Hindawi
Journal of Advanced Transportation
Volume 2020, Article ID 4294195, 18 pages
https://doi.org/10.1155/2020/4294195

Research Article

A Matheuristic Iterative Approach for Profit-Oriented Line Planning Applied to the Chinese High-Speed Railway Network

Di Liu,[1,2,3] Javier Durán Micco,[3] Gongyuan Lu,[1,2] Qiyuan Peng,[1,2] Jia Ning,[1,2] and Pieter Vansteenwegen[3]

[1]*School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China*
[2]*National United Engineering Laboratory of Integrated and Intelligent Transportation, Southwest Jiaotong University, Chengdu 610031, China*
[3]*KU Leuven Mobility Research Center-CIB, KU Leuven, Celestijnenlaan 300-box 2422, Leuven 3001, Belgium*

Correspondence should be addressed to Qiyuan Peng; qiyuan-peng@swjtu.edu.cn

Received 25 October 2019; Revised 13 March 2020; Accepted 26 May 2020; Published 1 July 2020

Academic Editor: Tao Liu

In this paper, a matheuristic iterative approach (MHIA) is proposed to solve the line planning problem, also called network design problem, and frequency setting on the Chinese high-speed railway network. Our optimization model integrates the cost-oriented and passenger-oriented objectives into a profit-oriented objective. Therefore, the passenger travel time is incorporated in the ticket price using a travel time value. As a result, transfers and detours will result in lower ticket prices and thus lower revenues for the operator. When evaluating the performance of a given line plan, the way in which passengers will travel through the network needs to be modelled. This passenger assignment is typically a time-consuming calculation. The proposed line planning approach iteratively improves the line plan using easy-to-determine indicators. During the process, a mixed integer linear programming model addresses the passenger assignment and optimizes the frequency setting in order to maximise the operational profit. Extensive computational experiments are executed to show the effectiveness of the proposed approach to deal with the real-world railway network line planning problem. Through extensive computational experiments on the small example network and real-world-based instances, the results show that the proposed model can improve the profits by 22.4% on average comparing to their initial solutions. When comparing to an alternative iterative approach, our proposed method has advantage of obtaining high quality of solutions by improving the profit 10.8% on average. For small, medium, and large size networks, the obtained results are close to the optimal solutions, when available.

1. Introduction

The Chinese high-speed railway (HSR) network has developed rapidly during the past decade. More railway lines will be constructed in 2020 to accomplish a comprehensive connection between 80% of the cities of China. Currently, the basic backbone of the HSR network contains 4 "vertical" and 4 "horizontal" tracks (4V4H). The practical HSR operation in China is different from Europe and Japan because, in China, a large number of long-distance HSR trains operate every day to satisfy as many passenger travel demand as possible. However, the average passenger travel distance is usually much shorter compared to the HSR line lengths. For instance, the average passenger travel distance on the two main HSR lines is about 558 km (Beijing-Guangzhou HSR) and 621 km (Beijing-Shanghai HSR), while the lengths of the lines are 2281 km and 1318 km, respectively [1]. This might lead to the inefficient use of railway resources such as train capacities and line capacities. For example, the average passenger load factor of HSR trains is less than 40% in some extreme cases [2].

The line planning for the HSR network is a complicated task because of the large-scale size, the high transportation demand, and the limited network capacity. Developing an efficient line plan to improve the whole network's operational performances is becoming urgent.

This paper aims to design a line plan for the 4V4H HSR network and to determine the frequency of the lines, optimizing operational cost and passenger travel time, while considering transfers when necessary. In order to obtain this, a profit-oriented objective function is applied. When the passenger demand (or potential) is considered as given and fixed, the operator's profit is determined by the operator costs and the revenue from selling train tickets. The (variable) operator costs are determined by the lines that are operated. The selling price of the train tickets is assumed to decrease when passengers need to make a transfer or detour to reach their destination. This is explained in detail in Section 3.3. Therefore, a trade-off will have to be made during line planning between, on the one hand, operating more and longer lines and, on the other hand, transfers and longer trips for the passengers.

This paper proposes an iterative approach combined with a mixed integer linear programming (MILP) model for maximizing the operator's profit during line planning. The iterative approach aims to determine better lines by heuristically modifying the current set of lines based on a fast evaluation of the current line plan. The MILP optimizes the frequency setting of the lines based on the expected routes the passengers will take, the so-called "passenger assignment" (or "transit assignment"). The two stages are optimized iteratively by what we call a matheuristic iterative approach.

The detailed contributions of this paper are as follows:

(i) A profit-oriented objective is proposed using a time value parameter in order to consider the travel time in the ticket price.

(ii) A matheuristic iterative approach is designed to solve the line planning problem.

(iii) Different local search improvements are considered to improve the current set of lines, such as extending a line, reducing a line, inserting a line, and removing a line. Fast and heuristic evaluation methods are designed to choose the most promising neighbourhood solution in order to obtain a better line plan.

(iv) MILP optimizes the frequency setting of the lines based on the expected passenger assignment.

(v) An alternative solution approach is also developed in order to illustrate the effectiveness of our approach.

(vi) A number of benchmark instances of different sizes are designed and made available together with detailed information about the best available solutions.

The remainder of this paper is structured as follows. In Section 2, the existing literature concerning the LPP is reviewed. After that, we present a mathematical model to define our profit-oriented line planning problem in detail in Section 3. This model will also be used to solve the passenger assignment and frequency setting problem. In Section 4, our matheuristic iterative approach is proposed. Several case studies and numerical experiments are shown in Section 5 to evaluate the performance and effectiveness of the proposed approach. Finally, the performance of our approach, our results for the Chinese HSR network, and our further work are summarized in Section 6.

2. Literature Review

The planning process in public transportation is typically divided into consecutive planning phases. Desaulniers and Hickman [3] consider network design, i.e., building the infrastructure, as the first phase, usually followed by line planning, timetabling, and then vehicle and crew scheduling. During operations, disturbances and disruptions might occur. Therefore, real-time rescheduling is required in order to minimize passenger inconvenience. As a crucial component of public transportation planning, the line planning problem (LPP) has attracted more and more attention recently [4–7]. Basically, the LPP decides which stops will be served by which line and in which order. Then, the frequency setting is about determining how often each line is operated. Many different variants of line planning, with different assumptions and objectives, are available in the state of the art. We will discuss a selection of the most relevant papers in this section and define the variants we will tackle mathematically in the next section.

Canca et al. [7] probably describe a problem closest to the problem discussed in this paper. However, the main differences are the components of the objective function and the planned time period. In [7], the objective function consists of ticket revenue, operational cost, and network infrastructure construction cost which are all based on the operator's point of view. Conversely, our model combines passenger-oriented and cost-oriented objectives into a profit-oriented objective. The method we used in our model considers a ticket price that depends on the passenger travel time which is not included in [7]. In terms of planning period, Canca et al. [7] consider revenue over a long period of time (i.e., years), while the LPP in this paper considers a much shorter duration (i.e., per day). Since the passenger demand scenario per day used in our model represents the regular pattern of the demand during a long period of time, our objective represents the profit over a longer period of time, typically three months, six months, or a year. If also the investment of building railway infrastructure is considered, as in [7], the considered time horizon is typically multiple years at least.

When comparing to [7], the similar components in the objective function are the revenue function, the variable operation cost, and the acquisition cost (i.e., fixed cost in our model). Our revenue function considers a ticket price that depends on the travel time, and this is not considered in [7]. Besides, the variable cost in both models include operation costs related to the line length, but crew cost is added in [7]. The crew cost shown in [7] is related to the line (i.e., frequency) and the yearly crew cost per train. We consider the crew cost as a fixed cost per train in our model. In addition, the acquisition cost in [7] formulates the purchase of each train model, while the fixed cost in our MILP includes the

depreciation expenses, material expense, fuel expense, crew cost, and other related cost per train.

In summary, our model not only considers ticket revenue but also includes the passenger travel time into the profit-oriented objective so as to minimize the passenger travel time and to make a better trade-off between the cost-oriented and passenger-oriented objectives. We consider this model to be more useful for profit-oriented line planning in practice.

Line planning models can be divided into two groups according to the objective function: a cost-oriented objective and a passenger-oriented objective. For the cost-oriented line planning problems, the objective function aims to minimize the operational cost [8–11]. The passenger-oriented line planning problems focus on maximizing the number of direct travellers [12, 13] or reducing the passenger travel time and/or the number of transfers [14, 15]. In those problems, the operator costs are considered as constraints, such as a limited number of lines. This is confirmed by Nachtigall and Jerosch [16], who propose that the cost-oriented objective and passenger-oriented objective can be considered by transforming one of them into constraints.

Obviously, some approaches try to combine the two aspects into a single problem. Pfetsch and Borndörfer [17] presented a weighted sum of the cost-oriented and passenger-oriented objective function. Rosalia [18] proposed an approach to optimize the operational cost and passenger travel time iteratively on a city road network.

A crucial challenge in passenger-oriented line planning is that, for evaluating the performance of the line plan, the passenger route choice or passenger assignment needs to be modelled. This will also determine the number of passengers on each line. This leads to a bilevel optimization problem which solves the line planning problem on the upper level and optimizes passenger assignment on the lower level [1, 5]. Friedrich et al. [19] investigated a cost-oriented line planning model with a passenger assignment evaluation process. Since the objective is cost-oriented, the solutions focus more on the operational cost, which has a negative effect on the service quality. Borndörfer and Karbstein [20] integrated line planning and passenger routing optimization by using a direct connection approach, which encourages direct connections and penalizes transfers. In addition, Karbstein [21] applied a variant of the 2-terminal Steiner connectivity problem to handle the transfers when integrating line planning and passenger routing. The complexity of the integrated line planning and passenger routing is investigated in Schmidt and Schöbel [22]. It is shown that the resulting problem is NP-hard even in very special cases.

When it comes to solving the LPP, an early approach uses a skeleton model, described by Silman et al. [23], which assembles routes from short pieces iteratively. After adding stops, the short pieces are connected by the shortest paths to form the line plan. Many approaches to line planning assume that a limited pool of possible lines is available or calculated beforehand, e.g., [14, 19]. Many models and algorithms to construct the line pool have been proposed and can be found in Kepaptsoglou and Karlaftis [24]. Other approaches to solve the LPP construct and modify lines instead of using a pool of lines [25–28]. These methods usually consider a (minimum and) maximum line length for each line. Considering heuristic approaches in the LPP becomes a tendency to reduce computation time while promising high quality solutions. The analysis of different heuristic methods applied on urban line planning is studied by Ahmed et al. [29]. Schmid [30] decomposed the bus rapid transit line planning using a large neighbourhood search to calculate the line planning design subproblem and using a linear programming model to obtain the results of passenger assignment and frequency setting. Several instances considering a single corridor are tested to show the efficiency of the proposed approach. Goerigk and Schmidt [31] used a bilevel optimization to model the line planning with passenger route choice and proposed two different techniques, binary variables and "big-M-constraints," to transfer the bilevel model into a single level model. But when the instance becomes large, say 250 stations, a genetic algorithm is required and performs well in determining a trade-off between passenger travel time and operational costs. In [7], an ALNS metaheuristic method is proposed to solve the line planning problem in railway rapid transit. At each iteration, a branch-and-cut algorithm is called to solve the passenger assignment given the results of operation information such as frequencies, train types, and fleet sizes. More details on the LPP can be found in the overviews by Schöbel [32] and Schmidt [33].

In this paper, instead of choosing between a cost-oriented and a customer-oriented objective, we propose a profit-oriented line planning model which maximizes the ticket price income minus the operational cost. The ticket price (and thus the operator revenues) is reduced when passengers need a transfer or a detour and have no direct train from their origin to their destination. This is to compensate a passenger facing a detour or transfer and an incentive for the railway operator to offer direct services as much as possible. Moreover, the operational costs consider fixed and length-dependent costs for operating the different lines. The profit-oriented line planning also turns the bilevel problem of line planning and passenger route choice into a single level problem. We have published preliminary results for profit-oriented line planning in a conference paper [34]. However, in that paper, passenger assignment is done heuristically, while this paper applies MILP to optimize the passenger assignment. Moreover, in the current paper, a better structure for implementing the local search operators allows to further improve the performance of the solution approach.

3. Profit-Oriented Line Planning

This section starts by discussing the assumptions of the line planning problem considered in this paper as well as the input data required. Then, the problem is defined mathematically with a mixed integer linear programming model. Finally, a small example network is introduced to illustrate the profit-oriented line planning problem.

3.1. Assumptions and Input Data. In this paper, the proposed profit-oriented line planning focuses on optimizing the line plan while considering passenger assignment and line plan design in a single model. The integrated model aims to make a trade-off between a cost-oriented objective, related to the number and length of the lines operated, and the passenger-oriented objective, related to minimising the travel inconvenience. This travel inconvenience is defined here as the additional travel time compared to the travel time of having a direct connection along the shortest path in the infrastructure network. In order to formulate the integrated model, the following assumptions are made throughout this paper:

(i) Stopping pattern: since only major stations are considered as nodes in the network, which attract the majority of the HSR passenger demand and are the backbone of the HSR network, the stopping pattern of the line plan is considered as an all-stop pattern for these major stations. The passenger demand of small stations can be assigned to the major stations in a precalculation phase.

(ii) Demand: passenger demand is assumed symmetrically. All demand in the network must be served with at most two transfers. In this network of limited size (only considering the major stations), two transfers should be more than enough. Transfers are penalized by a penalty time value.

(iii) Train type: trains are considered homogeneous, i.e., a double train set with 1000 seats, and its operation speed is 300 km/h. Including trains with different speeds is considered as future work.

(iv) Passenger route choice: passengers will always choose the shortest travel time path no matter what the price of the path is. Passengers of the high-speed railway normally pay more attention to the travel time rather than the ticket price. Moreover, many research papers on railway line planning [12, 15, 25, 31] assume that passengers travel according to their shortest path (with or without transfers).

(v) Line attributes: there is no limitation on the line length considered and lines can start and end in any station. Each line operates in both directions.

(vi) Feasible lines: all paths on the infrastructure network are allowed to be lines (except for cycles) and the possible passenger travel paths or train lines are not fixed in advance but determined during the line planning process. We consider no limitation on the number of lines (or their frequency) that can use a certain edge, but we take into account the capacity of the trains.

Actually, there are two types of trains operating on HSR. However, only few papers in literature [1, 7, 11, 34] consider a heterogeneous fleet during the line planning stage. Therefore, we will consider the different types of trains in future work. Moreover, the trend of HSR is to operate a single train speed so that the possibility of improving the HSR capacity utilization can be enhanced in the near future.

When considering the shortest path assignment for passenger route choice, more realistic models [35–37] have been proposed, but the LPP is already very challenging even with this simplification [22]. The additional assumption that passenger will not select a longer path (with detours and/or transfers) to get a reduction on the ticket price is based on the corresponding behaviour of most high-speed train customers. We assume they buy a more expensive high-speed ticket in order to get a shorter trip. In future work, these assumptions can be relaxed.

This is considered as input:

(i) Passenger OD matrix: the number of passengers traveling between any origin station and destination station is given in a symmetrical OD matrix. The passenger OD matrix represents the daily passenger demand.

(ii) HSR network topology: the available stations (nodes) and tracks (edges) are fixed, and the length of each edge between two stations is known beforehand.

The line planning solution is represented as a set of lines associated with certain frequencies. A line consists of a sequence of nodes.

3.2. Mathematical Model for Profit-Oriented Line Planning. A MILP model is presented to address the profit-oriented line planning. In order to keep the model understandable (and solvable in reasonable time), it assumes that a limited pool of possible lines is given. So the model will determine which lines from the pool to operate and at which frequency. It should be noted that the LPP we are solving for the Chinese HSR, and also the approach presented in Section 4, does not require such a limited pool of lines and allow to operate any feasible line.

3.2.1. Variables and Notations. The physical network topology is considered as the undirected graph $G = (S, E)$. The node set is defined as $S = \{s_1, s_2 \ldots, s_{|S|}\}$ and represents the stations. The edge set is described as $E = \{e, e \in S \times S\}$ and represents the connections of two stations in the network. When solving the LPP, a train service network (TSN) (Fu et al. [1]) is also needed to take the transfer times into account and to depict the itineraries of passengers. This is also called the Change & Go network in Schöbel and Scholl [14]. In this network, each station is duplicated per line it serves. See Figure 1(c) for an example of TSN. In the model, $[s_i, l]$ indicates the duplication of station s_i online l.

Let $W = \{w_1, \ldots, w_{|W|}\} \subseteq S \times S$ be the set of OD pairs $w_i = (s_i^{\text{org}}, s_j^{\text{des}})$. The number of passengers for a certain OD pair is denoted as d_{w_i}. A pool of possible lines L_{cur} is given as input to the MILP model. The length k_l of line $l \in L_{\text{cur}}$ is assumed to be known.

These variables are used in the model:

Inc: the total operational income

Cos: the total operational cost

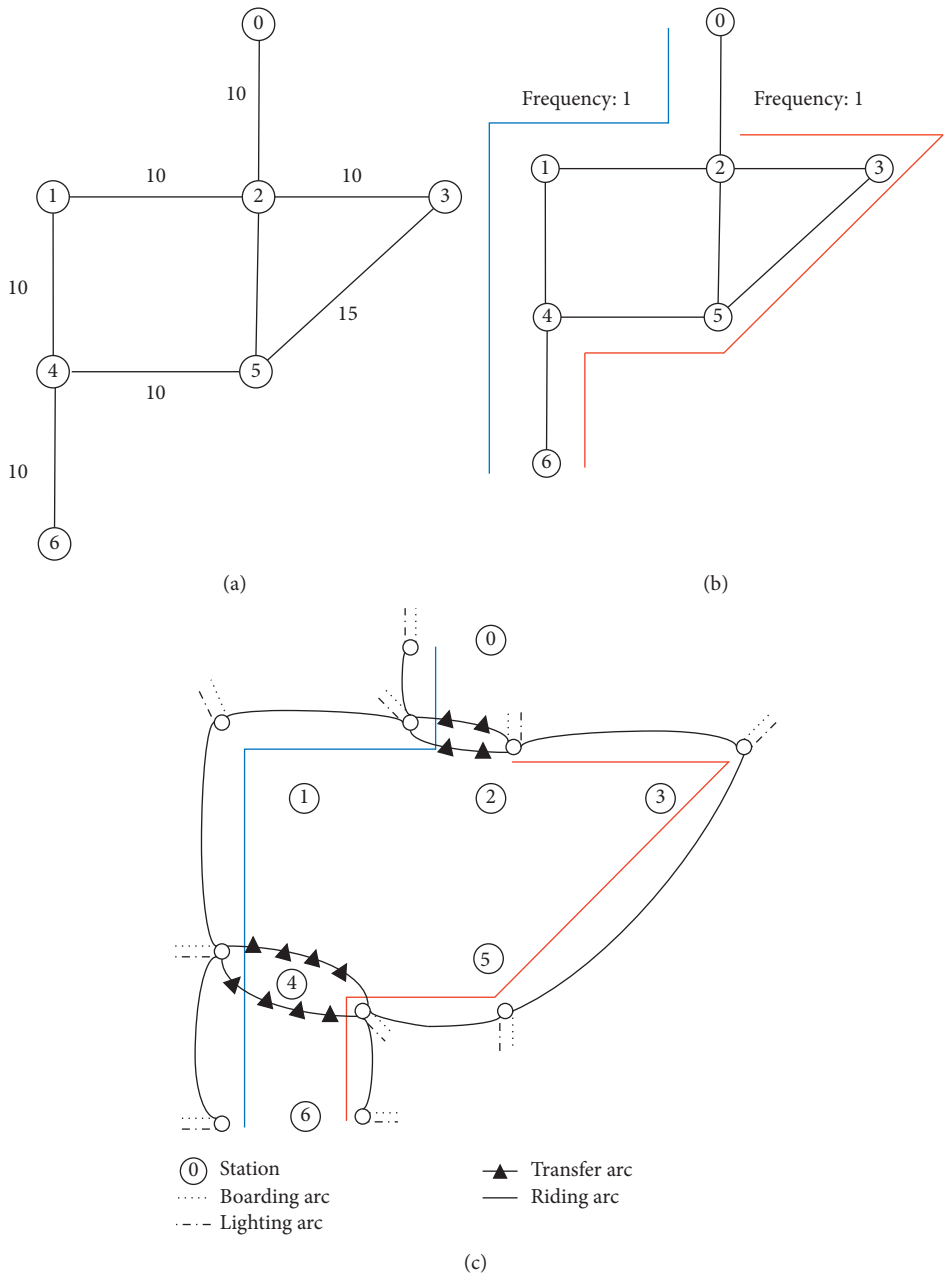

FIGURE 1: The infrastructure of the small example network (a); an example line plan with two lines (b); the corresponding TSN for this line plan (c).

Cap: the capacity of each train represented by the number of seats

L_{cur}: the pool of possible lines

C^{Fix}: the fixed cost for operating a line with frequency one

C^{Var}: the variable cost per line per kilometre for frequency one

IdeInc: the ideal income, if each passenger would have a direct connection on his/her shortest path in the physical network; i.e., IdeInc $= \sum_{w_i \in W} T_{w_i}^{\text{Phy}} * V_t * d_{w_i}$

d_{w_i}: the number of passengers for a certain OD pair w_i

$T_{w_i}^{\text{Phy}}$: the shortest path travel time of each OD pair w_i with respect to the physical network independent of the line plan

V_t: the travel time value (the ticket price per unit of time) to convert the passenger travel time into the ticket price

$T_{w_i}^{\text{TSN}}$: the shortest path travel time of each OD pair w_i on the TSN, including a fixed time penalty for each transfer

$T_{w_i}^{\text{TSN(dri)}}$: the total driving time and stopping time in $T_{w_i}^{\text{TSN}}$

$T_{w_i}^{\mathrm{TSN}(ts)}$: the total transfer time in $T_{w_i}^{\mathrm{TSN}}$

t_s: fixed stopping time at a station

t_{ts}: fixed transfer time between two lines at the same station

V_p: the penalty time value: the value of time for detours or transfers

The train service network (TSN) notations are as follows:

A_{dri}: set of driving arcs, $A_{\mathrm{dri}} = \{a = ([s_i, l], [s_j, l]), l \in L_{\mathrm{cur}}, s_i \text{ and } s_j \text{ in } S, \text{index } i < j\}$; the cost of each driving arc equals the travel time of the edge, i.e., $t_{e(a)}$

$t_{e(a)}$: the driving time of arc a on edge e

A_{ts}: set of transfer arcs, $A_{ts} = \{a = ([s_i, l], [s_i, l']): l \neq l', s_i \epsilon l, s_i \epsilon l', s_i \text{ in } S\}$; the cost of each transfer arc is set to a fixed penalty cost; i.e., t_{ts}

A_{org}: set of origin arcs, $A_{\mathrm{org}} = \{a = (s_i^{\mathrm{org}}, [s_i, l]): s_i \text{ in } S\}$; the travel time cost of each origin arc is 0

A_{des}: set of destination arcs, $A_{\mathrm{dest}} = \{a = ([s_j, l], s_j^{\mathrm{des}}), s_j \text{ in } S\}$; the travel time cost of each destination arc is 0

$A(w)$: arc set $A = \{A_{\mathrm{dri}} \cup A_{ts} \cup A_{\mathrm{org}} \cup A_{\mathrm{des}}\}$, the arcs from A used by OD pair w_i

$A_{\mathrm{dri}}(l)$: the driving arcs from A_{dri} used by line l

$y_{s_i}^l$: this parameter equals 1 when station s_i is covered by line l; otherwise, 0

$y_{e(a)}^l$: this parameter equals 1 when edge e is covered by line l as arc a; otherwise, 0

The decision variables are as follows:

f_l: the frequency of line l (which can be zero if the line is not actually operated)

$x_a^{w_i}$: binary variable equals 1 when the passenger OD pair $w_i = (s_i^{\mathrm{org}}, s_j^{\mathrm{des}})$ uses arc a; otherwise, 0

3.2.2. Objective. The objective is to maximise the operational profit, which equals the difference between revenues and operational costs. The operational cost consists of a fixed cost per line (of frequency one) and a variable cost related to the length of the line and its frequency. By introducing the travel time value parameter, the passenger travel time is converted into operational income. Thus, the operational income can be formulated as the passenger total travel time multiplied with the travel time value and minus the penalties for transfers and detours:

$$\max Z = \mathrm{Inc} - \mathrm{Cos}, \tag{1}$$

$$\mathrm{Inc} = \mathrm{IdeInc} - \sum_{w_i \in W} \left(T_{w_i}^{\mathrm{TSN}} - T_{w_i}^{\mathrm{Phy}}\right) * V_p * d_{w_i}, \tag{2}$$

$$\mathrm{Cos} = \sum_{l \in L_{\mathrm{cur}}} \left(C^{\mathrm{Fix}} + C^{\mathrm{Var}} * k_l\right) * f_l. \tag{3}$$

Equation (2) gives specific information about the revenue considering the penalties. The left side of the minus is the ideal income, obtained when each passenger travels along the shortest path in the physical network, with a direct connection. It should be noted that the ideal income is a fixed value and could be omitted from the objective function. However, since the ideal income is part of the profit components, we want to include it in order to make the objective function more readable and understandable. The right side of the minus is the penalty for transfers and detours the passengers require in the TSN. This assumes that the ticket price is reduced to compensate for the discomfort of having a longer travel time (than the ideal shortest path). The operational cost is presented as equation (3), which is related to the number of lines and its associated frequencies. An example of the objective function calculation will be presented in Section 3.3.

3.2.3. Constraints. We assume that all the lines in the given pool of lines are selected, but some lines might have a frequency of zero. The constraints included in the MILP model are listed below:

$$y_{e(a)}^l \geq x_a^{w_i}, \quad \forall l \in L_{\mathrm{cur}}, a \in A_{\mathrm{dri}}, w_i \in W, \tag{4}$$

$$y_{s_i}^{l_k} + y_{s_i}^{l_j} \geq 2x_a^{w_i}, \quad k \neq j, \forall s_i \in S, a \in A_{ts}, w_i \in W, \tag{5}$$

$$y_{s_i}^l \geq x_a^{w_i}, \quad \forall s_i \in S, l \in L_{\mathrm{cur}}, a \in A_{\mathrm{org}} \cup A_{\mathrm{des}}, w_i \in W, \tag{6}$$

$$\sum_{a \in A_{\mathrm{org}} \cap A(w)} x_a^{w_i} = 1, \quad \forall w_i \in W, \tag{7}$$

$$\sum_{a \in \delta^+(s_i) \cap A(W)} x_a^{w_i} - \sum_{a \in \delta^-(s_i) \cap A(w)} x_a^{w_i} = 0, \quad \forall s_i \in S \backslash \{s_i^{\mathrm{org}}, s_j^{\mathrm{des}}\}, a \in A(w), w_i \in W, \tag{8}$$

$$\sum_{a \in A_{\mathrm{des}} \cap A(w)} x_a^{w_i} = 1, \quad \forall w_i \in W, \tag{9}$$

$$\sum_{a \in A_{ts} \cap A(W)} x_a^{w_i} \leq 2, \quad \forall w_i \in W, \tag{10}$$

$$\sum_{a \in A_{\text{dri}} \cap A(W)} x_a^{w_i} \left(t_{e(a)} + t_s\right) - t_s = T_{w_i}^{\text{TSN(dri)}}, \quad \forall w_i \in W, \tag{11}$$

$$\sum_{a \in A_{ts} \cap A(W)} x_a^{w_i} \left(t_{ts} - t_s\right) = T_{w_i}^{\text{TSN}(ts)}, \quad \forall w_i \in W, \tag{12}$$

$$T_{w_i}^{\text{TSN(dri)}} + T_{w_i}^{\text{TSN}(ts)} = T_{w_i}^{\text{TSN}}, \quad \forall w_i \in W, \tag{13}$$

$$\sum_{w_i \epsilon W} \left(x_a^{w_i} \times d_{w_i}\right) \leq f_l \times \text{Cap}, \quad \forall a \in A_{\text{dri}}(l),\ l \in L_{\text{cur}}, \tag{14}$$

$$f_l \in \mathbb{N}, \quad \forall l \in L_{\text{cur}}, \tag{15}$$

$$x_a^{w_i} \in \{0, 1\}, \quad \forall a \in A,\ w_i \in W. \tag{16}$$

Equations (4)–(6) assure the TSN construction based on a line plan. The three constraints indicate that only nodes and edges that are covered by the possible lines are considered as the nodes and arcs of the TSN; i.e., edges belonging to the lines can be selected as the driving arc of passenger routes (4) and nodes (stations) covered by more than one line are chosen as potential transfer arcs (5). Origin and destination arcs are connected to each node covered by a line in the TSN (6). These constraints assume that driving arc a in the TSN corresponds to edge e in the physical network (4) that transfer arc a corresponds to a transfer between l_i and l_j (5) and that station s_i is an endpoint of an origin or destination arc a (6).

Equations (7)–(9) are network flow conservation constraints, which require that each passenger OD pair should have a feasible path on the TSN based on the given line plan. Equations (7) and (9) assure that only one origin arc and destination arc can be selected for each passenger OD pair. On the contrary, for each passenger OD pair, there must be one origin arc and one destination arc. Equation (8) ensures the conservation of passenger flows on the intermediate nodes. If there is no travel path for any of the OD pairs or the number of transfers is more than two, the model will turn out to be infeasible.

Equation (10) limits the number of transfers for each OD pair to 2. Equations (11)–(13) calculate the actual travel time for each OD pair. Equations (11) and (12) compute the actual total driving time and total transfer time for a certain OD pair. Equation (14) ensures the capacity of a line on a driving arc is sufficient to meet the passenger demand on that driving arc; i.e., all passengers are served taking into account the capacity of the trains. The frequency of the train is calculated based on the number of passengers taking that train. Constraint (15) and constraint (16) are variable value constraints.

The main purpose of this model is to clarify the profit-oriented line planning problem considered in this paper. Moreover, this model will be used in the line planning approach we present in Section 4. However, the pool of possible lines L_{cur} considered in the MILP model above will be replaced by the lines of the line plan under evaluation. As a result, it will optimize the passenger assignment and the frequencies of the line plan under evaluation.

3.3. Small Example Network. Here, a small network is introduced in Figure 1(a) which will be used to illustrate the calculation of the objective function. In this small example, the number besides each arc corresponds to both the distance (km) and the travel time (min). In this example, the data of Table 1 are assumed to be given.

A brief explanation of the objective function calculation is presented based on the example line plan in Figure 1(b). In order to consider the transfers and depict the itinerary of passengers, the TSN (see Figure 1(c)) is constructed based on the given line plan in Figure 1(b). As indicated in the figure, both lines have frequency 1 per unit time (for example, 1 per hour or 1 per day). The operational cost of the blue line is $(15,000 + 150 \times (10 + 10 + 10 + 10)) \times 1 = 21,000$ RMB. For the red line, this is 21,750 RMB, so the total operational cost is 42,750 RMB. The revenue is calculated by the ideal income minus the penalties caused by transfers and detours. In order to simplify the calculation process, we assume that the passenger demand of each OD pair is 100. The ideal income is calculated as the price that all the passengers pay for their direct connections on their shortest paths. For instance, the ideal income of passengers from nodes 1 to 3 (the path (1, 2, and 3)) is $(10 + 1 + 10) \times 2.5 \times 100 = 5\,250$ RMB. The total ideal income can be calculated as 105,750 RMB. When computing the penalty fees of transfers and detours, the penalty time equals the difference between the actual travel time and the shortest possible travel time. For example, the penalty fee of passengers from node 1 to node 3, requiring a transfer in node 2 (instead of a stop), equals $((10+5+10) - (10+1+10)) \times 0.55 \times 100 = 220$ RMB. Another example includes the passengers from node 2 to

Table 1: Data of the small example network.

Name	Value	Unit
Transfer time penalty	5	min
Station stopping time	1	min
Travel time value	2.5	RMB/person, min
Penalty time value	0.55	RMB/person, min
Fixed line operating cost	15,000	RMB/train
Variable line operating cost	150	RMB/train, km
Train capacity	1000	Seats/train

node 5, requiring a detour through node 3. The penalty time value multiplied with the detour time and the passenger demand equals $((10+1+15)-10)\times 0.55\times 100 = 880$ RMB. The total penalty fee of the passengers with transfers and detours (OD pairs 0-3, 0-5, 1-3, 1-5, and 2-5) is 2640. The final profit of this line plan is $105,750-2640-42,750 = 60,360$ RMB.

4. A Matheuristic Iterative Approach for Line Planning

After generating an initial line plan, our approach decomposes the LPP into two subproblems, improving the design of the lines and the passenger assignment together with frequency setting. The two subproblems are optimized iteratively based on a heuristic evaluation of possible neighbourhood solutions in the first subproblem and a MILP model to address the second subproblem. We call this approach MHIA, and it is illustrated in Figure 2 and discussed in detail in the next sections.

4.1. Initial Line Plan Generation. The basic idea of the initial line plan generation is to select those lines that serve directly as much passenger demand as possible. According to the classical approach for line planning introduced in Bussieck et al. [12], we first search the shortest paths for each OD pair on the physical network by the Floyd–Warshall algorithm [38]. This results in a set of candidate lines. The initial line plan generation consists of two stages, namely, heuristic construction and repair, determining which lines to include in the initial solution. During the heuristic construction process, the line set is built by selecting one line at a time from the candidate lines until all the nodes are covered. The choice of the lines obeys the rule that the selected line serves the most passengers on their shortest paths without transfers, not only from the starting towards the ending station of the line but also from and towards all stations in between on that line. When selecting the next line to include, all passengers already served directly by the already selected lines are no longer considered. For example, the first line is selected based on the calculation of the direct passengers served by each line. Then, the passengers served by that line are eliminated from the OD matrix. The next line is then selected from the candidate lines based on the remaining unserved passengers. The selection procedure continues until all the nodes of the network are covered.

Since the construction stage focuses on direct connections, it does not guarantee that all passengers have a path and have less than two transfers. For this reason, the MILP model of the previous section is used to check the line plan feasibility by using the Floyd–Warshall algorithm to calculate all the shortest paths for the OD pairs. If there exists any OD pair for which no path with two or less transfers is available in the TSN, then the shortest path on the physical network of that OD pair is selected as a new line and added to the initial line plan.

To illustrate the initial line plan generation, we return to the small example network of Figure 1(a). The list with the shortest path of each passenger OD pair (in one direction) is given in Table 2.

We assume each OD pair corresponds to 100 passengers. Based on serving directly as much demand as possible, the first line will be (0, 2, 1, 4, 6) or (0, 2, 5, 4, 6), both serving 1000 passengers. Then, those served passengers are removed from the OD matrix, and the second line is selected from the rest of the paths in the same way. If we choose (0, 2, 1, 4, 6) as the first line, then (3, 5, 4, 6) becomes the second line with 500 passengers served directly. Now, all the nodes on the network are covered by the selected lines. After the feasibility check (travel path, number of transfers, and node coverage), the selected lines meet all the constraints and an initial line plan is generated.

4.2. Line Plan Evaluation and Modification. In this line plan modification and evaluation process, illustrated in Algorithm 1, two subproblems are considered: line plan design (modification) and passenger assignment and frequency setting (evaluation). Clearly, the passenger assignment is a crucial part of the line plan evaluation. However, passenger assignment typically requires a lot of calculation time, and the smallest change in the line plan could significantly change the passenger assignment and thus the passenger travel times. Therefore, we try to limit the number of times the passenger assignment is calculated. It would be, for instance, too time-consuming to calculate the passenger assignment for every possible modification to the current line considered below. Therefore, we try to improve the line plan based on easier-to-calculate heuristic evaluation indicators. Only for the most promising modifications, the passenger assignment and frequency setting are applied using the MILP model discussed in Section 3.

In the plan design or modification part, a framework with four modification operators is developed to iteratively improve the current line plan. The detailed modification and evaluation process are illustrated in Figure 2. The four modification operators are reducing a line (*Reduction*), extending a line (*Extension*), removing a line (*Removal*), and inserting a line (*Insertion*). *Reduction* and *Extension* work as intensification or improvement modifications, while *Removal* and *Insertion* are used as diversification of the search. Each type of modification leads to a set of possible line plans, the neighbourhood of the current line plan.

The *Reduction* neighbourhood of a current line plan contains all line plans where one terminal node of one line is

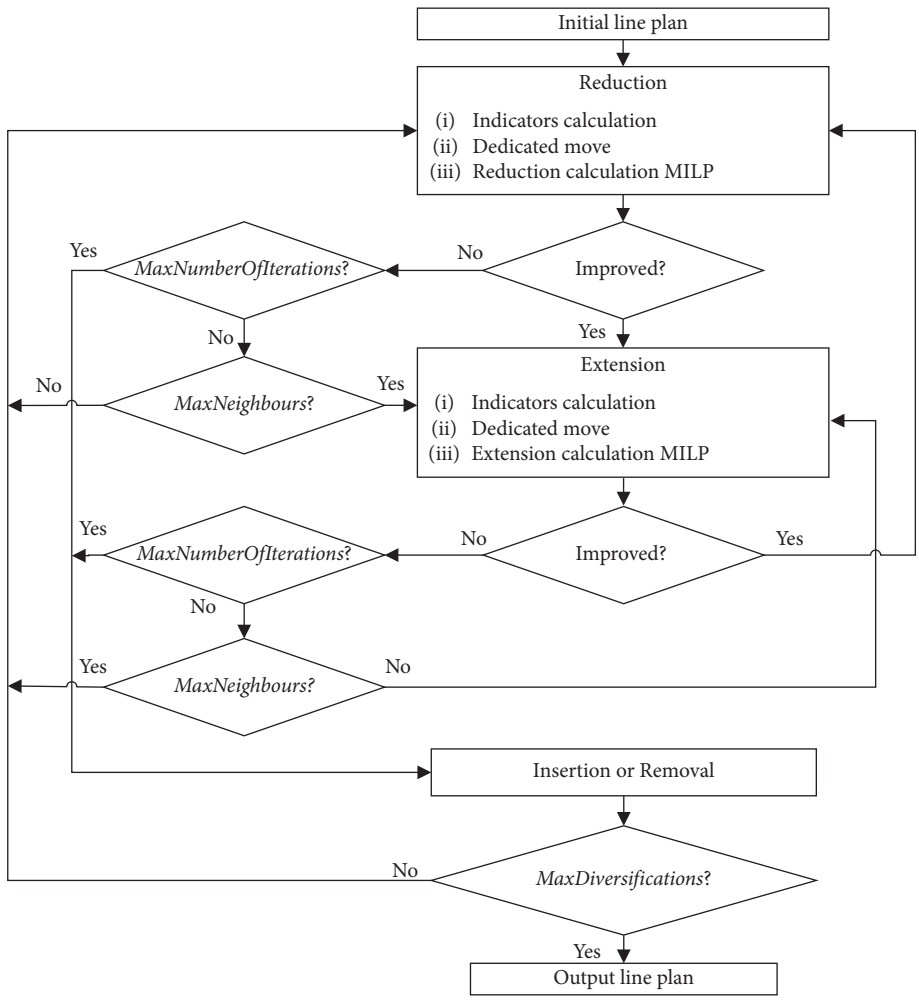

Figure 2: The process of MHIA for line planning optimization.

Table 2: The shortest paths for passenger OD pairs on the small example network.

O	D	The shortest paths
0	1	0, 2, 1
0	2	0, 2
0	3	0, 2, 3
0	4	0, 2, 1, 4/0, 2, 5, 4
0	5	0, 2, 5
0	6	0, 2, 1, 4, 6/0, 2, 5, 4, 6
1	2	1, 2
1	3	1, 2, 3
1	4	1, 4
1	5	1, 4, 5/1, 2, 5
1	6	1, 4, 6
2	3	2, 3
2	4	2, 1, 4/2, 5, 4
2	5	2, 5
2	6	2, 1, 4, 6/2, 5, 4, 6
3	4	3, 5, 4
3	5	3, 5
3	6	3, 5, 4, 6
4	5	4, 5
4	6	4, 6
5	6	5, 4, 6

removed. The *Extension* neighbourhood contains all line plans where one node, adjacent to a terminal node in the physical network, is added to one of the lines. When it comes to *Insertion*, each of the lines corresponding to OD pairs without a direct connection in the current line plan are considered to be inserted in the current line plan. For *Removal*, the neighbourhood contains all line plans where a line of the current line plan is removed. In each neighbourhood of the four operators, only feasible line plans are considered.

When considering *Reduction*, the load factors of terminal edges of each line are calculated as the evaluation indicator of the current line plan. The load factor is the actual passenger volume on the edge of a line divided by the total capacity (the number of seats) of the corresponding line. A low load factor might indicate a part of a line which is not profitable. The load factor is calculated for each terminal edge, and then, the terminal edge with the lowest load factor is removed from its line. The MILP model is then applied to optimize the frequencies and to optimally assign the passengers to the new line plan. Only when the total profit is actually increased by removing this edge, the new line plan is accepted and the algorithm continues with considering

```
(1) n_der = 0;
(2) Repeat
(3)   n_int = 0, n_red = 0, n_ext = 0;
(4)   Repeat
(5)     Reduction method
(6)       n_int = n_int + 1;
(7)       Select the L_nei and calculate the profit of L_nei by MILP;
(8)       if Z_nei > Z_cur do
(9)         L_cur ⟵ L_nei, Z_cur ⟵ Z_nei, n_red = 0;
(10)        if Z_nei > Z_best do
(11)          L_best ⟵ L_nei, Z_best ⟵ Z_nei;
(12)      else
(13)        n_red = n_red + 1;
(14)        if n_int > MaxNumberOfIterations
(15)          go to step 36;
(16)        else if n_red > MaxNeighbours
(17)          go to step 20;
(18)        else
(19)          go to step 5;
(20)    Extension method
(21)      n_int = n_int + 1;
(22)      Select the L_nei and calculate the profit of L_nei by MILP;
(23)      if Z_nei > Z_cur do
(24)        L_cur ⟵ L_nei, Z_cur ⟵ Z_nei, n_ext = 0;
(25)        if Z_nei > Z_best do
(26)          L_best ⟵ L_nei, Z_best ⟵ Z_nei;
(27)      else
(28)        n_ext = n_ext + 1;
(29)        if n_int > MaxNumberOfIterations
(30)          go to step 36
(31)        else if n_ext > MaxNeighbours
(32)        go to step 5;
(33)        else
(34)          go to step 20;
(35)  until the terminal conditions are satisfied
(36)  Disturb
(37)    Removal or Insertion
(38)      n_der = n_der + 1;
(39)      Select the L_nei and calculate the profit of L_nei by MILP;
(40)      L_cur ⟵ L_nei, Z_cur ⟵ Z_nei, go to step 3
(41) until n_der = Max Di versi fications
(42) Output the best line plan and objective profit
```

Algorithm 1: Pseudocode of the MHIA of line planning.

Extension. However, when removing this edge would decrease the profit, the reduction is cancelled and the algorithm considers to remove the edge with the next lowest load factor. This continues until an edge is found that actually improves the profit or until the number of neighbours considered reaches a predefined maximum number (*MaxNeighbours*). In both cases, *Reduction* is ended, and *Extension* is considered.

Extension evaluates how many passengers that currently need a transfer could be transported directly due to extending a line with an additional edge. The edge that can provide the most additional direct connections is added. Again, the MILP model is applied to optimize the frequencies and optimally assign the passengers to the new line plan. Only when the total profit is actually increased by extending this line, the new line plan is accepted and the algorithm continues by going back to *Reduction*. However, when adding this edge would decrease the profit, the extension is cancelled and the algorithm considers to add the next most promising edge. As with *Reduction,* this continues until an edge is found that actually improves the profit or until the number of neighbours considered reaches a predefined maximum number (*MaxNeighbours*). In both cases, *Extension* is ended, and *Reduction* is considered again.

In order to limit the total computation time of *Reduction* and *Extension*, we explicitly limit the total number of times the MILP model is used. When the MILP model is used *MaxNumberOfIterations* times or when no more improvements can be found by *Reduction*

or *Extension*, the algorithm continues with the diversification phase.

In order to diversify the algorithm, two ways of disturbing the search are implemented as well: *Removal* and *Insertion*. First, however, when some lines have a frequency of zero after the frequency setting, these nonoperated lines will be removed. Then, the new line plan is selected randomly from all possible *Removal* and *Insertion* neighbourhood solutions. The solution of the disturbance is always accepted. The number of diversification iterations (*MaxDiversifications*) is fixed beforehand as a stopping criterion for the algorithm. After that, the best solution obtained during the search process will be presented as the final solution.

An alternative implementation would be to consider all the four moves at the same time. However, this would be too time-consuming and not efficient. So we choose the aforementioned operator execution order to optimize the line plan. The *Reduction* and *Extension* are used to search for the approximate local optimal solution by small changes in the input line plan. These two operators are crucial operators in the good performance of the algorithm and are considered as intensification moves. However, the *Insertion* or *Removal* is selected randomly to diversify the solution search space. These should be regarded as large changes to the input line plan.

Since the initial line plan is generated by choosing the lines that can serve as many direct passengers as possible, this typically leads to long lines. In order to avoid such negative effects of the initial line plan, the *Reduction* is applied first. After a limited number of iterations with *Reduction* and *Extension*, the random *Insertion* or *Removal* provides other search directions.

Compared to using large neighbourhood search (LNS) [35, 39], the approach proposed in this paper is different in the following aspects. In LNS, a temporary solution is given by first applying a destroy method and then a repair method. For the proposed approach in this paper, the temporary solution is given separately by *Reduction*, *Extension*, *Removal*, or *Insertion*. Specifically, the destroy method in LNS destructs randomly a part of the current solution, and then, the repair method reconstructs the destroyed part. However, our proposed approach based on four operators, respectively, modifies the current solution by a heuristic evaluation and selects the most promising one as the temporary solution.

The notation used in Algorithm 1 is as follows:

Z_{cur}: the profit of the current line plan

Z_{best}: the best profit among the calculated line plans

Z_{nei}: the profit of a neighbourhood solution of the current line plan

L_{nei}: the selected neighbourhood line plan

A_{undir}: the set of OD pairs of passengers with a longer travel time than their ideal travel time

n_{der}: the current number of *MaxDiversifications* iterations

n_{red}: the current number of *Reduction*

n_{ext}: the current number of *Extension*

n_{int}: the current number of intensification iterations

4.3. An Alternative Iterative Approach. It is very time-consuming to obtain an optimal passenger assignment and frequency setting using the mathematical model when the network becomes larger. This is mainly due to the exponentially increasing number of possible lines that should be considered. So, an alternative iterative approach (AIA), avoiding this extensive use of the mathematical model, is now presented. It will be mainly used for evaluating the performance of MHIA. The AIA framework is shown in Figure 3.

AIA uses the same modification operators and heuristic evaluation methods as MHIA. However, the calculation time of the passenger assignment is significantly reduced by using the assignment results of the former line plan. We call this "heuristic passenger assignment." If passengers have a direct connection on their shortest path, the passenger paths are assumed to remain the same as in the former line plan. The algorithm only searches for shorter paths for those who do not have a direct connection or make a detour and then it calculates the required frequency of each line. Another difference is that the *MaxNeighbours* in AIA is limited to one, while the *MaxNeighbours* in MHIA is limited to ten. Therefore, only the most promising *Reduction* (*Extension*) is evaluated (and implemented when successful) before continuing to *Extension* (*Reduction*).

5. Computational Experiments

In this section, several experiments are executed to show the performance of the proposed MHIA on solving the profit-oriented LPP. The proposed method is implemented in C# and runs on an Intel(R) Core (TM) i7-3770 CPU 3.40 GHz and 16.0 GB computer. The MILP is implemented in CPLEX version 12.6.3, using the C# application program interface of the solver with the default parameter values. The input data include the network infrastructure, passenger demand OD matrix, and other operational parameters, such as track travel time (depending on the fixed train speed), fixed cost, and variable cost elements. The parameters used throughout the experiments are given in Table 3. It should be noted that some of these values are different (more realistic) compared to the values used in Table 1. In order to make the small example network more realistic, all link lengths were multiplied by 20. As a result, the distance between node 0 and node 2, for instance, is 200 km. Due to the lack of real-world data (which is confidential for the HSR network in China), the passenger demand OD matrices are generated randomly, and different demand scenarios are considered. All instance data are made available at https://www.mech.kuleuven.be/en/cib/lp/mainpage#section-4.

After preliminary experiments with different combinations of the maximum number of *MaxNeighbours*,

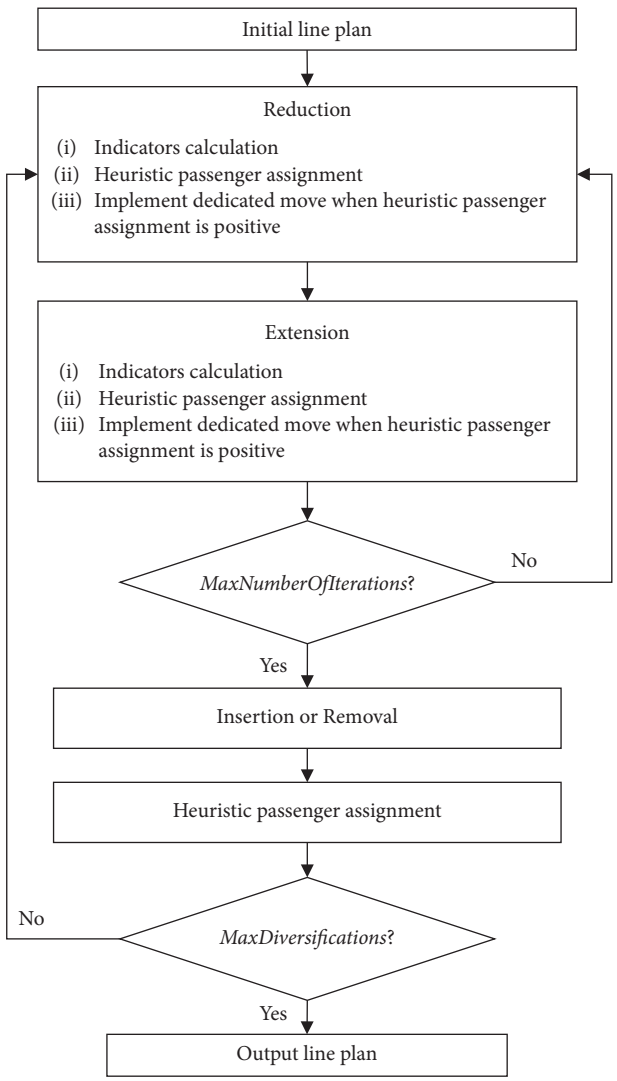

Figure 3: The alternative iterative approach framework for line planning optimization.

Table 3: Experiment parameter setting.

Name	Value	Unit
Train speed	300	km/h
Transfer time penalty	30	min
Station stopping time	3	min
Travel time value	2.5	RMB/person, min
Penalty time value	0.55	RMB/person, min
Fixed line operating cost	15,000	RMB/train
Variable line operating cost	150	RMB/train, km
Train capacity	1000	Seats/train

MaxNumberOfIterations, and *MaxDiversifications*, we set the number to 10, 20, and 30, respectively.

In the line planning problem, the transfer time is typically modelled as a penalty value [1, 7, 19–21, 29–31, 33–35]. Because the train departure times are not known during line planning, the exact transfer time is unknown and estimated by a transfer penalty. Here, we set the transfer time (penalty) to 30 minutes taking into account both the inconvenience of passengers and the practical situation in China high-speed railways.

The travel time value is used to calculate the ideal income. Here, we transform the ticket price of high-speed trains into a time-related price to link the operational revenue and the passenger travel time. However, the transfer penalty time value is used to compensate the inconvenience of making transfers.

5.1. Small Example Network Experiments. In this experiment on the small example network introduced in Section 3.3, the performance of MHIA is compared to the optimal solution obtained by using the MILP model presented in Section 3.2. In order to obtain the best possible profit-oriented line planning for this network, the MILP model starts with a pool of lines containing all 62 feasible lines that could be generated for this network. By doing this experiment, we can measure how close MHIA can get to the optimal solution for small networks.

This experiment considers three different passenger demand scenarios (PDS), based on three randomly generated OD matrices. Actually, the profit-oriented line planning problem as we propose it here has not been addressed in literature before. This makes a comparison with methods for literature [1, 7, 27, 28, 31] impossible.

The results of 10 runs of both AIA and MHIA are given in Table 4. The optimal solutions are calculated by the MILP model since all possible lines are considered. In all the tables, "PDS" refers to a passenger demand scenario. Table 4 reports the profit of the initial solution (IS) generated by AIA or MHIA, the best profit (BS) obtained by AIA and MHIA, the average profit (AveS) over all the runs, the average computation time (ACT) of one run of the approach, and the number of lines (Lines) included in the best result (BS). For MILP, BS is the optimal solution. In this experiment, the gap is the difference between the optimal solution and the best solution given by MHIA (or AIA), i.e., gap = $(BS_{MILP} - BS_{MHIA}$ (or $BS_{AIA}))/BS_{MILP}$.

From Table 4, it can be seen that the best results of MHIA are very close to the optimal solution, with an average gap of only 2.9%. Moreover, the best results of MHIA for this small example network are much better than the results given by the AIA. The computation time for all the approaches remains limited for this very small network, but, obviously, AIA is faster than MHIA and both are many times faster than MILP. We conclude here that, for a very small network, MHIA guarantees high quality solutions for the profit-oriented LPP in limited computation time. In this case, the heuristic passenger assignment used by AIA has the same initial results as the accurate passenger assignment applied in MHIA.

All detailed information about the best solutions obtained for each demand scenario (lines, frequencies, and profit calculations) and the 62 feasible lines are made available at https://www.mech.kuleuven.be/en/cib/lp/mainpage#section-4. Also the results discussed in the next sections are made available there.

Table 4: Comparison between different approaches on the small example network.

	PDS	IS	BS	AveS	ACT (s)	Lines	Gap (%)
MILP	1	—	638,670	—	26,950	4	—
	2	—	676,490	—	6680	4	—
	3	—	826,630	—	28,285	4	—
AIA	1	362,073	538,630	519,352	2	3	15.7
	2	98,773	576,724	541,066	0.9	5	14.7
	3	241,145	708,695	698,884	1	4	14.3
MHIA	1	362,073	621,244	582,370	33	5	2.7
	2	98,773	658,336	641,944	11	3	2.7
	3	241,145	799,111	759,788	16	4	3.3

5.2. Real-World HSR Network Experiments. After studying the performance of the MHIA on the small example network, we now apply it to three real-world HSR networks in China, a small network (11 nodes and 110 OD pairs), a medium network (26 nodes and 650 OD pairs), and a large network (34 nodes and 1122 OD pairs).

Given the complexity of the problem, the optimal solutions for profit-oriented line planning are not available and too time-consuming for these HSR networks, using the MILP. This makes it difficult to properly assess the performance of MHIA on these networks. Therefore, in order to test the performance of MHIA in this section, the best solutions found by MHIA will be compared with both the initial solutions found by MHIA and the best solutions found by AIA. In Section 5.3, the best solutions found by MHIA will be compared with the best results that can be obtained by solving the MILP model of Section 3.2 for these networks.

Figure 4 shows the topology of the small HSR network. We run all the experiments based on the parameters given in Table 3. The results of 10 runs of each approach are presented in Table 5. Here, the gap corresponds to gap = $(BS_{MHIA} - BS_{AIA})/BS_{AIA}$.

It should be noted that the initial solution for AIA and MHIA corresponds to the same line plan, but higher profits are calculated for MHIA due to the different passenger assignment calculations. Since AIA uses a heuristic calculation of the passenger assignment, its profits are lower than the optimal passenger assignment of MHIA on the small HSR network. Compared to AIA, MHIA can increase the operational profit by 10.7% on average while taking around ten times more computation time due to the accurate passenger assignment. The best results obtained by MHIA improve the initial solution with 40.0% on average. Taking passenger demand scenario 1 as an example, the detailed information about the best result of MHIA is shown in Figure 5 and Table 6. The numbers next to the coloured lines are the frequencies of those lines.

The high frequency on the line parts (4-9, 9-8, 8-7, and 7-6) in this solution corresponds nicely with the fact that 49% of the shortest paths include these line parts which indicate that passengers are more likely to choose these line parts. This illustrates that the best result found by MHIA is consistent with the passenger demand. The reason why line

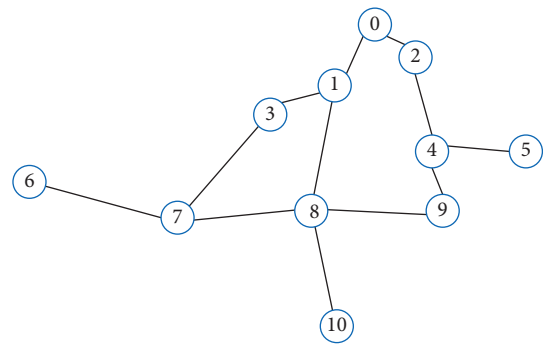

Figure 4: The topology of the small HSR network.

(4, 9, 8, and 10) has the highest frequency is that all the passengers who want to travel to destination 10 have to use that line. According to the passenger demand OD matrix, the number of passengers ending at station 10 is 2336. This indicates that 3 trains are required in order to serve all the passengers taking that train, considering the train capacity.

When the size of the considered network grows, the calculation time of looking for the optimal passenger assignment of the network will become very time-consuming. Therefore, we set the optimality gap parameter of the CPLEX solver to 5%, which means the solver will give the current best solution as a result when the gap between the current best solution and the current upper bound is lower than 5%. This significantly speeds up the passenger assignment and frequency setting. The network topology of the medium HSR network is given in Figure 6 (only the black edges), and the results of 5 runs are shown in Table 7. It should be noticed that the computation times are now expressed in minutes. The gap used here is gap = $(BS_{MHIA} - BS_{AIA})/BS_{AIA}$.

With the increasing size of the network, MHIA shows advantages in solving the profit-oriented LPP. In Table 7, the reason that the profit of the initial line plan of AIA is different from that of MHIA is again that AIA uses heuristic passenger assignment, while MHIA applies the optimal passenger assignment. On average, MHIA performs 8.3% better than AIA, and the best solution of MHIA improves the initial solution with 16.5%.

When applying MHIA to the large HSR network, corresponding to the entire 4V4H network (the black and grey edges in Figure 6), the calculation time becomes too long. Therefore, in order to make an evaluation of the calculation process, the sensitivity of algorithm parameters is tested for a single demand scenario. Since the *MaxDiversification* is one of the main parameters that determines the algorithm calculation time, different values for *MaxDiversification* are tested. In order to accelerate the algorithm calculation speed, another two values for *MaxDiversification* are considered: 10 and 20 (originally 30). The results of 5 runs for each value of *MaxDiversification* are given in Table 8. Here, gap = $(BS_{MHIA(30)} - BS_{MHIA(10\,or\,20)})/BS_{MHIA(30)}$.

The results in Table 8 show that the solution quality remains almost the same when reducing *MaxDiversification* of MHIA. At the same time, the computation time decreases from 12.0 hours to 4.1 hours, a reduction of 63%. So, we

Table 5: Comparison between AIA and MHIA on small HSR network.

	PDS	IS ($*10^6$)	BS ($*10^6$)	AveS ($*10^6$)	ACT (s)	Lines	Gap (%)
AIA	1	2.41	3.51	3.43	12	7	—
	2	2.05	3.40	3.27	10	5	—
	3	2.43	3.23	3.18	11	5	—
MHIA	1	3.10	3.94	3.81	113	5	12.3
	2	2.08	3.74	3.70	118	4	10.0
	3	3.08	3.55	3.51	91	5	9.9

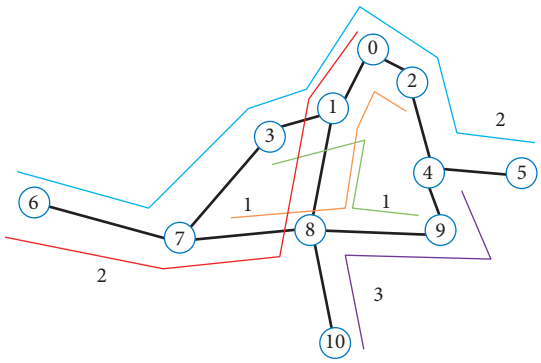

Figure 5: The best line plan of PDS 1 obtained from MHIA.

Table 6: Best line plan information of PDS 1 obtained by MHIA on the small HSR network.

Line	Distance (km)	Frequency
5, 4, 2, 0, 1, 3, 7, 6	2397	2
0, 1, 8, 7, 6	1608	2
2, 0, 1, 8, 7	1200	1
4, 9, 8, 10	1167	3
3, 1, 8, 9	891	1

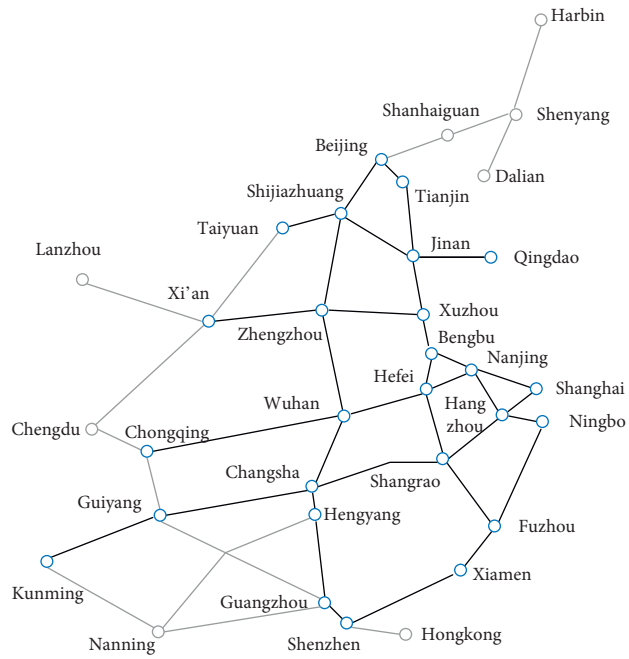

Figure 6: The topology of the medium (only black edges) and large (the black and grey edges) HSR network.

think it is appropriate to reduce the number of diversifications when solving the LPP on the large HSR network.

Finally, we run MHIA with the new parameter setting on the other two demand scenarios for the large HSR network. The parameters used in this experiment for MHIA are 5% optimality gap in CPLEX and 10 *MaxDiversification*. Also, for AIA, only 10 *MaxDiversification* are considered. The results of this setting are listed in Table 9.

For the large network case, MHIA shows an excellent ability in obtaining high quality solutions and controlling the computation time. The initial solutions of MHIA for all demand scenarios are better than the best results obtained by AIA. On average, the best results of MHIA are 13.3% better than the best results of AIA on the large HSR network. However, the computation times for AIA are clearly shorter.

These computation times mainly indicate the difficulty of the problem, not the efficiency of the algorithm. Actually, these kinds of computation times are normal for LPP of this size considering passenger assignment. Moreover, for practical applications, such computation times are acceptable, since the line planning is typically only changed in the long term.

Besides, since the problem we considered is new, the performance of the algorithm cannot be compared to other approaches from the state of the art. Therefore, we show the efficiency of the algorithm by comparing the results with optimal solutions obtained using CPLEX (only possible on smaller instances) and with AIA on all instances.

5.3. Comparison between MHIA and MILP on the HSR Network. As mentioned above, in this section, the best solutions found by MHIA will be compared with the best results that can be obtained by the MILP model of Section 3.2, for the small, medium, and large HSR network. To obtain these best results by the MILP model, all the lines obtained in the results of the different runs of MHIA (and AIA for the small HSR network) in Section 5.2 are included in the line pool for the MILP. Obviously, this does not guarantee that an optimal solution will be found, but at least a significant number of high quality lines are now considered in this line pool.

First, for the small HSR network and for each demand scenario, we use all the lines of the line plans from the result

Table 7: Comparison between AIA and MHIA on medium HSR network.

	PDS	IS ($*10^7$)	BS ($*10^7$)	AveS ($*10^7$)	ACT (min)	Lines	Gap (%)
AIA	1	1.66	2.24	2.21	13	7	—
	2	1.96	2.27	2.20	9	11	—
	3	1.76	2.14	2.11	9	10	—
MHIA	1	2.03	2.45	2.37	172	7	9.3
	2	2.18	2.40	2.37	214	7	5.7
	3	1.98	2.35	2.31	175	9	9.8

Table 8: Results for different values of *MaxDiversification* on the large HSR network.

MaxDiversification	BS ($*10^7$)	AveS ($*10^7$)	ACT (hour)	Lines	Gap (%)
10	6.79	6.64	4.1	11	−1.6
20	6.80	6.68	7.8	13	−1.8
30	6.68	6.61	12.0	11	—

Table 9: Comparison between AIA and MHIA on the large HSR network.

	PDS	IS ($*10^7$)	BS ($*10^7$)	AveS ($*10^7$)	ACT (hour)	Lines	Gap (%)
AIA	1	4.71	5.79	5.60	0.4	13	—
	2	5.16	6.11	6.02	0.3	13	—
	3	4.60	5.67	5.46	0.7	13	—
MHIA	1	6.05	6.79	6.64	4.1	11	17.3
	2	6.22	6.86	6.64	4.8	11	12.3
	3	5.71	6.25	6.09	5.2	11	10.2

of each of the 10 runs of MHIA and AIA as the line pool for the MILP model. The results are shown in Table 10.

The results show that the MHIA performs well in reaching the solutions calculated by MILP on the small HSR network for different demand scenarios. The gaps are very small. When the PDS 1 is taken as an example, there are 35 lines considered in the line pool and the detailed information about the results of both MHIA and MILP are given in Table 11. The column "Frequencies (MHIA)" presents the frequencies of the best solution found by MHIA, and the column "Frequencies (MILP)" refers to the best result the MILP model can obtain.

By comparing the frequencies in the second column and the third column, we can see the frequencies of the line in MHIA solution are typically reduced and replaced by slightly modified lines. For example, the frequency of the first line (4, 9, 8, 10) is reduced from 3 to 1, but instead, the MILP solution operates two very similar lines (4, 9, 8, 7, 6 and 4, 9, 8, 7). This could be an indication that the performance of MHIA can be improved by considering more (slightly) different lines instead of higher frequencies for fewer lines. Maybe these better solutions can also be obtained by increasing *MaxNeighbours*, *MaxNumberOfIterations,* or *MaxDiversification*. However, the quality improvement would be around 1.0%, while the computation time might significantly increase. For the current solutions, MHIA requires 113 s and MILP requires 261,445 s. We conclude that MHIA performs well in making a trade-off between the solution quality and the computation time. Moreover,

Table 10: Results of MHIA and MILP (with a given line pool) for the small HSR network.

	PDS	BS ($*10^6$)	ACT (s)	Lines	Gap (%)
MHIA	1	3.94	113	5	1.0
	2	3.74	118	4	0.8
	3	3.55	91	5	0.6
MILP	1	3.98	261,445	8	—
	2	3.77	5412	8	—
	3	3.57	6857	7	—

having more (slightly) different lines might also make it more difficult for the passengers to understand the network. Probably, the network obtained by MHIA is even more realistic than the result obtained by MILP.

Now we apply the same comparison between MHIA and MILP on the medium and large HSR networks. Here, only the lines included in the line plans obtained by MHIA in different PDS are included. Otherwise, the pool of lines becomes too large to solve the MILP model. The differences between MHIA and MILP are shown in Table 12. The optimality gap of the CPLEX solver used in MILP is the same as in MHIA, i.e., 5% for both HSR networks.

The numbers between brackets in Lines column indicate the number of lines considered in the line pool of MILP, while the regular number indicates the number of lines involved in the results. In theory, the solutions obtained by MILP should be no less than the results obtained by MHIA. However, due to the 5% optimality gap of the CPLEX solver,

Table 11: Results for the given line pool for MHIA and MILP on the small HSR network.

Lines in MILP	Frequencies (MHIA)	Frequencies (MILP)
4, 9, 8, 10	3	1
5, 4, 2, 0, 1, 3, 7, 6	2	1
0, 1, 8, 7, 6	2	1
2, 0, 1, 8, 7	1	0
3, 1, 8, 9	1	0
4, 2, 0, 1, 8, 10	—	2
4, 9, 8, 7, 6	—	1
4, 9, 8, 7	—	1
5, 4, 9, 8, 7, 6	—	1
3, 1, 8	—	1
5, 4, 2, 0, 1, 3, 7	—	0
2, 0, 1, 8, 7, 6	—	0
0, 1, 8, 10	—	0
0, 1, 8, 7	—	0
0, 2, 4	—	0
1, 8, 7	—	0
5, 4, 9, 8	—	0
8, 10	—	0
3, 1	—	0
4, 2, 0, 1, 8	—	0
7, 3, 1	—	0
4, 5	—	0
5, 4, 9, 8, 10	—	0
4, 2, 0, 1, 3, 7	—	0
0, 1, 8	—	0
1, 0, 2, 4	—	0
3, 1, 0, 2, 4	—	0
1, 8, 10	—	0
1, 8, 9	—	0
1, 0, 2, 4, 5	—	0
6, 7, 8	—	0
3, 1, 8, 10	—	0
0, 1, 8, 7, 3	—	0
4, 9, 8	—	0
3, 1, 0, 2, 4, 5	—	0

Table 12: Results of the given line pool for MILP on medium and large HSR networks.

HSR networks	Approaches	PDS	BS ($*10^7$)	ACT (min)	Lines	Gap (%)
Medium	MILP	1	2.40	2.7	12 (16)	—
		2	2.50	534.7	25 (34)	—
		3	2.37	5.3	17 (25)	—
	MHIA	1	2.45	172	7	−2.0
		2	2.40	214	7	4.0
		3	2.35	175	9	0.8
Large	MILP	1	6.69	6.6	17 (17)	—
		2	6.94	53	24 (27)	—
		3	—	—	(36)	—
	MHIA	1	6.79	246	11	−1.5
		2	6.86	288	11	1.2
		3	6.25	312	11	—

it can happen that the profit of MHIA is higher than the profit of MILP, leading to a negative gap. This happens for the first demand scenario both for the medium and large networks. The calculation of PDS 3 in large network by MILP could not be obtained since the solver runs out of memory due to the size of the network and the large line pool.

For both networks, we can see in Table 12 that the optimal results given by MILP are very close to the best results obtained by MHIA for all demand scenarios. It illustrates that MHIA obtains high quality results.

6. Conclusions

In this paper, we present a matheuristic iterative approach (MHIA) for profit-oriented line planning and frequency setting, applied to a high-speed railway (HSR) network. Profit-oriented line planning considers both the operator's cost and the passenger travel time. The passenger travel time is considered by reducing the ticket price (and thus operator revenues) in case of detours or transfers. A mathematical model is discussed to define the problem in detail. MHIA integrates heuristic improvements of the line plan with an exact approach for passenger assignment and frequency setting. Two intensification and two diversification moves are considered during the algorithm.

The performance of MHIA is assessed experimentally on networks of different sizes and with different passenger demand scenarios. On the smallest network with only 7 nodes, MHIA has an average gap with the optimal solution of 2.9%. On several real-world instances based on the Chinese HSR network, MHIA improves the profit with 10.7%, 8.3%, and 13.3% on average of different size networks compared to an alternative iterative approach (AIA) which does not use the exact passenger assignment and frequency setting. Besides, MHIA increases the initial solutions of small, medium, and large networks with 40.0%, 16.5%, and 10.7%, respectively. The average computation times for the small, medium, and large network instances are 117.3 seconds, 187 minutes, and 4.7 hours, respectively.

The experiments also provide useful insights into the parameter settings when the network becomes large. With adapted parameters for larger networks, MHIA acquires high quality solutions within reasonable computation times. This confirmed by comparing the performance with the exact solution approach MILP. The best profits for all three networks are very close to the optimal solutions using MILP, when all MHIA lines are included in the line pool. For the small, medium, and large HSR networks, the gaps between the best MHIA solution and best MILP solution are 0.8%, 0.9%, and −0.2%, respectively, when taking the high-quality lines as line pools.

The large HSR network with 34 major stations where tracks split or join is considered in this paper. Comparing to all the stations of this network (more than 800 stations), the number of stations considered is still small. But all the 4V4H high-speed tracks are involved. The intermediate stations between the major stations are not considered explicitly, but stopping there could be included in the travel time between

two major stations. This provides the possibility for further study on the optimal stopping pattern and other subsequent issues.

Further work could consider different train sizes and speeds and an optimization of the stopping patterns. It would also be interesting to find out how the profit-oriented line planning can be used or modified for operator-oriented or passenger-oriented line planning. Another possibility would be to integrate AIA and MHIA in order to find an approach that can further reduce the computation time while keeping the same solution quality.

Data Availability

The data used to support the findings of this study are available online at https://www.mech.kuleuven.be/en/cib/lp/mainpage#section-4.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (no. 2017YFB1200700) and the National Natural Science Foundation of China (no. U1834209). The author acknowledges the support of the China Scholarship Council (201707000074).

References

[1] H. Fu, L. Nie, L. Meng, B. R. Sperry, and Z. He, "A hierarchical line planning approach for a large-scale high speed rail network: the China case," *Transportation Research Part A: Policy and Practice*, vol. 75, pp. 61–83, 2015.

[2] F. Liu, Z. Sun, P. Zhang, Q. Peng, and Q. Qiao, "Analyzing capacity utilization and travel patterns of Chinese high-speed trains: an exploratory data mining approach," *Journal of Advanced Transportation*, vol. 2018, Article ID 3985302, 9 pages, 2018.

[3] G. Desaulniers and M. D. Hickman, "Public transit," *Handbooks in Operations Research and Management Science*, vol. 14, pp. 69–127, North Holland, Amsterdam, Netherland, 2007.

[4] R. Borndörfer, O. Arslan, Z. Elijazyfer et al., "Line planning on path networks with application to the istanbul metrobus," in *Proceedings of the 2016 Operations Research*, Springer, Cham, Switzerland, pp. 235–241, 2018.

[5] Y. Yang, J. Li, C. Wen, P. Huang, Q. Peng, and J. Lessan, "A bi-level passenger preference-oriented line planning model for high-speed railway operations," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 10, pp. 224–235, 2018.

[6] K. Steiner and S. Irnich, "Schedule-based integrated intercity bus line planning via branch-and-cut," *Transportation Science*, vol. 52, no. 4, pp. 882–897, 2018.

[7] D. Canca, A. De-Los-Santos, G. Laporte, and J. A. Mesa, "Integrated railway rapid transit network design and line planning problem with maximum profit," *Transportation Research Part E: Logistics and Transportation Review*, vol. 127, pp. 1–30, 2019.

[8] M. T. Claessens, N. M. van Dijk, and P. J. Zwaneveld, "Cost optimal allocation of rail passenger lines," *European Journal of Operational Research*, vol. 110, no. 3, pp. 474–489, 1998.

[9] J.-W. Goossens, S. Van Hoesel, and L. Kroon, "A branch-and-cut approach for solving railway line-planning problems," *Transportation Science*, vol. 38, no. 3, pp. 379–393, 2004.

[10] M. R. Bussieck, T. Lindner, and M. E. Lübbecke, "A fast algorithm for near cost optimal line plans," *Mathematical Methods of Operations Research (ZOR)*, vol. 59, no. 2, pp. 205–220, 2004.

[11] J.-W. Goossens, S. van Hoesel, and L. Kroon, "On solving multi-type railway line planning problems," *European Journal of Operational Research*, vol. 168, no. 2, pp. 403–424, 2006.

[12] M. R. Bussieck, *Optimal Lines in Public Rail Transport*, Verlag nicht ermittelbar, 1998.

[13] M. R. Bussieck, P. Kreuzer, and U. T. Zimmermann, "Optimal lines for railway systems," *European Journal of Operational Research*, vol. 96, no. 1, pp. 54–63, 1997.

[14] A. Schöbel and S. Scholl, "Line planning with minimal traveling time," in *Proceedings of the 5th Workshop on Algorithmic Methods and Models for Optimization of Railways (ATMOS'05)*, Schloss Dagstuhl-Leibniz-Zentrum für Informatik, Mallorca, Spain, October 2006.

[15] S. Scholl, *Customer-Oriented Line Planning*, Dissertation. de, Dissertation. de, 2005.

[16] K. Nachtigall and K. Jerosch, "Simultaneous network line planning and traffic assignment," in *Proceedings of the 8th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (ATMOS'08)*, Schloss Dagstuhl-Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2008.

[17] M. E. Pfetsch and R. Borndörfer, "Routing in line planning for public transport," in *Proceedings of the 2005 Operations Research*, Springer, Berlin, Heidelberg, Germany, pp. 405–410, 2006.

[18] R. Rosalia, "The line planning with simultaneous optimisation of operational cost and travel time," NTNU, 2017.

[19] M. Friedrich, M. Hartl, A. Schiewe, and A. Schöbel, "Integrating passengers' assignment in cost-optimal line planning," in *Proceedings of the 17th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2017)*, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, Vienna, Austria, September 2017.

[20] R. Borndörfer and M. Karbstein, "A direct connection approach to integrated line planning and passenger routing," in *Proceedings of the 12th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, Ljubljana, Slovenia, September 2012.

[21] M. Karbstein, "Integrated line planning and passenger routing: connectivity and transfers," in *Proceedings of the 2014 Operations Research*, Springer, Cham, Switzerland, pp. 263–269, 2016.

[22] M. Schmidt and A. Schöbel, "The complexity of integrating passenger routing decisions in public transportation models," *Networks*, vol. 65, no. 3, pp. 228–243, 2015.

[23] L. A. Silman, Z. Barzily, and U. Passy, "Planning the route system for urban buses," *Computers & Operations Research*, vol. 1, no. 2, pp. 201–211, 1974.

[24] K. Kepaptsoglou and M. Karlaftis, "Transit route network design problem: review," *Journal of Transportation Engineering*, vol. 135, no. 8, pp. 491–505, 2009.

[25] G. Laporte, A. Marín, J. A. Mesa, and F. Perea, "Designing robust rapid transit networks with alternative routes," *Journal of Advanced Transportation*, vol. 45, no. 1, pp. 54–65, 2011.
[26] Á. Marín and R. García-Ródenas, "Location of infrastructure in urban railway networks," *Computers & Operations Research*, vol. 36, no. 5, pp. 1461–1477, 2009.
[27] D. Canca, A. De-Los-Santos, G. Laporte, and J. A. Mesa, "An adaptive neighborhood search metaheuristic for the integrated railway rapid transit network design and line planning problem," *Computers & Operations Research*, vol. 78, pp. 1–14, 2017.
[28] R. Borndörfer, M. Grötschel, and M. E. Pfetsch, "A column-generation approach to line planning in public transport," *Transportation Science*, vol. 41, no. 1, pp. 123–132, 2007.
[29] L. Ahmed, C. Mumford, and A. Kheiri, "Solving urban transit route design problem using selection hyper-heuristics," *European Journal of Operational Research*, vol. 274, no. 2, pp. 545–559, 2019.
[30] V. Schmid, "Hybrid large neighborhood search for the bus rapid transit route design problem," *European Journal of Operational Research*, vol. 238, no. 2, pp. 427–437, 2014.
[31] M. Goerigk and M. Schmidt, "Line planning with user-optimal route choice," *European Journal of Operational Research*, vol. 259, no. 2, pp. 424–436, 2017.
[32] A. Schöbel, "Line planning in public transportation: models and methods," *OR Spectrum*, vol. 34, no. 3, pp. 491–510, 2012.
[33] M. E. Schmidt, *Integrating Routing Decisions in Public Transportation Problems*, Springer, New York, NY, USA, 2014.
[34] D. Liu, P. Vansteenwegen, G. Lu et al., "An iterative approach for profit-oriented railway line planning," in *Proceedings of the 8th International Conference on Railway Operations Modelling and Analysis (ICROMA), RailNorrköping*, Linköping University Electronic Press, Linköpings universitet, Norrköping, Sweden, June 2019.
[35] D. Canca, E. Barrena, A. De-Los-Santos, and J. L. Andrade-Pineda, "Setting lines frequency and capacity in dense railway rapid transit networks with simultaneous passenger assignment," *Transportation Research Part B: Methodological*, vol. 93, pp. 251–267, 2016.
[36] Z. Gao, H. Sun, and L. L. Shan, "A continuous equilibrium network design model and algorithm for transit systems," *Transportation Research Part B: Methodological*, vol. 38, no. 3, pp. 235–250, 2004.
[37] W. Wu, R. Liu, W. Jin, and C. Ma, "Stochastic bus schedule coordination considering demand assignment and rerouting of passengers," *Transportation Research Part B: Methodological*, vol. 121, pp. 275–303, 2019.
[38] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*, Prentice-Hall, Upper Saddle River, NJ, USA, 1993.
[39] D. Pisinger and S. Ropke, "Large neighborhood search," *Handbook of Metaheuristics*, pp. 399–419, Springer, Boston, MA, USA, 2010.