Univ.-Prof. Dr. Sven Müller
Business Analytics

Tasks and slide collection

Aachen, Germany, 9^{th} April, 2024

Part A

Environmental situation

Chapter 1

Environmental situation

1.1 The importance of PT

§2 RegG: (Act on the regionalization of PT) Public transportation is (...) generally accessible transportation of passengers by means of vehicles on regular service (...) in city, suburb or regional traffic. The majority of transportation cases involve a travel distance of up to 50 km and a maximum travel time of one hour.

The Federal Government considers an efficient and attractive public transportation system (PT) to be an indispensable contribution to the solution of current and future mobility requirements in cities and municipalities. Every day, 26 million citizens use public transportation. This public task of providing services of general interest is performed by more than 250,000 employees in around 6,000 private and municipal transportation companies. The associated preservation of mobility relieves the conurbations of individual traffic and guarantees equal living conditions in the regions. In addition, buses and trains help to relieve the environment and reduce climate-relevant emissions. In particular local rail passenger transportation is efficient in handling large commuter flows due to its system-related advantages (e.g. rail-based to avoid traffic). The quality of public transportation is increasingly becoming a location factor in the competition for investments and jobs. After all, buses and trains help to improve traffic safety by running on lanes separated from regular road traffic and thus alleviating traffic congestion.

The total public transfer payments for this sector amount to around 30 billion per year. The Federation alone contributes about 15 billion or 50% of this amount through the Regionalisation Act and the Municipal Transportation Financing Act. While public transportation falls under the jurisdiction of the states, the Federal Government demonstrates its commitment to supporting the sector with a substantial budget in the tens of billions of euros. (**?**)

1.2 General legal aspects

Legal framework

- Passenger Transportation Act (PBefG)
- Law on the Regionalization of Public Transportation (RegG)
- Law on Local Public Transportation in North Rhine-Westphalia (ÖPNVG)
- Regulation (EWG) No 1191/69 of the Council of 26 June 1969; the obligations of the member states in respect of transportation by rail, road and inland waterway inherent in the concept of a public service
- General Railway Act (AEG)
- Local transportation plan
- Social act (SGB)
- Municipal Transportation Financing Act (GVFG)

The German national legal framework and administrative practice to date are to be preserved as far as possible. The German licensing system distinguishes between two groups of transportation services: **commercial** resp. **socio-economic traffics**.

In the case of socio-economic traffics, a contract between the public authority and the operator will become the rule in the future, on the basis of a call for tenders, in order to compensate the public service accordingly. Commercial traffics will not be put out to tender, but will be subject to competition for the best bid. The government states that it must be ensured what is considered to be commercial and socio-economic traffic through clear and unambiguous definitions of the financial services provided by public authorities (Bundesregierung 2000, s.).

The local transportation plan (LTP; German: Nahverkehrsplan (NVP)) summarizes the transportation services to be provided as far as possible. The interaction of the parties involved is shown in the following figure:

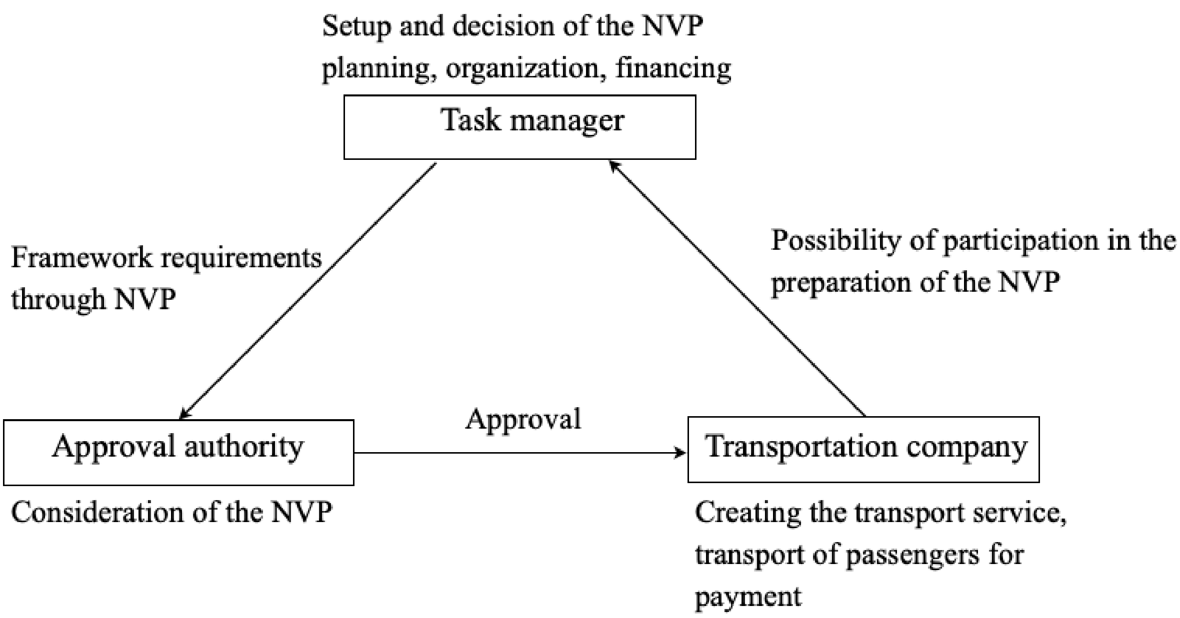


Figure 1.1: Process of creating the LTP

1.3 General development

- dismantling of grandfather rights
- increased competition for concessions
- acquisition of companies
- design of cooperations; in particular of transportation associations
- the possibility for transportation companies to participate in the development of local transportation plans
- demographic changes, especially declining student numbers
- expansion of flexible operating modes
- reduction of compensation payments for the transportation of trainees at preferential conditions (reduced time tickets)

1.4 Financing sources

Service revenues

- fare revenues
- compensation for school transportation (§45 PBefG)
- compensation for transportation of severely disabled persons (§145 SGB 9)

Subsidies

- funds according RegG
- funding according GVFG

Subsidies for operating costs

- vehicle tax exemption
- reduced VAT rate of 7%

1.5 Service creation process

The following (successive) planning process for the creation of the service is typical for transportation companies:

1. route network design
2. route planning
3. tariff design
4. schedule formation
5. circulation planning
6. duty scheduling
7. duty roster planning

Further important (operational) areas of the service creation process are

- dispatching and
- external and internal information management.

Chapter 2

Graph theory

2.1 Introduction:

Graph theory is a mathematical discipline that deals with the study of graphs, which are mathematical structures that represent relationships between objects. A graph is a collection of vertices (or nodes) connected by edges (or arcs). Graph theory has a wide range of applications in various fields, including transportation.

2.1.1 Basic Concepts of Graph Theory:

Vertices and Edges:

In graph theory, a vertex represents a point or a location, and an edge represents a connection or a link between two vertices. Vertices are often used to represent locations such as cities or intersections, and edges represent roads, railways, or other transportation routes.

Directed and Undirected Graphs:

In a directed graph, the edges have a direction, indicating a one-way connection between vertices. An edge with a direction is called arc. In an undirected graph, the edges have no direction, and connections are bidirectional. Directed graphs are often used to model transportation systems with one-way streets or traffic flow, while undirected graphs can represent systems with two-way roads or connections.

Weighted Graphs:

In some transportation problems, edges in a graph may have associated weights or costs that represent distances, travel times, or other relevant factors. Such graphs are called weighted graphs. Weighted graphs are commonly used to model transportation networks where the cost of travel along an edge is not uniform, such as road networks with varying road lengths or traffic congestion.

Paths and Cycles:

In graph theory, a path is a sequence of connected edges that form a route between two vertices. A cycle is a path that starts and ends at the same vertex, forming a loop. Paths and cycles are important concepts in transportation problems, as they represent possible routes or circuits in a transportation network.

Connectivity:

The connectivity of a graph refers to how vertices are connected to each other. A graph can be connected, meaning that there is a path between any two vertices, or it can be disconnected, meaning that there are isolated vertices or groups of vertices that are not connected to the rest of the graph. Connectivity is important in transportation problems, as it can affect the accessibility and efficiency of a transportation network.

Application of Graph Theory in Transportation Problems:

Graph theory has various areas of application in transportation problems, including...

Route Planning:

Graphs can be used to model transportation networks, such as road networks, airline routes, or public transportation systems. By representing these networks as graphs, graph theory can be used to find optimal or efficient routes between locations, considering factors such as distances, travel times, or costs associated with edges.

Network Design:

Network Design: Graph theory can be used to design or optimize transportation networks. For example, it can be used to determine the optimal locations for transportation hubs, such as airports or logistics centers, to minimize transportation costs or improve connectivity. It can also be used to design transportation networks that are robust against failures or disruptions.

Traffic Flow Analysis:

Traffic Flow Analysis: Graph theory can be used to model and analyze traffic flow in transportation networks. By representing roads or intersections as vertices and traffic flows as edges, graph theory can help understand traffic patterns, congestion, and optimize traffic flow in transportation networks.

Vehicle Routing Problems:

Graph theory can be used to solve vehicle routing problems, where a fleet of vehicles needs to visit a set of locations with specific constraints, such as capacity limitations, time windows, or multiple depots. Graph theory algorithms, such as the traveling salesman problem (TSP) or the vehicle routing problem (VRP), can be used to find optimal or near-optimal solutions to these problems, leading to more efficient transportation operations.

2.2 Remarks

Graph theory is a powerful mathematical tool for modeling, analyzing, and optimizing transportation problems. By representing transportation networks as graphs and using graph theory

Part B

Successive service provision

Chapter 3

Bus Stop Location

3.1 Introduction

Public transportation systems, integral to urban infrastructure, profoundly impact the daily commute of millions. The strategic placement of bus stops within these systems is crucial, as it directly influences accessibility, user satisfaction, and the overall efficiency of the transit service. This chapter explores the various aspects of bus stop location in public transportation operations, with a focus on optimizing accessibility, operational efficiency, and coverage.

3.1.1 Importance of Bus Stop Location

Bus stop location is a vital aspect of public transportation planning. The decisions made in locating bus stops affect a wide range of factors including:

- Operational costs and efficiency, as the placement of stops impacts route timings and service frequency.

- Accessibility for passengers, which is pivotal for a transit system's utility and attractiveness.

- User satisfaction, which is influenced by factors such as walking distance to stops, waiting times, and overall convenience.

This chapter aims to provide a comprehensive guide to the strategic planning of bus stop locations, incorporating current methodologies and practical case studies to illustrate key concepts.

3.2 Location Set Covering Problem (LSCP)

3.2.1 Framework and Assumptions

The Location Set Covering Problem (LSCP) is a well-known model in operations research, specifically in the field of facility location and public transportation planning. It seeks to identify the minimum number of facilities (in this context, bus stops) required to ensure that each demand point (e.g., residential area, commercial zone) is within a predefined distance from a facility. The LSCP is grounded on several key assumptions:

- Each demand point must be within a certain distance, known as the coverage radius, from at least one bus stop.
- Bus stops can only be placed at candidate locations, which are predetermined based on urban planning constraints and feasibility studies.
- The objective is to minimize the total number of bus stops needed to cover all demand points adequately.

This problem is particularly relevant in urban and suburban areas where public transportation authorities aim to provide comprehensive coverage while optimizing resource allocation.

3.2.2 Mathematical Formulation

Formally, the LSCP is defined over a set of demand points, I, and a set of potential stop locations, J. Each demand point, $i \in I$, must be within a maximum distance, r, from at least one stop $j \in J$ to be considered covered.

- **Set:**
  - Let $I = \{i_1, i_2, ..., i_m\}$ be the set of demand points.
  - Let $J = \{j_1, j_2, ..., j_n\}$ be the set of potential stop locations.
- **Parameters:**
  - Let d_{ij} be the distance from demand point i to stop location j.
  - Let a_{ij} be a binary parameter where $a_{ij} = 1$ if $d_{ij} \leq r$, and 0 otherwise.
- **Variables:**
  - Let x_j be a binary decision variable where $x_j = 1$ if a stop is placed at location j, and 0 otherwise.

The LSCP seeks the minimum number of stops needed to ensure that every demand point in the service area is covered. Therefore the objective function and constraints can then be expressed as follows:

$$\begin{aligned}
\text{Minimize} \quad & G = \sum_{j \in J} x_j \\
\text{subject to} \quad & \sum_{j \in J} a_{ij} x_j \geq 1 \quad \forall i \in I \\
& x_j \in \{0, 1\} \quad \forall j \in J
\end{aligned}$$

The constraints ensure that each demand point $i \in I$ is covered by at least one stop $j \in J$ where $a_{ij} = 1$.

3.2.3 Preprocessing and Reduction Rules (Option)

Prior to solving the LSCP, it is essential to determine whether a stop j can cover a demand point i based on a maximum distance threshold, $r \in \mathbb{R}_{\geq 0}$. This involves solving shortest path problems to find the distances between all pairs (i, j) and setting $a_{ij} = 1$ when the distance is within the threshold.

Furthermore, several reduction rules can streamline the problem:

- If a demand point i is exclusively covered by a single stop location j ($a_{ij} = 1$ and $a_{ik} = 0$ for all other $k \neq j$), then j is an essential stop, and we can set $x_j = 1$.
- If stop location j provides coverage equal to or greater than another stop k for all demand points, then j dominates k, and k can be removed.
- If covering a demand point i ensures coverage for another demand point s, then i dominates s, and s can be removed from consideration.

These rules reduce the number of variables and constraints, simplifying the LSCP without losing the integrity of the solution.

3.2.4 Example Application

An example scenario with nine demand points and potential stop locations illustrates the application of LSCP. Optimization solver, such as GAMS and Gurobi, can be utilized to find the optimal set of locations for public transport stops.

Consider a network with demand points and potential public transport stop locations. Distances between demand points i and potential stops j are given by d_{ij}.

If $d_{ij} \leq r$, where r is the maximum coverage radius, then $a_{ij} = 1$, indicating that stop j can service demand point i.

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$	$i=9$
$j=1$	714.56	651.04	514.37	327.79	160.49	240.45	434.40	597.73	695.84
$j=2$	419.66	403.21	345.37	258.03	170.13	155.39	233.08	324.95	391.87
$j=3$	300.00	297.16	288.66	277.32	268.81	267.67	273.91	285.26	294.90

Table 3.1: Distance matrix d_{ij} between demand points i and public transport stops j.

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$	$i=9$
$j=1$	0	0	0	0	1	1	0	0	0
$j=2$	0	0	0	1	1	1	1	0	0
$j=3$	1	1	1	1	1	1	1	1	1

Table 3.2: Binary matrix a_{ij} indicating if a public transport stop j can service demand point i within the distance limit.

By utilizing the GAMS code for the LSCP model provided in the appendix, the optimal solution yielded the following results:

```
MIP Solution:            1.000000
Final Solve:             1.000000

Best possible:           1.000000
Absolute gap:            0.000000
Relative gap:            0.000000

----     VARIABLE X.L

3 1.000


----     VARIABLE z.L                    =        1.000
```

The results indicate that the model has found an optimal solution with an objective value of 1. The variable ‘X.L‘ shows that stop 3 is to be opened to achieve this solution. The objective function value, represented by 'z.L', is equal to 1, confirming the number of stops opened in this optimal configuration. The zero absolute and relative gaps confirm that the solution is indeed optimal.

3.3 Maximal Covering Location Problem (MCLP)

3.3.1 Importance of the MCLP in Public Transport Planning

The Maximal Covering Location Problem (MCLP) is a pivotal model in the field of public transportation planning. It addresses the need to maximize service coverage within a restricted number of facilities. This problem becomes critical when public transport authorities face budget constraints but aim to ensure that the majority of the population is within a reasonable distance of a transit stop. The MCLP is an evolution of the Location Set Covering Problem (LSCP), which focuses on covering all demand points without the limitation on the number of stops.

3.3.2 Framework and Assumptions

The MCLP operates under a set of assumptions that guide the decision-making process for optimal stop placement:

- A fixed number of stops (p) can be established within the network.
- The goal is to maximize the coverage of demand within a service distance (l) under the constraint that every demand point must also be within a maximum distance (r) from a stop.
- Each demand point has an associated weight (c_i), usually representing the potential or existing passenger demand.

3.3.3 Mathematical Formulation

Formally, the MCLP is defined over a set of demand points, I, and a set of potential stop locations, J. It aims to maximize the number of demand points within the service distance l, ensuring that each is also within a mandatory coverage distance r, where l is less than r. Here, l represents the optimal service distance within which a demand point is ideally covered by a public transport stop for maximum efficiency, while r provides a broader mandatory coverage limit to ensure accessibility, reflecting a compromise between ideal coverage and resource constraints.

- **Sets:**

- Let $I = \{i_1, i_2, ..., i_m\}$ denote the set of demand points.
- Let $J = \{j_1, j_2, ..., j_n\}$ represent the set of potential stop locations.

- **Parameters:**
  - Let p be the fixed number of public transport stops to be opened.
  - Let c_i denote the number of potential passengers at demand point $i \in I$.
  - Let d_{ij} represent the distance between demand point i and stop location j.
  - Let a_{ij} be a binary parameter, where $a_{ij} = 1$ if $d_{ij} \leq r$, indicating that stop j can service demand point i within the mandatory coverage distance r, and 0 otherwise.
  - Let β_{ij} be a binary parameter, where $\beta_{ij} = 1$ if $d_{ij} \leq l$, indicating that stop j is within the service distance l for demand point i, and 0 otherwise.

- **Decision Variables:**
  - Let X_j be a binary decision variable where $X_j = 1$ if a stop is placed at location j, and 0 otherwise.
  - Let Y_i be a binary decision variable where $Y_i = 1$ if demand point i is covered by at least one public transport stop, and 0 otherwise.

The objective of the MCLP is to maximize the total covered demand. The mathematical model can be formulated as follows:

$$
\begin{aligned}
\text{Maximize} \quad & G = \sum_{i \in I} c_i \cdot Y_i \\
\text{subject to} \quad & \sum_{j \in J | a_{ij} > 0} X_j \geq 1 \quad \forall i \in I, \\
& \sum_{j \in J | \beta_{ij} > 0} X_j \geq Y_i \quad \forall i \in I, \\
& \sum_{j \in J} X_j = p, \\
& X_j \in \{0, 1\} \quad \forall j \in J, \\
& Y_i \in \{0, 1\} \quad \forall i \in I.
\end{aligned}
$$

This model ensures that a fixed number of p public transport stops are selected to maximize the coverage of demand points, considering both a mandatory coverage distance r and a service distance l, with the constraint that $l < r$. The constraints $\sum_{j \in J | a_{ij} > 0} X_j \geq 1$ guarantee that each demand point is within the mandatory coverage distance r from at least one selected public transport stop. In parallel, the constraints $\sum_{j \in J | \beta_{ij} > 0} X_j \geq Y_i$ ensure that as many demand points as possible are also within the more desirable service distance l. The objective function $\sum_{i \in I} c_i \cdot Y_i$ maximizes the total covered demand, effectively balancing the ideal coverage against the available resources.

3.3.4 Example Application

An example scenario with nine demand points and five potential stop locations illustrates the application of the MCLP. Optimization solvers, such as GAMS or Gurobi, can be utilized to find the optimal set of locations for public transport stops.

Consider a network with demand points and potential public transport stop locations. Distances between demand points i and potential stops j are given by d_{ij}. If $d_{ij} \leq r$, where r is the maximum coverage radius, then $a_{ij} = 1$, indicating that stop j can service demand point i.

Table 3.3: Distance matrix d_{ij} between demand points i and public transport stops j.

	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$	$i = 6$	$i = 7$	$i = 8$	$i = 9$
$j = 1$	714.56	651.04	514.37	327.79	160.49	240.45	434.40	597.73	695.84
$j = 2$	419.66	403.21	345.37	258.03	170.13	155.39	233.08	324.95	391.87
$j = 3$	300.00	297.16	288.66	277.32	268.81	267.67	273.91	285.26	294.90

The binary matrix a_{ij} represents whether a public transport stop j is within the mandatory coverage distance r to service demand point i. Similarly, β_{ij} indicates if a stop j is within the service distance l.

Additionally, the demand c_i at each demand point i is as follows:

Table 3.4: Demand c_i at each demand point i.

$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$	$i = 6$	$i = 7$	$i = 8$	$i = 9$
757	1765	1326	952	938	836	1025	1785	600

Similarly, if $d_{ij} \leq l$, where l is the service distance, then $\beta_{ij} = 1$, indicating a more desirable service range for efficiency. In this example, $l = 250$ meters.

Table 3.5: Binary matrix a_{ij} indicating if a public transport stop j can service demand point i within the distance limit r.

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$	$i=9$
$j=1$	0	0	0	1	1	1	0	0	0
$j=2$	0	1	1	1	1	1	1	1	0
$j=3$	1	0	0	0	1	1	1	1	1
$j=4$	1	1	1	1	0	0	0	0	1
$j=5$	1	0	0	0	0	0	0	1	1

Table 3.6: Binary matrix β_{ij} indicating if a public transport stop j can service demand point i within the service distance l.

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$	$i=9$
$j=1$	0	0	0	0	1	1	0	0	0
$j=2$	0	0	1	1	1	1	1	0	0
$j=3$	0	0	0	0	0	1	1	1	1
$j=4$	1	1	1	0	0	0	0	0	1
$j=5$	1	0	0	0	0	0	0	1	1

By utilizing the GAMS code for the MCLP model provided in the appendix, the optimal solution yielded the following results:

```
MIP   Solution:        8219.000000
Final Solve:           8219.000000

Best possible:         8219.000000
Absolute gap:             0.000000
Relative gap:             0.000000

---- VARIABLE x.L

2 1.000,    5 1.000

----VARIABLE Y.L

1 1.000, 3 1.000, 4 1.000, 5 1.000, 6 1.000, 7 1.000, 8 1.000, 9 1.000
```

The results indicate that the model has found an optimal solution with an objective value of 8219. The variables ‘x.L‘ show that stops 2 and 5 are to be opened to achieve this solution, and the ‘Y.L‘ values confirm that demand points 1, 3, 4, 5, 6, 7, 8, and 9 are covered, which corresponds to all demand points except for point 2. The zero absolute and relative gaps confirm that the solution is indeed optimal.

This GAMS formulation represents a binary linear program that addresses the MCLP by defining decision variables for the selection and coverage of potential public transport stops. The model maximizes the total demand coverage subject to the constraints of service and coverage distances. The solution to this model will provide the optimal placement of transport stops within the defined parameters.

3.4 Case Study and Network Extension

3.4.1 Introduction to Case Study

This case study explores the theoretical and practical applications of network optimization models in public transport systems. It aims to address the critical aspects of extending public transport networks to meet passenger demand efficiently and effectively. The study focuses on minimizing the distance or travel time between passenger demand points and public transport stops, considering both existing and potential new stops.

3.4.2 Assumptions

The foundation of our case study rests on several key assumptions to ensure the feasibility and relevance of the model:

- **Feasibility**: The number of public transport stops to be opened, denoted as p, is selected based on a comprehensive analysis to ensure that the proposed extension is feasible under current urban and financial constraints.

- **Service Distance** (l): This critical parameter defines the maximum allowable distance between a passenger demand point and the nearest public transport stop for the area to be considered adequately serviced. It reflects the balance between operational costs and service accessibility.

- **Strategic Expansion**: The assumption involves not just adding more stations but strategically placing them to maximize coverage and efficiency, taking into account urban planning constraints and future growth areas.

3.4.3 Network Extension: Minimizing Distance or Travel Time

Remark:

- The network is viewed as a discrete set of points, comprising passenger demand points and both existing and potential new public transport stops.
- The model allows for a predefined number of new stops, aiming to ensure that all demand points fall within a serviceable distance, thus improving the overall accessibility of the public transport system.
- Euclidean distances are used to measure the space between nodes. The serviceable distance criterion is employed to determine whether a demand point is considered covered by the network.
- The primary goal is to minimize the total distance or travel time that passengers must cover to access the public transport system, thereby enhancing the utility and attractiveness of public transport.
- Constraints included in the model ensure that every demand point is served, a fixed number of new stops are introduced, and binary decision variables dictate the opening of new stops and coverage of demand points.

3.4.4 Mathematical Model

The mathematical model provides a structured approach to identify optimal locations for new public transport stops. It includes definitions of sets, parameters, and variables critical to the model's formulation:

Sets:

I : set of passenger demand points.

$\mathcal{J}$: set of public transport stops.

J : set of candidate public transport stops.

J' : set of current public transport stops $J' \subset \mathcal{J}$.

$\mathcal{I}$: set of locations within the network, $\mathcal{I} = I \cup \mathcal{J}$. v, w

Parameters:

p : Number of public transport stops to open. $p \leq |J|$

d_{ij} : Distance between node $v \in \mathcal{I}$ and potential public transport stop $w \in \mathcal{I}$.

β_{ij} : 1 if public transport stop j is within the service distance l and is capable of servicing demand point i (i.e., $\beta_{ij} = 1$ when $d_{ij} \leq l$). 0 otherwise.

Variables:

X_j = Binary variable. 1 if candidate public transport stop $j \in J'$ is opened. 0 otherwise.

Y_{ij} = Binary variable. 1 if demand point $i \in I$ is covered by public transport station $j \in J'$. 0 otherwise.

The objective of this problem is to minimize the total distance coverage by optimally placing new stops. The mathematical model can be formulated as follows:

$$\text{Minimize} = \sum_{v \in \mathcal{I}} \sum_{j \in J} d_{vj} \cdot X_j + \sum_{i \in I} \sum_{j \in J} d_{ij} \cdot Y_{ij}$$

Subject to:

$$\sum_{j \in J} \beta_{ij} \cdot Y_{ij} = 1 \qquad \forall i \in I$$

$$\sum_{j \in J | j \notin J'} X_j = P$$

$$Y_{ij} - X_j \leq 0 \qquad \forall i \in I, j \in J$$

$$X_j \in \{0, 1\} \qquad \forall j \in J \mid j \notin J'$$

$$Y_{ij} \in \{0, 1\} \qquad \forall i \in I, j \in J$$

3.5 Literature

- Gkiotsalitis, K. (2022). Strategic Planning of Public Transport Services. In: Public Transport Optimization. Springer, Cham. https://doi.org/10.1007/978-3-031-12444-0**?**

- Schöbel, A. (2007). Optimization in public transportation: stop location, delay management and tariff zone design in a public transportation network (Vol. 3). Springer. https://doi.org/10.1007/978-0-387-36643-2**?**

Chapter 4

Network Design

4.1 Introduction

Framework conditions in practice:

- In most cases, a previously developed infrastructure already exists. a complete redesign of a route network is rather uncommon and therefore the question of a suitable route network extension or reduction is the rule.
- Renewal measures are of particular importance: For safety reasons, deteriorated sections of tracks require that vehicles are only allowed to travel at reduced speeds in such areas. By acceleration measures, the vehicles and the corresponding personnel are available for a return trip sooner, so savings in the operational (personnel, vehicles) area become possible.
- Typically, projects and measures (decisions) are evaluated and then implemented or rejected depending on political expectations. Political expectations influence the price of a decision. Thus a decision can be quantified by comparison with a model-based solution.
- Depending on the needs of customers, decision-makers can make a distinction between long-distance, regional, and local transport networks. The networks are usually planned separately, with regional transport generally aligned with long-distance and local and short-distance traffic to regional traffic.

4.2 Minimal spanning tree

Applications

- Construction of a network with minimum total costs for n locations such that all locations are directly or indirectly connected, with branching points of the network being situated only of locations.

Graphical illustration

- For each location we consider a node and for each possible direct connection an edge is considered weighted by construction costs or operating costs.

Procedure

- Determination of a minimal spanning tree using Kruskal's algorithm.

Minimal spanning tree

- V vertices or nodes
- E edges
- c cost (weight) of an edge
- Let $G = [V, E, c]$ be a contiguous (connected), weighted, undirected graph.
- Tree T^* from G is a minimal spanning tree of G, if there is no other spanning tree T' of G with smaller total cost (=sum of related edge cost).

Kruskal algorithm

Prerequisite: a weighted, contiguous, loop-free, undirected graph $G = [V, E, c]$ with n nodes and m edges; with set $\bar{E}$ edge quantity to be determined with minimal spanning tree $T = [V, \bar{E}]$.

Start: sort or number the edges k_i of G in order $k_1, k_2, \ldots, k_m$ after not decreasing weights $c(k_i)$, so that holds: $c(k_1) \leq c(k_1) \leq \ldots \leq c(k_m)$.

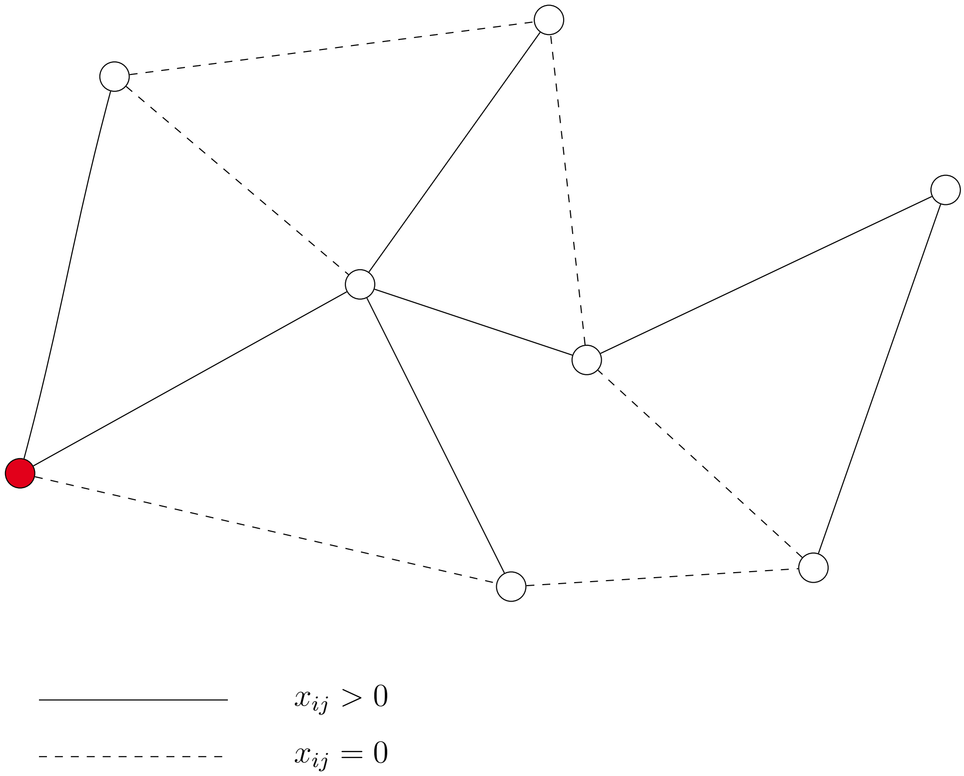
Set $\bar{E} := \emptyset$ and $T := [V, \bar{E}]$. **Iteration** $\mu = 1, 2, \ldots, m$**:**

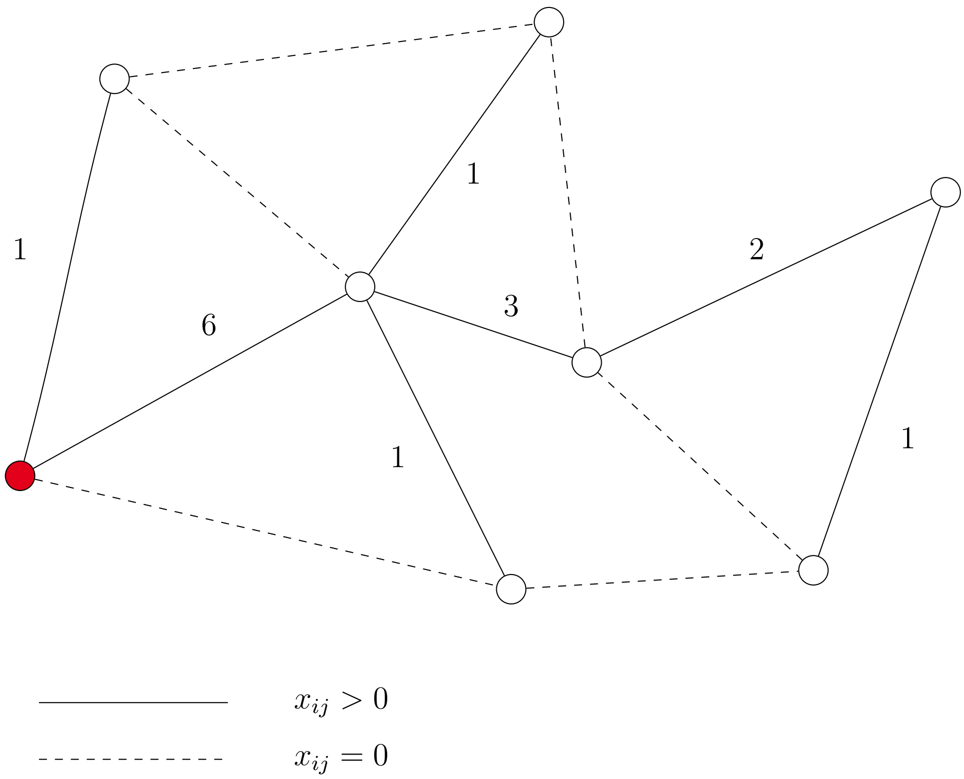
- choose edge k_μ and check if its inclusion in $T = [V, \bar{E}]$ creates a circle.
- if k_μ does not create a circle then set $\bar{E} := \bar{E} \cup \{k_\mu\}$.
- next iteration.

Termination: the procedure terminates as soon as $\bar{E}$ contains $n - 1$ edges.
Result: $T = [V, \bar{E}]$ is a minimal spanning tree of G.

Example 4.1 For an air traffic company, a route network that connects six cities, must be established. The possible flight connections and the associated costs can be seen in the following illustration:

Determine the minimal cost route network.


$x_{ij} > 0$
$x_{ij} = 0$


1
6
1
3
1
2
1
$x_{ij} > 0$
$x_{ij} = 0$

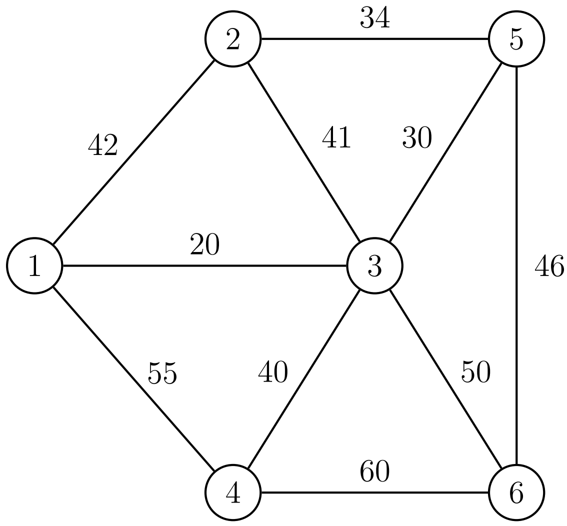


Figure 4.1: Numerical example

4.3 Network design problem

Assumptions

- Railway network for high-speed trains
- Investment costs for direct connection of two cities
- Minimization of travel times (or related cost)
- Graphical illustration
  - Nodes: cities, stations or stops
  - Edges: candidate connections
  - Edge weight: (periodized) fixed costs, (value of) travel time

Sets

V : sorted nodes (indices: i, j, u, v)

E : edges; with edge $[i, j] \in E$, where $i < j$ holds

Parameters

d_{uv} : expected number of passengers traveling from node u to node v, with $v \neq u$.
When $v = u$, d_{uv} represents the expected number of passengers leaving node $u = v$

t_{ij} : value of travel time for the direct trip from node i to node j with $[i,j] \in E$ and $t_{ij} = t_{ji}$

c_{ij} : fixed cost per period for the construction of an edge that directly connects the nodes i and j

B : maximum total cost allowed, where $B \geq F^{\star}{}_{1}$

Variables

Y_{ij} = 1, if edge $[i,j] \in E$ is built (0, otherwise)

X_{uij} = number of passengers starting their trip in node u and traveling from node i to node j with $[i,j] \in E$

Mathematical model

$$\text{Minimize } G = \sum_{[i,j]\in E} t_{ij} \cdot \sum_{u \in V} X_{uij} \tag{4.1}$$

such that

$$\sum_{[i,v]\in E} X_{uiv} - \sum_{[v,j]\in E} X_{uvj} = d_{uv} \qquad \forall u,v \in V \mid u \neq v \tag{4.2}$$

$$X_{uij} + X_{uji} \leq d_{uu} \cdot Y_{ij} \qquad \forall u \in V, [i,j] \in E \mid i < j \tag{4.3}$$

$$\sum_{[i,j]\in E \mid i<j} c_{ij} \cdot Y_{ij} \leq B \tag{4.4}$$

$$X_{uij} \geq 0 \qquad \forall u \in V, [i,j] \in E \tag{4.5}$$

$$Y_{ij} \in \{0,1\} \qquad \forall [i,j] \in E \mid i < j \tag{4.6}$$

$$\tag{4.7}$$

Remarks

- Objective function (4.1) minimizes the periodic infrastructure and travel time-induced opportunity costs.
- Constraints (4.2) ensure conservation of flow between departing and arriving passengers at each station.
- Constraints (4.3) ensures that a transport from i to j or j to i is only possible if the connection (edge) $[i,j]$ is established, i.e. such that $y_{ij} = 1$.

- Constraint (4.4) ensure the cost of the connections to built cannot exceed the available budget.
- Constraints (4.5) and (4.6) defines the domain of the variables.

Model critique

- The demand is independent of the actual route network. In general, however, it can be assumed that passengers will be displaced by motorised private transportation if the travel time by PT is significantly longer than the shortest travel time. The effect of passengers migrating will only be recorded as a lump sum by the opportunity costs.
- The maximum number of trips along an edge per time unit (TU) is limited. Therefore, edge-related capacity restrictions must be taken into account. It is conceivable to consider different expansion stages (single or double track).

Example 4.2 Five cities are to be connected by a high-speed rail network. The construction costs f_{ij} and the durations c_{ij} transformed into opportunity costs are shown in Figure 4.2.

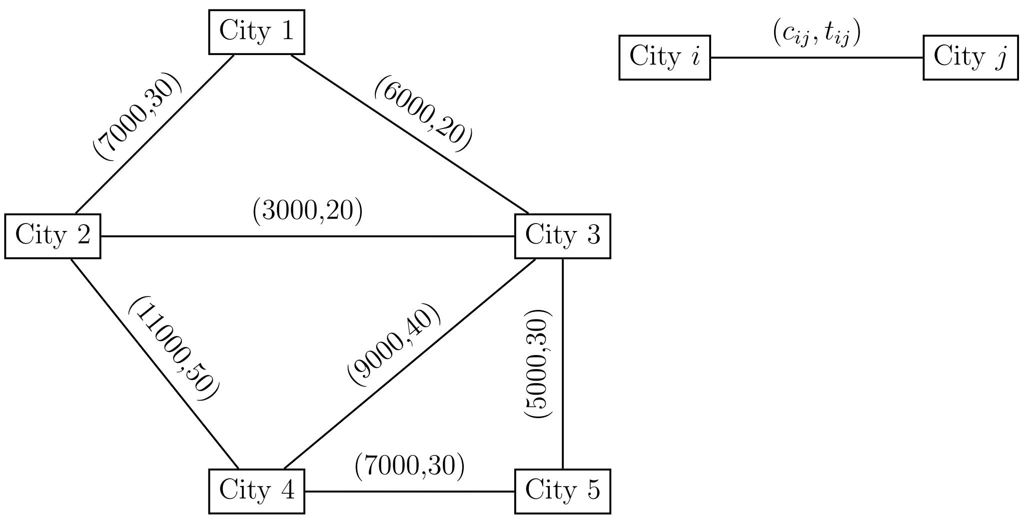


Figure 4.2: Example: Network design problem

A demand analysis shows the following OD matrix. Where, $u, v \in$ {City1, City2, City3, City4, City5 }. Calculation of d_{uu} when $v = u$ is given by:
$d_{uu} = \sum_{v \in V | u \neq v} d_{uv}, \forall u \in V$

$$(d_{uv}) = \begin{pmatrix} 900 & 250 & 400 & 100 & 150 \\ 250 & 400 & 500 & 250 & 400 \\ 400 & 500 & 1300 & 300 & 100 \\ 100 & 250 & 300 & 800 & 150 \\ 150 & 400 & 100 & 150 & 800 \end{pmatrix}.$$

Optimal solutions We solve the problem with GAMS/Cplex (see Appendix A.1) and get the following result:

```
Fixed costs         :  23000

Connections to built
---------------------------------
from city1 to city3   1
from city2 to city3   1
from city3 to city4   1
from city3 to city5   1

Connections loads
-----------------
from city1 to city3 900
from city3 to city1 900
from city2 to city3 1400
from city3 to city2 1400
from city3 to city4 650
from city4 to city3 650
from city3 to city5 650
from city5 to city3 650
from city4 to city5 150
from city5 to city4 150

Origin-related traffic flows
---------------------------------
from city1 from city1 to city3 900
from city1 from city3 to city2 250
from city1 from city3 to city4 100
from city1 from city3 to city5 150
from city2 from city2 to city3 1400
from city2 from city3 to city1  250
from city2 from city3 to city4  250
```

```
from city2 from city3 to city5  400
from city3 from city3 to city1  400
from city3 from city3 to city2  500
from city3 from city3 to city4  300
from city3 from city3 to city5  100
from city4 from city3 to city1  100
from city4 from city3 to city2  250
from city4 from city4 to city3  650
from city4 from city4 to city5  150
from city5 from city3 to city1  150
from city5 from city3 to city2 400
from city5 from city5 to city3 650
from city5 from city5 to city4 150
```

Traffic flows starting in City 1 ()$u = 1$):

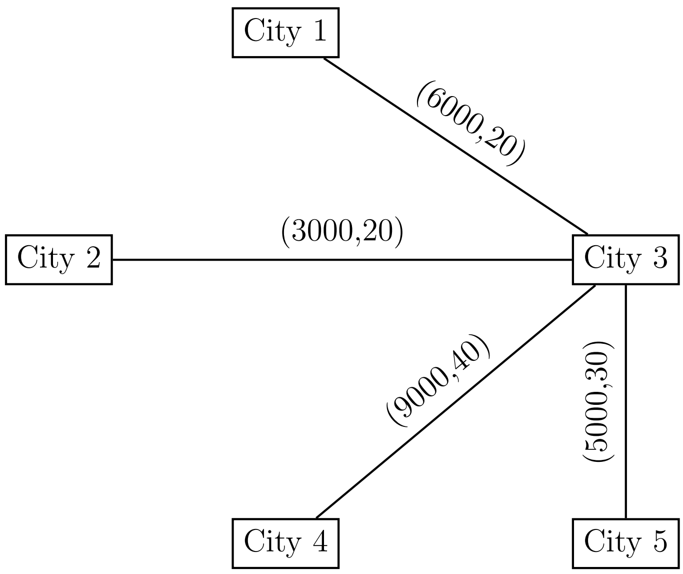


Figure 4.3: Example: Optimal solution

In node $v = 3$, 900 passengers arrive, of which 500 continue. So 400 passengers get off, which is exactly $od_{13} = 400$.

4.4 Network extensions

Assumptions

- Initial conceptual situation: we consider an existing tram network. To improve the public transport network, extension investments are expected to be carried out. An annual budget is available for this purpose. Possible extension investments are already known.
- Each passenger-kilometer provides a benefit (additional revenues, environment, safety, reduced operating costs). Therefore, the network extensions should be made in such a way as to maximize the sum of passenger-kilometers gained.
- Investment funds may be forwarded to the next year.

Sets

I : extension investment, candidate projects (index: i)

T : periods (years) of the planning period (index: t)

Parameters

d_i : duration of the construction period of extension i

k_{ij} : investment provided for extension i in year j ($j = 1, \ldots, d_i$)

B_t : budget in period t

c_{it} : utility of extension i at start of construction in period t

Variables

Y_{ij} = 1, if edge $[i, j] \in E$ is built (0, otherwise)

X_{uij} = number of passengers starting their trip in node u and traveling from node i to node j with $[i, j] \in E$

Mathematical model

$$\text{Maximize } F = \sum_{i \in I} \sum_{t=1}^{|T|-d_i+1} c_{it} y_{it} \tag{4.8}$$

such that

$$\sum_{t \in T} y_{it} \leq 1 \qquad i \in I \tag{4.9}$$

$$u_t + \sum_{i \in I} \sum_{j=1}^{d_i} k_{ij} y_{i,t-j+1} - u_{t-1} = B_t \qquad t \in T \tag{4.10}$$

$$u_t \geq 0 \qquad t \in T \tag{4.11}$$

$$y_{it} \in \{0, 1\} \qquad i \in I, t \in T \tag{4.12}$$

Remarks

- Equations (4.8) maximizes total utility during the planning period
- Constraints (4.9) ensure each extension can be made at most once.
- Constraints (4.10) ensure the total investments of the extensions decided in each period must not exceed the available budget for that period.
- Constraints (4.11) and (4.12) define the domain of the variables

Example 4.3 The following network extensions have to be scheduled for the tram network of the city of Dresden:

1. SCHÖNFELD/WEISSIG: Ullersdorfer Platz (Bühlau) – Lomnitzer Str. – Possendorfer Str. – Taubenberg – Heinrich-Lange-Str. – Heidestr. – Gasthof Weißig – Radeberger Str. – Zum Hutbergblick – Prießnitzaue (Weißig)
2. JOHANNSTADT: Güntzplatz – Böhnischplatz – Arnoldstr. – Hertelstr. – Fetscherstr.
3. STREHLEN-PROHLIS-NICKERN: H.-Bürkner-Str. – Teplitzer Str. – Dohnaer Str. – Spitzweg – Fritz-Bursch-Str. – M.-Wittig-Str. – Tornaer Str. – Gamigstr. – Fritz-Meinhardt-Str. – Langer Weg – Mechelisstr.
4. POSTPLATZ-UNIVERSITÄT: Webergasse – Josephienstr. – Budapester Str. – Schweizer Str. – Arbeitsamt – Chemnitzer Str. – Nürnberger Str. (Nürnberger Ei)
5. ALTSTADT-JOHANNSTADT: Synagoge – Gerichtsstr. – St. Benno Gymn. – Hans-Grundig-Str. – Stephanienstr. – Fetscherplatz

From studies such as System repräsentativer Verkehrsbefragungen 1998, the following socio-economic data can be derived :

Group		People	PT-users	
Students	12-16 years	46,550	29.20%	13,593
	12-18 years	25,400	37.10%	9,423
Professionals	with car	225,100	6.15%	13,837
	without car	120,000	40.32%	48,387
Non-professionals	with car	79,100	13,84%	10,950
	without car	111,900	34.39%	38,481
Pensioners	with car	141,300	33.50%	47,336
	without car	31,600	12.00%	3,792

Each route segment includes several stops, with each stop having a service area. In combination with socio-demographic data of the service areas, the potential passengers of the route network extensions were forecast. The results and further information on the investment opportunities are shown in the following table:

Candidate	Length (km)	Potential passengers	Length per trip (km)	Construction time (years)
1: Schönfeld/Weißig	3.30 km	2,020	2.5	3
2: Johannstadt	1.50 km	3,571	2.0	1
3: Strehlen-Prohlis-Nickern	4.25 km	4,481	3.1	4
4: Postplatz-Universität	2.45 km	4,400	2.0	2
5: Altstadt-Johannstadt	2.00 km	3,887	2.0	2

The required investment funds per km are estimated at 3 million euros. During the construction phase, the following financial resources [in million euros] are to be made available for the expansion measures:

Candidate	1^{st} year	2^{nd} year	3^{rd} year	4^{th} year
1: Schönfeld/Weißig	4.00	3.50	2.40	-
2: Johannstadt	4.50	-	-	-
3: Strehlen-Prohlis-Nickern	3.50	3.00	3.25	3.00
4: Postplatz-Universität	4.00	3.35	-	-
5: Altstadt-Johannstadt	3.50	2.50	-	-

The benefit per km is identical for all passengers. On average, a Dresden resident takes the OPNV 249 times a year. The transport company has an annual budget of 6 million euros for expansion investments. The planning period covers the years 2006 to 2013. Using the data provided, GAMS/Cplex provides the following optimal solution:

```
MIP Solution:     48421087.800000
Relative gap:            0.000000

Start of construction of an extension
---------------------------
Johann      in  2023
PostUni     in  2024
AltJohann   in  2024
SchoenWeiss in  2027
StreProNick in  2026
```

```
Period surpluses
-------------------
at the end of 2023  1.50 [Mill. Euro]
at the end of 2024  0.00 [Mill. Euro]
at the end of 2025  0.15 [Mill. Euro]
at the end of 2026  2.65 [Mill. Euro]
at the end of 2027  1.65 [Mill. Euro]
at the end of 2028  0.90 [Mill. Euro]
at the end of 2029  1.50 [Mill. Euro]
at the end of 2030  7.50 [Mill. Euro]
at the end of 2031 13.50 [Mill. Euro]
```

4.5 Literature

Ackermann, K., Schöppe, E., & Bradow, A. (1999). System repraesentativer Verkehrsbefragungen 1998-Erhebungsmethode und ausgewaehlte Ergebnisse. Straßenverkehrstechnik, 43(8).

Domschke, W., & Drexl, A. (2014). Logistik: standorte. In Logistik: Standorte. Oldenbourg Wissenschaftsverlag.

4.6 Exercises

Exercise 4.1 Let us look again at the example 4.2. The decision maker wants to know how sensitive the solution is to over- and underestimating opportunity costs. Therefore, determine the solutions that will result if the opportunity costs are 10%, 20% or 30% higher or lower, and interpret the results.

Exercise 4.2 A low-cost carrier would like to serve the cities of Hamburg, Bremen, Hanover, Berlin, Leipzig, Dresden, Mannheim, Saarbrucken, Nuremberg, Stuttgart and Munich. The following data are given:

City	x-Coordinate	y-Coordinate	Population
Hamburg	12.4 cm	15.6 cm	1,734,083
Bremen	9.6 cm	14.4 cm	544,853
Hanover	11.7 cm	12.8 cm	516,160
Berlin	20.4 cm	13.1 cm	3,388,477
Leipzig	18.1 cm	10.3 cm	497,531
Dresden	21.3 cm	9.5 cm	483,632
Mannheim	8.7 cm	5.9 cm	308,353
Saarbrücken	5.2 cm	5.3 cm	181,860
Nuremberg	14.9 cm	5.7 cm	493,553
Stuttgart	10.4 cm	4.2 cm	589,161
Munich	16.0 cm	2.6 cm	1,247,873

A length of 1 cm corresponds to 60 km. All cities should be connected (by changing trains or making stopovers).

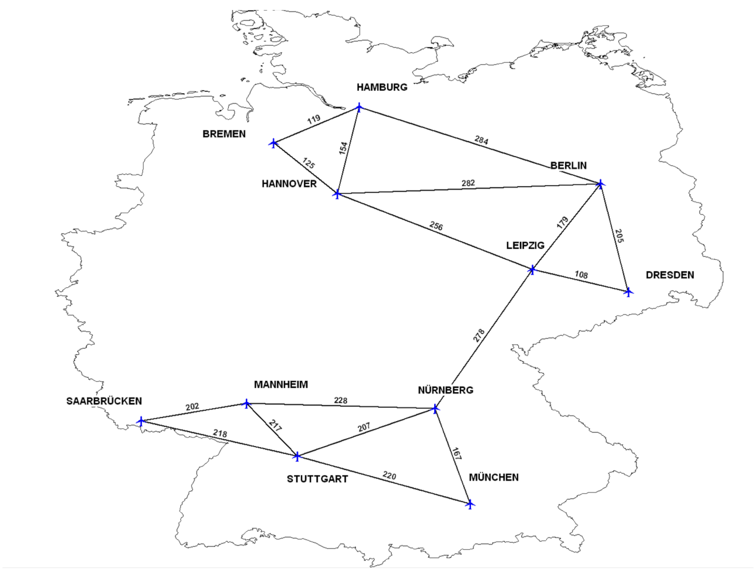


Figure 4.4: Potential connections

(a) It is expected that a resident will fly with the airline 0.05 times per year (outbound and return). The demand is distributed proportional to the number of inhabitants. Determine the OD matrix.

(b) Assume that the aircraft used must be refueled after 300 km at the latest. Determine the corresponding route network with minimum total length.

(c) Assume that the route network should ensure that a passenger makes only one stopover at most (with or without transfer). Formulate a suitable decision model for this purpose. Determine an optimal solution using GAMS.

(d) Formulate a general decision model to determine a route network with of a route network with three hubs. Each airport is to be assigned to exactly one hub and the hubs are to be connected directly.

Chapter 5

Tariff Zone Planning for Public Transport Service Providers

5.1 Introduction to Tariff Systems in Public Transportation

5.1.1 Introduction

Tariff systems are integral to public transportation, affecting everything from operational sustainability to user accessibility. This section sheds light on the intricate process of creating fare structures that reconcile the financial imperatives of service providers with the public's right to accessible transport. The focus is on tariff zone planning, a strategic approach to addressing the multifaceted challenges of fare management.

5.1.2 Importance

Designing effective tariff systems requires a nuanced understanding of economic, social, and logistical considerations. The collaboration with regional transport authorities underlines the complexities of aligning revenue objectives with the provision of equitable access. This importance section highlights the indispensable role of systematic tariff zone planning, leveraging real-world experiences to develop fare strategies that bolster public transport networks' efficiency and inclusivity.

5.2 Conceptual Framework for Tariff Systems

5.2.1 Theoretical Underpinnings

The design and implementation of tariff systems in public transportation are underpinned by various principles aimed at optimizing revenue, ensuring accessibility, and maintaining fairness for passengers. This section explores three primary models: distance tariffs, unit tariffs, and zone-based tariffs, highlighting their foundational rationale.

- **Distance Tariffs:** Fares are calculated based on the actual distance traveled, embodying the principle of proportional payment for service usage. This system is considered the most equitable, requiring accurate distance matrices for fare calculation.
- **Unit Tariffs:** A flat fare is applied to all trips, regardless of distance. This model prioritizes administrative simplicity and ease of understanding but may compromise on fare equity, especially across varying trip lengths.
- **Zone-based Tariffs:** The service area is divided into zones, with fares based on the number of zones traversed. This approach seeks to balance the granularity of distance-based pricing with the operational simplicity of unit tariffs, offering a pragmatic solution to fare structuring.

5.2.2 Comparative Analysis

A detailed comparison of the aforementioned tariff systems unveils a nuanced interplay of advantages and disadvantages, pivotal to their application in public transportation contexts.

- **Distance Tariffs:**
  - *Advantages:* High level of fare equity and direct linkage between fare and service usage.
  - *Disadvantages:* Operational challenges in maintaining distance matrices and potential passenger confusion over variable fares.
- **Unit Tariffs:**
  - *Advantages:* Operational simplicity and predictable revenue streams.
  - *Disadvantages:* Reduced fare equity, with shorter trips subsidizing longer ones, potentially discouraging short-distance public transport use.

- **Zone-based Tariffs:**
  - *Advantages:* Balances equity and simplicity, offering a more equitable structure than unit tariffs and reducing operational complexity compared to distance tariffs.
  - *Disadvantages:* Challenges in zone delineation and the potential for arbitrariness in zone boundaries.

In conclusion, each tariff system—distance tariffs, unit tariffs, and zone-based tariffs—has its distinct advantages and challenges. Distance tariffs offer fare equity but can complicate fare calculation. Unit tariffs simplify fare management but may not always fairly reflect journey lengths. Zone-based tariffs strike a balance, providing a practical compromise between fairness and operational simplicity. The choice of system depends on the transport provider's objectives, balancing financial sustainability, efficiency, and passenger equity.

5.3 Optimizing Revenue through Strategic Tariff Zone Pricing

Within the framework of public transport economics, the strategic implementation of tariff zone pricing plays a crucial role in balancing operational efficiency with revenue generation. This section encapsulates the essence of fare calculation, the intricacies of selecting pricing systems, and the methodologies for projecting expected revenues, grounded in the principles outlined in the provided beamer content.

5.3.1 Fare Calculation and Pricing Systems

The core of tariff zone pricing strategy lies in its straightforward yet effective fare calculation method. The fare for a journey is directly proportional to the number of tariff zones traversed, encapsulated by the formula:

$$\text{Fare} = t \cdot \pi_{pt},$$

where t represents the total number of zones visited from origin i to destination j, and π_{pt} indicates the price per zone under a specific pricing system p from the set of available pricing systems $\mathcal{P}$. An illustrative example is a journey spanning two tariff zones, calculated as $2 \cdot \pi_{pt}$ for $t = 2$.

To visually represent this concept, consider the following diagram illustrating a journey across two tariff zones:

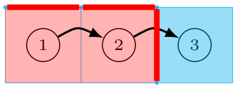


Figure 5.1: Illustration of a journey across two tariff zones.

The flexibility offered by the set of pricing systems, $\mathcal{P}$, allows for adaptive responses to varying demand patterns and operational objectives. Each system is defined by a unique π_{pt} value, enabling tailored pricing strategies that can enhance both passenger affordability and revenue predictability.

5.3.2 In-depth Aspects of Tariff Zone Planning

Tariff zone planning transcends the mere delineation of geographic areas for fare calculation; it integrates complex considerations to enhance both the operational efficiency of public transport systems and the commuting experience. This subsection explores these nuanced aspects further.

Integrated Approach to Fare and Zoning Challenges: Effective tariff zone planning necessitates a dual-focused approach, addressing:

- The fare structure, ensuring it remains equitable for users while sustainable for providers.
- The logical division of zones, which must simplify fare systems without compromising travel convenience.

Defining Districts and Tariff Zones: A clear distinction is made between districts—individual areas served by singular transport stops—and tariff zones, which are comprehensive sectors formed by connecting these districts. This distinction is crucial for:

- Streamlining fare calculations.
- Facilitating straightforward navigation for passengers.

Ensuring Contiguity and Connectivity: The planning process emphasizes the importance of contiguous tariff zones and well-connected stops, aiming to:

- Maintain an uninterrupted flow within and across zones.
- Avoid unnecessary complications in route planning for passengers.

These considerations are pivotal in crafting a tariff zone framework that not only aligns with the financial goals of transport authorities but also prioritizes the accessibility and satisfaction of passengers. By meticulously addressing each of these aspects, tariff zone planning can achieve a balance between operational demands and user needs, thereby contributing to a more efficient and user-friendly public transport system.

5.3.3 Illustrative Example: Selecting Price per Zone from a Finite Set

An essential aspect of tariff zone pricing strategy involves selecting the price per zone, π_{pt}, from a finite set of values. This approach allows for the strategic management of fare structures to address specific challenges associated with independent tariff zones. Here, we present an example to elucidate this concept further.

Consider a public transport system with a set of pricing systems, $\mathcal{P}$, wherein each system is designed to tackle distinct operational scenarios. The choice of π_{pt} within these systems directly impacts the fare calculation, adhering to the formula:

$$\text{fare} = t \cdot \pi_{pt},$$

where t signifies the number of tariff zones crossed during a journey. For demonstration purposes, let's examine two pricing systems ($p = 1$ and $p = 2$) and their respective fares across different numbers of tariff zones visited (t).

t	π_{pt} ($p = 1$)	π_{pt} ($p = 2$)
1	1.0	1.5
2	1.0	1.0
3	1.0	0.75
4	1.0	0.6

Table 5.1: Fare per tariff zone for different pricing systems.

Calculation Process for Fare Variation The fare variation across different numbers of tariff zones for the two pricing systems ($p = 1$ and $p = 2$) is calculated based on the predefined prices per zone for each system. Below is the calculation process outlined for each point plotted in the graph:

For pricing system $p = 1$, the fare for traversing t zones is calculated as follows:

- For $t = 1$, the fare is $1 \cdot 1.0 = 1.0$.
- For $t = 2$, the fare is $2 \cdot 1.0 = 2.0$.
- For $t = 3$, the fare is $3 \cdot 1.0 = 3.0$.
- For $t = 4$, the fare is $4 \cdot 1.0 = 4.0$.

For pricing system $p = 2$, a similar calculation process is followed:

- For $t = 1$, the fare is $1 \cdot 1.5 = 1.5$.
- For $t = 2$, the fare is $2 \cdot 1.0 = 2.0$.
- For $t = 3$, the fare is $3 \cdot 0.75 = 2.25$.
- For $t = 4$, the fare is $4 \cdot 0.6 = 2.4$.

These calculations form the basis for the data points plotted in the subsequent graph, visually representing the fare variations for each pricing system as the number of tariff zones increases.

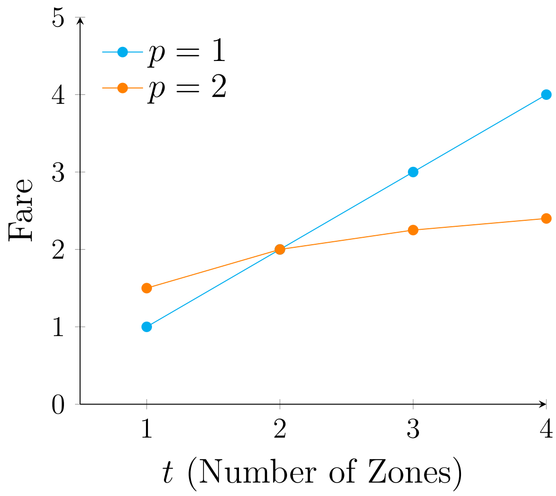


Figure 5.2: Fare variation across different numbers of zones for two pricing systems.

This example underscores the significance of choosing appropriate π_{pt} values from a predefined set, allowing transport authorities to fine-tune fare structures for optimal revenue generation and service accessibility.

5.3.4 Expected Revenues from Tariff Zones

Key Parameters in Revenue Estimation

- T_{ij}: **Maximum Number of Tariff Zones** - This parameter represents the maximum number of tariff zones encountered along the shortest path between any two stops i and j within the network. It sets a boundary for fare calculations and impacts the overall revenue estimation.

- $r_{ijt}(\pi_{pt})$: **Expected Revenue** - The expected revenue for traversing $t = 1, \ldots, T_{ij}$ tariff zones on the shortest path between stops i and j, under a given pricing system p. This calculation considers the fare applied per zone and the number of trips taken, factoring in the price sensitivity and travel preferences of passengers.

The revenue estimation is guided by the formula:

$$r_{ijt}\left(\pi_{pt}\right) = t \cdot \pi_{pt} \times f(t \cdot \pi_{pt}, \text{travel time}, \ldots),$$

where:

- $t \cdot \pi_{pt}$ represents the fare, calculated as the product of the number of tariff zones crossed (t) and the price per tariff zone (π_{pt}) for the selected pricing system (p).

- $f(t \cdot \pi_{pt}, \text{travel time}, \ldots)$ denotes a function that models the number of transit trips, incorporating variables such as the fare and travel time.

5.4 Tariff Zone Planning: Methodologies and Models

5.4.1 Mathematical Formulation of Tariff Zone Planning

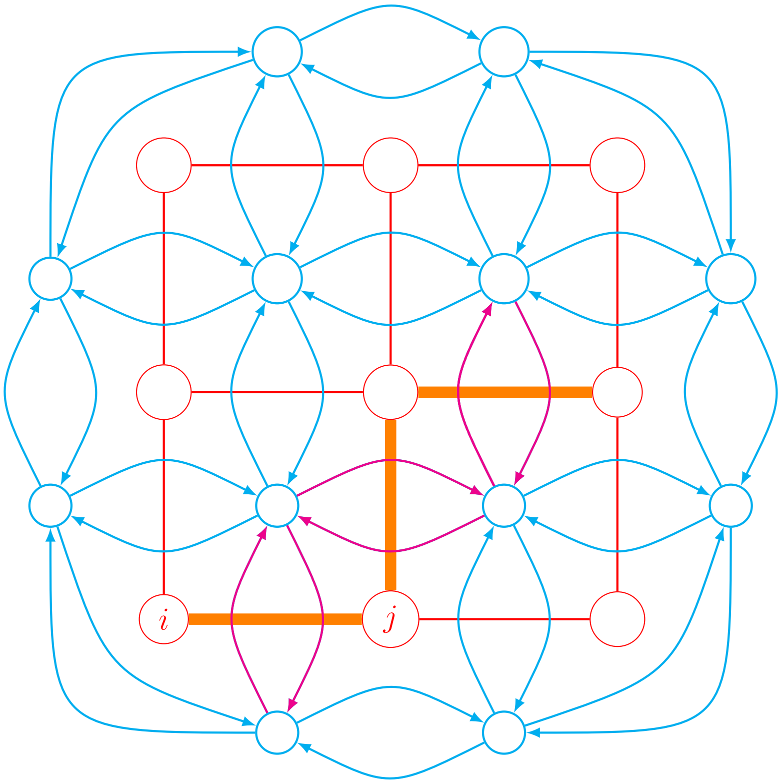
The mathematical formulation of tariff zone planning transcends basic operational considerations, encapsulating a sophisticated model that addresses the multifaceted dynamics of public transportation networks. This formulation is pivotal for strategizing optimal fare structures that not only maximize revenue but also enhance service accessibility and fairness. The development of this model is predicated on a rigorous definition of sets, parameters, decision variables, and constraints, each of which plays a vital role in delineating the optimal configuration of tariff zones.

Definition of Sets and Parameters

Sets The model begins with the definition of essential sets representing the structural components of the transportation network and the tariff zones:

- $\mathcal{I} = \{i_1, i_2, \ldots, i_n\}$: The set of all public transport stops within the network, serving as nodes for departure and arrival.
- $\mathcal{A} = \{(i_j, i_k) \mid i_j, i_k \in \mathcal{I}\}$: Directed arcs between stops, indicating possible routes.
- $\mathcal{N}$: Nodes at the periphery of districts, crucial for defining tariff zone boundaries.
- $\mathcal{B}$: Arcs connecting district border nodes, marking the physical limits of tariff zones.
- $\mathcal{D}_{ij}$: Dual arcs (n, m) and (m, n) in $\mathcal{B}$ corresponding to transport routes (k, l).
- $\mathcal{S}_{ij}$: Arcs $(k, l) \in \mathcal{A}$ forming the shortest path between i and j.

Two graphs are constructed from these sets, $\mathcal{G}^{\mathrm{PT}} : (\mathcal{I}, \mathcal{A})$, $\mathcal{G}^{\mathrm{BO}} : (\mathcal{N}, \mathcal{B})$, $\mathcal{S}_{ij}$ and $\mathcal{D}_{ij}$ representing the public transport network and the border organization, respectively. Their visual representations are as follows:



Parameters Key parameters are established to quantify the characteristics of the transportation network and passenger travel patterns:

- T_{ij}: Maximum number of tariff zones along the shortest path from $i \in \mathcal{I}$ to $j \in \mathcal{I}$.

- π_{pt}: Price per tariff zone, if t tariff zones are visited given pricing system p.

- $r_{ijt}(\pi_{pt})$: Expected revenue if $t = 1, \ldots, T_{ij}$ tariff zones are visited on the shortest path from $i \in \mathcal{I}$ to $j \in \mathcal{I}$.

- a_n: Feasible node degree of the border node $n \in \mathcal{N}$, indicating how many connections can be made to this node.

- b_n: Amount of artificial outflow or inflow at border node $n \in \mathcal{N}$, used to maintain flow balance at the network edges.

- u: Total sum of outflows over all border nodes $n \in \mathcal{N}$, calculated as $u = \sum_{n|b_n \geq 0} b_n$. This parameter helps ensure the network's flow conservation.

Decision Variables

The model introduces decision variables to capture the strategic choices regarding tariff zone configurations and fare structures:

- $X_{ijt} \in \{0, 1\}$: Indicates whether a trip from stop i to stop j that traverses t tariff zones is selected. This binary variable is essential for identifying the optimal path configuration under revenue maximization criteria.

- $Y_{nm} \in \{0, 1\}$: Determines the establishment of a tariff zone boundary along the border arc (n, m). The optimization process utilizes this variable to configure zone boundaries efficiently.

- $W_{nm} \in \{0, 1\}$: Represents the flow of passengers along the border arc (n, m), aiding in the enforcement of contiguous and logically structured tariff zones.

Objective Function and Constraints

The mathematical formulation of tariff zone planning aims to maximize the total revenue generated from fares across the transportation network, subject to operational and structural constraints. The complete model is presented below:

$$\max F(p) = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{I}} \sum_{t=1}^{T_{ij}} r_{ijt}(\pi_{pt}) \cdot X_{ijt}, \tag{5.1}$$

Subject to:

$$\sum_{t=1}^{T_{ij}} X_{ijt} = 1 \qquad \forall\ i, j \in \mathcal{I}, i < j, \tag{5.2}$$

$$\sum_{t=1}^{T_{ij}} t \cdot X_{ijt} - \sum_{(n,m) \in \mathcal{K}_{ij}} Y_{nm} = 1 \qquad \forall\ i, j \in \mathcal{I}, i < j, \tag{5.3}$$

$$\sum_{m \in \mathcal{N}} W_{nm} - \sum_{m \in \mathcal{N}} W_{mn} = b_n \qquad \forall\ n \in \mathcal{N}, \tag{5.4}$$

$$W_{nm} - W_{mn} = Y_{nm} \qquad \forall\ (n, m) \in \mathcal{B}, n < m, \tag{5.5}$$

$$\sum_{m \in \mathcal{B}} W_{nm} + \sum_{m \in \mathcal{B}} W_{mn} \leq a_n \qquad \forall\ n \in \mathcal{N}, \tag{5.6}$$

Detailed Explanation of the Model Components

- **Objective Function (5.2)**: Maximizes the total revenue, $F(p)$, by selecting paths (X_{ijt}) between stops that maximize revenue ($r_{ijt}(\pi_{pt})$) from fares across all trips in the network.

- **Constraint (5.2)**: Ensures that each trip between two stops is allocated exactly one path, thereby simplifying fare calculation and path selection for passengers.

- **Constraint (5.3)**: Aligns the number of tariff zones crossed with the tariff zone boundaries, ensuring that fare calculation accurately reflects the physical traversal of zones.

- **Constraint (5.4)**: Guarantees the conservation of passenger flows at district borders, maintaining a balance between the inflow and outflow of passengers, which is essential for network stability and efficiency.

- **Constraint (5.5)**: Links the establishment of tariff zone borders (Y_{nm}) with the flow of passengers (W_{nm}), facilitating the logical and efficient demarcation of tariff zones.

- **Constraint (5.6)**: Restricts the number of active connections at each border node to feasible levels, ensuring that the tariff zone configuration is operationally viable and adheres to physical and logistical constraints.

This comprehensive model facilitates the strategic planning of tariff zones, enabling transportation authorities to optimize fare structures for revenue maximization while ensuring operational feasibility and service quality.

5.5 Section 4: Empirical Insights and Practical Implementations

This section explores the practical application of tariff zone planning through the lens of numerical examples and discusses methodologies for deriving optimal solutions, thereby illustrating the intersection of theory and practice in public transportation planning.

Numerical Example

Consider a scenario with the following characteristics:

- Maximum number of tariff zones, $T_{ij} = 4$, for all pairs of stops i, j within the network.

- Two distinct pricing systems, $|\mathcal{P}| = 2$, illustrating the flexibility in tariff setting.
- Demand f is a function of both the tariff and travel time, capturing the dynamic nature of passenger preferences.
- Each border node ($n \in \mathcal{N}$) can have up to $a_n = 3$ connections with no flow imbalance ($b_n = 0$), simplifying network management.

In this example, the variation in demand from Stop 1 to Stop 5 under different pricing systems and tariff zones is presented, highlighting the impact of pricing decisions on travel behavior.

Demand Variation

p	Start Zone	End Zone	t	Demand f (public transport trips)
1	1	5	1	66.1997
1	1	5	2	60.8979
1	1	5	3	55.7475
2	1	5	1	65.1295
2	1	5	2	58.8170
2	1	5	3	52.7484

Price Systems The considered price systems are outlined, demonstrating the granularity of fare adjustments across different zones.

p	t	π_{pt}
1	1	1
1	2	2
1	3	3
1	4	4
2	1	1.2
2	2	2.4
2	3	3.6
2	4	4.8

By utilizing the GAMS code for the this model provided in the appendix, the optimal solution yielded the following results.

Optimal Solution The achievement of an optimal solution, with an objective function value of $F(p) = 100,649$, indicates the maximization of revenue within the selected tariff zone configuration. This solution was determined through 68 iterations, underpinning its optimality without any noted discrepancies.

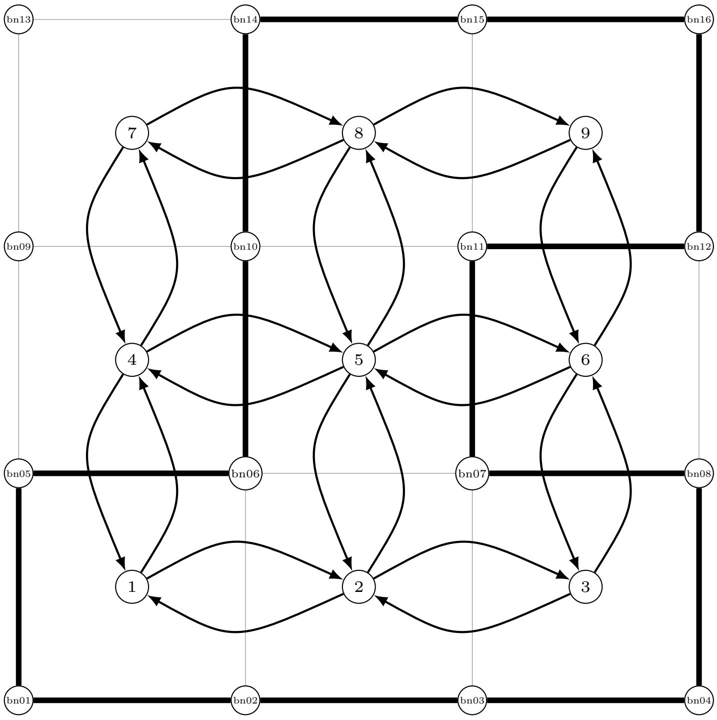
```
MIP Solution: 100649.000000 (68 iterations, 1 nodes)
Final Solve: 100649.000000 (0 iterations)
Best possible: 100649.000000
Absolute gap: 0.000000
Relative gap: 0.000000
```

Further exploration of the variable $Y.L$ delineates the boundaries of the optimal tariff zones, offering a strategic guide for the tariff system implementation.

`VARIABLE Y.L`

	bn01	bn02	bn03	bn04	bn05	bn06	bn07	bn10	bn11	bn12	bn14	bn15
bn02	1											
bn03		1										
bn04			1									
bn05	1											
bn06					1							
bn08				1			1					
bn10						1						
bn11							1					
bn12									1			
bn14								1				
bn15											1	
bn16										1		1

To visually represent these findings, a directed graph is constructed as follows:



This graphical illustration aids in understanding the optimal tariff zone configuration, facilitating informed decision-making for implementation.

5.5.1 Conclusion

The numerical examples and solution strategies presented in this section offer a practical perspective on the application of tariff zone planning models. Through rigorous analysis and optimization, transportation authorities can design fare systems that balance revenue objectives with the provision of equitable and efficient public transport services.

Chapter 6

Line Planning

6.1 Introduction

Line

- Regular service established between certain start and end points, on which passengers can get on and off at certain stations. It does not require the existence of a timetable with specific departure and arrival times or the establishment of intermediate stations
- For each regular line with a start and end station, there is another line with the opposite direction, from the end e and the start node
- Frequency of a line: number of trips per period
- The time interval between two trips of a regularly operated line is referred to as headway. It results from the reciprocal of the frequency. The frequency 'two trips per hour' provides a headway of 30 minutes.

Optimization problem

- Determine the number of lines and frequencies needed to meet passenger demand (in a given period) and optimize an objective
- A two-stage approach can be considered, first defining the lines and then determining the frequency for each line
- The headway influences expected travel times
- The travel times impact the demand
- An origin-destination matrix is an input parameter not given but is a variable to be defined

- Line planning with integrated demand is NP-hard

Objectives

- Service (Number of transfers)
- Travel time
- Costs
- Maximize expected demand

6.2 Maximization of direct trips

Objective: Maximize the number of passengers who do not have to change between lines

Assumptions:

- A uniform headway applies to all lines
- The vehicle size (in combination with the selected headway) is sufficiently large so that all passengers can be transported at all times

Sets

L	line candidates
V	sorted nodes (stops)
E	edges
L_{uv}	lines that connect the nodes (stops) $u \in V$ and $v \in V$ directly with $u < v$, i.e. passengers using the line $l \in L_{uv}$ can travel (without changing from $u \in V$ to $v \in V$ or the other way around); a line that indeed connects the nodes $u \in V$ and $v \in V$, but equally means a significant detour for passengers, does not have to be included in L_{uv}
$\tilde{L}_{ij}$	lines serving the edge $[i, j] \in E$

Paremeters

d_{uv}	demand (number) of passengers travelling from $u \in V$ to $v \in V$
k_l	operating costs of line $l \in L$
B	available budget
C_{ij}	maximal number of lines along edge $[i, j] \in E$

Variables

y_l	=1, if line l is selected (0, otherwise)
x_{uv}	=1, if a direct trip is feasible from u to v (0, otherwise)

Model

$$\max F = \sum_{u,v \in V | u<v} (d_{uv} + d_{vu}) \cdot x_{uv} \tag{6.1}$$

s.t.

$$\sum_{l \in L_{uv}} y_l \geq x_{uv} \qquad u,v \in V \mid u<v \tag{6.2}$$

$$\sum_{l \in \tilde{L}_{ij}} y_l \geq 1 \qquad [i,j] \in E \tag{6.3}$$

$$\sum_{l \in \tilde{L}_{ij}} y_l \leq C_{ij} \qquad [i,j] \in E \tag{6.4}$$

$$\sum_{l \in L} k_l y_l \leq B \tag{6.5}$$

$$y_l \in \{0,1\} \qquad l \in L \tag{6.6}$$

$$x_{uv} \in \{0,1\} \qquad u,v \in V \mid u<v \tag{6.7}$$

Remarks

- (6.14) represents the objective function, i.e. the number of direct trips must be maximized. It should be noted that each line is defined by a round trip.
- (6.15) ensures that direct trips from u to v are only included in the objective function if at least one line connects the two nodes.
- with (6.17) we avoid that more vehicles per time unit pass through an edge than is technically possible.
- compliance with the budget is ensured by (6.18).
- the permissible domain of the decision variables are defined by (6.19) and (6.20).

Construction mathematical model

Given Figure 6.1, the following is an example of the model conception:

Constraint 6.15 refer to:

$$\sum_{l \in L_{uv}} y_l \geq x_{uv} \qquad u,v \in V \mid u<v$$

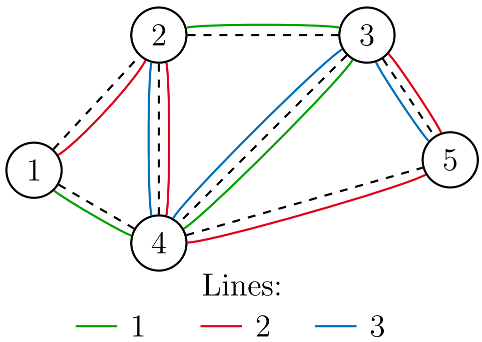


Figure 6.1: Constraints 6.15

- for $u, v = 4, 5$
- for $u, v = 3, 4$
- for $u, v = 1, 3$

Constraint 6.16 refer to:

$$\sum_{l \in \tilde{L}_{ij}} y_l \geq 1 \qquad [i, j] \in E$$

- for $i, j = 2, 4$
- for $i, j = 3, 4$
- for $i, j = 4, 5$

Constraint 6.17 refer to:

$$\sum_{l \in \tilde{L}_{ij}} y_l \leq C_{ij} \qquad [i, j] \in E$$

- for $i, j = 2, 4$
- for $i, j = 3, 4$
- for $i, j = 4, 5$

Example 6.1 Following network design graph is given:

The edge values represent the operative costs. Hence the operative line costs are the sum of the corresponding edge values. A total budget of 40000 is available. Maximally two lines may run along one edge. For a standard headway of 30 minutes, this specification corresponds to eight vehicles per hour (both directions). The number of potential direct trips between two nodes shows the following (d_{uv})

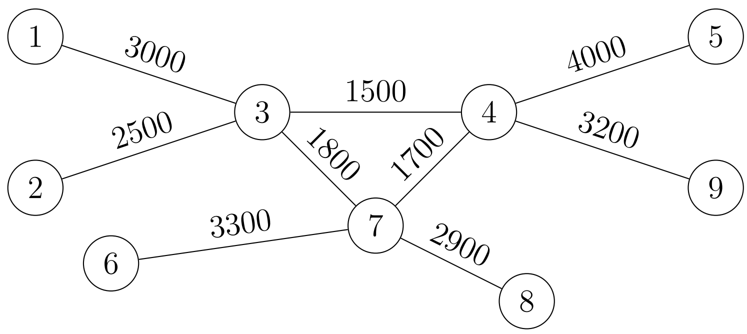


Figure 6.2: Exemplification of a public transport network

	1	2	3	4	5	6	7	8	9
1	0	5	25	40	4	3	50	6	4
2	5	0	30	60	8	9	70	7	6
3	25	30	0	15	55	65	12	45	50
4	40	60	15	0	75	60	10	38	80
5	4	8	55	75	0	5	46	7	8
6	3	9	65	60	5	0	77	9	2
7	50	70	12	10	46	77	0	82	43
8	6	7	45	38	7	9	82	0	8
9	4	6	50	80	8	2	43	8	0

According to a preliminary planning, following 21 line candidates were determined:

line-no.	node sequence (outward trip)
1	1-3-4-5
2	1-3-4-9
3	1-3-7-8
4	1-3-7-6
5	1-3-7-4-5
6	1-3-7-4-9
7	2-3-4-5
8	2-3-4-9
9	2-3-7-8
10	2-3-7-6
11	2-3-7-4-5
12	2-3-7-4-9
13	5-4-3-7-6
14	5-4-3-7-8
15	5-4-7-6
16	5-4-7-8
17	6-7-3-4-9
18	6-7-4-9
19	6-7-8
20	8-7-3-4-9
21	8-7-4-9

Optimal solution (see Appendix A.3):

```
MIP Solution:  1958  (95 iterations, 2 nodes)
Final Solve:   1958  (0 iterations)

Best integer solution possible:        1958.000000
Absolute gap:            0.000000
Relative gap:            0.000000

----    105 VARIABLE y.L  line selection decision

L2 1.000, L3 1.000, L7 1.000, L10 1.000, L15 1.000
```

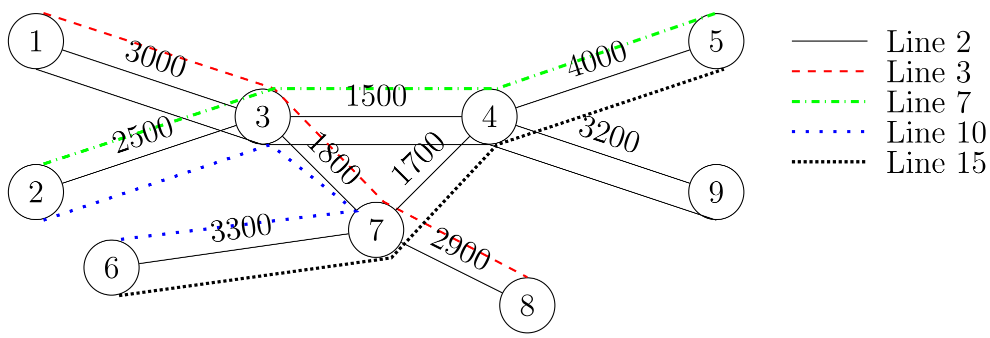


Figure 6.3: Optimal solution

6.3 Minimization of travel times

- Short travel time → attractive route network
- Travel time includes transfer times and thus waiting times
- Waiting times are calculated once the timetable of the travel lines is known
- Captured as a global sum in the predefined line planning:
  - Uniform constant time for transfer, 5 minutes per transfer
  - First boarding like transfer (min. 1 transfer per passenger)
- Lines are represented by set of arcs

Sets

L	Line candidates
V	Sorted nodes
$\vec{E}$	Arcs
$\vec{E}_l$	Arcs of line l

Parameters

d_{uv}	demand (number) of passengers traveling from u to v
k_l	operative cost of line l
B	available budget to cover the operative line costs
t_{lij}	duration of the direct trip from node i to node j when using line l; arc weight $(i,j) \in \vec{E}_l$

Variables

y_l =1, if line l is selected (0, otherwise)

x_{ulij} = number of passengers departing in u and traveling in a vehicle of line l along the arc (i, j)

Model

$$\min F = \sum_{u,l,(i,j)\in\vec{E}_l} t_{lij} \cdot x_{ulij} \tag{6.8}$$

such that

$$\sum_{l\in L} \sum_{(i,v)\in\vec{E}_l} x_{uliv} - \sum_{l\in L} \sum_{(v,j)\in\vec{E}_l} x_{ulvj} = d_{uv} \qquad u, v \in V \mid u \neq v \tag{6.9}$$

$$\sum_{v\in V} d_{uv} \cdot y_l \geq x_{ulij} \qquad u \in V, l \in L \quad (i, j) \in \vec{E}_l \tag{6.10}$$

$$\sum_{l\in L} k_l y_l \leq B \tag{6.11}$$

$$y_l \in \{0, 1\} \qquad l \in L \tag{6.12}$$

$$x_{ulij} \geq 0 \qquad u \in V, l \in L \quad (i, j) \in \vec{E}_l \tag{6.13}$$

Remarks

- Equation 6.21, objective function, minimising travel times
- Constraints 6.22 represents the flow conservation conditions which ensure that passengers from u reach their destination v.
- Constraints 6.23 ensure that a line can only be used if it has been selected as part of the public transport network.
- Constraint 6.24 ensures budget is not exceeded .
- Equations 6.25 and 6.26 define the domain of the variables.

Construction mathematical model
Network 6.4 represents the stations in line 4.
Table 6.3 shows the travel time of each arc

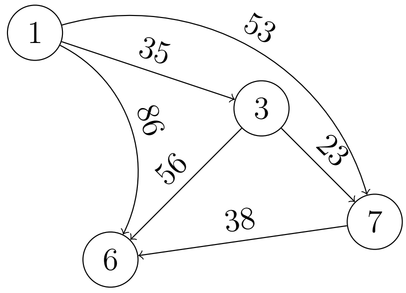


Figure 6.4: Example of arcs in a line

Arc	Travel time (minutes)
$[1,3]$	30
$[3,7]$	18
$[6,7]$	33

Network 6.5 represents all connections on line 4.

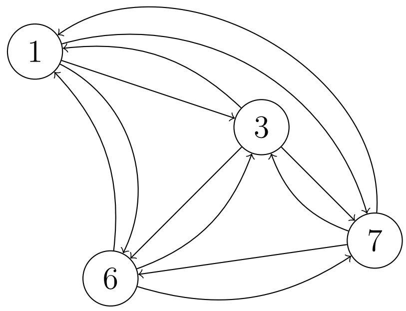


Figure 6.5: Example of connections in a line

Number of arcs of a line with n nodes: $2 \cdot (n-1+n-2+\ldots+1) = n \cdot (n-1) = n^2 - n$

Example 6.2 Given are the relevant data from Example 6.1. Assume that the cost of a line is perfectly related to the travel time. For example:

- Costs related to travel time
  travel time = 5+ 0.01 [TU/MU] * costs
  operative costs = 3000 MU $\Rightarrow$ travel time =30 TU

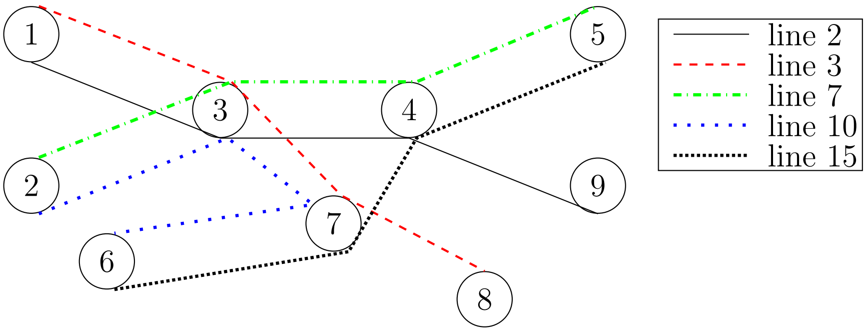


Figure 6.6: Optimal solution

```
MIP Solution:      111358.000000   (4573 iterations, 16 nodes)
Final Solve:       111358.000000   (125 iterations)

Best possible:     111358.000000
Absolute gap:           0.000000
Relative gap:           0.000000


----    153 VARIABLE y.L  line selection decision

L2  1.000,   L3  1.000,   L7  1.000,   L10 1.000,   L15 1.00
```

Graphical illustration of the solution:

The 137 passengers coming from node $u = 1$ are distributed as follows in the network design:

line	arcs	number of passengers
2	(1,3)	37
2	(1,4)	40
2	(1,9)	4
3	(1,7)	50
3	(1,8)	6
7	(3,5)	4
10	(3,2)	5
10	(3,6)	3

Table 6.3 presents by line and by arcs, the number of passengers who travel directly. Suppose we are now interested in calculating how many of these passengers have a direct trip, i.e., the passenger that travel without any transfer to reach his destination.

To calculate the number of passengers who travel directly. We do for each origin node u:

- Select from the Table 6.3 the connections with origin u. For examples:
  - Suppose origin u=1, we select the connections with origin node u= 1: {(1,3),(1,4),(1,9),(1,7),(1,8)}

Suppose we are interested in knowing how many passengers with origin at node u=1 travel directly to the destination node v=3, given that $x_{1,1,1,3}$=37 and $d_{1,3}$=25, so that of the 37 passengers, 25 passengers with origin u=1 travel directly through connection (1,3).

Taking d_{uv} and the outcome variable x_{ulij} as a reference, the number of passengers with a trip origin u who travel directly or have to transfer can be easily calculated. The Table 6.3 presents the number of passengers, for each origin node, who have to make at least one transfer along the route.

node	number of transfer passengers
1	0
2	0
3	40
4	132
5	0
6	0
7	124
8	0
9	0

6.4 Literature

- Borndörfer, R., Grötschel, M., & Pfetsch, M. E. (2008). Models for line planning in public transport. In Computer-aided systems in public transport (pp. 363-378). Springer, Berlin, Heidelberg.

- Bussieck, M. R., Kreuzer, P., & Zimmermann, U. T. (1997). Optimal lines for railway systems. European Journal of Operational Research, 96(1), 54-63.

- Schöbel, A. (2012). Line planning in public transportation: models and methods. OR spectrum, 34(3), 491-510.

6.5 Exercises

Exercise 6.1 Let us look again at the example 6.1. Analyze the dependency of the service level, that is, the number of direct drivers, on the budget. To do this, choose $B = 25000 + 5000 \cdot b$ with $b \in \{0, 1, \ldots, 5\}$. Graphically display the solution.

Exercise 6.2 Let's look again at the example 6.1. Assume that the operational costs are proportional to the length of an edge. Only lines should be allowed which at the same time represent a shortest trip.

1. Determine all possible lines.
2. Determine an optimal solution with GAMS.
3. Perform the same analysis as in task 6.1. How can any differences in comparison to the results in task 6.1 be explained?

Exercise 6.3 The line planning model with the aim of maximizing the number of direct travellers must be extended to include headways. Allowed headways are 30, 20, 15 or 10 minutes. Thus there are four execution modes for a line, of which at most one can be selected. When setting the mode, the capacity provided also varies, i.e. the number of passengers that can be transported along an edge. The model must therefore also represent the passenger flows.

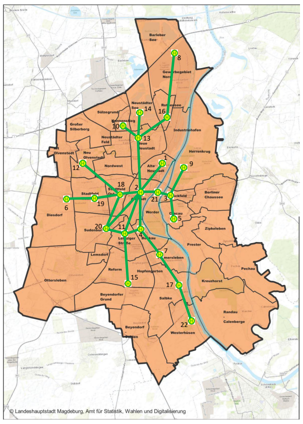
Exercise 6.4 We consider the line planning model with the objective of minimizing travel time. For the example 6.2, analyze the travel time depending on the budget provided. See also task 6.1.

6.6 Case study Magdeburg

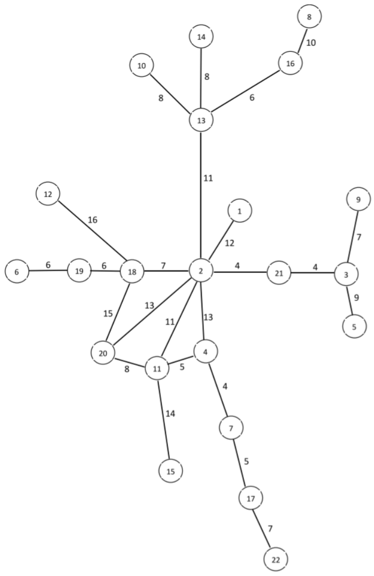
We take the instance of Magdeburg as an example. The provided data includes a passenger flow matrix and a simplified network of Magdeburg consisting of 22 stations and the corresponding undirected edges (6.7(a)). The underlying graph is connected and the edges are weighted by the travel times (6.7(b)). The travel times were given in the form of a matrix.

Passenger Demand - Origin Destination Matrix

For the first model, line planning, the demand of passengers traveling from one node to another is needed. The given data from the PT provider does not contain the origin and destination of the passengers. For this purpose, an origin-destination-matrix (OD-matrix) is derived from the available data **?**. Every entry



((a)) Initial Transportation Network



((b)) Transportation Network with Travel Times

Figure 6.7: Given Network Design

of an OD-matrix represents the number of passengers who want to travel from one destination to another within the given transportation network. The disadvantage of this approach is that the demand in the OD-matrix is highly aggregated and shows only a small number of the conducted trips in reality. Gathering the actual demand in the real world is cost-intensive and no alternative for deriving more exact data is known **?**. The underlying model is the gravity model, which originates from Newton's gravitational law **?**. The mathematical description can be found in **?**, **?** and **?**. For the following implementation, a simplified version of the gravity model is used to derive the OD-matrix. The corresponding GAMS code is listed in X and Figure X shows the resulting asymmetric OD-matrix (Appendix).

6.6.1 solution approach

We focus on solving problems under the customer-oriented and the cost-oriented perceptions:

- ◇ Maximization of service level
- ◇ Minimization of costs

For the customer-oriented view, this section presents two models - minimization of total travel times and maximization of direct trips. The minimization of travel times refers to the total travel times of all passengers. By minimizing the travel time, the convenience and service level for the passengers are increasing. The applied model does not account for the passengers' transfer- and waiting times. Therefore, if no direct trip is possible, a penalty takes place. In this way, the model incentivizes direct trips. For the second model, the number of direct trips is maximized. Direct trips refer to passengers who do not have to change lines to arrive at their destination. To make the results comparable, both measures are presented in percentages, alias the service level. When minimizing the travel times, the shortest path for each origin and destination pair is computed, assuming that no transfer is needed. The service level is then defined as the deviation of the obtained total travel time by the model to the "optimal" benchmark. Similarly, for the maximization of direct trips, the "optimal" result is that all passengers have direct trips. The obtained result from the model is then compared to this "optimal" benchmark as a percentage.

For the cost-oriented view, the operational costs of the operated lines are minimized. Deviating from **?** model, operational costs are considered to be solely variable costs per train per kilometer. When minimizing the costs, the resulting service level from the customer-oriented view serves as a lower bound and cannot be subceeded. For both models, the following assumptions are made:

- A uniform headway applies to all lines. Headway is defined as the time interval between two trips on one operated line (**?**). For example, a headway of 15 minutes produces 4 tips per hour

- The size of the transportation mode, also considering the headway, is large enough to meet the demand of the passengers

As an extension, a penalty for a high number of lines is included. With a high number of operated lines, the variable costs per vehicle are relatively high for serving a small number of nodes, resulting in a low cost-benefit ratio. Even though shorter lines are more reliable (**?**), they lead to pendulum tours and thus have to be avoided for logical reasons. In this model, the penalty is applied per additional line in case the threshold of operated lines is exceeded.

Maximization of direct trips

▷ **Objective**

- To maximize the number of passengers who do not have to change between lines

▷ **Assumptions**

- A uniform headway applies to all lines
- The vehicle size (in combination with the selected headway) is sufficiently large so that all passengers can be transported at all times

▷ **Sets**

- L: line candidates
- V: sorted nodes (stops)
- E: edges
- L_{uv}: lines that connect the nodes (stops) $u \in V$ and $v \in V$ directly with $u < v$, i.e. passengers using the line $l \in L_{uv}$ can travel (without changing from $u \in V$ to $v \in V$ or the other way around); a line that indeed connects the nodes $u \in V$ and $v \in V$, but equally means a significant detour for passengers, does not have to be included in L_{uv}
- $\tilde{L}_{ij}$: lines serving the edge $[i, j] \in E$

▷ **Parameters**

- d_{uv}: demand (number) of passengers travelling from $u \in V$ to $v \in V$
- k_l: operating costs of line $l \in L$
- B: available budget
- C_{ij}: maximal number of lines along edge $[i, j] \in E$

▷ **Variables**

- $Y_l = 1$, if line $l \in L$ is selected (0, otherwise)
- $X_{uv} = 1$, if a direct trip is feasible from $u \in V$ to $v \in V$ (0, otherwise)

▷ **Model**

$$\text{Maximize } F = \sum_{u,v \in V | u < v} (d_{uv} + d_{vu}) \cdot X_{uv} \tag{6.14}$$

such that

$$\sum_{l \in L_{uv}} Y_l \geq X_{uv} \qquad \forall u, v \in V \mid u < v \qquad (6.15)$$

$$\sum_{l \in \tilde{L}_{ij}} Y_l \geq 1 \qquad \forall [i,j] \in E \qquad (6.16)$$

$$\sum_{l \in \tilde{L}_{ij}} Y_l \leq C_{ij} \qquad \forall [i,j] \in E \qquad (6.17)$$

$$\sum_{l \in L} k_l \cdot Y_l \leq B \qquad (6.18)$$

$$Y_l \in \{0,1\} \qquad \forall l \in L \qquad (6.19)$$

$$X_{uv} \in \{0,1\} \qquad \forall u, v \in V \mid u < v \qquad (6.20)$$

▷ **Remarks**

- (6.14) represents the objective function, i.e. the number of direct trips must be maximized. It should be noted that each line is defined by a round trip.
- (6.15) ensures that direct trips from u to v are only included in the objective function if at least one line connects the two nodes.
- with (6.17) we avoid that more vehicles per time unit pass through an edge than is technically possible.
- compliance with the budget is ensured by (6.18).
- the permissible domain of the decision variables are defined by (6.19) and (6.20).

Minimization of travel times

Model approach

- Short travel time → attractive line network (lecture network design)
- Travel time includes transfer times and, thus, waiting times
- Waiting times are calculated once the timetable of the lines is known
  - Uniform constant time for transfer (for example, 5 minutes per transfer)
  - The first boarding is also considered as a transfer (min. 1 transfer per passenger)

Sets

L: Line candidates

V: Sorted nodes

$\vec{E}$: Arcs

$\vec{E_l}$: Arcs of line $l \in L$

Parameters

d_{uv} : demand (number) of passengers traveling from $u \in V$ to $v \in V$

k_l : operational cost of line $l \in L$

B : available budget to cover the operational line costs

t_{lij} : duration of the direct trip from node $i \in V$ to node $j \in V$ when using line $l \in L$; arc weight $[i,j] \in \vec{E_l}$

Variables

- $Y_l = 1$, if line $l \in L$ is selected (0, otherwise)
- X_{ulij} = number of passengers departing in u and traveling in a vehicle of line $l \in L$ along the arc $[i,j] \in \vec{E_l}$

Model

$$\text{Minimize}\, F = \sum_{u \in V} \sum_{l \in L} \sum_{[i,j] \in \vec{E_l}} t_{lij} \cdot X_{ulij} \tag{6.21}$$

such that

$$\sum_{\forall l \in L} \sum_{[i,v] \in \vec{E_l}} X_{uliv} - \sum_{\forall l \in L} \sum_{[v,j] \in \vec{E_l}} X_{ulvj} = d_{uv} \qquad \forall u, v \in V \mid u \neq v \tag{6.22}$$

$$X_{ulij} \leq M \cdot Y_l \qquad \forall u \in V, l \in L, [i,j] \in \vec{E_l} \tag{6.23}$$

$$\sum_{\forall l \in L} k_l \cdot Y_l \leq B \tag{6.24}$$

$$Y_l \in \{0,1\} \qquad \forall l \in L \tag{6.25}$$

$$X_{ulij} \geq 0 \qquad \forall u \in V, l \in L, [i,j] \in \vec{E_l} \tag{6.26}$$

Remarks

- Equation 6.21, objective function, minimising travel times
- Constraints 6.22 represents the flow conservation conditions which ensure that passengers from u reach their destination v.
- Constraints 6.23 ensure that a line can only be used if it has been selected as part of the public transport network.
- Constraint 6.24 ensures budget is not exceeded .
- Equations 6.25 and 6.26 define the domain of the variables.

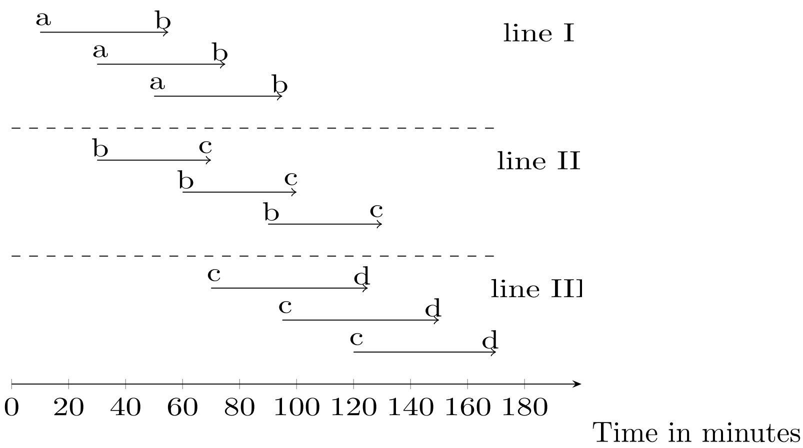
Chapter 7

Timetabling

7.1 Introduction

Objective: Minimization of waiting times when changing lines.
Remarks/assumptions:

- Line service with headway is given
- Waiting time when changing from line I to line II: Minimum { departure time line II - arrival time line I with departure time line II $\geq$ arrival time line I }.
- Transfer times are constant and therefore not relevant for decision making; departure time line II $\geq$ arrival time line I + transfer time.
- Time measured in minutes.
- Passengers transfer at most once
  $\Rightarrow$ simplified calculation of waiting times.



7.2 Minimization of waiting times

In the following model, each line trip is given a number, starting from zero with the numbering, i.e. the first trip is given the trip number zero, the second trip is given the number one, and so on. With a given headway, the departure times of the trips of a line clearly result from the start time of the first line trip. If, for example, we assume that the headway is 10 minutes and the first line trip (with the number zero) starts at 7:00 a.m., then the third trip of the line starts at 7:20 a.m.

Sets

L	Lines (indices: l, λ)
H	Stops (index: h)

Parameters

$u_{l\lambda h}$	number of passengers that change at stop h from line l to line λ
f_{lh}	travel time from the starting station of line l to stop h
τ_l	headway of line l

Variables

$w_{l\lambda h}$	waiting time at stop h when changing from line l to line λ
t_l	starting time of the first trip of line l
$z_{l\lambda h}$	number of the line trip of line l, which is chosen by passengers who leave line l at stop h in order to change to line λ.
$\tilde{z}_{l\lambda h}$	number of the line trip of line λ, which is chosen by passengers who leave line l at stop h in order to change to line λ.

Model

$$\min F = \sum_{l,\lambda,h|u_{l\lambda h}>0} u_{l\lambda h} w_{l\lambda h} + 0.0001 \sum_{l,\lambda,h|u_{l\lambda h}>0} z_{l\lambda h} \tag{7.1}$$

s.t.

$$\begin{aligned}
t_l + f_{lh} + \tau_l z_{l\lambda h} - (t_\lambda + f_{\lambda h} + \tau_\lambda \tilde{z}_{l\lambda h}) &= w_{l\lambda h} && l,\lambda,h \mid u_{l\lambda h} > 0 && (7.2)\\
t_l &\geq 0 && l \in L && (7.3)\\
w_{l\lambda h} &\geq 0 && l,\lambda,h \mid u_{l\lambda h} > 0 && (7.4)\\
z_{l\lambda h}, \tilde{z}_{l\lambda h} &\in Z^+ && l,\lambda,h \mid u_{l\lambda h} > 0 && (7.5)
\end{aligned}$$

Remarks

- (7.1) minimizes the waiting time of all passengers. The second term only serves to ensure that the lowest possible trip numbers are used for several optimal solutions.

- (7.2) determines the waiting time per passenger for each line- and stop-related relation.

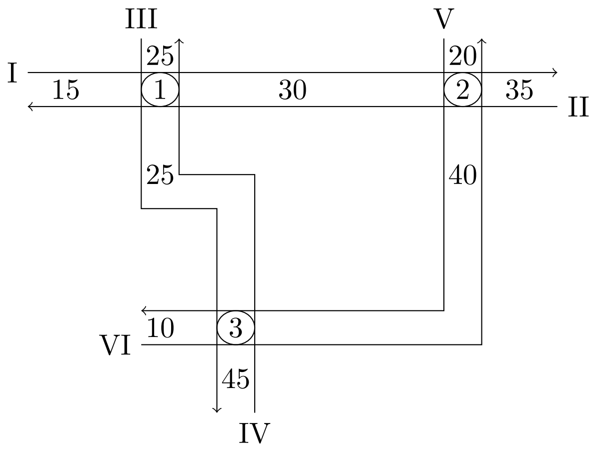
- (7.3) restricts the domain for the start time of the first trip to non-negative values for each line.

- (7.4) defines the non-negativity conditions for the waiting times. Thus no incompatible transfers can occur in (7.2).

- (7.5) ensures that integer and non-negative trip numbers are allocated.

7.3 Example

Example 7.1 The following network is given:



The arc weights correspond to travel times in minutes. Line II is the opposite direction of line I. The same applies to line pairs (III, IV) and (V, VI). The duration of a trip is independent of the direction of travel. Stops are included in the travel times as appropriate. A headway of 15 minutes applies to all lines. The number of transfer passengers from one line to another is shown in the following table.

from line	to line	stop	transfer passengers
I	IV	1	250
I	VI	2	175
II	IV	1	250
II	V	2	150
II	VI	2	125
III	I	1	150
III	II	1	100
III	V	3	125
IV	V	3	200
IV	VI	3	125
V	I	2	150
V	II	2	100
V	III	3	150
VI	III	3	150
VI	IV	3	100

GAMS provides the following solution (see Appendix A.5):

```
line starting time
    1      40
    2       5
    3      30
    4       0
    5      20
    6      35

from : from line
to: to line
stop: stop station
at  : arrival time
from : departure time
wt  : waiting time
zarr : trip number of the arriving line
zdep : trip number of the departing line

transfer relationships
from to stop transf at from wt zarr zdep
  1    4    1  250   70   70   0   0   1
  1    6    2  175   85   85   0   0   0
  2    4    1  250   70   70   0   0   0
  2    5    2  150   40   40   0   0   0
  2    6    2  125   85   85   0   0   3
  3    1    1  150   55   55   0   0   0
```

3	2	1	100	70	70	0	0	1
3	5	3	125	80	80	0	0	0
4	5	3	200	75	80	5	0	2
4	6	3	125	45	45	0	0	0
5	1	2	150	85	85	0	0	3
5	2	2	100	40	40	0	0	0
5	3	3	150	80	80	0	0	0
6	3	3	150	75	80	5	0	2
6	4	3	100	45	45	0	0	0

All in all, the transfer passengers must wait $200 \cdot 5 + 150 \cdot 5 = 1750$ minutes. This results in an average waiting time of 0.76 minutes.

7.3.1 Transfer times and dwell times

In an extension of the model formulation the transfer times of the passengers and the dwell times of the vehicles are included. The transfer times include the time needed to change platforms at the transfer stations and corresponding infrastructure issues (e.g. stairs). The arrival times of the passengers at the stations therefore increase by the needed transfer time. The dwell time refers to the difference between the arrival and departure of a vehicle, thus the time it stops at the platform. The departure time of the vehicle is therefore prolonged. The model is adjusted accordingly:

Parameters

$c_{l\lambda h}$	Transfer time when changing from line l to line λ at stop h
$a_{\lambda h}$	Dwell time of line λ at station h

The equation to calculate the waiting time 7.2 is adapted accordingly:

$$t_l + f_{lh} + a_{\lambda h} + \tau_l z_{l\lambda h} - (t_\lambda + f_{\lambda h} + c_{l\lambda h} + \tau_\lambda \tilde{z}_{l\lambda h}) = w_{l\lambda h} \qquad \forall l, \lambda, h \mid u_{l\lambda h} > 0 \quad (7.6)$$

7.4 Determination of a synchronized timetable

Sets

L	Lines (indices: l, λ)
H	Stops (index: h)

Parameters

τ	headway
$\delta_{l,\lambda,h}$	difference of arrivals within the headway interval of lines l and λ if both lines start at time zero ($T_l = T_\lambda = 0$)

Variables

$X_{l,\lambda,h}$ = 1, if headway shift of line l when changing from line l to line λ (0, otherwise)

$\tilde{X}_{l,\lambda,h}$ = 1, if headway shift of line λ when changing from line l to line λ (0, otherwise)

7.5 Literature

- Ceder (1987)

7.6 Exercises

Exercise 7.1 The planning begin ($t = 0$) is 6:00 a.m. For the example 7.1, determine the timetable of stop $h = 1$ for the period 10:00 a.m. - 12:00 p.m.

Exercise 7.2 Assume the travel time between two subsequent stops can be delayed by up to 2 minutes. In addition, transfer times between two lines should be explicitly taken into account, with a vehicle waiting time of 1 minute at a transfer node. How is the model formulation to be modified?

Exercise 7.3 Modify the problem such that the maximum waiting time is minimized.

Exercise 7.4 ASEAG operates multiple public transport lines in Aachen, and the headway between buses is fixed at 15 minutes. ASEAG needs to determine the optimal start times for each line and identify line connections that will minimize waiting times for passengers, enabling them to reach their destinations promptly. The number of passengers waiting at each station is provided in the table below. *Additional Information*: Considering traffic conditions, a team of engineers estimated that the travel times for each line to reach each station, as mentioned in the table, follow a uniform distribution with parameters 10 and 50 ($f_{lh} \sim U(10, 50)$).
Each item is solved independently.

1. Utilizing one of the models covered in the lecture, determine the optimal start times for each line and identify the (routes of) lines to be connected. Please report the optimal solution using GAMS. Note: Please ensure that the times each line starts operating must be less than or equal to the headway. I.e., the latest line to start operating should be at the 15th minute at the latest (if all public transport operations start at minute 0).
2. Evaluate the impact of a headway of 5 minutes and 20 minutes. Analyze the results and report your optimal solution.
3. Modify your model to minimize the maximum waiting time. Please report the optimal solution using GAMS.

l	λ	h	$u_{l\lambda h}$
1	6	2	2180
2	6	2	3170
	2	2	100
4	6	2	3748
	12	2	50
5	6	2	1648
6	11	2	692
10	9	11	100
	6	2	1375
11	12	5	1541
	13	2	491
12	7	13	467
13	2	2	129
	11	5	740
14	5	10	103
	6	13	2775

4. Now consider a headway of 15 min, a vehicle dwell time at the station of 1 min, and a passenger transfer time of 40 seconds. Utilizing one of the models covered in the lecture, determine the optimal start times for each line and identify the (routes of) lines to be connected. Please report the optimal solution using GAMS.

5. Determine the list of trips, considering all lines, and the time at which each trip will start running. Assume you have a planning period between 0:00 hr. and 7:30 hr. (night schedule) with a headway of 15 min.

6. Assume there is a special event this weekend in Aachen, and the number of passengers waiting (the number indicated in Table 1) cannot be estimated with certainty. Therefore, it is suggested that you generate five possible scenarios that represent five different samples of the parameter $u_{l\lambda h}$ for each scenario with the same probability of occurrence. You should create each scenario of the parameter $u_{l\lambda h}$ using an integer uniform distribution with parameters 800 and 3200. For each scenario, determine the minimum total waiting time (passenger-minutes) and the maximum passenger waiting time. Provide the optimal solution for each scenario using GAMS and determine the total expected waiting time (passenger-minutes) and

the maximum expected waiting time in the system. Note: Please use the data and assumptions that were utilized to solve point 1.