Chapter 1

Planning in Public Transit

In this chapter we give an overview of the planning process in public transit. We will describe the planning problems associated with the operation of a network of bus lines and present state of the art mathematical models and algorithms to solve them. We will show the general trend that the used planning scenarios become larger either by looking at larger planning units or by looking at more planning steps at once. Both approaches lead to potential better solutions by more degrees of freedom in the planning problems.

Some of the problems occurring in the planning process of bus traffic are similar to planning problems in other modes of public transportation, such as subways, trams, trains, or even airlines. We will comment on the details of the similarities and differences in the respective sections. Bussieck et al. [1997] describe the use of discrete optimization in the planning process of public rail transports. We concentrate here on recent models and algorithms. An overview about planning in public transport until 1994 can be found in the article of Odoni et al. [1994]. Borndörfer et al. [2006] highlight the increasing use of OR methods for planning problems in public transport and describe exemplarily some applications.

We use in this thesis the definitions and notations of Schrijver [2003] for graphs, linear programs (LPs), and mixed integer programs (MIPs). A short overview of the used symbols can also be found in the annex.

1.1 Introduction

In the last years, the budgets of the federal government, the states, and the municipalities in Germany were very tight. Therefore the federal government has cut the so called “Regionalisierungsmittel”, that is, subsidies for regional public transport in Germany, from €7.1 billion to €6.6 per year[1]. Also other subsidies were or will be reduced: Berlin has cut its subsidies to its public transport company BVG, by €100 millions from 2005 to 2006[2] and further reductions will likely follow. A study of Resch et al. [2006] reports reductions of subsidies of 32–44% of three German public transit companies[3]. Also more and more tenderings for the subsidies for public transit are put out instead of giving them directly to the local companies. Thus, public mass transit companies in Germany are under the pressure to reduce their costs. This can be accomplished by either discontinuing unprofitable lines, by lowering the wages of the personnel, or by increasing the efficiency of the schedules. All of these measures have been taken in the past. In the following we will examine where new mathematical methods may help to further improve the efficiency without disadvantages for the drivers or the passengers. We assume that even if computer based planning systems are already in use, new mathematical approaches are able to solve larger planning problems at once and help to integrate subsequent planning steps.

Besides the financial goals of the planning process also the general benefit of public transport, summarized as public welfare, is of concern. The following often opposing objectives are common for strategic and operational planning in public transport:

- increasing the attractiveness of public transport,
- reducing operation cost for a given service level,
- increasing the transport capacity for a given budget,
- reducing medium term capital investment (e.g., by reducing the number of buses, number of stops, or number of depots).

The improvement of the results of the planning process with respect to these objectives is not the only advantage of using optimization methods. The possibility of calculating alternative scenarios in short time is also of great interest for public transport companies because it backs up the decision process with reliable information.

[1] Haushaltsbegleitgesetz of Berlin, 2006
[2] business report BVG 2005
[3] The remaining subsidies are still about 9–21% of the total costs of the companies.

1.2 Classification of Planning Steps

Usually the planning process in public transport is divided into *strategic* and *operational* planning. Sometimes also tactical planning is mentioned as an intermediate step. Strategic planning consists of problems dealing with long term decisions about the infrastructure and the level of service, whereas operational planning handles the problems which occur in the operation of a given service.

We will concentrate in the sections below on the following planning steps:

1. Strategic planning:
   - network design,
   - line network planning,
   - time table planning.
2. Operational planning:
   - vehicle scheduling,
   - duty scheduling,
   - crew rostering.

We remark that the collection of passenger data as an input to strategic planning problems is a different problem on its own, which we do not discuss here. Often mentioned in this context are operational problems, such as the dispatching of vehicles or the recovery of planned schedules after delays or breakdowns. These problems need specialized algorithms due to their real time requirements, which are beyond the scope of this work. A recent overview of literature about the treatment of delays in vehicle and duty scheduling and an algorithm to deal with it can be found in Huisman [2004].

Another important planning step, which is in general not conducted by the public transit companies but by local authorities, is the planning of public tenderings of public transit for certain lines or regions. We cover this topic because it can use methods of the other planning steps.

1.3 Basic Models

We introduce two basic mathematical planning models whose variants are often used in public transport planning problems or other scheduling problems. We show the connections between those variants.

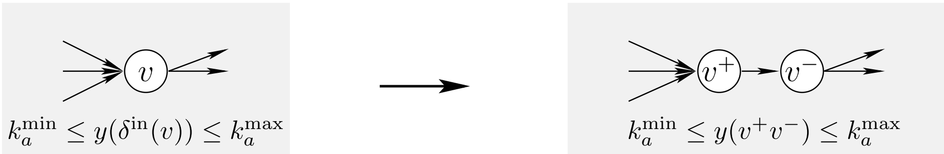

Figure 1.1: Modeling capacities on nodes

1.3.1 Flow Based Models

Many problems related to scheduling or transportation problems can be modeled as flow problems in the following general sense: Given is a directed network with a set of nodes V and a set of arcs A connecting the nodes, a source $s \in V$, and a sink $t \in V$. Additionally, we have a minimum capacity $k_a^{\min} \geq 0$ and a maximum capacity $k_a^{\max} \geq k_a^{\min}$ as well as a cost c_a per unit of flow over each arc $a \in A$. The goal is to find a minimum cost flow from s to t subject to the capacity constraints and eventually certain other constraints. This task can be formulated as the following *generalized minimum cost flow problem* (GMCF):

$$
\begin{array}{lll}
\text{(GMCF)} & \min c^{\mathsf{T}} y, & \\
& \text{s.t.} & \\
\text{(i)} & y(\delta^{\text{in}}(v)) - y(\delta^{\text{out}}(v)) = 0, & \forall v \in V \setminus \{s,t\}, \\
\text{(ii)} & k_a^{\min} \leq y_a \leq k_a^{\max}, & \forall a \in A, \\
\text{(iii)} & By = b \text{ or } By \leq b, & \\
\text{(iv)} & y_a \geq 0, & \forall a \in A, \\
\text{(v)} & y \in \mathbb{Z}^A. &
\end{array}
$$

Here, y_a is the flow variable of arc a giving the number of units of the flow over arc a. Because the decision variables are associated to arcs, we say (GMCF) is an *arc-based model*.

The flow conservation constraints (i) ensure that at every node the in-flow is equal to the out-flow. The capacity constraints (ii) guarantee that the arc capacities are satisfied. These constraints can also be used to enforce minimum or maximum flows over a node as follows: Replace a node by two new adjacent nodes and a new arc connecting these nodes. The lower capacity on the new arc is the desired minimum flow (see Figure 1.1). Additional properties of the flow can be modeled by constraints (iii). Here either equalities or inequalities or a mixture is possible, $B \in \mathbb{R}^{\mathcal{R} \times A}$ and $b \in \mathbb{R}^{\mathcal{R}}$

are an appropriate matrix and vector, and $\mathcal{R}$ is an index set. Inequalities (iv) ensure the non-negativity of the flow. Constraints (v) model integrality. This is required, e.g., to solve the problem of assigning integral resources like vehicles or drivers to certain activities.

The minimum cost flow problem (MCF) without constraints (iii) is solvable in polynomial time by special network algorithms (see Ahuja et al. [1993]) or by a specialized version of the simplex algorithm (see, e.g., Löbel [1996]).

The multi-commodity flow problem is a specialization of (GMCF), for which constraints (iii) of (GMCF) take the shape

$$\text{(iiia)} \quad y(\delta_g^{\text{in}}(v)) - y(\delta_g^{\text{out}}(v)) = 0 \quad \forall (v,g) \in V \setminus \{s,t\} \times G.$$

Here G is the set of *commodities*, $\delta_g^{\text{in}}(v) := \delta^{\text{in}}(v) \cap A_g$, and $\delta_g^{\text{out}}(v) := \delta^{\text{out}}(v) \cap A_g$, whilst $A_g, g \in G$ are disjoint subsets of A, and $\bigcup_{g \in G} A_g = A$. The equations (iiia) partition the network into $|G|$ subnetworks which each give rise to independent flow conservation constraints. Thus, equations (iiia) render equations (i) redundant. However, equations (i) are still useful when relaxing equations (iiia).

Multi-commodity flow problems occur in the planning of telecommunication networks, where each traffic gives rise to a single commodity, or as a subproblem of the network design problem (see next section). Also the multi-depot vehicle scheduling problem (see section 1.9) can be modeled as a multi-commodity flow problem.

Another specialization of (GMCF) is the minimum cost flow problem with resource constraints. For this problem we replace constraints (iii) by

$$\text{(iiib)} \quad Ry \leq \ell_r \quad \forall r \in \mathcal{R}.$$

Here $R \in \mathbb{R}^{\mathcal{R} \times A}$ is a matrix whose entry in the r-th row and a-th column gives the *resource consumption* of a resource $r \in \mathcal{R}$ by arc a. The vector $\ell \in \mathbb{R}^{\mathcal{R}}$ gives the maximal allowed consumption of the resources. A special case of (GMCF) with constraints (iiib) is the resource constraint shortest path problem. It occurs, e.g., in the duty generation subalgorithm of our duty scheduling algorithm (see Section 5.5.1).

Another practically relevant special case of (iiib) are the following "infeasible path constraints" (iiic), which are used to remove a set of forbidden paths $\mathcal{P}$ in feasible flows:

$$\text{(iiic)} \quad \sum_{a \in P} y_a \leq |P| - 1, \quad \forall P \in \mathcal{P}.$$

This type of constraints can be used to model complicated constraints on subflows, as they occur for example in duty scheduling problems or in vehicle scheduling problems with maintenance requirements, which are described below.

The model (GMCF) with constraints (iiic) is difficult to solve in practice, if the set $\mathcal{P}$ of infeasible paths is large in comparison to the set of all possible paths. This occurs, for example, in the duty scheduling problem since here most paths in the graph are not representing a valid duty (see Schlechte [2003]).

1.3.2 Path Based Models

To overcome the difficulties to solve (GMCF) with many constraints of type (iiic) the flow model can be transformed by Dantzig-Wolfe decomposition (Dantzig & Wolfe [1960]) into a path based model. The idea of this transformation is that each st-flow can be decomposed into a sum of st-paths and cycles. Thus, the problem (GMCF) can be reformulated as a problem of finding a (cost minimal) set of paths and cycles such that the resulting flow fulfills all capacity constraints. We call the result of this transformation *generalized path covering problem* (GPCP).

$$
\begin{array}{lll}
\text{(GPCP)} & \min d^{\mathsf{T}} x, & \\
 & \text{s. t.} & \\
\text{(i)} & k_a^{\min} \le \sum\limits_{P \ni a} x_P \le k_a^{\max}, & \forall a \in A, \\
\text{(ii)} & x_P \ge 0, & \forall P \in \mathcal{S}, \\
\text{(iii)} & x \in \mathbb{Z}^{\mathcal{S}}. &
\end{array}
$$

Here $\mathcal{S} \subset \mathcal{P}(A)$ is the set of feasible arc sets. In our applications of this model these sets form st-paths. The variables $x_P, P \in \mathcal{S}$ indicate how much flow is routed over path P. Inequalities (i) guarantee, like the constraints (GMCF)(ii), the compliance with the minimum and maximum capacities $k_a^{\min}$ and $k_a^{\max}$ on every arc $a \in A$. The notation $P \ni a$ denotes all arc sets $P \in \mathcal{S}$ that include arc a. Constraints (ii) and (iii) ensure non-negativity and integrality of the variable x_P, $P \in \mathcal{S}$. The cost $d_P, P \in \mathbb{R}^{\mathcal{S}}$, denotes the cost of a certain arc set P. Often $d_P = \sum_{a \in P} c_a$, that is, the cost of an arc set P is the sum of the cost of its arcs, or $d_P = 1$ for all $P \in \mathcal{S}$ if only the number of arc sets should be minimized.

In general, the set $\mathcal{S}$ of feasible arc sets is very large. Nevertheless models of this kind can be solved by column generation approaches (see Barnhart et al. [1998] or Section 3.1). Specializations of this problem are the path packing, path partitioning, as well as the path covering problem. In these problems the arc sets form st-paths. In the path covering problem all minimum capacities are one and the maximum capacities are infinite, i.e. $k_a^{\min} = 1$ and $k_a^{\max} = \infty$ for all $a \in A$, and in the path packing problem the maximum capacities are one ($k_a^{\max} = 1$) and the minimum capacities are zero ($k_a^{\min} = 0$) for each $a \in A$. The path partitioning problem finally has $k^{\min} = \mathbb{1} = k^{\max}$.

A slightly more generalized form of these problems are the well known set covering (SCP), set packing (SSP), and set partitioning (SPP) problems. In these problems the paths in a graph are replaced by subsets P of a basic set A. These problems can be formulated as follows:

$$
\begin{array}{lll}
(\mathsf{SCP}) & \min d^{\mathsf{T}} x, & \\
 & \text{s. t.} & \\
(\mathrm{i}) & \displaystyle\sum_{P \ni a} x_P \geq 1 & \forall a \in A, \\
(\mathrm{iii}) & x \in \{0,1\}^{\mathcal{S}}, &
\end{array}
$$

$$
\begin{array}{lll}
(\mathsf{SSP}) & \min d^{\mathsf{T}} x, & \\
 & \text{s. t.} & \\
(\mathrm{i}) & \displaystyle\sum_{P \ni a} x_P \leq 1 & \forall a \in A, \\
(\mathrm{iii}) & x \in \{0,1\}^{\mathcal{S}}, &
\end{array}
$$

and

$$
\begin{array}{lll}
(\mathsf{SPP}) & \min d^{\mathsf{T}} x, & \\
 & \text{s. t.} & \\
(\mathrm{i}) & \displaystyle\sum_{P \ni a} x_P = 1, & \forall a \in A, \\
(\mathrm{iii}) & x \in \{0,1\}^{\mathcal{S}}. &
\end{array}
$$

A survey about those problems can be found in Borndörfer [1998].

It is problem specific whether the formulation as a GMCF or a path oriented problem such as SSP, SCP, or SPP is appropriate. In our experience the path formulations are easier to solve if the network contains many infeasible paths. It may also be necessary to formulate a problem path based if the network is very large. This occurs regularly in multi-commodity flow problems with many commodities. If there are only a few additional constraints of type (iii) in (GMCF) the arc based formulation seems in general to give better results. A reason for this may be that for problems arising in applications the LP-relaxation is often already nearly integral.

1.4 Network Design

1.4.1 Description

The network design problem (NDP) consists of selecting routes that can be used by bus lines. The routes are meeting at the *transfer points*. At these points buses can change their route. The routes have to be selected in such a way that a given demand of traffic can be handled and that the cost for setting up the routes and the cost for using them are minimized.

Typically only the extension or modification of an existing network is considered since in most cases historically grown transportation networks exist that can not be modified easily, and because any alteration of lines involves expenses for removal or addition of bus stops, printouts of timetables, building new parking facilities, and marketing to inform the passengers about the changes.

1.4.2 Models

A framework for a slightly more general class of network design problems is presented in Kim & Barnhart [1997]. The model is based on a very similar model in Magnanti & Wong [1984]. We want to sketch here the mixed integer programming model of Kim & Barnhart [1997] and discuss its properties. Given is a network $N = (V, A)$. The nodes V represent potential end points and transfer points of bus lines which include also the origins and destinations of traffic. The set of arcs A models the physical links between these points called *routes*. The set of origin/destination pairs (OD-pairs) is denoted by P, while $O(p)$ and $D(p)$ for $p \in P$ are the origin or destination of p. Let b^p denote the demand of traffic from $O(p)$ to $D(p)$ (measured in number of passengers). The set of potential traffic modes is denoted by F. Typical traffic modes are different types of buses, such as articulated buses or double deckers, trams, or different types of trains. Let u_a^f denote the passenger capacity of one unit of traffic mode $f \in F$ on arc a. The cost $c_a^p, p \in P, a \in A$