Zuse Institute Berlin

Takustrasse 7
D-14195 Berlin-Dahlem
Germany

HEIDE HOPPMANN[1]

# An Extended Formulation for the Line Planning Problem*

[1] Zuse Institute Berlin, Takustr. 7, 14195 Berlin, Germany; Email hoppmann@zib.de

* This research was carried out in the framework of MATHEON supported by Einstein Foundation Berlin.

ZIB Report 16-08 (February 2016)

Zuse Institute Berlin
Takustrasse 7
D-14195 Berlin-Dahlem

Telefon: 030-84185-0
Telefax: 030-84185-125

e-mail: bibliothek@zib.de
URL: http://www.zib.de

ZIB-Report (Print) ISSN 1438-0064
ZIB-Report (Internet) ISSN 2192-7782

**Abstract** In this paper we present a novel extended formulation for the line planning problem that is based on what we call "configurations" of lines and frequencies. Configurations account for all possible options to provide a required transportation capacity on an infrastructure edge. The proposed configuration model is strong in the sense that it implies several facet-defining inequalities for the standard model: set cover, symmetric band, MIR, and multicover inequalities. These theoretical findings can be confirmed in computational results. Further, we show how this concept can be generalized to define configurations for subsets of edges; the generalized model implies additional inequalities from the line planning literature.

## 1 Introduction

*Line planning* is an important strategic planning problem in public transport. The task is to find a set of lines and frequencies such that a given demand can be transported. There are usually two main objectives: minimizing the travel times of the passengers and minimizing the line operating costs.

Since the late 90s, the line planning literature has developed a variety of integer programming approaches that capture different aspects, for an overview see Schöbel [10]. Bussieck, Kreuzer, and Zimmermann [5] propose an integer programming model to maximize the number of direct travelers. Operating costs are discussed for instance in the article of Goossens, van Hoesel, and Kroon [7]. Schöbel and Scholl [11] and Borndörfer and Karbstein [3] focus on the number of transfers and the number of direct travelers, respectively, and further integrate line planning and passenger routing in their models. Borndörfer, Grötschel, and Pfetsch [1] also propose an integrated line planning and passenger routing model that allows a dynamic generation of lines.

All these models employ some type of *capacity* or *frequency demand constraints* in order to cover a given demand. In this paper we propose a concept to strengthen such constraints by means of a novel extended formulation. The idea is to enumerate the set of possible *configurations* of line frequencies for each capacity constraint. We show that such an extended formulation implies general facet defining inequalities for the standard model.

We consider the following basic *line planning problem*: We are given an undirected graph $G = (V,E)$ representing the transportation network; a *line* is a simple path in $G$ and we denote by $\mathcal{L} = \{\ell_1,\dots,\ell_n\}$, $n \in \mathbb{N}$, the given set of lines. We denote by $\mathcal{L}(e) := \{\ell \in \mathcal{L} : e \in \ell\}$ the set of lines on edge $e \in E$. Furthermore, we are given an ordered set of *frequencies* $\mathcal{F} = \{f_1,\dots,f_k\} \subseteq \mathbb{N}$, $k \in \mathbb{N}$, such that $0 < f_1 < \dots < f_k$, and we define $\mathcal{F}_0 := \mathcal{F} \cup \{0\}$. The cost of operating line $\ell \in \mathcal{L}$ at frequency $f \in \mathcal{F}$ is given by $c_{\ell,f} \in \mathbb{Q}_{\geq 0}$. Finally, each edge in the network bears a positive frequency demand $F(e) \in \mathbb{N}$; it gives the number of line operations that are necessary to cover the demand on this edge.

A *line plan* $(\bar{\mathcal{L}}, \bar{f})$ consists of a subset $\bar{\mathcal{L}} \subseteq \mathcal{L}$ of lines and an assignment $\bar{f} : \bar{\mathcal{L}} \to \mathcal{F}$ of frequencies to these lines. A line plan is *feasible* if the frequencies of the lines satisfy the frequency demand $F(e)$ for each edge $e \in E$, i.e., if

$$\sum\nolimits_{\ell \in \bar{\mathcal{L}}(e)} \bar{f}(\ell) \geq F(e) \text{ for all } e \in E. \tag{1}$$

We define the cost of a line plan $(\bar{\mathcal{L}}, \bar{f})$ as $c(\bar{\mathcal{L}}, \bar{f}) := \sum_{\ell \in \bar{\mathcal{L}}} c_{\ell, \bar{f}(\ell)}$. The *line planning problem* is to find a feasible line plan of minimal cost.

## 2 Standard Model and Extended Formulation

A common way to formulate the line planning problem uses binary variables $x_{\ell,f}$ indicating whether line $\ell \in \mathcal{L}$ is operated at frequency $f \in \mathcal{F}$. In our case, this results in the following *standard model*:

$$\begin{aligned}
\text{(SLP)} \quad \min \quad & \sum_{\ell \in \mathcal{L}} \sum_{f \in \mathcal{F}} c_{\ell,f}\, x_{\ell,f} & & \\
\text{s.t.} \quad & \sum_{\ell \in \mathcal{L}(e)} \sum_{f \in \mathcal{F}} f\, x_{\ell,f} \geq F(e) & \forall\, e \in E & \quad (2)\\
& \sum_{f \in \mathcal{F}} x_{\ell,f} \leq 1 & \forall\, \ell \in \mathcal{L} & \quad (3)\\
& x_{\ell,f} \in \{0,1\} & \forall\, \ell \in \mathcal{L}, \forall\, f \in \mathcal{F}. & \quad (4)
\end{aligned}$$

Model (SLP) minimizes the cost of a line plan. The *frequency demand constraints* (2) ensure that the frequency demand is covered while the *assignment constraints* (3) guarantee that every line is operated at at most one frequency.

In the following, we give an extended formulation for (SLP) that aims at tightening the LP-relaxation. Our extended formulation is based on the observation that the frequency demand of an edge can also be expressed by specifying the minimum number of lines that have to be operated at each frequency. We call these frequency combinations *minimal configurations* and a formal description is as follows.

**Definition 1.** For $e \in E$ define the set of *(feasible) configurations of e* by

$$\bar{\mathcal{Q}}(e) := \left\{ q = (q_{f_1}, \ldots, q_{f_k}) \in \mathbb{Z}^{\mathcal{F}}_{\geq 0} : \sum\nolimits_{f \in \mathcal{F}} q_f \leq |\mathcal{L}(e)|, \sum\nolimits_{f \in \mathcal{F}} f\, q_f \geq F(e) \right\}$$

and the set of *minimal configurations of e* by

$$\mathcal{Q}(e) := \left\{ q \in \bar{\mathcal{Q}}(e) : (q_{f_1}, \ldots, q_{f_i} - 1, \ldots, q_{f_k}) \notin \bar{\mathcal{Q}}(e) \quad \forall\, i = 1, \ldots, k \right\}.$$

As an example, consider an edge with frequency demand of 9. Let there be three lines on this edge, that each can be operated at frequency 2 or 8. To cover this demand we need at least two lines with frequency 8 or one line with frequency 2 and one line with frequency 8.

We extend (SLP) using binary variables $y_{e,q}$ that indicate for each edge $e \in E$ which configuration $q \in \mathcal{Q}(e)$ is chosen. This results in the following formulation:

$$
\begin{aligned}
(\text{QLP}) \quad \min \quad & \sum_{\ell\in\mathcal{L}} \sum_{f\in\mathcal{F}} c_{\ell,f}\, x_{\ell,f} \\
\text{s.t.} \quad & \sum_{\ell\in\mathcal{L}(e)} x_{\ell,f} \geq \sum_{q\in\mathcal{Q}(e)} q_f\, y_{e,q} && \forall\, e\in E, \forall\, f\in\mathcal{F} && (5)\\
& \sum_{q\in\mathcal{Q}(e)} y_{e,q} = 1 && \forall\, e\in E && (6)\\
& \sum_{f\in\mathcal{F}} x_{\ell,f} \leq 1 && \forall\, \ell\in\mathcal{L} && (7)\\
& x_{\ell,f} \in \{0,1\} && \forall\, \ell\in\mathcal{L}, \forall\, f\in\mathcal{F} && (8)\\
& y_{e,q} \in \{0,1\} && \forall\, e\in E, \forall\, q\in\mathcal{Q}(e). && (9)
\end{aligned}
$$

The *(extended) configuration model* (QLP) also minimizes the cost of a line plan. The *configuration assignment constraints* (6) ensure that exactly one configuration is chosen for each edge, while the *coupling constraints* (5) guarantee that a sufficient number of lines is operated at the frequencies w.r.t. the chosen configurations.

## *2.1 Model Comparison*

The configuration model (QLP) provides an extended formulation of (SLP), i.e., the convex hulls of all feasible solutions—projected onto the space of the x variables—coincide. The same does not hold for the polytopes defined by the fractional solutions, since (QLP) provides a tighter LP relaxation.

Band inequalities were introduced by Stoer and Dahl [12] and can also be applied to the line planning problem. Given an edge $e\in E$, a *band* $f_{\mathcal{B}}:\mathcal{L}(e)\to\mathcal{F}_0$ assigns a frequency to each line containing $e$ and is called *valid band of e* if $\sum_{\ell\in\mathcal{L}(e)} f_{\mathcal{B}}(\ell) < F(e)$. That is, if all lines on the edge are operated at the frequencies of a valid band, then the frequency demand is not covered and at least one line needs to be operated at a higher frequency. Hence, the *band inequality*

$$
\sum_{\ell\in\mathcal{L}(e)} \sum_{f\in\mathcal{F}: f>f_{\mathcal{B}}(\ell)} x_{\ell,f} \geq 1 \qquad (10)
$$

is a valid inequality for (SLP) for all $e\in E$ and each valid band $f_{\mathcal{B}}$ of $e$. The simplest example is the case $f_{\mathcal{B}}(\ell)\equiv 0$, which states that one must operate at least one line on every edge, i.e., the *set cover* inequality $\sum_{\ell\in\mathcal{L}(e)}\sum_{f\in\mathcal{F}} x_{\ell,f}\geq 1$ is valid for $\mathrm{P_{IP}}(\text{SLP})$ for all $e\in E$. We call the band $f_{\mathcal{B}}$ *symmetric* if $f_{\mathcal{B}}(\ell)=f$ for all $\ell\in\mathcal{L}(e)$ and for some $f\in\mathcal{F}$. Note that set cover inequalities are symmetric band inequalities. We call the valid band $f_{\mathcal{B}}$ *maximal* if there is no valid band $f_{\mathcal{B}'}$ with $f_{\mathcal{B}}(\ell)\leq f_{\mathcal{B}'}(\ell)$ for every line $\ell\in\mathcal{L}(e)$ and $f_{\mathcal{B}}(\ell)<f_{\mathcal{B}'}(\ell)$ for at least one line $\ell\in\mathcal{L}(e)$. Maximal band inequalities often define facets of the single edge relaxation of the line planning polytope, see [9]. The symmetric ones are implied by the configuration model.

**Theorem 1.** *([8]) The LP relaxation of the configuration model implies all band inequalities* (10) *that are induced by a valid symmetric band.*

The demand inequalities (2) can be strengthened by the mixed integer rounding (MIR) technique [6]. Let $e \in E$, $\lambda > 0$ and define $r = \lambda F(e) - \lfloor \lambda F(e) \rfloor$ and $r_f = \lambda f - \lfloor \lambda f \rfloor$. The *MIR inequality*

$$\sum_{\ell \in \mathcal{L}(e)} \sum_{f \in \mathcal{F}} (r \lfloor \lambda f \rfloor + \min(r_f, r)) x_{\ell,f} \geq r \lceil \lambda F(e) \rceil \tag{11}$$

is valid for (SLP). These strengthened inequalities are also implied by the configuration model.

**Theorem 2.** *([8]) The LP relaxation of the configuration model implies all MIR inequalities* (11)*.*

The configuration model is strong in the sense that it implies several facet-defining inequalities for the standard model. However, the enormous number of configurations can blow up the formulation for large instances. Hence, we propose a mixed model that enriches the standard model by a judiciously chosen subset of configurations; only for edges with a small number of minimal configurations the corresponding variables and constraints are added. This provides a good compromise between model strength and model size. We present computational results for large-scale line planning problems in [2] and [8] that confirm the theoretical findings for the naive configuration model and show the superiority of the proposed mixed model. Our approach shows its strength in particular on real world instances.

## 3 Multi-Edge Configuration

In this section we generalize the concept of minimal configurations with the goal to tighten the LP relaxation further. The idea is to define configurations for a subset of edges. For this purpose we partition the lines according to the edges they pass in $\tilde{E}$, $\tilde{E} \subseteq E$. Let $E' \subseteq \tilde{E}$, then we denote by $\mathcal{L}(E')|_{\tilde{E}} := \{\ell \in \mathcal{L} : \ell \cap \tilde{E} = E'\}$ the set of lines such that $E'$ corresponds to the edges they pass in $\tilde{E}$ and define $\mathcal{E}(\tilde{E}) := \{E' \subseteq \tilde{E} : \mathcal{L}(E')|_{\tilde{E}} \neq \emptyset\}$. A *multi-edge configuration* specifies for each line set in this partition how many of them are operated at a certain frequency. The formal definition reads as follows:

**Definition 2.** For $\tilde{E} \subseteq E$, let $\bar{\mathcal{Q}}(\tilde{E}) \subseteq \mathbb{Z}_{\geq 0}^{\mathcal{E}(\tilde{E}) \times \mathcal{F}}$ be the set of *(feasible) multi-edge configuration of* $\tilde{E}$ with $Q \in \bar{\mathcal{Q}}(\tilde{E})$ if and only if

$$\sum_{f \in \mathcal{F}} Q_{E',f} \leq \left| \mathcal{L}(E')|_{\tilde{E}} \right| \qquad \forall E' \in \mathcal{E}(\tilde{E}), \tag{12}$$

$$\sum_{\substack{E' \in \mathcal{E}(\tilde{E}): \\ e \in E'}} \sum_{f \in \mathcal{F}} f \cdot Q_{E',f} \geq F(e) \qquad \forall e \in \tilde{E}. \tag{13}$$

We call a multi-edge configuration $Q \in \bar{\mathscr{Q}}(\tilde{E})$ *minimal* if there is no $\bar{Q} \in \bar{\mathscr{Q}}(\tilde{E})$ such that $\bar{Q}_{E'',f} \leq Q_{E'',f}$ for all $E'' \in \mathcal{E}(\tilde{E})$, $f \in \mathcal{F}$ and $\bar{Q}_{E'',f} < Q_{E'',f}$ for some $E'' \in \mathcal{E}(\tilde{E})$, $f \in \mathcal{F}$. The set of *minimal multi-edge configurations of* $\tilde{E}$ is denoted by $\mathscr{Q}(\tilde{E})$.

Let $\mathscr{E}$ be a cover of $E$, i.e., $\mathscr{E} \subseteq 2^E$ such that $\bigcup_{E' \in \mathscr{E}} E' = E$. We extend the standard model (SLP) with binary variables $y_{E',Q}$ indicating for each subset of edges $E' \in \mathscr{E}$ which minimal multi-edge configuration $Q \in \mathscr{Q}(\tilde{E})$ is chosen. The *multi-edge configuration model induced by the edge cover* $\mathscr{E}$ is defined as follows:

$$(\mathscr{E}\text{-QLP}) \quad \min \quad \sum_{\ell \in \mathcal{L}} \sum_{f \in \mathcal{F}} c_{\ell,f} x_{\ell,f}$$

$$\sum_{\ell \in \mathcal{L}(E')|_{\tilde{E}}} x_{\ell,f} \geq \sum_{Q \in \mathscr{Q}(\tilde{E})(\tilde{E})} Q_{E',f} \cdot y_{\tilde{E},Q} \qquad \forall \tilde{E} \in \mathscr{E}, \forall E' \in \mathcal{E}(\tilde{E}), \forall f \in \mathcal{F} \tag{14}$$

$$\sum_{Q \in \mathscr{Q}(\tilde{E})} y_{\tilde{E},Q} = 1 \qquad \forall \tilde{E} \in \mathscr{E} \tag{15}$$

$$\sum_{f \in \mathcal{F}} x_{\ell,f} \leq 1 \qquad \forall \ell \in \mathcal{L} \tag{16}$$

$$x_{\ell,f} \in \{0,1\} \qquad \forall \ell \in \mathcal{L}, \forall f \in \mathcal{F} \tag{17}$$

$$y_{\tilde{E},Q} \in \{0,1\} \qquad \forall \tilde{E} \in \mathscr{E}, \forall Q \in \mathscr{Q}(\tilde{E}). \tag{18}$$

The multi-edge configuration model ($\mathscr{E}$-QLP) minimizes the cost of a line plan. Since for each edge $e \in E$ in ($\mathscr{E}$-QLP) a minimal multi-edge configuration is chosen for the subset $\tilde{E} \in \mathscr{E}$ containing $e$, the frequency demand of $e$ is satisfied by every feasible solution of ($\mathscr{E}$-QLP). Model ($\mathscr{E}$-QLP) also provides an extended formulation for (SLP).

Let $\tilde{E} \subseteq E$ such that $\alpha(\tilde{E}) := \max\{|\ell \cap \tilde{E}| : \ell \in \mathcal{L}\} > 0$ and $F(E') = \sum_{e \in \tilde{E}} F(e)$ , then the aggregated frequency inequality (Bussieck [4])

$$\sum_{\ell \in \mathcal{L}(E')} \sum_{f \in \mathcal{F}} f x_{\ell,f} \geq \left\lceil \frac{1}{\alpha(\tilde{E})} F(E') \right\rceil \tag{19}$$

and the aggregated cardinality inequality

$$\sum_{\ell \in \mathcal{L}(E')} \sum_{f \in \mathcal{F}} x_{\ell,f} \geq \left\lceil \frac{1}{\alpha(\tilde{E})} |\tilde{E}| \right\rceil \tag{20}$$

are valid for (SLP). These inequalities are in general not valid for the LP relaxation of the standard model (SLP) and the configuration model (QLP). However, regarding the multi-edge configuration model ($\mathscr{E}$-QLP) we obtain the following result.

**Proposition 1.** *([8]) Let $\mathscr{E}$ be a cover of $E$ and $\tilde{E} \in \mathscr{E}$ s.t. $\alpha(\tilde{E}) > 0$. Then the aggregated frequency inequality* (19) *and the aggregated cardinality inequality* (20) *for $\tilde{E}$ are implied by the LP relaxation of the multi-edge configuration model* ($\mathscr{E}$-QLP).

## References


1. Ralf Borndörfer, Martin Grötschel, and Marc E. Pfetsch. A column-generation approach to line planning in public transport. *Transportation Science*, 41(1):123–132, 2007.
2. Ralf Borndörfer, Heide Hoppmann, and Marika Karbstein. A configuration model for the line planning problem. In Daniele Frigioni and Sebastian Stiller, editors, *ATMOS 2013 - 13th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems*, volume 33, pages 68 – 79, 2013.
3. Ralf Borndörfer and Marika Karbstein. A direct connection approach to integrated line planning and passenger routing. In Daniel Delling and Leo Liberti, editors, *ATMOS 2012 - 12th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*, volume 25, pages 47 – 57, 2012.
4. Michael Bussieck. *Optimal Lines in Public Rail Transport*. PhD thesis, Technische Universität Braunschweig, 1998.
5. Michael R. Bussieck, Peter Kreuzer, and Uwe T. Zimmermann. Optimal lines for railway systems. *Eur. J. Oper. Res.*, 96(1):54–63, 1997.
6. Sanjeeb Dash, Oktay Günlük, and Andrea Lodi. MIR closures of polyhedral sets. *Mathematical Programming*, 121(1):33–60, 2010.
7. Jan-Willem Goossens, Stan van Hoesel, and Leo Kroon. A branch-and-cut approach for solving railway line-planning problems. *Transportation Science*, 28(3):379–393, 2004.
8. Heide Hoppmann. A configuration model for the line planning problem. Master's thesis, Technische Universität Berlin, 2014.
9. Marika Karbstein. *Line Planning and Connectivity*. PhD thesis, TU Berlin, 2013.
10. Anita Schöbel. Line planning in public transportation: models and methods. *OR Spectrum*, pages 1–20, 2011.
11. Anita Schöbel and Susanne Scholl. Line planning with minimal traveling time. In Leo G. Kroon and Rolf H. Möhring, editors, *Proceedings of 5th Workshop on Algorithmic Methods and Models for Optimization of Railways*, 2006.
12. Mechthild Stoer and Geir Dahl. A polyhedral approach to multicommodity survivable network design, 1994.