Chapter 1

Line Planning Problem

Güvenç Şahin, Niels Lindner and Thomas Schlechte

Key words: public transport, transport planning, line planning, line optimization, integer programming

1.1 Introduction

A line in a public transit system is designated with an origin location, a destination location and a path designated with several intermediate stops between these two locations. The complete set of lines in the system makes up the skeleton of the transit services.

Public transportation planning requires various stages of planning that handle different but inherently interrelated decisions and cope with different temporal scope. Strategic planning covers long-

Güvenç Şahin

Sabanci University, Orhanli, Tuzla, 34956 Istanbul, e-mail: guvencs@sabanciuniv.edu

Amazon Web Services, Krausenstraße 38, 10117 Berlin, e-mail: guvencs@amazon.de

Niels Lindner

Freie Universität Berlin, Institut für Mathematik, Arnimallee 6, 14195 Berlin, e-mail: lindner@zib.de

Zuse Institute Berlin, Takustraße 7, 14195 Berlin, e-mail: lindner@zib.de

Thomas Schlechte

LBW Optimization GmbH, Englerallee 19, 14915 Berlin, e-mail: schlechte@lbw-optimization.de

term decisions in service design, among which network design, *line planning*, timetabling, and fare planning are included; such decisions have impact over a long planning horizon covering multiple years. The tactical planning level is concerned with attainment, allocation and usage of resources such as vehicles and crew; the temporal scope of tactical plans is typically several months that may extend to a year.

The operational plans are involved with actions on the day of operations; these include assignment of individual resource units to specific operations and services, traffic (and delay) management, and real-time scheduling of services. Line planning, as part of strategic decisions, lays out the foundation of all subsequent plans, decisions and operations. Lampkin and Saalmans [7] report the first systematic operational research approach for public transportation planning, refer to the line planning stage as choosing a set of routes, and identify it as the first task of a sequence of planning decisions/tasks.

The line plan of a public transportation (PT) system can be described as follows. Let a graph G represent the physical infrastructure of the PT network. Depending on the type of the PT, a node/vertex of G corresponds to a station, a terminal or a stop. Without loss of generality, a line corresponds to a walk on G from $s \in G$ to $t \in G$. Considering a line plan in a railway context, G could have stations as vertices and railway links between stations as edges. A railway line can be represented by a walk ℓ on G and a frequency f_ℓ over a pre-determined fixed length portion of the planning horizon. If $f_\ell = 2$ and frequencies are determined hourly, then line ℓ is supposed to run twice per every hour, as in Figure 1.1.

Definition 1 (Line Plan) A *line plan* on a graph G is a set $\mathcal{L}$ of walks (= *lines*) on G, together with positive integer *frequencies* $f_\ell \in \mathbb{N}$ for each $\ell \in \mathcal{L}$.

The aim of the *line planning problem* (LPP) is to find a feasible line plan providing both convenient travel for passengers and low operational costs. In practice, the feasibility of line plans can depend on various physical constraints associated with the infrastructure as well as the availability of resources such as vehicles and crew. For example, a railway infrastructure can only accommodate a certain maximum number of trains within a given time window. Meanwhile, line planning

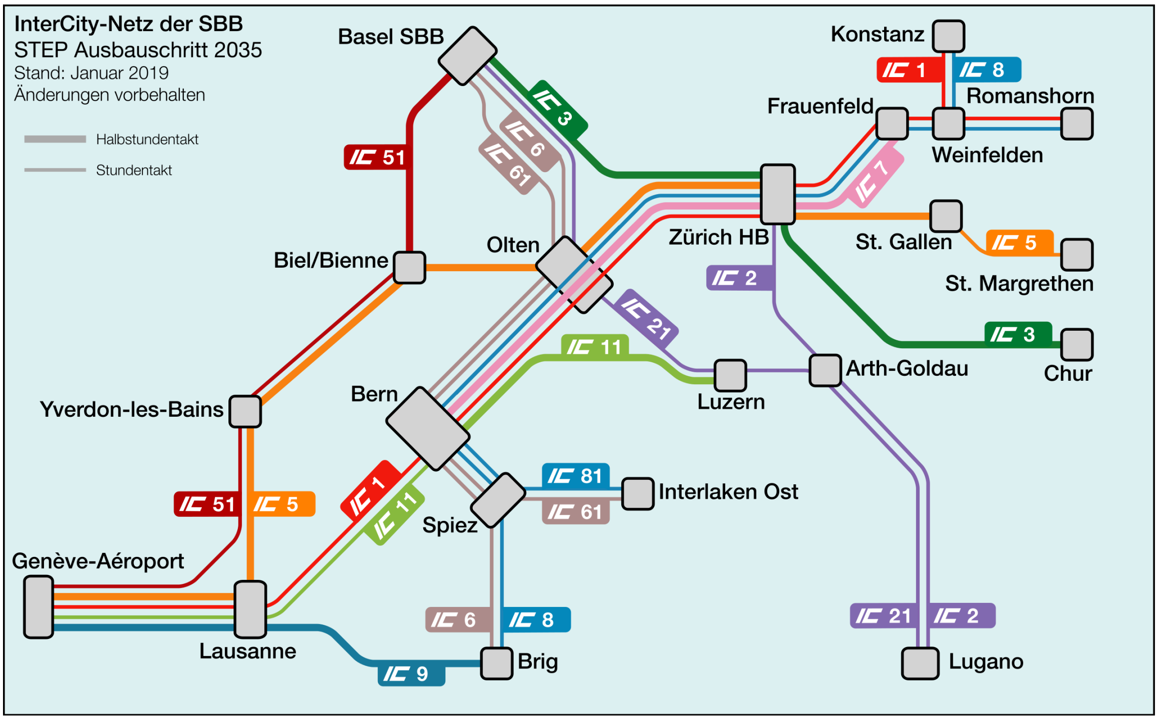

Fig. 1.1 A line plan of the projected future Swiss intercity network. Thick lines run twice per hour (frequency 2), thin lines only once (frequency 1). Source: Filzstift via commons.wikimedia.org, CC-BY-SA 4.0.

needs to balance two conflicting goals: The higher the frequencies, the shorter the travel times of the passengers, but the higher the operational expenses for fleet and crew. Indeed, the LPP is a prime example for a combinatorial optimization problem. Nevertheless, it was only in 1997 when Bussieck et al. [2] showed that the feasibility variant of the LPP is NP-complete by reducing the exact cover by 3-sets problem (X3C).

Bussieck et al. [2] developed the first integer programming (IP) problem formulation for the LPP. The objective function of this formulation is to maximize the number of passengers travelling from their origins to destinations directly, i.e., by using only one line. Schöbel [10] later classify this approach as passenger-oriented due to its objective function. In the early works of passenger-oriented line of research, Schöbel and Scholl [11] and Rittner and Nachtigall [9] minimize passengers total riding time. Alternatively, Klier and Haase [6] maximize the number of serviced passengers, i.e., the amount of transit demand.

The cost-oriented approach, in contrast to passenger-oriented approach, relies on minimizing the costs of the system mostly focusing on the operational costs. As the earliest example in this line of research, Claessens et al. [3] provide a non-linear IP formulation, while Goossens et al. [4] extend their formulation. Torres et al. [13] provide some theoretical complexity results on special cases of basic formulations based on a cost-oriented approach.

In their seminal work, Borndörfer et al. [1] present an IP formulation with an objective function minimizing both the passenger riding/travel times and operational costs, integrating the passenger-oriented and cost-oriented approaches. Building upon this work, Karbstein [5] tunes the model towards less transfers and connects it to the Steiner connectivity problem.

1.2 Models

Considering a finite-length planning horizon and the physical infrastructure of the PT system, the minimal product of the LPP is a set of lines and their corresponding frequencies which correspond to the number of times the line service is run during the planning horizon.

Definition 2 (Basic Line Planning Problem) Let G be a graph, let $\mathcal{P}$ be a set of walks on G (*line pool*), and let $f_e^{\min}, f_e^{\max} \in \mathbb{N}_0$ be lower resp. upper frequency bounds for each edge $e \in E(G)$. The *basic line planning problem* (BLPP) is to find a line plan $(\mathcal{L}, f)$ such that

$$\mathcal{L} \subseteq \mathcal{P} \quad \text{and} \quad \forall e \in E(G): \quad f_e^{\min} \leq \sum_{\ell \in \mathcal{L}} f_\ell \leq f_e^{\max}. \tag{1.1}$$

The BLPP is a feasibility problem, and the set of potential lines is constrained to be part of the pre-specified line pool $\mathcal{P}$. The BLPP is characterized with the following features:

- The line pool is used to reflect that not every walk in G is a feasible line due to various constraints including but not limited to infrastructure restrictions, e.g., missing switches or turning facilities and length restrictions, e.g, limited range for electric vehicles.

- It depends on the application context whether G should be directed or not, and similarly, whether lines should be (un)directed simple paths or be allowed to contain cycles.
- The lower frequency bounds $f^{\min}$ ensure a minimum service level, while the upper frequency bounds $f^{\max}$ limit the operational costs and model the infrastructure capacity of an edge.

It is straightforward to model the BLPP as an IP problem with the general integer variables $f_\ell \in \mathbb{N}_0$ for each $\ell \in \mathcal{P}$, with the interpretation that $f_\ell > 0$ if and only if $\ell \in \mathcal{L}$. However, the BLPP formulation doees not necessarily consider either the transit demand or the associated costs and therefore it is not realistic for practical purposes.

Cost-oriented Line Planning

In order to describe a cost-oriented version of the BLPP, a line in $\mathcal{P}$ is associated with an activation cost $C_\ell \in \mathbb{R}$, which is independent of the frequency, and an operating cost $c_\ell \in \mathbb{R}$. Accordingly, the cost-oriented BLPP can be defined as

$$\text{minimize} \quad \sum_{\ell \in \mathcal{L}} (C_\ell + c_\ell f_\ell) \tag{1.2}$$

$$\text{s.t.} \quad f_e^{\min} \leq \sum_{\ell \in \mathcal{L}} f_\ell \leq f_e^{\max}, \qquad \forall e \in E(G). \tag{1.3}$$

In practice, C_ℓ and c_ℓ can in general only be estimated, as line planning is typically performed before allocation and scheduling of resources such as vehicle and crew, and hence associated expenses such as fuel and driver costs are yet unknown.

Passenger-oriented Line Planning

The passenger demand is most commonly modeled by an origin-destination matrix (*OD matrix*) with entries d_{st} for all pairs (s, t) of vertices in G, estimating the number of passengers that want to travel from s to t. A pair (s, t) with $d_{st} > 0$ is called an *OD pair*, the set of all OD pairs is denoted by $\mathcal{D}$. In passenger-oriented line planning, the objective is then to find a line plan $(\mathcal{L}, f)$ satisfying (1.1) that minimizes the total travel time for each OD pair. However, it is quite challenging to assess the travel time as it first requires answering the following questions that will lead to a set of assumptions to frame the problem:

- *Fixed vs. integrated routing:* Do passengers always travel along the same s-t-path or do they rather take the shortest route offered by the line plan?
- *Capacitated vs. uncapacitated routing:* Can passengers take different routes depending on vehicle capacities?
- *Travel time vs. time in vehicle:* Should only in-vehicle time be considered, or should transfers and adaption times be taken in to account as well? Should waiting in vehicles be preferred over waiting at stops (perceived travel time)?

Pasenger-oriented approaches, at a minimum, require defining a travel time τ_p for a path $p \in P_{st}$ where P_{st} represents the set of all s-t-paths in G for OD pair (s,t). In essence, the passenger-oriented BLPP is to be formulated to minimize $\sum_{(s,t)\in\mathcal{D}} \sum_{p\in P_{st}} \tau_p y_p$ while the conditions on satisfaction of transit demand and dependency of the paths on the chosen set of lines require an elaborate modelling approach which could be represented as an IP problem formulation.

1.3 Formulations

In order to develop a path-based mixed-integer programming model for the LPP, it is necessary to predetermine a set of lines $\mathcal{L}$ from which a subset of lines is to be selected and a set of paths P_{st} for all OD pairs $(s,t) \in \mathcal{D}$. If $E(p)$ is the set of edges on path p and $E(l)$ is the set of edges covered by line l, the following problem formulation considers in-vehicle times and capacitated routing constraints [1]:

$$\text{minimize} \qquad \alpha \sum_{\ell\in\mathcal{L}} (C_\ell x_\ell + c_\ell f_\ell) + \beta \sum_{(s,t)\in\mathcal{D}} \sum_{p\in P_{st}} \tau_p y_p \tag{1.4}$$

$$
\begin{aligned}
\text{s.t.} \quad & \sum_{p \in P_{st}} y_p = d_{st} && (s,t) \in \mathcal{D} && (1.5)\\
& \sum_{(s,t)\in\mathcal{D}} \sum_{p \in P_{st}: e \in E(p)} y_p \leq \sum_{\ell \in \mathcal{L}: e \in E(\ell)} \kappa_\ell f_\ell && e \in E(G) && (1.6)\\
& \sum_{\ell \in \mathcal{L}: e \in E(\ell)} f_\ell \leq f_e^{\max} && e \in E(G) && (1.7)\\
& f_\ell \leq F x_\ell && \ell \in \mathcal{L} && (1.8)\\
& f_\ell \in \mathbb{N}_0 && \ell \in \mathcal{L} && (1.9)\\
& x_\ell \in \{0,1\} && \ell \in \mathcal{L} && (1.10)\\
& y_p \geq 0 && p \in P_{st}, (s,t) \in \mathcal{D} && (1.11)
\end{aligned}
$$

where the binary variable x_ℓ (1.10) indicates whether line ℓ is part of the chosen line plan or not, the integer variable f_ℓ (1.9) denotes the frequency of ℓ, and the continuous variable y_p counts the number of passengers from s to t using path p.

The objective (1.4) is a scalarized bicriteria objective function, weighing the operational costs by α and the passenger travel time by β. As passenger numbers are only estimates, it is reasonable not to require y_p to be integer. Of course, the total demand d_{st} must be distributed among all paths $p \in P_{st}$ (1.5). Introducing a vehicle capacity κ_ℓ for each line ℓ, the capacitated routing is established by (1.6): The total number of passengers using edge e must not exceed the total vehicle capacity on e. In particular, the line frequencies are driven by passenger demand. This coupling also replaces the lower frequency bound requirement from BLPP, the upper frequency bound remains (1.7). The upper bound F enforces that $f_\ell = 0$ whenever $x_\ell = 0$ (1.8).

Arc-based vs. path-based: The nature of the passenger demand is path-based, i.e., the transit demand for an OD pair (s,t) is distributed over the paths in P_{st}. This representation trivially allows the models to consider the distribution of the demand over the paths as part of the decisions underlying the line selection decision as demonstrated in the problem formulation (1.4)-(1.11). Since this inflates the degree of freedom in the solution space by allowing the passengers of the same (s,t) pair travel on different and the number of such paths can can in general be quite large, the resulting formulations become too large to solve to optimality. An alternative is to use an arc-based formulation where demand is satisfied by the most convenient path, i.e. usually the shortest or the fastest, and the transit demand over each arc can be calculated through the fixed routing scheme.

Line pool vs. unrestricted lines: In order to overcome the challenges inherent to the path-based formulations, a predetermined set of lines with a limited number of lines is used in the formulations while generating all possible lines over G of the PT leads to impractically large and therefore untractable formulations. Column generation techniques have been successfully applied to LPPs when lines are no longer restricted to be selected from a predetermined line pool. Though the pricing problem is a longest path problem, it is still solvable for practical instances. In this setting, it is possible to enforce length restrictions by constraints for practical purposes.

1.4 Conclusions

With the advancement in computational technology, the scope of line planning problem is now enriched with various practical considerations.

Integration with other stages

Line planning is intertwined with other public transport planning tasks on the strategic level, e.g., timetabling. An innovative approach to integrate line planning with more planning problems are the eigenmodel by Schöbel [12].

Multi-period planning

Public transport systems experience a large amount of variation and fluctuation in demand over the day. It is therefore desirable to consider line plans that are responsive to varying demand within the planning horizon. This multi-period line planning is able to outperform combining line plans for static demand [14].

Symmetric Line Plans

For special network layouts, e.g., those that adhere to a specific symmetry, one might consider line plans that respect this symmetry as well. It turns out that symmetric line plans are not always optimal, but approximation results can be provided [8].

See also

Branch and Price: Integer Programming with Column Generation, Modeling Difficult Optimization Problems.

References

[1] Borndörfer R, Grötschel M, Pfetsch ME (2007) A Column-Generation Approach to Line Planning in Public Transport. Transportation Science 41(1):123–132

[2] Bussieck MR, Kreuzer P, Zimmermann UT (1997) Optimal lines for railway systems. European Journal of Operational Research 96(1):54–63

[3] Claessens MT, van Dijk NM, Zwaneveld PJ (1998) Cost optimal allocation of rail passenger lines. European Journal of Operational Research 110(3):474–489

[4] Goossens JW, van Hoesel S, Kroon L (2006) On solving multi-type railway line planning problems. European Journal of Operational Research 168(2):403–424

[5] Karbstein M (2013) Line Planning and Connectivity. PhD thesis, Technische Universität Berlin

[6] Klier MJ, Haase K (2008) Line optimization in public transport systems. In: Kalcsics J, Nickel S (eds) Operations Research Proceedings 2007, Springer Berlin Heidelberg, Berlin, Heidelberg

[7] Lampkin W, Saalmans P (1967) The design of routes, service frequencies, and schedules for a municipal bus undertaking: A case study. Journal of the Operational Research Society 18(4):375–397

[8] Masing B, Lindner N, Borndörfer R (2022) The price of symmetric line plans in the Parametric City. Transportation Research Part B: Methodological 166:419–443

[9] Rittner M, Nachtigall K (2009) Simultane Liniennetz- und Taktfahrlagenoptimierung. Der Eisenbahningenieur 60(6):6–10

[10] Schöbel A (2012) Line planning in public transportation: models and methods. OR Spectrum 34(3):491–510

[11] Schöbel A, Scholl S (2006) Line planning with minimal traveling time. In: Kroon LG, Möhring RH (eds) 5th Workshop on Algorithmic Methods and Models for Optimization of Railways (ATMOS'05), Dagstuhl, Germany, OpenAccess Series in Informatics (OASIcs), vol 2

[12] Schöbel A (2017) An eigenmodel for iterative line planning, timetabling and vehicle scheduling in public transportation. Transportation Research Part C: Emerging Technologies 74:348–365

[13] Torres LM, Torres R, Borndörfer R, Pfetsch ME (2008) Line planning on paths and tree networks with applications to the quito trolebus system. In: Fischetti M, Widmayer P (eds) 8th Workshop on Algorithmic Methods and Models for Optimization of Railways (ATMOS'08), Dagstuhl, Germany, OpenAccess Series in Informatics (OASIcs), vol 9

[14] Şahin G, Ahmadi Digehsara A, Borndörfer R, Schlechte T (2020) Multi-period line planning with resource transfers. Transportation Research Part C: Emerging Technologies 119:102726