverse railway OD groups or OD markets (Sun et al., 2018). In other words, one OD encompasses multiple demand characteristics. For example, the OD pair from Beijing Nan (South) to Shanghai Hongqiao in the China HSR network is characterised as long-distance travel, high demand volume, the B-B (from Big station to Big station, the rest states similarly) station pair, and very profitable OD group (Hu et al., 2023, 2024; Li, 2020).

Among the various objective attributes, the station pairs derived from city tiers, along with travel distance demonstrate a stronger correlation with HSR demand volume. Table 3.1 and 3.2 show the OD percentage of a long-distance train and between station pairs (Tian, 2018). Train G1204 operates from Shanghai to Harbin, and the total operating distance is 2,422km.

Table 3.1 Passenger loading distribution during the itinerary of train G1204

Travel distance	**Passenger**	**Passenger percentage**	**person·koilmeter**	**person·koilmeter percentage**
0-200km	538	21.21%	51,798	3.06%
200-500km	650	25.36%	201,429	11.88%
500-1000km	639	25.20%	472,508	27.88%
1000-1500km	523	20.62%	627,505	37.02%
1500km-2000km	130	5.13%	219,648	12.96%
>2000km	56	2.21%	122,260	7.21%
Total	2,536	100.00%	1,695,148	100.00%

It can be concluded that ODs with travel distances between 200 to 1500 km have a higher percentage of passenger loading, and ODs with travel distances between 500 to 1500 km contribute the majority of turnover, representing profitable OD groups where ticket fares are calculated based on mileage. Passengers whose travel distance is over 2,000km contribute the least share of the train load. On the other hand, ODs associated with Big stations contribute significantly to the HSR profitability. Therefore, these two objective attributes are considered crucial in shaping future timetable structures and in determining the provision of railway capacity consumption. Firstly, station pairs represent the transportation generation and attraction driven by economic, population, administrative, and industrial factors between cities, which have been widely verified (Li, 2022; Li, 2015; Li, 2020; Tian, 2018). Secondly, the advantage of HSR travel distance also reflects the favour and tolerance of passengers for the length of travel time on HSR. It has already demonstrated that ultra-long-distance OD pairs (over 1500 km) have less advantage in travel time (Wang et al., 2020) compared to civil aviation and are less economical (affordable) than overnight sleeper trains (Sun et al., 2018).

Table 3.2 Daily bidirectional travels between station pairs on the Beijing-Shanghai HSR line

Station pair	**OD pairs**	**Passenger**	**Percentage**
B-B stations	15	77,710	53.04%
B-M stations	66	41,410	28.26%
B-S stations	36	17,580	12.00%
M-M stations	55	5,790	3.95%
M-S stations	66	3,615	2.47%
S-S stations	15	416	0.28%
Total	253	146,521	100.00%

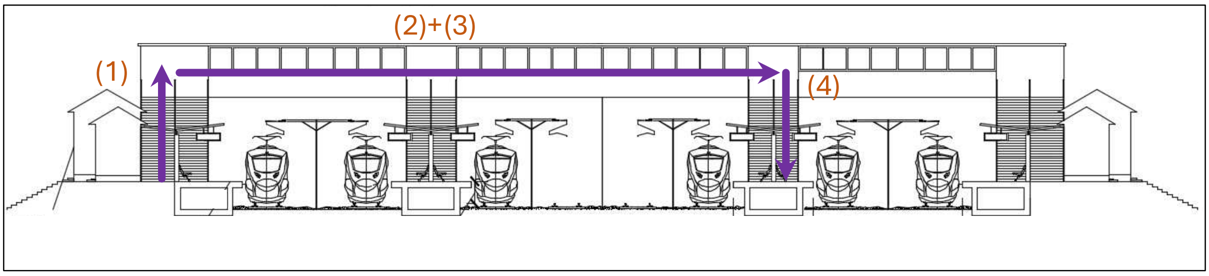
On the other side, from the perspective of transportation product supply, the multiple stops of a singular train create numerous direct OD transportation possibilities $C_n^2 = \binom{n}{2} = \frac{n(n-1)}{2}$, where n is the total stop number of a train. Each direct OD transportation changes in product attributes like arrival time, departure time, and travel time with the train operating. This means that one train could be a 'fast' service for some ODs, while act like a 'slow' service to other ODs. Compared to air travel, where an aircraft typically makes at most one stopover, resulting in three direct transportation products ($C_3^2 = \frac{3\times2}{2} = 3$), one train can offer a significantly greater variety of transportation products.

It should be acknowledged that, although OD groups can be further subdivided (e.g., by coach and seat class; geography topology of cities) and verified with actual data (Ma, 2018), this thesis argues that detailed categorisation of OD groups is better suited to tactical passenger demand predictions and dynamic ticketing studies, which is up to the 15-day ticketing window (Yuan, 2020). Based on station pairs and travel distance, sufficient distinct groups can be established. This allows for the prioritisation and ranking of capacity-related service qualities, such as travel speed and frequency, between OD groups, and providing a solid foundation for delivering a diverse and high-quality HSR service network with limited railway supply.

3.1.2 Understanding passenger transfer process

The latest investigation (Fu et al., 2024), from a pure enumeration perspective, shows that more than 90% of ODs in China HSR network do not have direct services. Although they might not account for the major travel parts of the China HSR, the service network needs to provide effective transfer services (mandatory transfer) if those passengers insist on choosing HSR. Scheduling interchange services involves determining the interchange location, the form of transfer (within the same station or between different stations), and the interchange connections between trains (Guo, 2015; Li, 2018). A typical transfer process is depicted as Figure 3.2. Wardman (2001) further defines the transfer time consisting of Figure 3.2 (1) and Figure 3.2 (4) as access time, Figure 3.2 (2) as walking and adjust time, and Figure 3.2 (3) as waiting time.

There are two types of passenger transfers: mandatory and voluntary transfers. Mandatory transfers occur when no direct service is available. Passengers opting for voluntary transfers typically do so to achieve a shorter total travel time or to benefit from higher frequencies of interchange trains.



(1) Access time: off-boarding and going upstairs,

(2) Walking & adjust time: Walking to the interchange train gate/travelling to other stations,

(3) Waiting time: suitable waiting,

(4) Access time: going downstairs and boarding again.

Figure 3.2 Transfer steps at an interchange station

Voluntary transfers can be easily affected by the following aspects of interchange environment (OECD/ITF, 2014; Wardman, 2001): reliability, information, accessibility, and comfort. Each of these plays a critical role in shaping passengers' transfer experiences and their willingness to engage in future interchanges.

Reliability pertains to passengers' confidence in successfully catching their connecting train. This is particularly significant as passengers often fear missing their connections, especially in infrequent services in the timetable. However, in networks operating periodic timetables with flexible arrangements (e.g., no mandatory seat reservations), passengers have the reassurance of being able to board later trains, thus reducing perceived risks.

Information refers to the clarity and accessibility of signage, maps, and digital guidance systems within the station. These tools are crucial for enabling efficient navigation, particularly for passengers unfamiliar with the station layout. Effective information systems minimise the time required to locate platforms or facilities and alleviate the stress associated with transferring.

Accessibility relates to the ease with which passengers can utilise station facilities and navigate transfer processes. For instance, the availability of lifts and escalators is essential for individuals with reduced mobility, heavy luggage, or accompanying young children. Additionally, the ticket re-checking process can significantly influence transfer efficiency—stations that require passengers to exit and re-enter may discourage transfers due to the added inconvenience.

Comfort is associated with the physical and psychological environment of the station. Factors such as the cleanliness of the station, the availability of seating areas, and the level of crowding all significantly affect passengers' overall experiences. Overcrowded or poorly maintained stations can leave a negative impression, potentially discouraging passengers from choosing transfers in the future.

By addressing these elements, railway operators can enhance the overall interchange condition, encouraging greater acceptance of transfers. However, according to the current investigation in this thesis, there is no systematic study that analyses passenger factors affecting the willingness to transfer. Nevertheless, it can be reasonably inferred that passengers carrying large luggage and /or travelling with children or elderly persons, those not experiencing significant travel time savings from transferring, or those with physical disabilities are less likely to opt for voluntary transfers.

3.1.3 Concluded assumptions of passenger and OD groups

Due to the lack of precise OD data for the whole China HSR, it is necessary to cross-reference existing literature (Li, 2022; Li & Pi, 2024; Li, 2020; Lu, 2021) to reasonably infer the demand characteristics of China HSR OD. The individual passenger behaviour and choices primarily pose the following concerns:

- The first stems from the fact that passengers can have multiple routes for the same destinations, and face attractions from other transport modes.
- The second arises from whether the HSR service network implicitly encourages voluntary and multiple transfers and passengers' attitudes towards transfers.

- The third is the inability to accurately capture schedule delays of PEI.
- The fourth is the travel time preference differ from ODs.

This thesis makes the following assumptions and responses: Firstly, the overall principle for scheduling the service network in this thesis is based on revealed preferences (RP), assuming known backend data. Purchase decisions align with passengers' expectations of service qualities, and the passenger route aligns with the train line itinerary; otherwise, it is presumed they would choose alternative transport modes. This also means that the any prediction regarding passenger behaviours and choice modelling is not included in this thesis. Voluntary transfers are treated as two or more direct travel demands. Lastly, according to Sections 2.4.4 and 2.4.5, the distinction of alternative routes, attitudes towards transfers, schedule delays, and travel time preference is highly related to OD groups differing from travel distance. For medium and long-distance travellers:

- The passenger routes can be regarded as limited or singular.
- Voluntary transfer might occur if considerable travel time saving can be achieved. Multiple transfers, however, are not considered as it strongly discourages passengers to choose HSR services.
- Schedule delays are considered within the goal of increasing daily frequency or the number of trains calling at the station.
- The time preference of the entire travel range is from 10.00 to 16.00 (Li, 2022; Y. Li, 2020), accounting for over 80% of daily travel.

The evidence supporting multiple transfer assumptions stems from two major reasons:

Firstly, the current service structure of China HSR network connects all major cities with direct fast trains. This arterial HSR is supplemented by regional HSR services, which function as capillaries to support the network. Such a system design inherently minimises the need for multiple transfers, as most origin-destination (OD) demands can be satisfied with either direct services or at most a single transfer.

Secondly, Guo (2015) conducted a reliable stated preference (SP) survey specifically investigating transfer behaviour on China HSR network, with a substantial sample size of 5,000 respondents. The survey findings (Table 3.3) reveal that for travel distances of 500–900 km, only 30%–32% of passengers were willing to accept a single transfer. For travel distances exceeding 1,100 km, the willingness to transfer starts at 40% but decreases incrementally by approximately 2% for every additional 200 km, eventually falling to just over 30%. These results strongly suggest that the willingness to accept multiple transfers would be even lower. Additionally, the same study examined transfer times, showing that even with transfer times set at 15 minutes and 30 minutes, 63% and 67% of passengers, respectively, still preferred direct train services.

Table 3.3 The proportion of passengers willing to transfer at different travel distances

Distance (km)	100	300	500	700	900	1100	1300	1500	1700
Proportion	46%	40%	32%	31%	34%	42%	40%	27%	32%

For short-distance travellers (see Figure 1.8), assumptions of these aspects are:

- The passenger routes can be multiple supported by regional HSR network.
- Frequent voluntary transfers, avoiding services running at 'artery' HSR lines, small schedule delays due to the high frequency of regional HSR services.
- Wider travel time preference range compared to medium and long-distance traveller.

The implications of travel time preference are: Firstly, this time range can be understood as the peak hours of HSR, during which high-intensity passenger flows may form in busy HSR network sections, potentially causing intense capacity utilisation. Secondly, any service that departs or arrives within this time period can be regarded as having high 'utility' or 'low' schedule delays. For service network scheduling, this travel time preference serves as an encouraging guideline rather than a mandatory requirement. For instance, the HSR service from Shanghai to Xi'an takes over 7 hours, making it impossible to satisfy both departure and arrival time within the time range of 10.00-16.00 indicated above.

3.1.4 Summary

This section explains the difference between the passenger and the railway backend perspectives regarding OD demands, concluding that station pairs and travel distance are the two most significant factors in identifying distinct OD groups. The existing literature review commonly concludes that there are obvious distinctions between short-distance and medium-to-long railway passengers in various aspects. Regarding transfer and interchanges, and the general finding is that passengers are more likely to accept interchange when service frequency is higher or shorter travel time can be achieved, and less so when it is lower. Consequently, short-distance passengers are best served by regional HSR services with periodic services. For medium-to-long-distance travel, an analysis reveals that HSR services show limited competitiveness for journeys over 2,400 km. These findings offer valuable insights for prioritising different origin-destination groups or developing appropriate service strategies.

3.2 Operational factors

3.2.1 Railway supply indicators (RSI): delivering service qualities

Section 2.4.1 has listed passenger service expectations or desire with 11 indicators classified in three aspects. To align these expectations with the actual services provided, corresponding railway supply indicators (RSI) are also presented below, detailing the quality of services from the railway operator's perspective. Clarifying this differentiation helps prevent confusion in the later sections. For example, transfer mainly refers to passenger behaviours, and interchange mainly refers to the facilities and trains provided by the railway operators.

Table 3.4 The mapping between PEI and RSI

Aspect	Passenger experience indicators	Railway supply indicators
Availability	(1) Coverage	1. Reachability
	(2) Direct travel	2. Direct service
	(3) Transfer	3. Interchange train
	(4) Service hours	4. Operating hours
	(5) Daily/weekly frequency (6) Hourly frequency	5. Service frequency
Travel time	(7) In-vehicle time	6. Transport speed 7. Stop plan 8. Dwell time
	(8) Transfer time cost (walk and wait)	9. Platforming and routing
	(9) Schedule delays/adaption time	10. Service frequency
Reliability	(10) Punctuality (11) Cancellation	11. Resilience and robustness

To schedule a good passenger railway network, railway operators and planners need to consider as many PEIs as possible at every stage of railway planning. RSIs are the direct information from a proposed timetable, and how PEIs of all journeys are satisfied can be analysed via RSI. Not all PEIs can perfectly be satisfied for each journey because the railway capacity of a network is always limited. Prioritising PEIs according to different contexts and determining the appropriate RSIs as the operators' goals can fully realise a high-quality utilisation of network capacity.

3.2.2 General approaches of the service network management

There are several styles of managing the service network: handling the network as the single unit, dividing the network into sub-networks, regarding the network as nodes and lines, and treating differently on intense and underutilised parts differently.

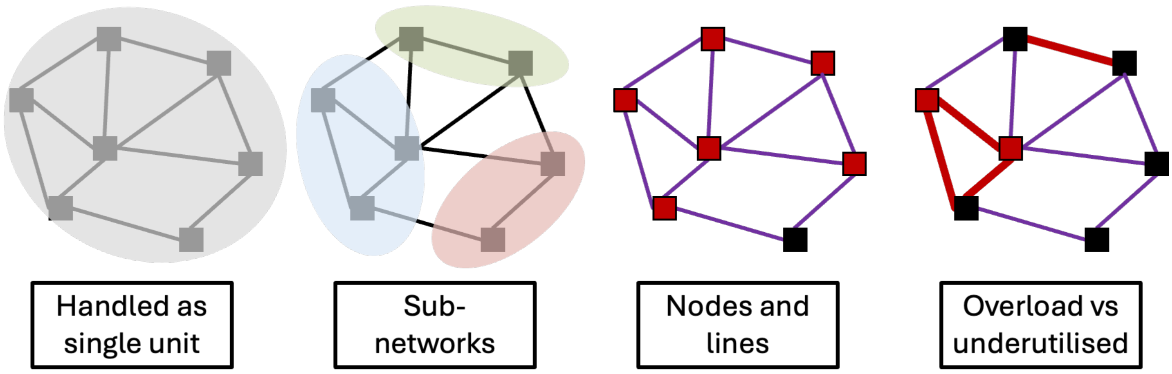


Figure 3.3 Service network classifications

In real railway operations, these approaches are often interwoven. In smaller railway networks with relatively balanced demand between OD pairs, managing the network as a single unit ensures that service improvements are equitably distributed across all potential OD travels. This approach allows for consistent service standards, especially in networks where demand patterns do not vary significantly. China regional HSR is a typical example of this approach. In contrast, dividing the network into sub-networks is often employed in larger and/or more complex railway systems. This method reduces the computational and operational complexity of managing constraints across the entire network. By focusing coordination efforts primarily at the boundaries of these sub-networks, railway operators can optimise operations and service within each sub-network. Dividing into sub-networks is a common practice in real-world railway operations that, such as China arterial HSR, whilst five dispatching regions which further divide into 14 routes in Great Britan's railway.

The latter two approaches are often discussed together. In railway operations, node stations are often the focus, rather than railway lines, due to their role in temporarily holding trains (i.e. inventory), adjusting train precedence, and managing train traffic released on the sequential railway lines from or to several directions. The availability of sufficient sidings is therefore critical. Furthermore, emphasising the congested parts of the network typically implies that capacity research focuses on high-density network parts, as well as the closely associated nodes and lines, thereby reducing the scope of analysis. This is a logical approach, as underutilised areas offer the flexibility to accommodate diverse services or absorb adjustments stemming from high-density areas, especially for China HSR where the train density is highly uneven (See Appendix A).

3.2.3 Operational attributes

1. Network interdependency

From a microscopic perspective, network interdependency refers to train movement relationships established by precedence constraints (Parbo et al., 2016), thus forming routing conflicts and the 'operational coupling' relationships among trains in railway operation. From a macroscopic perspective, dedicated railway line or platform users (i.e. railway operators) reduce network interdependency and negative capacity effects, such as delay propagation (Vansteenwegen & Oudheusden, 2006). Take the overground part of Tokyo station (Figure 3.4) as an example: platforms 1-10 are assigned to JR (Japan Railway) East, where at least seven types of trains are running separately. Platforms 14-19 and 20-23 are set for running HSR trains (Shinkansen) operated by JR East and JR Central, respectively.

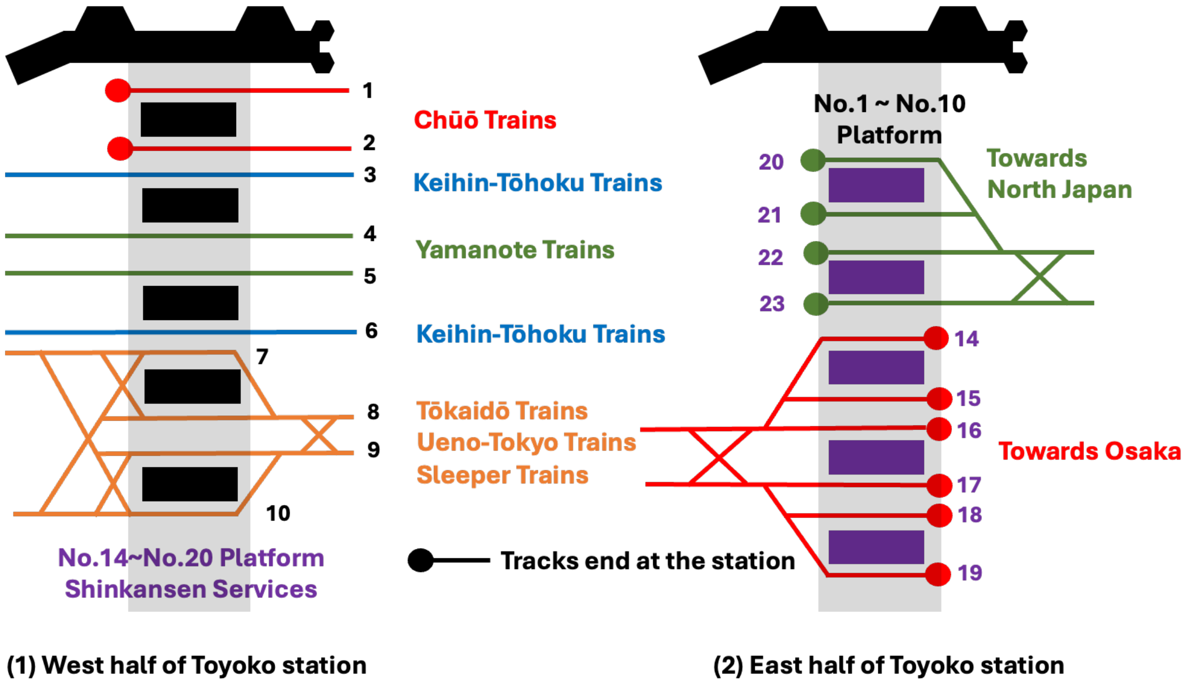


Figure 3.4 Track layouts and users of Tokyo stations (dedicated platforms)

However, if lines in a railway network are fully connected and trains in all directions are allowed to run on the network, a higher network interdependency is expected, and more types and numbers of routing conflicts are generated near station areas. Take Changsha Nan station at China HSR network as an example (Figure 3.5); if there is a train running from Kunming to Guangzhou through the connection point E, its departure route conflicts with other departures from Beijing to Guangzhou, arrivals from Guangzhou to Beijing, and passing routes running on the Beijing-Guangzhou line. Therefore, larger headways need to avoid these conflicts, causing negative effects on railway capacity utilisation.

In short, network independence not only reduces the scale of delay propagation during a disruption but also reduces routing conflicts within each independent network part. This arrangement allows scheduling more trains in the network, thereby possibly increasing overall railway capacity.

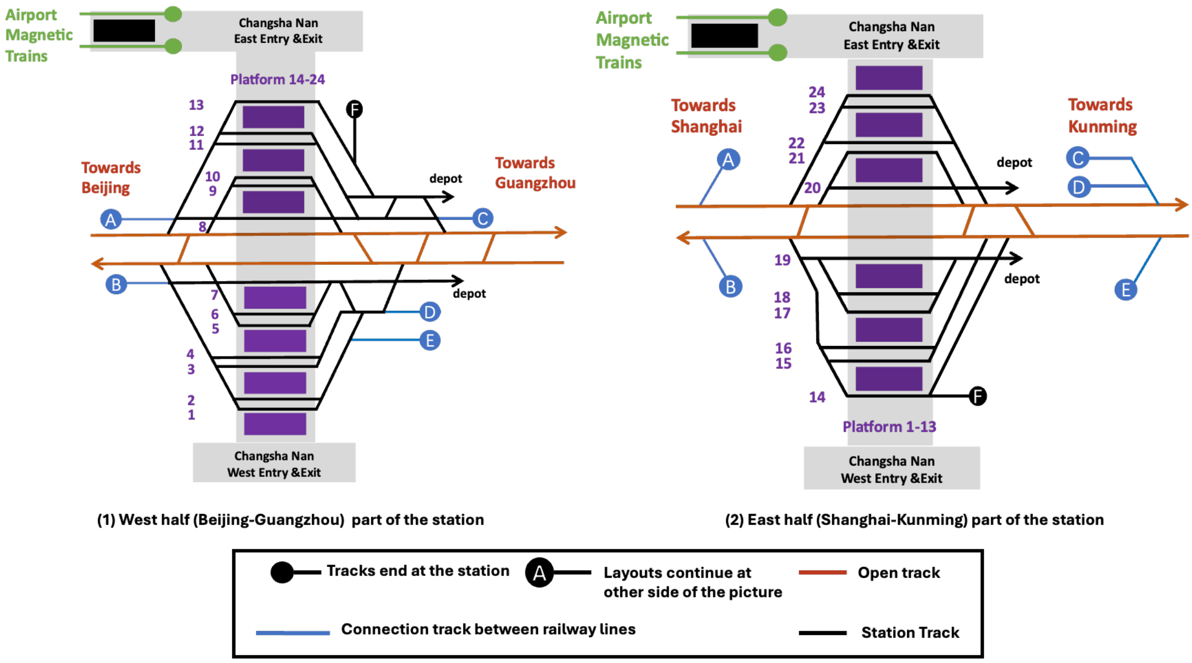


Figure 3.5 Track layouts of Changsha Nan HSR stations (mixed-use platforms)

2. Service type (e.g., direct services or requiring interchange)

Section 1.2.3.1 partially explained the trade-off between operating direct trains and services requiring interchanges. Figures 3.3 and 3.4 further illustrate that Tokyo Station is expected to have a higher volume of transfers due to its more independent railway sub-networks. In contrast, Changsha Nan HSR Station can accommodate a broader variety of direct trains due to providing more directional track connections, thereby reducing the need for passenger transfers. This contrast is even more apparent when comparing underground metro systems and long-distance railway networks, where metro systems typically require more frequent transfers than long-distance railway services.

3. Train speed levels and mix types

Some HSR trains, such as the German ICE, operate on upgraded tracks that can also accommodate conventional passenger trains and freight trains. Freight trains, due to their heavier weight, require longer acceleration and deceleration times and generally travel at lower speeds compared to passenger trains. Consequently, the headway between freight and passenger trains must be greater. In the timetable generation stage, passenger trains, particularly HSR trains, are typically given priority for preferred time slots, as seen with Schnellfahrstrecken (high-speed railway infrastructure) in Germany. Conversely, the United States is one of the few countries where freight trains often have higher priority.

In newly built HSR networks, such as those in Japan and China, only passenger trains utilise the railway infrastructure. In these networks, higher headway values are necessary due to the operation of passenger trains at different speeds—350 km/h, 300 km/h, and 250 km/h. In summary, the fundamental principle in both scenarios is that trains operating at different speeds require longer headways, thus resulting in lower levels of railway capacity utilisation.

4. The periodicity of the timetable

Section 1.2.2.3 already compared different railway systems regarding the periodicity of the timetable around the world. Some common features shared by countries adopting periodic timetables, such as Germany, the Netherlands, and Switzerland (Lei Nie, 2019), are believed to a positive effect on capacity utilisation. Firstly, the stop plans of trains are simple and hierarchical (e.g., fast, semi-fast, slow trains; Intercity Express, Intercity, Interregional), leading to more homogeneous train traffic and a more balanced service level throughout the day. Secondly, these countries also have good interchange facilities at their stations (See Figure 3.6 Switzerland railways), enabling a high coverage of train services by simplifying the proposed train line types.

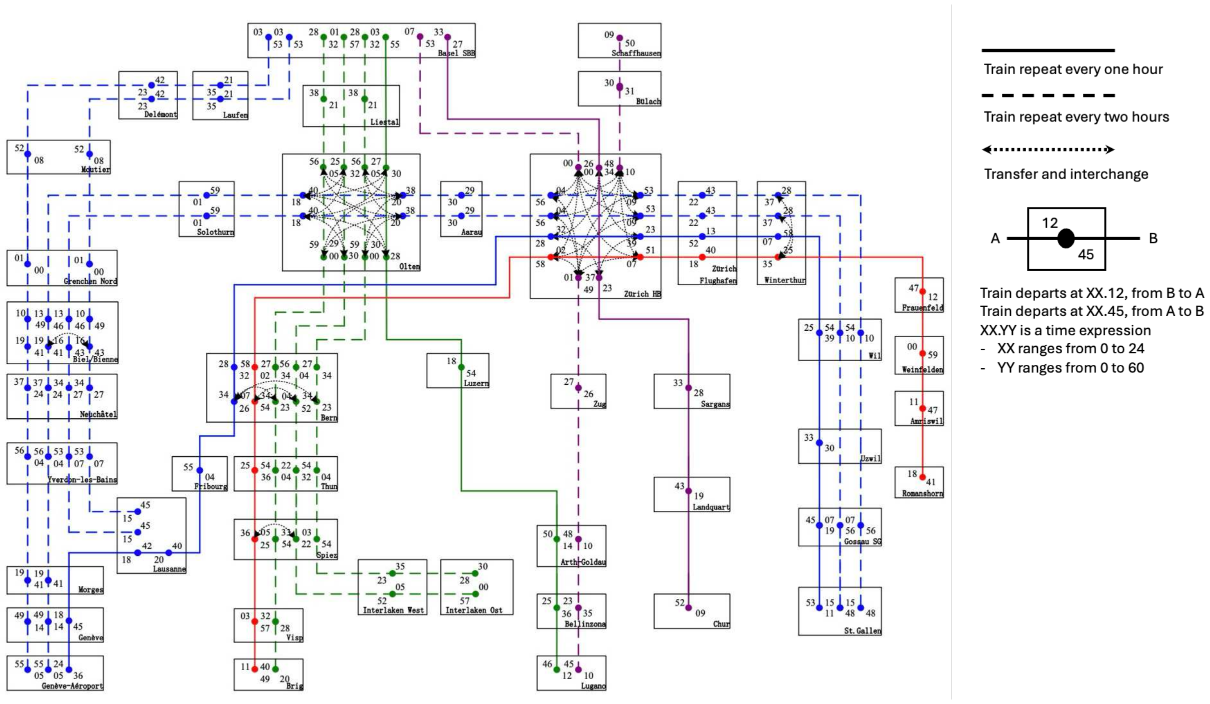


Figure 3.6 The timetable of IC and ICN in Switzerland (Guo, 2015)

Comparatively, aperiodic timetable allows more flexible stopping patterns so that more ODs can be covered by direct services (Fu et al., 2015; Hu et al., 2023, 2024). Besides, there is an ongoing debate regarding which type of timetable can achieve higher capacity utilisation. Proponents of periodic timetables (Yan et al., 2019; Yan & Goverde, 2019) argue that homogeneous train traffic enables lower headway values. In contrast, advocates of aperiodic timetables (Li, 2022; Liao, 2021; Zhang, 2019; Zhang et al., 2019) suggest that avoiding rigid time windows for aperiodic train lines allow for greater flexibility, potentially leading to a more compressed timetable.

3.2.4 Timetabling details

In this thesis, timetabling refers to the process of determining all operational details of a railway system: the types of train services to be offered (train line planning), the schedules for these services (timetable generation), and their platform arrangement (platforming and routing). In practice, additional constraints arise from infrastructure maintenance, rolling stock circulation, and crew assignment (Nie et al., 2021). Once a detailed timetable is established, the railway capacity utilisation of a railway network can be represented microscopically as a time-space graph.

1. The stop plans of a line plan

Apart from determining the frequency, the core content of line planning is the stop plan of trains, which contains the origin (starting) and terminus (end) stations (OT stations) and intermediate stops (calling points). Those stops form the itinerary of a train line, indicating which railway line to occupy (Fuchs et al., 2022). As previously discussed at Section 1.2.3.2, there is a trade-off between the total stop number (to achieve larger OD coverage through direct services) and maintaining acceptable or even attractive travel time cost. Previous works (Hu et al., 2023, 2024) suggest that passengers must often make a compromise: either endure slower trains with numerous stops or make a transfer at a major station to possibly achieve a shorter total travel time.

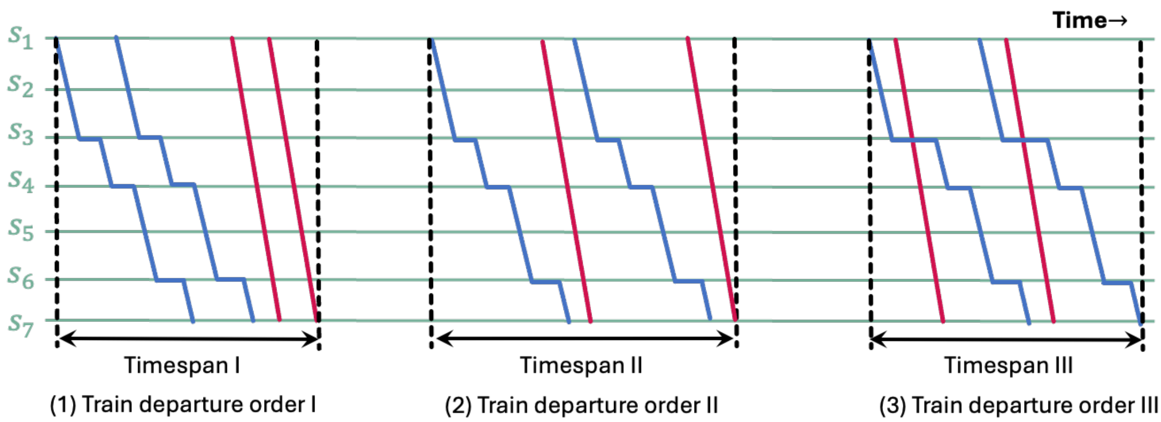


Figure 3.7 Capacity utilisations (timespans) of different train departing orders

Stopping at a station requires more time for a train to accelerate and decelerate compared to passing through. The impact of a train dwell on railway capacity depends on the position of adjacent trains on the time-space graph. If trains with different stopping patterns are adjacent to each other on the time-space graph, a larger headway is needed to isolate these trains, resulting in a higher capacity utilisation value, see the following paragraph and Figure 3.7 (1) and (2) as supplementary explanation.

2. Train traffics of a timetable

Train departure order at originating stations, precedence changes (overtaking), and dwell time allowance are the key elements of train traffic. In railway networks without separate platform tracks, the train departure order determines the whole train traffic for the entire timetable. In this case, if the departure precedence is optimally arranged, the time cost is minimised, resulting in low-capacity utilisation value (Figure 3.7 (3)).

In contrast, in networks with separate platform tracks, train precedence can change within different sections of the railway (i.e. overtaking). As a result, the timetable can lead to lower capacity utilisation value (Figure 3.8).

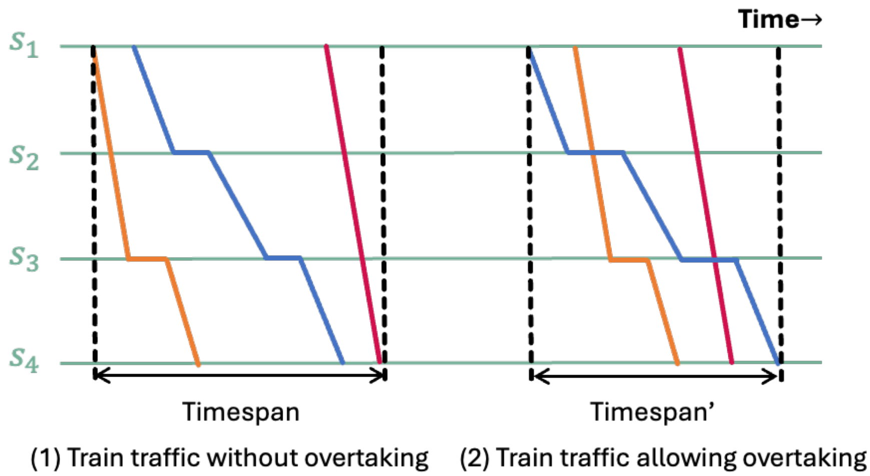


Figure 3.8 Effects of overtaking behaviours on capacity utilisations (timespans)

Moreover, whether adding dwell time allowance can impact the following train and the train traffic. A minimum dwell time is necessary to complete passenger boarding and alighting. However, if a stopping train adheres strictly to the minimum dwell time and the following train is passing, the following passing train requires a higher headway to ensure safety. Conversely, allowing dwell time allowance allows the following train to overtake the stopping train while maintaining a 'compressed' timetable (Figure 3.9). Similarly, as discussed in Section 2.3.1.1, dwell time allowance and platform numbers can jointly affect the station inventory and capacity utilisation of the station.

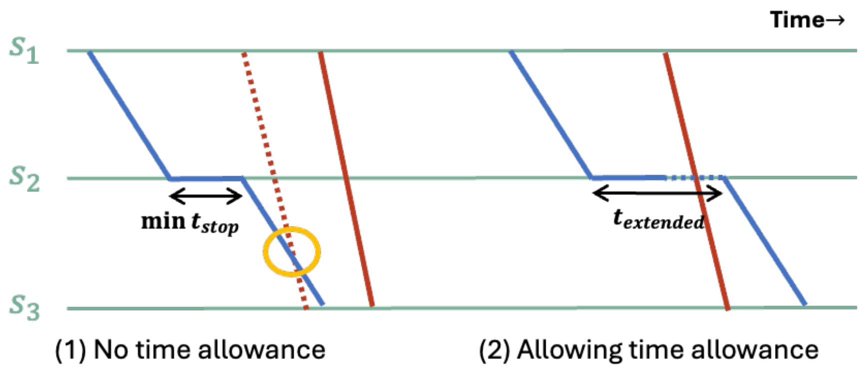


Figure 3.9 Effect of time allowance on capacity utilisations (timespans)

3. Platforming and routing conflicts

The objective of platforming is to minimise routing conflicts (Schlechte, 2012; Wang et al., 2023; Zhang et al., 2020) and arrange the maximum number of parallel conflict-free routes (Li, 2019), which are significantly influenced by station types and track layouts. During timetable generation, it is crucial to assess whether a railway station has sufficient platforms and an efficient switch layout design. Insufficient platforms can prevent a station from accommodating many stopping trains within a short period. Additionally, an inefficient switch design can increase the number and types of routing conflicts, consequently, extending the train headways.

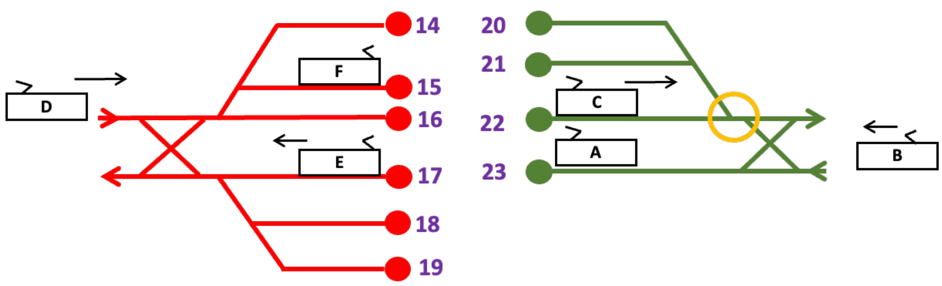


Figure 3.10 HSR track layout parts of Tokyo station

The opposite conflict, where trains travelling in opposite directions request the same switch occupation, is known to have longer headway compared to other types of headways (Liao, 2021; Zhang, 2022). A simple example can be observed at the HSR platforms of Tokyo Station (Figure 3.10). The tracks at platforms 14-19 and platforms 20-23 have similar switch layout settings. It is evident that there is a routing conflict between the arrival of Train B and the departure of Train C and Train A, whereas the arrival of Train D does not conflict with the departure of Train E, which is a conflict-free route. If one of the tracks at platforms 20 and 21 were located on the lower side of platform 23, the conflict between Train B, Train C, and Train A would not occur, like the layout at platforms 14-19.

4. Joining and splitting train sets

There are two straightforward rolling stock utilisation measures that can help alleviate intense capacity utilisation or bottlenecks during peak hours. The first measure is to couple multiple train sets to increase seat capacity. The second is to couple train sets with different terminus stations that partially share the same itinerary. This operational mode requires only one train path in the busy parts of the network, thus releasing space for other trains and is commonly implemented in the UK, Germany, Switzerland, and other Western countries. Figure 3.11 shows an example.

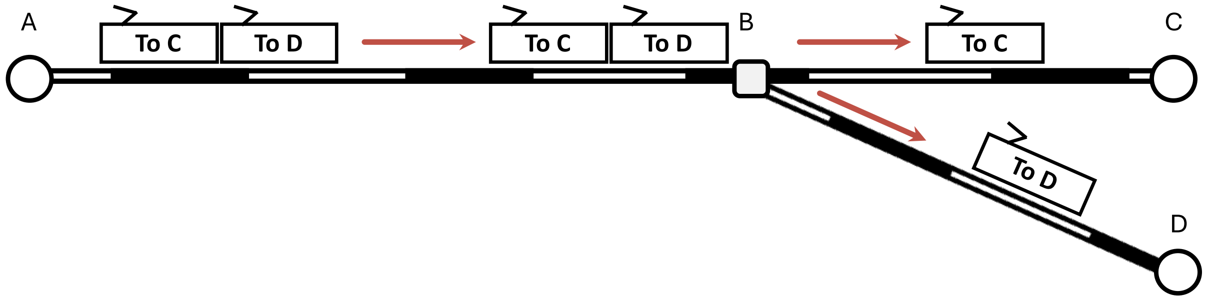


Figure 3.11 Joined train sets with different terminus stations

3.2.5 Organising interchanges

For an HSR service network that values more on direct transportation, replacing some direct services for cross-line passenger flows with services that contain transfers can help eliminate complex route conflicts of trains in the switch areas. The cumulative effect of this change is very likely to enhance the overall network capacity, such as potentially arranging more trains, improving the independency level of network parts (Nie et al., 2021) and therefore more stable timetable performance (Parbo et al., 2016), as previously discussed. Following this motivation, transfers between different stations within the city are not included in this thesis because trains in different stations are regarded as network independency.

1. Principles of deciding interchange location

Many scholars (Guo, 2015; Hu et al., 2024; Ke et al., 2024; Li, 2018; Nie et al., 2021;Nie, 2019; Zhang, 2019) argue that it is not advisable to arrange too many interchanges in a large-scale intercity railway service network, as creating feasible transfers essentially introduces additional constraints into timetabling, which increases the likelihood that the desired timetable will be unfeasible. Reasonable prerequisites, such as the selection of interchange locations (Schiewe, 2020), are recommended to be set in advance. The interchange location is recommended to have the following conditions:

- Sufficient platforms or equipped with depot to accommodate trains.

- High frequency of train services to provide multiple interchange trains.
- Extensive interchange directions to ensure convenient connections among HSR lines.

From the perspective of optimal capacity utilisation, sufficient platforms can allow trains to dwell for extended periods to support smoother interchange. Similarly, if a train terminus at an interchange location, equipping a depot can enable quick train clearances. Both can maintain a good level of station inventory. Besides, an interchange location should have a high frequency of train services to maintain a stable level of interchange, ensuring that even if passengers miss a connection, they can still catch the next available service. Likewise, an interchange location should be situated in the network accommodating trains from multiple directions, providing interchange opportunities across different railway lines. While interchanges help manage capacity, direct services should still be prioritised on railway network parts with low train density.

2. Deciding transfer types based on cross-line train direction

From the perspective of analysing train running directions, it is important to identify and to limit those opposite conflicts when operating cross-line trains (i.e. the opposing conflicts depicted in Figure 3.12). Figure 3.12 shows how connection tracks form the conflict types at Changsha Nan (South) station.

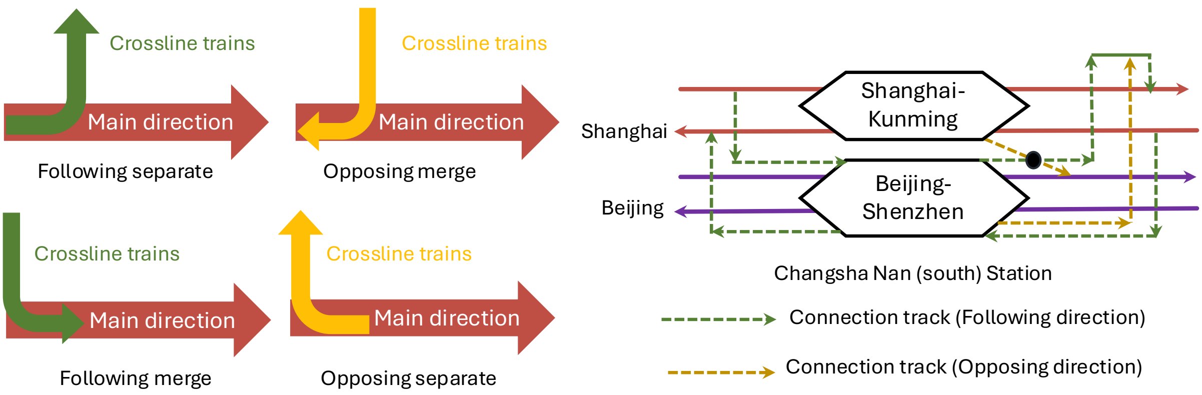


Figure 3.12 Routing conflict types by operating cross-line trains

The following direction is defined in this thesis as the scenario where cross-line trains (green arrows in left side of Figure 3.12) run in the same direction with original-line trains (orange arrows in the left side of Figure 3.12) when arriving at or departing from the node station, resulting in smaller headways compared to opposing conflict routes. The opposing direction is defined as the scenario where the cross-line trains (yellow arrows in left side of Figure 3.12) move in reverse order with original-line trains when arriving at or departing from the node station.

This section further explains why the difference between cross-line train movements in the following scenario and original-line trains train movements is smaller. Referring to Figure 3.13, it illustrates the sequential arrivals of two original-line trains. The headway between them is safeguarded by the speed protection curve, ensuring a separation of two blocks. When Train A fully exits the open track and enters the switch areas, the automatic signalling system updates the movement authority, with the minimum speed set to 75 km/h. It is important to note that the headway in this section is three minutes, which allows Train A ample time to come to a complete stop at the station. According to Figure 3.5, as long as Train A and Train B are allocated platforms that are sufficiently distant from one another, the number of conflicting switches between them is minimised. Li (2019) verified that the vast majority of stations in China are capable of handling arrival tasks within a three-minute window.

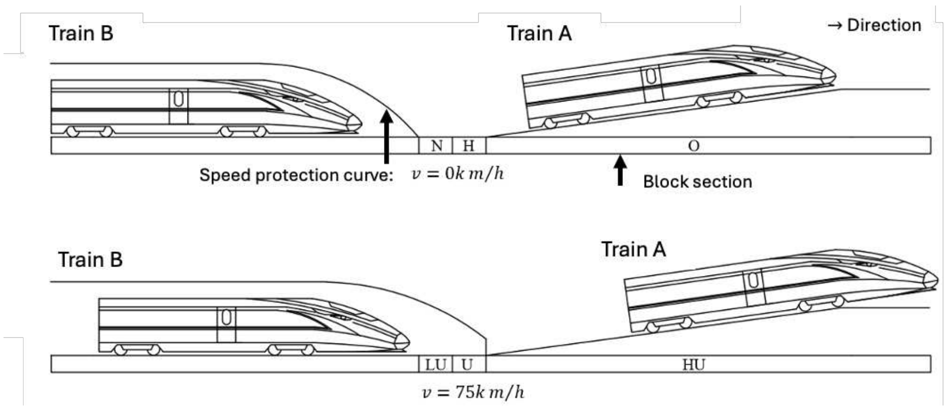


Figure 3.13 A pair of arrivals between two original-line trains

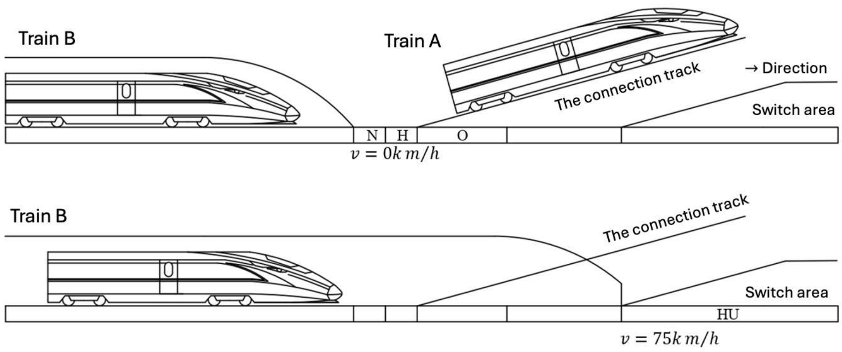


Figure 3.14 A pair of arrivals between a following cross-line trains and an original-line train

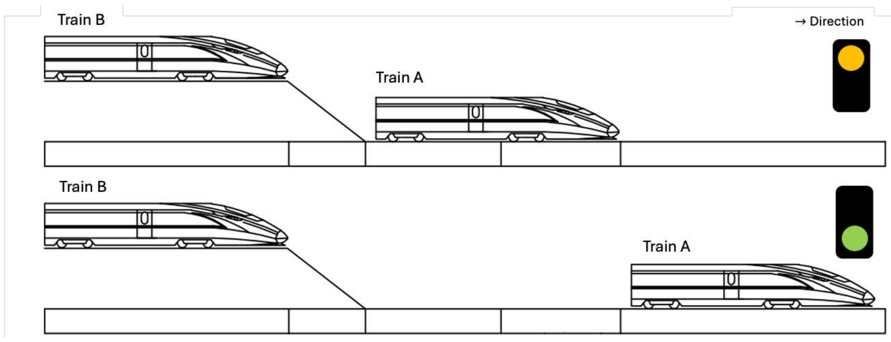


Figure 3.15 A pair of departures in China HSR

Similarly, connection tracks are often separated from the open track at a considerable distance before the switch areas, allowing cross-line trains (Train A in Figure 3.14) to exit the open track even earlier. As a result, the following train, Train B, is unlikely to be affected.

Table 3.5 The signal profile of China Train Control-level 3

Code	**Meaning**
LU	Two available block sections ahead
U	One available block section ahead
H/HU	Stop by the end of this block section
N/Null	Entry forbidden
O	Occupied

For following departures, the process is relatively straightforward. Regardless of the complexity of the departure switch area, all departing trains must return to the open track or access the open track of another railway line via the connection track. If the connection track to an alternative HSR line is positioned sufficiently far from the original HSR line, the departure operations can be compressed, or even synchronised, allowing for the simultaneous handling of two departure tasks.

However, in both cases, conflict-free routes can still be arranged, as illustrated at Point E in Figure 3.16, where cross-line trains (Green arrows) arriving at Platforms 1 and 2 operate independently of other arrivals at Platforms 4 to 7 (Purple line sections). However, an excessive number of cross-line arrivals can be problematic due to the limited availability of only two platforms from Point E and another platform 14 from Point F.

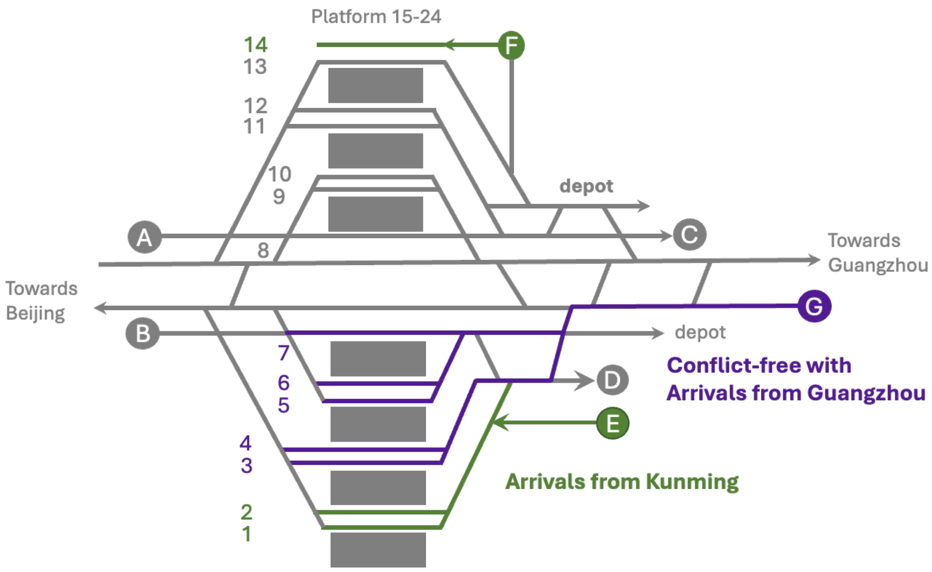


Figure 3.16 Conflict routes analysis: cross-line trains arrive from Kunming.

In contrast, the opposing direction not only extends headways between trains but also causes joint impacts on related platforms. In Figure 3.17, departures from Platform 1 and 2 (yellow arrows) render Platforms 3 to 7 (orange line sections) unusable for arrivals and open track usable for passing through, whereas departures from Point F exhibit fewer such negative effects.

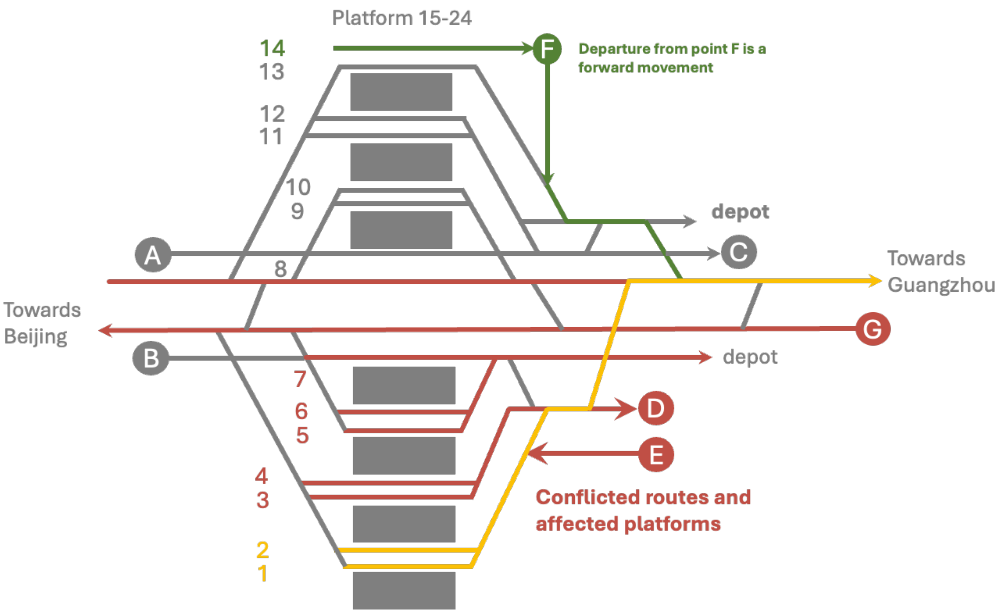


Figure 3.17 Conflict routes analysis: cross-line trains depart towards Guangzhou

In summary, following direction of cross-line train movements cause less variation in headways, similar to operations of original-line train. Direct service can be reasonably scheduled based on passenger demand, even if it might require a high rate of cross-line train. Joining rolling stock operations at the interchange locations can be considered where appropriate. The opposing direction of cross-line train movements not only requires longer headways but also, due to the track layout, causes associated arrivals or departures unable to proceed simultaneously or requiring a longer headway. Therefore, it is advisable to arrange transfers where cross-line train directions moving opposing directions within the stations.

3. Analysing passenger transfer routes within the station hall

From a mathematical modelling perspective, forming a feasible interchange relationship essentially requires that the time difference between the arrival of the leading train and the departure time of the interchange train must be greater than the total transfer time (Guo, 2015; Li, 2018). As described in Sections 3.2.3 and 3.1.2, the following four types of in-station transfer can be established. As illustrated in Figure 3.18, it can be observed that the higher the 'degree of freedom' in the interchange relationship, the stronger the interdependency between the trains.

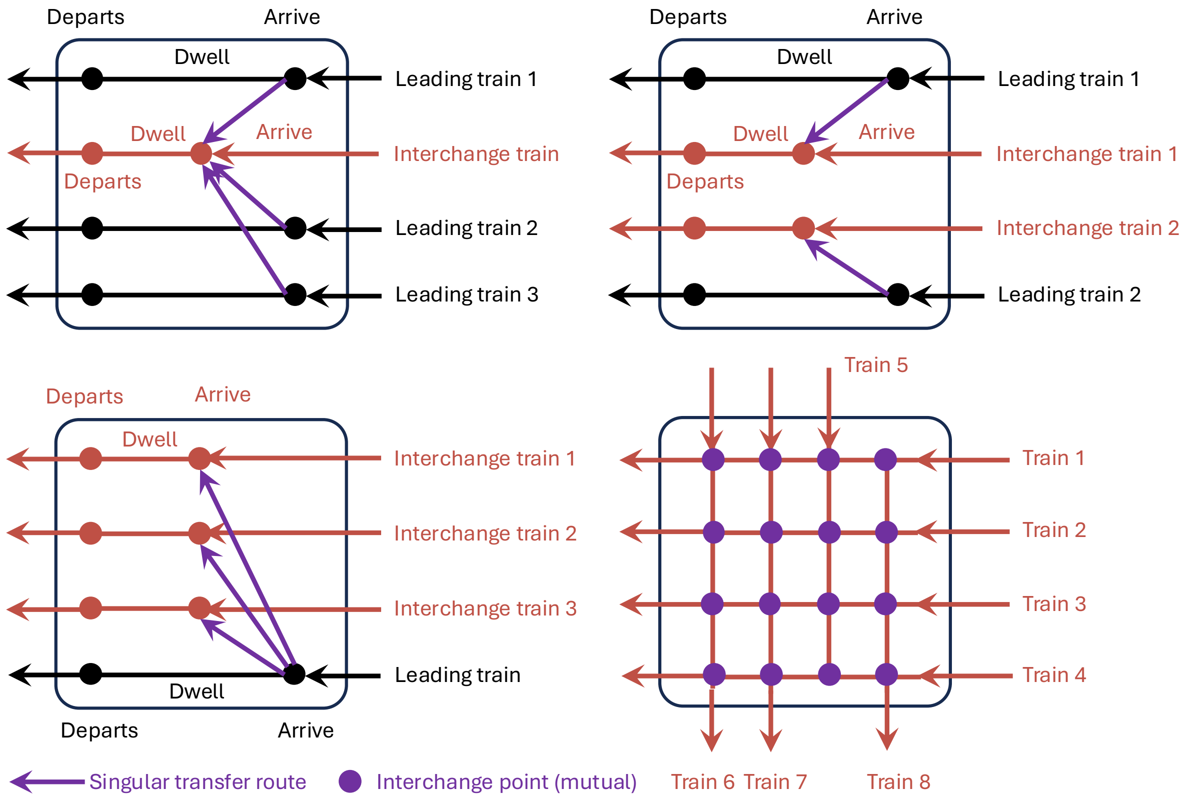


Figure 3.18 Different forms of transfer routes at the interchange station

Regarding platforming and routing, arranging the leading train and the interchange train on adjacent platforms seems to be the most convenient measure for passengers. Although a route can be designed from the track layout perspective to achieve this, according to Section 1.2.2.1, the platform tracks of China HSR are arranged on both sides of the open tracks, with each side's platform track aligned in the same direction as one open track. This means that arranging adjacent platforms either leads to excessive opposite route conflicts (i.e. opposing directions) or necessitates that the leading train and the interchange train travel in the same direction (i.e. following direction), increasing the likelihood that transfer passengers will be attracted to direct service.

In summary, organising transfers at busy node stations requires an analysis of the track layout to identify potential conflict forms between trains from various directions. From the perspective of establishing constraints, it is essential to consider the types and scales of transfers. Current qualitative analysis concludes that if arranging adjacent platforms in the context of China HSR does not significantly reduce route conflicts, it contradicts the motivation for arranging transfers: increasing and improving the capacity.

3.2.6 Summary

This section (Section 3.2) first introduces 11 railway supply indicators that correspond to 12 capacity-related PEIs across three aspects (availability, travel time, reliability), further demonstrating the central role of line planning and timetable generation in railway capacity utilisation. The analysis of operational attributes reveals that a railway operator's attitude towards transfer and interchange significantly influences the service network's interdependency and the decision to implement a periodic timetable. Passenger travels containing flexible transfers typically occur in short-distance, high-frequency networks with simple, easily memorable stopping patterns, which are related features of periodic timetable of real-world cases.

The analysis of timetabling details indicates that the scheduling styles behind stop plans in line planning heavily relies on a deep understanding of passenger flow and OD groups characteristics. Key factors in determining whether a timetable can be compressed, and thus achieve lower capacity utilisation values, include dwell time allowances and overtaking behaviours.

For networks focused more on direct services, considering interchange is often regarded as a compromise on the provision of increasing capacity. This section examines three critical related factors: interchange location, cross-line train directions, and passenger routes within the station hall. It concludes that discussions should focus on whether direct or interchange services are appropriate for cross-line trains operating in the opposing direction within busy node stations.

3.3 Concluded capacity-related scheduling outline for China HSR

Instead of the traditional geographically managed by railway bureaus, this thesis investigates for centralising congested HSR networks, particularly around the Beijing-Shanghai and Beijing-Guangzhou-Shenzhen-HK line. For major Chinese cities, clearly defining directional roles is crucial to prevent an overconcentrated of arrivals and departures at a single station, which can lead to longer headways and negative impacts on capacity utilisation, as seen at Beijing Nan (South) station. By contrast, the shared roles of Shanghai and Shanghai Hongqiao stations, with a total of 41 platforms, successfully manage arrivals from the north China and regional HSR around the Shanghai-Hangzhou-Nanjing area.

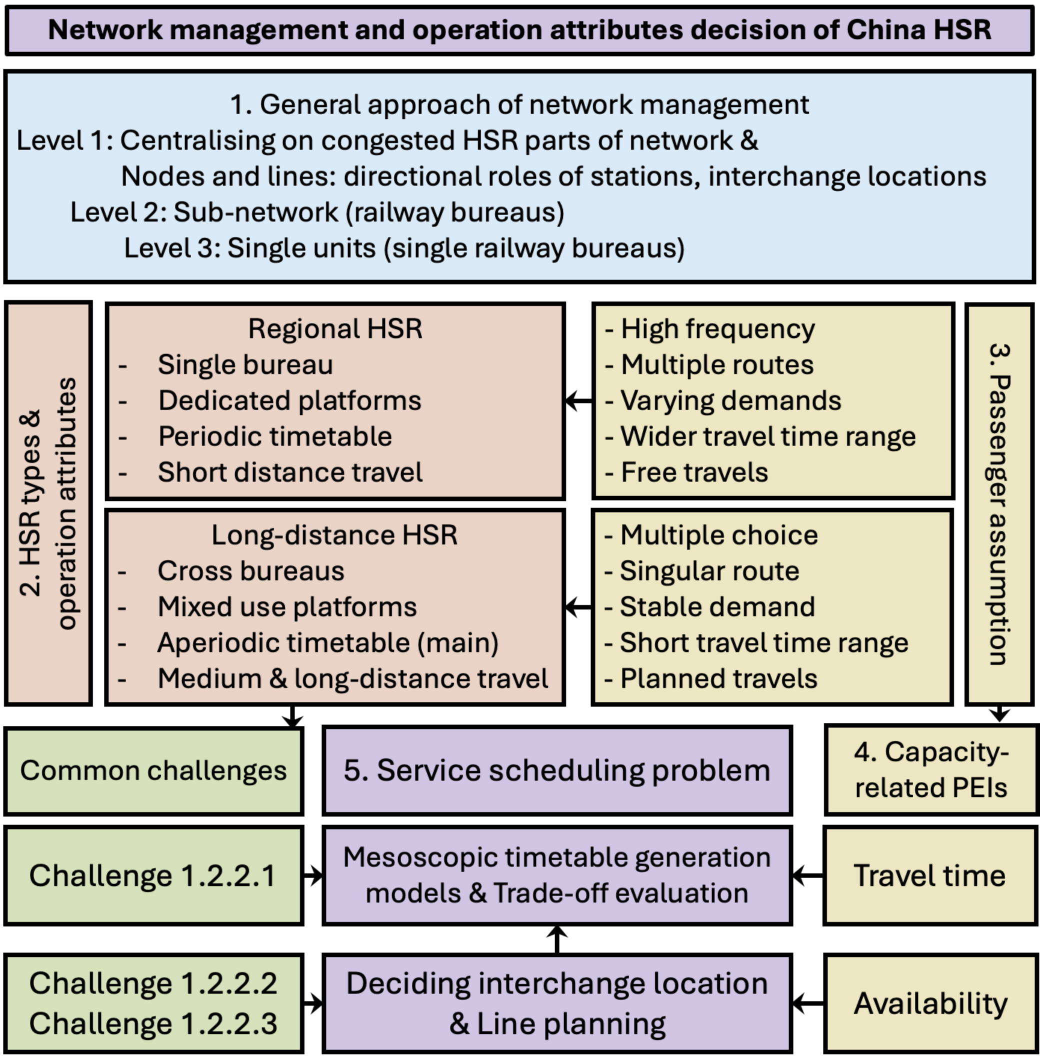


Figure 3.19 The outline of China HSR network capacity utilisation and optimisation workflow

Node stations on these two HSR lines such as Shijiazhuang, Zhengzhou Dong (East), Changsha Nan (South), and Nanjing Nan (South) serve as potential interchange locations. As discussed in Section 4.2, optimising OT stations of cross-line train can further reduce the number of necessary interchange points. According to the demand feature analysis of different types of passengers, regional HSR lines should focus on short-distance services, while long-distance HSR routes should handle medium to long-distance travel demands. Long-distance HSR reflects better regarding the three challenges identified in Section 1.2.3, which are commonly shared across railway contexts. These challenges, from a passenger perspective, are denoted as availability and travel time, capacity-related PEIs. To address these challenges and providing HSR services satisfying PEIs, the service scheduling problem is structured as a two-stage models: line planning and then timetable generation, which allow for the examination of capacity utilisation and capacity improvement strategies through various line planning approaches.

3.4 Methodology of capacity-related service scheduling problem

Instead of forming an integrated problem, the solution for the service scheduling problem defined in Section 1.2.1 is divided into a line planning problem (Chapter 4) and a mesoscopic timetable generation problem (Chapter 5). Adhering to the conventional passenger railway planning stages is believed to allow the separate model structures to better align with more railway operation realities, more effectively meet demand features, and better answer the challenges summarised in Section 1.2.3.

It should be noticed that this thesis focuses more on improving the availability and travel time aspects of PEIs, which align with the general motivation in Section 1.1 while placing less emphasis on resilience. This consideration is based on several considerations:

- Firstly, this thesis relies on the fact that the determination of headway time shall consider certain levels of stability thus absorbing light errors and delays (Ning et al., 2009).
- Secondly, China HSR operates with conservative headway values and achieves high levels of station-to-station punctuality (Ministry of Transport of the People's Republic of China, 2020) thereby inherently reducing the immediate need for additional stability, robustness and resilience. Thirdly, by reducing network interdependency, it is expected that the overall stability of the network will improve, as fewer interdependencies mean that delays or disruptions are less likely to propagate across the network.

3.4.1 Principles of line planning model

In response to challenge 1.2.3.2 (coverage, attractiveness versus capacity), the proposed line planning model is established with the idea of providing extensive direct services through a line plan with smaller number of stops and train line combinations. This approach not only reduces passengers expected total travel time and minimises the number of mandatory transfers, thereby enhancing the attractiveness of train lines. Given the fewer stops and train combinations, it is expected that the capacity utilisation value and total travel time of these proposed line plans will be low.

In response to challenge 1.2.3.3 (service scheduling styles versus capacity), line planning categorises OD groups based on objective attributes (e.g., station pairs, travel distance indicated in Figure 3.1) and test corresponding scheduling styles, thereby offering diverse service needs to OD groups. The macro-operational characteristics of different countries essentially provide targeted service scheduling styles for different OD groups or railway contexts.

In short, line planning primarily delivers the availability aspect of PEIs. Meanwhile, as the stop plan is also determined, the expected In-vehicle time for each OD can be calculated and will be finally confirmed in the timetable generation stage.

3.4.2 Principles of timetable generation model

In response to challenge 1.2.3.1 (trade-off between direct services and transfers), the timetable generation model uses a known line plan as input and is set at a mesoscopic level. This approach ensures that the model's complexity remains manageable while considering the platforming and routing problem. This decision is based on the fact that, in the real-time traffic management stage, operations, a route is formed by signal aggregation. The model employs the objective functions, aiming to minimise railway capacity utilisation.

The timetable generation model further reduces capacity utilisation or improves service quality through the following measures: First, overtaking is allowed to ensure that the timespan of the timetable shorten. Second, for the known line plan, attempts are made to add stops to increase the direct reachability of more ODs, thereby testing the sensitivity of capacity utilisation changes.

It should be noted that this thesis arranges transfers due to the high proportion of cross-line trains in the network, which creates numerous route conflicts near node stations. To address this, it is necessary first to determine interchange stations based on current HSR train density situation and analyse the types of route conflicts at these stations (Figure 3.10). For following-type conflicts, they are similar to original-line trains because the conflict type is not opposite. Therefore, Chapter 5 mainly explores the chance of combing train lines and implementing joining operations. For opposing-type conflicts, Chapter 5 develops a trade-off evaluation method based on a 'full-timetable' scenarios.

Chapter 4 Providing multiple levels of service: a railway supply-based line planning method

This chapter introduces a railway supply-based line planning method designed to optimise service levels across OD pairs. First, an interchange location decision method is developed based on a conceptual framework of deciding the interchange locations, ensuring efficient transfers and minimising unnecessary interchanges. Second, a line planning model is proposed to provide multiple service levels, ensuring that most OD pairs have access to fast train options. The adaptability of this model to various railway contexts is also discussed. Finally, different scheduling styles are tested and compared with real-world case studies, offering insights into the practical application and effectiveness of the proposed method.

4.1 Railway operation terminology

1. A train line is a train itinerary including a stop plan (i.e. calling point list), containing the originating (i.e. starting, O), terminus (i.e. ending, T) stations, and intermediate stop stations, rolling stock set form (8-car, 16-car), and a frequency.
2. The line pool is a train line set containing train lines as candidate. Each line plan proposal is the subset of the line pool.
3. Railway stations are classified into big (B-), medium (M-), or small (S-) stations, and this thesis does not analyse the specific methodology for determining station category. This distinction is made based on factors such as economic status, population, geographical location, and position within the railway network (e.g., a hub or a node).
4. Train types are classified into Fast train (F-train), Semi-fast train (S-train), and Local train (L-train). F-trains, S-trains, and L-trains essentially represent three distinctive stopping patterns. F-trains only stop at the big stations, S-trains stop at big and a subset of medium stations, and L-train stop at big and a subset of medium and small stations. Each train here represent the railway supply indicator (RSI) in terms of In-vehicle time and reachability.
5. A passenger is defined an origin-destination (OD) demand, and passenger groups as OD groups. The station pairs refer to the station category of origin and destination, such as B-B, B-M, and B-S. Passengers expecting to take F-trains, S-trains, and L-trains are referred as F-passengers, S-passengers, and L-passengers. The OD associated with B, M, and S stations refers to the travel from or to B, M, or S stations.
6. The scheduling style refers that the general structure of a line plan shows a clear feature, as similarly stated in Table 1.2.
7. The marketing policy refers to a certain seat allocation priority to defined OD groups.

8. A parallel stretch (forward) is a list of consecutive stations occurring in different train lines consecutively in the same order (Schiewe, 2020).
9. An anti-parallel stretch (opposing) is a list of stations occurring in two trainlines consecutively in reverse order (Schiewe, 2020).

4.2 Deciding the minimum interchange locations

In actual railway operations, operators typically predefine certain basic train lines and primarily determining the origin and terminus (OT) stations of some train lines, along with the interchange locations. This determination process is often qualitative and subjective, as suggested in Section 3.2.5. However, if this process is generalised, it can be addressed using graph theory methods. This thesis presents a solution of Selecting Interchange Locations based on the Weighted Maximum Spanning Tree (SILWMST) as Section 4.2, functioning as preliminary step in the overall line planning solution, allowing railway operators to decide, based on their own railway contexts, whether to adopt this step.

4.2.1 Solution inspiration

1. The original model of deciding interchange location

Schiewe (2020) introduced an optimisation model that decides the minimum interchange station number based on a given line plan. Define that $x_s = 0, 1$ whether a station s is selected as interchange location. Other definitions: the stretch set $P = (s_1, s_2, .., s_n)$ covered by train line l_1 and l_2, and $s_1, s_2, .., s_n$ are the station s index. The model is structured as:

$$min \sum x_s$$

s.t. $x_s = 1$ In this constraint, the x_s shall satisfy the following:
If P is the max. anti-parallel stretch of $l_1 \cap l_2 \neq \emptyset$
s is the first station in P with $|P| \geq 1$
s is not the origin and terminus station of l_1 or l_2

$\sum_{s \in P} x_s \geq 1$ In this constraint, the x_s shall satisfy the following:
If P is the max. parallel stretch of $l_1 \cap l_2 \neq \emptyset$
$l_1 \cap l_2 \neq l_1$ or l_2
l_1 or l_2 do not both originate or terminus in station in P

The Figure 4.1 shows some examples of parallel and anti-parallel stretches, and the minimum interchange station selection model is formed based on the following principles, and a full proof progress is provided by Schiewe (2020):

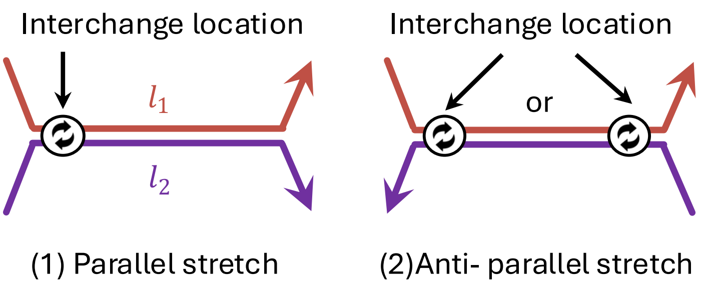


Figure 4.1 Graph description of stretch types and interchange location

- *There is at most one interchange location between* l_1 and l_2 *in a parallel stretch.*
- *There is no interchange between* l_1 and l_2 *in parallel stretch if either itinerary of* l_1 or l_2 *can be covered by each other or both lines originates or terminus in parallel stretch.*
- *There is no interchange between* l_1 and l_2 *in anti-parallel stretch if either* l_1 or l_2 *originates or terminus at the first station of parallel stretch.*
- *If there is an interchange location between* l_1 and l_2 *in anti-parallel stretch, it has to be the first station in the anti-parallel stretch.*

2. Interference

It can be anticipated that in more complex scenarios (Figure 4.2), there is a trend towards further optimisation of train line itineraries, leading to a reduction in interchange locations. Firstly, when train lines operate bi-directionally, the number of interchange locations decreases compared to operating singular direction train lines, as shown in Figure 4.2 (1). Secondly, if train line itineraries are extended, as demonstrated in the transition from Figure 4.2 (2) to Figure 4.2 (3), interchange locations are further reduced, potentially eliminating the need for any transfers.

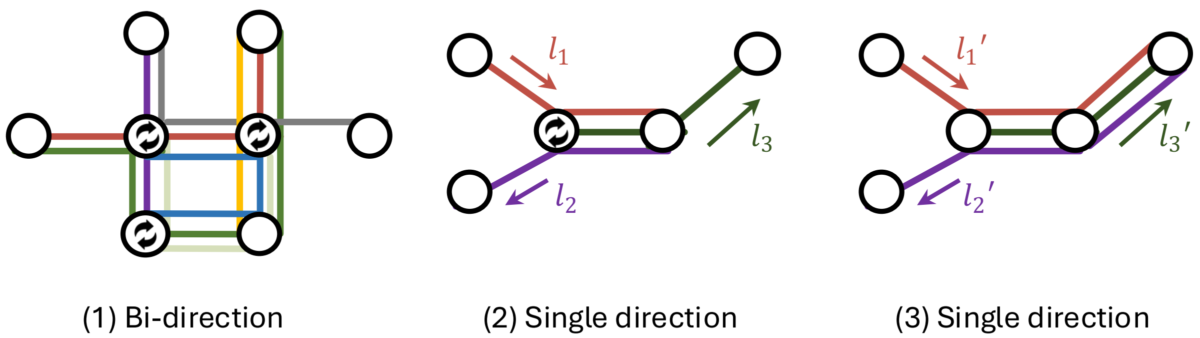


Figure 4.2 Interchange locations at more complex scenarios

4.2.2 A solution process based on spanning tree generation (M1)

Based on this reflective analysis, the interchange location decision problem can be defined as follows: Given a railway infrastructure network, identify a combination of basic train lines with optimal origin and terminus stations that minimises the number of interchange locations.

Considering that the primary goal of HSR services is to provide fast links from/to/between major cities of a country, the reflective analysis shares the essential principle with the spanning tree.

Step 1: General solution process

Assume a given railway infrastructure network as undirected graph $G = (V, E)$, where (v, e) represents the edge (i.e. HSR section) connecting vertex v and e, and $w(v, e)$ represent the weight of this edge, if there exist a subset $T \subseteq E$ such that (V, T) forms a tree, and the total weight is: $w(T) = max \sum_T w(v, e)$.

then, the T is the weighted maximum spanning tree (WMST) of railway infrastructure network G. The solution process of WMST can adopt cycle-elimination algorithm (Zhao et al., 2023), which is the process of identifying and removing any edge that creates a closed loop (cycle) within the spanning tree structure, thereby ensuring the tree remains acyclic.

Step 2: Deciding the weight of HSR section

For a simplified consideration, the weight $w(v, e)$ is set as the distance between HSR sections. If further combining advanced geographic factors, it is necessary to consider a method that allows for the determination of weights of these diverse factors under non-dimensional conditions. For example, the population of a city, GDP, administrative rank, OD data, all have different dimensions. This thesis considers adopting Grey Relational Analysis (GRA) (Li, 2015) as an illustrative example. The reasons are listed as follows:

- **Handles multiple indicators with different units effectively** – GRA normalises data efficiently, making it easier to integrate economic, demographic, and administrative factors into a single relational metric without complex conversions.
- **Ability to handle small sample sizes and incomplete information** – Since intercity relationships involve multiple factors that may not always have uniform datasets, GRA is better suited to working with limited or uncertain data.
- **Captures relative importance rather than absolute variation** – Unlike entropy weighting, which relies heavily on statistical variance, GRA measures the similarity between different factors, allowing for a more direct evaluation of how GDP, population, and administrative rank collectively influence city connectivity.

- **Dynamic adaptability to evolving data –** Since city linkages evolve over time (e.g., shifts in economic centres or administrative upgrades among cities), GRA is well-suited to tracking changing patterns without requiring full-scale model recalibrations.

Define X_{ij} that there is i HSR sections of tree T and j factors considered, if some factors could be regarded as 'benefit', the normalisation calculation formula is:

$$X_{ij} = \frac{x_{ij} - min(x_j)}{max(x_j) - min(x_j)}$$

Comparatively, if some factors are 'cost'. the normalisation calculation formula is:

$$X_{ij} = \frac{max(x_j) - x_{ij}}{max(x_j) - min(x_j)}$$

Where X_{ij} is the normalised value of i-th HSR section for the j factor. Then, define Δ_{ij}= $|X_o(j) - X_i(j)|$ is the absolute difference between ideal value and the actual value of j, $min(\Delta)$ and. $max(\Delta)$ are the minimum and maximum values Δ_{ij}, d_i is the distance of HSR section, calculate the grey relational coefficients $\mathcal{E}_{ij}$:

$$\mathcal{E}_{ij} = \frac{min(\Delta) + max(\Delta)}{\Delta_{ij} + max(\Delta)}$$

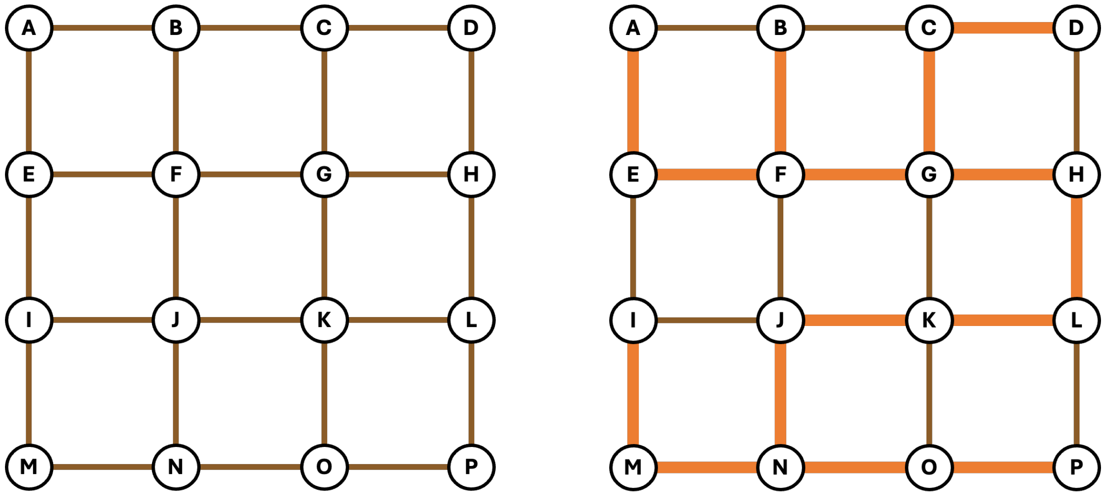
grey relational grade r_{ij}:

$$r_{ij} = \frac{1}{n}\sum_{j=1}^{n} \mathcal{E}_{ij}$$

and the assigned weights w_i:

$$w_i = \frac{r_{ij}}{\sum_{i=1}^{m} \mathcal{E}_{ij}} \times d_i$$

Step 3. Forming the basic train line type and OT stations



(1) The assumed infrastructure network (in brown colour)

(2) The maximum weighted spanning tree (in orange colour)

Figure 4.3 An assumed maximum weighted spanning tree example

Following steps 1 and 2, assuming that a maximum-weight spanning tree of the railway network is generated, as illustrated in Figure 4.3(2), based on the given infrastructure network in Figure 4.3(1). The OT station candidates are A, B, D, I, and P. Further define the railway line A-P as the stem line, and an algorithm process which helps generating the basic train line set is designed as follows:

Table 4.1 Generating basic train line set on a maximum weighted spanning tree network

Step 1: Selecting the longest railway line of the maximum weighted spanning tree

Step 2: Setting the longest railway line as the first basic train line itinerary.

Step 3: Forming new train lines that connect two branches (i.e., a branch pair) near the OT stations and sharing part of the stem itinerary.

Step 4: Repeating Step 3 until no further branch is left.

Taking back Figure 4.3 as the example, the longest railway line of the maximum weighted spanning tree is A-P, and the basic train line set are (Figure 4.4):

- train line l_1: from A to P, is selected after Step 2
- train line l_2: from B to I, via F and N, and
- train line l_3: from D to G and can either extend towards P or I.

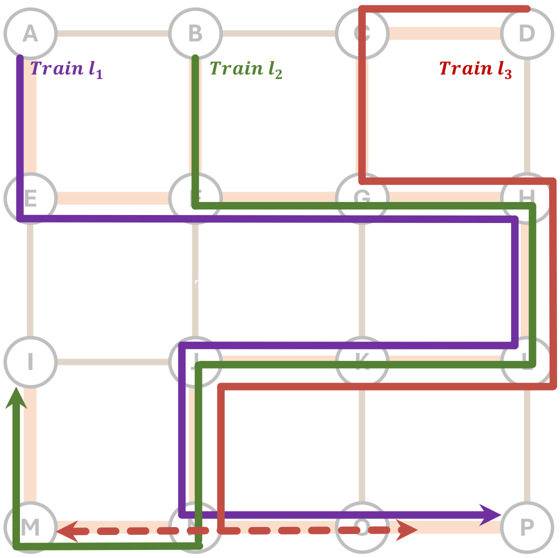


Figure 4.4 the basic train line set for the assumed railway infrastructure network.

4.2.3 Discussion and implication

Essentially, Schiewe (2020) concludes a fundamental principle that the optimal interchange location shall be the point which cover the most overlapped train line itineraries. In terms of methodological differences, Schiewe (2020)'s approach to determining interchange locations relies on predefined train line itineraries, resulting in fixed interchange points. In contrast, within the SILWMST framework, the selection of origin and terminus stations remains flexible, allowing for a more adaptive and optimised approach to line planning.

However, some clear drawbacks of SILWMST can be observed. For instance, there are many OD pairs for which shorter itineraries are available, but no train lines have been generated. The method of Selecting Interchange Locations based on the Weighted Maximum Spanning Tree (SILWMST) tends to cause certain parts of the network (i.e., the stem itinerary) to become overly congested, leading to intense capacity utilisation, while other parts of the network remain underutilised.

From another perspective, this actually aligns with many real-world railway scenarios. As shown in the Beijing-Shanghai HSR line (Appendix A), train density is unevenly distributed across the China high-speed network, with certain railway lines often becoming overloaded and congested. Admittedly, the basic train lines and OT stations produced by SILWMST are not the final choices for real railway operation. However, this method still holds significant implications in several areas:

- SILWMST is able to identify busy or critical parts of the network at an early stage of railway planning. The basic train line set ensures that all stations in the network are connected to major cities, which meet the primal goal of HSR services.

- For the overloaded and congested parts of the network, SILWMST provides an effective way to identify the minimum interchange location, and the basic train line set through a relatively small-scale iteration, which minimises the number of interchange locations while ensuring broad direct service reachability.
- SIWMST suggests that, as a potential for capacity utilisation improvement, congested and overloaded railway lines or network sections should be the focus of optimisation and adjustments, rather than targeting and prioritising cross-line trains.

4.3 Problem description of the line planning method

4.3.1 Fundamental problem: set covering problem

The solution process of train line planning is essentially the Set Covering Problem: given a matrix $A = (a_{ij})_{m\times n}$, where each element $a_{ij} \in \{0,1\}$, and each column corresponds to a cost $c_j (j = 1,2,3,\dots,n)$, with $a_{ij} = 1$ indicating the j th column covers the i th row, the objective is to select a subset of columns that cover all rows while minimising the total cost. An example is given:

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}, c = [10, 12, 8, 9, 30, 25, 20]$$

Selecting the subset of columns where each row is covered at least twice (i.e. the constraint description), the minimum cost solution is $x = [1, 1, 1, 1, 0, 0, 1]$. By defining each column as train line l_j and each row as station s_i, the matrix A is therefore the line pool and $a_{ij} = 1$ represent train l_j stops at s_i. Therefore $x = [1, 1, 1, 1, 0, 0, 1]$ is a line plan solution for this line planning problem. The line planning problem here is restated as deciding the subset from the train line pool such that where each station has at least two trains (i.e. ensuring a minimum frequency). Similarly, the cost c can be the matrix with additional dimensions, and the row can be defined as ODs, railway sections, and other railway operation elements. The modelling perspectives are extended and modified in such manners.

4.3.2 Railway operation scenario

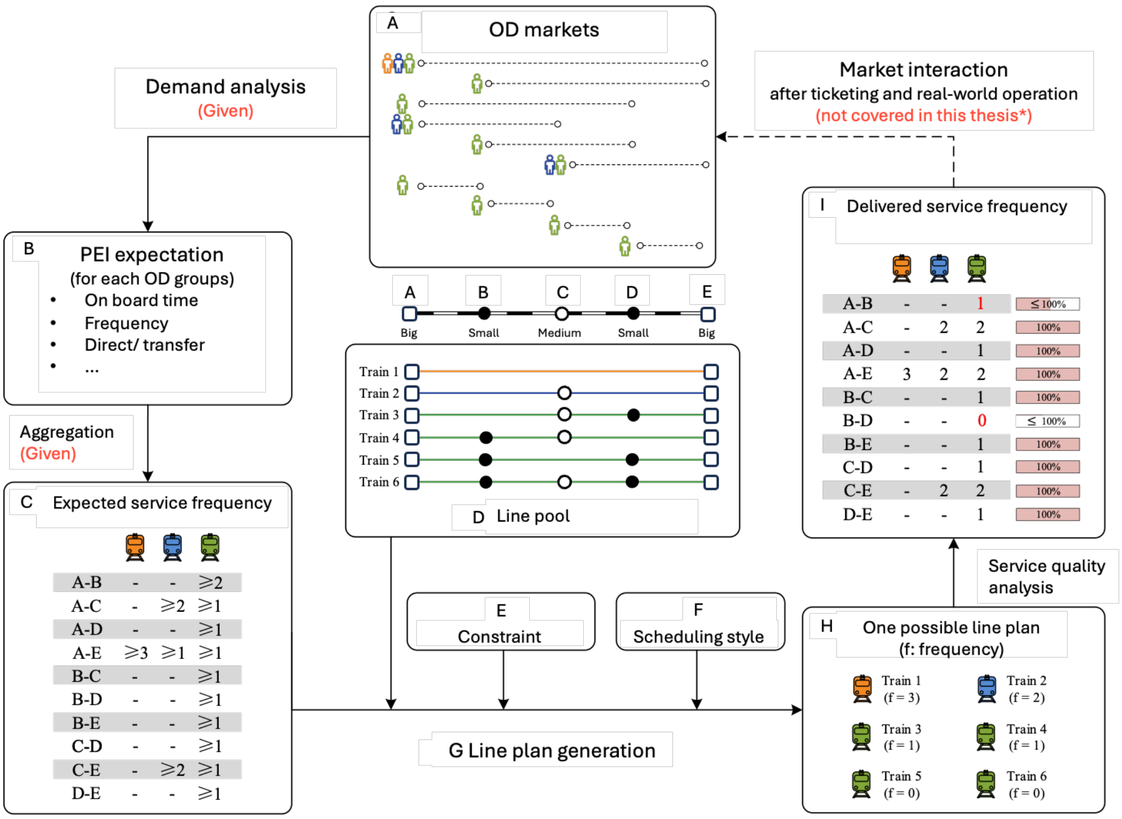


Figure 4.5 The line planning process defined in this thesis

Figure 4.5 illustrates the railway operation scenarios applied in this Chapter. The same OD groups, service expectations vary, with different passengers expecting different types of trains - specifically, orange, blue, and green trains representing F-, S-, and L-trains respectively. This is represented by orange (i.e. F-train), blue (i.e. S-trains), and green (i.e. L-trains) passengers travelling from A to E. The PEIs of OD groups (Step B) are clustered and aggregated to form the expected service frequency list (Step C). Assuming that a proposed scheduling style (Step F) prioritises orange and blue passengers and the capacity upper limit is seven trains, one of possible line plan shown on Step H reveals that the service L-passengers of A-B and B-D cannot be satisfied, while passengers of A-E and C-E receive more train services than the expected frequency (Step I). The reason for this phenomenon is that the service satisfaction constraints can be presented in a non-rigid manner and an iterative approach stops by reaching an approximate solution (Hu et al., 2024). As a degraded alternative, passenger of B-D will take $Train\ 3$ first and then the interchange $Train\ 4$ at Station C. Note that the line planning only shows the spatial relationship regarding transfer.

It is evident that arranging $Train\ 6$ in Step D can provide direct service for every OD, thereby representing the equal transport scheduling style. However, whether the line plan remains sufficiently 'attractive' is left to the railway operators. This thesis proposes aims for aggregating different OD demand features into the requirements for F-, S-, and L-trains. The structure of the model shall enable railway companies to implement flexible and diverse service scheduling styles, rather than arguing that a specific railway context is suited to only one service scheduling style. Additionally, this thesis assumes that Step C is predetermined and does not analyse the specific methodologies for concluding Steps A and B. Consistent research finding suggests that any prediction of OD or train service demand, incorporating elements such as city economic and political rankings, as well as population factors (Li, 2020; Zhang, 2012) must ultimately be converted into service frequency demand as input.

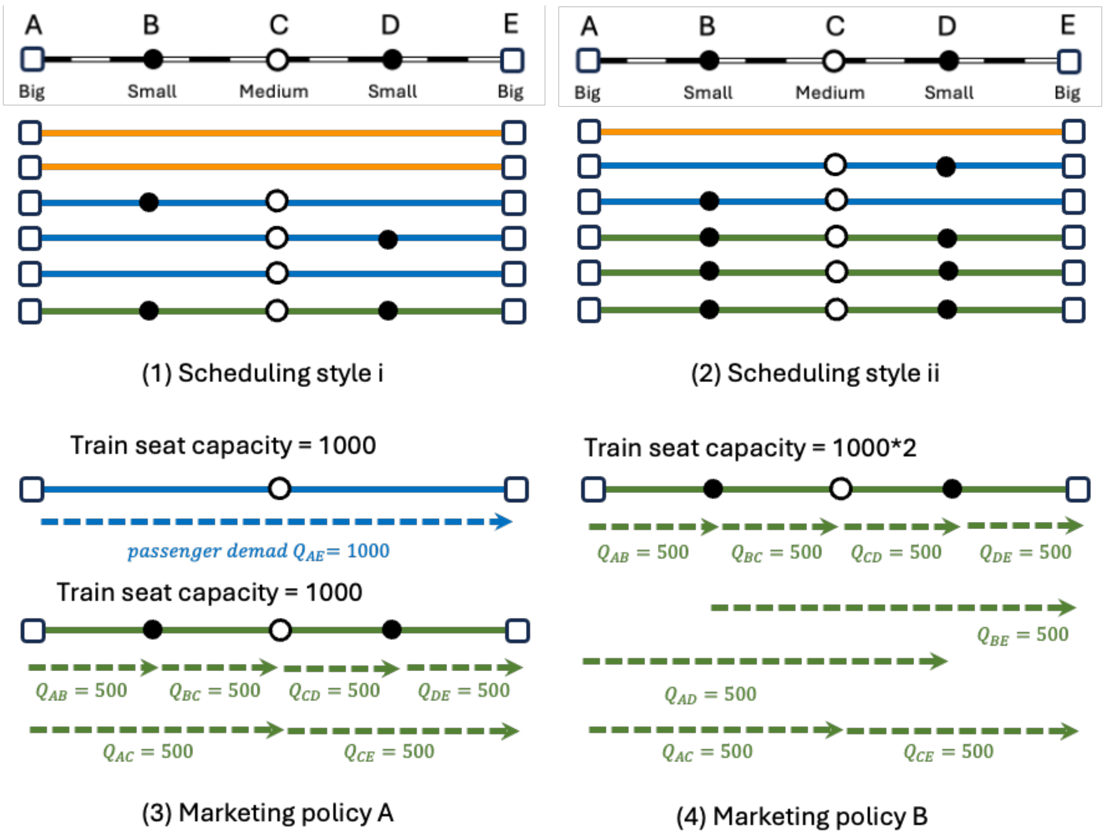


Figure 4.6 Graph difference between scheduling styles and marketing policies

Besides, aggregating PEIs of OD groups as the frequency requirements for different types of trains shows several advantages over line planning which relies solely on passenger assignments (Schmidt & Schöbel, 2024). The effectiveness of line planning with passenger assignments (LPPA), however, depend heavily on the accuracy of OD data and the principles underlying passenger assignment methods (Zhang et al., 2020). In essence, applying different marketing policies to line planning reflects varying priorities in assigning OD groups to train lines (Figure 4.6). For instance, if a railway operator aims to implement a line plan that maximises loading efficiency, certain ODs will be concentrated on a minimal number of train lines (Figure 4.6 (3)). Conversely, if the goal is to create a more 'equal' service network, additional stops will be added, making long-distance travel less attractive (Figure 4.6 (4)). A more evident drawback of LPPA is the potential for travel inequality among ODs: some OD pairs may receive significantly fewer services than high-demand pairs. Conversely, a more 'equal' service network could result in overcrowding during peak hours, as high-demand passengers compete for limited seating capacity.

Comparatively, aggregating PEIs of OD groups as the frequency requirements for different types of trains not only avoid the above drawbacks associated with passenger assignment models but also respect the diverse service requirements of various OD groups (Hu et al., 2023). Presenting as the form of RSI, as previously stated in Section 3.1, the operation of each train line inherently involves multiple categories and attributes of transportation products. From a modelling perspective, this approach constitutes part of the constraints of railway supply modelling, simplifying the general model structure and letting the railway operator evaluate various trade-offs between different line plans without significantly decreasing service quality of OD groups.

4.3.3 Formal problem statement

The line planning problem is defined as a multi-objective set covering problem: Given a predefined line pool and an expected service list, the task is to identify a subset of the line pool that forms a proposed line plan, which satisfies the service constraints either in rigid constraints or non-rigid manner, while minimising the objective functions with acceptable gaps.

4.4 Model formulation

4.4.1 Notation

Table 4.2 Sets and indexes

Set	Description
S	Stations, indexed by $s = 1,2,3,\dots,o,\dots,d,s,s+1,\dots$
S^B	Big stations, $S^{\mathrm{B}} \subset S$
S^M	Medium stations, $S^M \subset S$
S^S	Small stations, $S^S \subset S$
E	Railway sections, indexed by $e = (s, s+1)$
P	Stopping list, indexed by p
U	Station pairs and passenger OD, from o to d, indexed by $u = (o, d)$
U^{st}	ODs of B-B (from big to big) stations, $U^{st} \subset U$

Set	Description
U^{nd}	ODs of B-M, M-M, and B-S stations, $U^{nd} \subset U$
U^{rd}	ODs of M-S and S-S stations, $U^{rd} \subset U$
K	Special ODs that must be provided with train line (s), indexed by k
Q_e	Passenger demand volume at section e
Q_e^{st}	Passenger demand volume of U^{st} at section e, $Q_e^{st} \subset Q_e$
Q_e^{nd}	Passenger demand volume of U^{nd} at section e, $Q_e^{nd} \subset Q_e$
Q_e^{rd}	Passenger demand volume of U^{rd} at section e, $Q_e^{rd} \subset Q_e$
L	Train line candidates in the line pool, indexed by l
L^{Fast}	Fast trains in the line pool, F-trains, $L^{Fast} \subset L$
L^{Semi}	Semi-fast trains in the line pool, S-trains, $L^{Semi} \subset L$
L^{Local}	Local trains in the line pool, L-trains, $L^{Local} \subset L$
L_u	Trains in the line pool that provide direct service to OD u, $L_u \subset L$
L_e	Trains in the line pool that pass the section e, $L_e \subset L$
L_s	Trains in the line pool that stops at the station s, $L_s \subset L$
L_k	Trains in the line pool with specific requirements, $L_k \subset L$
L_u^{fast}	A representative example, all F-train lines can serve OD u, the rest is analogous.

Table 4.3 Fixed parameters of the models (Haitong International Securities Group Limited, 2022)

Notation	Description
$v_{l,short}$	Seat capacity of 8-car train set, $v_{l,short} = 580$ passengers
$v_{l,long}$	Seat capacity of 16-car train set, $v_{l,long} = 1{,}190$ passengers
$\omega_{l,short}^{fixed}$	Fixed cost, $\omega_{l,short}^{fixed} = 63{,}000\ CNY/train$
$\omega_{l,short}^{run}$	Cost of train set lease, electricity, and dispatch $\omega_{l,short}^{run} = 115\ CNY/km$
$\omega_{l,short}^{stop}$	Service charges (i.e. water refill, trash drop-off, security) $\omega_{l,short}^{stop} = 200\ CNY/stop$

Notation	Description
$\omega_{l,long}^{fixed}$	Fixed cost, $\omega_{l,long}^{fixed} = 84{,}000\ CNY/train$
$\omega_{l,long}^{run}$	$\omega_{l,long}^{\text{run}} = 1.5 \times \omega_{l,short}^{\text{run}} = 172.5\ CNY/km$
$\omega_{l,long}^{stop}$	$\omega_{l,long}^{\text{run}} = 1.5 \times \omega_{l,short}^{stop} = 300\ CNY/stop$
d_e	Distance mileage of section e, decided by the data input
d_l	Total operating distance of train line l, decided by the train line
t_u	In-vehicle time (i.e. station to station) of OD u, calculated by mileages
Δt_u	Expected additional time cost when stopping at a station, $\Delta t_u = 7min/stop$.
n_l	Total stop number limit of train line l , decided by the train line l
n_l^u	Total intermediate stop number of OD u when take train line l

Table 4.4 Adjustable parameters for sensitivity analysis

Notation	Description
ξ_e^{Fast}	The proportion of F-trains on section e
ξ_e^{Semi}	The proportion of S-trains on section e
ξ_e^{Local}	The proportion of L-trains on section e
N_e	Upper-bound of train amounts (i.e. capacity) on section e, positive integer
F_u^{Fast}	Expected frequency for passenger OD u serviced by F-trains, natural number
F_u^{Semi}	Expected frequency for passenger OD u serviced by S-trains, natural number
F_u^{Local}	Expected frequency for passenger OD u serviced by L-trains, natural number
ΔF_u^{Fast}	Frequency change, representing the multiple service supply, positive integer
ΔF_u^{Semi}	Frequency change, representing the multiple service supply, positive integer

Table 4.5 Decision variables

Notation	Description
x_l	0-1 variables, =1 if train line l is selected, =0 otherwise

Notation	Description
f_l	service frequency of train line l
$y_{l,s}$	0-1 variables, = 1 if train line l stops at station s

4.4.2 Objective functions

This thesis mainly considering the following objectives:

- The objective function (4-1) aims to minimise the operation cost of a line plan.

$$min\, Z_1 = \sum_{l\in L}(\omega_l^{fixed} + \omega_l^{stop} \cdot n_l + \omega_l^{run} \cdot d) \cdot x_l \cdot f_l \quad (4\text{-}1)$$

- The objective function (4-2) aims to minimise the expected additional travel time cost incurred due to the intermediate stops during the travel of OD u.

$$min\, Z_2 = \sum_{u\in U}\sum_{l\in L_u} n_l^u \cdot \Delta t_u \cdot x_l \cdot f_l \quad (4\text{-}2)$$

4.4.3 The basic model structure

A simple line planning model constraints usually cover the following aspects: frequency, seat capacity, railway section capacity, and delivering customised services.

$$\sum_{l\in L_u} x_l \cdot f_l \geq F_u \;\; \forall u \in U \quad (4\text{-}3)$$

$$\sum_{l\in L_e} x_l \cdot f_l \cdot v_l \geq Q_e \;\; \forall e \in E \quad (4\text{-}4)$$

$$\sum_{l\in L_e} x_l \cdot f_l \leq N_e \;\; \forall e \in E \quad (4\text{-}5)$$

$$\sum_{l\in L_k} x_l \cdot f_l \geq 1 \quad (4\text{-}6)$$

Constraint (4-3) makes sure the service frequency can be satisfied. All passengers can be guaranteed with direct services by setting $F_u \geq 1$. Constraint (4-4) is the seat capacity restriction. Constraint (4-5) imposes upper-bound train amount limits on every section. Constraint (4-6) represents customised services, such as compulsory service links for less-developed areas, can be generated.

4.4.4 The modified model structure (M2)

The conventional line planning model does not represent the match between demands of clustered passenger groups and multiple levels of services. Therefore, this thesis is hereby improving constraint (4-3), (4-4), and (4-5) accordingly.

$$\sum_{l \in L_u^{Fast}} x_l \geq F_u^{Fast} - \Delta F_u^{Fast} \quad \forall u \in U^{st} \tag{4-7}$$

$$\sum_{l \in L_u^{Semi}} x_l \geq F_u^{Semi} - \Delta F_u^{Semi} + \Delta F_u^{Fast} \quad \forall u \in U^{nd} \tag{4-8}$$

$$\sum_{l \in L_u^{Local}} x_l \geq F_u^{Local} + \Delta F_u^{Semi} \quad \forall u \in U^{rd} \tag{4-9}$$

Constraint (4-3) is extended and modified into constraints (4-7), (4-8), and (4-9). Constraint (4-7) regulates that the expected F-train frequencies that serve U^{st} group, with any unsatisfied frequencies (ΔF_u) potentially served by the lower-level semi-fast trains, as indicated in Constraint (4-8). Constraints (4-8) and (4-9) follow the same logic explaining S-train and L-train frequency.

$$\sum_{l \in L_e^{Fast}} x_l \cdot v_l \geq Q_e^{st} - \Delta Q_e^{st} \quad \forall e \in E \tag{4-10}$$

$$\sum_{l \in L_e^{Semi}} x_l \cdot v_l \geq Q_e^{rd} - \Delta Q_e^{rd} + \Delta Q_e^{st} \quad \forall e \in E \tag{4-11}$$

$$\sum_{l \in L_e^{Semi}} x_l \cdot v_l \geq Q_e^{nd} - \Delta Q_e^{nd} + \Delta Q_e^{rd} \quad \forall e \in E \tag{4-12}$$

Accordingly, Constraint (4-4) is extended and modified into constraints (4-10), (4-11), and (4-12), and the total amounts of unsatisfied passenger groups ($\Delta Q_e^{st}, \Delta Q_e^{rd}, \Delta Q_e^{nd}$) either potentially served by the lower-level trains (ΔQ_e^{st} in Constraint (4-11), and ΔQ_e^{rd} in Constraint (4-12)) or take a focused transfers (ΔQ_e^{nd}in Constraint (4-12)).

$$\sum_{l \in L_e^{Fast}} x_l \leq \left\lceil \xi_e^{fast} \cdot N_e \right\rceil \forall e \in E \tag{4-13}$$

$$\sum_{l \in L_e^{Semi}} x_l \leq \left\lceil \xi_e^{Semi} \cdot N_e \right\rceil \quad \forall e \in E \tag{4-14}$$

$$\sum_{l \in L_e^{Local}} x_l \leq \left\lceil \xi_e^{Local} \cdot N_e \right\rceil \quad \forall e \in E \tag{4-15}$$

Next, constraint (4-6) is extended and modified into constraints (4-13), (4-14), and (4-15) collectively, which regulate the proportions of F-trains, S-trains, and L-trains in the line plan.

$$\xi_e^{fast} + \xi_e^{semi} + \xi_e^{Local} = 1 \tag{4-16}$$

$$U_e = U_e^{st} \cup U_e^{nd} \cup U_e^{rd} \quad \forall e \in E \tag{4-17}$$

$$v_l = v_{l,long} \vee v_{l,short} \tag{4-18}$$

$$\omega_l^{fixed} = \omega_{l,long}^{fixed} \vee \omega_{l,short}^{fixed} \tag{4-19}$$

$$\omega_l^{run} = \omega_{l,long}^{run} \vee \omega_{l,short}^{run} \tag{4-20}$$

$$\omega_l^{stop} = \omega_{l,long}^{stop} \vee \omega_{l,short}^{stop} \tag{4-21}$$

$$n_l^u = \sum_{s=o}^{d} y_{l,s} - 2 \quad \forall u \in U \tag{4-22}$$

$$x_l \leq \sum_{s=1}^{d} y_{l,s} \quad \forall s \in S, \forall l \in L \tag{4-23}$$

Constraints (4-16) - (4-23) are all logic constraints. Constraints (4-16) is the proportion equation among F-trains, S-trains, and L-trains. Constraint (4-17) calculate the leftover passenger demands. Constraint (4-18), (4-19), (4-20), and (4-21) are the train set selection (8-car or 16-car) equations. Constraint (4-22) is the calculation method of intermediate stop for an OD u. Constraint (4-23) establishes the logic connection between train line l and its stop behaviour.

4.4.5 Rationale for the modified model structure

1. How the proposed model manages the railway capacity and general service quality.

In the series of work by Zhang et al. (2020) and Zhang & Nie (2016), they testified that that a line plan with the requested lower capacity occupation value is based on achieving the specified service level with minimum possible number of total stops. However, the objective function (4-2) can also achieve the same goal by minimising the additional travel time. A shorter travel time means that train lines with a smaller number of stops will be selected, so that most ODs will have fewer intermediate stops.

2. How the proposed passenger group classification differs with other studies.

In previous studies (Hu et al., 2023; Li, 2015; Zhang et al., 2020), passenger groups have been classified into B-related, M-related, and S-related categories, where combinations like M-S and B-S are considered lower-level groups. However, when checking Table 3.2, the passenger of B-M and B-S is underestimated. Re-classifying passenger groups in this thesis can enhance the travel opportunities for passengers specially from S-stations by providing faster HSR services to B-stations and increasing the potential for interchanges at B-stations.

3. How the proposed model manages the structure of generated line plan.

Section 4.3.4 is the major extended parts of the line planning models. For a perspective of general line planning compositions, the proportion of F-trains, S-trains, and L-trains can be customised as numerical experiments. Moreover, the selective constraint (4-6) helps in generating train lines that may not necessarily be profitable, but they hold significant importance for social responsibility and improving 'equality', such as for providing HSR links to less developed areas.

4. How the proposed model reflects OD groups and train preference

From the constraints (4-17), (4-8), and (4-9) regulate the train preference for diverse types of passengers, passenger from $u \in U^{st}$ prefer F-trains and S-trains, while $\in U^{nd} \cup U^{rd}$ are arranged by S-trains and L-trains.

5. How the proposed model manages direct service or next-level service

The constraints do not directly control the level of direct service provision for different passenger OD flow groups; instead, this is managed by the parameters F_u^{Fast}, F_u^{Semi}, and F_u^{Local} in Constraints (4-7), (4-8), and (4-9). As long as the values of equations like F_u and $F_u - \Delta F_u$ are not zero (Hu et al., 2023, 2024), OD u are assured with direct services. Based on the percentage trend of OD groups in China HSR, there is a clear trend that M-S or S-S passengers might receive less direct services.

6. How the proposed model manages focused transfer travels.

This model does not precisely reflect all passenger transfer behaviours. However, Constraints (4-8) ensure that mandatory transfer travels will have direct trains to Big stations where multiple interchange trains are available. The exact transfer times will need to be confirmed during the timetable generation stage. At line planning stage, only interchange location is regulated at Big stations, supported by high frequency of services.

7. How the proposed model manages the line planning at the network levels.

Although certain efforts have been made regarding network line planning (Han et al., 2019), network line planning does not exhibit any essential or fundamental differences compared to railway line-level planning. The solution for network line planning will be divided into stages, with line plans developed for individual railway lines. However, during the timetable generation stage, infrastructure and headway conflicts, as well as constraints, will be considered, making network timetable generation more challenging than line-level timetable generation.

8. How the proposed model manages periodic or non-periodic line planning.

The differences between non-periodic and periodic line planning are relatively minor in most cases. Generally, periodic planning only requires determining the train lines for a one- or two-hour interval, setting the capacity limit ($N_e = \frac{60\ or\ 120\ min}{min\ headway}$) based on the system's theoretical minimum headway. As described in Chapter 3, the challenge in periodic line planning lies in adopting limited stop list combinations to ensure a sufficiently large scale of direct services. Typically, railway companies prioritise stop-by-stop trains to ensure direct transport. In short, the proposed model can be directly adopted in periodic operation scenarios.

9. How the proposed model be applied in regional HSR line planning.

The difference between regional HSR line planning and arterial HSR line planning is that the stopping pattern in regional HSR is much simpler, with the stop list showing far fewer variations compared to arterial HSR line plans. Passenger demand in regional HSR lines exhibits clear fluctuations, particularly during rush hours on weekdays, weekends, and other holidays. As a result, there will be temporary stop-skipping operations during rush hours, similar to those in underground and urban railways. The proposed model can still be applied to regional HSR, as the train line candidates generated from the line plan will adhere to constraints (4-10), (4-11), and (4-12). This will result in some train lines not stopping at certain stations, effectively performing 'stop-skipping'.

10. The relationship between proposed model and ticketing.

The complexity of passenger assignment shares the same logic with ticketing process, which is explained and argued in Section 4.2. It should be remarked that the constraint (4-4), (4-10), (4-11), and (4-12) do not represent the final ticket allocation result. The primary role of Constraint (4-4) to complement constraint (4-3) in determining the required number of trains and the correct train type (16-car or 8-car). This ensures that the total seating capacity of all trains meets the passenger demand at each railway section. Constraint (4-10), (4-11), and (4-12) plays similar roles. In a non-reservation seating environment, constraint (4-4), (4-10), (4-11), and (4-12) can either be ignored or applied with a coefficient exceeding 100% to account for potential overloading scenarios.

4.5 Solution approach

4.5.1 Discussion between business solver and heuristic algorithm

In the efficiency debate between business solvers and heuristic algorithms, it is widely recognised that business solvers (e.g., Gurobi, Cplex) are highly effective for small to medium-sized optimisation problems, but their performance significantly diminishes when applied to large-scale models. This is primarily due to the exponential increase in computational complexity as the problem size grows, leading to prohibitively long solution times or an inability to reach a feasible solution within practical time limits. (Hu et al., 2023, 2024) already proved when the train line scale $L \geq 200{,}000$ commercial solvers fail to return a solution within five hours. Such a scale is easily reached; for example, a railway line with 23 stations has a total combination of $C_{23}^{7} = \binom{23}{7} = 245{,}157$ possible train routes that only stop at 7 stations.

In particular, for research problems like line planning, which heavily depend on the interpretation of railway scenarios, the internal processes of business solvers acting like grey box cannot be intervened. Solvers treat every variable and every enumeration equally, even when some enumerations are ineffective. By designing a rule-based heuristic algorithm, it is possible to selectively narrow the scope of enumeration while effectively controlling the acceptable gaps.

4.5.2 A rule-based heuristic algorithm

The model structure implicitly includes two features that need to be incorporated into the solution process: the first is to ensure that train line do not have unnecessary stops, and the second is to ensure that passenger loading is optimised to minimise leftover passengers. Based on this, the solution process has been designed in two stages.

Table 4.6 The outline of the line planning solution

Step 1: Input data

- OD data of passenger groups U , Beijing-Shanghai HSR section data E,
- Interpretating scheduling styles as $F_u^{Fast}, F_u^{Semi}, F_u^{Local}$
- Setting given $\Delta F_u^{Fast}, \Delta F_u^{Semi}$, N_e

Step 2: Generate initial train lines

- Afflicted elements are also decided: $v_l, \omega_l^{fixed}, \omega_l^{run}, \omega_l^{stop}$ (constraint (4-18)- (4-21))

Step 3: Optimise the itinerary to minimise stops

- **Step 3.1: Ensure coverage:**

  Ensure that each medium and small station is connected to at least a set number of trains that links it to the next Big station in the route by checking

  - $x_l \leq \sum_{s=1}^{d} y_{l,s}$
  - $\sum_{l \in L_e^{Fast}} x_l \cdot v_l \geq Q_e^{st} - \Delta Q_e^{st}$
  - $\sum_{l \in L_e^{Semi}} x_l \cdot v_l \geq Q_e^{rd} - \Delta Q_e^{rd} + \Delta Q_e^{st}$
  - $\sum_{l \in L_e^{Semi}} x_l \cdot v_l \geq Q_e^{nd} - \Delta Q_e^{nd} + \Delta Q_e^{rd}$

- **Step 3.2: Remove unnecessary stops, for each train:**
  - Start with an initial itinerary that includes only Big stations.
  - **Step 3.2.1:** Add Medium or Small stations to the itinerary only if:
    - The station is necessary to ensure connectivity between itinerary section.
    - The station serves a significant demand by ranking $Q_e^{st}, Q_e^{nd}, Q_e^{rd}$.
    - The frequency change shall be smaller than ΔF_u^{Fast}

- **Step 3.3: Sort the itinerary**

Ensure the itinerary is sorted in the correct sequence.

Step 4: Calculate total intermediate stops and cost

- **Step 4.1:** For each generated and optimised itinerary $\forall u \in U$, calculate the number of all intermediate stops for all ODs:

$$n_l^u = \sum_{s=o}^{d} y_{l,s} - 2.$$

- **Step 4.2:** Sum up the total number of intermediate stops across all trains:

$$Z_2 = \sum_{u \in U} \sum_{l \in L_u} n_l^u \cdot \Delta t_u \cdot x_l \cdot f_l$$

- **Step 4.3:** calculating the total cost of all train lines:

$$Z_1 = \sum_{l \in L} (\omega_l^{fixed} + \omega_l^{stop} \cdot n_l + \omega_l^{run} \cdot d\,) \cdot x_l \cdot f_l$$

Step 5: Evaluate and refine

- **Step 5.1:** Evaluate whether the total number of intermediate stops meets the desired minimum, if not, head to Step 5.2.
- **Step 5.2:** further by iterating through **Steps 3 and 4**, removing any additional unnecessary stops while ensuring coverage, updating Z_1, Z_2 and n_l^u.

Step 6: Finalise itinerary

- Return the optimised train itineraries with minimised intermediate stops and confirmed coverage of all required stations and exact Z_1, Z_2 and n_l^u.
- End the itineration when the cost reduction gap is less than 15% or the iteration step reaches 50.

4.6 Case study

4.6.1 Background information

The China Beijing-Shanghai HSR line, also known as the Jinghu HSR line, is one of the busiest HSR lines in the world. The stations, railway sections, and OD are ordered according to the direction from Beijing to Shanghai. Therefore, BJN is indexed as s_1, SHHQ as s_{23}, and the set of station $S = \{s_1, s_2, \ldots, s_{23}\}$. The BJN-LF section is indexed as e_1, thus the set of railway section $E =$

$\{e_1, e_2, \ldots, e_{22}\}$. The O-D pairs are ordered as follows: BJN-LF (u_1), BJN-TJN (u_2), ...BJN-SHHQ (u_{21}), ..., KSN-SHHQ (u_{253}), and the set of ODs $U = \{u_1, u_2, \ldots, u_{253}\}$.

An average train density of Beijing-Shanghai HSR line is over 120 trains per day, and it accommodate trains from all other provincial capitals only except Hangzhou (Zhejiang), Nanchang (Jiangxi province), Changsha (Hunan province), Guiyang (Guizhou province), Kunming (Yunnan province), Guangzhou (Guangdong province & the Great Bay area), and Fuzhou (Fujian province). The busiest section, between Bengbu (s_{13}) and Xuzhou (s_{11}), sees over 160 trains daily (Li & Pi, 2024). The Jinghu HSR line spans 1,318 km, includes 23 stations and 22 railway sections, serves 253 ODs (original-line) and over 160,000 passengers daily in 2019 (Haitong International Securities Group Limited, 2022), with a maximum train speed of 350 km/h. The OD classifications are: B-B (Big to Big stations): 15 OD pairs, B-M (Big to Medium stations): 66 OD pairs, B-S (Big to Small stations): 36 OD pairs, M-M (Medium to Medium stations): 55 OD pairs, M-S (Medium to Small stations): 66 OD pairs, and S-S (Small to Small stations): 15 OD pairs.

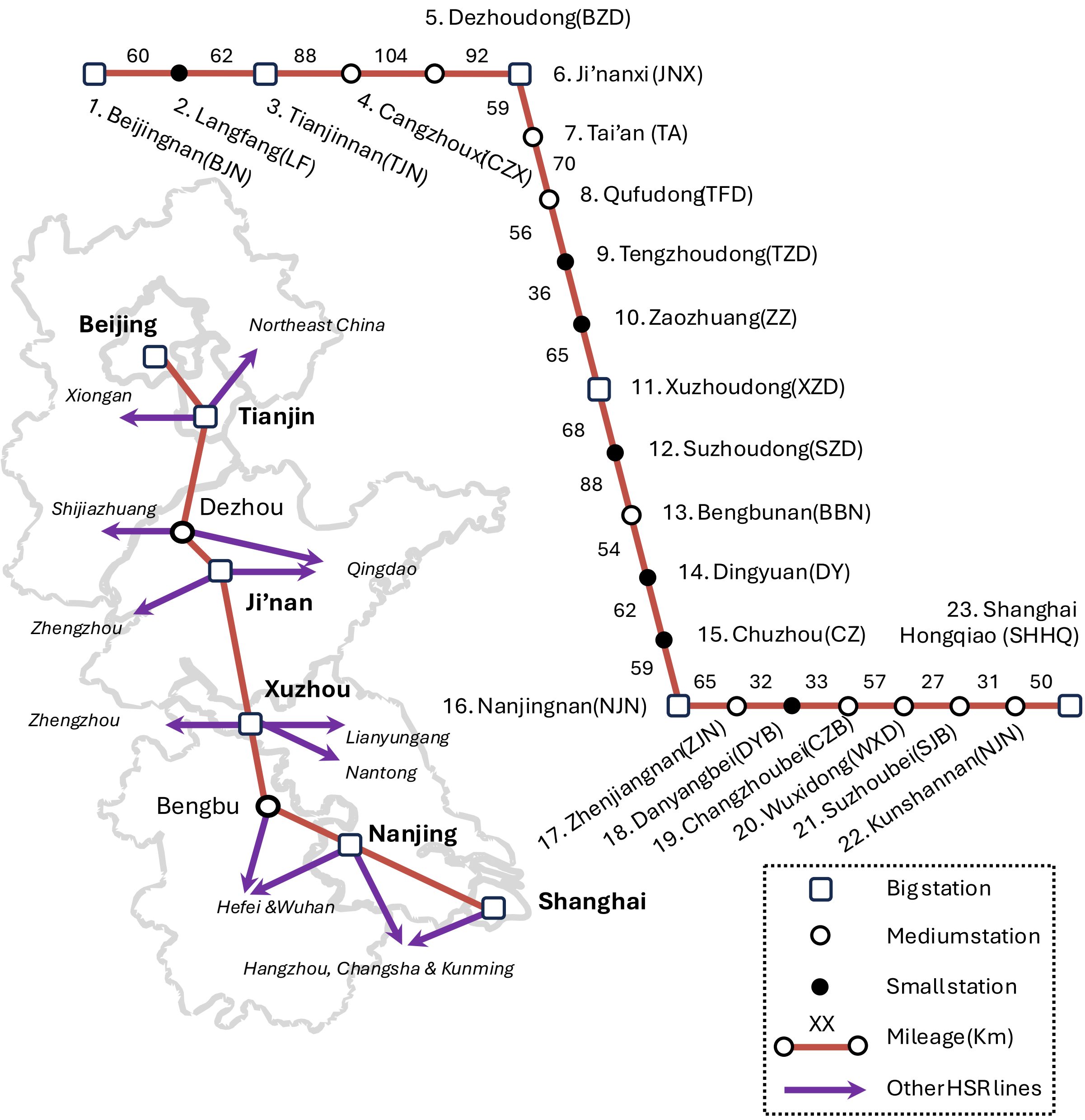


Figure 4.7 Beijing-Shanghai HSR line and other HSR lines

It should be noted that there are three regional HSR areas which play significant roles in short-distance travel: the Beijing-Tianjin intercity, the Shanghai-Nanjing-Hangzhou intercity, and the Shandong provincial circle. This suggests that passengers may prefer regional HSR for certain origin-destination (OD) pairs. These travels typically occur in section between Beijing and Tianjin ($s_1 \sim s_3$), Dezhou and Zaozhuang ($s_5 \sim s_{10}$), and Nanjing and Shanghai ($s_{16} \sim s_{23}$).

Regarding classifying OD groups based on travel distance, Section 1.2 summarises the definitions of long-distance railway passengers across countries. However, none of the existing definitions fully reflect travel features in China HSR. Statistics only indicate a relatively clear feature that travel distances within 350 km exhibit minimal variations in travel time (i.e., within 45 minutes) compared to other ODs, regardless of the train's stopping patterns. Consequently, this thesis defines 350 km as the boundary between short-distance and medium-to-long-distance travel. Accordingly, the minimum travel time reference for short-distance travel is 90 minutes. As representative examples, every adjacent Big station on Beijing-Shanghai HSR belongs to short-distances.

4.6.2 Scenario pre-processing and rules for line pool

1. General passenger behaviours.

The demand of F-passengers, S-passengers, and L-passengers are the result of known market analysis and assumed as planned rail travels, which implies that passengers are not expected to switch to other transport modes, regardless of variations in the proposed line plans. Ticket prices rank from high to low is F-trains, S-trains, and L-trains and from short travel time to long travel time (Pengpai News, 2024).

2. Inferred and assumed original-line OD data of Beijing-Shanghai HSR line.

Ma (2018) and Tian (2018) provide detailed data on OD flows between all Big stations of Beijing-Shanghai HSR lines, together with Table 3.2. The appendix of (Y. Li, 2020) contains passenger flow data for a specific hour on the Beijing-Shanghai HSR line. By synthesising the information from these three sources, a comprehensive OD dataset can be constructed, accounting for potential systematic errors. The data is rounded up to the nearest multiple of 10. After aggregating the OD data, the results align with the overall passenger flow characteristics of the original-line travels on Beijing-Shanghai HSR line after cross verifications (Haitong International Securities Group Limited, 2022; Li, 2020; Ma, 2018; Tian, 2018). The case study only takes the direction from Beijing to Shanghai direction as the example.

3. HSR train line category and line plan scale: 70 trains

The HSR trains are categorised as F-trains, S-trains, and L-trains where the stopping patterns are explained in Section 1.2.2.: The F-train stops at most or all Big stations, S-train stops at Big stations and a subset of Medium Stations, and L-Train stops at the Big stations and subset of Medium and small stations. The total stop number of F-trains, S-trains, and L-trains are not fixed and will be tested in the following cases. Note that only Big station is allowed to set as Origin or Terminus stations. The long-type and short-types accounts for 80% and 20%, respectively.

4. ODs from other railway lines: cross-line operation.

According to the statistics provided by (Xu, 2020), passengers who travel from other HSR lines account for 20%-50% of passenger loading on cross-line trains, which implies that once cross-line trains running on the Beijing-Shanghai HSR line, the proportion of original-line travels served by these cross-line trains ranges between 50% and 80%. Besides, the majority of cross-line travel destinations are Big stations on Beijing-Shanghai HSR line. The related demands are aggregated as setting node stations as destinations. For example, passengers travel from Qingdao is regarded as the demand associated with Ji'nanxi (s_6).

Due to the track layout design of Beijing-Shanghai HSR, there is no opposing direction of cross-line movements. For each node station, considering the direction from Beijing to Shanghai and track layouts, the available cross-line train directions are listed as follows:

- Tianjin nan: accepting trains from Northeast China to Shanghai
- Dezhou dong: accepting trains leaving Jinghu HSR line for Qingdao, from Shijiazhuang to Shanghai.
- Ji'nan xi: accepting trains from Qingdao to Shanghai.
- Xuzhou dong: accepting trains from Zhengzhou to Shanghai, leaving Jinghu HSR for Lianyungang.
- Bengbu nan: accepting trains leaving for Hefei.
- Nanjing nan: accepting trains leaving for Hangzhou, from Hefei.

5. Applying SILWMST on Beijing-Shanghai HSR line.

Combining the assumed OD data, the scenario of Beijing-Shanghai HSR line suits the SILWMST. The case study only set distance as the weight of spanning tree, and the OT stations of cross-line trains, which centralised on the node station here, are determined as follows:

- From North-East China to Hefei, and Hangzhou: Tianjin nan (O), Nanjing Nan (T)
- From Shijiazhuang & Qingdao to Hefei, and Hangzhou: Ji, nan (O), Xuzhou (T), Nanjing Nan (T)
- From Zhengzhou to Hefei and Hangzhou: Xuzhou Dong (O), Nanjing Nan (T)

4.6.3 Scheduling styles and modelling results

The testing scheduling styles are listed as follows:

- Scheduling style A: different maximum stop number of S-trains and L-trains.
- Scheduling style B: different maximum stop number gap between S-trains and L-trains.
- Scheduling style C: different maximum stop number gap between S-trains and L-trains between each B-B railway section.
- Scheduling style D: different percentage of F-trains, S-trains, and L-trains.

Challenges 1.2.3.2 and 1.2.3.3 are specifically represented in the line plan structure as Schedules A, B, and C. Combing Beijing-Shanghai HSR contexts and scheduling style A, B, C help understanding the potential trends or relationships between reachability, timetable train traffic, and railway capacity. After analysing the effects of Scheduling style of A, B, and C. A more comprehensive case result based on Scheduling style D is proceed. Firstly, for a single train, the setting of stop numbers will lead to discussions around coverage (reachability) and minimum train types required. From the perspective of multiple trains, if the gap in stop numbers is too large, the timetable generation phase will see an increase in overtaking behaviour, further prolonging the overall travel time for trains with more stops. Furthermore, from a ticketing perspective, this could result in some trains becoming overcrowded while others are less attractive to passengers. Similarly, the stop number gap could be observed from a broader railway section perspective. Interestingly, it should be notice that if arranging UK-style line plans: trains sectionally terminus at s_3, s_6, s_{11}, s_{16}, and s_{23}. The discussion of style C could be removed.

To begin with, it is reasonable for excluding the F-trains when discussing Style A, B, C, which is obvious because there are limited stop list variations if train only dwelling at Big stations. Next, to build up a foundation of discussion, the goal of scheduling style A is to testify the minimal numbers of S-trains and L-trains to ensure the minimum direct service to its next Big stations. The result shows as Table 4.6. It can be concluded that the minimum number of train types required to provide direct service to Big stations decreases significantly as the number of stops increases. In the case of L-trains, the extent of direct service to Big stations is determined by the smaller number of stops allocated to M-stations and S-stations.

Upon further analysis, the largest stop number gap between S-trains and L-trains is seven stations, while the smallest gap could be zero (Scheduling style B). The advantage of having larger station number gaps is that B-M passengers could experience significantly shorter travel times by taking S-trains, with at least 49 minutes saved. However, this approach results in more heterogeneous train traffic, as at least 10 S-trains would be required to achieve minimal reachability. Conversely, a smaller stop number gap reduces the difference between S-trains and L-trains, resulting in more balanced and equal travel times for both B-M and B-S passengers, and increase the direct service opportunities of M-M, M-S, and S-S travels.

Table 4.7 The minimal train line types to support B-M direct service.

Minimal types of S-trains		**Minimal types of L-trains**		
Stop number	**B-M direct service**	**Stop number**	**B-S direct service**	**B-M & B-S direct service**
6	10	9	7	$Max\{10, 7, 5, 4, 3\}$
8	5	11	4	
10	4	13	3	

Examining Sections $s_2 \sim s_8$, $s_8 \sim s_{16}, s_{16} \sim s_{23}$, it is obvious that M-stations and S-stations are successively distributed on Beijing-Shanghai HSR line, respectively, rather than distributed in a staggered manner (Figure 4.8). The implication for railway operations is that reducing the stop number gap between S-trains and L-trains increases the likelihood of generating an 'evenly-distributed' stop list combination of line plan (Scheduling style C), which in turn leads to more homogeneous train traffic and a lower capacity utilisation value.

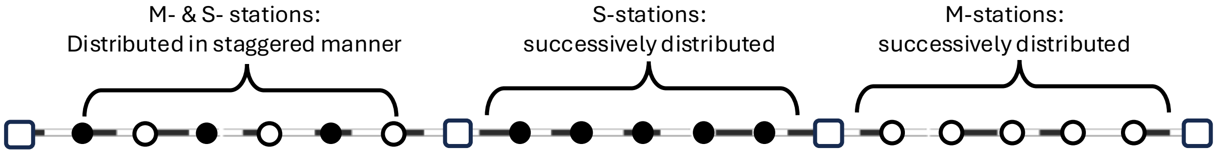


Figure 4.8 Distribution pattern of M-station and S-station

Based on the concluded observations above, the minimum stop number for testing Scheduling style D is three for F-trains, while eight for S-trains, and ten for L-trains. Besides, Scheduling style D examines only 50% of the OD data, with the line plan designed to accommodate approximately 40% of the daily demand density (Peng et al., 2024), equivalent to 70 trains. The intention for this is using a minimal stop list combination to meet relatively higher OD demand volumes—if the scale of the line plan were sufficiently large, train lines would naturally be expected to meet all demands. The format of table cell explains as: $\xi_e^{fast} + \xi_e^{semi} + \xi_e^{local} = 1$. The experiments of Scheduling style D further define two conditions as the details of constraint (4-9):

- Condition 1: for each M-station and S-stations, setting at least five trains linking to its next Big station along the itinerary, and reduce the unnecessary intermediate stops.
- Condition 2: for each M-station and S-stations, setting at least ten trains linking to its next Big station along the itinerary, and reduce the unnecessary intermediate stops.

The results are shown as Table 4.7 and 4.8. The implication of table items is explained as following: The total cost represents the operational costs of railway transport, corresponding to objective (4-1). The total intermediate stop and time cost refers to the total number of intermediate stops that all OD services provided by all trains must pass through, reflecting objective (4-2), which concerns as one type of 'efficiency' (i.e. fast travel) of all OD travels. Clearly, the fewer the intermediate stops, the more efficient the travels. For different station pairs, the table introduces four new concepts, listed from top to bottom: total direct service, average frequency, average intermediate stops, and average tickets assigned to each OD.

Total direct service reflects the overall railway supply of transportation products. Average frequency indicates the extent to which services are evenly distributed throughout the day for each OD. Average intermediate stops reflect the efficiency of specific OD journeys, similar to the previous definition. Lastly, the average tickets assigned to each OD is calculated by dividing the demand volume for that station pair by the total direct services, indicating the 'demand intensity' or 'ticketing pressure' that each train line must 'endure'. Since the proposed model aims to remove unnecessary and redundant stops, the total railway supply for each OD is reduced, thereby increasing the 'pressure' on each train. The implication for railway operators is that if this value is high, it suggests the need to increase railway supply: representing as adding more stops or train lines. This thesis defines this effect as 'ticketing pressure'.

As general trends, requiring more compulsory links contributes to an increase in the total number of direct services. Second, the trends for various station pairs generally follow a proportional pattern across F-trains, S-trains, and L-trains. Thirdly, there is a less intermediate stop variations across these six style cases, which could be attribute to the layout of the case study.

To be more specific, The RSI performance of B-B stations significantly outperforms other station pairs across all items. This is due to the stopping rules set for F-trains, S-trains, and L-trains, where almost all trains are required to stop at Big stations. However, the calculation results align with the provided OD data. Despite the ample railway supply, each train still faces significant ticketing pressure, with requirements ranging from 48 to 53 tickets for each OD pair. This represents the second-highest demand among all station pairs. The overall supply for B-M is similar in scale to B-B. However, due to the higher number of B-M combinations, the average frequency calculation result is lower. Nevertheless, ticketing pressure is the least compared to B-B and B-S, indicating that the line planning results are more favourable for both railway operators and passengers.

Regarding B-S travels, the overall supply for B-S is significantly lower than B-B and B-M, leading to the lowest average frequency calculation result and the highest ticket pressure. The average travel of S-S travels receives fewer intermediate stops, which is the least among all OD groups. Comparatively, this figure for M-S travels is the highest. This result can account for the layout of S- and M- stations on the Beijing-Shanghai HSR line rather than the optimisation results. Due to the minimal ticket pressure on S-S pairs, they can be deprioritised during the ticketing process. As a commonly shared feature, the reachability of ticketing pressure of M-M, M-S, and S-S are all low. A fair compensation from the railway operation reality is that the intercity regional HSR provide high-frequency of services serving these M-M, M-S, and S-S pairs.

Regarding the long-type and short-type trains, long-type trains are clearly favoured, constituting the majority of the line plan cases. This preference arises from the fact that, while incurring only 1.5 times the cost, long-type trains provide 2.05 times the total seating capacity. However, a potential drawback of this arrangement lies in the limited flexibility for joining and splitting operations (Section 3.2.3), as the platforms can accommodate a maximum of only 18 carriages.

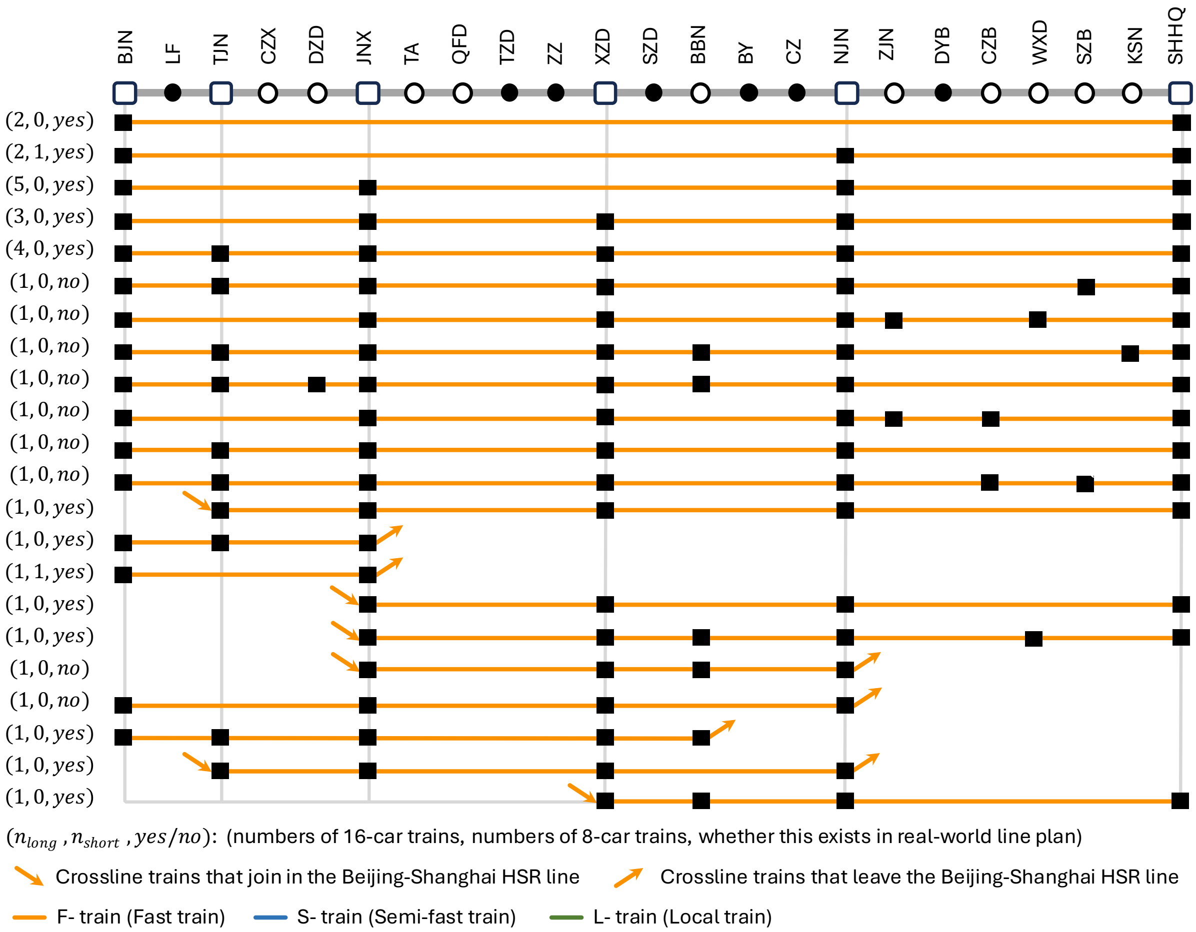


Figure 4.9 The optimised line plan (F- trains) on Beijing-Shanghai HSR (Case D2.2)

Table 4.8 Railway supply indicator performance in different percentage (Condition 1)

Style	**Line plan structure**	**Total cost**	**Total intermediate stops (Travel time lost)**	**Rows (from top to bottom): total direct service/average frequency/ average intermediate stops/average tickets assigned to each OD**					
				B-B	**B-M**	**B-S**	**M-M**	**M-S**	**S-S**
D2.1	30%+30%+40%*	17,692,967.5 CNY	3,450 (26,117 minutes)	738	782	136	122	85	12
				50.8	11.84	4.17	2.36	1.27	1.2
				1.97	2.49	3.00	2.96	2.93	1.20
				53	27	65	24	22	18
D2.2	50%+20%+30%	16,682,702.5 CNY	3,410 (19,607 minutes)	804	717	111	136	61	15
				53.06	10.86	3.08	2.40	0.77	0.93
				2.16	2.22	3.75	3.04	3.03	1.15
				48	29	80	21	30	21
D2.3	20%+40%+40%	16,814,982.5 CNY	3,620 (26, 872 minutes)	728	793	136	94	68	22
				48.53	12.02	4.29	1.71	1.02	1.20
				1.64	2.11	2.75	3.48	3.00	1.30
				54	26	63	30	30	18

Table 4.9 Railway supply indicator performance in different percentage (Condition 2)

Style	Line plan structure	Total cost	Total intermediate stop (Travel time lost)	Rows (from top to bottom): total direct service/average frequency/ average intermediate stops/average ticket assigned to each ODs					
				B-B	**B-M**	**B-S**	**M-M**	**M-S**	**S-S**
D2.1	30%+30%+40%	17,692,967.5 CNY	3,500 (26,817 minutes)	810	840	150	130	85	23
				52.00	12.72	4.17	2.36	1.27	5.3
				2.20	2.63	3.25	2.11	2.00	1.73
				48	25	59	23	22	16
D2.2	50%+20%+30%	16,682,702.5 CNY	3,420 (22,925 minutes)	850	785	135	100	70	9
				56.67	11.89	3.75	1.82	1.05	0.87
				2.25	2.35	2.58	1.99	1.67	1.11
				45	27	37	29	25	23
D2.3	20%+40%+40%	16,814,982.5 CNY	3,630 (26, 467 minutes)	780	840	145	110	80	34
				52.00	12.72	4.03	2.00	1.21	1.20
				2.00	2.63	2.92	1.95	1.83	1.40
				50	25	60	27	23	15

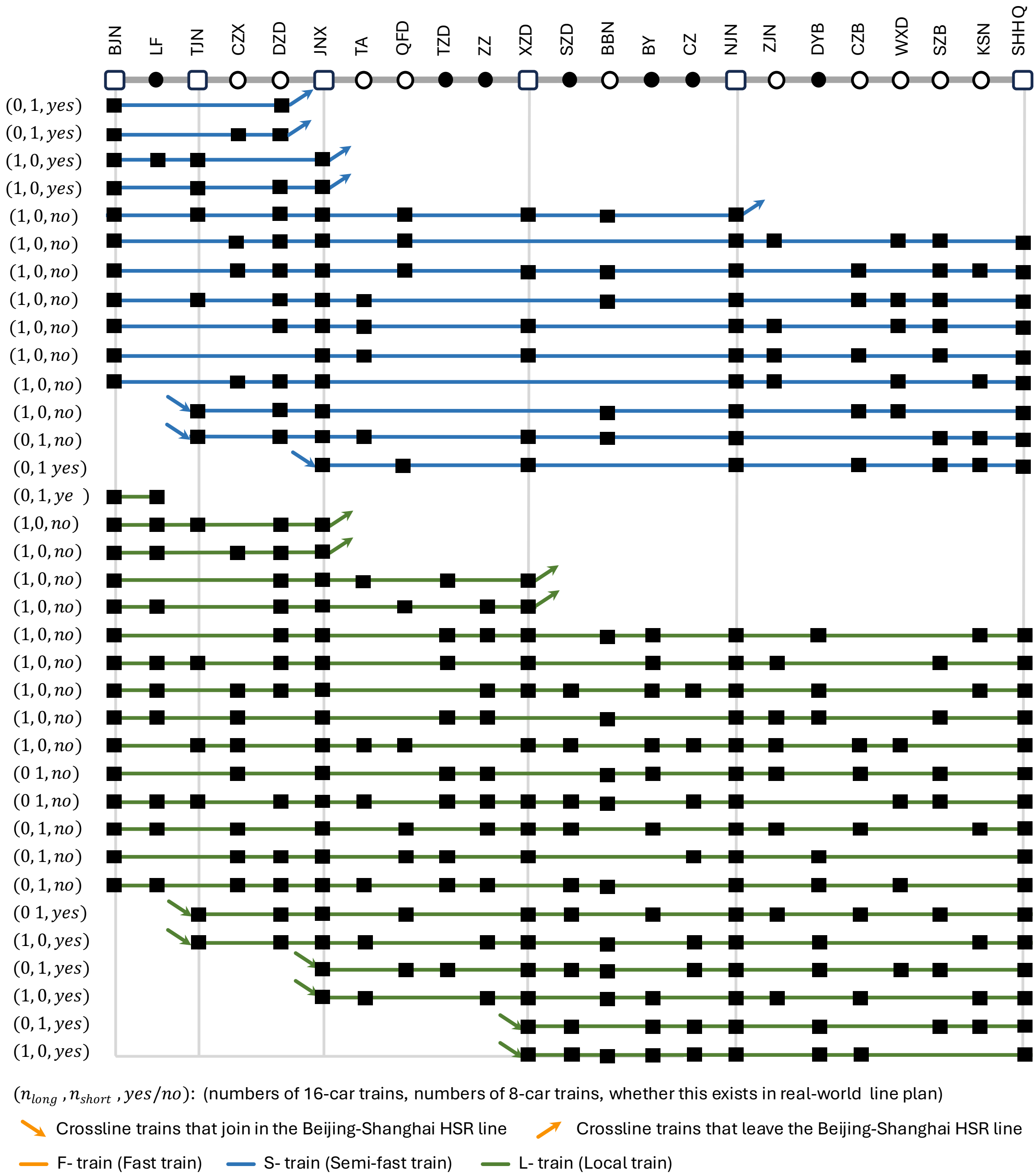


Figure 4.10 The optimised line plan (S-trains and L- trains) on Beijing-Shanghai HSR (Case D2.2)

4.7 Reflective summary

Facing the issue of line planning based on passenger assignment models—such as increased model complexity and the exponential growth in potential train assignments—this chapter proposes a solution through the aggregation of PEIs as different train frequency requirements. This approach significantly reduces the difficulty and variety of enumeration, as a single train inherently comprises multiple transportation products. From the perspective of railway supply, increasing frequency is a straightforward method, but it offers passengers the clear benefit of greater flexibility of travel choices.

Based on this problem-solving approach, an innovative railway supply-based line planning model is introduced that can provide multiple service levels. The chapter defends and argues the model's versatility in terms of periodic line planning, network level line planning, and regional HSR line planning. The model aims to minimise both operation costs and the number of intermediate stops, allowing for comprehensive OD coverage with a smaller set of stop combinations. This chapter interprets Challenge 1.2.2.2 and Challenge 1.2.2.3 at the level of three scheduling styles, explaining the mathematical principles behind minimal stops and reachability, train traffic, and capacity. The optimal stopping strategy must be tailored to the specific railway context to determine the differences between S-trains and L-trains. This chapter also identifies the potential drawbacks of using S-trains and L-trains as train service types: when M- and S-stations are not alternately distributed, both S-trains and L-trains may make consecutive stops on certain railway sections, which could have a negative impact capacity utilisation. Reducing the distinction between S-train and L-train stops would help to establish a more evenly distributed stopping pattern across all major HSR sections.

The selection of objective functions comes with both advantages and disadvantages. The primary advantage is that these two objective functions aim to reduce unnecessary stops, thereby lowering the total number of stops and intermediate stops. This can result in shorter travel times for most ODs, potentially more homogeneous train traffic, and reduced capacity utilisation. However, while equality in the total direct services may not be achieved due to demand disparities between different types of station pairs, these objective functions can ensure equality in terms of average intermediate stops. Another advantage is that adopting these two objective functions avoids the inherent shortcomings associated with turnover-based and loading efficiency-based objective functions. Specifically, the turnover-based objective function tends to favour medium- to long-distance passengers, whereas the loading efficiency-based objective function focuses more on the total number of passengers a train can serve. As a result, L-trains and short- to medium-distance passengers are given preferential treatment, as illustrated in Figures 4.5 (3) and (4). Not to mention that the ticketing process is another branch of optimisation problem.

Table 4.8 Comparison between optimised line plan and the existing plan

	B-B	**B-M**	**B-S**	**M-M**	**M-S**	**S-S**
Case D2.1	738	782	136	122	85	12
Existing plan	578	618	152	237	197	23
Comparison	Increased	Increased	Similar	Reduced	Reduced	Similar

Compared with the existing line plan in the current document, the service frequencies for B-B and B-M have increased significantly, while the frequency of B-S services has remained largely unchanged. In contrast, the service levels for M-M, M-S, and S-S have experienced varying degrees of decline. For instance, in the case of D.2.1, a comparison with the existing plan is presented in Table 4.8. Additionally, the existing line plan only cover the 92.8% of the M-S station pairs and 83.4% of the S-S pair, the Case D2.1 make sure all station pairs get direct train service.

However, the direct services for M-M, M-S, and S-S might be low because there is no constraint guaranteeing related direct service. Another drawback is that a higher ticketing pressure is shown for each train. From the further calculation from Table 4.7 and 4.8, the average passenger each train required to serve is over 4,500, which is much higher than Table 3.1, indicating that adopting 40% of trains to serve 50% passenger demand is challenging and suggesting there could be a motivation for additional stops. In general, the methodology of line planning heavily depends on the chosen solution approach, the interpretation of railway operation scenarios, and the consideration of handling passenger ODs.

Chapter 5 Integrated railway capacity utilisation: a mesoscopic timetable generation model

This chapter explores timetable generation models by comparing different research pathways and establishing a mesoscopic model that integrates railway operation rules and train overtaking behaviours. The effectiveness of the mesoscopic model is demonstrated through a comparison with macroscopic models, highlighting its advantages in capturing railway operation complexities. Finally, four timetabling approaches for capacity improvement are tested: capacity variations affected by overtaking behaviours and dwell time adjustment, adding extra stops, splitting train sets, and analysing opposing conflicts of crossline trains.

5.1 Managing capacity-constrained network

5.1.1 A summary of research pathways and layovers

The research pathway refers to the structured process or methodological steps adopted. There are several layers involved in managing capacity-constrained service network. Firstly, regarding the general approach of the service network management, all four approaches are discussed and analysed in Section 3.2.2. Liao et al. (2024) examined a relatively small network of the China HSR, with one station (Zhengzhou Dong) located as the central station and evaluated the capacity with random train departure orders. Zhang (2019) proposed that sub-networks operate independently and developed timetables for these sub-networks separately. Zhang (2022) analysed the impact of platform changes (e.g. by rendering certain platform tracks unavailable) to assess train schedule delays, with desired departure time windows predefined. Zhang (2019) and Li et al., (2023) focused on the extended Beijing-Shanghai HSR network. Zhang (2019) treated cross-line train operations as time window constraints, while Li et al. (2023) evaluated the impact on capacity by introducing different proportions of cross-line trains.

Secondly, regarding generating a timetable with a minimum timespan, introducing sequential steps for timetable generation followed by the implementation of a compression (Lange et al., 2011; Pouryousef & Lautala, 2015; Zhang, 2012), as stated in Section 2.3.1, is one approach. Another approach is generating timetables by strictly following minimum headways to ensure closely arranged train paths, which is more common compared to the first approach (Petering et al., 2016; Sparing & Goverde, 2017; Zhang & Nie, 2016).

Thirdly, regarding the maintenance of service levels and timetable capacity, there are several pathways that can be identified. One is integrated line planning and timetable generation where the adjustments, stop-skipping, or change to partial stops mainly are made to meet time-dependent demand and increasing average travel speed. Example works includes Jamili & Pourseyed Aghaee (2015); Parbo et al., (2018); and Xu et al. (2021), and the scenarios are mostly in urban and regional railways. Yao et al. (2022) designed the train precedence change scheme to testify further timetable compression. The second approach is modular construction of optimal line plan or timetable structures where the accumulation of local optima leads to a globally optimal solution. Representative works include Liao et al. (2021) who schedule the time-space graph by adding up 'closely-arranged' train path pieces. Zhou (2022) schedules the time-space graph by adopting the minimum unused time-space pieces, and Wu (2020) schedules the time-space graph by adding. more indented train path pairs. The third approach is to insert additional train paths (Burdett & Kozan, 2010; Li, 2022) to directly increase travel opportunities for most ODs, but potentially leading to large-scale timetable adjustments.

Fourthly, regarding train traffic optimisation and adjustments, certain efforts have been devoted to relaxing rigid constraints on microscopic operational elements. For example, Li (2018) relaxed the running time constraints between railway sections to identify optimal transfer links through a greedy algorithm. Li (2022) and Li (2015) focused on relaxing the schedule delay constraints for train lines. However, similar to research on inserting additional train paths, managing the structure and computational complexity of the model becomes highly challenging when such constraints are relaxed.

5.1.2 Identifying the research pathway in this thesis

Given the uneven capacity situations across China HSR network shown in Appendix A, this thesis emphasises the congested parts of the network, which means this thesis focuses on the congested areas of the network. The capacity analysis is, therefore, centred around the congested HSR lines and their associated HSR parts. The rationale behind this focus is that the underutilised areas of the network offer flexibility to absorb operational adjustments stemmed from congested parts.

Building on the research pathway outlined in Chapter 4, which aims to provide a broader range of direct services through a line plan with fewer stops and train line combinations, this thesis proposes the addition of stops informed by a capacity sensitivity analysis. The motivation behind this proposal is that a relatively small increase in the timespan can generate a significant increase in OD direct services, and such partial timetable structure can be modularly arranged. In relation to the microscopic operational elements, this thesis sets dwell time allowances as an approximation for relaxing time-related cost elements. The rationale here is that any extra running time or associated costs would ultimately affect the arrival and departure times at a station and can therefore be considered as 'absorbed' dwell time.

Furthermore, this thesis examines the capacity evaluation considering platforming constraints. There are well-established macroscopic timetabling models that consider overtaking behaviours aimed at reducing capacity utilisation (i.e., minimising the timespan), such as the works of Yan & Goverde (2017); Zhang & Nie (2016). However, a significant concern arises when stations do not have sufficient platforms to handle successive arrivals, in which case the model may fail to accurately represent the true timespan. Besides, unintended overlapping scenarios may also occur if platform availability is neglected (Harrod, 2011; Zhang, 2022). Moreover, as outlined in Section 3.2.4, the allowance for dwell time is a crucial factor in enabling overtaking behaviours. Consequently, the proposed model in this chapter shall provide a more comprehensive analysis of these issues. The time, space, train relationships can be graphed as Figure 5.1.

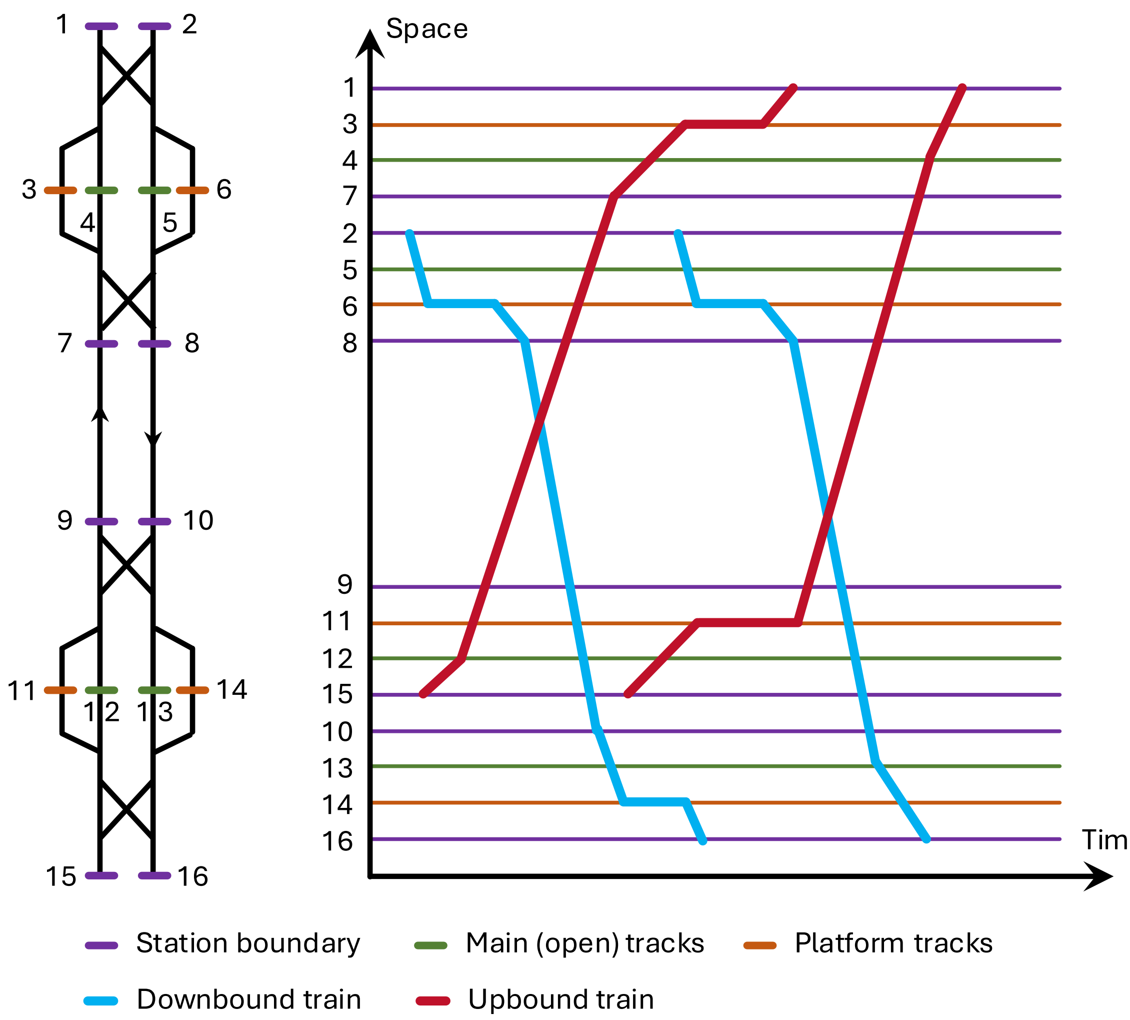


Figure 5.1 Modelling principles of mesoscopic timetabling: space-time diagram

5.1.3 Formal problem statement

The timetable generation stage in this chapter is a mesoscopic and integrated timetable generation and platforming problem, which is stated as: given the layout of a railway network and a train line proposal, the objective is to determine the optimal precedence of trains across railway sections and within stations and to generate the time-space graph (i.e., timetable) that minimises the overall timespan (i.e. makespan): $min\ F = timespan$

5.1.4 General assumptions

- This thesis does not consider the train trajectory variations caused by driver skill errors, rolling stock grouping (long or short types), and electronic multiple units features.
- This thesis assumes that the valuation of headways, minimum dwell time and other necessary time cost have already considered the initial stability by adding adequate buffer time (Ning et al., 2009).

5.2 A modular modelling framework

5.2.1 Notation

Table 5.1 Space-time network denotations

Symbol	Description
i, j	Physical point: station boundary, platform track, open track.
t, τ	Discrete time instants, the unit is one minute.
$v \in (i; t)$	Vertex of the space-time network
$a \in (i, j; t, \tau)$	Arc of a space-time movement: from $(i; t)$ to $(j; \pi)$

Table 5.2 Sets, index, and supportive parameter

Set	Description
A	Set of arcs, indexed by a.
F	Set of trains, indexed by f, f'.
S	Set of stations, indexed by s

Set	Description
D_f	Set of stop list of train f
E	Set of stations, indexed by e
T	Set of time windows
V^s, U^s	Set of platform tracks and open tracks at station s, indexed by p
P^s	Set of and all tracks at station s, indexed by p, $P^s = V^s \cup U^s$
A^+, A^-	Approaching and departing direction
A_f^R	Set of arcs representing the running process of train f in sections, stations
A_f^D	Set of arcs representing the arriving and departing process of train f
A_f^{Dw}	Set of arcs representing the dwelling process of train f
A_f^P	Set of arcs representing the passing process of train f without stop
A_f^{Or}	Set of arcs representing the departing process of train f at its origin station
A_f^{Te}	Set of arcs representing the arrival process of train f at its terminus station
A_s	set of arcs associated with station s
A_v	Set of arcs related to the vertex v, $A_f^R, A_f^D, A_f^P, A_f^{Or}, A_f^{Te}, A_s \subseteq A_v$
$\pi_{max}^{Dw}, \pi_{min}^{Dw}$	The upper and lower bound of dwell time
π_e^A	The additional time cost if stop at the end of section e
π_e^D	The additional time cost if depart from the start at the beginning of section e
π_e^R	The pure running time cost in section e
π_s^W	The dwell time cost occurring at station s
π_s^B	The random buffer time added to a time cost, representing an initial variation.
M	Sufficiently large positive number, set as $24 \times 60 = 1440$ in this case

Table 5.3 Decision variables

Symbol	Description
x_a^f	0-1 variables, =1representing an arc is selected

Symbol	Description
y_f^s	0-1 variables, =1 representing a platform is selected when entering the station s
z_f^s	0-1 variables, =1 representing if train stops at the station s
$b_{f,f'}^s$	0-1 variables, =1 representing both train f and f' required the same platform
$o_{f,f'}^s$	0-1 variables, =1 representing train f arrives before train f' at station s

Table 5.4 Headway set (H) within the station (unit: minutes) (Li et al., 2023)

The front train	The later train	Arrival headway	Departure headway
Stop	Pass	$h_{ar}^{sp} = 5$	$h_{de}^{sp} = 5$
Stop	Stop	$h_{ar}^{ss} = 3$	$h_{de}^{ss} = 4$
Pass	Stop	$h_{ar}^{ps} = 3$	$h_{de}^{ps} = 3$
Pass	Pass	$h_{ar}^{pp} = 3$	$h_{de}^{pp} = 3$

5.2.2 Train movement illustration

In the subsequent paragraph, all time-related concepts follow a consistent description: a time window represents a specific t and τ and denoted as $T = [t, \tau]$. Time refers to discrete time instances. The time cost does not have time stamp attributes and is regarded as a 'cost' (e.g., headway, dwell time, running time). The timespan, or the makespan, represents one form of capacity utilisation value. The sentence that τ is greater than t can also be stated as τ is later than t,or $t < \tau$. The mathematical illustration of train movements is shown as Figure 5.2.

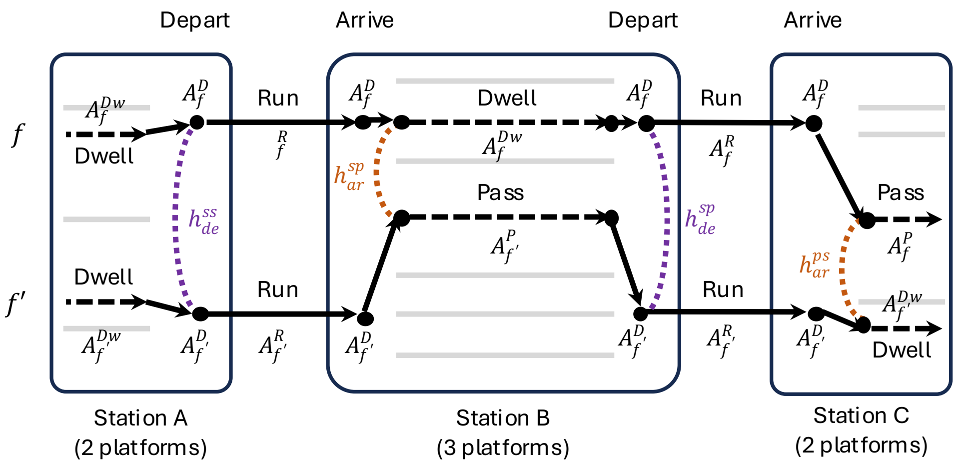


Figure 5.2 An illustrative diagram of timetable generation modelling

1. For all time windows that shall link to the decision variables, the formulation is:

$$t \times x_{a_f} \leq \tau \times x_{a_f}$$

2. For train f running movements $\forall a_f(i,j;t,\tau) \in A_f^R$, the time window T is formulated as:

$$\tau - t = \pi_{ij}^R + \pi_j^A \times z_f^{s-1} + \pi_i^D \times z_f^s + \pi_s^B$$

3. For train f passing movements $\forall a_f(i,j;t,\tau) \in A_f^P$, the time window T is formulated as:

$$\tau - t = \pi_{ij}^R \times (1 - z_f^s)$$

4. For train f dwelling movements $\forall a_f(i,j;t,\tau) \in A_f^{Dw}$, the time window T is formulated as:

$$z_f^s \times \pi_{min}^{Dw} \leq \tau - t \leq \pi_{max}^{Dw} \times z_f^s$$

For all τ and t, their assignment is inherently associated with a specific set $A_f^R, A_f^D, A_f^{Dw}, A_f^P$, ensuring that each τ and t holds a distinct interpretation within the given arcs a_f.

5.2.3 The model structure (M3)

Objective function: $min\ F = timespan$

Group I: train movements within the sections-flow balance constraints

$\sum_{(1)} x_{a_f} = 1$	$\forall f$	(1): $a_f \in A_f^{or-}$	(5-1)
$\sum_{(2)} x_{a_f} = \sum_{(3)} x_{a_f}$	$\forall f, v$	(2): $a_f \in A_v^+$. (3): $a_f \in A_v^-$	(5-2)
$\sum_{(4)} x_{a_f} = 1$	$\forall f, v$	(4): $a_f \in A_f^{Te+}$	(5-3)
$\sum_{\forall f} \sum_{(5)} x_{a_f} \leq 1$	$\forall f, s, e$	(5): $a_f \in A_f^R, A_f^D, A_f^P \Longrightarrow T$	(5-4)

Group II: train movements within the stations

$x_{a_f} = z_f^s \times \sum_{V^s} y_f^s$	$\forall f, s$	$a_f \in A_f^D$	(5-5)
$x_{a_f} = (1 - z_f^s) \times \sum_{U^s} y_f^s$	$\forall f, s$	$a_f \in A_f^P$	(5-6)

Constraints (5-1), (5-2) and (5-3) are the flow balance constraints representing an integrated train path that is defined for each train. Constraint (5-4) represents that when a train path is running between sections (A_f^R), arriving at, stopping at, passing through, and departing from stations (A_f^D, A_f^P), there should be a corresponding headway or blocking time to ensure the safety. The literal explanation is that, for any arc associated with A_f^R, A_f^D, A_f^P, within the specific time window T, only one train f is assigned to this train path. Constraint (5-5) represents that if the stopping train should be assigned with a platform from V^s. Constraint (5-6) represents that if the passing should be assigned with open track from U^s.

Group III: Mapping space-time variables with physical timeline

$$timespan \geq \tau \qquad \forall f, e, s \qquad (5\text{-}7)$$

$$t \times x_{a_f} \leq \tau \times x_{a_f} \qquad \forall a_f(i,j;t,\tau) \in A_f^R, A_f^D, A_f^{Dw}, A_f^P \qquad (5\text{-}8)$$

$$\tau - t = \pi_{ij}^R + \pi_j^A \times z_f^{s-1} + \pi_i^D \times z_f^s + \pi_s^B \qquad \forall a_f(i,j;t,\tau) \in A_f^R: \{(i,j) \in e; (t,\tau) \in T\} \qquad (5\text{-}9)$$

$$z_f^s \times \pi_{min}^{Dw} \leq \tau - t \leq \pi_{max}^{Dw} \times z_f^s \qquad \forall a_f(i,j;t,\tau) \in A_f^{Dw}: \{(i,j) \in s; (t,\tau) \in T\} \qquad (5\text{-}10)$$

$$\tau - t = \pi_{ij}^R \times (1 - z_f^s) \qquad \forall a_f(i,j;t,\tau) \in A_f^P: \{(i,j) \in s; (t,\tau) \in T\} \qquad (5\text{-}11)$$

Constraint (5-7) represents the timespan constraint, which tracks the very last time instant of the timetable. Constraint (5-8) represents the timeline mapping for a train path. Constraint (5-9) represents the running process of a train path. Constraint (5-10) represents the dwell time of a stopping train. Constraint (5-11) represents the running time costed by a passing train within the station.

Group IV: Mapping the precedence between a pair of trains

$$\tau - \tau' + (1 - o_{f,f'}^s) \times M \geq 0 \qquad \forall a_f \cup a_{f'} \in A_{s,e}^R \qquad (5\text{-}12)$$

$$o_{f,f'}^s + o_{f',f}^s = 1 \qquad \forall a_f \cup a_{f'} \in A_{s,e}^R \qquad (5\text{-}13)$$

$$b_{f,f'}^s \times (1 - o_{f,f'}^s) \times (t' - \tau) + b_{f,f'}^s \times o_{f,f'}^s \times (t - \tau') \geq H \qquad \forall a_f \cup a_{f'} \in A_s^D \cup A_s^P \qquad (5\text{-}14)$$

Constraints (5-12) and (5-13) establish the logic coupling between precedence variables and time instants of trains f and train f'. Constraint (5-14) represents that a precedence decision and corresponding headway must be made whenever there is a possible route conflict between trains f and train f'.

5.3 Solution approach and model analysis

5.3.1 Computing efficiency of the business solver

This section utilises Gurobi as the optimisation solver to validate the effectiveness of the model. A series of random small case studies will be conducted by following established rules: The number of stations will increase by 10 at each step, starting from 10 and ending at 40, reflecting the number of stops on the Beijing-Guangzhou HSR line. The number of trains will begin at 10 and increase up to 40, adding 10 trains each time. The number of platforms at each station will be generated ranging from two to four. The section running time between stations ranges from 10 to 20 minutes. The minimum dwell time is set at three minutes, with a maximum dwell time of 30 minutes to accommodate potential overtaking. The train lines are randomly generated, with the number of stops ranging from two to the total number of stations. The case studies are conducted in a single direction only. To manage computation time, the acceptable optimality gap is set at 10%, and the maximum computation time is limited to 500 seconds. Each case will be run five times, as some configurations may result in infeasibility. The total number of case studies is 80.

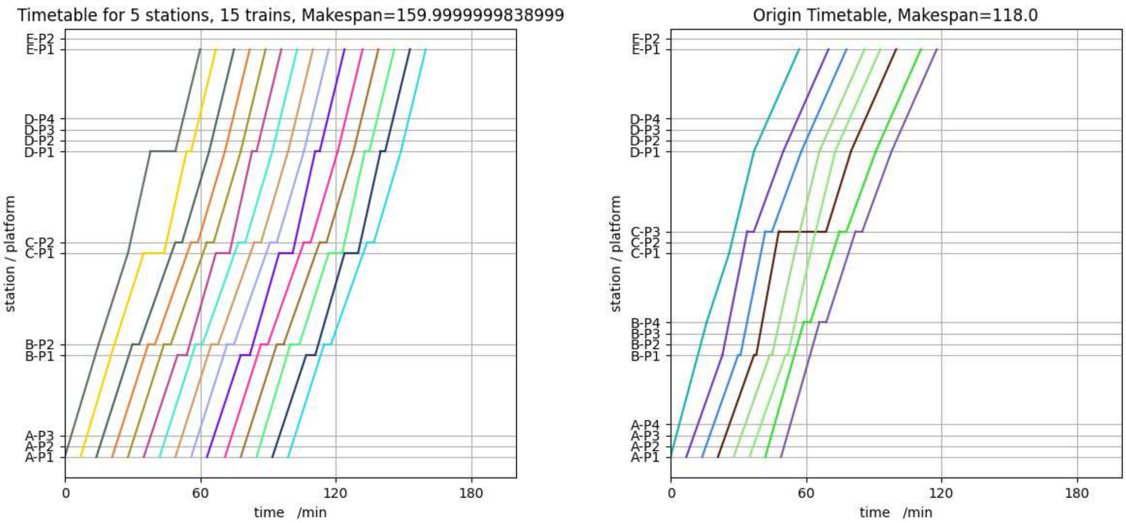


Figure 5.3 Examples of a timetable output

However, this series of random experiments only returned four feasible solutions. These include one case for 10 stations with 10 trains, two cases for 30 stations with 10 trains, and one case for 40 stations with 10 trains. The respective timespans were 235 minutes, 595 minutes, 697 minutes, and 794 minutes. Admittedly, extending the iteration time or increasing the number of random test cases could yield more feasible solutions. Nevertheless, since the experiments were conducted randomly, they provide a more accurate representation of the inefficiency of the business solver when dealing with large-scale problems.

5.3.2 An insertion-inspired heuristic algorithm

Facing the inefficiency by business solver experiments and getting inspired by the train intensity heuristic (Liao et al., 2021). This thesis designed an insertion-inspired heuristic, which belongs to the modular timetable generation method. The general principle is arranging train paths close to each other without violating the headway and maximum dwell time constraints. We define two additional sets which are: $F_{draw}, F_{on-hold}$.

Table 5.5 The outline of the timetable generation solution

Input: railway network data, train line plans, parameter

Step 1: Initialisation

Initialise F_{draw} as the set for feasible train paths. Let $F_{draw} = \emptyset$.

Initialise $F_{on-hold} = F$ as the set for un-scheduled train lines.

Read railway data: S, E

Read headway set: H

Read Running time and other time cost: π

Step 2: Drawing the time-space graph

For each train line $f \in F_{on-hold}$, read: the stopping list: D_f

Solve the model M3 at Section 5.2:

- For each section and station of the train line f

  Running between station:

$$\sum_{(2)} x_{a_f} = \sum_{(3)} x_{a_f}, \text{2): } a_f \in A_v^+. \text{ (3): } a_f \in A_v^-$$

$$t \times x_{a_f} \leq \tau \times x_{a_f}$$

$$-t = \pi_{ij}^R + \pi_j^A \times z_f^{s-1} + \pi_i^D \times z_f^s + \pi_s^B$$

  - if dwelling at the station:

$$x_{a_f} = z_f^s \times \sum_{V^s} y_f^s$$

$$z_f^s \times \pi_{min}^{Dw} \leq \tau - t \leq \pi_{max}^{Dw} \times z_f^s$$

  - if passing at the station:

$$x_{a_f} = (1 - z_f^s) \times \sum_{U^s} y_f^s$$

$$\tau - t = \pi_{ij}^R \times (1 - z_f^s)$$

  - if succeed:

    - Update the $x_{a_f}, y_f^s, z_f^s, b_{f,f'}^s$, and $o_{f,f'}^s$

    - Update the $timespan$ value.

    - Next train line, satisfying:

$$\tau - \tau' + (1 - o_{f,f'}^s) \times M \geq 0$$

$$o_{f,f'}^s + o_{f',f}^s = 1$$

$$b_{f,f'}^s \times (1 - o_{f,f'}^s) \times (t' - \tau) + b_{f,f'}^s \times o_{f,f'}^s \times (t - \tau') \geq H$$

  - else:

    - skip this train line.

Step 3: Postprocessing:

Update the feasible timetable and time-space graph

Appending the $F_{draw} \longleftarrow F_{draw} \cup \{f\}$.

Removing the $F_{on-hold} \longleftarrow F_{on-hold}/\{f\}$.

Step 4: Repeat Step 3 and Step 4

End if $F_{on-hold} = \emptyset$.

Step: Return the feasible timetable and time-space graph

5.3.3 Comparing integrated platforming vs non-platforming approaches

To demonstrate the practical significance of the model, specifically how platforming impacts the accuracy of capacity assessment, this section compares the results of macroscopic and mesoscopic timetable generation models: The experiment is set with a total of 10 stations and 10 trains. Notably, in the mesoscopic timetable generation, the number of platforms is randomly assigned between 1 and 3. Both models are tested through random experiments until five valid solutions are obtained for each.

The results show that the macroscopic model's results range between 210 and 230 minutes, while the mesoscopic model's results range from 229 to 245 minutes. In particular, when the minimum number of platforms is set to 1, the resulting timespans consistently exceed 240 minutes, whereas when the minimum platform value is set to 2, the timespans range between 225 and 237 minutes. The macroscopic model assumes an ample number of platforms, allowing overtaking to occur freely at every station and consecutive arrivals. This results in an overly optimistic compression of the timetable, leading to a shorter timespan, while this phenomenon is constrained by mesoscopic model structure at Section 5.2. Therefore, it is obvious that the mesoscopic models, which is defined as the integrated timetable and platforming problem, fits better on real railway operation scenarios.

However, the experimental results suggest a hidden trend which previous studies might not argue before: when the number of platforms at a station on a single direction is two or more, the discrepancy between capacity assessments from macroscopic and mesoscopic models could be less significant. In other words, there are certain conditions under which the inclusion or exclusion of platforming considerations in the timetable generation model has little to no significant impact on the final timespan. This can be cross validated using queuing theory principles: specifically, when the minimum dwell time of a train is shorter than the headway, one or two platforms are generally sufficient to accommodate various precedence relationships between trains.

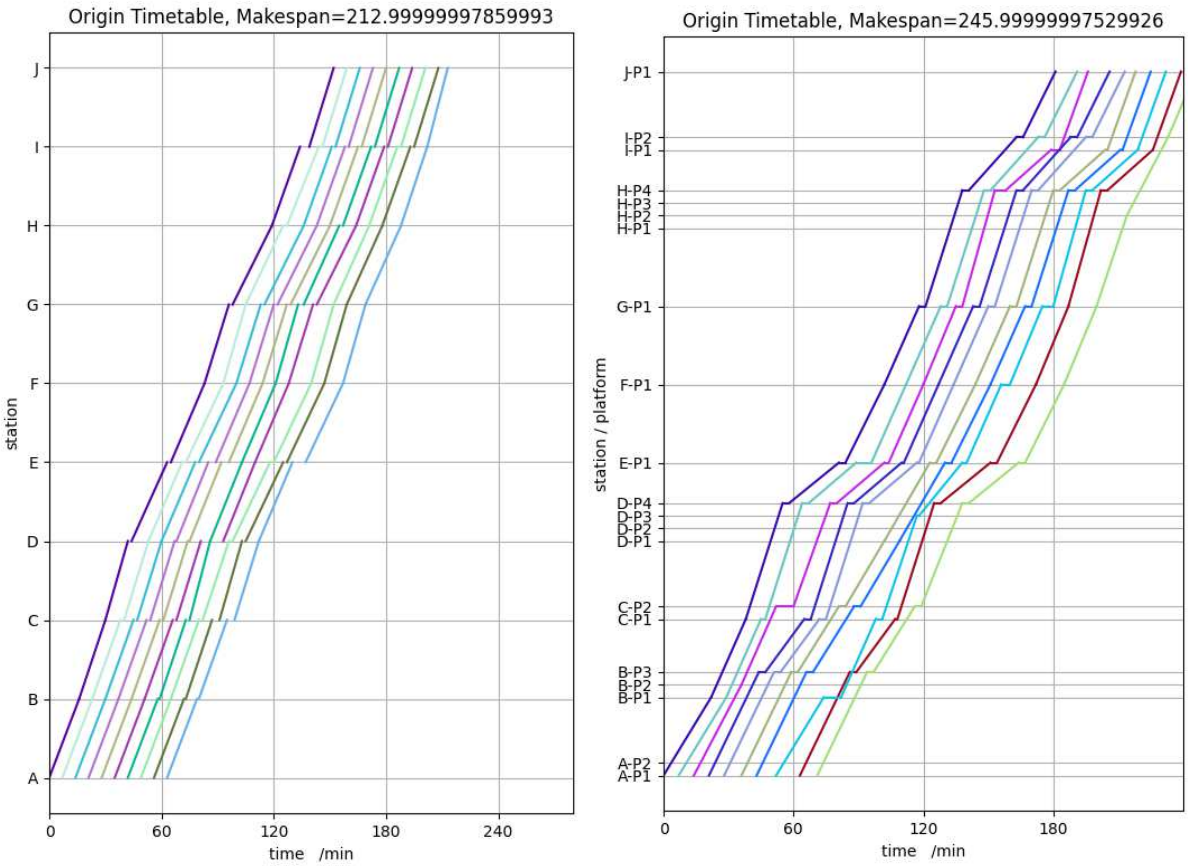


Figure 5.4 Comparison between macroscopic vs mesoscopic models

5.4 Improving the capacity utilisation

Generally, Chapter 4 concludes a line planning which minimises the number of intermediate stations for all OD pairs while meeting the required minimum service frequency. Acting as input, Section 5.3 and Section 5.2 optimise in a feasible timetable with the shortest timespan, and the feasible timetable can be put into practice. However, several representative studies, such as Li (2022); Zhang (2019), suggest that there is still potential for improving capacity utilisation, for example, by adjusting stops (i.e., jointly adding or removing stops) or flexibly altering departure times at origin stations. Both studies are based on macroscopic level models that overlook infrastructure constraints, raising doubts about their validity, as verified in Section 5.3.3. Moreover, while removing stops aims to reduce capacity utilisation, it often ignores the loss of direct travel opportunities for certain OD pairs.

Liao et al. (2024) and Zhang (2019) both question the prioritisation of long-distance and cross-line trains in China's HSR service scheduling, studies such as Li (2020); Zhang & Nie (2017), that discuss the cross-line train operations, also remain at a macro level and adopt two basic methods: either ensuring a specific headway or fixing an unchangeable time window, without addressing the detailed categories of cross-line train movements, as stated in Section 3.2.4.

Thus, capacity improvement focuses on finding additional benefits for passengers and operators after achieving the core capacity goal (i.e., timespan). The working contents of this section are:

Firstly, as highlighted in Section 2.2, the most critical infrastructure factors affecting capacity are open tracks and the number of platforms at each station, which are prerequisites for determining how many trains can be accommodated within a given time window, successive train arrivals and departures, and the occurrence of overtaking behaviours. At the mesoscopic modelling scale, it is necessary to explore the relationships between these factors.

Secondly, with a given line plan, the sensitivity of the timespan is tested by increasing the number of stops under a series of rules and reasoning. If the timespan increase is insignificant, the additional stops can be considered as extra benefits for certain passengers. Admittedly, more stops may increase In-vehicle time for some OD pairs due to intermediate stops; however, whether to add stops is ultimately decided by the railway operator.

Finally, in terms of cross-line and long-distance trains, two scenarios are considered. In the first, where cross-line train movements are following, the proportion of cross-line trains should be determined by passenger demand, rather than being determined by capacity utilisation rate. In the second, where cross-line train movements are opposing, the primary focus is on analysing how the arrival and departure processes affect other original-line trains and whether conflict-free routes exist. By jointly analysing the actual train density, passenger OD feature, and the cross-line train movements across the China HSR network, the Changsha Nan (South) station on the Kunming–Guangzhou direction is used as an example. This section explores the key factors that should be considered when organising transfers under such conditions.

5.4.1 Exploring the timetable structure: extending dwell time allowance

The experiment is conducted on a sample network consisting of five stations and ten train lines to examine how varying dwell times influence the overall timespan. Each dwell time scenario (ranging from 2 to 6 minutes) is tested across five cases (marked as 'X' in Figure 5.5). It is important to note that this experiment does not account for any headway reductions resulting from signalling system upgrades. Additionally, the arrival and departure headways remain constant at 3 minutes in most cases.

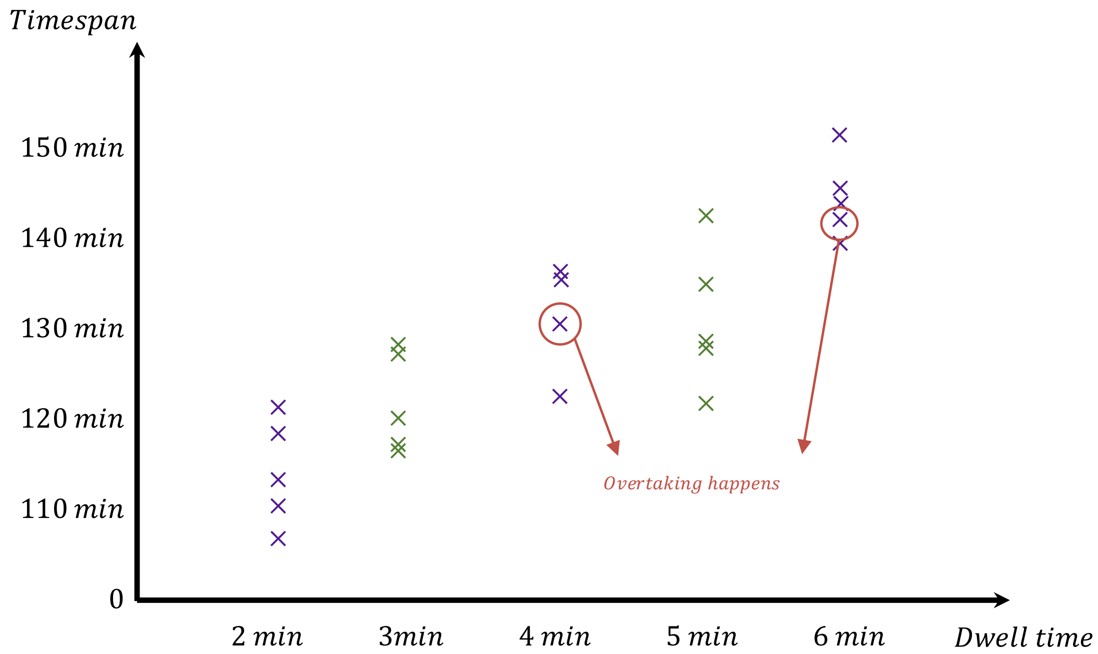


Figure 5.5 Timespan experiment results and the trend graph

Firstly, when dwell time is less than the headway (i.e., 3 minutes), overtaking behaviour does not occur, and the timespan is primarily influenced by the headway, resulting in a relatively short time span. When dwell time equals or exceeds 3 minutes, the timespan becomes variable and intends to increase. The occurrence of overtaking behaviour depends on the stopping patterns of the preceding and following trains (results indicated by the red circles), as well as the maximum permissible dwell time. While overtaking behaviour can reduce the timespan, its impact remains limited. The results indicate that the longest timespans are generally associated with stations that have only one platform which only has one siding, suggesting the importance of platform availability.

Secondly, Increasing the minimum number of platforms per station from one to two immediately reduces the timespan. As dwell time increases, the variation in timespan is primarily influenced by the number of platforms rather than by train overtaking behaviour. The experiment concludes at a maximum dwell time of 6 minutes, beyond which further testing was not conducted.

The results of this experiment are consistent with queueing theory principles. Assume there is a system with n platforms, the average arrival headway is a minutes, and the average departure headway is d minutes. There are m sources for arrivals, but only one source for departure. The average stopping time is s minutes. When s minutes can be extended and m sources can be changed, the inventory rate is calculated as (Zhao et al, 2023):

$$\rho = \frac{m \times max\{s, d\}}{n \times a}$$

As can be seen, the number of platforms contributes significantly to reducing the inventory rate. In this context, the overtaking behaviour here can be regarded as a train taking a 'virtual' platform, in fact occupying an open track, to leave the system immediately. While overtaking theoretically shortens the departure headway, the overall inventory rate remains governed by $max\{s, d\}$.

As a result, extending the dwell time allowance increases the probability of overtaking occurrences, thereby shortening the total timespan to some extent. However, the number of platforms plays a more decisive role in minimizing timespan compared to train overtaking strategies.

5.4.2 Increasing more direct services: adding extra stops on train lines

The rationale for this section is based on a "tooth-shaped" time-space traffic feature (Wu, 2020): when the front train (blue train in Figure 5.6) stops at a distant station, it allows the following later train (orange train in Figure 5.6) to stop at a nearby station, thus reducing the timespan, improving the capacity utilisation by saving a time-space-slots (light orange parts in Figure 5.6). The new train traffic (right side of Figure 5.6) is described as a 'tooth-shaped'. Particularly, (Zhou, 2022) establishes the minimum waste as a timetable with the minimum area of unused-time-space-slot (light orange part in Figure 5.6). The description of Figure 5.5 suggests the principle for adding stops to train lines:

- Begin by increasing stops on train lines with fewer stops, such as the F-train, and then extend to S-train and L-train.
- Furthermore, by analysing the list of total service frequency at each station, decisions can be made on where to add stops.

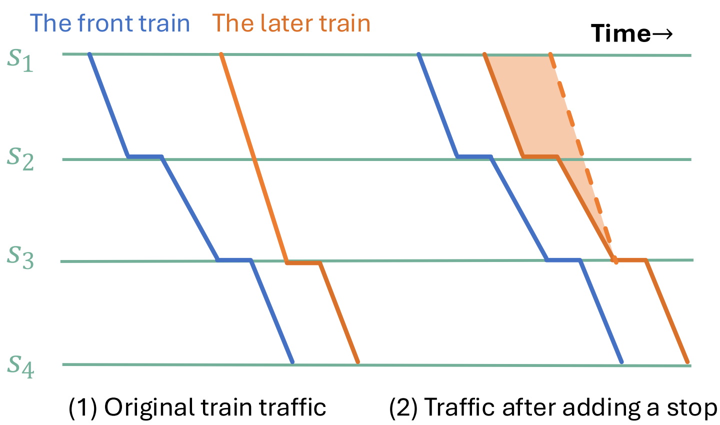


Figure 5.6 An example of capacity improvement by adding stops

The assumed case is set as follows, a railway line contains 11 stations where the Big station is A, G, and L. Medium stations are C, E, J, and K, and Small Stations are B, D, F, H, I, and J. Each B-stations has 4 platform for a singular direction, each M-stations has three platforms, and each S-station has two platforms. Orange, blue, and green trains are F-, S-, and L-trains respectively. The line plan shown below has ensured that at least there is one direct train serving each station pair. The total stop number is 39 excluding OT stations. The running time between each station is set as 20 to 30 minutes. The fast train service of the whole journey is around 330 minutes, while the slowest could be around 400 minutes, which equals to a medium and long-distance railway line in real world.

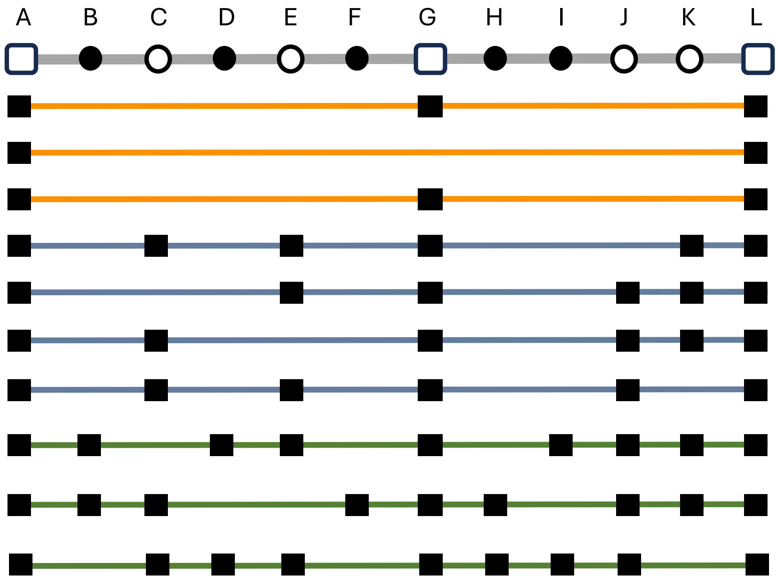


Figure 5.7 The case set for capacity sensitivity analysis

The adding stops process is essentially changing Orange trains (F-trains) slowly into Blue trains(S-trains) and finally all Green trains (L-trains), the discussion of these line plan changes is already mentioned on Section 2.1.4. The testing results are shown as:

The x-axis in the Figure 5.8 is composed of letters and numbers: letters represent specific line plan features, named as scenario, while numbers indicate the total number of stops across 10 trains. Initially, Scenario A corresponds to the reference line plan depicted in Figure 5.5, with an average timespan of 387 minutes. In scenario B, six additional stops were introduced on the F-train, leaving only one orange train in the line plan and increasing the average timespan to 416 minutes. In scenario C, when the line plan is entirely made up of Blue trains and Green trains, the timespan decreases to 398 minutes. For scenario D, the line plan includes only three Blue trains, with the remainder being Green trains, which results in the timespan rising again to approximately 407 minutes. Finally, in scenario E, with all trains being Green trains, the timespan stabilises around 423 minutes.

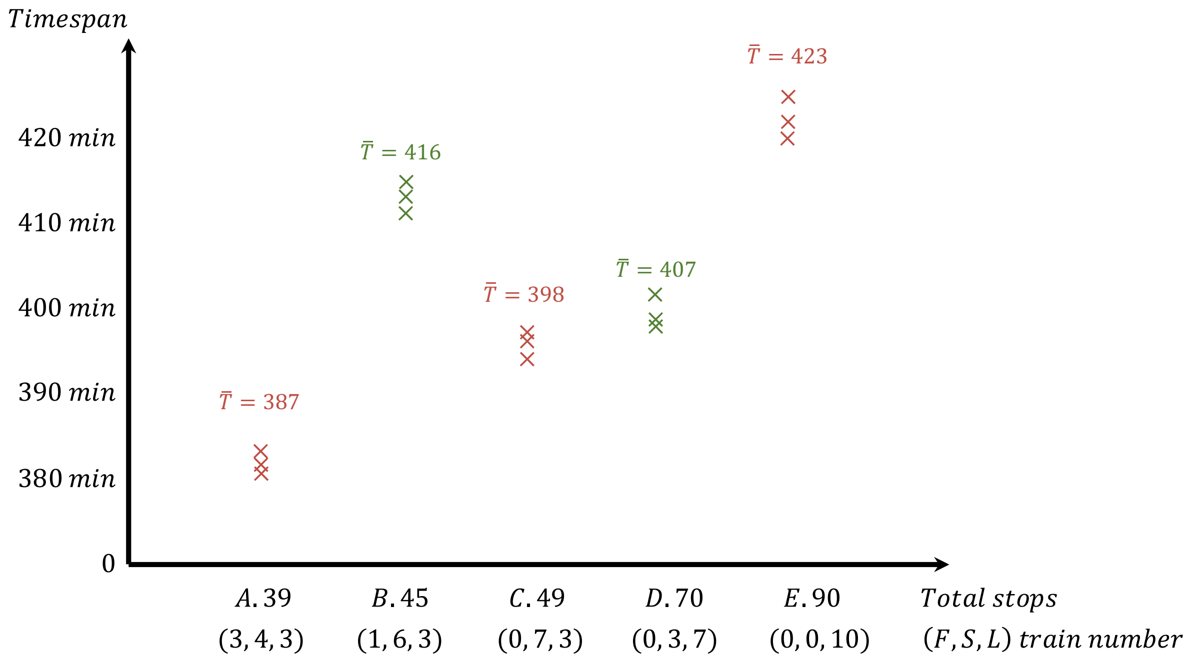


Figure 5.8 Capacity sensitivity results and the trend graph

In summary, the experiment illustrates how changes in the composition of train types (orange, blue, and green trains) impact the overall timespan of the timetables. As extra stops are added to the line plan, the overall trend is that the timespan increases. However, this experiment has confirmed the hypothesis of this section: a small increase in timespan can indeed lead to an improvement in service quality, as demonstrated by the change from scenario A to scenario C, with a timespan increase of 2.8%, service frequency grows 25.6%. Regarding implications, it might be considered to switch from scenario A to scenario C during major holidays or other short periods of high passenger demand, allowing capacity utilisation to remain at a similar level.

5.4.3 Addressing network timetable issue: managing long-distance and cross-line trains

Section 3.2.4 concluded that following cross-line trains can be approximately operated as original-line trains terminating at a node stations, and the track layout of Beijing-Shanghai HSR satisfy all this feature. A brief track denotation and some train line candidates graphed as Figure 5.9 for further explanation, and the detailed mesoscopic track layout of Beijing-Shanghai HSR can be seen in Appendix B.

To be more detailed, C1 in Figure 5.8 represents cross-line train from Northeast China to Shanghai. C2 represent cross-line train from Zhengzhou to Shanghai. C3 represents cross-line train from Northeast China to Lianyun gang. C4 represents cross-line trains from Shijiazhuang to Shanghai. C5 represents cross-line trains from Zhengzhou to Hangzhou. All 5 types of cross-line trains represent major cross-line ODs. According to Section 4.1 by adopting SILMST, Cross-line

ODs which only occupies short railway sections, should be avoided, such as Qufu E. to Xuzhou E. section and Xuzhou E. to Bengbu S. sections. C1 and C4 are the typical cross-line trains which could be handled as original-line trains. More train traffic could be added based on C2, C3, and C5:

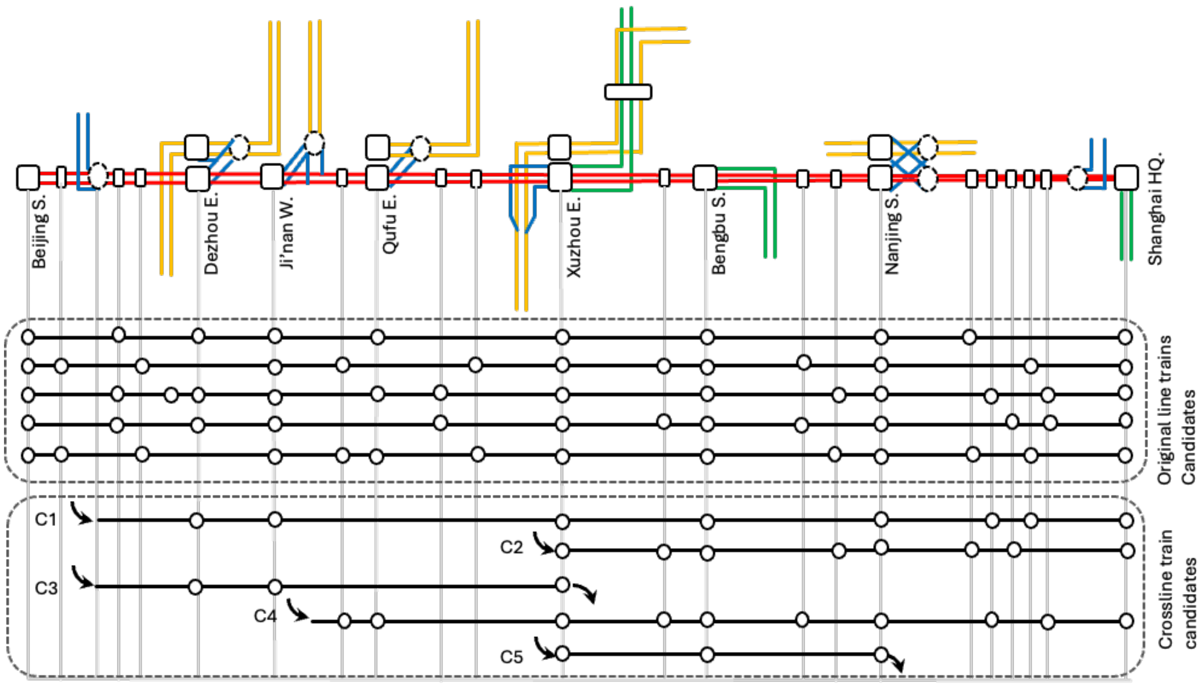


Figure 5.9 A train line plan example for following cross-line train operation

For C1 and C4: a flexible delayed time window can be proposed, within which selected original-line train services can potentially become cross-line trains, extending to major cities on other HSR lines. For example, on the Xi'an to Shanghai Hongqiao journey, the corresponding node station is Xuzhou Dong (E.). The average HSR travel time from Xi'an to Xuzhou Dong is 2 hours and 5 minutes. According to the assumptions in Section 3.1.3, the preferred travel time for passengers is between 10:00 and 16:00, corresponding to a time window at Xuzhou Dong between 12:05 and 18:05, with possible adjustments. The same principle applies to trains travelling in other directions. Such train line is recommended to adopt 16-car type rolling stock.

By adopting the concept of delayed time windows, long-distance trains can extend their journey to cover greater distances, thereby increasing their operational range and increasing OD direct service coverage. For example, Train G87 from Beijing Xi to Changsha Nan continues crossing the Shanghai-Kunming line and extending its itinerary to Shaoyang via Loudi City since 2019.

For C2, C3, and C5: joining train sets could be considered, this measure only works when: 1) two trains are 8-car rolling stock, and 2) either arrival end or departure end of switch area satisfy the following train movements. The entire process **(M4)** could be described as:

Table 5.6 Train line pool updating process considering joining operation

Stage 1: Update the train line pool: $L \leftarrow L \cup L_{group}$

- Introduce l as the joined train line and l', l'' as the selected lines, $l, l', l'' \in L_{group}$
- introduce the mapping relationship among these three lines:

$$y_{l',s} + y_{l'',s} - 1 \leq y_{l,s} \leq y_{l',s} + y_{l'',s} \; \forall s \in S_B \tag{5-15}$$

$$x_{l'} + x_{l''} - 1 \leq x_l \leq x_{l'} + x_{l''} \tag{5-16}$$

$$x_l \cdot v_l = x_{l'} \cdot v_{l'} + x_{l''} \cdot v_{l''} \tag{5-17}$$

Stage 3: Resolving the M2:

S.t. Constraints (4-7) ~ (4-23), Constraint (5-15), (5-16), and (5-17)

Stage 4: Solving M3

S.t. Constraints (5-1) - (5-14)

An iterative-search study has been implemented on Beijing-Shanghai HSR, and some rolling stock joining suggestions are calculated and recommended as:

Table 5.7 Sample calculation for joined-train operation

Front train	**Rear train**	**Splitting at**	**Reference train code**
Beijing-Qingdao	Beijing-Shanghai	Dezhou Dong	G1983, G9003
Qingdao-Hefei	Qingdao-Hangzhou	Bengbu Nan	G6988, G2817
Xuzhou- Hangzhou	Xuzhou- Shanghai	Nanjing Nan	G7107, G1884
……			

In the operation of train splitting, only a limited number of train services form as feasible solutions. This is primarily due to the relatively small proportion of 8-car trainsets concluded by M2 model. Specifically, within the context of the Beijing–Shanghai-based HSR, no feasible solution exists for the joining operation approach. This is because all cross-line train services, in order to ensure sufficient seating capacity, are operated exclusively with 16-car trains.

5.4.4 Discussing trade-off: between transfer and cross-line train movements

The trade-off between transfer services and direct services from the perspective of passengers involves a necessity of sufficient interchange trains when passengers opt for transfer. Conversely, from the perspective of railway operators, the implementation of direct services has to accept that arrangements typically result in a lower overall frequency of service. From this perspective, the amount of train paths is the common focus, which is therefore a better capacity denotation.

Therefore, this thesis defines the trade-off as the potential affected train number (i.e. railway capacity loss). To describe the trade-off more clearly, the trade-off calculation is based on a capacity sensitivity analysis as well.

1. A simple test case

Starting with a simple question based on Figure 1.11: Station B and Station C each contain two sub-station areas, designated for servicing the line A-B-C-D and line F-B-C-E routes, respectively. The infrastructure configuration closely resembles that of the China HSR. Selecting A→B→C→D and F→B→C→E as experiment example (Figure 5.10).

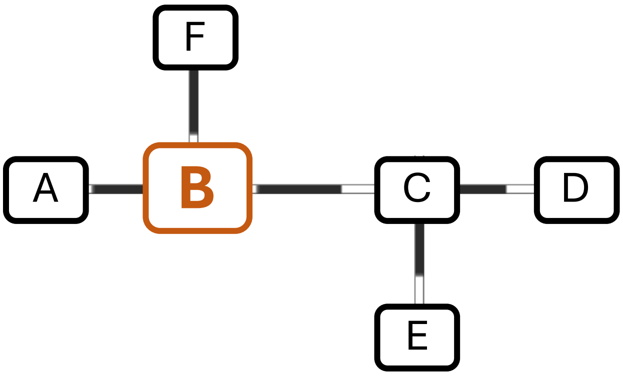


Figure 5.10 A case study of station B permitting multi-directional of train movements

Assuming that the itinerary A→B→C→D, D→C-B→A, F→B→C→E, and E→C→B→F are designated as following directions (original-line), this entails a reduced headway set at 3 minutes. Conversely, the routes A→B→F, F→B→A, D→C→E, and E→C→A are designated as opposing directions (movements are regarded as cross-line trains), utilising connection tracks with a longer headway of 4 minutes. It is important to note that the use of a train in the opposing direction precludes simultaneous use of the switch from other sub-areas due to potential route conflicts (Li, 2019), the affected time is 1 minute. Again, this setting shares a great similarity with China HSR operation at node stations.

Furthermore, stations A, D, F, and E are designated as either O or T stations. The travel time between each station edge ranges from 16 minutes, with a station dwell time fixed at 3 minutes. The total dwell time which contains a turnover at station B and C is set at 10 minutes. The objective is to develop a two-hour timetable that accommodates the maximum train paths from these six train lines. The objective function and constraints are briefly stated as **(M5):**

$$max \sum \sum x_{a_f}$$

S.t. Constraints (5-1) ~ (5-6), (5-8) ~ (5-14)

Similarly, introducing a 'full-timetable-oriented' algorithm

Table 5.8 The solution outline of train path adding

Input: railway network data, train line plans, parameter

Step 1: Initialisation

Initialise F_{draw} as the set for feasible train paths. Let $F_{draw} = \emptyset$.

Initialise $F_{on-hold} = F$ as the set for un-scheduled train lines.

Initialise the total iteration index i

Step 2: Adding train path into the network

For each train $f \in F_{on-hold}$:

Solve the model M3 via Gurobi:

- if succeed:
  - Update the $x_{a_f}, y_f^s, z_f^s, b_{f,f'}^s$
  - Next train line.
- else:
  - skip this train line.

Step 3: Postprocessing:

Update the feasible timetable and

Appending the $F_{draw} \leftarrow F_{draw} \cup \{f\}$.

Removing the $F_{on-hold} \leftarrow F_{on-hold}/\{f\}$.

Step 4: Repeat Step 3 and Step 4

End if $F_{on-hold} = \emptyset$

Return $len[F_{draw}]$ (the number of elements inside a set)

By varying the total number of cross-line trains, set the platforms are sufficient for accommodating all stopping trains and the time window is set as 120 minutes. The X-axis of Figure 5.11 represents the independent variable, which is the increasing number of crossline trains operating on route A-B-F. The Y-axis indicates the number of train paths that Station B can accommodate within a 120-minute time window, serving as a measure of railway capacity.

The series of tests are shown as below (Figure 5.10): there is a clear trend that the number of opposing cross-line train movements will have negative impacts on railway capacity. With the direct service of Train A-B-F increase, the trade-off price (the gap from 80 trains) that a railway operator needs to consider is 3 trains, 8 trains, 10 trains, 15 trains, and 19 trains.

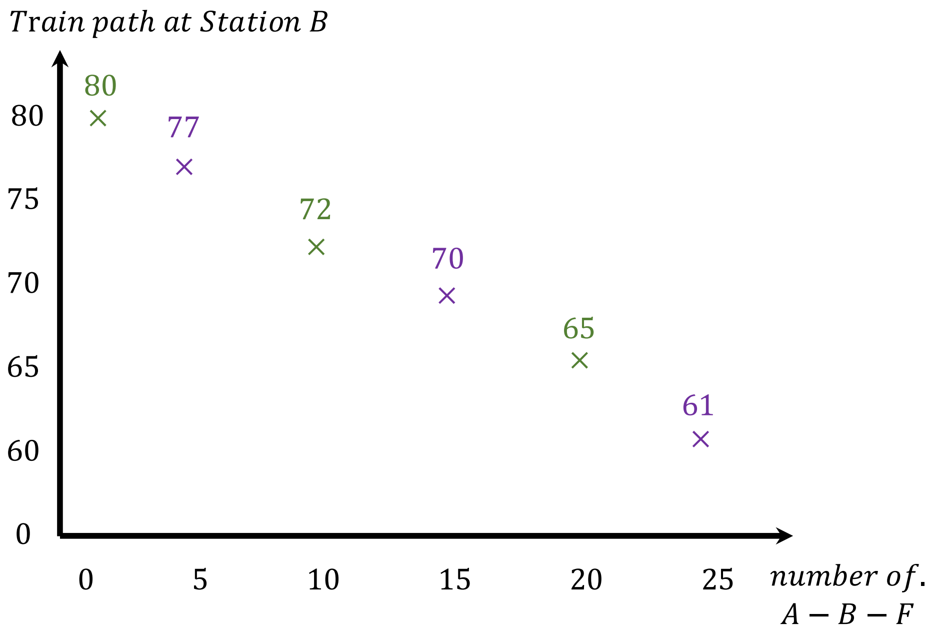


Figure 5.11 Total available train path with different rate of cross-line train (route) A-B-F

The rest of testing results are summarised as follow:

- Keeping the same condition that the platform number is sufficient and dwell time of Train A is extended, there is no difference between dwell time of 20 minutes, 25 minutes, and 30minutes, the result from left to right are: 80, 76, 73, 70, 66, and 63 trains.
- Changing the platform number for train A-B-F to 2, 3, and 4, the maximum train A-B-F can be arranged within 120 minutes is 12 trains, 18 trains, and 24 trains, and the lowest total train path that Station B reduce from 80 is 76, 74 and 72 trains, respectively.

2. Implication and observation:

Inspiring from the testing result above, the cross-line train platform at Changsha Nan railway station is three. Therefore, the most likely affected original-line train number range from 4 to 8 trains within the 120 minutes windows.

In general, when the platform number is sufficient for opposing cross-line trains, the total feasible cross-line train that a station can accommodate increase within a set time window, which is the worst case for network capacity as the potential affected original-line train number would be increase, such trade-off would encourage railway operators to organising transfer and interchange.

5.5 Reflective summary

Given a line plan as input which satisfies a minimum levels of service frequency for all ODs and a systematic comparison of research pathways, an integrated timetable generation and platforming model based on mesoscopic levels (M3) is established, the capacity utilisation is set as the timespan. By introducing train path consistency, time mapping and precedence constraints, the proposed model is capable of scheduling a feasible timetable with the minimum timespan. Compared to macroscopic models, the proposed model shows a more accurate capacity measurement result when the testing result involve stations that only has one platform. By introducing an insertion-based heuristic algorithm, a good feasible timetable can be established when applying on a large-scale network containing a large train line scale.

As further capacity improvement, this section has experimented the relationship between timespan and overtaking behaviours, platform numbers under the condition of dwell time changes, the relationship between timespan and extra stops, and analysing the operation approaches of cross-line trains. Firstly, Results show that platform accounts perform more significant roles on timespan reduction, timespan only shows reduction when all Orange train (F-trains) are changed into Blue trains (S-trains).

Secondly, regarding both splitting and joining operation approaches, an integrated solution process—referred to as M4—was developed. This process conducts itinerary matching based on a solution space composed of 8-car train lines, from which a limited number of feasible splitting solutions were identified. Although this approach can effectively reduce the number of requested train paths in the timetable, it may also result in an overall increase in passengers' travel time due to inconsistencies in the stopping patterns of the front and rear portions of the train. For instance, the front train may stop at a particular station while the rear train does not, leading to mismatches in service continuity.

Thirdly, regarding cross-line train movements within node stations, M5 was developed, using train paths as representations of potentially affected railway capacity (i.e. railway capacity loss). The analysis demonstrates that opposing cross-line movements exert a negative impact on railway capacity, as evidenced by a reduction in the total number of train services that could otherwise have been scheduled in the timetable within the station. This implies a loss in service opportunities for passengers due to the opposing crossline train movements.

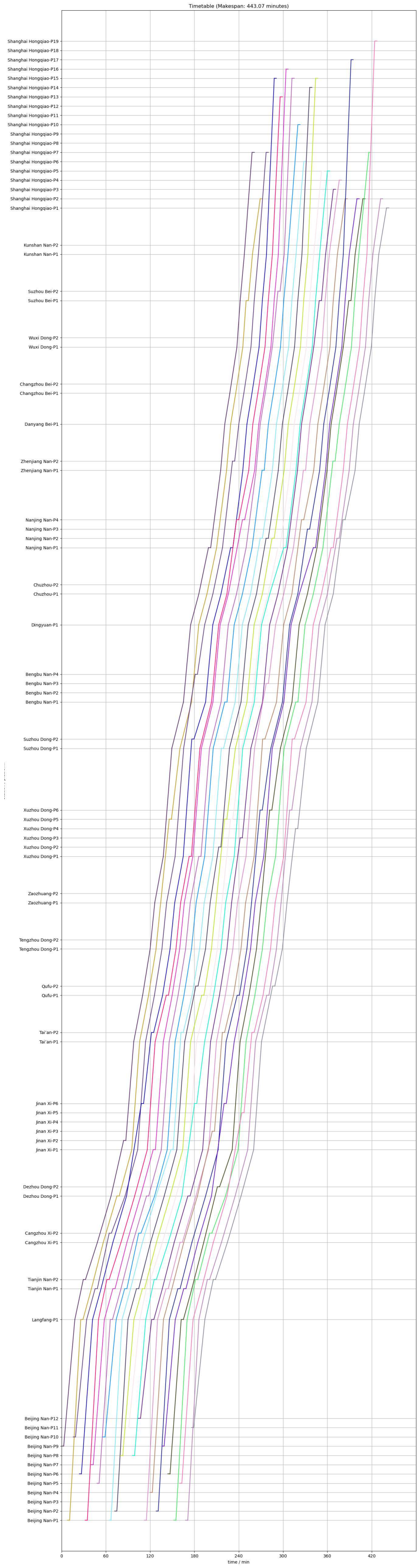


Figure 5.12 An optimised time-space graph on Beijing-Shanghai HSR

Chapter 6 Conclusion

6.1 Evaluation of research objectives

This section evaluates the extent to which each research objective has been addressed in the thesis and how the findings contribute to the field of railway capacity planning and optimisation.

1. Identifying passenger and railway operator factors affecting railway capacity

This objective is addressed through a comprehensive research scope analysis (Chapter 1), literature review (Chapter 2) and further expanded in the theoretical analysis (Chapter 3).

Firstly, Chapter 1 analyses the roles and functions of all railway planning stages within the context of railway capacity studies. It identifies line planning and timetable generation as the core planning stages that most significantly affect railway capacity. Secondly, Chapter 2 identifies key infrastructural factors—such as the number of open tracks, platform tracks, and their layout—as critical elements that differentiate national railway infrastructures. Thirdly, Chapter 3 introduces four network management approaches and concludes that centralising on congested HSR parts better aligns with HSR's primary goal of fast transport to major cities. At the same time, it divides general operational features into operational attributes and timetabling details, thereby clarifying capacity comparisons across different railway contexts.

2. Assessing key capacity-related service qualities in passenger railways

This objective is addressed through the literature review as Chapter 2 and the theoretical analysis as Chapter 3.

Chapter 2 systematically summarises theories and research pathways related to passenger experience evaluation, concluding that capacity-related Passenger Experience Indicators (PEIs) and time-based evaluation functions are most directly correlated with capacity studies in passenger railways.

Chapter 3 explores the relationship between individual passenger decisions and cluster OD groups, resulting in a series of well-founded passenger assumptions. It identifies station pairs and distance as the most relevant factors in distinguishing OD groups. From a detailed level, it summarises the Railway Supply Indicator (RSIs) matching the capacity-related PEIs.

Additionally, Chapter 3 outlines key capacity management frameworks, providing railway operators and infrastructure managers with a structured approach to identifying their core tasks. The details are presented in Figure 6.1.

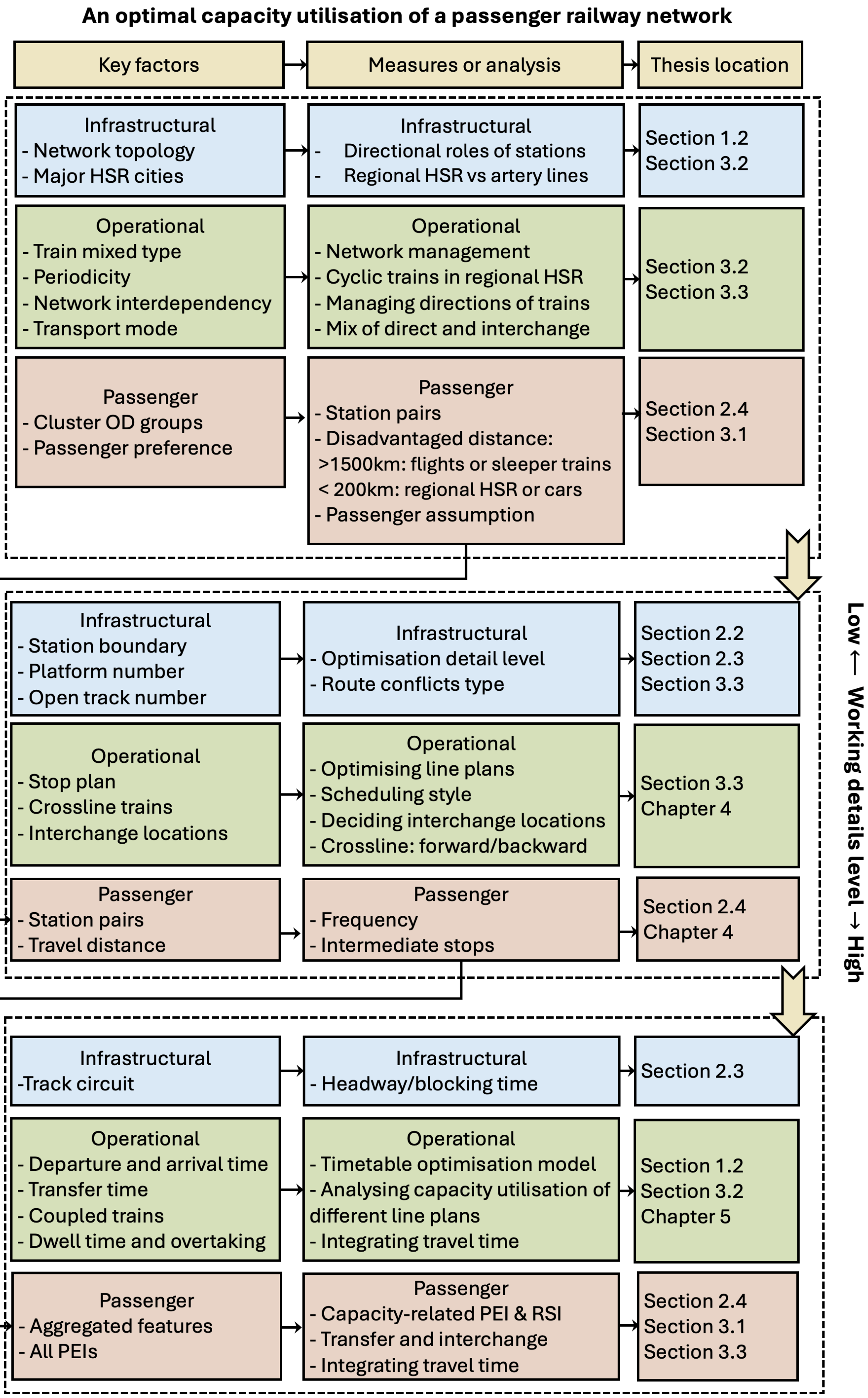


Figure 6.1 Key working contents of optimal railway capacity management

3. Developing a line planning method for testing various service scheduling styles

This objective is addressed in Chapter 4. Chapter 4 begins by dividing the line planning process into two stages. The first stage focuses on minimising the number of interchange locations using a weighted maximum spanning tree approach, ensuring minimal interchange stations from a pure graph theory perspective and providing valuable reference even for other transport modes. It should be noticed that manual setting for OT (origin-terminus) station of cross-line trains needs to be respected first.

The second stage addresses the service frequency for OD pairs, formulating a Weighted Set Covering problem to plan the train lines. The PEI expectations are interpreted as frequency requirements for different types of train lines, allowing flexibility in passenger choice. For a given passenger flow, the objective is to minimise both the number of intermediate stops for passengers and total operational costs of line plans. The inclusion of intermediate stops ensures that each OD pair has more equal travel opportunities and short travel time, benefiting from the train's multi-attribute product offerings: a single train could provide 'fast' or 'semi-fast' services for different OD pairs.

4. Developing a capacity utilisation measurement method for a given train line plan

This objective is addressed by a literature analysis in Chapter 4 and a mesoscopic timetable generation model proposed in Chapter 5.

Chapter 2 first analysed timetable-based capacity measurement methods and argues that mesoscopic-level modelling could be manageable in terms of complexity. Chapter 5 then summarises the research pathways of the timetabling studies that presents an integrated timetable generation and platforming mesoscopic model via train path flow balance, time-mapping, and precedence constraints. The superiority of the model's structure is verified through comparison with macroscopic models and can flexibly change into other modelling frameworks. An insertion-based heuristic algorithm is designed which enables a rapid generation of feasible timetables for a large-scale network with a large number of train lines. The travel time of the PEIs is concluded by the computing results of the model.

5. Optimising timetables to balance capacity utilisation and service quality

This objective is comprehensive addressed by the case studies and experiments designed in Chapter 4 and Chapter 5.

Chapter 4 introduces two key principles to balance capacity utilisation and service quality. Firstly, it proposes minimising the number of intermediate stops to ensure that most OD pairs can access short travel time train services. This is further supported by a service frequency constraint to prevent excessive stop removal. Additionally, a lower total number of stops leads to shorter overall capacity occupation time. Secondly, the proposed line planning model provides multiple levels of service to accommodate the diverse needs of passengers. For instance, some passengers who prioritise later arrival times may be more willing to choose train services with a higher number of intermediate stops. From a model structure perspective, chapter 4 justifies and highlights the model's versatility across periodic line planning, network-level line planning, and regional HSR line planning. It enables railway operators to customise scheduling styles to accommodate varying demand characteristics according to railway contexts.

Chapter 5 explores a number of operational approaches to capacity improvement. These include analysing the impact of overtaking and extended dwell times on capacity, assessing the capacity effect of adding extra stops to the line plan, and examining the impact on the maximum number of trains a station can accommodate by introducing crossline trains with opposing directional movements. Additionally, it investigates the potential of splitting train sets to create extra time-space slots. The results shows that F-train (Fast train) services have a limited impact on the timespan, and no case has been identified where the timespan is significantly reduced under the condition of increased stopping. To the best of our understanding, this thesis is the first to propose classifying cross-line train movements into following and opposing types, offering targeted strategies based on specific station track layout. The trade-off between transfer and direct cross-line trains are defined as affected original-line trains, and the affected capacity is evaluated via a space-time modelling framework that share most of the timetable generation constraints.

Overall, this thesis systematically achieved its research objectives through theoretical development, methodological innovations, and applied case studies, contributing to both academic knowledge and practical railway planning. Future research can extend these findings by incorporating adaptive AI-driven scheduling, real-time capacity monitoring, and multimodal integration.

6.2 Contribution

Compared to previous academic work, this thesis makes a significant improvement and contribution compared to previous works in the following areas:

1. Integration of passenger experience and railway capacity:

Previous studies have often either narrowly focused on improving certain aspects of railway passenger travel quality or on how to achieve a high-capacity occupation rate in a particular line or small-scale network. This thesis systematically connects passenger experience with railway capacity with capacity-related PEIs and RSIs. By making reasonable assumptions, it develops line plans and timetables guided by appropriate scheduling strategies, thereby bridging the gap between these two critical aspects.

2. Systematic compilation of railway capacity knowledge:

This thesis systematically compiles and organises the concepts, interpretations, related research pathways, and methods associated with railway capacity across various research sub-topics. It integrates passenger experience into the discussion of passenger railway capacity. This offers valuable guidance for quickly understanding railway capacity and its related and extended studies.

3. Comparative analysis of railway capacity across different contexts:

Unlike the UIC 406 capacity balance indicators, which are largely theoretical and infrastructure-oriented, this research develops a comparative framework that captures real-world operational characteristics. By incorporating infrastructure, operational, and passenger factors, the thesis provides a more practical tool for evaluating railway capacity in different railway networks.

4. Practical service scheduling solution with capacity improvement:

This thesis redefines railway capacity utilisation and optimisation as a phased service scheduling problem involving line planning and timetable generation in stages. This methodology introduces several key advancements:

- **Shared itineraries for original-line and crossline trains:** Enhancing operational efficiency by reducing the number of interchange locations while maintaining train traffic similarity.
- **Minimising intermediate stops across all trains:** Ensuring fair optimisation across different OD pairs rather than favouring only high-demand or long-distance OD flows.
- **Introducing a mesoscopic timetable generation model to balance complexity and practicality –** Allowing for greater adaptability in railway operations while capturing variations in capacity across different network settings.
- **Revising crossline train proportions in China HSR**: Challenging the widely cited 20-40% cross-line train threshold (Li, 2020), by demonstrating that only opposing train movements significantly impact node station capacity, while following movements have a minor effect.
- **Optimising stop additions:** Demonstrating that in scenarios where line plans already meet the minimum service frequency requirement, additional stopping patterns can improve capacity without compromising direct OD services (detailed in Section 5.1).

6.3 Implication and advice

This section refines the practical implications of this thesis for OD group management, railway operation, and infrastructure management by translating key research findings into actionable recommendations.

1. OD Group management

To improve effectiveness of demand forecasting and passenger experience evaluation, it is recommended that both national and regional stated preference (SP) surveys be conducted. These surveys would complement the revealed preference (RP) data currently used in this thesis, which captures actual travel behaviour but does not account for potential demand that might be overlooked due to existing supply limitations. As demonstrated in Chapter 4, integrating line planning with the ticketing process enables more precise seat allocation strategies, ensuring that OD demand characteristics directly inform service design.

However, it is essential to ensure that during the survey design process, passengers are provided with appropriate transport knowledge. This helps prevent passengers from making stated but unrealistic expectations and demands due to a lack of understanding of railway operations.

2. Railway operation management

For timetable generation, the prioritisation of long-distance and cross-line trains—especially those covering extensive routes—should be carefully re-evaluated. Instead of allocating time-space slots to all these trains from the outset, priority should be given to high-demand corridors, such as the Beijing-Shanghai and Beijing-Guangzhou-Shenzhen-HK HSR lines. Once time slots for these busy sections are secured, long-distance and cross-line services can be integrated into the schedule hierarchically. Additionally, suitable dwell time extensions at key node stations can improve operational flexibility by:

- Facilitating efficient transfers, thereby reducing excessive interchange waiting times;
- Enabling train overtaking, which minimises schedule disruptions caused by speed differentials between different service levels;
- Reducing interdependencies between parts of an interconnected of the service network, mitigating operational coupling effects.

Direct train services remain preferable if passenger demand is sufficiently large, even if they introduce opposing route conflicts at node stations. Conversely, in less congested railway corridors, long-distance services should be integrated with regional HSR networks, increasing intermediate stops to enhance direct service accessibility.

3. Railway infrastructure management

At busy node stations, track layout modifications should prioritise:

- Adding extra platform tracks or sidings to accommodate train overtaking;
- Reducing the number of cross-line trains with opposing directional movements, which currently cause unnecessary congestion by occupying multiple platform tracks simultaneously.

For other highly congested HSR parts (i.e. lines, sections, or corridors), increasing the number of open tracks is advisable to prevent excessive route conflicts at major interchange stations. In high-demand HSR cities in China, developing multiple stations with directional specialisation should be encouraged. This approach would alleviate congestion at single terminals, distribute HSR train traffic more evenly, and reduce station-area bottlenecks that negatively impact capacity utilisation.

6.4 Limitation and future works

1. Passenger preference assumptions:

The assumptions regarding passenger preferences in this thesis are based on available literature and inferred data. However, these assumptions require firmer empirical evidence to enhance the reliability. Future research should focus on conducting comprehensive surveys (e.g., SP investigation) and studies to gather richer data that can validat and refine these assumptions, and accurately reflect passenger behaviour in different contexts.

2. Passenger OD assignment and transfer consideration:

This thesis does not consider a detailed passenger OD assignment (distribution), and the passenger transfers between HSR stations within major Chinese HSR cities. The lack of detailed OD assignment limits the ability to fully understand and model passenger flow patterns. Additionally, the exclusion of transfer dynamics within the HSR station overlooks a critical aspect of passenger experience. Future research should incorporate detailed OD data to examine transfer behaviours within these China major HSR cities and develop a more comprehensive model.

3. Comprehensive 'value' determination method:

The comprehensive 'value' determination method could be further extended and applied to individual HSR train and passenger travel time. This approach would enable a more detailed analysis of each train's contribution to the overall network, providing insights into optimising railway supply and improving passenger overall satisfaction. Future research should explore the application of this method on a per-train basis, potentially leading to more practical recommendations for service improvements.

4. Advanced model structures and efficient algorithms:

The model structure and algorithms used in this thesis, while effective for the scope of this research, could be enhanced to improve efficiency and scalability. More advanced model structures, coupled with effective algorithms, could provide faster and more accurate solutions to the complex problems addressed in this study. Future work should explore the development and implementation of these advanced techniques, with the aim of improving computational efficiency and expanding the applicability of the model to larger and more complex rail networks.

Appendix A China and Japan HSR density map

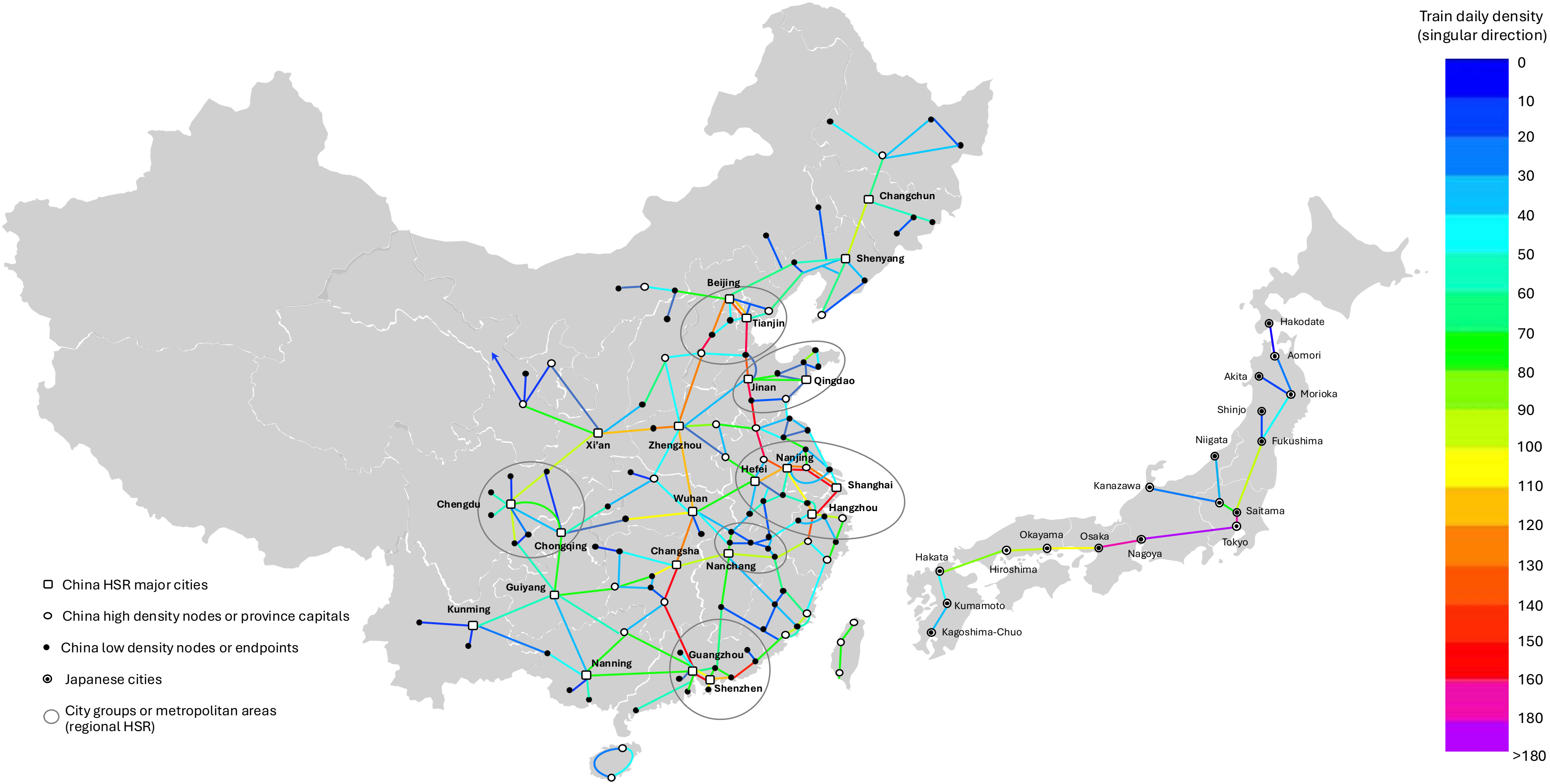
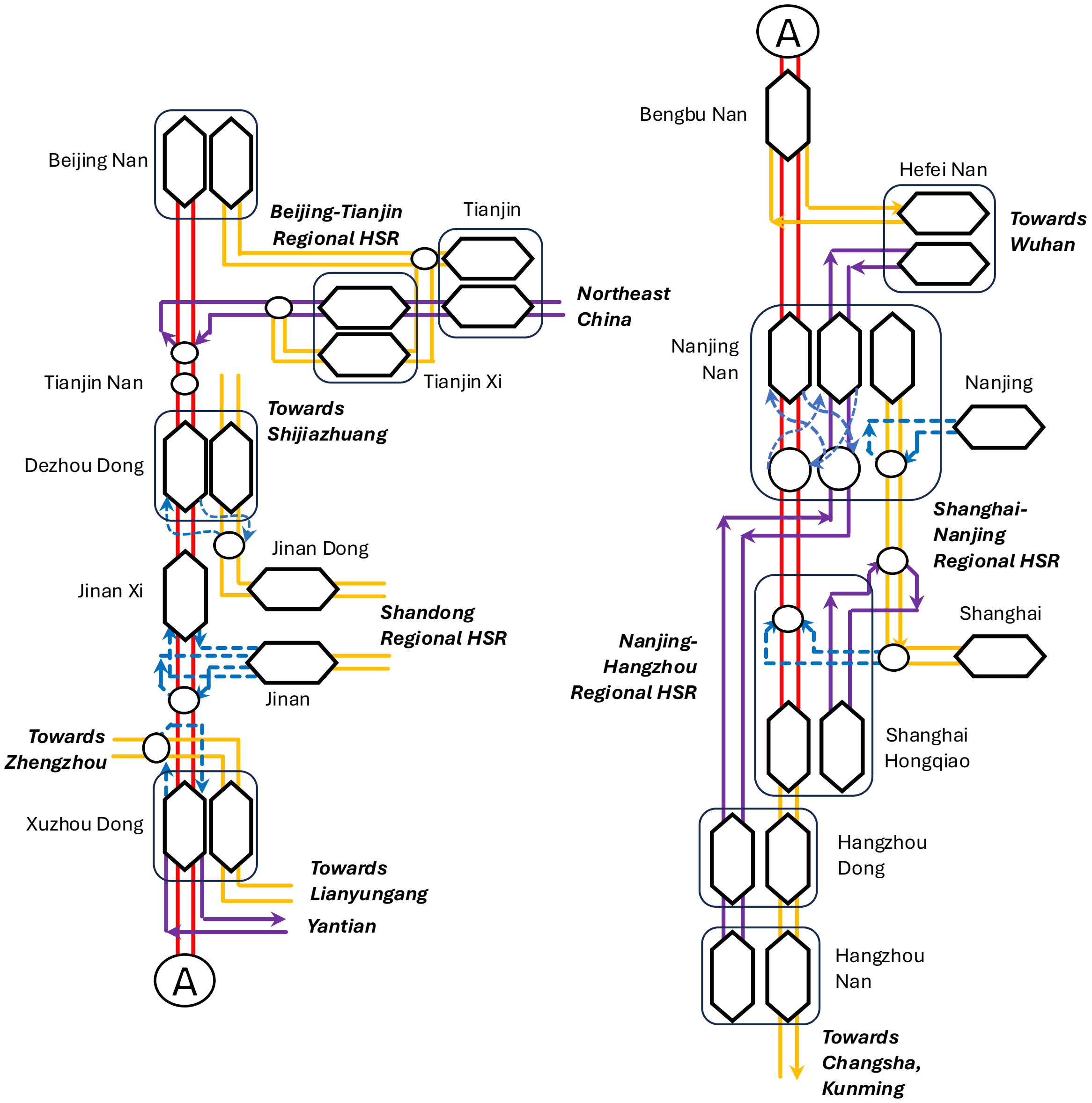


Figure source revised from: (Li & Pi, 2024)

Appendix B Track layouts of Beijing-Shanghai HSR-based network



The figure is graphed based on Google map.

Appendix C Beijing-Shanghai HSR line data

Station index	Station name	Platform number each side	Section index	Distance (km)	Pure running time cost (unit: s)	
					350km/h	300km/h
s_1	Beijing Nan	12	-	0	0	0
s_2	Langfang	1	e_1	60	930	975
s_3	Tianjin Nan	2	e_2	62	675	176
s_4	Cangzhou Xi	2	e_3	88	945	1065
s_5	Dezhou Dong	2	e_4	104	1095	1275
s_6	Jinan Xi	6	e_5	92	1035	1155
s_7	Tai'an	2	e_6	59	645	750
s_8	Qufu	2	e_7	70	735	855
s_9	Tengzhou Dong	1	e_8	56	600	675
s_{10}	Zaozhuang	2	e_9	36	375	435
s_{11}	Xuzhou Dong	6	e_{10}	65	675	795
s_{12}	Suzhou Dong	2	e_{11}	68	705	825
s_{13}	Bengbu Nan	4	e_{12}	88	930	1065
s_{14}	Dingyuan	1	e_{13}	54	585	660
s_{15}	Chuzhou	2	e_{14}	62	675	765
s_{16}	Nanjing Nan	4	e_{15}	59	780	810
s_{17}	Zhengjiang Nan	2	e_{16}	65	795	870
s_{18}	Danyang Bei	1	e_{17}	32	315	360
s_{19}	Changzhou Bei	2	e_{18}	33	435	390
s_{20}	Wuxi Dong	2	e_{19}	57	600	705
s_{21}	Suzhou Bei	2	e_{20}	27	285	330
s_{22}	Kunshan Nan	2	e_{21}	31	345	390
s_{23}	Shanghai Hongqiao	19	e_{22}	50	615	735

Appendix D Detected daily ODs from Beijing to Shanghai HSR (2018)

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	3200	3300	2900	2500	1340	1600	1600	450	720	6600	420	420	50	200	6500	700	100	1000	1000	2500	350	13900
2		350	210	140	680	30	30	50	30	100	30	20	40	30	300	30	40	30	180	140	70	700
3			300	500	930	400	400	400	500	300	200	500	900	300	600	400	400	500	400	360	300	4000
4				80	380	100	120	30	150	350	30	30	40	60	350	120	80	40	50	100	50	400
5					400	40	140	80	100	340	40	60	50	70	400	80	160	60	90	200	80	450
6						400	370	360	350	3600	370	200	350	500	1280	100	260	240	300	250	120	1800
7							40	50	20	600	40	50	60	20	400	20	70	60	80	180	160	570
8								30	30	400	60	40	30	20	200	30	60	80	70	200	160	1000
9									100	250	30	20	10	10	500	10	10	10	10	180	40	220
10										600	10	50	20	10	600	20	30	10	20	100	50	800
11											200	400	200	90	3500	130	100	120	150	380	220	9000
12												10	20	16	500	20	10	5	10	20	10	470
13													200	300	1400	40	80	90	100	250	210	1300
14														10	400	30	10	40	40	70	10	150
15															800	30	50	50	30	120	60	400
16																140	90	180	240	400	300	9000
17																	10	190	130	190	100	1800
18																		20	10	30	20	130
19																			140	240	160	1300
20																				300	260	1900
21																					240	1860
22																						1260

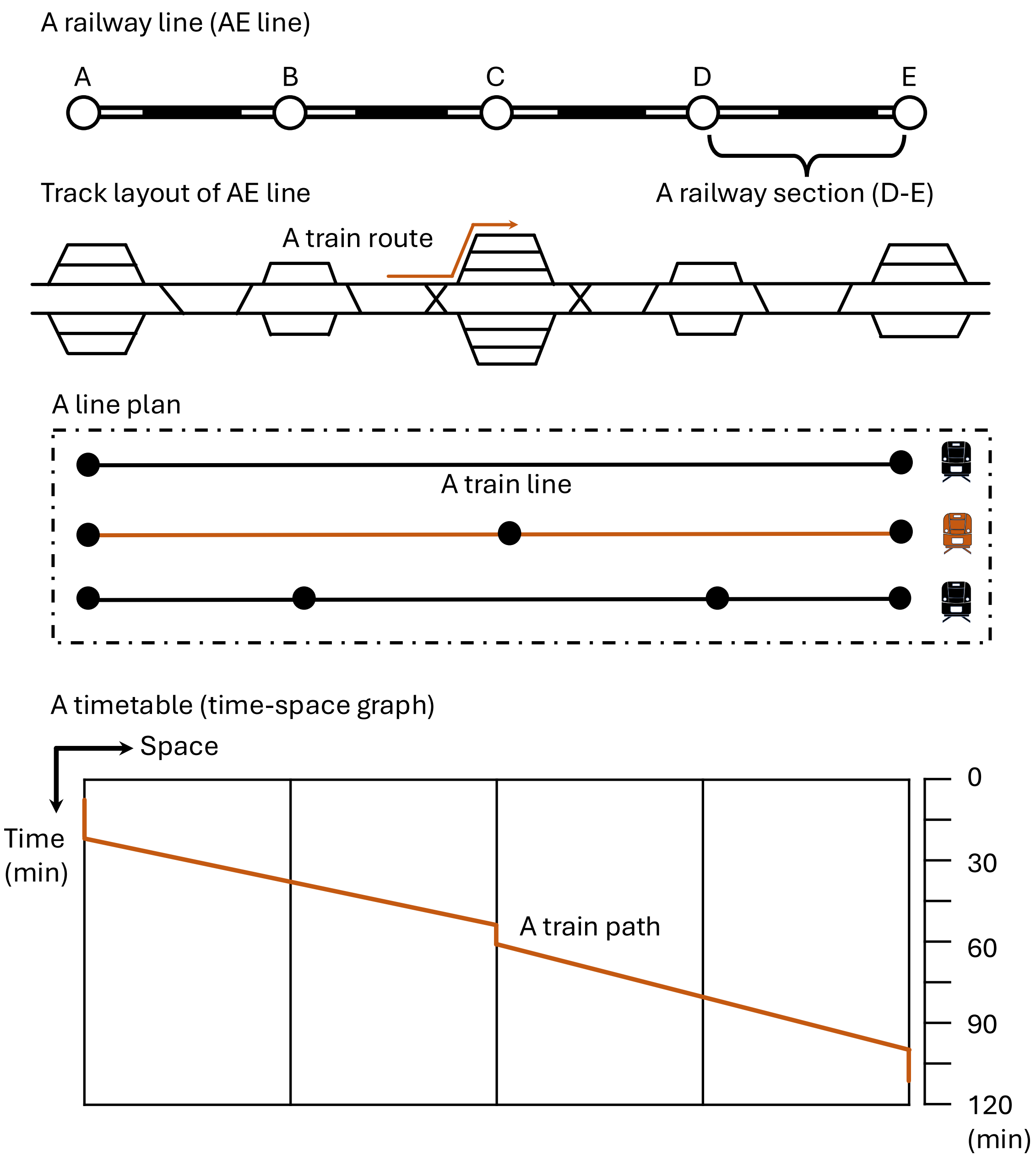
*The first column represents the origin (O), the firs row represents the destination (D), the station code starts with Beijing Nan (s_1)

Appendix E Daily Service Frequency from Beijing to Shanghai HSR (2018)

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	21	54	33	60	93	24	22	15	19	54	15	21	2	14	58	11	5	13	15	22	6	43
2		7	10	10	18	4	3	3	3	9	3	4	1	1	5	0	1	2	2	3	0	4
3			24	48	66	11	18	10	11	39	7	15	1	10	34	8	3	9	12	18	6	24
4				24	51	18	12	11	13	34	10	10	3	8	31	11	3	7	11	15	4	23
5					75	14	17	11	14	42	9	15	1	10	40	7	4	10	14	19	5	28
6						41	44	23	38	91	22	38	3	19	84	18	6	22	28	36	10	61
7							10	10	16	37	8	18	1	6	29	9	2	8	11	14	3	22
8								4	14	38	9	18	2	10	30	8	2	9	11	13	3	22
9									14	3	5	18	4	1	7	8	5	5	13	14	3	5
10										33	8	13	0	7	27	8	3	12	8	15	4	20
11											24	43	4	21	74	16	3	22	30	34	11	52
12												3	11	28	4	1	10	11	13	6	18	3
13													4	6	3	2	6	5	7	4	8	4
14														29	8	5	11	11	17	6	21	29
15															32	9	42	51	60	61	118	32
16																32	9	42	51	60	61	118
17																	1	17	25	24	8	32
18																		5	4	8	2	9
19																			27	35	12	39
20																				33	19	47
21																					18	56
22																						63

*The first column represents the origin (O), the firs row represents the destination (D), the station code starts with Beijing Nan (s_1)

Glossary of Terms



Railway Infrastructure

1. A railway line, in most cases, is a pair of tracks. Each track serves a singular direction, A railway line connects two endpoints (stations and/or cities) or forms a circle containing intermediate stations. Railway lines can be named by planning files or railway operators.
2. A railway section is a part of the railway line linking two any mentioned nodes (i.e. stations and/or junctions).
3. A railway corridor refers to several railway lines connecting two same endpoints (cf. UIC 406 corridor: a strategic, high-capacity railway route, often emphasising major revenue-generating axes or international/transnational freight and passenger routes.)

4. A block section is the basic track circuit unit for signalling control.
5. A railway network consists of railway lines connected by included nodes.
6. A node station is the station connecting at least two railway lines in the network.
7. A platform track positioned alongside a passenger platform for train dwell behaviours.
8. A siding is set for train dwell behaviours without adjacent passenger platforms.
9. A switch is a track facility for a train direction change.
10. A switch area is a group of switches set next to stations.
11. A (flyover/flat) junction is a location several railway lines meet, cross, or diverge, which can be isolated part far away from the station.
12. A connection track links different mentioned railway lines in China HSR (a flyover junction).
13. A signalling junction is a signalling-controlled operational point positioned at the boundary of block sections, often used to manage traffic movements.
14. The open tracks (main tracks) allow trains can run at their maximum permitted speed. The open tracks are located between stations and/or sidings. In station areas, main tracks can also act as platform tracks for stopping trains, depending on station layouts and rules.
15. Terminal layout VS terminus station (destination station of a train)
16. Through layout VS passing train
17. Block section VS blocking time

Describing train relationship in different contexts

18. Front/rear train: physical grouping (e.g., 8+8 car grouped trains) (cf. join and split operation)
19. Leading/(observed)/following train: control, safety, signalling
20. Predecessor/(observed) /successor train: timetabling
21. Up/down/South/East/West/Northbound trains/directions

Describing trains in railway operation

22. A train line is a train itinerary including a stop plan (i.e. calling point list), containing the originating (i.e. starting, O), terminus (i.e. ending, T) stations, and intermediate stop stations, rolling stock group form (8-car, 16-car), and a frequency.
23. A train path is a line segment set on the time-space graph (timetable) denoting a train line.
24. A train route is a track, switch, and platform allocation of a train path at a mentioned station or junction.
25. Following/opposing conflicts: route conflict types
26. The line pool is a train line set containing train lines as candidates. Each line plan proposal is the subset of the line pool.
27. A stopping train dwells at the mentioned station.
28. A passing train does not stop at the mentioned station.

29. An originating train starts the service at the mentioned station.
30. A terminating train ends the service at the mentioned station.
31. A cross-line train, compared to an original-line train, is a type of train service whose train path needs to occupy several railway lines. A railway line must be defined first when discussing cross-line trains.
32. An original-line train, compared to a cross-line train, is a type of train service whose train path occupies only one railway line. This term is introduced to distinguish such trains from cross-line trains. A railway line must be defined first when discussing original-line trains.
33. The train order is the sequence departing from their origin stations.
34. The train precedence is the possible sequence change when they are running in the network.

Describing passenger

35. Passenger boarding/alighting
36. Transfer refers to passenger movements changing between trains at a station.
37. Interchange refers to the facility, equipment, and train services offered by railway operators to support the transfer.

Describing time

38. A time window represents specific time instances t and τ and denoted as $T = [t, \tau]$.
39. The time cost does not have time stamp attributes and is regarded as a 'cost' (e.g., headway, dwell time, running time).
40. τ is later than t: τ is greater than t, or $t < \tau$. (cf. j is the next station of i: $i < j$.)
41. Travel time is the duration of the journey.
42. Arrival time and departure time are the exact clock times.
43. Travel time preference include arrival time, departure time, and travel time

Bibliography

Abril, M., Barber, F., Ingolotti, L., Salido, M. A., Tormos, P., & Lova, A. (2008). An assessment of railway capacity. *Transportation Research Part E: Logistics and Transportation Review*, *44*(5), 774–806. https://doi.org/10.1016/j.tre.2007.04.001

Anderegg, L., Barber, F., Bessas, A., Büsing, C., Cicerone, S., Clausen, J., Corman, F., D'Angelo, G., D'Ariano, A., Delling, D., Diepen, G., Di Stefano, G., Fischetti, M., Flier, H., Frigioni, D., Gatto, M., Gelashvili, R., Goverde, R. M. P., Graffagnino, T., … Zaroliagis, C. (2009). Robust and Online Large-Scale Optimization. In *Lncs 5868* (Vol. 5868).

Andersson, E. V., Peterson, A., & Törnquist Krasemann, J. (2013). Quantifying railway timetable robustness in critical points. *Journal of Rail Transport Planning & Management*, *3*(3), 95–110. https://doi.org/10.1016/J.JRTPM.2013.12.002

ARAFER. (2016). *The French Passenger Rail Transport Market*.

Armstrong, J., Blainey, S., Preston, J., & Hood, I. (2011). *Developing a CUI-based Approach to Network Capacity Assessment*. *June 2014*, 16–18.

Armstrong, J., & Preston, J. (2017). Capacity utilisation and performance at railway stations. *Journal of Rail Transport Planning & Management*, *7*(3), 187–205. https://doi.org/10.1016/J.JRTPM.2017.08.003

Autorite De Regulatyion de Transports. (2019). *Rail Transport Market in France*.

Ball, M., Barnhart, C., Nemhauser, G., & Odoni, A. (2007). *Handbooks in Operations Research and Management Science* (Vol. 14, Issue C). https://doi.org/10.1016/S0927-0507(06)14001-3

Bešinovic, N. (2017). *Integrated capacity assessment and timetabling models for dense railway networks* [Delft University of Technology]. https://doi.org/10.4233/uuid:9083a9cc-64a1-4676-9134-9f8652d629e0

Bešinović, N. (2020). Resilience in railway transport systems: a literature review and research agenda. *Transport Reviews*, *40*(4), 457–478. https://doi.org/10.1080/01441647.2020.1728419

Besinovic, N., & Goverde, R. M. P. (2015). Robust routing in station areas with reducing capacity utilisation. *Caspt 2015*, 1–36.

Bešinović, N., Goverde, R., Nicholson, G., & Steele, H. (2021). *Review of railway timetabling developments: planning phases, goals and future directions*. *July*, 0–33.

Borndörfer, R., & Lamorgese, L. (2018). *Handbook of Optimization in the Railway Industry* (Vol. 268).

Brännlund, U., Lindberg, P. O., Nõu, A., & Nilsson, J. E. (1998). Railway timetabling using Lagrangian relaxation. *Transportation Science*, *32*(4), 358–369. https://doi.org/10.1287/trsc.32.4.358

Bundesnetzaguntur. (2022). *Railway Market Analysis in Germany*. https://www.bundesnetzagentur.de/EN/Areas/Rail/Companies/Publications/MarketAnalysis/MarketAnalysis-node.html

Burdett, R. L., & Kozan, E. (2006). Techniques for absolute capacity determination in railways. *Transportation Research Part B: Methodological*, *40*(8), 616–632. https://doi.org/10.1016/j.trb.2005.09.004

Burdett, R. L., & Kozan, E. (2010). A disjunctive graph model and framework for constructing new train schedules. *European Journal of Operational Research*, *200*(1), 85–98. https://doi.org/10.1016/j.ejor.2008.12.005

Burggraeve, S., Bull, S. H., Vansteenwegen, P., & Lusby, R. M. (2017). Integrating robust timetabling in line plan optimization for railway systems. *Transportation Research Part C: Emerging Technologies*, *77*, 134–160. https://doi.org/10.1016/J.TRC.2017.01.015

Cacchiani, V., Galli, L., & Toth, P. (2015). A tutorial on non-periodic train timetabling and platforming problems. *EURO Journal on Transportation and Logistics*, *4*(3), 285–320. https://doi.org/10.1007/S13676-014-0046-4

Caimi, G., Fuchsberger, M., Laumanns, M., & Schüpbach, K. (2011). Periodic railway timetabling with event flexibility. *Networks*, *57*(1), 3–18. https://doi.org/10.1002/net.20379

Caimi, G., Kroon, L., & Liebchen, C. (2017). Models for railway timetable optimization: Applicability and applications in practice. *Journal of Rail Transport Planning and Management*, *6*(4), 285–312. https://doi.org/10.1016/j.jrtpm.2016.11.002

Canca, D., De-Los-Santos, A., Laporte, G., & Mesa, J. A. (2019). Integrated Railway Rapid Transit Network Design and Line Planning problem with maximum profit. *Transportation Research Part E: Logistics and Transportation Review*, *127*(September 2018), 1–30. https://doi.org/10.1016/j.tre.2019.04.007

Caprara, A., Fischetti, M., & Toth, P. (2002). Modeling and solving the train timetabling problem. *Operations Research*, *50*(5). https://doi.org/10.1287/opre.50.5.851.362

Caprara, A., Monaci, M., Toth, P., & Guida, P. L. (2006). A Lagrangian heuristic algorithm for a real-world train timetabling problem. *Discrete Applied Mathematics*, *154*(5 SPEC. ISS.), 738–753. https://doi.org/10.1016/j.dam.2005.05.026

Chen, J. (2009). *Research of Stable Timetabling and Evaluation on High-speed Railways*. Beijing Jiaotong University.

Chen, K. (2017). *Research on the Influence of Train Operation Structure on the Section Carrying Capacity for High Speed Railway*. Beijing Jiaotong University.

Cheng, Y. (2018). *Dynamic Transit Assignment Model of Urban*. Tongji University.

China Government. (2020). *Railway Target for 2035 and 2050 of China*. https://www.gov.cn/xinwen/2020-08/13/content_5534685.htm

China Railway. (2020). *Measures for the Administration of the names of railway lines, stations, Bridges and tunnels*. China Railway.

Corman, F., D'Ariano, A., Pacciarelli, D., & Pranzo, M. (2010). A tabu search algorithm for rerouting trains during rail operations. *Transportation Research Part B: Methodological*, *44*(1), 175–192. https://doi.org/10.1016/j.trb.2009.05.004

Corman, F., Goverde, R., & D'Ariano, A. (2009). Rescheduling dense train traffic over complex station interlocking areas. *Robust and Online Large-Scale Optimization : Models and Techniques for Transportation Systems*, 369–386.

D'Ariano, A., Pacciarelli, D., & Pranzo, M. (2007). A branch and bound algorithm for scheduling trains in a railway network. *European Journal of Operational Research*, *183*(2), 643–657. https://doi.org/10.1016/j.ejor.2006.10.034

DB. (2023). *Marktsegmete SPNV*. https://www.dbinfrago.com/web/schienennetz/leistungen/trassen/trassenpreise/Trassenpreissystem/marktsegmente-spnv-11161980#

Dewilde, T., & Sels, P. (2011). Defining robustness of a railway timetable. *… Annual Conference of …*, 1–20.

Diaz de Rivera, A., & Dick, C. T. (2021). Illustrating the implications of moving blocks on railway traffic flow behavior with fundamental diagrams. *Transportation Research Part C: Emerging Technologies*, *123*(November 2020), 102982. https://doi.org/10.1016/j.trc.2021.102982

Dicembre, A., & Ricci, S. (2011). Railway traffic on high density urban corridors: Capacity, signalling and timetable. *Journal of Rail Transport Planning and Management*, *1*(2), 59–68. https://doi.org/10.1016/j.jrtpm.2011.11.001

Dick, C. T. (2023). Influence of traffic complexity on railway gravity hump classification yard performance and capacity. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *237*(5), 576–587. https://doi.org/10.1177/09544097221127064

Drábek, M., Mazel, D., & Pospíšil, J. (2021). High-speed rail network and periodic timetable: A comparative analysis of operational concepts. *Acta Polytechnica CTU Proceedings*, *31*, 5–9. https://doi.org/10.14311/APP.2021.31.0005

Fu, H., Liu, X., Hu, H., Nie, B., & Zhang, K. (2024). Optimization Method for Train Connection Plan in Timetable of High Speed Railway Network. *China Railway ScienceRailway*, *45*(3), 149–158. https://doi.org/10.3969/j.issn.1001-4632.2024.03.14

Fu, H., Nie, L., Meng, L., Sperry, B. R., & He, Z. (2015). A hierarchical line planning approach for a large-scale high speed rail network: The China case. *Transportation Research Part A: Policy and Practice*, *75*(2015), 61–83. https://doi.org/10.1016/j.tra.2015.03.013

Fuchs, F., & Corman, F. (2019). An open toolbox for integrated optimization of public transport. *MT-ITS 2019 - 6th International Conference on Models and Technologies for Intelligent Transportation Systems*, 1–7. https://doi.org/10.1109/MTITS.2019.8883335

Fuchs, F., Trivella, A., & Corman, F. (2022). Enhancing the interaction of railway timetabling and line planning with infrastructure awareness. *Transportation Research Part C: Emerging Technologies*, *142*(July), 103805. https://doi.org/10.1016/j.trc.2022.103805

Gao, G. (2022). *Research on designing theory and method of express/local train mixed operation for regional rapid rail transit*. Beijing Jiaotong University.

Gattermann, P., Großmann, P., Nachtigall, K., & Schöbel, A. (2016). Integrating passengers' routes in periodic timetabling: A SAT approach. *OpenAccess Series in Informatics*, *54*, 3.1-3.15. https://doi.org/10.4230/OASICS.ATMOS.2016.3/-/STATS

Goerigk, M., & Schmidt, M. (2017). Line planning with user-optimal route choice. *European Journal of Operational Research*, *259*(2), 424–436. https://doi.org/10.1016/j.ejor.2016.10.034

Gong, C., Shi, J., Wang, Y., Zhou, H., Yang, L., Chen, D., & Pan, H. (2021). Train timetabling with dynamic and random passenger demand: A stochastic optimization method. *Transportation*

Research Part C: Emerging Technologies, *123*(December 2020), 102963. https://doi.org/10.1016/j.trc.2021.102963

Goverde, R. M. P. (1998). Max-plus algebra approach to railway timetable design. *Proceedings of the International Conference on Computer Aided Design, Manufacture and Operation in The Railway and Other Advanced Mass Transit Systems*, *34*, 339–350.

Goverde, R. M. P. (2005). Punctuality of Railway Operations and Timetable Stability Analysis. In *PhD Thesis, Delft University of Technology* (Issue May).

Goverde, R. M. P. (2007). Railway timetable stability analysis using max-plus system theory. *Transportation Research Part B: Methodological*, *41*(2), 179–201. https://doi.org/10.1016/j.trb.2006.02.003

Goverde, R. M. P., Bešinović, N., Binder, A., Cacchiani, V., Quaglietta, E., Roberti, R., & Toth, P. (2016). A three-level framework for performance-based railway timetabling. *Transportation Research Part C: Emerging Technologies*, *67*, 62–83. https://doi.org/10.1016/j.trc.2016.02.004

Goverde, R. M. P., & Hansen, I. A. (2013). Performance indicators for railway timetables. *IEEE ICIRT 2013 - Proceedings: IEEE International Conference on Intelligent Rail Transportation*, 301–306. https://doi.org/10.1109/ICIRT.2013.6696312

Guo, G. (2015). *Theory and Method of Cyclic Train Connection plan for High Speed Railway*. Beijing Jiaotong University.

Haase, K., & Müller, S. (2014). A comparison of linear reformulations for multinomial logit choice probabilities in facility location models. *European Journal of Operational Research*, *232*(3), 689–691. https://doi.org/10.1016/J.EJOR.2013.08.009

Haitong International Securities Group Limited. (2022). *Beijing-Shanghai High Speed Railway Report*.

Han, P., Nie, L., Fu, H., Gong, Y., & Wang, G. (2019). A multiobjective integer linear programming model for the cross-track line planning problem in the Chinese high-speed railway network. In *Symmetry* (Vol. 11, Issue 5). https://doi.org/10.3390/sym11050670

Harrod, S. (2011). Modeling network transition constraints with hypergraphs. *Transportation Science*, *45*(1), 81–97. https://doi.org/10.1287/trsc.1100.0337

Harrod, S. (2012). A tutorial on fundamental model structures for railway timetable optimization. *Surveys in Operations Research and Management Science*, *17*(2), 85–96. https://doi.org/10.1016/j.sorms.2012.08.002

Hartleb, J., & Schmidt, M. (2022). Railway timetabling with integrated passenger distribution. *European Journal of Operational Research*, *298*(3), 953–966. https://doi.org/10.1016/j.ejor.2021.06.025

Hartleb, J., Schmidt, M., Friedrich, M., & Huisman, D. (2019). A Good or a Bad Timetable: Do Different Evaluation Functions Agree? *SSRN Electronic Journal*, 1–42. https://doi.org/10.2139/ssrn.3368839

Hartleb, J., Schmidt, M., Huisman, D., & Friedrich, M. (2023). Modeling and Solving Line Planning with Mode Choice. *Transportation Science*, *57*(2), 336–350. https://doi.org/10.1287/trsc.2022.1171

Heydar, M., Petering, M. E. H., & Bergmann, D. R. (2013). Mixed integer programming for minimizing the period of a cyclic railway timetable for a single track with two train types. *Computers and Industrial Engineering*, *66*(1), 171–185. https://doi.org/10.1016/j.cie.2013.06.003

Högdahl, J., & Bohlin, M. (2023). A Combined Simulation-Optimization Approach for Robust Timetabling on Main Railway Lines. In *Transportation Science* (Vol. 57, Issue 1). https://doi.org/10.1287/trsc.2022.1158

Hu, H., Yue, Y., Fu, H., & Li, J. (2023). Improving service qualities of cyclic line plans considering heterogeneous passenger origin-destination (OD) flow groups. *Transportation Planning and Technology*, 1–24. https://doi.org/10.1080/03081060.2023.2226117

Hu, H., Yue, Y., Fu, H., & Li, J. (2024). How to optimize train lines for diverse passenger demands: A line planning approach providing matched train services for each O-D market. *Transportation Research Part A: Policy and Practice*, *186*(January), 104154. https://doi.org/10.1016/j.tra.2024.104154

Huisman, T., Boucherie, R. J., & Van Dijk, N. M. (2002). A solvable queueing network model for railway networks and its validation and applications for the Netherlands. *European Journal of Operational Research*, *142*(1), 30–51. https://doi.org/10.1016/S0377-2217(01)00269-7

Iliasov, A., Lopatkin, I., & Romanovsky, A. (2013). The SafeCap platform for modelling railway safety and capacity. *Lecture Notes in Computer Science (Including Subseries Lecture Notes* *in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *8153 LNCS*, 130–137. https://doi.org/10.1007/978-3-642-40793-2_12

International Union of Railways. (2004). *UIC 406* (1st editio).

Isaai, M. T., & Singh, M. G. (2001). Hybrid applications of constraint satisfaction and meta-heuristics to railway timetabling: A comparative study. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, *31*(1), 87–95. https://doi.org/10.1109/5326.923271

ITF. (2019). Efficiency in Railway Operations and Infrastructure Management. *International Transport Forum*, 1–35.

Jamili, A., & Pourseyed Aghaee, M. (2015). Robust stop-skipping patterns in urban railway operations under traffic alteration situation. *Transportation Research Part C: Emerging Technologies*, *61*, 63–74. https://doi.org/10.1016/j.trc.2015.09.013

Japan Railpass. (2023). *Tokyo Rail Map*. https://www.jrailpass.com/blog/tokyo-train-station

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*(April 2023). https://doi.org/10.1016/j.trc.2023.104464

Khadem Sameni, M. (2012). Railway Track Capacity: Measuring and Managing [University of Southampton]. In *University of Southampton*. https://doi.org/10.1016/0041-2678(70)90288-5

Khadem Sameni, M., & Moradi, A. (2022). Railway capacity: A review of analysis methods. *Journal of Rail Transport Planning and Management*, *24*(July), 100357. https://doi.org/10.1016/j.jrtpm.2022.100357

Khan, M. B., & Zhou, X. (2010). Stochastic optimization model and solution algorithm for robust double-track train-timetabling problem. *IEEE Transactions on Intelligent Transportation Systems*, *11*(1), 81–89. https://doi.org/10.1109/TITS.2009.2030588

Kozan, E., & Burdett, R. (2005). A railway capacity determination model and rail access charging methodologies. *Transportation Planning and Technology*, *28*(1), 27–45. https://doi.org/10.1080/0308106052000340378

Lai, Y. C., & Barkan, C. P. L. (2011). Comprehensive Decision Support Framework for Strategic Railway Capacity Planning. *Journal of Transportation Engineering*, *137*(10), 738–749. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000248

Lai, Y. C., Liu, Y. H., & Lin, Y. J. (2015). Standardization of capacity unit for headway-based rail capacity analysis. *Transportation Research Part C: Emerging Technologies*, *57*, 68–84. https://doi.org/10.1016/j.trc.2015.05.013

Lamorgese, L., Mannino, C., & Natvig, E. (2017). An exact micro–macro approach to cyclic and non-cyclic train timetabling. *Omega (United Kingdom)*, *72*, 59–70. https://doi.org/10.1016/j.omega.2016.11.004

Landex, A. (2008). Methods to estimate railway capacity and passenger delays. In *Technical University of Denmark*. Technical University of Denmark.

Landex, A., & Wittrup, L. (2023). Capacity gains with virtual sub-sections in the ETCS Signaling System. *Rail Belgrade 2023*.

Lange, J., & Werner, F. (2018). A comparison of approaches to modeling train scheduling problems as job-shops with blocking constraints. *Journal of Scheduling*, *21*(2), 191–207. https://doi.org/10.1007/s10951-017-0526-0

Lange, J., Werner, F., Kozan, E., Burdett, R. L., Lindfeldt, O., Mussone, L., Wolfler Calvo, R., Abril, M., Barber, F., Ingolotti, L., Salido, M. A., Tormos, P., Lova, A., Yaghini, M., Nikoo, N., Ahadi, H. R., Bešinović, N., Goverde, R. M. P., Quaglietta, E., … Crawford, I. (2011). Mathematical solutions for solving periodic railway transportation. *Transportation Research Part B: Methodological*, *1*(1), 14–32. https://doi.org/10.1155/2009/728916

Lee, J.-H. (2014). Analyzing the Difference between the Stated Preference and the Revealed Preference before/after the High-speed Rail Service in Korea. *International Journal of Railway*, *7*(1), 24–33. https://doi.org/10.7782/ijr.2014.7.1.024

Lestari, P. D. R., Liu, R., & Batley, R. (2024). Emerging Trends in Railway Line Planning and Timetabling: A Comprehensive Review. *University Transport Study Group Annual Conference*.

Leutwiler, F., & Corman, F. (2023). A review of principles and methods to decompose large-scale railway scheduling problems. *EURO Journal on Transportation and Logistics*, *12*(December 2021), 100107. https://doi.org/10.1016/j.ejtl.2023.100107

Li, J. (2015). *Research on Optimising Cyclic Train Line Planning for High-Speed Railways*. Beijing Jiaotong University.

Li, J. (2019). *Research on Compressed Headways Calculation Method Considering Train Dynamics*. University College London.

Li, J. (2022). *Identification and Releasing Method of High-speed Railway Capacity Bottleneck* [Southwest Jiaotong University]. https://doi.org/10.27414/d.cnki.gxnju.2022.000101

Li, J., Preston, J., Armstrong, J., & Yuan, W. (2023). Evaluating Timetables ’ Capacity Utilisation with An Extended Event-activity Network Method. *Rail Belgrade 2023*, 84.

Li, L., & Pi, D. (2024). *High-speed Train Density of China and Japan*. https://mp.weixin.qq.com/s/ljeP0pR00LjG9MK7EQ7oTQ

Li, T. (2018). *Model and Algorithm for Integrated Optimization on yclic Highspeed Railway Connection Planning Problem*. Beijing Jiaotong University.

Li, W. (2019). *Research on Theory and Method of High-speed Railway EMU circulation plan Optimization*. Beijing Jiaotong University.

Li, Y. (2020). *Capacity Calculation and Utilization of High-speed Railway Corridor*. Tongji University.

Liao, Z. (2021). *Resource-Based Railway Capacity Theory and Estimation Method* [Beijing Jiaotong University]. https://doi.org/10.26944/d.cnki.gbfju.2021.000301.

Liao, Z., Li, H., Miao, J., & Corman, F. (2021). Railway capacity estimation considering vehicle circulation: Integrated timetable and vehicles scheduling on hybrid time-space networks. *Transportation Research Part C: Emerging Technologies*, *124*(December 2020), 102961. https://doi.org/10.1016/j.trc.2020.102961

Liao, Z., Li, H., Miao, J., & Meng, L. (2024). Estimating the Railway Network Capacity Utilization with Mixed Train Routes and Stopping Patterns: A Multiobjective Optimization Approach. *Journal of Advanced Transportation*, *2024*. https://doi.org/10.1155/2024/5467767

Liao, Z., Miao, J., Meng, L., & Li, H. (2020). A resource-oriented decomposition approach for train timetabling problem with variant running time and minimum headway. *Transportation Letters*, *00*(00), 1–14. https://doi.org/10.1080/19427867.2020.1824310

Liebchen, C., & Möhring, R. H. (2008). The modeling power of the periodic event scheduling problem: Railway timetables - and beyond. *Lecture Notes in Economics and Mathematical Systems*, *600*, 117–150. https://doi.org/10.1007/978-3-540-73312-6_7

Lindfeldt, A. (2015). *Railway capacity analysis delays and infrastructure Anders Lindfeldt*.

Liu, J. (2012). *Transfer-based Flow Assignment Modeling for Urban Rail Transit Network* [Beijing Jiaotong University]. https://doi.org/10.7666/d.Y2221994

Liu, S. Q., & Kozan, E. (2009). Scheduling trains as a blocking parallel-machine job shop scheduling problem. *Computers and Operations Research*, *36*(10), 2840–2852. https://doi.org/10.1016/j.cor.2008.12.012

Löbel, F., Lindner, N., & Borndörfer, R. (2020). *The Restricted Modulo Network Simplex Method for Integrated Periodic Timetabling and Passenger Routing*. 757–763. https://doi.org/10.1007/978-3-030-48439-2_92

Lu, B. (2023). *High-speed EMU tenders 55 units for the second time in the year*. China Buiness Journal. https://new.qq.com/rain/a/20231128A03RP300

Lu, G., Ning, J., Liu, X., & Nie, Y. (Marco). (2022). Train platforming and rescheduling with flexible interlocking mechanisms: An aggregate approach. *Transportation Research Part E: Logistics and Transportation Review*, *159*(February), 102622. https://doi.org/10.1016/j.tre.2022.102622

Lu, Y. (2021). *Research on Cross-line Passenger Flow Assignment Oriented to Capacity Releasing of High-speed Railway Bottleneck Section*. Beijing Jiaotong University.

Luijt, R. S., van den Berge, M. P. F., Willeboordse, H. Y., & Hoogenraad, J. H. (2017). 5 years of Dutch eco-driving: Managing behavioural change. *Transportation Research Part A: Policy and Practice*, *98*, 46–63. https://doi.org/10.1016/J.TRA.2017.01.019

Lusby, R. M., Larsen, J., & Bull, S. (2018). A survey on robustness in railway planning. *European Journal of Operational Research*, *266*(1), 1–15. https://doi.org/10.1016/j.ejor.2017.07.044

Ma, L. (2015). *Research of High Speed Railway Capacity Simulation Based on Cellular Automata*. Lanzhou Jiaotong University.

Ma, P. (2018). *Research on the Influence on Passenger Flow by Improving the Service Level of High-Speed Railway Line Planning*. Beijing Jiaotong University.

Mei, Y., Gu, W., Cassidy, M., & Fan, W. (2021). Planning skip-stop transit service under heterogeneous demands. *Transportation Research Part B: Methodological*, *150*(July), 503–523. https://doi.org/10.1016/j.trb.2021.06.008

Meng, L., & Zhou, X. (2014). Simultaneous train rerouting and rescheduling on an N-track network: A model reformulation with network-based cumulative flow variables. *Transportation Research Part B: Methodological*, *67*, 208–234. https://doi.org/10.1016/j.trb.2014.05.005

Meng, L., & Zhou, X. (2019). An integrated train service plan optimization model with variable demand: A team-based scheduling approach with dual cost information in a layered network. *Transportation Research Part B: Methodological*, *125*, 1–28. https://doi.org/10.1016/j.trb.2019.02.017

Ministry of Transport of the People's Republic of China. (2020). *Government Year Report 2019*.

Mussone, L., & Wolfler Calvo, R. (2013). An analytical approach to calculate the capacity of a railway system. *European Journal of Operational Research*, *228*(1), 11–23. https://doi.org/10.1016/j.ejor.2012.12.027

NetworkRail. (2013). *Long Term Planningh Process: Long Distance Market Study*.

NetworkRail. (2020). *Unblocking the Croydon bottleneck – The Croydon Area Remodelling Scheme*. https://www.networkrail.co.uk/running-the-railway/our-routes/sussex/upgrading-the-brighton-main-line/unblocking-the-croydon-bottleneck/

Nie, L. (2019). *Theory and Approaches of Cyclic Train Timetable for Large Scale Complex High-speed Railway Network*.

Nie, L., Zhu, J., Zhang, Y., & Li, W. (2021). *High-speed Railway Transport Organisation*. China Railway Publishing House Co., Ltd.

Ning, B., Tao, T., K, Q., C, G., & Q, Wang. (2009). CTCS—Chinese Train Control System. In *WIT Transactions on State of the Art in Science and Engineering* (Vol. 34). https://doi.org/10.2495/978-1-84564

Niu, H. M. (2021). Literature Review on Rail Train Timetabling. *Jiaotong Yunshu Xitong Gongcheng Yu Xinxi/Journal of Transportation Systems Engineering and Information Technology*, *21*(5), 114–124. https://doi.org/10.16097/j.cnki.1009-6744.2021.05.011

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. https://doi.org/10.1016/j.trc.2013.08.016

Odijk, M. A. (1996). A constraint generation algorithm for the construction of periodic railway timetables. *Transportation Research Part B: Methodological*, *30*(6), 455–464. https://doi.org/10.1016/0191-2615(96)00005-7

OECD/ITF. (2014). Valuing Convenience in Public Transport. In *Valuing Convenience in Public Transport*.

ORR. (2023). *ORR Data Portal*.

Ortúzar, J. de D., & Willumsen, L. G. (2011). Modelling Transport. In *Modelling Transport*. John Wiley and Sons. https://doi.org/10.1002/9781119993308

Parbo, J., Nielsen, O. A., & Prato, C. G. (2016). Passenger Perspectives in Railway Timetabling: A Literature Review. *Transport Reviews*, *36*(4), 500–526. https://doi.org/10.1080/01441647.2015.1113574

Parbo, J., Nielsen, O. A., & Prato, C. G. (2018). Reducing passengers' travel time by optimising stopping patterns in a large-scale network: A case-study in the Copenhagen Region. *Transportation Research Part A: Policy and Practice*, *113*(February), 197–212. https://doi.org/10.1016/j.tra.2018.04.012

Pätzold, J., & Schöbel, A. (2016). A matching approach for periodic timetabling. *OpenAccess Series in Informatics*, *54*, 1.1-1.15. https://doi.org/10.4230/OASICS.ATMOS.2016.1/-/STATS

Peeters, L. (2003). Cyclic Railway Timetable Optimization [Elsevier Ltd]. In *Transportation Research Part B: Methodological* (Vol. 128, Issue 1). https://doi.org/10.1016/j.trb.2019.08.001

Peng, W., Teng, J., & Wang, H. (2024). Understanding Heterogeneous Passenger Route Choice in Municipal Rail Transit with Express and Local Trains: An Empirical Study in Shanghai. *Urban Rail Transit*, *10*(2), 122–143. https://doi.org/10.1007/s40864-024-00214-8

Pengpai News. (2024). *The price hike of high-speed trains has sparked a heated debate over whether to improve service quality or increase travel costs*. China Central Television. https://news.cctv.com/2024/05/05/ARTImyxdIJMSt0peWuAl6eqA240505.shtml

Petering, M. E. H., Heydar, M., & Bergmann, D. R. (2016). Mixed-integer programming for railway capacity analysis and cyclic, combined train timetabling and platforming. *Transportation Science*, *50*(3), 892–909. https://doi.org/10.1287/trsc.2015.0652

Pinedo, M. (2007). Scheduling Theory, Algorithms, and Systems. In *Springer Optimization and Its Applications*. Springer US. https://doi.org/10.1007/978-0-387-78935-4

Polinder, G.-J., Cacchiani, V., Schmidt, M., & Huisman, D. (2020). An Iterative Heuristic for Passenger-Centric Train Timetabling with Integrated Adaption Times. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.3629576

Pouryousef, H., & Lautala, P. (2015). Hybrid simulation approach for improving railway capacity and train schedules. *Journal of Rail Transport Planning and Management*, *5*(4), 211–224. https://doi.org/10.1016/j.jrtpm.2015.10.001

Quaglietta, E., Wang, M., & Goverde, R. M. P. (2020). A multi-state train-following model for the analysis of virtual coupling railway operations. *Journal of Rail Transport Planning and Management*, *15*(November 2019), 100195. https://doi.org/10.1016/j.jrtpm.2020.100195

Rabot, A. (2023). *Train Lines of UK Train Operation Companies*. https://drive.google.com/drive/folders/1cZ0-_dELkFm738fLrs0wjUtk_3mzm_TR

Rail Net Europe. (2021). *Capacity Model*. https://rne.eu/sales-timetabling/

Railway Delivery Group. (2018). *Passenger Demand Forecasting Council*.

Robenek, T., Maknoon, Y., Azadeh, S. S., Chen, J., & Bierlaire, M. (2016). Passenger centric train timetabling problem. *Transportation Research Part B: Methodological*, *89*, 107–126. https://doi.org/10.1016/j.trb.2016.04.003

Rosell, F., & Codina, E. (2020). A model that assesses proposals for infrastructure improvement and capacity expansion on a mixed railway network. *Transportation Research Procedia*, *47*, 441–448. https://doi.org/10.1016/j.trpro.2020.03.119

Ryes, P., Danaher, A., Walker, M., Nichols, F., Carter, B., Ellis, E., Cherrington, L., & Bruzzone, A. (2013). Transit Capacity and Quality of Service Manual, Third Edition. In *Transit Capacity and Quality of Service Manual, Third Edition* (Issue January 2013). https://doi.org/10.17226/24766

Schiewe, P. (2020). Integrated optimization in public transport planning. In *Springer Optimization and Its Applications* (Vol. 160).

Schlechte, T. (2012). *Railway track allocation: Models and methods* [Technische Universität Berlin]. https://doi.org/https://doi.org/10.14279/depositonce-3124

Schlechte, T., Borndörfer, R., Erol, B., Graffagnino, T., & Swarat, E. (2011). Micro-macro transformation of railway networks. *Journal of Rail Transport Planning and Management*, *1*(1), 38–48. https://doi.org/10.1016/j.jrtpm.2011.09.001

Schmidt, M., & Schöbel, A. (2024). *Planning and Optimizing Transit Lines*.

Schöbel, A. (2012). Line planning in public transportation: Models and methods. *OR Spectrum*, *34*(3), 491–510. https://doi.org/10.1007/s00291-011-0251-6

Sels, P., Meisch, K., Møller, T., Parbo, J., Dewilde, T., Cattrysse, D., Vansteenwegen, P.(2015). *Towards a Better Train Timetable for Denmark Reducing Total Expected Passenger Time*.

Shih, M. C., Dick, C. T., & Barkan, C. P. L. (2015). Impact of passenger train capacity and level of service on shared rail corridors with multiple types of freight trains. *Transportation Research Record*, *2475*, 63–71. https://doi.org/10.3141/2475-08

Sparing, D., & Goverde, R. M. P. (2013). An optimization model for periodic timetable generation with dynamic frequencies. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, *Itsc*, 785–790. https://doi.org/10.1109/ITSC.2013.6728327

Sparing, D., & Goverde, R. M. P. (2017). A cycle time optimization model for generating stable periodic railway timetables. *Transportation Research Part B: Methodological*, *98*(434), 198–223. https://doi.org/10.1016/j.trb.2016.12.020

Sun, Y., Jiang, Z., Gu, J., Zhou, M., Li, Y., & Zhang, L. (2018). Analyzing high speed rail passengers' train choices based on new online booking data in China. *Transportation Research Part C: Emerging Technologies*, *97*, 96–113. https://doi.org/10.1016/J.TRC.2018.10.015

Szpigel, B. (1973). Optimal train scheduling on a single line railway. *Operational Research*, 343–352.

Tian, Y. (2018). *Research on Cyclic Train Line Planning for High Speed Rail Based on Line Pool of Train Stop Pattern*. Beijing Jiaotong University.

Train, K. E. (2002). *Discrete Choice Methods with Simulation*.

Transportation Research Board of the National Academies. (2013). Transit Capacity and Quality of Service Manual, Third Edition. In *National Academies Press* (3rd ed.). https://doi.org/10.17226/24766

Trenolab. (2022). *Trenolab*. https://www.trenolab.com/about/

UIC. (2013). *Capacity* (Issue 2nd edition).

UIC. (2015). *HIgh speed rail: fast track to sustainble mobility*.

UIC. (2018). A global vision for railway development. In *Dairy Industries International* (Vol. 83, Issue 5). https://doi.org/10.1038/443148a

UIC. (2022). High-Speed Rail Atlas 2021. In *International Union of Railway*.

van Wee, B., Bliemer, M., Steg, L., & Verhoef, E. (2008). Pricing in Road Transport: A Multi-Disciplinary Perspective. In *Pricing in Road Transport: A Multi-Disciplinary Perspective*. Edward Elgar Publishing. https://doi.org/10.5040/9781501333040.ch-c

Vansteenwegen, P., & Oudheusden, D. Van. (2006). Developing railway timetables which guarantee a better service. *European Journal of Operational Research*, *173*(1), 337–350. https://doi.org/10.1016/j.ejor.2004.12.013

Vieira, A. P., Christofoletti, L. M., & Vilela, P. R. S. (2018). *Analyzing Railway Capacity Using a Planning Tool*. https://doi.org/10.1115/JRC2018-6160

Wang, E., Yang, L., Li, P., Zhang, C., & Gao, Z. (2023). Joint optimization of train scheduling and routing in a coupled multi-resolution space–time railway network. *Transportation Research Part C: Emerging Technologies*, *147*(December 2021). https://doi.org/10.1016/j.trc.2022.103994

Wang, J., Huang, J., & Jing, Y. (2020). Competition between high-speed trains and air travel in China: From a spatial to spatiotemporal perspective. *Transportation Research Part A: Policy and Practice*, *133*(December 2019), 62–78. https://doi.org/10.1016/j.tra.2019.12.030

Wang, P., & Goverde, R. M. P. (2016). Multiple-phase train trajectory optimization with signalling and operational constraints. *Transportation Research Part C: Emerging Technologies*, *69*, 255–275. https://doi.org/10.1016/j.trc.2016.06.008

Wang, Y., Song, R., He, S., Song, Z., & Chi, J. (2022). Integrated train routing and timetabling problem in a multi-station high-speed railway hub. *International Journal of Rail Transportation*, *00*(00), 1–40. https://doi.org/10.1080/23248378.2022.2093280

Wang, Y., Tang, T., Ning, B., van den Boom, T. J. J., & De Schutter, B. (2015). Passenger-demands-oriented train scheduling for an urban rail transit network. *Transportation Research Part C: Emerging Technologies*, *60*, 1–23. https://doi.org/10.1016/J.TRC.2015.07.012

Wardman, M. (2001). A review of British evidence on time and service quality valuations. *Transportation Research Part E: Logistics and Transportation Review*, *37*(2–3), 107–128. https://doi.org/10.1016/S1366-5545(00)00012-0

Wardman, M. (2012). Review and meta-analysis of U.K. time elasticities of travel demand. *Transportation*, *39*(3), 465–490. https://doi.org/10.1007/S11116-011-9369-2/TABLES/9

Wardman, M. (2022). Meta-analysis of British time-related demand elasticity evidence: An update. *Transportation Research Part A: Policy and Practice*, *157*(February), 198–214. https://doi.org/10.1016/j.tra.2022.02.001

Weik, N., Niebel, N., & Nießen, N. (2016). Capacity analysis of railway lines in Germany – A rigorous discussion of the queueing based approach. *Journal of Rail Transport Planning and Management*, *6*(2), 99–115. https://doi.org/10.1016/j.jrtpm.2016.06.001

Wheat, P., & Wardman, M. (2017). Effects of timetable related service quality on rail demand. *Transportation Research Part A: Policy and Practice*, *95*, 96–108. https://doi.org/10.1016/j.tra.2016.11.009

Wong, R. C. W., Yuen, T. W. Y., Fung, K. W., & Leung, J. M. Y. (2008). Optimizing Timetable Synchronization for Rail Mass Transit. *Https://Doi.Org/10.1287/Trsc.1070.0200*, *42*(1), 57–69. https://doi.org/10.1287/TRSC.1070.0200

Wu, H. (2020). *Study on the High-speed Railways Capacity Utilisation Based on Stopping Plan Optimisation*. Southwest JIaotong University.

Xu, W. (2020). *Research on Optimization of High-speed Railway Cross-line Train Operation Based on Capacity Utilization Improvement*. Beijing Jiaotong University.

Xu, X., Li, C. L., & Xu, Z. (2021). Train timetabling with stop-skipping, passenger flow, and platform choice considerations. *Transportation Research Part B: Methodological*, *150*(April), 52–74. https://doi.org/10.1016/j.trb.2021.06.001

Yan, F., Bešinović, N., & Goverde, R. M. P. (2019). Multi-objective periodic railway timetabling on dense heterogeneous railway corridors. *Transportation Research Part B: Methodological*, *125*, 52–75. https://doi.org/10.1016/j.trb.2019.05.002

Yan, F., & Goverde, R. M. P. (2017). Railway timetable optimization considering robustness and overtakings. *5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems, MT-ITS 2017 - Proceedings*, 291–296. https://doi.org/10.1109/MTITS.2017.8005683

Yan, F., & Goverde, R. M. P. (2019). Combined line planning and train timetabling for strongly heterogeneous railway lines with direct connections. *Transportation Research Part B: Methodological*, *127*, 20–46. https://doi.org/10.1016/j.trb.2019.06.010

Yang, H., Dobruszkes, F., Wang, J., Dijst, M., & Witte, P. (2018). Comparing China's urban systems in high-speed railway and airline networks. *Journal of Transport Geography*, *68*(April), 233–244. https://doi.org/10.1016/j.jtrangeo.2018.03.015

Yang, H., Nie, L., Zhou, L., & Le, Yi. (2011). *Railway Operation and Organisation*. China Railway Publishing House Co., Ltd.

Yao, Z., Nie, L., & Fu, H. (2024). Railway line planning with passenger routing: Direct-service network representations and a two-phase solution approach. *Transportation Research Part B: Methodological*, *June*. https://doi.org/10.1016/j.trb.2024.102989

Yao, Z., Nie, L., & He, Z. (2022). A genetic algorithm for heterogeneous high-speed railway timetabling with dense traffic: The train-sequence matrix encoding scheme. *Journal of Rail Transport Planning & Management*, *23*, 100334. https://doi.org/10.1016/J.JRTPM.2022.100334

Yao, Z., Nie, L., Yue, Y., He, Z., Ke, Y., Mo, Y., & Wang, H. (2023). Network periodic train timetabling with integrated stop planning and passenger routing: A periodic time–space network construct and ADMM algorithm. *Transportation Research Part C: Emerging Technologies*, *153*(July 2022), 104201. https://doi.org/10.1016/j.trc.2023.104201

Yin, J., Pu, F., Yang, L., D'Ariano, A., & Wang, Z. (2023). Integrated optimization of rolling stock allocation and train timetables for urban rail transit networks: A benders decomposition approach. *Transportation Research Part B: Methodological*, *176*, 102815. https://doi.org/10.1016/J.TRB.2023.102815

Yuan, J., & Hansen, I. A. (2007). Optimizing capacity utilization of stations by estimating knock-on train delays. *Transportation Research Part B: Methodological*, *41*(2), 202–217. https://doi.org/10.1016/J.TRB.2006.02.004

Yuan, W. (2020). *Research on the Optimization of Railway Ticket Pre-allocation and Dynamic Pricing*. Beijing Jiaotong University.

Yuan, W., Nie, L., Fu, H., Wu, X., & Ke, Y. (2018). Simulation-Based Method for Seat Inventory Control Optimization under Heterogeneous Customer Behavior: A Case Study of the China Railway. *CICTP 2018: Intelligence, Connectivity, and Mobility - Proceedings of the 18th COTA International Conference of Transportation Professionals*, *October*, 788–798. https://doi.org/10.1061/9780784481523.079

Zhang, J. (2012). *Research on Theory & Method of Capacity Calculation & Evaluation for High Speed Railway*. Beijing Jiaotong University.

Zhang, J. (2015). Analysis on line capacity usage for China high speed railway with optimization approach. *Transportation Research Part A: Policy and Practice*, *77*, 336–349. https://doi.org/10.1016/j.tra.2015.04.022

Zhang, Q. (2019). *The Study of Theory and Methodology of Train Operational Drawing Preparation for High Speed Railway Based on Regional Coordination*. Southwest Jiaotong University.

Zhang, Q. (2022). *Theory and Method for Optimization of Technical Operation in High- speed Railway Station*. Beijing Jiaotong University.

Zhang, Q., Lusby, R. M., Shang, P., & Zhu, X. (2020). Simultaneously re-optimizing timetables and platform schedules under planned track maintenance for a high-speed railway network. *Transportation Research Part C: Emerging Technologies*, *121*(November), 102823. https://doi.org/10.1016/j.trc.2020.102823

Zhang, Q., Lusby, R. M., Shang, P., & Zhu, X. (2022). A heuristic approach to integrate train timetabling, platforming, and railway network maintenance scheduling decisions. *Transportation Research Part B: Methodological*, *158*(February), 210–238. https://doi.org/10.1016/j.trb.2022.02.002

Zhang, R., Ma, Y., Zhao, B. R., & Luan, W. X. (2016). Passenger choice behavior of high-speed rail and airline between Beijing and Shanghai. *Journal of Transportation Systems Engineering and Information Technology*, *16*(1), 223–228.

Zhang, X. (2019). *Theory and Method of Cyclic Line Planning for High Speed Railway Based on Capacity Utilization* [Beijing Jiaotong University]. https://doi.org/10.26944/d.cnki.gbfju.2019.000711.

Zhang, X., & Nie, L. (2016). Integrating capacity analysis with high-speed railway timetabling: A minimum cycle time calculation model with flexible overtaking constraints and intelligent enumeration. *Transportation Research Part C: Emerging Technologies*, *68*, 509–531. https://doi.org/10.1016/j.trc.2016.05.005

Zhang, X., & Nie, L. (2017). *Railway Capacity Analysis based on Train Time Window by Utilizing Macroscopic yclic Timetabling Model*.

Zhang, X., Nie, L., & Ke, Y. (2019). *The Comparison of Three Strategies in Capacity - oriented Cyclic Timetabling for High - speed Railway*. 1212–1228.

Zhang, X., Nie, L., Wu, X., & Ke, Y. (2020). How to Optimize Train Stops under Diverse Passenger Demand: a New Line Planning Method for Large-Scale High-Speed Rail Networks. *Networks and Spatial Economics*, *20*(4), 963–988. https://doi.org/10.1007/s11067-020-09506-5

Zhang, Y. (2020). *Study on Optimisation Models and Algorithms for Raiwlay Train Timetabling Problems Under Complex and Typical Scenarios* [Southwest Jiatong Univeristy]. https://doi.org/10.27414/d.cnki.gxnju.2020.002545

Zhang, Y., Peng, Q., Yao, Y., Zhang, X., & Zhou, X. (2019). Solving cyclic train timetabling problem through model reformulation: Extended time-space network construct and Alternating Direction Method of Multipliers methods. *Transportation Research Part B: Methodological*, *128*, 344–379. https://doi.org/10.1016/j.trb.2019.08.001

Zhao, P., Sun, W., Hu, T., & Zhang, X. (2023). *Operational Research Guidance*. Beijing Jiaotong University Press Ltd.

Zhou, H. (2022). *Reserch in Calculation Methods of High-Speed Railway Capacity Based on Abstract Train Timetable*. Bei.

Zhu, Y. (2019). Passenger-Oriented Timetable Rescheduling in Railway Disruption Management [Delft University of Technology]. In *TRAIL Research School*. https://doi.org/https://doi.org/10.4233/uuid:10f84fa1-f5a2-4c3b-b40b-914d8858f536