Line Planning in Public Transport

CO@Work Berlin

Marika Neumann
Ralf Borndörfer, Marc Pfetsch

10/02/2009
Outline

Introduction

Line Planning with Fixed Passenger Routes

A Column Generation Approach to Line Planning

Line Planning in Practice
Outline

Introduction

Line Planning with Fixed Passenger Routes

A Column Generation Approach to Line Planning

Line Planning in Practice
Planning Steps in Public Transport

- network design
- line planning
- timetabling
- fare planning
- duty scheduling
- vehicle scheduling
Service Design

network design

line planning

timetabling

fare planning

demand
Passenger Demand

Data acquisition

▶ passenger interviews (purpose of trip, start and end,...)
▶ automatic counting in vehicles
   → reflects aggregated passenger demand of current transportation system

Data representation

▶ Origin Destination Matrix (OD Matrix)
▶ passenger volume for each edge

There are statistical and mathematical programming methods for estimating OD matrices from edge counts.
OD Matrix

- public transport area divided into different districts
- district represented by OD node
- OD Matrix – number of passengers traveling between each two OD nodes
OD Matrix

- public transport area divided into different districts
- district represented by OD node
- OD Matrix – number of passengers traveling between each two OD nodes

Features

- aggregated (usually given for a whole day)
- give snapshot type of view
- representation of reality questionable
- industry standard for estimating transportation demand
- no relevant alternative in sight
OD Matrix – Potsdam
Traffic Area Divided in Districts (OD-Nodes)
Line Planning Problem (LPP)

Given:

- network
- demands $d_{st}$ (OD-matrix)
- operating costs and traveling times

**Line:** path (list of stations) with special start and end nodes, frequency

**Problem:** Design lines to satisfy demand.

**Goals:**

- minimize traveling times or number of transfers
- minimize costs of line plan
Literature Overview I

Heuristics:

- Build lines from smaller pieces
  - Remove lines from a “complete” line plan:
    - Patz, 1925
    - Lampkin and Saalmans 1967;
    - Dubois, Bel, and Llibre 1979;
    - Sonntag 1979

- Enumeration of lines:
  - Ceder and Wilson 1986

- Local search:
  - Mandl 1980

- Quadratic covering model:
  - Ceder and Israeli 1992, 1995
Mixed integer programming methods:

- **Fixed Passenger Routes – System Split**
  - Minimize cost:
    - Claessens, van Dijk, and Zwaneveld 1995;
    - Bussieck, Lindner, and Lübbecke 2002
  - Maximize direct travelers:
    - Bouma, Oltrogge 1994;
    - Bussieck, Kreuzer, and Zimmermann 1997

- **Free Routing of Passengers**
  - Minimize transfers/transfer time:
    - Scholl 2005;
    - Schöbel and Scholl 2005
  - Minimize travel time and cost (weighted sum):
    - Borndörfer, Grötschel, Pfetsch, 2005, 2007

(for rail transport)
Bouma and Oltrogge 1994

Idea: Split network into different means of transport (fast train, local train; bus, tram, subway)
→ find line plan for each network, independently

Assumptions on behavior of passengers:
- “choose shortest path”,
- “change to faster system as early as possible”,
- “change to slower system as late as possible”.

→ distribution of passengers to different paths
→ passenger traveling paths are known
Some Notation

Notation for line planning problems (LPP)

- number $M$ of transportation modes (bus, tram, subway,...)
- undirected multigraph $G = (V, E) = (V, E_1 \cup \ldots \cup E_M)$
- terminals: set of nodes $T_1, \ldots, T_M$ where lines can start and end
- OD matrix $d_{st} \in \mathbb{Q}_+^{V \times V}$
- $D = \{(s, t) \in V \times V \mid d_{st} > 0\}$ set of OD pairs
- $\mathcal{L}$ set of lines (simple paths)
- $\mathcal{F}_\ell$ set of frequencies for each line
Outline

Introduction

Line Planning with Fixed Passenger Routes

A Column Generation Approach to Line Planning

Line Planning in Practice
Assumptions

- number of passengers $\rho_e$ for each edge known
- only one mode given, i.e., capacity for all lines equal
- $\mathcal{L}$ pool of predefined lines
Assumptions

- number of passengers $\rho_e$ for each edge known
- only one mode given, i.e., capacity for all lines equal
- $\mathcal{L}$ pool of predefined lines

Definition  
Let

- $G = (V, E)$ be a public transportation graph,
- $\mathcal{L}$ a set of simple line paths in $G$ with capacity $K$
- $\mathcal{F}$ a set of possible frequencies
- $\rho_e$ transportation demand for each edge $e \in E$.

The Feasible Line Plan Problem is to find a set of lines $\mathcal{L}' \subseteq \mathcal{L}$ and frequencies $f_\ell \in \mathcal{F}$ for all $\ell \in \mathcal{L}'$ such that

$$\sum_{\ell \in \mathcal{L}', e \in \ell} K \cdot f_\ell \geq \rho_e \quad \forall e \in E.$$
Example

- public transport network with given demand on edges
- capacity of a line $K = 10$
- possible frequencies $\mathcal{F} = \{1, 2, 3\}$
Example

- public transport network with given demand on edges
- capacity of a line $K = 10$
- possible frequencies $\mathcal{F} = \{1, 2, 3\}$
Definition

Given

- a public transportation graph $G = (V, E)$,
- a set of simple paths $\mathcal{L}$ defined in $G$,
- and a set of edges $E' \subseteq E$ with positive transportation demand.

The Minimum Line Cover Problem is to find a minimum set of lines $\mathcal{L}' \subseteq \mathcal{L}$ that cover all “demand edges” $E'$. 
Subproblem: Minimum Line Cover

Definition

Given

- a public transportation graph \( G = (V, E) \),
- a set of simple paths \( \mathcal{L} \) defined in \( G \),
- and a set of edges \( E' \subseteq E \) with positive transportation demand.

The *Minimum Line Cover Problem* is to find a minimum set of lines \( \mathcal{L}' \subseteq \mathcal{L} \) that cover all “demand edges” \( E' \).
The minimum line cover problem can be formulated as a set covering problem.

\[
\begin{align*}
\min & \quad \sum_{\ell \in \mathcal{L}} x_\ell \\
\text{s.t.} & \quad \sum_{\ell : e \in \ell} x_\ell \geq 1 & \forall e \in E' \\
& \quad x_\ell \in [0, 1] & \forall \ell \in \mathcal{L}
\end{align*}
\]
The minimum line cover problem can be formulated as a set covering problem.

\[
\begin{align*}
\text{min} & \quad \sum_{\ell \in \mathcal{L}} x_{\ell} \\
\text{s.t.} & \quad \sum_{\ell : e \in \ell} x_{\ell} \geq 1 \quad \forall e \in E' \\
\end{align*}
\]

\[x_{\ell} \in [0, 1] \quad \forall \ell \in \mathcal{L}\]

Proposition

*The minimum line cover problem is NP-hard.*

Proof: Reduction from set covering problem.
$S = \{a, b, c, d, e\}, \quad (\{a, c\}, \{b, d\}, \{b, c\}, \{c, e\}, \{a, d, e\})$
$S = \{a, b, c, d, e\}, \quad (\{a, c\}, \{b, d\}, \{b, c\}, \{c, e\}, \{a, d, e\})$
$S = \{a, b, c, d, e\}, \quad (\{a, c\}, \{b, d\}, \{b, c\}, \{c, e\}, \{a, d, e\})$
Cost Minimization Model (Claessens et al.)

\[
\min \sum_{\ell \in \mathcal{L}} \left[ \frac{f_{\ell} T_{\ell}}{T} \right] (C^t + C^c z_{\ell}) + d_{\ell} f_{\ell} (c^t + c^c z_{\ell})
\]

s.t. \[
\Lambda_e \leq \sum_{\ell \in \mathcal{L}_e} f_{\ell} \leq \overline{\Lambda}_e \quad \forall e \in E
\]
\[
\sum_{\ell \in \mathcal{L}_e} K_{\ell} f_{\ell} z_{\ell} \geq \rho_e \quad \forall e \in E
\]
\[
z \leq z_{\ell} \leq \overline{z}
\]
\[
f_{\ell}, z_{\ell} \in \mathbb{Z}_+
\]

Variables: \[ z_{\ell} \quad \text{number of carriages of line} \ \ell \]
\[ f_{\ell} \quad \text{frequency of line} \ \ell \]

Parameter: \[ K \quad \text{capacity of one carriage} \]
\[ \Lambda_e, \overline{\Lambda}_e \quad \text{lower, upper bound on frequency} \ (\overline{\Lambda}_e = \lceil \frac{\rho_e}{z \cdot K} \rceil) \]
\[ z, \overline{z} \quad \text{lower, upper bound on number of carriages} \]
### Objective Function

\[
\begin{aligned}
\left[ \frac{f_\ell T_\ell}{T} \right] (C^t + C^c z_\ell) &+ d_\ell f_\ell (c^t + c^c z_\ell) \\
\quad \text{# trains} &\quad \text{fixed costs} &\quad \text{operating costs}
\end{aligned}
\]

**Parameter:**
- \( T \): time horizon
- \( T_\ell \): turn around time for line \( \ell \)
Objective Function

\[
\left\lceil \frac{f_\ell T_\ell}{T} \right\rceil (C^t + C^c z_\ell) + d_\ell f_\ell (c^t + c^c z_\ell)
\]

Parameter:
- \( T \): time horizon
- \( T_\ell \): turn around time for line \( \ell \)
- \( C^t \): fixed cost for one train
- \( C^c \): fixed cost for one carriage
Objective Function

$$\left\lceil \frac{f_\ell T_\ell}{T} \right\rceil \left( C^t + C^c z_\ell \right) + d_\ell f_\ell \left( c^t + c^c z_\ell \right)$$

**Parameter:**
- $T$: time horizon
- $T_\ell$: turn around time for line $\ell$
- $C^t$: fixed cost for one train
- $C^c$: fixed cost for one carriage
- $d_\ell$: length of line $\ell$
- $c^t$: operating cost for one train per distance
- $c^c$: operating cost for one carriage per distance
Cost Minimization Model

\[
\min \sum_{\ell \in \mathcal{L}} \left[ \frac{f_{\ell} T_{\ell}}{T} \right] \left( C^t + C^c z_{\ell} \right) + d_{\ell} f_{\ell} \left( c^t + c^c z_{\ell} \right)
\]

s.t.

\[
\Lambda_e \leq \sum_{\ell \in \mathcal{L}_e} f_{\ell} \leq \overline{\Lambda}_e \quad \forall e \in E
\]

\[
\sum_{\ell \in \mathcal{L}_e} K f_{\ell} z_{\ell} \geq \rho_e \quad \forall e \in E
\]

\[
z \leq z_{\ell} \leq \overline{z} \quad \forall \ell \in \mathcal{L}
\]

\[
f_{\ell}, z_{\ell} \in \mathbb{Z}_+ \quad \forall \ell \in \mathcal{L}
\]
Cost Minimization Model

\[
\begin{align*}
\min & \quad \sum_{\ell \in \mathcal{L}} \left[ \frac{f_\ell T_\ell}{T} \right] (C^t + C^c z_\ell) + d_\ell f_\ell (c^t + c^c z_\ell) \\
\text{s.t.} & \quad \Lambda_e \leq \sum_{\ell \in \mathcal{L}_e} f_\ell \leq \bar{\Lambda}_e \quad \forall e \in E \\
& \quad \sum_{\ell \in \mathcal{L}_e} K f_\ell z_\ell \geq \rho_e \quad \forall e \in E \\
& \quad z \leq z_\ell \leq \bar{z} \quad \forall \ell \in \mathcal{L} \\
& \quad f_\ell, z_\ell \in \mathbb{Z}_+ \quad \forall \ell \in \mathcal{L}
\end{align*}
\]

Linearization

- $\mathcal{F}$ set of feasible frequencies, e.g., $\mathcal{F} = \{1, \ldots, F\}$
- $\mathcal{C}$ set of feasible numbers of carriages, e.g., $\mathcal{C} = \{3, 4, 5\}$
- $\mathcal{R} = \mathcal{L} \times \mathcal{F} \times \mathcal{C}$
Linearized Cost Model (Goossens et al.)

\[
\min \sum_{r \in \mathcal{R}} \left( \left[ \frac{f_{r\ell} T_{r\ell}}{T} \right] (C^t + C^c r_z) + d_{r\ell} r_f (c^t + c^c r_z) \right) \cdot y_r
\]

s.t. \( \Lambda_e \leq \sum_{r \in \mathcal{R}: e \in r_{\ell}} r_f y_r \leq \bar{\Lambda}_e \) \quad \forall e \in E

\sum_{r \in \mathcal{R}: e \in r_{\ell}} K r_f r_z y_r \geq \rho_e \quad \forall e \in E

\sum_{r \in \mathcal{R}: r_{\ell} = \ell} y_r \leq 1 \quad \forall \ell \in \mathcal{L}

\]

\( y_r \in \{0, 1\} \quad \forall r \in \mathcal{R} \)

Variables: \( y_r \) choosing combination of \( r = (r_{\ell}, r_f, r_z) \in \mathcal{R} \) (line frequency and number of carriage)

Solving with preprocessing and branch-and-cut methods.
Proposition

The cost minimizing line planning approach is NP-hard.

Proof.

Setting

- \( z = \bar{z} \), (i.e., fixed number of carriages),
- \( F = 1 \), (i.e., fixed frequency),
- \( K = \max\{\rho_e \mid e \in E\} \),
- \( \Lambda_e = 1, \overline{\Lambda}_e = \infty \)
- \( C^t = 1, C^c = c^c = c^t = 0 \)

leads to a minimum line cover problem.
Cutting Plane – Example

\[ \mathcal{F} = \mathcal{C} = \{1\}, \mathcal{L} = \{1, 2, 3\}, K = 10 \]

Consider capacity constraint

\[
\sum_{r \in R: e \in r} K r f r z y_r \geq \rho_e \quad \forall e
\]
Cutting Plane – Example

\[ \rho_{e_1} = 10 \]
\[ \rho_{e_2} = 10 \]
\[ \rho_{e_3} = 10 \]

\[ \mathcal{F} = \mathcal{C} = \{1\}, \mathcal{L} = \{1, 2, 3\}, K = 10 \]

Consider capacity constraint

\[ \sum_{r \in \mathcal{R} : e \in r} K r_f r_z y_r \geq \rho_e \quad \forall e \]

1. \[ 10 \cdot y_1 + 10 \cdot y_2 \geq 10 \quad \{e_1\} \]
2. \[ 10 \cdot y_1 + 10 \cdot y_3 \geq 10 \quad \{e_2\} \]
3. \[ 10 \cdot y_2 + 10 \cdot y_3 \geq 10 \quad \{e_3\} \]
Cutting Plane – Example

\[ \rho_{e_1} = 10 \]
\[ \rho_{e_3} = 10 \]
\[ \rho_{e_2} = 10 \]

- \( \mathcal{F} = \mathcal{C} = \{1\}, \mathcal{L} = \{1, 2, 3\}, K = 10 \)
- Consider capacity constraint:

\[
\sum_{r \in \mathcal{R}: e \in r} K r_f r_z y_r \geq \rho_e \quad \forall e
\]

- \( 10 \cdot y_1 + 10 \cdot y_2 \geq 10 \quad \{e_1\} \)
- \( 10 \cdot y_1 + 10 \cdot y_3 \geq 10 \quad \{e_2\} \)
- \( 10 \cdot y_2 + 10 \cdot y_3 \geq 10 \quad \{e_3\} \)

- \( y_1 = y_2 = y_3 = 0.5 \) is solution
- \( \Rightarrow y_1 + y_2 + y_3 \geq 2 \) valid
Cutting Plane – Example

Proposition

Let

- $E' \subseteq E$ such that $\rho_e > 0 \ \forall \ e \in E'$,
- $\alpha_{E'}^\ell := |\ell \cap E'|$, and
- $\alpha_{E'}^{\text{max}} := \max\{\alpha_{E'}^\ell | \ell \in \mathcal{L}\}$.

Then

$$\sum_{r \in \mathcal{R}, \alpha_{E'}^{r\ell} \geq 1} y_r \geq \left\lceil \frac{|E'|}{\alpha_{E'}^{\text{max}}} \right\rceil$$

is a valid inequality.
Discussion of the Model

- detailed cost function based on following assumption
  - no switching of rolling stock between lines
  - line is operated by same trains (same number of carriages)
  - timetable is periodic (e.g. repeated every hour)
- many variables (every possible combination of frequency and number of carriages)
  however, reduction by preprocessing
- only one transportation mode considered
- passenger paths are fixed
- line pool
Outline

Introduction

Line Planning with Fixed Passenger Routes

A Column Generation Approach to Line Planning

Line Planning in Practice
Example – Free Passenger Routing

- public transport network with following demand (OD pairs):
  - $a \rightarrow c$: 10
  - $a \rightarrow d$: 10
  - $d \rightarrow c$: 10
  - $c \rightarrow f$: 20
  - $d \rightarrow f$: 20

- capacity of a line $K = 10$, frequencies $\mathcal{F} = \{1, 2, 3\}$
- public transport network with following demand (OD pairs):
  - $a \rightarrow c$: 10
  - $a \rightarrow d$: 10
  - $d \rightarrow c$: 10
  - $c \rightarrow f$: 20
  - $d \rightarrow f$: 20

- capacity of a line $K = 10$, frequencies $\mathcal{F} = \{1, 2, 3\}$
Example – Free Passenger Routing

- public transport network with following demand (OD pairs):
  - $a \rightarrow c$: 10
  - $a \rightarrow d$: 10
  - $d \rightarrow c$: 10
  - $c \rightarrow f$: 20
  - $d \rightarrow f$: 20

- capacity of a line $K = 10$, frequencies $\mathcal{F} = \{1, 2, 3\}$
- $\rightsquigarrow$ directed graph $G = (V, A)$ for passenger paths

\[ \begin{array}{cccc}
   & a & c & d & f \\
 a & 0 & 10 & 10 & 0 \\
 c & 0 & 0 & 0 & 20 \\
 d & 0 & 10 & 0 & 20 \\
 f & 0 & 0 & 0 & 0 \\
\end{array} \]
Multi Commodity Flow Model  

\[
\min_{\lambda} \sum_{\ell} \left( C_{\ell} x_{\ell} + c_{\ell} f_{\ell} \right) + (1 - \lambda) \sum_{p} \tau_p y_p
\]

i) \[\sum_{p \in P_{st}} y_p = d_{st} \quad \forall (s, t) \in D\] transport all passengers

ii) \[\sum_{p \ni a} y_p \leq \sum_{\ell: e(a) \in \ell} \kappa_{\ell} f_{\ell} \quad \forall a \in A\] arc capacity constraints

iii) \[\sum_{\ell \ni e} f_{\ell} \leq \Lambda_e \quad \forall e \in E\] frequency bounds

iv) \[f_{\ell} \leq F x_{\ell} \quad \forall \ell \in L\] coupling constraints

\[y_p \in \mathbb{R}_+ \quad \forall p \in P\] passenger flow

\[x_{\ell} \in \{0, 1\} \quad \forall \ell \in L\] choose line \(\ell\)

\[f_{\ell} \in \mathbb{R}_+ \quad \forall \ell \in L\] frequency of line \(\ell\)
Discussion of the Model

Properties of the model:

▶ no system split
▶ continuous frequencies
▶ other linear constraints possible
▶ system optimum = user equilibrium

Advantages of the model:

▶ Traveling paths of passengers are not fixed a priori.
▶ Lines can be generated dynamically (column generation).

Disadvantages of the model:

▶ Some passengers may use long paths.
   possible solution: length constraints
▶ Transfers between lines of same type cannot be controlled.
Discussion of the Model

Properties of the model:
- no system split
- continuous frequencies
- other linear constraints possible
- system optimum = user equilibrium

Advantages of the model:
- Traveling paths of passengers are not fixed a priori.
- Lines can be generated dynamically (column generation).
Discussion of the Model

Properties of the model:

- no system split
- continuous frequencies
- other linear constraints possible
- system optimum = user equilibrium

Advantages of the model:

- Traveling paths of passengers are not fixed a priori.
- Lines can be generated dynamically (column generation).

Disadvantages of the model:

- Some passengers may use long paths.
  possible solution: length constraints
- Transfers between lines of same type cannot be controlled.
Transfers

Lines of different type:
Time penalties for transfers

Lines of same type:
Capacity constraints do not distinguish between these lines:

\[
\sum_{p \ni a} y_p \leq \sum_{\ell : e(a) \in \ell} \sum_{f \in F_{\ell}} \kappa_{\ell}^f x_{\ell}^f
\]

Solution: expansion of the graph:

Problem: Symmetries
$$\begin{align*}
\text{min} & \quad \sum_{\ell \in \mathcal{L}} (C_{\ell} x_{\ell} + c_{\ell} f_{\ell}) + \sum_{p \in \mathcal{P}} \tau_p y_p \\
\text{s.t.} & \quad \sum_{p \in \mathcal{P}_{st}} y_p = d_{st} \quad \forall (s, t) \in D \\
& \quad \sum_{p : a \in p} y_p \leq \sum_{\ell : e(a) \in \ell} \kappa_{\ell} f_{\ell} \quad \forall a \in \mathcal{A} \\
& \quad \sum_{\ell : e \in \ell} f_{\ell} \leq \Lambda_e \quad \forall e \in \mathcal{E} \\
& \quad f_{\ell} \leq F x_{\ell} \quad \forall \ell \in \mathcal{L} \\
& \quad x_{\ell} \in [0, 1] \quad \forall \ell \in \mathcal{L} \\
& \quad f_{\ell} \geq 0 \quad \forall \ell \in \mathcal{L} \\
& \quad y_p \geq 0 \quad \forall p \in \mathcal{P}.
\end{align*}$$

Solve LP relaxation with column generation.

Proposition

The computation of the optimal value of (LP) is NP-hard.
(LP) \( \min \sum_{\ell \in \mathcal{L}} \gamma_\ell f_\ell + \sum_{p \in \mathcal{P}} \tau_p y_p \)

s.t.

\[ \sum_{p \in \mathcal{P}_{st}} y_p = d_{st} \quad \forall (s, t) \in D \]

\[ \sum_{p : a \in p} y_p - \sum_{\ell : e(a) \in \ell} \kappa_\ell f_\ell \leq 0 \quad \forall a \in \mathcal{A} \]

\[ \sum_{\ell : e \in \ell} f_\ell \leq \Lambda_e \quad \forall e \in \mathcal{E} \]

\[ f_\ell \geq 0 \quad \forall \ell \in \mathcal{L} \]

\[ y_p \geq 0 \quad \forall p \in \mathcal{P}. \]
LP Relaxation

\[
\text{(LP)} \quad \min \sum_{\ell \in \mathcal{L}} \gamma_{\ell} f_{\ell} + \sum_{p \in \mathcal{P}} \tau_p y_p
\]

s.t. \[
\sum_{p \in \mathcal{P}_{st}} y_p = d_{st} \quad \forall (s, t) \in D
\]
\[
\sum_{p : a \in p} y_p - \sum_{\ell : \ell(a) \in \ell} \kappa_{\ell} f_{\ell} \leq 0 \quad \forall a \in A
\]
\[
\sum_{\ell : e \in \ell} f_{\ell} \leq \Lambda_e \quad \forall e \in E
\]
\[
f_{\ell} \geq 0 \quad \forall \ell \in \mathcal{L}
\]
\[
y_p \geq 0 \quad \forall p \in \mathcal{P}.
\]

Solve LP relaxation with column generation.

Proposition

The computation of the optimal value of \((LP)\) is NP-hard.
The dual model is:

\[
\text{(DLP)} \quad \min \sum_{(s,t) \in D} d_{st} \pi_{st} - \sum_{e \in E} \Lambda_e \eta_e
\]

s.t.

\[
\begin{align*}
\pi_{st} - \sum_{a \in p} \mu_a & \leq \tau_p \quad \forall p \in \mathcal{P}_{st}, (s,t) \in D \\
\sum_{e \in \ell} (\kappa_{\ell}(\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e) & \leq \gamma_{\ell} \quad \forall \ell \in \mathcal{L} \\
\eta_e & \geq 0 \quad \forall e \in E \\
\mu_a & \geq 0 \quad \forall a \in A
\end{align*}
\]
The dual model is:

\[
(DLP) \quad \min \sum_{(s,t) \in D} d_{st} \pi_{st} - \sum_{e \in E} \Lambda_e \eta_e \\
\text{s.t.} \quad \pi_{st} - \sum_{a \in p} \mu_a \leq \tau_p \quad \forall p \in \mathcal{P}_{st}, \ (s, t) \in D \\
\sum_{e \in \ell} (k_\ell (\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e) \leq \gamma_\ell \quad \forall \ell \in \mathcal{L} \\
\eta_e \geq 0 \quad \forall e \in E \\
\mu_a \geq 0 \quad \forall a \in A
\]

Reduced cost $\bar{\tau}_p$ for $y_p$, $p \in \mathcal{P}_{st}, \ (s, t) \in D$:
Pricing of Passenger Paths

The dual model is:

\[\text{(DLP)} \quad \min \sum_{(s,t) \in D} d_{st} \pi_{st} - \sum_{e \in E} \Lambda_e \eta_e\]

s.t.

\[\pi_{st} - \sum_{a \in p} \mu_a \leq \tau_p \quad \forall p \in \mathcal{P}_{st}, (s, t) \in D\]

\[\sum_{e \in \ell} \left(\kappa_{\ell}(\mu_{a(e)} + \mu_{\bar{a}(e)}) - \eta_e\right) \leq \gamma_{\ell} \quad \forall \ell \in \mathcal{L}\]

\[\eta_e \geq 0 \quad \forall e \in E\]

\[\mu_a \geq 0 \quad \forall a \in A\]

Reduced cost \(\overline{\tau}_p\) for \(y_p, p \in \mathcal{P}_{st}, (s, t) \in D\):

\[\overline{\tau}_p = \tau_p - \pi_{st} + \sum_{a \in p} \mu_a = -\pi_{st} + \sum_{a \in p} (\mu_a + \tau_a)\]

\[\overline{\tau}_p < 0 \iff \sum_{a \in p} (\mu_a + \tau_a) < \pi_{st}\]
Pricing of Passenger Paths

The dual model is:

\[
\text{(DLP) } \min \sum_{(s,t) \in D} d_{st} \pi_{st} - \sum_{e \in E} \Lambda_e \eta_e
\]

s.t.

\[
\pi_{st} - \sum_{a \in p} \mu_a \leq \tau_p \quad \forall p \in \mathcal{P}_{st}, (s,t) \in D
\]

\[
\sum_{e \in \ell} \left( \kappa_{\ell}(\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e \right) \leq \gamma_{\ell} \quad \forall \ell \in \mathcal{L}
\]

\[
\eta_e \geq 0 \quad \forall e \in E
\]

\[
\mu_a \geq 0 \quad \forall a \in A
\]

Reduced cost \( \overline{\tau}_p \) for \( y_p, p \in \mathcal{P}_{st}, (s,t) \in D \):

\[
\overline{\tau}_p = \tau_p - \pi_{st} + \sum_{a \in p} \mu_a = -\pi_{st} + \sum_{a \in p} (\mu_a + \tau_a)
\]

\[
\overline{\tau}_p < 0 \iff \sum_{a \in p} (\mu_a + \tau_a) < \pi_{st}
\]

\( \rightsquigarrow \) shortest path problem
Pricing of Line Paths

\[(\text{DLP}) \quad \min \sum_{(s,t) \in D} d_{st} \pi_{st} - \sum_{e \in E} \Lambda_e \eta_e \]

s.t.
\[\pi_{st} - \sum_{a \in p} \mu_a \leq \tau_p \quad \forall p \in \mathcal{P}_{st}, (s,t) \in D\]
\[\sum_{e \in \ell} \left( \kappa_{\ell} \left( \mu_a(e) + \mu_{\bar{a}}(e) \right) - \eta_e \right) \leq \gamma_{\ell} \quad \forall \ell \in \mathcal{L}\]
\[\eta_e \geq 0 \quad \forall e \in E\]
\[\mu_a \geq 0 \quad \forall a \in A\]

Reduced cost $\overline{\gamma}_{\ell}$ for $f_{\ell}$, $\ell \in \mathcal{L}$, $(s,t) \in D$: 
Pricing of Line Paths

\[(\text{DLP}) \quad \min \sum_{(s,t)\in D} d_{st} \pi_{st} - \sum_{e\in E} \Lambda_e \eta_e \]

s.t. \[
\pi_{st} - \sum_{a\in p} \mu_a \leq \tau_p \quad \forall p \in \mathcal{P}_{st}, (s,t) \in D
\]
\[
\sum_{e\in \ell}(\kappa_\ell(\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e) \leq \gamma_\ell \quad \forall \ell \in \mathcal{L}
\]
\[
\eta_e \geq 0 \quad \forall e \in E
\]
\[
\mu_a \geq 0 \quad \forall a \in A
\]

Reduced cost $\overline{\gamma}_\ell$ for $f_\ell$, $\ell \in \mathcal{L}$, $(s,t) \in D$:
\[
\overline{\gamma}_\ell = \gamma_\ell - \sum_{e\in \ell}(\kappa_\ell(\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e)
\]
\[ \bar{\gamma}_\ell = \gamma_\ell - \sum_{e \in \ell} (\kappa_\ell (\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e) \]
\[ \bar{\gamma}_\ell = \gamma_\ell - \sum_{e \in \ell} (\kappa_\ell (\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e) \]

\[ = \frac{C_\ell}{F} + c_\ell - \sum_{e \in \ell} (\kappa_\ell (\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e) \]
\[ \bar{\gamma}_\ell = \gamma \ell - \sum_{e \in \ell} (\kappa \ell (\mu_a(e) + \mu \tilde{a}(e)) - \eta_e) \]

\[ = \frac{C_\ell}{F} + c \ell - \sum_{e \in \ell} (\kappa \ell (\mu_a(e) + \mu \tilde{a}(e)) - \eta_e) \]

\[ = \frac{C_i}{F} + \sum_{e \in \ell} c^i_e - \sum_{e \in \ell} (\kappa_i (\mu_a(e) + \mu \tilde{a}(e)) - \eta_e) \]
Pricing of Line Paths

$$\bar{\gamma}_\ell = \gamma_\ell - \sum_{e \in \ell} (\kappa_\ell (\mu_a(e) + \mu_\bar{a}(e)) - \eta_e)$$

$$= \frac{C_\ell}{F} + c_\ell - \sum_{e \in \ell} (\kappa_\ell (\mu_a(e) + \mu_\bar{a}(e)) - \eta_e)$$

$$= \frac{C_i}{F} + \sum_{e \in \ell} c_i^e - \sum_{e \in \ell} (\kappa_i (\mu_a(e) + \mu_\bar{a}(e)) - \eta_e)$$

$$= \frac{C_i}{F} - \sum_{e \in \ell} (\kappa_i (\mu_a(e) + \mu_\bar{a}(e)) - \eta_e - c_i^e)$$
\[
\bar{\gamma}_\ell = \gamma_\ell - \sum_{e \in \ell} (\kappa_\ell (\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e) \\
= \frac{C_\ell}{F} + c_\ell - \sum_{e \in \ell} (\kappa_\ell (\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e) \\
= \frac{C_i}{F} + \sum_{e \in \ell} c^i_e - \sum_{e \in \ell} (\kappa_i (\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e) \\
= \frac{C_i}{F} - \sum_{e \in \ell} (\kappa_i (\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e - c^i_e)
\]

\[
0 > \bar{\gamma}_\ell \iff \sum_{e \in \ell} (\kappa_i (\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e - c^i_e) > \frac{C_i}{F}
\]
\[ \bar{\gamma}_\ell = \gamma_\ell - \sum_{e \in \ell} (\kappa_\ell (\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e) \]
\[ = \frac{C_\ell}{F} + c_\ell - \sum_{e \in \ell} (\kappa_\ell (\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e) \]
\[ = \frac{C_i}{F} + \sum_{e \in \ell} c_e^i - \sum_{e \in \ell} (\kappa_i (\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e) \]
\[ = \frac{C_i}{F} - \sum_{e \in \ell} (\kappa_i (\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e - c_e^i) \]

\[ 0 > \bar{\gamma}_\ell \iff \sum_{e \in \ell} (\kappa_i (\mu_a(e) + \mu_{\bar{a}}(e)) - \eta_e - c_e^i) > \frac{C_i}{F} \]

\[ \sim \text{longest path problem (NP-hard)} \]
Pricing of Line Paths

Let $n$ be the number of nodes.

**Theorem**

*If the lengths of paths are $O(\log n)$, one can solve the longest path problem in polynomial time.*

**Corollary**

*If the lengths of lines are $O(\log n)$, one can solve the LP relaxation in polynomial time.*

Alternative method: Find lines by *enumeration*. 
Outline

Introduction

Line Planning with Fixed Passenger Routes

A Column Generation Approach to Line Planning

Line Planning in Practice
Current Practice

- sophisticated simulation tools, e.g., VISUM (but no mathematical optimization methods)
- experience of practitioners
- political requirements
Cooperation with: ViP Potsdam

network of Potsdam:

- inhabitants: 150,000
- travels in morning traffic: 42,973
- number of bus lines: 15 + 8
- number of tram lines: 6
- nodes: 872 (1643)  
  edges: 2462 (5470)
- OD-nodes: 385  
  nonzeros: 12,787
Variation of (Grötschel, Borndörfer, Pfetsch)

\[
\min \quad \lambda \sum_{\ell \in \mathcal{L}} \sum_{f \in \mathcal{F}} c_{\ell}^{f} x_{\ell}^{f} + (1 - \lambda) \sum_{p \in \mathcal{P}} \tau_{p} y_{p}
\]

i) \[ \sum_{p \in \mathcal{P}_{st}} y_{p} = d_{st} \quad \forall \ (s, t) \in D \] transport all passengers

ii) \[ \sum_{p \ni \alpha} y_{p} \leq \sum_{\ell : e(\alpha) \in \ell} \sum_{f \in \mathcal{F}_{\ell}} \kappa_{\ell}^{f} x_{\ell}^{f} \quad \forall \ \alpha \in \mathcal{A} \] arc capacity constraints

iii) \[ \sum_{f \in \mathcal{F}_{\ell}} x_{\ell}^{f} \leq 1 \quad \forall \ \ell \in \mathcal{L} \] one frequency per line

\[ y_{p} \in \mathbb{R}_{+} \quad \forall \ p \in \mathcal{P} \] passenger flow

\[ x_{\ell}^{f} \in \{0, 1\} \quad \forall \ \ell \in \mathcal{L} \] line and frequency

- pricing problem for passenger paths similar as before
- pricing of line paths? (exercise)
How to present the solution?
Things to Do Besides Optimization

How to present the solution?

- list of lines with stations and frequencies, basic visualization
line name cycle time in minutes:
- BVB134R;32088;U 30
- BVB234H;10723;U 30
- BVB534R;35166;U 30
- HVG612;180287;U 30
- N01RB20H;612362;U 60
- N01RB21R;610169;U 60
- VIB692H;530239;U 30
- VIT92;92_KA-MJ;U 20

list of stations for each line:

BVB134R;32088;U Hottengrund (Berlin)
BVB134R;32088;U Kaserne Hottengrund (Berlin)
BVB134R;32088;U Temmeweg (Berlin)
BVB134R;32088;U Gößweinsteinergang (Berlin)
BVB134R;32088;U Parnemannweg (Berlin)
BVB134R;32088;U Alt-Kladow (Berlin)
BVB134R;32088;U Finnenhaus-Siedlung (Berlin)
BVB134R;32088;U Neukladower Allee (Berlin)
BVB134R;32088;U Krankenhaus Havelhöhe (Berlin)
BVB134R;32088;U General-Steinhoff-Kaserne (Berlin)
BVB134R;32088;U Weg nach Breitehorn (Berlin)
BVB134R;32088;U Breitehornweg (Berlin)
BVB134R;32088;U Helleberge (Berlin)
BVB134R;32088;U Am Graben (Berlin)
BVB134R;32088;U Alt-Gatow (Berlin)
BVB134R;32088;U Gatow Kirche (Berlin)
BVB134R;32088;U Pfirsichweg (Berlin)
BVB134R;32088;U Emil-Basdeck-Str. (Berlin)
BVB134R;32088;U Biberburg (Berlin)
BVB134R;32088;U Zur Haveldüne (Berlin)
BVB134R;32088;U Gatower Str./Weinmeisterhornweg (Berlin)
BVB134R;32088;U Sandheideweg (Berlin)
BVB134R;32088;U Gatower Str./Breitehornweg (Berlin)
How to present the solution?

- list of lines with stations and frequencies, visualization with matlab
  *advantage*: easy
  *disadvantage*: very rudimental, e.g., no switching through lines
Things to do besides Optimization

How to present the solution?

▶ list of lines with stations and frequencies, visualization with matlab
  advantage: easy
  disadvantage: very rudimentary, e.g., no switching through lines

▶ visualization tool implemented by M. Kinder (student at ZIB)
Things to do besides Optimization

How to present the solution?

- list of lines with stations and frequencies, visualization with matlab
  advantage: easy
  disadvantage: very rudimental, e.g., no switching through lines

- visualization tool implemented by M. Kinder (student at ZIB)
  advantage: tool was ready
  disadvantage: no geographic map, restricted evaluation of solution
Things to do besides Optimization

How to present the solution?

- list of lines with stations and frequencies, visualization with matlab
  advantage: easy
  disadvantage: very rudimental, e.g., no switching through lines

- visualization tool implemented by M. Kinder (student at ZIB)
  advantage: tool was ready
  disadvantage: no geographic map, restricted evaluation of solution

- visualization with map
How to present the solution?

- list of lines with stations and frequencies, visualization with matlab
  advantage: easy
  disadvantage: very rudimental, e.g., no switching through lines

- visualization tool implemented by M. Kinder (student at ZIB)
  advantage: tool was ready
  disadvantage: no geographic map, restricted evaluation of solution

- visualization with map
  advantage: geographic map, part of software used by Potsdam
  disadvantage: restricted evaluation of solution
Things to do besides Optimization

How to present the solution?

- list of lines with stations and frequencies, visualization with matlab
  advantage: easy
  disadvantage: very rudimental, e.g., no switching through lines

- visualization tool implemented by M. Kinder (student at ZIB)
  advantage: tool was ready
  disadvantage: no geographic map, restricted evaluation of solution

- visualization with map
  advantage: geographic map, part of software used by Potsdam
  disadvantage: restricted evaluation of solution

- visualization with VISUM
  advantage: geographic map, evaluation of solution possible
  disadvantage: expensive
Other Things to Do

Data

- some data missing, incomplete, contains errors
  ⇒ long iterative process to get needed data

Parameter

- cost function, operating cost for each line
- capacities of lines
- choose terminal nodes (endpoint of line)
- weighting of cost and travel time (choosing $\lambda$)
  ⇒ computation of Pareto curve
Pareto Curve

Fahrzeit [m]

Borndörfer, Neumann, Pfetsch ()
Line Planning in Public Transport
10/02/2009 56 / 61
column generation of passenger paths, predefined line pool
(computations in chronological order):

| Instance  | $|\mathcal{L}|$ | $|\mathcal{F}|$ | #line var | #constr. | time$^1$ |
|-----------|----------------|----------------|-----------|----------|---------|
| potsdam1  | 30 232         | 4              | 120 592   | 44 577   | 349     |
| potsdam2  | 29 142         | 2              | 58 268    | 43 419   | 125     |
| potsdam3  | 1172           | 3              | 3 486     | 15 491   | 24      |
| potsdam4  | 623            | 3              | 1 755     | 14 939   | 20      |
| potsdam5  | 3861           | 3              | 11 471    | 21 199   | 377$^2$ |

$^1$in minutes, after solving root node including separators and heuristics; gap < 5%

$^2$for adjusted network
Problems

General Problems

- data contains information relevant for transportation company but not relevant for optimization tools (e.g., nodes of network not only stops but also crossings, track switches, turnouts,..)

Problems concerning our solutions

- too many too short lines
- service (frequency) on some stations to small
- the importance of tram not represented by our solution (important for tourism, environment, prestige)
- lines over a train crossing – no robust timetable
- “curious” bus lines (no stations shared with tram, regional traffic)
Problems – Solutions

- relevant vs. irrelevant data
Problems – Solutions

- relevant vs. irrelevant data
  ~ preprocessing
- too many too short lines
Problems – Solutions

- relevant vs. irrelevant data
  \(\Rightarrow\) preprocessing

- too many too short lines
  Idea: include fixed cost

- service (frequency) on some stations to small
Problems – Solutions

- relevant vs. irrelevant data
  ⇒ preprocessing

- too many too short lines
  Idea: include fixed cost

- service (frequency) on some stations too small
  ⇒ bound on minimal frequency for serving a station

- the importance of the tram not represented by our solution
Problems – Solutions

- relevant vs. irrelevant data
  - preprocess

- too many too short lines
  Idea: include fixed cost

- service (frequency) on some stations to small
  - bound on minimal frequency for serving a station

- the importance of the tram not represented by our solution
  Idea: condition of covering all tracks of tram

- lines over train crossing
Problems – Solutions

- relevant vs. irrelevant data
  → preprocessing

- too many too short lines
  Idea: include fixed cost

- service (frequency) on some stations to small
  → bound on minimal frequency for serving a station

- the importance of the tram not represented by our solution
  Idea: condition of covering all tracks of tram

- lines over train crossing
  Idea: high cost for using crossing in line; prohibit the use of crossing

- “curious” bus lines (no stations shared with tram, regional traffic)
Problems – Solutions

- relevant vs. irrelevant data
  ⇝ preprocessing

- too many too short lines
  Idea: include fixed cost

- service (frequency) on some stations to small
  ⇝ bound on minimal frequency for serving a station

- the importance of the tram not represented by our solution
  Idea: condition of covering all tracks of tram

- lines over train crossing
  Idea: high cost for using crossing in line; prohibit the use of crossing

- “curious” bus lines (no stations shared with tram, regional traffic)
  Idea: generate only bus lines that contain a station shared with tram or regional traffic
Work in Progress...
Line Plan – Solutions

Work in Progress...

- first step to establish optimization methods for line planning in practice
- line planning optimization not completely solved
- hope for future: optimization in (service design of) public transport as decision support