Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

MARIKA KARBSTEIN

Integrated Line Planning and Passenger Routing: Connectivity and Transfers

Supported by the DFG Research Center MATHEON “Mathematics for key technologies”, project B15.

ZIB Report 14-42 (October 2014)

Herausgegeben vom
Konrad-Zuse-Zentrum für Informationstechnik 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

Integrated Line Planning and Passenger Routing: Connectivity and Transfers[*]

Marika Karbstein[**]

1 Introduction

The *integrated line planning and passenger routing problem* is an important planning problem in service design of public transport.

The infrastructure of the public transport system is represented by a graph where the edges correspond to streets and tracks and the nodes correspond to stations/stops. We are further given point-to-point demands, i. e., the number of passengers that want to travel from one point in the network to another point. A line is a path in the network, visiting a set of stops/stations in a predefined order. Passengers can travel along these lines and they can change from one line to another line in a stop/station if these lines intersect. Bringing capacities into play, the task is to find paths in the infrastructure network for lines and passengers such that the capacities of the lines suffice to transport all passengers. There are two main objectives for a line plan, namely, minimization of line operation costs and minimization of passenger discomfort measured in, e. g., travel times and number of transfers.

In general, the computed line system should be connected, i. e., one can travel from one station to any other station along the lines. Associating cost with the lines and searching for a cost minimum set of lines such that all stations are connected, gives rise to a combinatorial optimization problem which covers the connectivity aspect of integrated line planning and passenger routing. We denote this problem by the *Steiner connectivity problem*. The solution of a Steiner connectivity problem gives a lower bound on the costs of a line plan.

In this paper we introduce some results for the Steiner connectivity problem and show that they can be used to handle the transfer aspect for the line planning problem. In Section 2, the Steiner connectivity problem is introduced in more detail. Here, we focus on the special case to connect two nodes (terminals) via a set of lines or paths. In Section 3, we propose a new model for the integrated line planning and passenger routing problem that handles transfers in an approximative way and involves a type of the 2-terminal Steiner connectivity problem as pricing problem. We briefly discuss computational results in Section 4 and the optimized line plan for ViP Potsdam of the year 2010 in Section 5.

[*]Supported by the DFG Research Center MATHEON "Mathematics for key technologies"
[**]Zuse Institute Berlin, Takustr. 7, 14195 Berlin, Germany; Email karbstein@zib.de

Figure 1: Example of a Steiner connectivity problem. *Left:* A graph with four terminal nodes ($T = \{a, d, e, f\}$) and six paths $\big(\mathcal{P} = \{p_1 = (ab, bc, cd), p_2 = (ef, fg), p_3 = (ae), p_4 = (ef, fc), p_5 = (gd), p_6 = (fg, gc, cd)\}\big)$. *Right:* A feasible solution with three paths ($\mathcal{P}' = \{p_3, p_4, p_6\}$).

2 Steiner Connectivity Problem

The Steiner connectivity problem is a generalization of the well-known *Steiner tree problem*. Given a graph with costs on the edges, the Steiner tree problem is to find a cost minimum set of edges that connects a subset of nodes. The Steiner connectivity problem is to choose a set of paths instead of edges. Steiner trees are fundamental for network design in transportation and telecommunication; see Dell'Amico, Maffioli, and Martello [1] for an overview. In fact, the Steiner tree problem can be seen as the prototype of all problems where nodes are connected by installing capacities on individual edges or arcs. In the same way, the Steiner connectivity problem can be seen as the prototype of all problems where nodes are connected by installing capacities on *paths* which is exactly the case in line planning. Hence, the significance of the Steiner connectivity problem for line planning is similar to the significance of the Steiner tree problem for telecommunication network design.

A formal description of the Steiner connectivity problem (SCP) is as follows. We are given an undirected graph $G = (V, E)$, a set of *terminal nodes* $T \subseteq V$, and a set of elementary *paths* $\mathcal{P}$ in G. The paths have nonnegative costs $c \in \mathbb{R}^{\mathcal{P}}_{\geq 0}$. The problem is to find a subset of paths $\mathcal{P}' \subseteq \mathcal{P}$ of minimal cost $\sum_{p \in \mathcal{P}'} c_p$ that *connect the terminals*, i. e., such that for each pair of distinct terminal nodes $t_1, t_2 \in T$ there exists a path q from t_1 to t_2 in G such that each edge of q is covered by at least one path of $\mathcal{P}'$. We can assume w.l.o.g. that every edge is covered by a path, i. e., for every $e \in E$ there is a $p \in \mathcal{P}$ such that $e \in p$; in particular, G has no loops. Figure 1 gives an example of a Steiner connectivity problem and a feasible solution.

Main results about complexity, approximation, integer programming formulations, and polyhedra can be generalized from the Steiner tree problem to the Steiner connectivity problem, see [5, 7]. In the following we want to consider the two-terminal case of the Steiner connectivity problem, see also [3]. The problem is to find a minimum set of paths connecting two given nodes s and t.

We call a set of paths $\mathcal{P}' \subseteq \mathcal{P}$ an *st-connecting set* if s and t are connected in the subgraph $H = (V, E(\mathcal{P}'))$, i. e., $\mathcal{P}'$ is a solution for the Steiner connectivity problem with $T = \{s, t\}$. A set $\mathcal{P}' \subseteq \mathcal{P}$ is *st-disconnecting* if $\mathcal{P} \setminus \mathcal{P}'$ is not an *st*-connecting set.

Algorithm 1: Primal-dual minimum *st*-connecting set algorithm.

Input : A connected graph $G = (V, E)$, a set of paths $\mathcal{P}$ with costs $c \in \mathbb{R}^{\mathcal{P}}_{\geq 0}$ that covers all edges E, $s, t \in V$.

Output: The value of a minimum cost *st*-connecting set.

$d(s) := 0$, $d(v) := \infty \; \forall v \in V \setminus \{s\}$
all nodes are unmarked
while t is unmarked **do**
    Choose v with $v = \text{argmin}\ \{d(u) : u \text{ unmarked}\}$
    for all $p \in \mathcal{P}$ with $v \in p$ **do**
        for unmarked w with $w \in p$ **do**
            if $d(w) > d(v) + c_p$ **then**
                $d(w) := d(v) + c_p$
            end
        end
    end
    mark v
end
return $d(t)$

Algorithm 1 computes the cost of a minimum *st*-connecting set. It generalizes Dijkstra's algorithm to our setting. The distances from node s are stored in node labels $d(v)$. The algorithm can be extended such that it also determines the minimum *st*-connecting set $\mathcal{P}'$.

A cut formulation for the SCP with $T = \{s, t\}$ is:

$$
\begin{array}{llll}
\text{(MCS)} & \min & \displaystyle\sum_{p \in \mathcal{P}} c_p x_p & \\
\text{(i)} & \text{s.t.} & \displaystyle\sum_{p \in \mathcal{P}_{\delta(W)}} x_p \geq 1 & \forall\, s \in W \subseteq V \setminus \{t\} \\
 & & x_p \in \{0, 1\} & \forall\, p \in \mathcal{P}.
\end{array}
$$

Here, x_p is a $0/1$-variable that indicates whether path p is chosen ($x_p = 1$) or not ($x_p = 0$). Furthermore, $\mathcal{P}_{\delta(W)} := \{p \in \mathcal{P} : \delta(W) \cap p \neq \emptyset\}$ is the set of all paths that cross the cut $\delta(W) = \{\{u, v\} \in E : |\{u, v\} \cap W| = 1\}$ at least one time.

Theorem 1. *The inequality system of (MCS) is TDI.*

This can be shown by extending Algorithm 1 to a primal-dual algorithm that defines integer solutions for (MCS) and its dual program.

Setting $c \equiv 1$ in Algorithm 1 and interpreting the set of paths $\mathcal{P}$ as lines and s and t as origin and destination stations, then the algorithm computes the minimum number of lines that are necessary to connect s and t. This number corresponds to the minimum number of transfers minus 1 that are necessary to travel from s to t. The calculation of the minimum number of transfers is the basic idea of the model introduced in the next section.

3 A Transfer Model for Line Planning

In this section we want to propose a model for line planning and passenger routing that account for the number of *unavoidable transfers*. Each passenger path is associated with its number of minimum transfers with respect to the given set of all possible lines. More precisely, considering a certain passenger path, it may not be possible to cover this path by a single line or even by two lines, i. e., in any definition of a line plan, passengers on the path under consideration have to transfer at least once or twice, respectively. We call such transfers *unavoidable*.

We use the following notation. Consider a public transportation network as a graph $N = (V,E)$, whose nodes and edges correspond to stations and connections between these stations, respectively. Denote by $\mathscr{L}$ the line pool, i. e., a set of paths in N that represent all valid lines and by $\mathscr{F} \subseteq \mathbb{N}$ the set of possible frequencies at which these lines can be operated. If line ℓ is operated with frequency f, $\kappa_{\ell,f} \in \mathbb{Q}_+$ denotes the capacity and $c_{\ell,f} \in \mathbb{Q}_+$ the operation cost of this line. Let further $(d_{st}) \in \mathbb{Q}_+^{V\times V}$ be an origin-destination (OD) matrix that gives the travel demand between pairs of nodes, and denote by $D = \{(s,t) \in V \times V : d_{st} > 0\}$ the set of all OD-pairs with positive demand. Derive a directed passenger routing graph $\bar{N} = (V,A)$ from N by replacing each edge $e \in E$ with two antiparallel arcs $a(e)$ and $\bar{a}(e)$. Denote by $\mathcal{P}_{(s,t)}$ the set of all possible directed (s,t)-paths in $\bar{N}$ for $(s,t) \in D$, and by $\mathscr{P} = \bigcup_{(s,t)\in D} \mathscr{P}_{(s,t)}$ the set of all such paths; these represent travel routes of passengers. Associated with each arc $a \in A$ and path $p \in \mathscr{P}$ are travel times $\tau_a \in \mathbb{Q}_+$ and $\tau_p = \sum_{a\in p} \tau_a$, respectively, and with each transfer a (uniform) penalty $\sigma \in \mathbb{Q}_+$. Let k_p be the minimum number of transfers that passengers must do on path p if all lines in $\mathscr{L}$ would be built. A path $p \in \mathscr{P}$ with k_p unavoidable transfers has travel and transfer time $\tau_{p,k} = \tau_p + k_p\sigma$. Let $e(a)$ be the undirected edge corresponding to $a \in A$, and let us interpret a(n undirected) line in N in such a way that passengers can travel on this line in both directions in $\bar{N}$. The unavoidable transfer model is then

$$\text{(UT)} \quad \min \lambda \sum_{\ell\in\mathscr{L}} \sum_{f\in\mathscr{F}} c_{\ell,f}\, x_{\ell,f} + (1-\lambda)\Big(\sum_{p\in\mathscr{P}} \tau_{p,k_p}\, y_{p,k_p}\Big)$$

$$\sum_{p\in\mathscr{P}_{st}} y_{p,k_p} = d_{st} \qquad \forall (s,t) \in D \tag{1}$$

$$\sum_{p\in\mathscr{P}:a\in p} y_{p,k_p} \le \sum_{\ell\in\mathscr{L}:e(a)\in\ell} \sum_{f\in\mathscr{F}} \kappa_{\ell,f}\, x_{\ell,f} \qquad \forall a \in A \tag{2}$$

$$\sum_{f\in\mathscr{F}} x_{\ell,f} \le 1 \qquad \forall \ell \in \mathscr{L} \tag{3}$$

$$x_{\ell,f} \in \{0,1\} \qquad \forall \ell \in \mathscr{L}, \forall f \in \mathscr{F} \tag{4}$$

$$y_{p,k_p} \ge 0 \qquad \forall p \in \mathscr{P}. \tag{5}$$

Model (UT) minimizes a weighted sum of line operating costs and passenger travel times. We use binary variables $x_{\ell,f}$ for the operation of line $\ell \in \mathscr{L}$ at frequency $f \in \mathscr{F}$. The continuous variables y_{p,k_p} account for the number of passengers that travel on path $p \in \mathscr{P}$ doing *at least* k_p transfers. Equations (1) enforce the passenger flow. Inequalities (2) guarantee sufficient total transportation capacity on each arc. Inequalities (3) ensure that a line is operated at one frequency at most.

Algorithm 1 can be extended such that it computes a travel-time minimal path from a given node $s \in V$ to all other nodes including a uniform transfer penalty $\sigma \in \mathbb{Q}_+$ for each transfer w. r. t. a given set of lines $\mathscr{L}$. More precisely, replacing c_p by the travel time on line ℓ from v to w in lines 7 and 8 of the algorithm and adding a σ for $v \neq s$ in the same lines, yields the following proposition.

Proposition 2. *The pricing problem for the passenger path variables in model (UT) can be solved in polynomial time.*

The number k_p accounts for the minimum number of transfers w. r. t. all lines $\mathscr{L}$. In a final line plan usually only a small subset of lines $\mathscr{L}' \subseteq \mathscr{L}$ is established, i. e., the number of necessary transfers on a path p might be much larger. Since offering direct connections is a major goal in line planning, we included constraints to ensure enough capacities for passenger paths considered as direct connections. Let $\mathscr{L}_{st}$ be the number of lines supporting a direct connection from s to t, $\mathscr{L}_{st}(a) = \{\ell \in \mathscr{L}_{st} : a \in \ell\}$ be the direct connection lines for (s,t) containing arc a, and $\mathscr{P}^0_{st} = \{p \in \mathscr{P}_{st} : k_p = 0\}$ be the set of all passenger paths from s to t with 0 unavoidable transfers. Then we can define *direct connection constraints* for each arc and each OD pair

$$\sum_{(u,v)\in D} \sum_{p\in\mathscr{P}^0_{uv}: a\in p,\, \mathscr{L}_{uv}(a)\subseteq\mathscr{L}_{st}(a)} y_{p,0} \leq \sum_{\ell\in\mathscr{L}_{st}(a)} \sum_{f\in\mathscr{F}} \kappa_{\ell,f}\, x_{\ell,f} \quad \forall a \in A, (s,t) \in D. \tag{6}$$

These constraints are a combinatorial subset of the so-called *dcmetric inequalities* [4] that enforce sufficient transportation capacity to route all st-paths with 0 transfers via arc a. For each path $p \in \mathscr{P}^0 = \cup_{(s,t)\in D} \mathscr{P}_{st}$ we then have an additional variable $y_{p,1}$ which come into play if the associated direct connection constraints for $y_{p,0}$ are not satisfied. Then the path can still be chosen in the optimization model but it is associated with at least one transfer and incurs one transfer penalty.

4 Computational Results

We made computations for several instances, e. g., a SiouxFalls instance from the Transportation Network Test Problems Library of Bar-Gera, a Dutch instance for the train network introduced by Bussieck in the context of line planning [6], an artificial China instance based on the 2009 high speed train network and some real world instances provided by our cooperation partner Verkehr in Potsdam GmbH.

For the SiouxFalls, Dutch, and China instances it turned out that it already suffice to distinguish passenger paths on direct connections and passenger path with one transfer and to consider the direct connection constraints. Indeed, evaluating the computed line plans shows that each passenger path of these instances is either a direct connection path or involves exactly one transfer. Since the Potsdam instances are real multi-modal public transportation networks, there exist several passenger path containing two or more transfers. However, modeling transfers between different transportation modes via transfer arcs (including a transfer penalty) and distinguishing direct connection paths from paths with at least one transfer for paths of one transportation mode via the direct connection constraints (6) yields a tractable model also for the Potsdam instances that estimates the travel times and transfers quite accurately.

5 Line Plan for Potsdam 2010

A reorganization of the line plan in Potsdam became necessary when the public transport company of Potsdam, ViP Verkehrsbetriebe Potsdam GmbH, took over six additional bus lines that were formerly operated by Havelbus Verkehrsgesellschaft mbH. The new line plan should minimize the travel time at a same cost level, and ViP emphasized the importance of a minimal number of transfers.

Our mathematically optimized solution for the Potsdam line plan 2010 minimizes the total number of transfers by around 5% in comparison to a "hand made" plan on the basis of experience, see [2]. It further reduces the cost by around 4% and the perceived travel time by around 6%. ViP also certified that this line plan was indeed practicable and established a slightly modified version of our optimized solution.

References

[1] Anantaram Balakrishnan, Thomas L. Mangnanti, and Prakash Mirchandani. Network design. In Mauro Dell'Amico, Francesco Maffioli, and Silvano Martello, editors, *Annotated Bibliographies in Combinatorial Optimization*, chapter 18, pages 311–334. Wiley, Chichester, 1997.

[2] Ralf Borndörfer, Isabel Friedow, and Marika Karbstein. Optimierung des Linienplans 2010 in Potsdam. *Der Nahverkehr*, 30(4):34–39, 2012.

[3] Ralf Borndörfer, Nam Dũng Hoàng, Marika Karbstein, Thorsten Koch, and Alexander Martin. How many Steiner terminals can you connect in 20 years? In Michael Jünger and Gerhard Reinelt, editors, *Facets of Combinatorial Optimization; Festschrift for Martin Grötschel*, pages 215–244. Springer, 2013.

[4] Ralf Borndörfer and Marika Karbstein. Metric inequalities for routings on direct connections. ZIB-Report 14-04, Zuse Institute Berlin, 2014.

[5] Ralf Borndörfer, Marika Karbstein, and Marc E. Pfetsch. The Steiner connectivity problem. *Mathematical Programming*, 142(1):133–167, 2013.

[6] M. R. Bussieck. *Optimal lines in public rail transport*. PhD thesis, TU Braunschweig, 1997.

[7] Marika Karbstein. *Line planning and connectivity*. PhD thesis, TU Berlin, 2013.