Doctoral Thesis

Investigation of space-time structures in public transport networks and their optimization

Author(s):
Alt, Bernhard

Publication Date:
2010

Permanent Link:
https://doi.org/10.3929/ethz-a-006229432

Rights / License:
In Copyright - Non-Commercial Use Permitted
Investigation of space-time structures in public transport networks and their optimization

A dissertation submitted to

ETH ZURICH

for the degree of

Doctor of Sciences

presented by

BERNHARD CORNELIUS ALT

Diplom-Maschinenbauingenieur, Universität Siegen

born 11th January, 1975

citizen of Germany

accepted on the recommendation of

Prof. Dr. Ulrich Weidmann, examiner

Prof. Dr. Markus Friedrich, co-examiner

2010
"Die Zukunft, die wir wollen, muss erfunden werden.
Sonst bekommen wir eine, die wir nicht wollen"

Joseph Beuys
Abstract

A method for public transport (PT) network optimization was developed that considers the entire PT chain from door-to-door. The relatively small number of network evaluations of the approach enables the use of comparable realistic transport models. Consequently, PT systems can be evaluated more realistically than in current approaches without significantly increasing computing times.

The approach optimizes different speed levels, e.g., bus and tram, regional train, etc., sequentially, starting with the fastest service. To reduce computing times, each speed level for the geographic area under consideration is divided into several planning areas. If computing times for planning areas are short enough, computing times for network design in entire areas can be handled as well since they are only linearly dependent on the total size of the area based on the suggested approach. For each planning area, the approach uses a network reduction process (also called guided stochastic search heuristic) that requires comparatively few network evaluations. The network reduction process starts with a network of the shortest lines. Then, lines are deleted, merged or shortened sequentially using the ant colony optimization algorithm. A genetic algorithm simultaneously optimizes service frequencies and vehicle sizes. During the network reduction process, total operating and travel time costs are minimized. For network evaluations, a headway-based stochastic multiple route assignment is used.

The reduction approach was compared to existing approaches by applying it to Mandl’s Swiss benchmark problem. Based on the comparison, it can be stated that the approach developed shows promising results in terms of optimized fleet size and user costs.

The multiple area approach and the reduction process were tested together in a larger case study in the city of Winterthur, Switzerland. The radial PT network was updated in 2007 using a planner’s experience. Mainly, four radial lines were merged into through lines. The reduction process was able to additionally reduce total costs by around 3%.

To ensure that such small changes can bring sound improvements, the accuracy of the OD matrix and the assignment model has to be sufficient. The relative error of passenger flows at maximum load sections of PT lines should not be higher than 5% for each speed level. In the future, data available from electronic ticketing systems will make it possible to validate and improve assignment models and achieve such a higher level of assignment quality.
While keeping total costs at one level, the variability of time costs (ca. 7%) and of operating costs (ca. 20%) could be used by politics to decide how the different costs of PT are covered and by whom. Hereby, the effects on modal split should be considered.

The nearly shortest PT lines are suited to high-demand relations. For lower demand relations, circuitous PT lines with still acceptable headway are advantageous. It turned out that a third class of PT lines should be considered: connecting important transfer nodes with each other. This class of PT lines shows forms similar to the magnetic streamlines of a field that connects several magnetic poles.

The multiple area approach was tested by optimizing the slowest speed level (bus level). Several speed levels could not be optimized due to the long computing times for network evaluations. Results from optimizing one single planning area are only slightly better than from just sequentially optimizing the same area in two steps (two planning areas). Furthermore, it could be shown that it is helpful to have a kind of “ideal infrastructure network” to compare and evaluate PT networks on given infrastructures.

For future research, it can be stated that schedule optimization is an important part of PT network design and should be fully integrated. In this case, the integration of stop placement can show its full effects. In placing stops, travel times can be adjusted. The integration of schedule optimization allows using schedule-based assignments to model passenger behavior. First, the assignment becomes more accurate. Second, travel times can be reduced while offering coordinated transfers, especially for less loaded lines with longer headways. Schedule-based assignment allows the consideration of more frequent service close to the most loaded section of a line. One important condition for realistic schedule-based assignments is that delay distributions should be considered, at least on an average level.

Numerous PT network and schedule design studies have been conducted to date. Due to the complexity of PT network design and the variety of approaches to be performed, it is difficult to define benchmark problem sets and standard transport models that allow different solutions to be compared. However, simplifications in network design methods and in the transport models used during network design require the application of benchmark problems in order to compare the results of different network design methods. Therefore, the basis for all benchmark problems should be an accurate transport model. Benchmark problem sets and standard transport models should be simple enough to keep the amount of required data low, but complex enough to allow a comparison of PT network design methods under realistic conditions. For the future, the development of new benchmark problems will be of great interest. In particular, larger benchmark problems and benchmark problems that cover the entire PT chain from door-to-door should be considered.
To summarize, the reduction approach in combination with the ant colony optimization is promising, as can be seen from the results for Mandl’s benchmark problem, as well as the solutions for the more realistic study in Winterthur. However, in the case of Winterthur, network evaluations were about a factor 10 too slow to be able to get satisfying results. In combination with the accuracy of the transport model, this still prevented being able to answer questions concerning, for example, optimal parameters for a given demand pattern such as the number of speed levels, the line distances and the transfer node distances.
Zusammenfassung


Der Netzreduktions-Prozess wurde durch einen Test an Mandls Benchmarkproblem mit anderen Ansätzen verglichen. Bezüglich der Anzahl an Fahrzeugen und der Nutzerkosten wurden viel versprechende Ergebnisse erzielt.


Um sicher zu gehen, dass solch kleine Verbesserungen in der Tat Verbesserungen darstellen, muss die Genauigkeit der Quell-Ziel-Matrix und die des Umlegungsmodells hoch genug sein. Die Fahrgastflüsse an den maximal belasteten Querschnitten sollten für jede Geschwindigkeitsebene einen durchschnittlichen relativen Fehler von unter 5% aufweisen. In
Zukunft werden solche Genauigkeiten der Verkehrsmodelle dank elektronischer Ticketsysteme erreichbar sein.

Solange die Gesamtkosten auf einem ähnlichen Niveau gehalten werden, steht die Variabilität der Zeitkosten (ca. 7%) und die der Betriebskosten (ca. 20%) als Entscheidungsspielraum der Politik zur Verfügung. Diese kann entscheiden, wie die Kosten des ÖV verteilt werden und wer sie trägt. Allerdings sollten hierbei die Effekte auf den Modal-Splitt Berücksichtigung finden.


Bis heute wurde bereits eine Vielzahl an Studien zum Design von ÖV-Netzen unternommen. Aufgrund der Komplexität des Designs von ÖV-Netzen und der Vielfalt an möglichen Designansätzen ist es schwierig, Benchmarkprobleme und Standardverkehrsmodelle zu entwickeln mit denen die Lösungen der verschiedenen Ansätze miteinander verglichen werden können.

Acknowledgement

First of all, I am very grateful to my supervisor Prof. Ulrich Weidmann for his interest to my initial idea of this project and for having given me the opportunity to develop it further in his group at ETH Zurich’s Institute for Transport Planning and Systems (IVT). His patience and support accompanied me through this “marathon”. The discussions we had together often inspired me so I could go a step further in the research. His suggestions especially helped me to focus on the main issues of public transport systems.

Second, I want to thank my co-supervisor Prof. Markus Friedrich for his interest in the details of this approach, for the discussions and his suggestions. He helped me to make the thesis more interesting for readers who want to reproduce some single steps of the approach, and to give a deeper insight into the results of the developed and implemented network design process.

Beside the help of my two supervisors I got support in various forms:

- The many published works, presentations at conferences and meetings and other information is invaluable not only for this thesis. Thanks to all, contributing to the wisdom of mankind. A small selection of links to it can be found in the literature of this thesis.

- The data about the behavior of passengers and other transport users is the basis to model and improve transport systems. That is why I want to thank the citizen of the Canton of Zurich and of Switzerland for providing their travel data to statistics in answering surveys. Further, thanks go to the Canton of Zurich for allowing me to use its transport model for the study, to Thomas Kreyenbühl and the Verkehrsbetriebe Glattalbahn for getting some data from passenger counts at Winterthur, to the PTV AG for the research license of their tool VISUM [PTV07] and to Prof. Kay W. Axhausen, Dr. Michael Balmer, Dr. Philipp Fröhlich, Dr. Michael Löchl, Nadine Schüssler (soon with a doctor degree as well), Dr. Milenko Vrtic for creating the transport model of the Canton Zurich and/or helping me to understand the model, to disaggregate it, to use the tool VISUM or to develop the software tool MatSim, a dynamic transport model showing the future of transport planning. Thanks go as well to Prof. Christoph Mandl who created a sample problem for network design in 1979 which later became, thanks to others using it, the Swiss benchmark problem. Mr. Mandl told me where the network of the problem is situated in Switzerland. Finally, thanks go to Linjuan Yu try-
ing together with me to understand the probabilities behind the headway-based stochastic multiple route assignment [PTV07].

- To make this thesis readable, several people were involved. For improving the English and for suggestions concerning the organization of the report multiple thanks go to Samuel Alt, Prof. Markus Friedrich, Dr. Sonja L. Beperling-Kurz, Andy Nash, Milena Scherer, Prof. Ulrich Weidmann and Linjuan Yu.

- Thanks go to my colleagues and the students of the group transport systems and the IVT, especially to my office colleague Milena Scherer for discussion, Alex Erath and Remo Jucker for cooperating in using the programming language python and the later for testing the ant colony optimization.

- For discussing with me the transport related issues, for motivating me and for showing understanding for the time I spent on this thesis I thank my family, my friends and especially my wife and partner Linjuan Yu.
# Table of Content

## 1 INTRODUCTION AND OBJECTIVES ................................................................. 1

1.1 Introduction .............................................................................................................. 1
   1.1.1 Public Transport Systems ................................................................. 1
   1.1.2 Major Questions ............................................................................. 2
   1.1.3 Software Prototype ........................................................................ 2
   1.1.4 Organization of the Chapter .......................................................... 3

1.2 Design Objectives and Generalized Costs .......................................................... 3
   1.2.1 Overview ......................................................................................... 3
   1.2.2 Customers ..................................................................................... 4
   1.2.3 Service Providers .......................................................................... 6
   1.2.4 Competing Modes ......................................................................... 6
   1.2.5 Politics ......................................................................................... 7

1.3 Objectives and basic Methods of the Study ......................................................... 9
   1.3.1 Overview ......................................................................................... 9
   1.3.2 Generalized Costs Model for this Study and Constraints .............. 9
   1.3.3 Simplifications and Limitations of the Generalized Costs Model .. 11

1.4 Overview of the Dissertation .............................................................................. 14

## 2 LITERATURE REVIEW OF LINE NETWORK DESIGN .............................. 15

2.1 Introduction ............................................................................................................. 15
   2.1.1 General Aims ............................................................................... 15
   2.1.2 Basic Mathematical Approaches .................................................. 16

2.2 Analytical Network Generation Studies ............................................................. 17

2.3 Simultaneous Network Generation Studies ....................................................... 18
   2.3.1 Overview ....................................................................................... 18
   2.3.2 Genetic Algorithm ....................................................................... 19
   2.3.3 Integer Programming .................................................................... 19
   2.3.4 Simulated Annealing .................................................................... 20

2.4 Sequential Network Generation Studies ............................................................ 21
   2.4.1 Overview ....................................................................................... 21
   2.4.2 Line Network Reduction ............................................................. 22
   2.4.3 Line Network Composition .......................................................... 22

2.5 Synthesis and Conclusion .................................................................................... 24
4.3 Line Network Design ................................................................. 68
4.3.1 Introduction ........................................................................... 68
4.3.2 Start Condition ...................................................................... 69
4.3.3 Adjustment of Stop Densities on Links and OD matrix Aggregation ............................................. 72
4.3.4 Headway Optimization using a Genetic Algorithm .......................................................................... 73
4.3.5 Network Reduction Process ..................................................... 73
4.3.6 Ant Colony Optimization ......................................................... 79
4.3.7 Local OD Matrix during Ant Colony Optimization ................................................................. 85
4.3.8 Final Stop Placement on Lines .................................................. 86

5 CASE STUDIES ............................................................................. 88
5.1 Overview on the Case Studies ..................................................... 88
5.2 Introduction to the Case Study of Winterthur ........................................... 88
5.2.1 PT Network of Winterthur 2008 ................................................. 88
5.2.2 Transport Model of the Canton of Zurich ...................................... 90
5.2.3 Adaptation of the Model to Needs of Network Design ................................................... 91
5.2.4 Parameters of the Objective Function for the Case Study Winterthur ............................................ 95
5.3 Case Study Winterthur .................................................................. 97
5.3.1 Network 2008 .......................................................................... 97
5.3.2 Headway Optimization for Winterthurs PT Network ................................................................. 98
5.3.3 New Design of Winterthurs PT Network ....................................... 102
5.3.4 Sequential PT Network Design Using two Overlapping Planning Areas ......................................... 106
5.3.5 Final Stop Placement .................................................................. 108
5.4 Design Parameter Variation ........................................................ 109
5.4.1 Variation of the Cost Balance Factor and of Maximum Bus Capacities ........................................ 109
5.4.2 Variation of Maximum Bus Speed and Artificial Infrastructure Network ..................................... 112
5.5 Summary of the Case Study Winterthur ........................................ 114
5.6 Comparison with Results of Other Approaches ............................... 117

6 CONCLUSIONS ........................................................................... 128
6.1 Summary of the Work ................................................................. 128
6.2 Further Research ........................................................................ 129
6.3 Final Remarks ............................................................................... 130

7 REFERENCES ............................................................................... 132
List of Tables

Table 1  Objectives of transport policy in agglomerations........................................ 8

Table 2  Mathematical approaches for network design, listed here by help of 9 criteria (see as well Table 3). .................................................................26

Table 3  Mathematical approaches for network design, listed here by help of 9 criteria (see as well Table 2). ....................................................................27

Table 4  Mathematical approaches for network design, listed here by help of 9 criteria (see as well Table 5). .................................................................28

Table 5  Mathematical approaches for network design, listed here by help of 9 criteria (see as well Table 4). .................................................................29

Table 6  Existing solutions for the review criteria (numerical approaches only). ......30

Table 7  Mean travel speeds for urban PT depending on stop distances (source: [Weid08]). ................................................................................................. 35

Table 8  Properties of nearly shortest lines and NSR if fractions of nearly shortest lines: sample start network for a test problem (see Figure 25 and Figure 26). Terminal times of at least 3 minutes at both terminal stations of a line are mandatory. They are part of the cycle time. The “maximum headway” is the longest headway still complying with vehicle capacities. ................71

Table 9  Selection probabilities of lines to be chosen for reduction for a test problem and after the 8th reduction step (see Figure 31 and Figure 32); $h_{\text{max}} = 30$; $\beta = 3$; $r_p = 0.7613 (=14'321/18'811)$ .........................................................75

Table 10 Line properties: after the second reduction step for the test problem (see Figure 25 and Figure 26). .......................................................................77

Table 11 Line properties: after the third reduction step for the test problem (see Figure 25 and Figure 26). .......................................................................78
Table 12  Network properties for all reduction steps (with changes in the network) of the test problem (see Figure 25 and Figure 26). ........................................81

Table 13  Line properties: after the 5th reduction step for the test problem (see Figure 25 and Figure 26). .................................................................................82

Table 14  Line properties: after the 7th reduction step for the test problem (see Figure 25 and Figure 26). .................................................................................83

Table 15  Line properties: after the 8th reduction step for the test problem (see Figure 25 and Figure 26). .................................................................................84

Table 16  Profit and loss statement in parts of STADTBUS Winterthur 2005 [CHF] source: STADTBUS Winterthur. (ZVV: “Zürcher Verkehrsverbund”, transport association of the Canton of Zurich) ...........................................90

Table 17  Average demand at maximum load sections (see Figure 39) during morning peak hour. The sum of the demand at the maximum load sections of all 6 lines was made identical in adjusting total demand during the morning peak hour. ...........................................................................................................94

Table 18  Costs for infrastructure investments and external costs for Switzerland in percentage of operating costs (see [BUND06]). ........................................95

Table 19  Accepted line headways ...........................................................................96

Table 20  Vehicle capacities and estimated costs for service composition kilometers (assumed to be 1/3 of the total costs) and service composition hours (assumed to be 2/3 of the total costs) for standard compositions in the Canton of Zurich (compare [Weid08] (1) section 2.4.1.3.6 and [Frank08]). 97

Table 21  Network properties of the design cases for the PT network of Winterthur. Ca. 15’000 CHF are estimated for the operation of the rail lines, ca. 110’000 minutes are estimated for the in-vehicle travel times on the rail network and ca. 280’000 minutes are estimated for perceived waiting times for trains and perceived transfer time from or towards rail. .................................99

Table 22  Case 0: Line properties of the PT network of Winterthur realized in 2008 (Figure 40). ...................................................................................................100
<table>
<thead>
<tr>
<th>Table</th>
<th>Case Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>Case 2: Line properties of the PT network of Winterthur realized in 2008 after headway optimization, for the optimized lines only (see as well Figure 41). Assignment is based on <strong>line nodes</strong> and <strong>link corridors</strong>.</td>
<td>101</td>
</tr>
<tr>
<td>24</td>
<td>Case 4: Line properties of the PT network, designed with existing lines after long calculation (see Figure 42). Assignment is based on <strong>line nodes</strong> and <strong>link corridors</strong>.</td>
<td>103</td>
</tr>
<tr>
<td>25</td>
<td>Case 6: Line properties of the PT network design without existing lines (see Figure 43).</td>
<td>104</td>
</tr>
<tr>
<td>26</td>
<td>Case 5: Line properties of the sequential PT network design of two overlapping areas with existing lines (see Figure 46).</td>
<td>108</td>
</tr>
<tr>
<td>27</td>
<td>Network properties. Stop placement comparison of four cases. In two cases the cost balance factor has been changed.</td>
<td>109</td>
</tr>
<tr>
<td>28</td>
<td>Network properties. Variation of the cost balance factor for the PT network of Winterthur and of vehicle capacities.</td>
<td>110</td>
</tr>
<tr>
<td>29</td>
<td>Case 9: Line properties of a new PT network designed with existing lines (maximum vehicle load is two standing passengers/ m²); ( f_{cost} = 6 ).</td>
<td>110</td>
</tr>
<tr>
<td>30</td>
<td>Case 10: Line properties of the new PT network, designed with existing lines (maximum vehicle load is one standing passenger/ m², see Figure 47); ( f_{cost} = 6 ).</td>
<td>111</td>
</tr>
<tr>
<td>31</td>
<td>Network properties. Variation of maximum bus speed and use of an artificial infrastructure network; ( F : 2/3 ) of average car speed.</td>
<td>112</td>
</tr>
<tr>
<td>32</td>
<td>Case 7: Line properties of the PT network, designed with existing lines (Maximum bus speed is 50km/h).</td>
<td>113</td>
</tr>
<tr>
<td>33</td>
<td>Case 8: Line properties of the PT network design on an artificial infrastructure (see Figure 48).</td>
<td>114</td>
</tr>
<tr>
<td>34</td>
<td>Design cases. Assignment is based on <strong>line nodes</strong> and <strong>link corridors</strong> (( v_{max} = 2/3 ) ( v_{car} ); ( f_{cost} = 1 ); max. capacity = 2 passenger/m²).</td>
<td>116</td>
</tr>
</tbody>
</table>
Table 35  Mandl’s Swiss benchmark problem [Mandl79] (see Figure 51): Symmetric OD matrix for an average day. .................................................................119

Table 36  Comparison of line network layouts (compare Table 37 to Table 43) created by different search- or optimization methods for Mandl’s benchmark problem (see Figure 51) as expanded by Baaj and Mahmassani. Comparison 1 compares networks with 75 or more vehicles, comparison 2 with 68 or more. .....................................................................................120

Table 37  Line properties of GSSH comparison 1 network design solution displayed in Figure 52. The “maximum headway” is the longest headway still complying with vehicle capacities. ............................................................................121

Table 38  Line properties of Lee’s [Lee98] network design solution displayed in Figure 53...........................................................................................................122

Table 39  Line properties of Case 2 of Shih and Mahmassani’s [Shih94] network design solution displayed in Figure 54....................................................123

Table 40  Line properties of Baaj and Mahmassani’s [Baaj91] Case 1 network design solution displayed in Figure 55............................................................ 124

Table 41  Line properties of Mandl’s [Mandl79] network design solution displayed in Figure 56.....................................................................................................125

Table 42  Line properties of GSSH comparison 2 network design solution displayed in Figure 57..............................................................................................126

Table 43  Line properties of Zhao and Gan’s [Zhao03] Case S&M2 network design solution displayed in Figure 58...............................................................127

List of Figures

Figure 1  Basic network design approaches. ..........................................................15

Figure 2  Network Design Schema. .......................................................................33
Figure 3  Illustration of a multilevel network (compare: [Nes02]).................................36

Figure 4  Examples of transport routes using different levels of speed for private car (left-hand side) and PT (right-hand side); (source: [Nes02]).................................36

Figure 5  Basic matrix OD-zones for the slowest speed level in a three level network (exemplary). For a planning area and its neighboring planning areas, where the grid is finest, the squares are hectares. More far away of these planning areas the grid is getting rougher........................................................................................................38

Figure 6  Functionality of a schedule with integrated cycle times (Timed PT-System) to reduce transfer waiting times (source: [Weid08]).................................39

Figure 7  Assumed access/egress walking path from hectare zone to a link corridor. The parallel part of the access path has an average length of 25% of the on-link stop distance. ........................................................................................................41

Figure 8  Possible access/egress walking paths from a hectare zone to line nodes and link corridors. For the black squared area there are 5 possibilities to access the PT network, three accesses via line nodes and two via link corridors. The OD-zones of link corridors are displayed as red circles, the OD-zones of line nodes as blue triangles and the OD-zones of higher level networks as small black squares (only Winterthurs main station is displayed). ........................................................................................................42

Figure 9  Taxonomy of search and optimization algorithms depending on objective function and strategy. Typical samples are listed as well (compare [Stüm99]). ........................................................................................................45

Figure 10 Structure of a basic evolutionary algorithm. The Greek letters are standing for the number of individuals. For each generation the process (main circle) is repeated. ........................................................................................................48

Figure 11 Two fitness functions; left hand: assignment proportional to the values of the objective function (the gradient of the straight line depends in case of the normalized fitness assignment on the objective function values of the entire population), right hand: the assignment of the truncation-selection (the μ best Individuals of the population get all the fitness value 1, the others are discarded). ........................................................................................................49
Figure 12 Selection: To each individual a fitness proportional segment on a roulette wheel is assigned. In case of roulette selection (left hand), the bowl is operated $\mu$-times. In case of stochastic universal sampling (right hand) an annulus of $\mu$ bowls is operated simultaneously to get several results.  

Figure 13 Recombination of two Individuals $i$ and $k$ (single point crossover).  

Figure 14 Discrete recombination: Recombinants $R_i$ are positioned at the vertexes of the hypercube which is spanned by the parent variables $E_k$ (in two dimensions the hypercube is a rectangle). In case of headway optimization, $x$ and $y$ both are headways for a corresponding line $L_x / L_y$ of the PT network.  

Figure 15 Possible results of variable mutation on an individual. In case of headway optimization, $x$ and $y$ both are headways for a corresponding line $L_x / L_y$ of the PT network.  

Figure 16 Demand potential: demand at each hectare with the demand of its corresponding catchment area. Subdivision of the demand pattern (see Figure 37) from Winterthur, filtered by a responsiveness function (see Figure 19 a). The yellow area contents values over 1'000 moves per day (the maximum is below 5'000).  

Figure 17 Possible line node positions (dark raster squares). Filtered demand (see Figure 16) should be higher than a lower bound demand (here 20 moves per day) in the catchment areas of the line nodes (Demand pattern from Winterthur). Otherwise the raster squares are white.  

Figure 18 Start line node plan 1. Line nodes are set here starting from the left lower corner, filling column after column. Minimal line node distances (between 350 and 800m) are complied with. Other start line node plans are created similar. The background shows the filtered demand (compare Figure 16).  

Figure 19 Responsiveness of people to PT. It is estimated by Gaussian distributions (see [Bränd78]) dependent on walking distances (here linear distances are used) to each potential line node. Line nodes of higher transport speed levels tend to have larger catchment areas (see [Wal73]). a) lowest level, b) middle level and c) highest speed level. The demand pattern (see Figure 37) is equivalent to 100%. A detour-factor of 1.24 [Bränd78] is considered.  

61
Figure 20  Graphical recombination of two line node plans. The red rectangle shows the exchanged area. If parts of the random chosen rectangle leaves the line node plan area, the rectangle is continued at the opposite side (become two rectangles). New line nodes are black, nodes of one old plan green and of the other old plan blue. Around the rectangle some gaps occur, which only partly can be filled..........................62

Figure 21  Mutation of line nodes in a circular area. The red circle shows the area where new line nodes are created around a stochastic placed centre line node. New line nodes are black, line nodes of the old plan green and the circle centre is marked red. Around the circle some gaps occur, which only partly can be filled..........................64

Figure 22  Possible result of single line nodes mutation. New nodes are black and nodes of the old plan are green. The background shows the filtered demand (compare Figure 16)...............................65

Figure 23  Resulted plan after optimizing line nodes. Line nodes (black) in the red rectangle are saved. The other nodes are discarded, to keep influence of borders low..............................66

Figure 24  Result after line node placement. The area of Winterthur has been divided into two speed levels (blue circular and black quadratic line nodes) and several planning areas. The triangle grid is added only for the bus-level, to get an acceptable infrastructure network as base for line network design. The doted rail network including stops was kept..........................67

Figure 25  Test problem (compare [Lee98]). Bus network layout with node numbers and run times on links in minutes. Lengths are not considered. The link connecting node 6 and 9 is a one direction link for bus but a two direction link for light rail. Only the 6 dark black nodes (rings) are potential terminal stations. Demand between all nodes is assumed to be 30 people in the analysis period. Perceived travel times are assumed to be according to section 5.2.4. Operating costs are 130.0 CHF/h for bus and 200.0 CHF/h for light rail. Kilometer costs are not considered. Terminal times of a line are at least 3 minutes..........................69

Figure 26  Sample start network of nearly shortest lines for a test problem (see Figure 25). Only the 6 dark black nodes (rings) are potential terminal stations. Travel times on lines differ from shortest routes by a maximum of 5%. This
makes 36 \textit{nearly shortest routes} (NSR) and 10 lines in the start network (see Table 8).

Figure 27 In a first reduction step for the test problem (see Figure 25) line T1 is kept in a second step the lines B2 and B6 are merged using ACO (see below) to a new line B2.

Figure 28 In a third reduction step for the test problem (see Figure 25) lines B1 and B8 are deleted.

Figure 29 In a 5th reduction step for the test problem (see Figure 25) line B7 and B4 are merged to a new line B7.

Figure 30 In the 7th reduction step for the test problem (see Figure 25) lines B3 and B9 are merged to a new loop line B3.

Figure 31 In the 8th reduction step for the test problem (see Figure 25) line B7 is deleted.

Figure 32 Passenger flows in both directions of the lines of the 8th reduction step for the test problem (see Figure 25).

Figure 33 Possible on-link stop positions (dark raster squares, compare Figure 17). Filtered demand (see Figure 16) should be higher than a lower bound demand (here 10 moves per day) in the \textit{catchment areas} of the possible stop positions (Demand pattern from Winterthur). Otherwise the raster squares are white. Around the red \textit{line node} circles minimum stop distances are considered.

Figure 34 Existing stop placement for the PT line network 2008 with passenger flows ($f_{\text{cost}} = 1.0$, 165 stops).

Figure 35 New automated stop placement on \textit{link corridors} for the PT line network 2008 with passenger flows ($f_{\text{cost}} = 1.0$, 136 stops).

Figure 36 Passenger load and transfers in the PT network of Winterthur realized in 2008 during the morning peak hour, using headway-based assignment.
Figure 37 Demand pattern of Winterthur. “Demand” is here the sum of trip origins and trip destinations at each hectare. The OD matrix of Winterthur is derived from the transport model of the Canton of Zurich [Vrtic05].

Figure 38 General load profiles over the daytime according to trip purposes for car and PT together in the city of Zurich (generated from [CEOZ07]).

Figure 39 Most loaded sections of the lines north-west branches during morning peak hour (PT network 2007, source STADTBUS Winterthur).

Figure 40 Case 0: The PT network of Winterthur, realized in 2008.

Figure 41 Case 2: The PT network of Winterthur, realized in 2008 after headway optimization (see also Table 23). Assignment is based on line nodes and link corridors.

Figure 42 Case 4: PT network design with lines from network 2008 after long calculation (see also Table 24).

Figure 43 Case 6: Network design without existing lines (see also Table 25).

Figure 44 Magnetic streamlines of the field displayed with small iron particles (source: Practical Physics, publ. 1914 by Macmillan and Company).

Figure 45 Case 5: Sequential network design using two overlapping planning areas, with lines from network 2008. Firstly a PT network was designed in Planning area 1 on the left side of the entire area.

Figure 46 Case 5: Sequential network design using two overlapping planning areas, with lines from network 2008 (see also Table 26). In a second step a network for nearly the entire area was designed in Planning area 2. Lines on the shaded part at the left side are partly fixed.

Figure 47 Case 10: Network design with existing lines (see also Table 30). Cost balance factor $f_{cost} = 6$. Maximum vehicle load is one standing passenger/m².

Figure 48 Case 8: Network design on an artificial infrastructure (see also Table 33); Maximum bus speed is 50km/h.
Figure 49  Total travel time on the bus network versus number of vehicles \( (v_{\text{max}} = 2/3 \ v_{\text{car}}, f_{\text{cost}} = 1) \). The design cases are listed in Table 34..............................115

Figure 50  Cost comparison \( (v_{\text{max}} = 2/3 \ v_{\text{car}}, f_{\text{cost}} = 1) \). The design cases are listed in Table 34................................................................................................. 115

Figure 51  Mandl’s Swiss benchmark problem [Mandl79]. a) Network layout with node numbers and run times on links in minutes. b) OD matrix (Table 35) assigned to network links on shortest routes. If routes have the same travel time, demand is assigned to one of the routes..............................118

Figure 52  GSSH Comparison 1; network design solution for Mandl’s sample network (see Figure 51)..........................................................................................................................121

Figure 53  Lee’s [Lee98] network design solution for Mandl’s sample network (see Figure 51). ...........................................................................................................................122

Figure 54  Case 2 of Shih and Mahmassani [Shih94]; network design solution for Mandl’s sample network (see Figure 51). ..............................................................123

Figure 55  Baaj and Mahmassani [Baaj91] Case 1; network design solution for Mandl’s sample network (see Figure 51). .............................................................124

Figure 56  Mandl [Mandl79]; network design solution for Mandl’s sample network (see Figure 51). ........................................................................................................125

Figure 57  GSSH Comparison 2; network design solution for Mandl’s sample network (see Figure 51).....................................................................................126

Figure 58  Zhao and Gan’s [Zhao03] Case S&M2; network design solution for Mandl’s sample network (see Figure 51).........................................................127
Abbreviations

ACO ant colony optimization (see sections 3.5.4 and 4.3.6)
ARE “Bundesamt für Raumentwicklung”, federal office for spatial development
ae access, egress (time; costs)
aw access wait
BFS “Bundesamt für Statistik”, federal statistical office
COM component object model
e external (costs)
GA genetic algorithm (see section 3.5.3)
GSSH guided stochastic search heuristic (see section 4.3)
IV in-vehicle (time; costs)
m other modes (costs)
NSR nearly shortest routes (see section 4.3.2)
OD origin-destination
p passenger (costs)
PT public transport (also called “transit”)
PTS potential terminal stations (see section 4.3.2)
sp service provider (costs)
Transf transfer
TransfW transfer wait
UpB upper bound
w Wait
ZVV “Zürcher Verkehrsverbund”, transport association of the Canton of Zürich
Glossary

Basic zone hectare, area: 100m long and 100m wide

Catchment area a catchment area is the area from which the PT service attracts a specific amount of customers (synonym: coverage area)

Cycle time is the scheduled time interval between a departure of a PT vehicle at a terminal station and its next departure at the same terminal after serving the line in both directions. Terminal times are included (synonyms: operating time, over all travel time)

Line A line is defined by its terminal stations, the route or a sequence of routes and the stops.

Line node at a line node it is possible that different PT lines come together, cross each other or separate enabling passenger to transfer. (synonym to “Transfer Node”)

Link corridor a link corridor is the catchment area of a PT link between two line nodes

Nearly shortest route NSR (see section 4.3.2)

Optimization period Optimization period is the basic period (the longest headway; e.g. one hour). This could be either an average peak period or an average off-peak period.

Origin-destination OD zone: An area or node in the transport model where trips start or end. OD matrix: Matrix of flows between OD-zones in a defined time period.

Potential terminal station (PTS, see as well section 4.3.2) At a PTS it is possible for PT vehicles to change to the opposite direction.

Route a route connects two potential terminal stations. Each route has its respective opposite directed route (see as well “Nearly Shortest Route”)

Stop PT Station (location where passengers can board or alight PT vehicles)

Transfer node Stop with possibilities to transfer between PT lines (synonym to “Line Node”)
1 Introduction and Objectives

1.1 Introduction

1.1.1 Public Transport Systems

Public transport (PT) is an essential means of achieving sustainable economic development for cities and regions. PT reduces energy use, pollution and the space required for mobility. In densely populated areas, PT provides a high level of mobility at low external costs. Due to different reasons most PT systems cannot cover their full operating costs through fares. Therefore, PT systems must improve their efficiency and become more economical. An important technique for improving PT efficiency is to optimize the public transport network structure, thus enabling PT service providers to offer high-quality service at low cost.

Designs of existing PT networks are very complex due to the interplay of market orientation, operating costs and investments. For economic reasons, services are bundled spatially into lines and temporally into frequencies or schedules. The result is a network that is highly space and time dependent.

The complexity of PT systems as part of urban areas makes it extremely difficult to design attractive and efficient PT networks. First, there are four main influences on PT (compare [Weid08] (2) section 3.2.5.1.1):

- Customers want a fast and reliable journey from anywhere to everywhere at anytime at a low price with a minimum number of transfers.

- Service providers (operators, authorities, etc.) want to minimize costs while satisfying transport demand as much as possible.

- Competing modes can be chosen as alternatives and thus define the required level of PT service to achieve a desired PT market share.

- Policy-makers consider many different goals, including economic, social, environmental and urban planning, when making PT planning decisions.

Second, urban areas are places with existing transport networks and urban patterns. These constraints cannot be ignored when designing totally new transport networks. In addition to
political obstacles to change, previous investments in existing infrastructure would be lost (sunk costs). Third, the multi-objective function of PT network design with many degrees of freedom is non-linear, combinatorial, partly discrete and continuous (compare section 3.5.1). Finally, the reaction of people to an updated PT supply, as well as the costs for new PT infrastructures, can only be estimated and therefore involves uncertainties.

### 1.1.2 Major Questions

Knowledge of optimal network structures in PT should be extended in answering following questions:

- What a PT network structure would look like, if infrastructure could be newly designed?

- What kinds of line alignments are optimal in PT networks if only few restrictions are applied?

- How line headways are interrelated with network structures?

- How many speed levels (i.e. different fast PT modes; see section 3.3.1) would be optimal for an existing demand pattern and what speeds would be favorable?

### 1.1.3 Software Prototype

Aims of the method are:

- Cost optimal PT network designs (functionality).

- Modular structure of the algorithm, allowing to start the process from different start conditions including intermediate results (flexibility).

- Some input/output-data are optional and can be included in different forms (flexibility).

- Computing times are adjustable, to get results in a reasonable amount of time. Herby, the quality of results is influenced on the other side as well.

The software prototype can be downloaded from http://www.ivt.ethz.ch/docs/institute/sr150.zip.
1.1.4 Organization of the Chapter

In the following, some main design objectives concerning PT as part of the entire transport system are introduced and a generalized costs model is composed based on those objectives (section 1.2). In section 1.3 a reduced generalized costs model is composed for this study including constraints. Limitations of this study are discussed. Finally a short overview of the dissertation is given in section 1.4.

1.2 Design Objectives and Generalized Costs

1.2.1 Overview

The four main influences on PT as listed in the introduction entail a multitude of objectives (indicators). The objectives contradict each other in parts. Therefore it is necessary to find a compromise.

To pick out only one important objective would cause biased PT networks (compare [Weid08] (2) section 3.2.5.1.5). If, e.g., only

- average travel times are minimized it results in a huge number of PT vehicles providing short headways and direct connections, if only
- direct traveling passengers are maximized it results in detours and high operating costs and if only
- operating costs are minimized it results in no PT service at all.

Such extreme results can be avoided in setting constraints for the other objectives. To look for suitable constraints is another optimization problem though and the search space is being reduced in setting constraints.

A pareto-optimal solution is only possible for objective pairs which do not contradict each other in the entire search space.

The only way to find a true compromise is to make the different objectives comparable. The difficulty is to find reasonable weights or weight-functions for the single objectives. The advantage is that the search space is not restricted.
The last mentioned approach is used in this study. Each single objective is transferred to a monetary dimension in a defined time period, and so a generalized cost model may be created which should be minimized:

\[ C_{\text{total}} = C_p + C_{sp} + C_e + C_m \rightarrow \text{minimize} \]  

(1)

(notation: \( p \) – passenger; \( sp \) – service provider; \( e \) – external; \( m \) – other modes)

The defined time period is necessary to compare long term costs like investments with short term costs like costs for travel time.

The main input in PT network evaluation is the transport behavior of passengers, imitated by the transport model. It can not be influenced directly by the transport supply. It depends on the requirements of passengers and the available transport possibilities. It is important for PT system optimization to model passenger behavior accurately, because passenger time costs as well as service provider costs in PT strongly depend on it.

The system size of the transport model defines on the one hand the objectives which can be included into PT optimization on the other hand it is responsible for the complexity of the optimization problem (required data, computing times, possible accuracy of the results). For example it is possible to roughly model the interactions between land use, transportation, and public policy to get some impressions about urban long term developments [Opus09]. Or it is possible to only focus on PT neglecting modal split and so to get more accurate results on PT-systems themselves without considering interactions with the larger context in which PT is embedded.

The following subsections introduce the main cost fractions in PT network design. In section 1.3 the selected cost fractions in this study are presented. To select all of them would make the analyzed system too large. This would have exceeded the study’s scope of work and the computability of the optimization.

1.2.2 Customers

Objectives

Besides more qualitative goals like a comfortable ride, customers’ primary objectives are a fast and reliable journey at a low price and with few transfers. They do not want to spend a lot of time in studying schedules, on purchasing tickets and on check-in procedures. If problems occur, they want to be well informed and get some alternative solutions about how to arrive at their destination. Qualitative goals from customer side can be found more detailed in
Investigation of space-time structures in public transport networks and their optimization

the European Standard EN13816 [CEN02]. Objectives relevant for PT network design are included into the generalized costs.

**Generalized Costs**

Generalized passenger costs are expressed here as the weighted sum of all single passenger costs:

\[ C_p = \sum_{\text{passenger}} \left\{ n_{\text{time}} \left( a_{ae} \cdot t_{ae} + a_{aw} \cdot t_{aw} + a_{iv} \cdot t_{iv} + a_{T_{\text{TransfWalk}}} \cdot t_{T_{\text{TransfWalk}}} + a_{T_{\text{TransfW}}} \cdot t_{T_{\text{TransfW}}} + a_{T_{\text{TransfPen}}} \cdot t_{T_{\text{TransfPen}}} \right) \right\} + C_{\text{ticket}} + C_{\text{rely}} \]  

(2)

(notation: \( n_{\text{time}} \) – value of travel time savings; \( t \) – time; \( a \) – weight factor; \( ae \) – access/egress; \( aw \) – access wait; \( iv \) – in-vehicle; \( T_{\text{Transf}} \) – transfer; \( T_{\text{TransfWalk}} \) – transfer walk; \( T_{\text{TransfW}} \) – transfer wait; \( T_{\text{TransfPen}} \) – transfer penalty; \( rely \) – reliability)

Each person evaluates components of journey time in a different way, depending on availability of time, trip purpose, total journey time, mode, etc. A common way to evaluate travel time is to assign to each time component a different weight or weight function. Waiting or walking times are considered as longer than in-vehicle travel times unless it is very crowded in a vehicle (see elasticity functions for travel attributes [Vrtic06] and [Vrtic05]).

Time costs compared to other (dis-) utility costs are evaluated individually as well. The ratio of income to availability of time is quite important. None the less it is common to survey average cost values of travel time savings \( n_{\text{time}} \) (for Switzerland see [Axhau06]). Such values, despite being only virtually, can be used to compare differences in passenger costs of entire transport systems due to changes of transport supply.

Variations of expected arriving times, the so called reliability, can also be expressed in disutility costs \( C_{\text{rely}} \) (for Switzerland see [König05]) being composed of one part for the delay amount (additional costs due to the delay) multiplied by a factor for the delay probability. Fares \( C_{\text{ticket}} \) are already expressed in money. In case of single trip tickets the trip costs are directly given by its prices. In case of flat fares for a defined time period they can be calculated in several ways according to available data and average PT use.

For other quality measures (e.g. see the European standard EN13816 [CEN02]) it is more difficult to find appropriate cost values. Beside the effect of crowding, local influences of such measures on network design are quite low. In inter modal comparisons the most important of
them should be considered though. They could be added to the generalized costs function as penalty terms.

1.2.3 Service Providers

Objectives

Service providers (operators, authorities etc.) want to minimize costs while satisfying transport demand as much as possible and to optimize their cost-benefit ratio.

Generalized Costs

Total costs of services can be calculated easily. They are listed in annual reports. In the planning phase service provider costs can only be estimated.

\[ C_{sp} = \sum_{\text{vehicles}} \left( C_h + C_{km} \right) + C_{\text{infrastructure}} \]  

(3)

(notation: \( sp \) – service provider; \( h \) – vehicle hour; \( km \) – vehicle kilometer)

Total operating costs consist mainly of wages for operating personnel, energy costs, rolling stock acquisition and maintenance, insurances, taxes and overhead costs. For network optimization purposes it is not necessary to reproduce the entire structure of a transit agency. In most cases it is sufficient to assign typical costs to a service vehicle km and a service vehicle hour (compare [Frank08]). In this case, investment costs for rolling stocks and overhead cost are included in the cost rates per revenue service kilometer and per service hour.

A main problem of estimating network costs is that (capital) investments can differ a lot, depending on the planned transport infrastructure. If a tunnel for example is necessary, costs would be more than 10 times higher than for a rail track on level (compare [Weid07a]). During cycle time calculations it can be necessary to consider further constraints like minimum rest times of the crew [Lieb08]. In doing so, later cost surprises during planning shifts for the crew (crew scheduling) can so be avoided.

1.2.4 Competing Modes

Effekts on PT

For short distance trips people mostly walk or use bikes. The longer distances are the more people go by car, PT and for the longest distances by airplane.
Competing modes can be chosen by potential customers as alternatives and so they define the level of PT service needed to achieve a desired market share. If changes in modal split should be considered in PT network design, competing modes have to be included into the transport model. Long term effects such as changes in consumer behavior or changes caused by land use regulation could influence transport demand as a whole.

**Generalized Costs**

Costs for car use are operating costs, time costs (values of travel time savings; for Switzerland see [Axhau06]) and infrastructure costs. Infrastructure costs can differ a lot, similar to infrastructure costs of PT.

\[
C_{car} = \sum_{car\_user} \left[ v_{time} \cdot \left( a_{ae} \cdot t_{ae} + a_{iv} \cdot t_{iv} \right) + C_{\text{rely}} \right] + \sum_{\text{vehicles}} \left( C_h + C_{km} \right) + C_{\text{infrastructure}}
\]  

\( (\text{notation: } v_{time} \text{ – value of travel time savings; } t \text{ – time; } a \text{ – weight factor; } ae \text{ – access/egress; } IV \text{ – in-vehicle; } Transf \text{ – transfer; } TransfWalk \text{ – transfer walk; } TransfW \text{ – transfer wait; } TransfPen \text{ – transfer penalty; rely \text{ – reliability; } h \text{ – vehicle hour; } km \text{ – vehicle kilometer}) \) 

Analogue general costs can be found for walking and cycling.

**1.2.5 Politics**

**Objectives**

Politics have a powerful influence on transport. In Friedrich’s words [Fried06]: “It needs to be accepted that major changes in mode choice can only be achieved by political decisions which for example discourage the use of private cars in urban areas or for long distance travel by prices or parking limitations”.

Policy-makers consider many goals including economic, social, environmental and urban planning oriented ones. The most relevant goals for transport planning and operation are listed in Table 1 with the respective indicators. Similar goals can be found e.g. in the “Green Paper – Towards a new culture for urban mobility “by the European Commission [COM07].

Some of the goals should be integrated into line network design as constraints; others could be part of the objective function. The question of the weight of every single goal is though a political one.
<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure</th>
<th>Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Security policy</strong></td>
<td>Reduce traffic deaths</td>
<td>Traffic deaths/a (no crimes, no accidents, emergency management)</td>
</tr>
<tr>
<td></td>
<td>Reduce traffic victims</td>
<td>Traffic victims/a</td>
</tr>
<tr>
<td></td>
<td>Reduce vandalism and aggressive attacks to people</td>
<td>Reported cases/a</td>
</tr>
<tr>
<td><strong>Economic policy</strong></td>
<td>Enhance the performance of transport infrastructure</td>
<td>People/h at all relevant cross sections (all modes) relative to performance.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Means to enhance combined mobility (park + ride, bike + ride, information offer etc.)</td>
</tr>
<tr>
<td></td>
<td>Ensure reliability</td>
<td>Congestion hours in car transport</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Punctuality in PT/ constant average speed in PT</td>
</tr>
<tr>
<td><strong>Social policy</strong></td>
<td>Enhance quality of life in public space</td>
<td>Public space and space for slow modes</td>
</tr>
<tr>
<td></td>
<td>Enhance availability of PT</td>
<td>Amount of people (residents/working places etc.) in catchment areas</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transit service in catchment areas (operating hours, operating frequencies, transfer connections)</td>
</tr>
<tr>
<td><strong>Health policy</strong></td>
<td>Comply with air pollution threshold</td>
<td>Reduction of air pollution (PM10, CO₂, etc.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Amount of affected people (% of population in area)</td>
</tr>
<tr>
<td></td>
<td>Comply with noise threshold</td>
<td>Amount of affected people (% of population in area)</td>
</tr>
<tr>
<td><strong>Environment policy</strong></td>
<td>Reduce emissions (exhaust, noise, vibration)</td>
<td>Share of vehicles with little exhaust emissions (%), share of car free households (%)</td>
</tr>
<tr>
<td></td>
<td>Reduce energy use</td>
<td>Average energy use for mobility per person kilometer</td>
</tr>
<tr>
<td><strong>Land use regulation policy</strong></td>
<td>Steer land use regulation</td>
<td>Ratio: transport area to settlement area (%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Population density (residents/ha)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transport area per transport mode (m²)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>From PT covered settlement area per quality category (m²)</td>
</tr>
<tr>
<td><strong>Finance policy</strong></td>
<td>Improve microeconomic costs</td>
<td>Microeconomic revenue cost ratio (all transport modes)</td>
</tr>
<tr>
<td></td>
<td>Adequate transport expenditures</td>
<td>Transport expenditures per resident</td>
</tr>
</tbody>
</table>

Table 1 Objectives of transport policy in agglomerations.
1.3 Objectives and basic Methods of the Study

1.3.1 Overview

The objective of this research project is to develop a method for PT network optimization which considers the entire PT chain from door-to-door. It aims at extending knowledge of optimal network structures in PT. For this purpose, a software prototype for planning and improving PT network design should be created.

The approach shall make several improvements over existing ones.

1. It shall integrate stop placement and scheduling into line planning. Several scientific papers make recommendations for further research in this direction (e.g. Bussieck et al. 1997 [Bus97], Beilner et al. 2001 [Beil01]).

2. It shall be able to use more accurate transport models without considerably increasing computing times.

3. It shall expand the design process to include more than one speed level (i.e. different fast PT modes; see section 3.3.1).

4. It shall be based on a PT related total cost approach.

Based on the objectives of section 1.2, a reduced generalized cost model with constraints is composed for this study in the following section. In section 1.3.3 the limitations of this study are discussed.

1.3.2 Generalized Costs Model for this Study and Constraints

The objective function for the proposed approach is to minimize the PT system’s generalized costs. Variables are the paths of the lines in the PT network (line alignments \( l \)) and the time intervals between two services offered on the lines (headways \( h \)). The total cost function (compare function (1)) is based upon passenger costs \( C_p \) and service provider costs \( C_{sp} \) in a defined time period:

\[
C_{total}(l, h) = C_p \sqrt{f_{cost}} + C_{sp} \rightarrow \text{minimize} \quad (5)
\]

Consequently, demand for PT is fixed. Competing modes (see section 1.2.4) are not considered. Although demand, and thus modal split, are fixed, it is none the less possible to ratio-
nalize PT networks under a given demand structure. Assignment of the origin-destination (OD) matrix to the PT lines, however, can be influenced by PT system design. This is done to focus on PT systems and to reduce computing times for network evaluations.

Some costs are not considered in the total cost function (5). On the service provider side, it is mainly costs for infrastructure investments and external costs that are neglected (compare next section as well). To keep the ratio of passenger costs and service provider costs in the right balance, a cost balance factor $f_{\text{cost}}$ is introduced (for values see section 5.2.4). The cost balance factor can also be used as weight factor to change the importance of service provider costs compared to passenger costs, which is a political question.

Passenger costs are reduced to perceived travel times $a \cdot t$ [min] weighted by the value of travel time savings $\nu_{\text{time}}$ [CHF/min].

$$C_p(l, h) = \nu_{\text{time}} \cdot \sum_{\text{passenger}} \left( a_{ae} \cdot t_{ae} + a_{aw} \cdot t_{aw} + a_{iv} \cdot t_{iv} + a_{\text{TransfWalk}} \cdot t_{\text{TransfWalk}} + a_{\text{TransfW}} \cdot t_{\text{TransfW}} + b_{\text{TransfPen}} \cdot n_{\text{Transf}} \right)$$  \hspace{1cm} (6)

(notation: $ae$ – access/egress; $aw$ – access wait; $iv$ – in-vehicle; $\text{Transf}$ – transfer; $\text{TransfWalk}$ – transfer walk; $\text{TransfW}$ – transfer wait)

Travel times $(t_{ae}, t_{aw}, t_{iv}, t_{\text{TransfWalk}}, t_{\text{TransfW}})$ and the number of transfers $n_{\text{Transf}}$ are calculated by the assignment method (see section 3.4.3 and [PTV07]). The same weight factors $(a_{ae}, a_{aw}, a_{iv}, a_{\text{TransfWalk}}, a_{\text{TransfW}})$ are used for evaluations and during the assignments. They take psychological aspects into account, e.g., that time is perceived to pass slower while waiting for a vehicle than while riding. The transfer penalty $b_{\text{TransfPen}}$ penalizes the inconvenience of transfers. The average density of standing passengers must not exceed a given limit for the lines section of maximum load. The evaluation neglects PT fares and stochastic variations in passenger arrival times.

Service provider costs are reduced to service vehicle-kilometer costs $C_{km}$ and service vehicle-hour costs $C_h$.

$$C_{sp}(l, h) = \sum_{\text{vehicles}} \left( C_h + C_{km} \right)$$  \hspace{1cm} (7)

The goal to guarantee access to PT stops within a given limit can be complied with (if infrastructure allows it) setting a maximum linear distance of OD-zones to stops.
Line length margins and *cycle time* margins can be given. Possible line headways as well as vehicle sizes, vehicle kilometer costs and vehicle hour costs can be given as a list of discrete values. Average link speeds of the mixed traffic are a necessary input.

1.3.3 **Simplifications and Limitations of the Generalized Costs Model**

The following assumptions are made for this study:

- **Schedule-based assignment**

  *Assumption:* The assignment of OD-pairs to the resulting line network used in this study is a headway-based stochastic multiple route assignment (see section 3.4.3 and [PTV07]).

  *Reason:* A formal schedule-based assignment would require a schedule optimization preceding any network evaluation to get a fair, unbiased comparison between network variants. In addition to the slower schedule-based assignment, schedule optimization would increase computing times.

  *Consequences:* Due to headway-based travel time calculation, only schedule independent (uniformly distributed) passenger access and no schedule coordination at transfer points is considered (compare [Weid07b]). The demand model could become more accurate if schedule-based assignment would be used. In contrast using headway-based assignment has the advantage that the resulting schedules are more robust to delays. Schedules of different lines don not need to be synchronized in any case. While analyzing different network designs, schedule coordination should be taken into account. Suggested network solutions of this study are more of theoretical value as long as headways of several multiple connected lines could be synchronized with each other.

- **Reliability**

  *Assumption:* Using a headway-based assignment delays are not considered, but headways should be kept.

  *Reason:* private transport is not part of the transport model

  *Consequences:* they can be neglected because a schedule is made for average *cycle times*, and interconnected lines are not explicitly considered (headway-based assignment of demand). In case of schedule-based assignment stochastic varia-
tions of expected arriving times would be important for network evaluations (time costs).

- **Trip profile over the daytime**

  *Assumption:* Optimization period is the basic period (the longest headway; e.g. one hour). This could be either an average peak period or an average off-peak period.

  *Reason:* This is to save computing time, which would nearly double if two optimization periods were to be considered.

  *Consequences:* vehicle scheduling can not be included into PT network optimization. Supply can not directly be adapted to the trip profile over the daytime.

- **Real access and egress paths**

  *Assumption:* Instead of real access and egress paths, linear access and egress distances are used multiplied by a detour factor.

  *Reason:* it is still difficult to get digital maps with accurate footpaths

  *Consequences:* if direct footpaths are not available, OD-pairs could be assigned to different PT stations than in reality with consequences for time costs and vehicle loads.

- **Fares and subsidies**

  *Assumption:* It is assumed that operating costs can not totally be covered by fares. The difference is covered by subsidies. So, fares plus subsidies are per definition equal to operating costs.

  *Reason:* As operating costs, fares and subsidies are already included in the total cost function (5).

  *Consequences:* fares and subsidies are important for financing PT systems. In case of area fees, they do not have a strong spatial influence on PT network designs. Modal split is fixed thus fares are not necessary for influencing modal split.
• **Costs for investments in infrastructure**

  *Assumption:* costs for investments in infrastructure could be included implicitly as operational costs (average costs for service km).

  *Reason:* investments in infrastructure vary by at least a factor 10 depending on circumstances and types (compare [Weid07a]) but are not easily to be estimated. They strongly depend on the local situation. A general estimation whether new bridges or tunnels are necessary or not requires diligent investigations in the context of network design. Amortizations and costs for maintenance are not considered either as long as they are not included into km costs as track fees. Given these obstacles, this study has been restricted to the operational costs. Investments in infrastructures have been excluded.

  *Consequences:* investments in infrastructure are an important part of PT system costs, especially in case of railroads, and can not be easily neglected in PT network design. While analyzing different network designs, variations of infrastructure investments should be taken into account. If possible, average investments or better estimations of infrastructure costs of all potential infrastructure elements (if known) should be added into the evaluation of network designs. Suggested network solutions based upon the method of this study have to be evaluated in later steps with regard to their infrastructure needs.

• **Deadhead (non-revenue) kilometers**

  *Assumption:* Deadhead kilometers are not considered

  *Reason:* to estimate vehicles deadhead kilometers the entire vehicle scheduling must be performed including adapting the schedules to load profiles over the daytime (see Figure 38).

  *Consequences:* compared to the amount of revenue service kilometers deadhead kilometers can be neglected if the locations of vehicle depots are chosen well.

• **Minimal rest times of the crew**

  *Assumption:* minimal rest times of the crew are not considered

  *Reason:* minimal rest times of the crew in the Canton of Zurich (between 40 and 360 minutes) are too long to be considered directly in the schedule. If necessary, line lengths can be limited to allow flexible crew scheduling.
Consequences: If minimal rest times of the crew are quite long they can be neglected.

The above mentioned limitations could influence, in the worst case, the balance between passenger costs and service provider costs. Variations of infrastructure investments and neglecting schedule coordination could have a strong spatial influence on PT network design though and should be kept in mind while analyzing different network designs.

1.4 Overview of the Dissertation

Till now already many methods for PT network design have been suggested and documented. That is why a short review is undertaken in chapter 2.

The methodological approach of this study is discussed in chapter 3. In section 3.2 the conceptual procedure for the design process is presented, which is the base for the subprocesses described in chapter 4. The space and time approach of this study is presented in section 3.3. In section 3.4.3 a short introduction to the assignment methods is given which are used for this study. One is the headway-based stochastic multiple route assignment which is part of the PTV tool VISUM [PTV07]. The proposed network design optimization approach shall apply stochastic optimization techniques i.e. principles from genetic algorithm and ant colony optimization algorithm. The according basic heuristics are introduced in section 3.5. In section 3.6 the implementation of the PT network design process is outlined.

To evaluate the optimum, reached by using this integrated design procedure, the optimum can be compared with an ideal PT network i.e. one with direct point-to-point links and in which stops can be placed anywhere, with an existing network or with a network created by a different procedure. The design process for an ideal PT infrastructure network is presented in section 4.2. The design process for a PT line network based on an existing or beforehand designed infrastructure and a headway-based assignment is presented in section 4.3.

The design process is tested by applying it on case studies, which are presented in chapter 5. The design process is applied with variable settings to an urban area: the city Winterthur, Canton of Zurich and compared with the actual network. For this sample the PT model of the Canton of Zurich [Vrtic05] is adapted. Further, the core of the design process is compared with the results of different approaches using the small benchmark problem of Mandl [Mandl79].

In chapter 6 some conclusions are drawn.
2 Literature Review of Line Network Design

2.1 Introduction

2.1.1 General Aims

Intuition and experience of planners and practical guidelines still dominate real world network and schedule generation until today (see [Weid08] (2) section 3.2.5). Planners use more and more software for network analysis, vehicle scheduling and crew scheduling though. This for several reasons: besides considering many different requirements to transit systems a planner is able to evaluate line plans and schedules from the passenger, as well as from the service provider’s point of view. Even if total costs of a numerically designed PT network are low, it might be e.g. confusing from passenger’s perspective (usability). Often knowledge of local PT features is not systematically collected and therefore not available for mathematic optimization.

Figure 1 Basic network design approaches.

This study focuses on mathematical approaches (compare Figure 1) for network design which should provide the following advantages:
• Solutions should only depend on network design methods and their constraints excluding other influences, especially historical traditions of network design.

• They should be able to better cope with a huge search space compared to manual planning.

• They should provide similar good solutions for the entire area of investigation. A planner dealing with complex problems tends to focus on sub problems.

2.1.2 Basic Mathematical Approaches

There are two basic mathematical approaches for network design: analytical and numerical (see Figure 1).

An analytic approach can with some simplifications be used to optimize single PT lines and their specific parameters ([Nes02] and [Schäf05a]). However, it has not been possible to obtain feasible results applying this approach for designing multiple line networks (i.e. to solving the problem in two dimensions). Although such approaches are strongly simplified, they can nevertheless provide general information on how to design efficient multi-line networks but they can not be used for PT network design directly.

Numerical approaches, in contrast, are generally used to design networks of multiple PT lines. These can be divided into two groups.

• The first consists of optimizing the entire network in one step (([Biel02], [Born08a], [Ritt09], [Zhao08])).

• The second consists of designing PT networks in a rather local heuristic optimization procedure ([Ceder07], [Lee98]).

Numerical approaches should be able to provide the three advantages of mathematical approaches as listed above.

Former numerical approaches for designing PT networks generally focused on sequentially optimizing different sub-problems, i.e., stop placement, line network design, scheduling, etc.). Beilner et al. [Beil01], e.g., developed and implemented a sequential method for line network design using at first an evolution strategy for stop placement, than the heuristic line building method of Sonntag [Sonn77] for line network design and finally the method of Weigand [Wei81] for scheduling. For every sub-problem there are further papers and reports describ-
Recent approaches try to include several sub-problems into one optimization procedure. Rittner and Nachtigall [Ritt09] and Zhao and Zeng [Zhao08] e.g. combine line network design and scheduling.

Some selected studies are reviewed in the following by applying 9 criteria (see Table 2 to Table 3):

1. Objective criteria and constraints (costs of travel time, transport fares, operating costs, infrastructure costs, modal split, etc.)

2. Integrated sub-problems (stop placement, scheduling, scheduling of vehicles and crew)

3. Origin-destination (OD) matrix (fixed, dynamic)

4. Used assignment method (all-or-nothing, multiple route, stochastic, logit, etc.)

5. Precision in space (address fine, hectare fine, stop fine, intersection fine, etc.)

6. Precision in time (system-based assignment, headway-based assignment, schedule-based assignment)

7. Optimization method (heuristic, search technique, exact optimization)

8. Integrated levels of speed

9. Ability to include given lines or line elements.

Reviews of additional network design studies can be found in [Axhau84], [Ceder07], [Gui08], [Kep09], [Zhao03] and others.

### 2.2 Analytical Network Generation Studies

Ulrich Schäffeler evolved in his PhD thesis [Schäf05b] an analytical method for PT network design. This method is more general. It was aimed at getting optimal design parameters of two parallel lines.
Objective criterion is to maximize subjective supply quality (passenger share). To maximize supply quality means to reduce

- access and egress time,
- headway and
- PT to car transport door-to-door travel time ratio.

Out of these objectives a logit model is composed, which split modes (car and PT). Operator costs are considered by fixing the number of vehicles (fixed vehicle density).

Potential customers are assumed to live uniformly distributed in catchment areas. Due to headway-based travel time calculation only schedule independent (uniformly distributed) passenger access to the stations is considered. These simplifications conjoint with the only partly realistic objective function limit this models application to rather theoretical samples. The exact integration over the catchment areas though is quite interesting for some basic investigations. Stop setting and headway optimization are integrated into line parameter optimization and express services are discussed.

Main results of Schäffelers work are:

- Optimal distances between stops on a line should generally be shorter than distances between PT lines.
- PT network design should be based on few routes with high speeds and short headways.
- To get efficient PT systems, demand density should exceed a certain level and land use should be adapted to the networks structure.

### 2.3 Simultaneous Network Generation Studies

#### 2.3.1 Overview

If a set of PT lines is given in advance, an optimal selection can be found in a simultaneous optimization process. For such optimization several optimization methods can be used. In the next sections, examples with genetic algorithm, integer programming and simulated annealing are presented.
2.3.2 Genetic Algorithm

Maurizio Bielli et al. [Biel02] evolved and implemented a numerical method for PT network design using a genetic algorithm. As multi-objective function they used the sum of (negative and positive) weighted indicators (i.e.: number of passengers, total travel time, total number of transfers, crowds index, vehicle number, etc.). The transport model calculates modal split between public transit and walking. A network equilibrium model (modes: walking and PT) is used for assignment of the OD matrix to lines. A pool of PT lines is defined.

In an example using the city of Parma (Italy), the given line pool consists of 80 lines. It consists of the existing lines, of shortest lines connecting the main OD-Pairs, (manually designed) half circular through-lines and (manually designed) ring lines. A small number of additional lines can easily be included. The genetic algorithm optimizes line selection and line headways.

Limitation: The used network equilibrium model (modes: walking and PT) may only be appropriate if crowding in-vehicles is considered and relevant in the considered network. Otherwise a network equilibrium model is in the best case an all-or-nothing assignment. Due to headway-based travel time calculation only schedule independent (uniformly distributed) passenger access and no schedule coordination at transfer points is considered. The genetic algorithm varies line selection and headways. If the line pool is large enough, scheduling is considered to a certain extend. That means a balance between headway reduction and shorter routes is achievable in order to reduce travel times. In the used case study, the small size of the given line pool (80 lines, [Biel02]) indicates that it takes a long time to compute the objective function. With more lines in the line pool results could further be improved.

2.3.3 Integer Programming

Ralf Borndörfer et al.

Ralf Borndörfer et al. [Born08] evolved and implemented a numerical method for PT network design using mixed integer linear programming (see also [Born04]). Objective function is the weighted sum of total riding times and service provider costs. On a basic level operating and infrastructure costs are considered. The OD matrix is fixed. Several speed levels can be considered. Stops and potential terminal stations are considered as given. Line length margins are optional. First a pool of PT lines is designed.

In a case study using the city of Potsdam (Germany), the pool consists of the 86 existing lines. Additionally lines are added connecting up to 56 stops. If only the shortest paths are considered 1537 lines are added. If the allowed deviation in length from the shortest route is
increased the line pool contains more and more lines, finally up to 30’000. Computing times are quite short.

Limitation: Especially the PT system-based assignment according to system optimum makes the transport model artificial. If assignment is PT system-based headways or schedules are not considered. Assignment is performed according to system optimum. If the objective function can be reduced, passengers are even assigned to circuitous routes. In the used case study, Borndörfers et al.’s method can handle 111 OD-zones and up to 30’000 lines in the line pool (computing times of around 12 hours). This allows it theoretically to minimize total travel time with a combination of direct and little circuitous routes combined with short headways. Due to simplifications, the assignment does not realistically reproduce the real route choice of passengers, and transfer waiting times can only be modelled between different PT systems.

Michael Rittner and Karl Nachtigall

Michael Rittner and Karl Nachtigall’s numerical method for PT network design using mixed integer linear programming as well is presented here. During linearization they do not make some simplifications which Borndörfer et al. [Born08] make. Objective function is the minimum of the weighted sum of riding times and operating costs. Riding times are weighted with an evaluation function according to the deviation from the shortest path time. Stops and potential terminal stations are considered as given. Line length margins, vehicle and passenger flow margins and stop capacity margins (important for scheduling) are optional. In a pre-process, the existing network can be reduced and new lines can be created. Scheduling is included into line network design. The OD matrix is fixed. The used assignment method is a schedule-based all-or-nothing assignment. The flows are assigned more accurate than by the headway-based all-or-nothing method. Waiting times are calculated in minute wise steps. Several speed levels may be considered. It is possible to include given lines. An application using a case study has not been published yet.

Limitation: If the line pool is small it is only partly possible to minimize total travel times with a combination of direct and little circuitous routes combined with more circuitous routes but with short headways.

2.3.4 Simulated Annealing

Fang Zhao [Zhao06] and Xiaogang Zeng [Zhao08] evolved and implemented a numerical method for PT network design using an integrated simulated annealing, tabu search and greedy search method. Passenger costs are minimized. In setting a penalty for not covered demand, the coverage of the demand in the service area is simultaneously maximized. Beside the line network layout, the schedule is variable in the optimisation process too. The OD
matrix is fixed. The schedule-based assignment method assigns all-or-nothing. Trips are neglected and a penalty is added to the objective function value beyond an upper bound of one or two transfers. Consequently longer detours are preferred to avoid the upper bound for transfers. Several constraints are set: fixed number of vehicles, headway margins, route directness and maximum line length. It is possible to include or extend given lines. To reduce computing time, two hierarchies of nodes are defined: a smaller number of “key-nodes” and a bigger number of “local-nodes”. Around every key-node, local-nodes can be defined. The farther away a local-node is located from its corresponding key-node, the lower is its order. An infrastructure node could be key-node and local-node (related to another key-node) at the same time. At first, potential lines are designed connecting key-nodes with a shortest route. Later the alignments can be varied, replacing a key-node by a corresponding local-node. The new line alignment allows also shortest routes between the new set of nodes. The more key-nodes are defined and the lower the order of the local-nodes is chosen, the longer computing times become.

In a case study, using the service area of the Miami-Dade Transit (MDT, USA), the considered number of possible stops (OD zones) with 2'800 is quite high. If one transfer is maximally allowed during assignment, the computing time is still around 24 hours [Zhao08].

Limitation: If the number of transfers in a trip exceeds a certain limit (one or two), trips are either assigned with less transfers or neglected. For network evaluation such neglected trips are evaluated with a maximum travel time.

### 2.4 Sequential Network Generation Studies

#### 2.4.1 Overview

There are several possibilities to create line networks step by step.

- Lines are successively created, changed or removed.

- Line networks are reduced, starting from a minimum in-vehicle travel time network.

- A line network can be composed step by step using lines of a given line pool.

If lines are created, prolonged or changed step by step ([Sonn77], [Simo81]) they are more and more used by passengers. Line frequencies have to be increased with increasing passenger flows. If lines interact with other lines, passengers can transfer and use shortcuts. So, some lines should later be split again or even be removed. Then, the step by step process makes no sense anymore, because the later checking process (for splits and removals)
needs more and more computation time the larger the network is growing. That is why such approaches are not further analyzed.

### 2.4.2 Line Network Reduction

If line networks are reduced, starting from a minimum in-vehicle travel time network, remaining lines are more and more used by passengers and headways should be reduced in order to cope with the passenger flows. Frequencies are rising if lines with few passengers are merged with other lines.

Young-Jae Lee created and implemented in his PhD thesis [Lee98] (see [Lee06] as well) a numerical method for PT network design (called “TRANED: TRAnsit NEtwork Design”), following this type of network reduction process. Objective criterion is minimum travel time. Constraints can be a recovery ratio i.e., the ratio of revenues and operating costs, or the minimum number of routes. The stop based OD matrix can be dynamic (logit model for modal split between public transit and car). For the headway-based assignment of passengers to PT all-or-nothing assignment is used. Lee’s heuristic reduces a shortest line network in deleting lines or merging line pairs iteratively, combined with assignment and headway adjustment. To some extent Lee’s approach combines line network design and scheduling. It takes into account that a shorter headway can reduce travel times in some cases more than a shorter route to the same operating costs. As a result, major travel flows have direct connections while minor travel flows have shorter waiting times on more circuitous routes. Optionally express services can be included. The test problems had less than 20 nodes.

Limitation: Due to the all-or-nothing assignment passengers are partly assigned to other lines than the lines they would chose realistically. Due to small successive changes in the PT line network, it is possible that the optimisation process ends in local optima.

### 2.4.3 Line Network Composition

Similar to the simultaneous network generation studies it is also possible to use a heuristic, which successively builds up a line network by small changes using lines of a pre-created line pool.

**Avishai Ceder and Yechezkel Israeli**

Such approach has been introduced e.g. by Avishai Ceder and Yechezkel Israeli (compare [Ceder07] and [Ceder02]). The numerical method for PT network design replaces successively lines in a network which are evaluated as “worst”. If a PT line has been deleted, it is replaced by other lines. To find new lines, a “set-covering problem” is solved which mini-
mizes the number of lines and transfers such that the connectivity between OD-zones is guaranteed and the networks total deviation from shortest line network is minimized as well.

Objectives are to reduce total waiting time, total unused seat capacity, in-vehicle travel time and fleet size. Outputs of the calculation are Pareto optima. Constraints are maximum load of each vehicle, maximum allowed deviation from shortest path and maximum number of transfers. The OD matrix between stop-zones is fixed. The headway-based assignment is conducted according to the frequency share of a connection (the higher the frequency of a connection the more it is chosen by passengers compared to others) and partly according to travel speeds too. Headways are set according to maximum load sections. In a case study Ceders method can handle 91 OD-zones and 550 lines in the line pool (see [Ceder02]).

Limitation: Due to small successive changes in the line network, it is possible that the optimisation process stops in local optima.

Mathias Michaelis and Anita Schöbel

Another approach has been introduced by Mathias Michaelis and Anita Schöbel [Micha09]. The numerical method for PT network design creates lines, connecting successively neighboured stops in a PT network following various rules. The most important of the network design rules are the following:

- All lines should have a slightly shorter *cycle time* than a multiple of the basic period.

- Stops which are not served by other lines yet are more likely to be chosen for a new line.

- Stops which are already part of the line network are less likely to be added to a new line.

With random chosen frequencies and natural numbers, multiplied by the basic period defining the *cycle time*, a set of lines is created. Then, in a second step, the as best pre-evaluated line out of the set is added to the PT network. This line design process is repeated, till all vehicles of a predefined pool are used.

Objective is to maximize the attractiveness of a PT network from passengers' and from potential passengers' point of view. The attractiveness of the PT network defines the OD matrix between stop-zones depending on in-vehicle travel times, on waiting times and on travel times by car. Constraints are the total number of vehicles for the network and the capacity of each vehicle. Further the capacity of stops concerning the number of vehicles can be restricted. The used assignment method is a schedule-based all-or-nothing assignment. For
pre-evaluations during line building, the schedule is not optimized. If all vehicles are used, the line network is fixed and a schedule optimization is performed. The schedule optimization simultaneously synchronizes pairs of lines with each other successively increasing the number of already synchronized lines in a line group. For building line pairs matching techniques are used. During synchronization of line pairs, schedules at possible transfer stations are varied and it is assured that the stop capacity always is complied with. In a case study Michaelis and Anita Schöbels method can handle 248 OD-zones and a number of 46 buses on a star shaped network structure of Göttingen (Germany, see [Micha09]).

Limitation: It is crucial for this design procedure that the parameters and restrictions during line building are chosen well, to get optimal properties for the entire PT network. The pre-evaluation of lines is biased since no schedule optimization is performed.

2.5 Synthesis and Conclusion

2.5.1 Objektive Criteria, Constraints and OD Matrix

The objective criteria of PT network design are normally a combination of passenger costs (time costs) and operating costs (service provider costs). When operating costs are not included into the objective function they are often set as constraints (compare Table 2 to Table 5).

Some approaches include mode choice modeling, e.g. walking, car or PT, in generating the OD matrix, while others start with a given public transport OD matrix. Most approaches optimize one PT speed level (often bus), although some include speed levels with higher speeds, e.g., rail or limited-stop bus routes. Stop placement and coverage of the demand area are mostly neglected in line network design.

2.5.2 Assignment Methods and Integration of Schedule Optimization

Only two methods ([Ritt09] and [Zhao08]) combine line network design and scheduling by using a formal schedule-based assignment. Schedule-based assignment requires a schedule optimization preceding any network evaluation to get a fair, unbiased comparison between the different network variants. In addition to the slower schedule-based assignment, schedule optimization increases computing times.

Instead of a formal schedule-based assignment, a headway-based assignment is often used. Due to headway-based travel time calculation, only schedule independent (uniformly distributed) passenger access and no schedule coordination at transfer points is considered.
Investigation of space-time structures in public transport networks and their optimization 2010

(compare [Weid07b]). The demand model could become more accurate if schedule-based assignment would be used. In contrast, using a headway-based assignment has the advantage that the resulting schedules are more robust regarding delays. Schedules of different lines don’t need to be synchronized in any case.

Several of the numerical approaches for PT network design optimization developed in the last few years do include scheduling, at least partly ([Ceder07], [Lee98] and [Zhao06]). In this research, scheduling is considered to be included if at least the assignment of an OD matrix to routes is headway-based and sufficient detour alignments are possible. In this case, the lack of a direct route can be compensated for by shorter headways on more circuitous routes.

2.5.3 Headway Optimization and Precision in Space

Some approaches set headways based only on vehicle capacity and estimated passenger flows at maximum load sections. However, there are numerous other possible headway settings that might reduce PT system costs. For example, if the frequency of a given line is increased (headway decreased), it could unburden other lines and lead to a reduction in system-wide PT costs by making it possible to decrease frequencies on other lines. That is the reason why headway optimization is helpful in PT line network design ([Biel02] and [Zhao06]).

The network sizes of important urban areas can only be handled by a few of the described approaches, and most of these approaches simplify parts of the transport model to achieve acceptable computing times. Such simplification includes using an all-or-nothing assignment of the OD matrix to routes or an even simpler assignment method, such as methods ignoring transfers. However, a schedule-based all-or-nothing assignment is already more accurate concerning the passenger flows than the headway-based all-or-nothing assignment. Integer programming and heuristic methods both are able to design large networks. For examples, see Borndörfers mixed integer linear programming approach ([Born08a] and [Born08b]) and Zhao et al.’s simulated annealing approach ([Zhao06] and [Zhao08]).

2.5.4 Conclusions

In summary, due to the complexity of designing urban public transport networks, there are currently no really satisfying solutions for numerical network design [Fried06]. All the approaches mentioned above simplify the PT system in several ways. Therefore, even if an approach is able to reach a “global optimum”, such an optimum will be sub-optimal in most cases given the simplifications made in the approach. The goal of this research is to develop a new approach that reduces the number of simplifications and thereby improves the quality of network design. The following chapters describe this new approach.
### 2. Integrated Sub-problems

If the line pool is large enough, scheduling (frequency setting) is considered to be included to a certain extent.

### Source (I)

<table>
<thead>
<tr>
<th>1a. Objective Criteria</th>
<th>1b. Constraints</th>
<th>2. Integrated Sub-problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum of weighted indicators (i.e.: number of passengers, total travel time, total number of transfers, crowds index, vehicle number, etc.)</td>
<td>Passenger flow according to modal split and assignment, a line pool is given</td>
<td>No</td>
</tr>
<tr>
<td>Weighted sum of total ride time and service provider costs</td>
<td>Passenger flow and headways according to assignment, potential terminal stations are given, line length margins (optional)</td>
<td>Scheduling (frequency setting according to maximum load section) is considered to be included to a certain extend</td>
</tr>
<tr>
<td>Reduce total waiting time, reduce total unused seat capacity, reduce in-vehicle travel time, reduce fleet size</td>
<td>Passenger flow and headways according to assignment, minimum frequency, max. load on each vehicle, maximum allowed deviation from shortest path, max. number of transfers</td>
<td></td>
</tr>
<tr>
<td>Reduce travel time</td>
<td>Passenger flow and headways according to modal split and assignment, recovery ratio: minimum revenue per operating costs (optional), minimum number of routes (optional)</td>
<td>Scheduling (frequency setting) is considered to be included to a certain extend</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>---------------------------------------------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>Maurizio Bielli et al. [Biel02]</td>
<td>Genetic Algorithm variables: line use, frequencies</td>
<td>Dynamic (modal split between public transit and walking)</td>
</tr>
<tr>
<td>Ralf Borndörfer et al. [Born08], [Born04]</td>
<td>Mixed integer linear programming (solver: SCIP 1.0) variables: line use</td>
<td>Fixed</td>
</tr>
<tr>
<td>Avishai Ceder, Yechezkel Israeli [Ceder07]</td>
<td>Heuristic</td>
<td>Fixed</td>
</tr>
<tr>
<td>Young-Jae Lee [Lee98], [Lee06]</td>
<td>Heuristic (reducing a shortest line network through deleting lines or merging line pairs iterative combined with modal split, assignment &amp; headway adjustment)</td>
<td>Can be dynamic (logit model for modal split between public transit and car).</td>
</tr>
</tbody>
</table>

Table 3: Mathematical approaches for network design, listed here by help of 9 criteria (see as well Table 2).
### Table 4: Mathematical approaches for network design, listed here by help of 9 criteria (see as well Table 5).

<table>
<thead>
<tr>
<th>Source (II)</th>
<th>1a. Objective Criteria</th>
<th>1b. Constraints</th>
<th>2. Integrated Sub-problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathias Michae- lis, Anita Schöbel [Micha09]</td>
<td>Maximize the attractiveness of a PT network from passengers’ and potential passengers’ point of view.</td>
<td>Passenger flow and headways according to assignment, the total number of vehicles for the network, the capacity of each vehicle and the stop capacities (concerning the number of vehicles)</td>
<td>Scheduling in the end (pre-evaluation without scheduling)</td>
</tr>
<tr>
<td>Michael Rittner, Karl Nachtigall [Ritt09]</td>
<td>Weighted sum of total ride time and operating costs</td>
<td>Passenger flow and headways according to assignment, line length margins, vehicle and passenger flow margins and stop capacity margins (important for scheduling) are optional</td>
<td>Scheduling</td>
</tr>
<tr>
<td>Ulrich Schäffeler [Schäf05b]</td>
<td>Maximize subjective supply quality (passenger share); reduce access and egress time, headways and the public to car transport door-to-door travel time ratio.</td>
<td>Passenger share according to modal split, fixed number of vehicles (fixed vehicle density)</td>
<td>Stop setting and headway optimization</td>
</tr>
<tr>
<td>Fang Zhao et al. [Zhao08], [Zhao06], [Zhao03]</td>
<td>Maximize demand in the service area (penalty for not covered demand) and minimize passenger costs</td>
<td>Passenger flow according to assignment, fixed number of vehicles, headway margins, route directness, maximum line length.</td>
<td>Scheduling</td>
</tr>
</tbody>
</table>
### Mathematical approaches for network design, listed here by help of 9 criteria (see also Table 4).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathias Michaels, Anita Schöbel [Micha09]</td>
<td>Heuristic</td>
<td>Dynamic (modal split between public transit and car)</td>
<td>All-or-nothing assignment</td>
<td>Stop fine</td>
<td>Schedule-based assignment</td>
<td>Several levels possible</td>
<td>Possible</td>
</tr>
<tr>
<td>Michael Rittner, Karl Nachtigall [Ritt09]</td>
<td>Mixed integer linear programming (LP-solver); variables: line use, schedules</td>
<td>Fixed</td>
<td>All-or-nothing assignment</td>
<td>Stop fine</td>
<td>Schedule-based assignment</td>
<td>Several levels possible</td>
<td>Possible</td>
</tr>
<tr>
<td>Ulrich Schäffeler [Schäf05b]</td>
<td>Adaptive recursive Simpson’s rule [Gan00]</td>
<td>Dynamic (potential customers are assumed to live uniformly distributed in catchment area; modal split between PT and car)</td>
<td>Logit model: split modes according to the objective function</td>
<td>Exact</td>
<td>Headway-based travel time calculation</td>
<td>Express services (optional)</td>
<td>Not possible</td>
</tr>
<tr>
<td>Fang Zhao et al. [Zhao08], [Zhao06], [Zhao03]</td>
<td>Integrated simulated annealing, tabu search and greedy search; variables: line use, schedules</td>
<td>Fixed</td>
<td>All-or-nothing assignment</td>
<td>Street node fine (2800 nodes)</td>
<td>Schedule-based assignment</td>
<td>One speed level</td>
<td>Possible</td>
</tr>
</tbody>
</table>
Table 6 Existing solutions for the review criteria (numerical approaches only).

For some of the nine criteria solutions exist already in a satisfying way (compare Table 6). Other points still make problems, especially in combination with each other, and prevent a holistic solution for PT network design. Therefore, the current approaches for PT network design meet the needs of real world PT planning only in parts yet.

The main points which should be deepened are:

- Beside line network design and scheduling, stop placement as well as scheduling of vehicles and crew should be considered in order to complete line network design (a full inclusion of scheduling and an inclusion of vehicle scheduling into line network design cannot be considered here. This study concentrates on integrating stop placement and only in parts integrates scheduling).

- Computing times are rising while either improving accuracy of the transport models, enlarging the number of considered speed levels or enlarging the considered network areas. Therefore, a strategy is necessary which copes with some uncertainties in the same degree of accuracy. The accuracy could then be adjusted to achieve accepta-
ble computing times. Existing approaches for large network sizes simplify the transport model to achieve moderate computing times. An improved approach should be able to use more realistic transport models without increasing computing times too much.

- Questions such as the optimal number of speed levels (speed levels of different transport modes, for example local and express services) and parameters such as line distances, transfer node distances, optimal speeds etc. are not considered at present.
3 Methodological Approach

3.1 Overview

As could be seen in chapter 2, existing approaches for large network sizes simplify the transport model more or less to achieve moderate computing times. In this study, a new approach for PT system optimization is being developed. The approach shall integrate stop placement and scheduling into line network design and additionally be able to use more realistic transport models without increasing computing times too much.

The main elements of this PT network design optimization approach such as

- sequential optimizing planning areas and
- reducing the number of network evaluations

are combined to a conceptual procedure which is presented in section 3.2. In the following two sections (section 3.3 and 3.4) the basic principles and main methods behind this procedure are discussed more in detail. In section 3.3 the basic principles of PT are discussed. In section 3.4 the chosen assignment methods are summarized. In section 3.5 the choice of the optimization algorithms is outlined. Further on, the chosen form of the genetic algorithm (as can be found in literature) is presented to understand the differences of the adapted genetic algorithm (section 4.2). In section 3.6 the implementation of the PT network design process is outlined.

3.2 Conceptual Procedure for the Design Process

This thesis presents a stochastic multiple area approach for public transport network design optimization. The approach enables users to model the transport system more realistically in the network design process than previous approaches. It enables users to more accurate assign passengers to specific PT lines, as well as considering stop placement and scheduling in the network design process. The new approach is illustrated in Figure 2.
Investigation of space-time structures in public transport networks and their optimization

Network Design Input
- OD matrix on a hectare base
- transport modes and available road and rail networks with parameters
- network development projects (extension/break-off/upgrading)

Network Design Schema

a) if infrastructure is not or only partly given: placement of line nodes at OD centres of gravity in optimal line distances, considering coverage and travel time. Between line nodes a triangular infrastructure network is built.

b) build a line network on the base of nearly shortest routes (NSR)

c) aggregate OD matrix from hectares to line nodes and line corridors; adjust stop densities, access/egress from hectares, dwell times and so adjust ride times

d) headway and vehicle size optimization using GA

e) select one line

f) reduce line (delete, shorten, merge with an other line) with ACO (local OD matrix)

g) headway based calculation of travel times/generalised costs between line nodes and line corridors and assignment of demand (comply with vehicle capacities) for all in block f produced alternatives

h) select best alternative if an improvement, otherwise keep old solution

i) stop placement on lines

Network Design Output
- schedule optimization
- comparison of several different networks

Figure 2 Network Design Schema.
The stochastic multiple area approach for PT network design consists of the following three loops:

Loop 1: The network design process is run for each PT speed level (bus and tram, regional train, etc.) from fastest to slowest.

Loop 2: Each speed level is divided into several geographic planning areas. This enables the approach to achieve faster computing times by optimizing smaller area PT networks.

Loop 3: The guided stochastic search heuristic (GSSH) used in this process requires comparatively few network evaluations.

The multiple area approach (Figure 2, loop 2) only makes sense if the speed levels are optimized separately (Figure 2, loop 1). Due to the top-down process, PT networks of the lower speed levels are adapted to line networks of the higher speed levels. The used guided stochastic search heuristic (Figure 2, loop 3) is applied sequentially to the planning areas. After completing line network design for all planning areas of one PT speed level, stops are placed definitely on the lines (see Figure 2, block i). Previously, only stop densities on link corridors were given (see Figure 2, block c).

The schedule gives the most important information about a PT network. It defines the transport service both spatially and temporally, as well as the connections between lines. Therefore the schedule should be optimized together with its associated network, at least in the end.

After an entire network has been developed, i.e., the line alignments for all the PT speed levels have been defined for all the planning areas, it is possible to repeat the stochastic process to create due to the stochastic nature several different networks. These can then be compared with the previous ones to identify the best of the thus created local optima.

A public transit network can be adapted best to a given demand pattern if existing infrastructures are not considered. In this case the transit network design can start from scratch and is not influenced by already grown transport networks (see Figure 2, block a). Line nodes can be placed at OD-centers of gravity in optimal line distances, considering coverage and travel time. Between line nodes a triangular infrastructure network is built. If existing road and rail networks are considered, possible line distances are already fixed and placing of line nodes is not necessary anymore. In that case, these steps are omitted.
3.3 Space and Time Approach

3.3.1 Hierarchies in Transport Networks and Levels of Speed

Because it is not possible to diffuse through spaces in land transport (that means always choose the linear distance as route), transportation needs definite transport corridors. So, road and railroad infrastructures are necessary. Roads occupy space, they cut through landscapes and separate them, roads require capital investment and maintenance and the traffic emits noise, vibration and exhaust gases. All this total costs make it desirable to reduce the total length of roads or the network density respectively in transport networks, till an equilibrium with network utilities is reached.

Low level roads and railroads allow only few vehicles per time period to use them, generally at low speeds. At the same time they are less costly to be built and maintained than high level roads (or railroads, allowing fast traffic on them), and due to low speeds, it is easier to access, egress or to cross them. High level roads allow generally higher speeds, are more costly including accesses and egresses and due to longer acceleration times and access and egress costs, distances between access and egress points (in case of PT: stops) are long. Example values for PT are listed in Table 7. A PT line with short distance stops allows low average travel speeds only. A line with long distance stops allows faster travel speeds.

<table>
<thead>
<tr>
<th>Specification of Means of Transport</th>
<th>Assumed Dwell Time [s]</th>
<th>Stop Distance [m]</th>
<th>Travel Speed [km/h]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed traffic, usual accelerations, $v_{\text{max}} = 30$ km/h (e.g. light rail, bus)</td>
<td>20</td>
<td>300</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>500</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>700</td>
<td>22</td>
</tr>
<tr>
<td>Separated traffic, $v_{\text{max}} = 60$ km/h (e.g. light rail, metro)</td>
<td>20</td>
<td>500</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>700</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>900</td>
<td>34</td>
</tr>
<tr>
<td>Rail-systems, $v_{\text{max}} = 130$ km/h (e.g. regional train)</td>
<td>20 – 40</td>
<td>2'500</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>20 – 40</td>
<td>3'000</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>20 – 40</td>
<td>3'500</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 7 Mean travel speeds for urban PT depending on stop distances (source: [Weid08]).

From travelers’ point of view, shorter OD-distances allow to use slower means of transport to move from origin to destination in a certain time, whereas longer distances require the use of faster means of transport to reach the destination in an acceptable time.
Both effects lead together to a hierarchy of transport networks according to road and railroad characteristics and speeds on them (compare Figure 3). The densest network is the network of footpaths. A dense basic network of low level roads (e.g. back roads, roads for bus and light rail) ensures the coverage of origin and destination areas. It is superimposed by less dense networks of high level roads (e.g. main roads, railroads for subway and regional train) to ensure faster connections between origins and destinations more far away from each other (compare [Nes02]).

Figure 3 Illustration of a multilevel network (compare: [Nes02]).

Due to cost reasons, the possible number of speed levels is limited, the flow on each level should exceed a certain lower limit (economies of scale). Additionally, the distances between roads of the same level should be a multiple of road distances of the denser and lower levels (superimposition of levels to reduce network costs). That is because different levels can partly share their infrastructures or at least different levels can be bundled and so their separation effect on the surroundings can be combined (reduced).

Figure 4 Examples of transport routes using different levels of speed for private car (left-hand side) and PT (right-hand side); (source: [Nes02]).

In PT networks, passenger routes follow PT-lines, where, due to the necessary transfers between vehicles (compare Figure 4), passenger costs for transfers result. Transfer costs generally are independent from speed levels (if transfer facilities are comparable and headways
of the two lines do not differ too much). None the less, transfers reduce the optimal number of speed levels in PT another time.

### 3.3.2 Design Sequence of Speed Levels

With this approach PT line networks are sequentially designed and optimized for each speed level starting with the fastest (Figure 2, loop 1). This simplification has several advantages:

1. Each level can be optimized separately, in order to reduce computing time.

2. During the design of the fastest speed level the access/egress costs using the slower levels can be already roughly considered. If later networks for lower levels are designed, they can be adapted to the higher level networks with only few additional costs for the PT system.

3. During the design of the slower speed levels the faster exists already. All “long” distance trips can be roughly assigned using a simplified transport chain (approximated access/egress). Consequently, the assignment of demand to routes of the slower speed levels can include the longer distance trips as well.

However, the multiple area approach could not be tested on optimizing several speed levels due to the slow computing times for network evaluations (see section 5.3.4).

### 3.3.3 Organization of Planning Areas

The multiple area method (Figure 2, loop 2) enables the approach to achieve faster computing times by optimizing smaller area PT networks. The planning area should be defined large enough so that most through-trips start or end in the neighboring planning areas or use a higher speed level. Consequently, the approach uses more planning areas in the analysis of slower speed levels than for the faster speed levels.

Inside the planning area being optimized and its neighboring areas, trips are assigned precisely, whereas trips using higher speed levels are assigned less precisely the further away they are from the planning area (see Figure 5). This means that computing times are only linearly dependent on total area size.
If advantageous, PT lines in the area being planned can be merged with lines developed earlier in the neighboring areas, as long as the line alignments (sequences of nodes on a line) in the neighboring areas are not changed. Once the PT line network inside a planning area is optimized by the guided stochastic search heuristic (see section 4.3), only the lines in the middle of the planning area are considered as definitely set (for an example, see section 5.3.4).

3.3.4 Time Coordination at Transfer Points

Instead of a schedule-based assignment, in this study a headway-based stochastic multiple route assignment (see section 3.4.3 and [PTV07]) of OD-pairs to the PT network is used.

Due to headway-based travel time calculation, only schedule independent (uniformly distributed) passenger access and no schedule coordination at transfer points is considered (compare [Weid07b]). Timed transfers (see Figure 6) are not considered. The demand model could become more accurate if schedule-based assignment would be used (compare Table 17).

In contrast to use headway-based assignment has the advantage that the resulting schedules are more robust regarding delays (it is assumed, that schedule-based assignment leads to good connections, which could be broken however in case of delays). Schedules of different lines don’t need to be synchronized in any case.
If schedule-based assignment would be considered,

- firstly the demand model could be more accurate,

- secondly integrated cycle times (timed transfers, see Figure 6) could reduce waiting times during transfers and

- finally synchronized schedules could similar to intercity railway services in Switzerland lead to a co-evolution of network and schedule to integrated cycle times and a network with similar distances of transfer nodes.
3.4 Assignment Methods

3.4.1 Introduction

Objectives that can be approximated during the optimization depend first on the model of the system to be optimized, i.e., the model of the PT system, and, second, on the optimization process itself. One of the main parts of network evaluation is to get exact user costs and passenger flows on PT lines. Accurately modeling passenger behavior is especially complex in network design since the network itself is being developed. For example, during the network design process, the exact schedule might not be given, e.g., due to limited computing time. In such a case, the behavior of passengers has to be estimated based on headway information only. Passenger behavior is modeled as assignment of OD-pairs to the PT network.

In this approach two different assignment methods are used in combination. The headway-based stochastic multiple route (HBSMR) assignment method models the behavior of passengers in the PT network quite accurately (see section 3.4.3 and [PTV07]). To use hectares as OD-zones in the HBSMR assignment method (see section 3.4.3) would require too much computing time. That is why only line nodes and link corridors which cover demand are considered as OD-zones in the HBSMR assignment method. To be able to assign passenger access from hectares (basic zones) to those OD-zones, a more simple assignment method is used (see the following section).

During ant colony optimization a headway-based all-or-nothing assignment is used to assign the local OD matrix to one single line. The headway-based all-or-nothing assignment assigns the demand according to shortest paths. Average waiting times preceding passenger boarding are half the lines headway. In case the new line alignment contains loops, the precise waiting time for transfers between two branches of the line is calculated according to the cycle times of the new line.

In section 3.4.4 the schedule-based assignment used in the transport model of the Kanton Zürich [Vrtic05] is outlined.

3.4.2 Assignment of the Basic Zones to Line Nodes and Link Corridors

Depending on access and egress walking times and on perceived PT travel times (see brackets in equation (6)), demand of each basic zone (hectare; 100m * 100m) is assigned best way (all-or-nothing) to line nodes and link corridors which become the OD-zones for the headway-based stochastic multiple route assignment (see next section and Figure 8). Perceived PT travel times are calculated beforehand by the HBSMR assignment.
The linear access and egress distances are multiplied by a detour factor of 1.24 [Bränd78]. Between hectares and line nodes the linear distances are given directly. In case of the distances between hectares and the virtual on-link stops on the link corridors, linear distances are estimated from the orthogonal distance between the hectare and the link corridor and from the stop density on the respective link corridor (see Figure 7). The parts parallel to the PT road have an average length of 25% of the on-link stop distance (only the average stop distance on the respective link corridor is given not the exact position). OD-zones of link corridors are situated at their centroid regarding the demand potentials (see Figure 16) on the link. For the city centre of Winterthur and one hectare the possibilities to access the PT network are illustrated in Figure 8. OD-zones are displayed as well.
3.4.3 Headway-Based Stochastic Multiple Route Assignment

The assignment method used in this study to simulate the route choice of passengers between OD-zones (see Figure 8) is a headway-based stochastic multiple route (HBSMR) assignment (see [PTV07]). It is described in the following. Stochastic refers to the probability based waiting times and path choices of passengers.

Waiting times during accesses and transfers are not given and have to be estimated. Arriving times of passengers at stations are assumed to be uniformly distributed. Departure times of PT lines at stations are assumed to be independent from each other and uniformly distributed as well. Then, in case of one available PT line, average waiting time is half of the lines headway. If two PT lines offer the same in-vehicle travel time and the same headways for a direct connection, then average waiting time is one third of the two lines headways. This is due to the fact that these two lines are not necessarily coordinated and that their departure
times at the station are independent from each other and uniformly distributed. If the two lines were coordinated such, that the two lines together build a new common headway which is half of the single headways, then average waiting time would become one forth of the two line headways. In the general case average waiting time in the time period \( I \) at the decision point \( p \) is according to [PTV07]:

\[
\begin{align*}
\mathcal{w}_p^I &= \sum_{\ell \in \mathcal{L}(p)} \frac{1}{\tau^I_{\ell}} \int_{\tau^I_{\ell}}^{\infty} x \prod_{\nu} \left( 1 - \frac{\left( x + \nu_{\ell,p}^I - \nu_{\nu,p}^I \right)}{\tau^I_{\ell}} \right)^x dx \\
\tau^I_{\ell} &= \text{headway of line } \ell \text{ in the time period } I \\
\nu_{\ell,p}^I &= \text{impedance (generalized costs of passengers) of the passenger route after the decision point } p \text{ if line } \ell \text{ is used without waiting time at decision point } p. \\
(\cdot)^+ &= \text{only positive values are considered}
\end{align*}
\]

The passenger share of line \( \ell \) in the time period \( I \) at the decision point \( p \) is:

\[
\begin{align*}
\alpha_{\ell,p}^I &= \frac{1}{\tau^I_{\ell}} \int_{\tau^I_{\ell}}^{\infty} \prod_{\nu} \left( 1 - \frac{\left( x + \nu_{\ell,p}^I - \nu_{\nu,p}^I \right)}{\tau^I_{\ell}} \right)^x dx \\
\end{align*}
\]

All route choices follow the same principle: “The probability of choosing a specific option is equal to the probability of this option being the best of all alternatives” [PTV07]. There is a specific probability that one option with a longer in-vehicle travel time could lead to a faster connection if the trip started earlier than the alternative connection. At least the differences of in-vehicle travel times should then be compensated. For all route choices this principle is applied independent from the route decisions made previously (Markov property).

The route options are evaluated with their generalized costs from customers point of view (see section 1.3.2). Further it is assumed that passengers know headways and in-vehicle travel times of all lines and that they can get information about the different trip options at all decision points (stop points).

### 3.4.4 Schedule-Based Assignment

The assignment method originally used in the transport model of the Kanton Zürich [Vrtic05] is a schedule-based assignment (for more details see [PTV07]). For each acceptable and
“dominant” connection the impedances are calculated according to the perceived user times (brackets in equation (6)). A connection dominates others if at least one of its attributes shows better values and no of its attributes show worse. With access waiting time:

\[ t_{aw} = \min\left(h^e, t_{aw\text{ max}}\right) \]  

(10)

(where: \( t \) – time; \( aw \) – access wait; \( h \) – headway; \( e \) – exponent, in [Vrtic05] \( e = 0.6 \), \( t_{aw, \text{ max}} \) – maximum waiting time, in [Vrtic05] \( t_{aw, \text{ max}} = 15 \text{ min} \))

In using equation (10) it is assumed that passengers know the schedule at the origin the more the longer the headways are.

The choice of a connection \( p_j \) (the percentage of the demand using the connection for a given time interval) is calculated according to Lohse [PTV07]:

\[
    p_j = \frac{\beta \left( \frac{R_j}{R_*} - 1 \right)^2}{\sum_j N_j \beta \left( \frac{R_j}{R_*} - 1 \right)^2}
\]

(11)

(where: \( \beta \) – impedance parameter, in [Vrtic05] \( \beta = 4 \); \( R \) – impedance of a connection – perceived user times (brackets in equation (6); also fares and the temporal utility of a connection could be included, \( R_* \) – minimal impedance of a connection)

### 3.5 Optimization Algorithms

#### 3.5.1 Introduction

Since the optimization of PT network design is complex and the solution space quite large it is not possible to get the global optimum for larger network sizes in an acceptable time. Due to a relative “flat” objective function (compare [Schäf05b]) many local optima are close to the global. Since the optimality criteria (total costs) are general costs and the input data are not perfectly accurate, it is also not necessary to find the global optimum. It is sufficient to achieve a sub-optimal solution close to the global optimum.

The multi-objective function of PT network design is non-linear, combinatorial, partly discrete and continuous. A non-linearity is e.g. the recursive connection between the capacity constraint (maximum headways of lines) and the route choice of passengers according to head-
Investigation of space-time structures in public transport networks and their optimization _____________________ 2010

The combinatorial nature of PT network design is caused by the huge number of possible line and schedule combinations. Most of the variables are discrete like e.g. line selection, line alignment and headways; others are continuous like e.g. departure times of vehicles and stop positions. That is why search techniques (see Figure 9) are appropriate to solve such problems. Search techniques are “robust” optimization methods in the following sense: they only operate with objective function values, they do not need a derivation of the objective function, a steady objective function or a convex search space.

Figure 9 Taxonomy of search and optimization algorithms depending on objective function and strategy. Typical samples are listed as well (compare [Stüm99]).

Linearization of the objective function would be an alternative [Ritt09] to the costs of simplification and discretization though. The combinatorial and discrete natures of the problem still remain.

In the presented network design optimization approach, the stochastic search techniques of ant colony optimization (ACO) and genetic algorithm (GA) are applied. ACO algorithms are inspired by the behavior of ants, marking their way with pheromones. With such marking, ants lead other ants of the colony to goals such as food sources. ACO algorithms are useful, if interim spatial results should be stored and locally used for further search steps. GAs belong to the evolutionary algorithms. They are an abstract model of the biologic evolution. GAs can handle well stochastic multi objective optimization problems. They are appropriate for complex objective functions or for problems where no better optimization methods are available for.
Other search techniques like simulated annealing also can be used successfully for PT network design [Zhao08]. Often, it is more a question of how to implement a search technique successfully than a question of which method should be used. ACO is especially appropriate to be combined with the reduction approach of Young-Jae Lee [Lee98]. The evolutionary algorithms are quite efficient, if the objective function is of stochastic nature (see [Alt03]) and if neighborhoods can be used. The GA is better than the evolution strategy suited to handle discrete problems.

In the following, the stochastic search technique GA is presented as can be found in literature. The presentation starts with the population concept GAs have in common with ACOs. In chapter 4 the algorithms adaptation and application to the network design optimization approach is presented.

3.5.2 Population Concept

**General Population Concept**

Evolutionary algorithms and ant colony optimization algorithms both use a population concept. Individuals $x_p^{(g)} (g = 0, \ldots, \gamma-1 \text{ [generation number]; } p = 1, \ldots, \lambda \text{ [population number]})$ of a population are a set of variables (the analogon in biology is a haploid chromosome). In case of headway optimization, an individual defines for each line of the PT network a corresponding headway. So, an individual is a set of line headways.

First, a start population is created. In case of headway optimization all headways of the individuals could be set to the same value or to values taken from already existing PT networks. During the optimization it is the goal to get populations i.e. individuals with better properties.

start population: $x_1^{(0)}$, $x_2^{(0)}$, ..., $x_{\lambda}^{(0)}$

\[\downarrow \text{development} \]

1. generation: $x_1^{(1)}$, $x_2^{(1)}$, ..., $x_{\lambda}^{(1)}$

\[\downarrow \text{($\gamma$-2)-times development} \]

($\gamma$-1). generation: $x_1^{(\gamma-1)}$, $x_2^{(\gamma-1)}$, ..., $x_{\lambda}^{(\gamma-1)}$
Generally an objective function $F(x)$ is used as quality characteristic, in case of headway optimization equation (5). For optimization the existence of a more or less strong causal relationship between variables and objective function values is important.

**Population Concept of Evolutionary Algorithms**

The structure of a basic evolutionary algorithm (see Figure 10) is listed below (compare [Alt03]):

1. Generate start population (generation 0).
2. Evaluate all individuals (PT networks) of the start population using an objective function
3. **Variation:**
   a. Selection of certain individuals from population of generation $i$ depending on their objective function values.
   b. Recombination of individuals (in case of headway optimization exchange some corresponding headways between two headway settings of a PT network)
   c. Mutation of individuals (stochastic changes of PT networks headways)
4. **Conservation** (parallel to variation):
   a. Selection of certain individuals from population of generation $i$ depending on their objective function values

5. Evaluation of all varied individuals of the population of generation $i$ using an objective function (in case of headway optimization equation (5)).
6. Combination of the varied individuals after evaluation and of the (unchanged) individuals selected in the conservation step to a new population of generation $i+1$.
7. If the stop condition is not met, go back to variation (step 3) and conservation (step 4), otherwise quit.
3.5.3 Genetic Algorithms (GA)

Introduction

Genetic Algorithms (GA) are the most famous evolutionary algorithms. In their original form they are developed by John Holland [Hol75]. In the meantime some better operators exist.

The basic GA is presented here following the structure of a basic evolutionary algorithm (see Figure 10; compare [Alt03]): selection, recombination and mutation. Some explanations follow about conservation of individuals.

Selection

After evaluation by an objective function (e.g. equation (5)), the individuals (PT networks) are selected by stochastic universal sampling proportional to their objective function values. To assure that all values are positive and to influence the selection process in a positive way, the values of the objective function are transformed first by a fitness function.
Figure 11 Two fitness functions; left hand: assignment proportional to the values of the objective function (the gradient of the straight line depends in case of the normalized fitness assignment on the objective function values of the entire population), right hand: the assignment of the truncation-selection (the $\mu$ best Individuals of the population get all the fitness value 1, the others are discarded).

Subtraction of the smallest objective value from all objective values is a simple form of scaling [Gol89] and leads to positive fitness values. Absolute values of the objective function are not important anymore (see Figure 11). Two fitness functions are used here (see Figure 11): assignment proportional to values of the objective function developed by Goldberg [Gol89] and the assignment of truncation-selection. For variation (step 3 in “population concept of evolutionary algorithms”) the proportional assignment has the advantage, that not only the as best evaluated individuals give their properties further, but theoretically all individuals to a certain probability. This prevents early convergence to a sub-optimal solution. For conservation (step 4 in “Population Concept of Evolutionary Algorithms”) the truncation-selection has the advantage that a small number of the as best evaluated individuals are always kept in the population. Those individuals steer the direction of the optimization.

After values of the objective function have been transformed into fitness values, the individuals are selected by Stochastic Universal Sampling [Bak87] (see Figure 12) proportional to their fitness. The fitness of an individual is its selection probability. If its fitness is higher than one, the corresponding individual can be selected several times. So, it is able to proliferate. In case of several selections, results of stochastic universal sampling (see Figure 12 at the right side) deviate less than results of the simple roulette selection (see Figure 12 at the left side) due to the shorter random distance.
Selection: To each individual a fitness proportional segment on a roulette wheel is assigned. In case of roulette selection (left hand), the bowl is operated $\mu$-times. In case of stochastic universal sampling (right hand) an annulus of $\mu$ bowls is operated simultaneously to get several results.

Recombination

A discrete Recombination (see Figure 13) can be achieved through exchange of the parent variables. In case of headway optimization, some corresponding headways between two headway settings of a PT network are exchanged.

Recombination of two Individuals $i$ and $k$ (single point crossover).
According to biology the individual's variables are stored in data chains, as "chromosomes". If two variables are close to each other situated in the chromosome, the probability to be separated during recombination is small. This can be used in headway optimization to be-queath headways of lines with the same stop sequences together.

The crossover operation is applied to two randomly chosen parents (choice without putting back). Analogue to biology, both recombinants are positioned at the vertexes of the hypercube which is spanned by the parent variables. (see Figure 14).

![Figure 14](image)

**Figure 14** Discrete recombination: Recombinants $R_i$ are positioned at the vertexes of the hypercube which is spanned by the parent variables $E_k$ (in two dimensions the hypercube is a rectangle). In case of headway optimization, $x$ and $y$ both are headways for a corresponding line $L_x/L_y$ of the PT network.

**Mutation (Mutation of variables)**

A mutation operator for real variables is the mutation operator of the breeder GA ([Müh93], [Müh94]):

$$x_i^{\text{Mut}} = x_i + s_i \cdot r_i \cdot a_i \quad i \in \{1,2,\ldots,n\}$$  

(12)

Each variable is mutated with probability of the mutation rate. In case the mutation rate is 0.5 only every second variable is changed according to equation (12).
notation:

\[ s_i \in \{-1, +1\} \text{ uniformly distributed} \]

\[ r_i = r \cdot \text{Domain}_i \]

\[ r \text{ – mutation range.} \]

\( \text{Domain}_i \) – domain (defined range of values) of variable \( i \).

\[ a_i = 2^{-uk} \]

\[ u \in [0,1] \text{ uniformly distributed} \]

\( k \) – mutation precision.

In case of headway optimization and a given list of accepted headways, the list index can be mutated instead of directly mutate the headway itself.

![Figure 15 Possible results of variable mutation on an individual. In case of headway optimization, x and y both are headways for a corresponding line \( L_x / L_y \) of the PT network.](image)

It happens that the search space of accepted values is left during variable-mutation. In such case, it is possible to set the individual e.g. perpendicularly back to the border of the search space.
Conservation (Generation gap)

In stochastic optimization it is often helpful to let one part of population unchanged respectively proliferate it. The ratio of the changed individuals to that of the parents’ generation is called generation gap. In conservation, individuals are selected like in variation (see the section “selection” above).

3.5.4 Ant Colony Optimization (ACO)

Ant colony optimization (ACO) algorithms were first developed by Marco Dorigo in 1992 [Dori92]. ACO algorithms use the population concept in a different way compared to the evolutionary algorithms. Individuals (e.g. PT lines) are not directly varied like individuals in evolutionary algorithms. PT lines depend on the “pheromone values”, interim spatial results of previous generations, and other spatial information. These pheromone values are adapted during optimization. The best PT lines, which are generated by the help of the (adapted) pheromone values, are stored.

The ACO has been directly adapted to the characteristics of PT line generation. The adapted algorithm is presented in detail in section 4.3.6.

3.6 Implementation of the PT Network Design Process

The fast prototyping language Python is used to implement the design process. The implementation can be downloaded from http://www.ivt.ethz.ch/docs/institute/sr150.zip. The implementation embeds the PTV tool VISUM [PTV07]. VISUM stores the network supply data and provides the assignment procedure. It is included via a COM- (component object model) interface. Additionally, VISUM can be used for network visualization and manual editing of stops, lines and schedules.

The PTV tool VISUM is used for several reasons.

- The transport model of the Canton of Zurich [Vrtic05] is realized in VISUM, its OD matrix and route choice parameters could be used.

- Several important components like assignment methods, network visualization and manual editing (of stops, lines and schedules) are already implemented.

- VISUM offers a COM-interface to transfer data and to start VISUM processes from other programs in an automatic way.
The fast prototyping language Python was chosen to implement the design process, because it is a fully object oriented programming language, but offer simplified data handling especially in combination with the COM-interface. Python enables ad hoc programming ("scripting"), database applications and GUI applications. The Python code is up to 5 times shorter than in Java and more readable. Disadvantage is the around 30 times longer computing times of programs written in Python compared to Java or C++.

Unfortunately the COM-interface to connect python and VISUM is much slower than python itself. So, it is necessary to reduce the direct communication between python and VISUM to a minimum. The headway-based assignment in VISUM for example has the advantage to schedule-based assignment that only headways are transferred to VISUM via the COM-interface during headway adaptation and optimization and not the departure times of all vehicles at the corresponding terminal stations. The later consumes a huge amount of computing times and makes it nearly impossible even for Mandles Swiss benchmark problem (see section 5.6) to use schedule-based assignment during the reduction process. Due to the slow COM-interface, schedule-based assignment can only be used after the line network design process has been finished. This could be changed from PTV AG though.
4 Network Design

4.1 Overview

A public transit network can be adapted best to a given demand pattern if existing infrastructures are not considered. In this case the PT network design can start from scratch and is not influenced by already grown transport networks. In a later step, a PT network on the existing infrastructure can be compared with such infrastructure independent PT network and network development processes could be launched if necessary. An approach for such infrastructure independent link network design is presented in section 4.2. In section 4.3 the PT network design method (the guided stochastic search heuristic; Figure 2, loop 3) is presented either starting from an artificial infrastructure independent link network or from a given infrastructure network.

4.2 Line Node Placement

4.2.1 Introduction

To evaluate PT networks on given infrastructures it can be helpful to have a kind of “ideal network” as comparison (compare [Fried05]). An “ideal network” is an infrastructure network with which it is possible to minimize total transport costs. In “ideal networks” the infrastructure can be flexibly adapted to supply and demand characteristics (see Figure 2, block a). In this case the existing infrastructure is not considered.

One basic parameter of infrastructure networks is their density. Beside the possibility to influence network density, it is important while designing “ideal networks” to place potential transfer nodes of different lines close to demand. Potential transfer nodes of different lines are called “line nodes”. So, in placing line nodes, both parameters are being fixed: distances between lines and catchment of demand. In a second step links are added between stops building a triangular infrastructure network.

Line node placement is performed without considering existing networks. Existing stops may be considered though, if there are any important. In case existing road and rail networks would be considered, possible line distances would be already fixed and placing of line nodes would not be necessary anymore.
The area under investigation can be developed with independent PT networks of different speed levels (see section 3.3.2). Line nodes of higher levels are automatically line nodes of lower levels in order to enable transfers. The investigated area is divided into several planning areas for each speed level (analogous to line network design, compare section 3.3.3), to locally optimize line node placement. This keeps computation time low and only linear dependent from total area size. According to the considered number of speed levels and planning areas, the process of line node placement is repeated several times, starting with the highest level.

The resulting infrastructure network is not perfect (not globally optimal), but can be used as a lower bound for evaluating public transit networks on given infrastructures. Creating lines before creating an infrastructure network could allow more direct lines being created. But here no appropriate approach exists till now.

### 4.2.2 Optimization of Line Node Positions with a Genetic Algorithm

A genetic algorithm (GA, see section 3.5.3), adapted to the problem, is used to optimize line node positions in each of the planning areas separately. The GA adapted to the line node placement is presented below, following the structure of a basic evolutionary algorithm (see Figure 10) and using the demand pattern of Winterthur (see Figure 37) as case study. In case of line node optimization a “population” consists of different line node plans (“individuals”) which are optimized during several iterations.

In the following the parameters and the distance function are chosen to demonstrate the procedure of line node placement. It is not the aim of the line node placement step to come already close to a global optimum in the sense of the generalized cost model. For this purpose the line node distance function (see below) has to be adjusted. At first the distance restriction is presented below.

### 4.2.3 Minimum Distances between Line Nodes as Constraint

The minimum distance between two neighboring line nodes depends on the smaller of the two demand potentials of these nodes. The demand potential of a line node is obtained through filtering the demand pattern (see Figure 37) in its surroundings by a responsiveness function (see Figure 19a) independent from other line nodes around:

\[
DP_{node,i} = \sum_{j=1}^{n} \sum_{k=1}^{n} D_{patt,j,k} \cdot R_{j,k} \quad \text{with} \quad n = \frac{d_{catch}}{l_{rast}}
\] (13)
notation:

\[ D_{\text{node}, i} \] demand potential at line node \([j, k = 0]\) compare Figure 16

\[ D_{\text{patt}, j, k} \] demand pattern at raster \(j, k\)

\[ R_{i, j} \] responsiveness according to Figure 19 a at raster \(j, k\)

\[ d_{\text{catch}} \] catchment distance

\[ l_{\text{rast}} \] length of a raster cell (in case of the slowest speed level a raster cell is a hectare)

So, it is taken into account that demand of PT declines the longer people have to walk to stations. In Figure 16 a subdivision of the filtered demand (see Figure 37) of Winterthur can be seen.

Then, minimum distances are calculated between minimal and maximal line node distances (linear combination):

\[
d_{\text{min}} (DP)_{\text{node}, i} = \begin{cases} 
  d_{\text{min}, tb} & \text{if } D_{\text{node}, i} > D_{\text{ub}} \\
  d_{\text{min}, tb} + (d_{\text{min}, ub} - d_{\text{min}, tb}) \cdot \frac{D_{\text{node}, i} - D_{\text{ub}}}{D_{\text{ub}} - D_{\text{tb}}} & \text{if } D_{\text{ub}} > D_{\text{node}, i} > D_{\text{tb}} \\
  d_{\text{min}, ub} & \text{if } D_{\text{node}, i} < D_{\text{tb}}
\end{cases}
\]

notation:

\[ d_{\text{min}} (DP)_{\text{node}, i} \] minimal distance of the next neighbor line node to line node \(i\) dependent on \(DP_{\text{node}, i}\)

\[ d_{\text{min}, tb} \] lower bound of minimal line node distance

\[ d_{\text{min}, ub} \] upper bound of minimal line node distance

\[ D_{\text{ub}} \] lower bound of demand potential

\[ D_{\text{ub}} \] upper bound of demand potential

Other functions could be applied instead of linear combination as well, e.g. quadratic or exponential relations. So, according to economies of scale, in more dense populated areas and areas with a high workplace density, line distances may be smaller than in less dense populated areas, assumed headways of the compared lines are the same.
4.2.4 Start Population

Before starting the optimization of line node plans, node positions with a demand potential (compare Figure 16) below a lower bound are excluded (see Figure 17). Minimum distances to existing line nodes (see section 4.2.3) are complied with. Existing line nodes could be the result of previous optimizations of the faster speed levels and of neighboring planning areas or the result of manual stop settings.

Figure 16 Demand potential: demand at each hectare with the demand of its corresponding catchment area. Subdivision of the demand pattern (see Figure 37) from Winterthur, filtered by a responsiveness function (see Figure 19 a). The yellow area contents values over 1’000 moves per day (the maximum is below 5’000).
Figure 17  Possible line node positions (dark raster squares). Filtered demand (see Figure 16) should be higher than a lower bound demand (here 20 moves per day) in the catchment areas of the line nodes (Demand pattern from Winterthur). Otherwise the raster squares are white.

To create a first sample of line node plans, line nodes are set starting from each of the four corners of the coordinate system filling either column by column or line by line, complying with the minimal line node distances (see Figure 18). That makes 8 different line node plans. If more than 8 plans are advantageous, they are created in mutating circular areas (see section about mutation below) of the already existing 8 line node plans. Similar as the number of generations, the size of the population depends on the size of the planning areas and the required quality of the solution. The size of the population and the number of generations should be balanced. Optimally would be a convergence of the GA in the end. The number of generations allows convergence. The number of individuals guarantees variability and prevents too early convergence.
4.2.5 Evaluation of Line Node Plans

After completing line node placement, line node plans can be evaluated by the generalized costs model (see section 1.3.2) together with a PT network designed on them. During line node placement, line node plans are pre-evaluated based on the assumption that the responsiveness of people to PT in the surroundings of a line node follows a Gaussian distribution (see [Bränd78]) depending on access distances (here linear distances with a detour factor of 1.24 [Bränd78] is used) to each line node (see Figure 19). Therefore, the demand pattern (see Figure 37) around each node of a line node plan is multiplied by a Gaussian distribution as filter function (see function (13)).
Investigation of space-time structures in public transport networks and their optimization 2010

Figure 19 Responsiveness of people to PT. It is estimated by Gaussian distributions (see [Bränd78]) dependent on walking distances (here linear distances are used) to each potential Line node. Line nodes of higher transport speed levels tend to have larger catchment areas (see [Wal73]). a) lowest level, b) middle level and c) highest speed level. The demand pattern (see Figure 37) is equivalent to 100%. A detour-factor of 1.24 [Bränd78] is considered.

Line node plans are (pre-) evaluated summing up the demand potentials (function (13)) of all nodes of the line node plan. The responsiveness $R_{i,j}$ of higher level stops (regional train and interregional train) is according to Figure 19 b and c at raster $i,j$.

If catchment areas of several adjacent line nodes overlap, the estimated demands of the corresponding hectare – line node pairs are not simply added to the objective value (demand potential) of a line node plan. The highest demand of such hectare - line node pair is added to 100% to the objective value, the second highest demand (a pair of the same hectare – but different line node) to 50% and so on. Though there can be some hectares with a demand contribution of more than 100%, the minimum distance between adjacent line nodes guarantees, that the demand contribution of a hectare can not be much higher than 100% (this mistake can be neglected).

If exact walking distances between hectares would be available they could be used as well. The relative mistake made in using linear distances with a detour factor of 1.24 [Bränd78] is in average plus/minus 20% walking time (compare [Jer04]).

4.2.6 Selection

After pre-evaluation (see above), the line node plans are selected by stochastic universal sampling proportional to their objective values (see section 3.5.3).
4.2.7 Recombination

Analogue to evolution, recombination consist in the combination of two “parent” line node plans to “children” line node plans. The “parent” pairs are chosen randomly. In recombination of two dimensional plans of line nodes, most of the adjacent line nodes should stay together and the minimum distances should additionally be guaranteed. That is why the classical discrete recombination of a GA (see 3.5.3) should not be used. Instead, the line node plans are recombined graphically (see Figure 20).

Figure 20 Graphical recombination of two line node plans. The red rectangle shows the exchanged area. If parts of the random chosen rectangle leaves the line node plan area, the rectangle is continued at the opposite side (become two rectangles). New line nodes are black, nodes of one old plan green and of the other old plan blue. Around the rectangle some gaps occur, which only partly can be filled.
Thus connected parts of the two-dimensional search space are recombined (in the terminology of evolution theory: recombination of phenotypes not of genotypes). The recombination “rectangle” is chosen randomly.

If the rectangle leaves the planning area, it is continued at the opposite side (of the planning area), to treat the entire planning area with a similar probability. At the edge of the exchanged areas overlapping line nodes appear during recombination. The minimum distances are guaranteed in deleting one of two overlapping line nodes. Hereby at the edge of the exchanged areas gaps occur. These gaps are filled, if possible, with new line nodes. Otherwise, to increase line node density, they can only be reduced by mutation (see below). In that case and if possible, line nodes are densified.

Position and size of the rectangle are chosen randomly. Its size varies between a lower and an upper bound (between e.g. 20% and 80% of a planning area edge). The rectangle should not become too small and not too large to prevent that no real recombination is performed.

### 4.2.8 Mutation

**Overview**

Analogue to biologic evolution, mutation is the stochastic change of line node plans. Like in recombination (see above), in mutation of line node plans most of the adjacent line nodes should stay together and the minimum distances also should be guaranteed.

To be able to mutate connected areas of line nodes as well as single line nodes, two methods have been created.

**Mutation of Line Nodes in a Circular Area**

This form of mutation is similar to recombination (see above). Differences are that

- “children” line node plans are randomly changed inside a circular area and that
- new nodes are created in setting them circular around a stochastic chosen circle center complying with minimal line node distances (see Figure 21).

The minimum distances are guaranteed in deleting one of two overlapping nodes. Hereby, at the edge of the circular area gaps occur. These gaps are filled, if possible, with new line nodes. Otherwise, to increase line node density, they can only be reduced by mutation of
single line nodes (see below). In that case and if possible, line nodes are densified. Only if gaps are small they can not be filled.

**Mutation of Single Line Nodes**

The mutation of single line nodes is no stochastic mutation. If possible, a single line node is moved to a place of higher demand in its direct neighborhood. The search radius is hereby restricted to the distance of the closest adjacent node. The minimum distances are guaranteed. In a random sequence, the algorithm tries to mutate (shift) each single line node once.

---

**Figure 21** Mutation of line nodes in a circular area. The red circle shows the area where new line nodes are created around a stochastic placed centre line node. New line nodes are black, line nodes of the old plan green and the circle centre is marked red. Around the circle some gaps occur, which only partly can be filled.
4.2.9 Conservation (Generation gap)

In stochastic optimization it is helpful to leave some of the line node plans unchanged respectively proliferate them. The ratio of the changed individuals (here line node plans) to that of the parents' generation is called “generation gap”. In conservation several best individuals are selected (truncation selection; see section 3.5.3).
4.2.10 Results of Line Node Placement

After repeating the steps of the GA (see above) several times, the best line node plan in one area can be seen in Figure 23. Compared with the best start plan, the GA leads to a higher demand potential of the planning area between 5 and 10 percent. After processing several planning areas of the slowest speed level (bus level), following result has been obtained (see Figure 24).

![Figure 23 Resulted plan after optimizing line nodes. Line nodes (black) in the red rectangle are saved. The other nodes are discarded, to keep influence of borders low.](image)

To get an infrastructure network, with direct connections between neighboring nodes, but no intersections of links, a triangular grid is chosen.
The triangulation for the infrastructure network follows three rules:

1. shortest possible links are chosen first
2. a lower bound is given for the acute angles
3. no intersections are allowed in one speed level

As already mentioned above, the so designed network is not perfect in the sense of minimized total costs (equation (5)). It can be used indeed, after adjusting the density-dependent minimum distance function of line nodes, as a lower bound for evaluating and optimizing public transit networks on given infrastructures.

Figure 24 Result after line node placement. The area of Winterthur has been divided into two speed levels (blue circular and black quadratic line nodes) and several planning areas. The triangle grid is added only for the bus-level, to get an acceptable infrastructure network as base for line network design. The doted rail network including stops was kept.
4.3 Line Network Design

4.3.1 Introduction

This section outlines the guided stochastic search heuristic (GSSH) that forms the core of the new approach for designing PT networks (see blocks b through h in Figure 2). The guided stochastic search heuristic is applied sequentially to the planning areas.

- A network reduction process is used for the network design process in each planning area. It starts with a network of nearly shortest lines (see section 4.3.2), but not with a network of optimal line frequencies and not optimal service provider costs. In addition to the shortest lines between potential terminal stations, nearly shortest lines are also included (see Figure 2, block b). If a nearly shortest line is already part of another line, it is not considered. Other advantageous lines or given lines can be considered as well in the network design process.

- The hectare-based OD matrix is aggregated to link corridors and line nodes depending on the local transport supply (see Figure 2, block c and section 3.4.2) on the network of nearly shortest lines. The number of stops and transport speeds on links are then adjusted while locally minimizing total transport costs.

- Preceding each reduction step, frequencies and vehicle sizes are optimized by a genetic algorithm (see Figure 2, block d).

- The network reduction process is generalized from Lee’s [Lee98] approach. Merging or shortening of lines is only possible, if the total costs of the PT system, i.e., passenger costs and operating costs, are minimized simultaneously. Directness and number of lines are only reduced, if total costs are reduced as well. Lines are merged or shortened using a combination of the ant colony heuristic and a local OD matrix. Alternatively, lines can be deleted.

- All alternatives of the line network produced in the reduction step are evaluated using a headway-based stochastic multiple route assignment (see section 3.4.3 and [PTV07]). The best alternative is selected as input for another reduction step until no further improvements are possible.

The approach is demonstrated in the following using the test problem of Figure 25.
Figure 25 Test problem (compare [Lee98]). Bus network layout with node numbers and run times on links in minutes. Lengths are not considered. The link connecting node 6 and 9 is a one direction link for bus but a two direction link for light rail. Only the 6 dark black nodes (rings) are potential terminal stations. Demand between all nodes is assumed to be 30 people in the analysis period. Perceived travel times are assumed to be according to section 5.2.4. Operating costs are 130.0 CHF/h for bus and 200.0 CHF/h for light rail. Kilometer costs are not considered. Terminal times of a line are at least 3 minutes.

4.3.2 Start Condition

The process starts (see Figure 2 block c and Figure 26) with a network of (nearly) shortest lines between a given set of potential terminal stations (PTS). This network does not include optimal line frequencies or optimal service provider costs. If potential terminal stations are not manually specified, all line nodes and other given stops are set to be potential terminal stations by default.

The infrastructure network can be any road or rail network consisting of nodes, links, turns and potential terminal stations defined as either open or closed for specific PT modes. The infrastructure network can either be artificial (compare section 4.2) or a really existing one.
Figure 26 Sample start network of nearly shortest lines for a test problem (see Figure 25). Only the 6 dark black nodes (rings) are potential terminal stations. Travel times on lines differ from shortest routes by a maximum of 5%. This makes 36 nearly shortest routes (NSR) and 10 lines in the start network (see Table 8).

A PT line consists of a sequence of nearly shortest routes (NSR) in both directions. Nearly shortest routes start and end in different potential terminal stations. Generally, several different nearly shortest routes exist between two potential terminal stations. Alternative routes can have detours which make their run time a specified percentage longer than the shortest run time. Each route has a corresponding route in its opposite direction. A route and its opposite route should have as many common nodes as possible. This constraint restricts the use of different route alignments for each direction to prevent a confusing PT network.

The nearly shortest line network contains only NSRs that are not included in longer NSRs (similar to Lee’s approach; [Lee98]). From passengers’ point of view, lines should be as long as possible to reduce the need of transfers. From operators point of view line lengths should be restricted by margins, though. If cycle times of lines are long, it is necessary to monitor schedule deviations over the entire line. Long cycle times additionally lead to long shifts,
which make crew scheduling more difficult and eventually cause higher costs. Short cycle times lead to high transfer costs for passengers and high terminal times. In setting margins, short lines are not considered during optimization and computing times can be reduced.

Generally, there are more lines in the start network than the optimal number. During the reduction process, lines can be shortened easily. By merging two lines, lines may be prolonged.

<table>
<thead>
<tr>
<th>Name</th>
<th>Headway [min]</th>
<th>Max Headway [min]</th>
<th>Cycle Time (for NSR Run Time) [min]</th>
<th>Terminal Times [min]</th>
<th>Number of Vehicle</th>
<th>Vehicle Capacity</th>
<th>Node Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>30 UpB.</td>
<td>60</td>
<td>3 2 50</td>
<td></td>
<td>[3, 6, 9], [9, 8, 6, 3]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td>30 UpB.</td>
<td>60</td>
<td>10 2 50</td>
<td></td>
<td>[7, 8, 9], [9, 8, 7]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B3</td>
<td>20 UpB.</td>
<td>80</td>
<td>6 4 50</td>
<td></td>
<td>[3, 6, 8, 7], [7, 8, 6, 3]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B4</td>
<td>20 UpB.</td>
<td>80</td>
<td>6 4 50</td>
<td></td>
<td>[3, 5, 7], [7, 5, 3]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B5</td>
<td>12 UpB.</td>
<td>84</td>
<td>8 7 50</td>
<td></td>
<td>[1, 4, 8, 9], [9, 8, 4, 1]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B6</td>
<td>20 UpB.</td>
<td>80</td>
<td>6 4 50</td>
<td></td>
<td>[1, 5, 9], [9, 5, 1]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B7</td>
<td>30 UpB.</td>
<td>60</td>
<td>10 2 50</td>
<td></td>
<td>[1, 4, 7], [7, 4, 1]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B8</td>
<td>30 UpB.</td>
<td>90</td>
<td>15 3 50</td>
<td></td>
<td>[1, 3, 6], [6, 3, 1]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B9</td>
<td>10 UpB.</td>
<td>70</td>
<td>5 7 50</td>
<td></td>
<td>[1, 5, 6], [6, 5, 1]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>20 UpB.</td>
<td>40</td>
<td>10 2 115</td>
<td></td>
<td>[6, 9], [9, 6]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8 Properties of nearly shortest lines and NSR if fractions of nearly shortest lines: sample start network for a test problem (see Figure 25 and Figure 26). Terminal times of at least 3 minutes at both terminal stations of a line are mandatory. They are part of the cycle time. The “maximum headway” is the longest headway still complying with vehicle capacities.

All links in the infrastructure network that cover sufficient demand build a set $L$. Each link out of $L$ should be part of at least one NSR. If any link out of $L$ is not part of a NSR, then the detour percentage is increased by one step (for example 5%), which adds slower routes to the set of NSRs. This process is continued until at least all links from set $L$ are included in the set of NSRs. In the example (see Figure 26) not all links out of $L$ are included in the set of NSRs to keep the set small.

All line nodes should be served. Between some line nodes, several routes exist with the same run time. If all routes would be accepted, which are up to 20% longer than the shortest route, than instead of 36 NSR 46 NSR would be generated. One of them has e.g. the node
sequence: \([1, 5, 8, 9]\). With a run time of 40 minutes it is 17.65% longer than the shortest route with 34 minutes between node 1 and node 9.

Terminal stations of a line should only be served once in its alignment (in order to avoid confusion of passengers). Ring lines are possible. Each link out of \(L\) has to be included into at least one line. No line should include a link more often than twice.

### 4.3.3 Adjustment of Stop Densities on Links and OD matrix Aggregation

For each planning area, preceding the reduction process, stop densities on links are adjusted. While adjusting stop densities on links, generalized costs are locally minimized. Afterwards the OD matrix is aggregated to link corridors and line nodes depending on the local transport supply (see section 3.4.2).

While adjusting stop densities on links, transport speeds are reset depending on the dwell times and maximal allowed/possible link speeds. For each link corridor,

- the access and egress walking times \((t_{\text{ae}})_{\text{link}}\) for those passengers which are boarding or alighting at the links stops,

- the passengers on-link in-vehicle travel times \((t_{\text{iv}})_{\text{on-link}}\) and

- the on-link operating costs

are minimized.

The access and egress walking paths are assumed to be orthogonal and calculated according to section 3.4.2. Two constraints are that all PT lines serve all stations of each link and that all lines of the same mode have the same on-link travel speed. The second constraint is just a simplification to reduce complexity of data handling. According to the average percentage of boarding and alighting passengers per stop, the dwell time \(t_{\text{dwell}}\) is set linear between \(t_{\text{dwell, min}}\) and \(t_{\text{dwell, max}}\):

\[
t_{\text{dwell}} = \max \left( \min \left( \frac{n_{\text{pass, av.}}}{c_{\text{vehicle}}} \cdot t_{\text{dwell, max}}, t_{\text{dwell, max}} \right), t_{\text{dwell, min}} \right)
\]

(15)

notation:
Acceleration of vehicles is assumed to be $1\text{m/s}^2$. For buses, maximum speed is assumed to be two thirds of the average link speed of cars. According to the number of stops on links, the average dwell times and the maximum link speeds, the average travel speeds on the corresponding links are faster or slower.

### 4.3.4 Headway Optimization using a Genetic Algorithm

Integrating headway optimization in the network reduction process (see Figure 2, block d) helps to design more efficient public transport networks ([Biel02] and [Zhao06]). The basis for headway optimization is a list of accepted headways. Accepted headways are defined based on passenger requirements. From the passengers’ perspective, headways should not be too long and they should be evenly divisible into 60. This makes schedules easier for passengers to memorize since vehicles arrive each hour at the same minute.

The PT network design approach proposed in this paper uses the breeder GA (see section 3.5.3) to perform the headway optimization. The breeder GA directly processes real variables rather than the binary variables used in many GAs. The GA assigns one headway (variable) out of the list of accepted headways to each line of the PT network (individual) and then varies them. The GA stores and manipulates variables (headways) as data chains called “chromosomes”. Neighborhoods of headways in these chains are organized according to the number of identical stops in the corresponding line alignments.

The maximum headway for a PT line is the headway (from the list of accepted headways) short enough to cope with the passenger flow at the lines maximum load section, using vehicles with the largest available capacity at the same time. During headway optimization, the headway on a given line can be decreased to unburden other lines, if the list of accepted headways allows this. Vehicle capacities are always chosen as small as possible but large enough that they can still cope with the passenger flow at the maximum load section.

### 4.3.5 Network Reduction Process

The start condition generates a large number of potential lines for the given geographic area. The next step in the process consists of reducing the number of lines based on the objective function (5). If some lines should not be changed, they are never chosen for reduction or for merging with “multi-connected” lines. Two lines are defined to be multi-connected if they have at least two common nodes or one common node close to their ends. The larger the
area and the smaller the demand, the more the network is reduced. The network reduction process is generalized from Lee’s method [Lee98]. During the reduction process, the directness of trips generally declines. The increased in-vehicle travel time of passengers is over-compensated though by shorter waiting times due to shorter headways.

Following Lee’s process, the line with the smallest demand at its maximum load section is chosen for reduction. All overlapped lines, which have a common part, called trunk, with the chosen line, are candidates for merging. In Lee’s method, different branches of two overlapped lines are combined into one new line. All possibilities for combining the nodes of two branches into a new line are compared with each other, based on average network travel time. The common trunk is kept. The merged line with the shortest average travel time is chosen as new line.

For the approach proposed in this research, a line is chosen for shortening or merging according to the following probability (for an example see Table 9):

$$p_l = \frac{\exp \left( -\frac{f(t_{iv,l}, t_{cycle,l}, h_l)}{f(t_{iv,l}, t_{cycle,l}, h_{max})} \right)^{\beta} \right)}{\sum_{l=1}^{n_l} \exp \left( -\frac{f(t_{iv,l}, t_{cycle,l}, h_l)}{f(t_{iv,l}, t_{cycle,l}, h_{max})} \right)^{\beta}$$

(16)

where:

$$f(t_{iv,l}, t_{cycle,l}, h_l) = \sum_{j=1}^{n_{pass}} t_{iv,l,j} \cdot \left[ \frac{r_p + \frac{h_l}{h_{max}} \cdot (1 - r_p)}{t_{cycle,l}} \right]$$

(17)

$t_{iv,l,j}$ in-vehicle travel time on line $l$ of passenger $j$ out of all $n_{pass}$ passengers

$t_{cycle,l}$ cycle time of line $l$

$h_l$ headway of line $l$

$h_{max}$ maximum (longest) headway of all lines,

$\beta = 3$

The probability $p_l$ of line $l$ out of $n_l$ lines to be chosen for reduction is calculated by using function (16) on the result of function (17) $f(t_{iv,l}, t_{cycle,l}, h_l)$ of line $l$. Function (16) is used to increase the probability that lines evaluated as “bad” will be chosen for shortening or merging. The
highest value of equation (17) \( f(t_{iv}, t_{cycle}, h)_{max} \) is the line evaluated as "best" out of all lines in the PT network and the planning area. The cycle time \( t_{cycle,l} \) is the scheduled time interval between a departure of a PT vehicle at the terminal station and its next departure at the same terminal after serving line \( l \) in both directions. The ratio \( r_p \) of passenger costs \( C_p \) to total costs \( C_{total(l,h)} \) is derived from a preceding network evaluation according to equation (5).

The first part of equation (17) evaluates lines from the passengers’ point of view. If a line is attractive, average use is high. The second part of equation (17) reflects the operator’s point of view. A line is attractive from the operator’s perspective if average vehicle occupancy is high. Therefore, equation (17) multiplies average use of a line by its headway, which is normalized by the longest headway of all lines.

<table>
<thead>
<tr>
<th>Line Name</th>
<th>( p_l )</th>
<th>( f(t_{iv}, t_{cycle}, h) )</th>
<th>( \frac{\sum_{j=1}^{n_{max}} t_{iv,l,j} [h]}{t_{cycle,l}} )</th>
<th>( h_l )</th>
</tr>
</thead>
<tbody>
<tr>
<td>B2</td>
<td>0.1824</td>
<td>1.0958</td>
<td>104.25</td>
<td>80</td>
</tr>
<tr>
<td>B3</td>
<td>0.0500</td>
<td>1.9277</td>
<td>458.50</td>
<td>200</td>
</tr>
<tr>
<td>B5</td>
<td>0.2383</td>
<td>0.0431</td>
<td>78.67</td>
<td>75</td>
</tr>
<tr>
<td>T1</td>
<td>0.5293</td>
<td>0.0093</td>
<td>12.33</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 9 Selection probabilities of lines to be chosen for reduction for a test problem and after the 8th reduction step (see Figure 31 and Figure 32); \( h_{max} = 30; \beta = 3; r_p = 0.7613 (=14'321/18'811) \).

One problem with this approach is that lines are sequentially chosen for reduction (sequence problem). Therefore, this method does not always reduce the line evaluated as the “worst”. Instead it chooses a line according to its probability given by equation (16), so that it increases the variability of the resulting networks and makes the influence of the sequence problem more visible.

If a line is chosen for reduction, there are five possibilities to continue:

1. The line is deleted.
2. The line is deleted together with a multi-connected line.
3. The line is merged with a multi-connected line.
4. The line is shortened.
5. The line is kept.
All multi-connected lines of the line chosen for reduction are checked for merging or deletion. The best alternative of the five possibilities is chosen for adjusting the line network according to objective function (5). Therefore, for all alternatives, at least two assignments of the OD matrix to the network are calculated: one or more to adjust headways and another to calculate the objective function (5). The best alternative of the five possibilities is selected as input for another reduction step until no further improvements are possible.

To find new line alignments, an ant colony optimization algorithm (see section 4.3.6) is used in combination with a local OD matrix. All potential terminal stations of the line to be shortened or of the two multi-connected lines are potential terminal stations for the new line alignment. Not every station needs to be included in a new line. Some stations may not have enough demand in a particular direction or the demand may already be sufficiently covered by other lines.

Merging, shortening or deletion of lines is only allowed if total costs for the PT system are minimized as a result of the reduction, i.e., the approach allows the directness of routes and number of routes to be reduced only if the total PT system costs are reduced as well.

In the beginning of the network reduction process, there may be many more lines than optimal, depending on the density of the infrastructure network, the demand and the size of the planning area. In this case, some of the worst lines defined based on equation (17) can be deleted directly without adapting headways in advance (only alignments are compared, not headways); this reduces computing time. However, if there are only a small number of lines, headways can be optimized before deleting or keeping a line (chosen according to equation (16)). After another reduction sequence, all five possibilities for reduction are available.

The results for the test problem are displayed in Table 12, Figure 31 and Figure 32. Total costs were reduced in every step. Compared to the start network total costs were reduced by 7%. The lowest “total passenger travel time” appeared after reduction step 5, the smallest total fleet size after reduction step 7. After reduction step 8 no reduction was advantageous any more. During the test problem, the network was evaluated less than 1'000 times. The reduced evaluations during ACO were performed less than 10'000 times.
In a first reduction step for the test problem (see Figure 25) line T1 is kept in a second step the lines B2 and B6 are merged using ACO (see below) to a new line B2.

Table 10  
Line properties: after the second reduction step for the test problem (see Figure 25 and Figure 26).
Figure 28  In a third reduction step for the test problem (see Figure 25) lines B1 and B8 are deleted.

<table>
<thead>
<tr>
<th>Line Name</th>
<th>Headway [min]</th>
<th>Max Headway [min]</th>
<th>Cycle Time (Run Time)</th>
<th>Terminal Times [min]</th>
<th>Number of Vehicle</th>
<th>Vehicle Capacity</th>
<th>Node Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>B2</td>
<td>20</td>
<td>30</td>
<td>80</td>
<td>3</td>
<td>4</td>
<td>50</td>
<td>[7, 8, 9, 5], [5, 9, 8, 7]</td>
</tr>
<tr>
<td>B3</td>
<td>7.5</td>
<td>12</td>
<td>75</td>
<td>3.5</td>
<td>10</td>
<td>50</td>
<td>[3, 6, 8, 7], [7, 5, 6, 3]</td>
</tr>
<tr>
<td>B4</td>
<td>20</td>
<td>30</td>
<td>80</td>
<td>6</td>
<td>4</td>
<td>50</td>
<td>[3, 5, 7], [7, 5, 3]</td>
</tr>
<tr>
<td>B5</td>
<td>7.5</td>
<td>12</td>
<td>75</td>
<td>3.5</td>
<td>10</td>
<td>50</td>
<td>[1, 4, 8, 9], [9, 8, 4, 1]</td>
</tr>
<tr>
<td>B7</td>
<td>30</td>
<td>UpB.</td>
<td>60</td>
<td>10</td>
<td>2</td>
<td>50</td>
<td>[1, 4, 7], [7, 4, 1]</td>
</tr>
<tr>
<td>B9</td>
<td>10</td>
<td>12</td>
<td>70</td>
<td>5</td>
<td>7</td>
<td>50</td>
<td>[1, 5, 6], [6, 5, 1]</td>
</tr>
<tr>
<td>T1</td>
<td>30</td>
<td>UpB.</td>
<td>30</td>
<td>5</td>
<td>1</td>
<td>115</td>
<td>[6, 9], [9, 6]</td>
</tr>
</tbody>
</table>

Table 11  Line properties: after the third reduction step for the test problem (see Figure 25 and Figure 26).
4.3.6 Ant Colony Optimization

Ant colony optimization (ACO) is an appropriate method for optimizing the alignment of a single PT line. The ACO combines NSRs (section 4.3.2) between potential terminal stations into new lines. This problem has similarities to the well-known traveling salesman problem, which was one of the first applications. The main difference is that the objective is not the shortest line between a set of potential terminal stations, but the lowest generalized costs (see equation (5)) for the entire PT network.

Several lines are built and tested during one iteration of the ACO, later called a “generation”. At first, a start terminal station is chosen for each line. Then, NSRs are chosen successively. A certain number of potential terminal stations (PTS) out of a set \( \mathbf{A} \) have to be served once. It is possible to vary the sequence of the service and the NSR between the PTS. NSRs in each step are chosen randomly. The probability for an NSR to be chosen as the next part of a line \( p \) is calculated as follows:

\[
p_{q,j}^{(g)}(k) = \begin{cases} 
\frac{\tau_{q,j}^{(g)} \cdot \eta_{q,j}}{\sum_{j \notin \text{tabu}_{p,k}} \tau_{q,j}^{(g)} \cdot \eta_{q,j}} & \text{if } j \notin \text{tabu}_{p,k} \\
0 & \text{else}
\end{cases}
\]  

(18)

notation:

\( p_{q,j}^{(g)}(k) \) Probability of NSR between PTS \( i \) and \( j \) to be added to line \( p \) of generation \( g \) in step \( k \), dependent on the left PTS of set \( \mathbf{A} \) ( \( j \notin \text{tabu}_{p,k} \) ) which still should be served.

\( \tau_{q,j}^{(g)} \) pheromone value of NSR between PTS \( i \) and \( j \) in generation \( g \) (see below).

\[ \eta_{q,j} = \frac{(n_{\text{PTS}} - 1)}{t_{\text{cycle},q}} \]  

(19)

\( t_{\text{cycle},q} \) cycle time on (“length” of) NSR \( q \)

\( n_{\text{PTS}} \) number of PTS of set \( \mathbf{A} \) along NSR \( q \).

\( \text{tabu}_{p,k} \) PTS of set \( \mathbf{A} \) already included into line \( p \) at step \( k \).

The design of new lines depends on the pheromone values \( \tau_{q,j}^{(g)} \) and on other spatial information. They are adapted during the ACO. Before starting, the pheromones for each NSR are initialized to the same value between 0 and 1. After one generation all lines are evaluated by the objective function (5) using a local OD matrix (see section 4.3.7). Pheromone
values $\tau^{(g)}_{q,i,j}$ are updated according to the values of the objective function $C_{total}(l,h)^{(g)}_p$ as follows:

$$
\tau^{(g+1)}_{q,i,j} = \begin{cases} 
\rho \cdot \tau^{(g)}_{q,i,j} + \omega^{(g)}_{q,i,j} \cdot \phi^{(g)}_{q,i,j} & \text{if } \text{NSR}_q \text{ is part of any line in generation } g \\
\tau^{(g)}_{q,i,j} & \text{else}
\end{cases}
$$

(20)

$$
\omega^{(g)}_{q,i,j} = \frac{1}{n_{NSR_q \in \text{lines}^{(g)}}} \sum_{p=1}^{i} \left( - C_{total}(l,h)^{(g)}_p \right)_{\text{normalized}} 
\begin{cases} 
\text{if } \text{NSR}_q \in \text{lines}^{(g)} \\
0 & \text{else}
\end{cases}
$$

(21)

$$
\phi^{(g)}_{q,i,j} = \max \left( b_{low}, \frac{1}{n_{\text{link}_{q,i}}} \sum_{\text{link}_{q,i}=1}^{n_{\text{link}_{q,i}}} Q_{q,\text{link}_{q,i}} \right)_{\text{normalized}}
$$

(22)

notation:

$\rho$ reduces the pheromone value of last generation (evaporates the pheromones) $\rho \in (0,1)$.

$\lambda$ number of lines in one generation.

$C_{total}(l,h)^{(g)}_p$ values of the objective function for line $p$ (negative: to change the smallest values into largest)

$n_{NSR_q \in \text{lines}^{(g)}}$ number of NSR_q \in \text{lines}^{(g)}

lines^{(g)} \text{ in generation } g \text{ generated lines}

\text{normalized} \text{ linear transferred to values between 0 and 1. In each generation } g \text{ the highest value equals 1.}

$n_{\text{link}}$ number of links on NSR_q

$Q_{q,\text{link}_{q,i}}$ passenger flow on \text{link}_{q,i} on NSR_q (passengers per basic period)

$b_{low}$ lower bound, $b_{low} > 0$

Without equation (22), pheromone values would depend only on the evaluation of the entire line $p$ (global information of equation (21)). In PT most passengers do not travel from one terminal of a line to its opposite terminal. Evaluation of the single NSRs as parts of a line themselves should influence pheromone values as well to also include information about local route quality into line building. As a local quality measure during pheromone update the average passenger flow of all lines^{(g)} on a NSR_q is used (equation (22)). For example, Yang et al. [Yang07] included e.g. a factor for local direct traveler density as local quality measure.
If a station is a terminal station of a line evaluated as “good”, it’s probable this station will be chosen in the next generations as a start terminal station. The selection probability is calculated analogously to pheromone updates (equation (20)). However, only information about entire lines (equation (21)) is considered for this evaluation. The genetic algorithms (GA) stochastic universal sampling method [Baker87] is used to select start terminal stations in proportion to their selection probability.

Once a specified number of generations have been performed, the process is stopped and the best line is compared with other solutions of the ACO algorithm, as well as with possibilities to keep a line or delete it. The closer the network reduction process comes to its end (until no further improvement is possible), the more important the merging of lines becomes (in the beginning, lines are deleted rather than merged). Therefore, at the end of the process, the number of generations of the ACO could be increased. The absolute number of generations and the number of individuals (designed lines) in one generation depend on the available computing time and on the expected quality of the results of ACO.

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Total passenger travel time [min]</th>
<th>Total fleet size</th>
<th>Number of lines</th>
<th>Sum of Terminal Times (minimum, if all terminal times are 3 minutes) [min]</th>
<th>0-transfer trips [%]</th>
<th>1-transfer trips [%]</th>
<th>2-transfer trips [%]</th>
<th>Total passenger travel time costs [CHF]</th>
<th>Operating costs [CHF]</th>
<th>Total costs [CHF]</th>
<th>Network illustrated in</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (Start network)</td>
<td>65’496</td>
<td>36</td>
<td>10</td>
<td>460 (186)</td>
<td>66.0</td>
<td>33.9</td>
<td>0.1</td>
<td>15’391</td>
<td>4’750</td>
<td>20’141</td>
<td>Figure 26</td>
</tr>
<tr>
<td>1</td>
<td>62’395</td>
<td>38</td>
<td>9</td>
<td>344 (198)</td>
<td>71.5</td>
<td>28.1</td>
<td>0.4</td>
<td>14’662</td>
<td>5’010</td>
<td>19’672</td>
<td>Figure 27</td>
</tr>
<tr>
<td>3</td>
<td>60’603</td>
<td>38</td>
<td>7</td>
<td>286 (192)</td>
<td>68.5</td>
<td>31.4</td>
<td>0.0</td>
<td>14’241</td>
<td>5’010</td>
<td>19’251</td>
<td>Figure 28</td>
</tr>
<tr>
<td>5</td>
<td>59’148</td>
<td>39</td>
<td>6</td>
<td>272 (180)</td>
<td>74.4</td>
<td>25.4</td>
<td>0.2</td>
<td>13’899</td>
<td>5’140</td>
<td>19’039</td>
<td>Figure 29</td>
</tr>
<tr>
<td>7</td>
<td>60’942</td>
<td>34</td>
<td>5</td>
<td>162 (102)</td>
<td>80.8</td>
<td>19.2</td>
<td>0.0</td>
<td>14’321</td>
<td>4’490</td>
<td>18’811</td>
<td>Figure 30</td>
</tr>
<tr>
<td>8</td>
<td>60’090</td>
<td>34</td>
<td>4</td>
<td>156 (108)</td>
<td>83.6</td>
<td>16.3</td>
<td>0.0</td>
<td>14’121</td>
<td>4’490</td>
<td>18’611</td>
<td>Figure 31</td>
</tr>
</tbody>
</table>

Table 12 Network properties for all reduction steps (with changes in the network) of the test problem (see Figure 25 and Figure 26).
Figure 29 In a 5th reduction step for the test problem (see Figure 25) line B7 and B4 are merged to a new line B7.

<table>
<thead>
<tr>
<th>Line Name</th>
<th>Headway [min]</th>
<th>Max Headway [min]</th>
<th>Cycle Time (Run Time) [min]</th>
<th>Terminal Times [min]</th>
<th>Number of Vehicle</th>
<th>Capacity</th>
<th>Node Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>B2</td>
<td>30</td>
<td>UpB.</td>
<td>90</td>
<td>8</td>
<td>3</td>
<td>50</td>
<td>[7, 8, 9, 5], [5, 9, 8, 7]</td>
</tr>
<tr>
<td>B3</td>
<td>7.5</td>
<td>12</td>
<td>75</td>
<td>3.5</td>
<td>10</td>
<td>50</td>
<td>[3, 6, 8, 7], [7, 8, 6, 3]</td>
</tr>
<tr>
<td>B5</td>
<td>7.5</td>
<td>15</td>
<td>75</td>
<td>3.5</td>
<td>10</td>
<td>50</td>
<td>[1, 4, 8, 9], [9, 8, 4, 1]</td>
</tr>
<tr>
<td>B7</td>
<td>15</td>
<td>30</td>
<td>120</td>
<td>6</td>
<td>8</td>
<td>50</td>
<td>[1, 4, 7, 5, 3], [3, 5, 7, 4, 1]</td>
</tr>
<tr>
<td>B9</td>
<td>10</td>
<td>12</td>
<td>70</td>
<td>5</td>
<td>7</td>
<td>50</td>
<td>[1, 5, 6], [6, 5, 1]</td>
</tr>
<tr>
<td>T1</td>
<td>30</td>
<td>UpB.</td>
<td>30</td>
<td>5</td>
<td>1</td>
<td>115</td>
<td>[6, 9], [9, 6]</td>
</tr>
</tbody>
</table>

Table 13 Line properties: after the 5th reduction step for the test problem (see Figure 25 and Figure 26).
In the 7th reduction step for the test problem (see Figure 25) lines B3 and B9 are merged to a new loop line B3.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>B2</td>
<td>20</td>
<td>30</td>
<td>80</td>
<td>3</td>
<td>4</td>
<td>50</td>
<td>[7, 8, 9, 5], [5, 9, 8, 7]</td>
</tr>
<tr>
<td>B3</td>
<td>10</td>
<td>12</td>
<td>200</td>
<td>6</td>
<td>20</td>
<td>50</td>
<td>[5, 6, 3, 6, 7, 4, 1, 5], [5, 1, 4, 7, 8, 6, 3, 6, 5]</td>
</tr>
<tr>
<td>B5</td>
<td>15</td>
<td>30</td>
<td>75</td>
<td>3.5</td>
<td>5</td>
<td>50</td>
<td>[1, 4, 8, 9], [9, 8, 4, 1]</td>
</tr>
<tr>
<td>B7</td>
<td>30</td>
<td>UpB.</td>
<td>120</td>
<td>6</td>
<td>4</td>
<td>50</td>
<td>[1, 4, 7, 5, 3], [3, 5, 7, 4, 1]</td>
</tr>
<tr>
<td>T1</td>
<td>30</td>
<td>UpB.</td>
<td>30</td>
<td>5</td>
<td>1</td>
<td>115</td>
<td>[6, 9], [9, 6]</td>
</tr>
</tbody>
</table>

Table 14 Line properties: after the 7th reduction step for the test problem (see Figure 25 and Figure 26).
In the 8th reduction step for the test problem (see Figure 25) line B7 is deleted.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>B2</td>
<td>10</td>
<td>20</td>
<td>80</td>
<td>3</td>
<td>8</td>
<td>50</td>
<td>[7, 8, 9, 5], [5, 9, 8, 7]</td>
</tr>
<tr>
<td>B3</td>
<td>10</td>
<td>12</td>
<td>200</td>
<td>6</td>
<td>20</td>
<td>50</td>
<td>[5, 6, 3, 6, 8, 7, 4, 1, 5], [5, 1, 4, 7, 8, 6, 3, 6, 5]</td>
</tr>
<tr>
<td>B5</td>
<td>15</td>
<td>30</td>
<td>75</td>
<td>3.5</td>
<td>5</td>
<td>50</td>
<td>[1, 4, 8, 9], [9, 8, 4, 1]</td>
</tr>
<tr>
<td>T1</td>
<td>30</td>
<td>UpB.</td>
<td>30</td>
<td>5</td>
<td>1</td>
<td>115</td>
<td>[6, 9], [9, 6]</td>
</tr>
</tbody>
</table>

Table 15  Line properties: after the 8th reduction step for the test problem (see Figure 25 and Figure 26).
4.3.7 Local OD Matrix during Ant Colony Optimization

Many different PT line alignments are evaluated during the ACO using equation (5). To reduce computing times, evaluations during the ACO are based on a local OD matrix and on a headway-based all-or-nothing assignment (see section 3.4.1). For evaluations in the case of loops in the line alignment, transfer penalties are added to transfer times (compare equation (6)). The OD matrix is reduced to the demand, which uses the lines being tested for shortening or merging. Furthermore, OD zones are limited to the nodes of these lines.

Not every station needs to be served by the new line alignment. For PT network evaluation in such cases, travel times between not served OD relations are obtained from the entire line network without the lines chosen for shortening or merging. The additional service provider costs due to shorter headways of other lines are added evenly as additional time costs. Because of this simplification, only lines with a similar number of potential terminal stations are compared with each other during the ACO using the local OD matrix.

The PT lines evaluated as “best” using the local OD matrix are included in the entire line network and then evaluated more accurately in a second step using the OD matrix of the corresponding planning area. Lines with a different number of potential terminal stations are compared with each other using the more accurate evaluation of the entire line network.
4.3.8 Final Stop Placement on Lines

After completing line alignments (see Figure 2, block b through h), the line node placement method of section 4.2 can be used to place stops on the existing transport infrastructure network according to the given on-link stop densities (see section 4.3.3) and according to the distribution of the demand.

The minimum on-link stop distances are in general shorter than the line node distances. On each link corridor, the minimum distances are adjusted till the given stop density could be achieved at least on a straight line.

The PT network of Winterthur (see section 5.2 - 5.3) is used to demonstrate the final stop placement. Possible on-link stop positions are illustrated in Figure 33. In Figure 34 the existing stop positions are displayed, in Figure 35 the new stop positions after a final stop placement can be seen. Minimal stop distances (between 200 and 500m) are complied with. The results are discussed in section 5.3.5.

Figure 33 Possible on-link stop positions (dark raster squares, compare Figure 17). Filtered demand (see Figure 16) should be higher than a lower bound demand (here 10 moves per day) in the catchment areas of the possible stop positions (Demand pattern from Winterthur). Otherwise the raster squares are white. Around the red line node circles minimum stop distances are considered.
Figure 34 Existing stop placement for the PT line network 2008 with passenger flows ($f_{cost} = 1.0$, 165 stops).

Figure 35 New automated stop placement on link corridors for the PT line network 2008 with passenger flows ($f_{cost} = 1.0$, 136 stops).
5 Case Studies

5.1 Overview on the Case Studies

The PT network design processes were applied to two test problems in Switzerland:

- the city Winterthur, Canton of Zurich, and to
- a small benchmark problem from Mandl [Mandl79].

For the first case study on the bus network of Winterthur, the transport model of the Canton of Zurich [Vrtic05] was used. The model is presented in section 5.2.2. In section 5.2.4 the corresponding parameters of the objective function are specified. In sections 5.3 and 5.4 the results of the first case study are discussed. The case study is summarized in section 5.5. The core of the design process is compared in section 5.6 with the results of different approaches using the small benchmark problem of Mandl.

5.2 Introduction to the Case Study of Winterthur

5.2.1 PT Network of Winterthur 2008

The city Winterthur, Canton of Zurich, has about 100’000 inhabitants. Many people commute to Winterthur from its rural surroundings. Additionally, strong commuter relations exist between Winterthur and Zurich, at the morning mainly in the direction of Zurich. Its size and its rural surroundings make Winterthur suitable for first tests on a middle size PT network design problem. However, Winterthurs radial PT network with little tangential demand makes it a challenge to improve the line alignments.

To simplify the optimization a little and to increase the search space, the existing trolleybus network (Lines 1, 2, 3 and 4) is treated as if all busses use diesel engines and thus are independent from catenaries. The regional bus lines (Lines 9, 610, 611, 615, 660, 665, 670, 676, 680) including the respective headways are taken as given (fixed). In contrast to the network 2008, regional busses stop at all city bus stops.

The PT network of Winterthur, realized in 2008, is illustrated in several Figures. In Figure 36 the passenger flows are shown including the number of transfers. At Winterthurs main station nearly 5000 people transfer in the morning peak hour between 7.00 and 8.00 am. The fre-
Investigation of space-time structures in public transport networks and their optimization

The PT network of Winterthur and its region is operated by STADTBUS Winterthur, belonging to the cities municipality. The organizations profit and loss statement is shown in Table 16. The rail network is operated by the Swiss Federal Railways (SBB). Both cooperate together with other operators in the Canton of Zurich’s transport association (ZVV).
Revenues

<table>
<thead>
<tr>
<th>Description</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income from reserve</td>
<td>22'000</td>
</tr>
<tr>
<td>Income from fares (ca. 52%) and subsidies (ca. 48%) from ZVV</td>
<td>37'348'000</td>
</tr>
<tr>
<td>Auxiliary income (advertisement, rental income, etc.)</td>
<td>2'353'000</td>
</tr>
<tr>
<td><strong>Total revenue</strong></td>
<td><strong>39'723'000</strong></td>
</tr>
</tbody>
</table>

Expenses

<table>
<thead>
<tr>
<th>Description</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct labor</td>
<td>24'381'000</td>
</tr>
<tr>
<td>Non-wage labor</td>
<td>381'000</td>
</tr>
<tr>
<td>Vehicle costs</td>
<td>3'151'000</td>
</tr>
<tr>
<td>Administrative expense</td>
<td>1'093'000</td>
</tr>
<tr>
<td>Markdown and interests</td>
<td>3'953'000</td>
</tr>
<tr>
<td>Handover of auxiliary income to ZVV</td>
<td>2'353'000</td>
</tr>
<tr>
<td>Other expenses (real estate, materials, insurance, etc.)</td>
<td>2'513'000</td>
</tr>
<tr>
<td><strong>Total expenses</strong></td>
<td><strong>37'825'000</strong></td>
</tr>
</tbody>
</table>

**Annual result** 1'888'000

Table 16 Profit and loss statement in parts of STADTBUS Winterthur 2005 [CHF]
source: STADTBUS Winterthur.
(ZVV: “Zürcher Verkehrsverbund”, transport association of the Canton of Zurich)

5.2.2 Transport Model of the Canton of Zurich

The quality of the OD matrix is, besides the assignment method (see section 3.4), the other important part of the transport model. The OD-data and some parameters of the transport model of the Canton of Zurich [Vrtic05] were used in the sections 5.3 and 5.4 as sample for line network planning.

The transport model of Canton of Zurich's PT network (see [Vrtic05]), with raw data from SBB, ZVV, ARE and BFS, is a macroscopic model for an average working day. It follows the classic sequence of the four sub-models: trip generation, trip distribution, modal split and assignment. Trip distribution and modal split were performed simultaneously though using a Nested-Logit-Model.

The area of the Canton of Zurich was divided in to 878 OD-zones with additional 26 border zones. Demand of each zone was assumed to be concentrated in its center of population. Each center is connected with one or more PT stations. 17 OD-groups were built and each assigned to one out of five trip purposes: work, education, business, shopping and leisure.
The assignment method for PT is schedule-based, the underlying route choice model is performed with the PTV tool VISUM according to Lohse (see section 3.4.4 and [PTV07]).

After calibration of the PT transport model there were still some relative errors in the flow calculation of 8% in average (compared to passenger counts on links, see [Vrtic05]). These relative errors are especially high for PT links with small demand i.e. for bus and light rail.

5.2.3 Adaptation of the Model to Needs of Network Design

Instead of the schedule-based assignment method used in the transport model of the Canton of Zurich [Vrtic05] a headway-based model (see section 3.4.3 and [PTV07]) was used for this study to avoid calculating a schedule optimization preceding any network evaluation.

Spatial resolution of the transport model of the Canton of Zurich [Vrtic05] is not fine enough to accurately model access and egress to PT stations. Resolution of each zone was increased, using the hectare information about resident population (Volkszählungsdaten 2000) and working places (all working places, working places in education and working places in retail trade; Betriebszählungsdaten 2005). Hereby the historical segregation of PT and car users according to PT accessibility was neutralized. Totally 2241 basic zones (hectares) were built, connected with each other by 469’234 OD-relations. If all PT trips are summed up for every hectare, which starts or ends there, for the city of Winterthur the pattern of Figure 37 occurs. The basic zones were assigned according to section 3.4.2 to 138 line nodes and link corridors or in case of the existing stops to 165 stops.
Figure 37  Demand pattern of Winterthur. “Demand” is here the sum of trip origins and trip destinations at each hectare. The OD matrix of Winterthur is derived from the transport model of the Canton of Zurich [Vrtic05].

Figure 38  General load profiles over the daytime according to trip purposes for car and PT together in the city of Zurich (generated from [CEOZ07]).
In the transport model of the Canton of Zurich, the OD matrices are symmetric. The flow directions between residence areas and activity areas are not given. They were estimated using OD-potentials of every OD-relation. It is possible with this information to estimate the trip directions according to day times.

Together with the general trip profile over the daytime according to trip purpose and transport mode of the city of Zurich (see Figure 38), it is possible to estimate demand e.g. during the peak hour.

Table 17 shows the accuracy of the transport model at the maximal loaded sections during morning peak hour. Only north-west branches of lines from “Hauptbahnhof Winterthur” are analyzed. For the regional lines (660, 665, 670 and 676) passenger counts were not given. Line 11 is parallel to line 660 and not in service for the entire day. Line 12 is only in service at weekends. Several different assignment methods are compared with average vehicle occupancy data from passenger counts.

Figure 39 Most loaded sections of the lines north-west branches during morning peak hour (PT network 2007, source STADTBUS Winterthur).
Investigation of space-time structures in public transport networks and their optimization

Data source

<table>
<thead>
<tr>
<th>Line 1</th>
<th>Line 2</th>
<th>Line 3</th>
<th>Line 4</th>
<th>Line 5</th>
<th>Line 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passenger counts during morning peak hour (sample from automatic counts; source: STADTBUS Winterthur spring 2007)</td>
<td>393</td>
<td>550</td>
<td>218</td>
<td>63</td>
<td>91</td>
</tr>
</tbody>
</table>

Model of Canton of Zurich [Vrtic05]:

<table>
<thead>
<tr>
<th>Line 1</th>
<th>Line 2</th>
<th>Line 3</th>
<th>Line 4</th>
<th>Line 5</th>
<th>Line 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>The original schedule-based assignment (see section 3.4.4) 11.5% demand of a workday</td>
<td>319</td>
<td>470</td>
<td>319</td>
<td>179</td>
<td>47</td>
</tr>
<tr>
<td>Deviation from passenger counts: weighted mean 32%</td>
<td>-19%</td>
<td>-15%</td>
<td>+46%</td>
<td>+184%</td>
<td>-48%</td>
</tr>
<tr>
<td>Changed to headway-based assignment (see section 3.4.3) 10.5% of a workday</td>
<td>252</td>
<td>401</td>
<td>386</td>
<td>206</td>
<td>28</td>
</tr>
<tr>
<td>Deviation from passenger counts: weighted mean 52%</td>
<td>-36%</td>
<td>-27%</td>
<td>+77%</td>
<td>+28%</td>
<td>-69%</td>
</tr>
</tbody>
</table>

Adapted model on a hectare basis with fixed headways:

<table>
<thead>
<tr>
<th>Line 1</th>
<th>Line 2</th>
<th>Line 3</th>
<th>Line 4</th>
<th>Line 5</th>
<th>Line 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schedule-based assignment (see section 3.4.4) between existing stops. 13.9% of workday commuters and 2.3% other trip purposes (adjusted values between 7 and 8 AM according to Figure 38).</td>
<td>233</td>
<td>521</td>
<td>237</td>
<td>106</td>
<td>149</td>
</tr>
<tr>
<td>Deviation from passenger counts: weighted mean 28%</td>
<td>-41%</td>
<td>-5%</td>
<td>+9%</td>
<td>+69%</td>
<td>+64%</td>
</tr>
<tr>
<td>Headway-based assignment (see section 3.4) between existing stops. 12.5% of workday commuters and 2.1% other trip purposes (adjusted values between 7 and 8 AM according to Figure 38).</td>
<td>278</td>
<td>448</td>
<td>281</td>
<td>76</td>
<td>162</td>
</tr>
<tr>
<td>Deviation from passenger counts: weighted mean 32%</td>
<td>-29%</td>
<td>-19%</td>
<td>+29%</td>
<td>+21%</td>
<td>+78%</td>
</tr>
</tbody>
</table>

Adapted model on a hectare basis after headway optimization:

<table>
<thead>
<tr>
<th>Line 1</th>
<th>Line 2</th>
<th>Line 3</th>
<th>Line 4</th>
<th>Line 5</th>
<th>Line 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Headway-based assignment (see section 3.4) between existing stops. 12.6% of workday commuters and 2.1% other trip purposes (adjusted values between 7 and 8 AM according to Figure 38).</td>
<td>246</td>
<td>461</td>
<td>275</td>
<td>69</td>
<td>189</td>
</tr>
<tr>
<td>Deviation from passenger counts: weighted mean 35%</td>
<td>-37%</td>
<td>-16%</td>
<td>+26%</td>
<td>+9%</td>
<td>+107%</td>
</tr>
<tr>
<td>Headway-based assignment (see section 3.4) between line nodes and link corridors. 12.0% of workday commuters and 2.0% other trip purposes (adjusted values between 7 and 8 AM according to Figure 38).</td>
<td>192</td>
<td>560</td>
<td>237</td>
<td>44</td>
<td>231</td>
</tr>
<tr>
<td>Deviation from passenger counts: weighted mean 32%</td>
<td>-51%</td>
<td>+2%</td>
<td>+9%</td>
<td>-29%</td>
<td>+154%</td>
</tr>
</tbody>
</table>

Table 17  Average demand at maximum load sections (see Figure 39) during morning peak hour. The sum of the demand at the maximum load sections of all 6 lines was made identical in adjusting total demand during the morning peak hour.

It can be seen in Table 17 that disaggregating zones to hectares and optimizing headways, both have a positive effect on the accuracy of the headway-based assignment (compare
weighted means). However, the weighted means of the model deviations from passenger counts are still too high with around 30%. The aim should be to get mean deviations lower than 5%. Otherwise, only fundamental network improvements evaluated by simplified transport models could be implemented in reality.

The use of the PT lines 5 and 7 by passengers, both are overestimated after disaggregating the model of the Canton of Zurich to hectares. Both lines take passengers over from other lines (line 1 and line 2). Especially in connection with link corridors as new OD-zones, line 5 has a higher load, since the stop density on the link corridor close to the main station is assumed to be higher than stop density in the network 2008 (compare Figure 39).

The PT lines B1 and B2, both with a short headway, are according to the transport model less loaded than in reality. The transport model could be improved in considering the transport supply while building the basic hectare fine OD matrix (compare section 3.4.2). If distances to stops and headways offered there are short, the modal split of PT increases. While increasing the modal split for hectares 300m or closer to lines, with headways 7.5 minutes or shorter, the average deviations from passenger counts in the morning peak hour could be reduced from 28% to 19%. Since the focus of this thesis lies not on improving the transport model and it is not planned to directly implement the results in Winterthur, the OD matrix for the case studies was not adjusted.

### 5.2.4 Parameters of the Objective Function for the Case Study Winterthur

**Cost Balance Factor**

Costs for infrastructure investments and external costs are implicitly included into the generalized costs by a cost balance factor $f_{\text{cost}}$ in equation (5). The cost balance factor could increase the operating costs by the factor up to 1.719 (compare Table 18) depending on what kind of costs should be considered and which PT mode is used. However, in addition to a cost balance factor of $f_{\text{cost}} = 1.0$, in section 5.4 also a cost balance factor of $f_{\text{cost}} = 6.0$ is used to shift user costs and operating costs to the same magnitude.

<table>
<thead>
<tr>
<th>PT mode</th>
<th>Costs for infrastructure investments</th>
<th>External costs</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus, light rail</td>
<td>4.6%</td>
<td>4.6%</td>
<td>9.2%</td>
</tr>
<tr>
<td>Rail</td>
<td>63.1%</td>
<td>8.8%</td>
<td>71.9%</td>
</tr>
</tbody>
</table>

Table 18 Costs for infrastructure investments and external costs for Switzerland in percentage of operating costs (see [BUND06]).
User Costs

The average value of travel time savings represents the value of time $\nu_{\text{time}}$ for passengers. This value is 14.10 CHF/hour in Switzerland [Axhau06]. The objective function and the assignment method weights each part of the passenger trip differently based on passenger sensibility (see [Vrtic05]). Therefore walking time is weighted as 1.2 times in-vehicle travel time ($\alpha_{\text{ae}} = \alpha_{\text{TransferWalk}} = 1.2$ [Vrtic05]); access waiting time is weighted 2.5 times IVTT ($\alpha_{\text{aw}} = 2.5$ [Vrtic05]); each transfer is penalized by adding 11 minutes ($\alpha_{\text{TransferPen}} = 11$ min [Vrtic05]); and additional transfer waiting time is weighted by 0.23 ($\alpha_{\text{TransferW}} = 0.23$ [Vrtic05]). All together build the “perceived journey time” $t_{pj}$.

Operating Costs, Vehicle Capacities and Link Speeds

Total operating costs are calculated from service vehicle kilometers and service vehicle hours. Respective cost parameters are listed in Table 20. Here, it would be in principle possible to include infrastructure maintenance costs per revenue service kilometer. Average link speeds of the mixed traffic are given by the transport model of the Canton of Zurich [Vrtic05]. Maximum bus speeds are assumed to be 2/3 of the mixed traffic speed.

Other Parameters

Accepted line headways are given as a list of discrete values (see Table 19). In case of bus lines, headways of 30 and 60 minutes lead hardly to competitive total travel times compared to car (compare [Schäf05b]). Such long headways just provide a basic connection.

<table>
<thead>
<tr>
<th>Headways [min]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
</tr>
</tbody>
</table>

Table 19 Accepted line headways.

To guarantee access to PT stop facilities for passengers (residents/workers), the required maximum distance is set according to the “Angebotsverordnung” (supply order) of the Canton of Zurich [KZH99] to 400m linear distance. Additionally, it is necessary that a served settlement has at least 300 people (not explicitly considered in the following case studies).

It is guaranteed that at the maximum load section of a line, the average density of standing passengers during one hour does not exceed 2 passengers/ m². The dwell times $t_{dwell}$ are set for each link corridor according to equation (15) linear between $t_{dwell,min} = 15$s and $t_{dwell,max} = 60$s. For line nodes the dwell time $t_{dwell} = 20$s. Line length margins and cycle time margins could be given. Restrictions are not set, to make the different effects visible. If line align-
mements become confusing from passengers perspectives, lines can be divided into several parts. In that case, vehicles still serve the entire line.

<table>
<thead>
<tr>
<th>Composition</th>
<th>Number of seats</th>
<th>Total capacity (2 passengers/m²)</th>
<th>Costs/km [CHF]</th>
<th>Costs/h [CHF]</th>
<th>Maximum speed [km/h]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midi bus</td>
<td>16</td>
<td>28</td>
<td>2.3</td>
<td>80</td>
<td>60</td>
</tr>
<tr>
<td>Standard bus</td>
<td>35</td>
<td>53</td>
<td>2.7</td>
<td>92</td>
<td>60</td>
</tr>
<tr>
<td>Articulated bus</td>
<td>43</td>
<td>79</td>
<td>3.0</td>
<td>104</td>
<td>60</td>
</tr>
<tr>
<td>Double articulated bus</td>
<td>60</td>
<td>93</td>
<td>3.3</td>
<td>116</td>
<td>60</td>
</tr>
<tr>
<td>Light rail “2000” with low-floor partition</td>
<td>71</td>
<td>116</td>
<td>4.0</td>
<td>128</td>
<td>70</td>
</tr>
<tr>
<td>Light rail “cobra”</td>
<td>96</td>
<td>167</td>
<td>4.7</td>
<td>140</td>
<td>70</td>
</tr>
<tr>
<td>Regional train</td>
<td>162</td>
<td>246*</td>
<td>5.3</td>
<td>300</td>
<td>125</td>
</tr>
<tr>
<td>Heavy rail (one four carriage unit)</td>
<td>378</td>
<td>600*</td>
<td>6.0</td>
<td>450</td>
<td>125</td>
</tr>
<tr>
<td>Heavy rail (2 times a four carriage unit)</td>
<td>756</td>
<td>1’200*</td>
<td>6.7</td>
<td>525</td>
<td>160</td>
</tr>
<tr>
<td>Heavy rail (3 times a four carriage unit)</td>
<td>1’134</td>
<td>1’800*</td>
<td>7.3</td>
<td>600</td>
<td>160</td>
</tr>
</tbody>
</table>

Table 20 Vehicle capacities and estimated costs for service composition kilometers (assumed to be 1/3 of the total costs) and service composition hours (assumed to be 2/3 of the total costs) for standard compositions in the Canton of Zurich (compare [Weid08] (1) section 2.4.1.3.6 and [Frank08]).

* In local trains most passengers should be able to sit. For calculations only 2/3 of this capacity should be assumed.

5.3 Case Study Winterthur

5.3.1 Network 2008

Using the parameters given in the last section, the main indicators of the network 2008 during the morning peak hour (7.00 – 8.00 am) are derived and displayed by Case 0 in Table 21. A total of 615’280 minutes of perceived travel time are spent in the chosen section of the PT network. About 110’000 minutes thereof are spent in trains and regional busses. Another 280’000 perceived minutes are spent for waiting for the trains and regional busses. Perceived total passenger time costs are 144’591 CHF/h. The operating costs are 23’050 CHF/h, 15’000 CHF/h thereof are estimated (according to [Weid08] (1) section 2.4.1.3.6) for the operation of the rail lines. The total costs are 167’641 CHF/h according to equation (5). So, ca. 1/3 of the network costs are costs for the bus network.
5.3.2 Headway Optimization for Winterthurs PT Network

Frequencies of the lines B1-B14 (without B9) are varied according to section 4.3.4 and Table 19 during headway optimization. The optimization period was one average morning peak hour. The results are displayed in Table 21 (Case 1 and 2), Figure 41 and Table 23.

The larger number of vehicles (71 and 68) after headway optimization leads to shorter travel times and smaller vehicles to be used. In Table 23 it can be seen that only in case of line B1 the articulated bus is necessary. Often the midi bus provides sufficient capacity to cope with passenger flows. The high time cost factor of passengers allows high operating costs and a large fleet size (compare section 5.4.1).

In Table 21, two results of headway optimization can be compared (Case 1 and Case 2). The first case is based on given stops and the second is based on line nodes and link corridors. It can be seen that the difference in total perceived passenger travel time mainly derive from different access and egress times, but also from shorter cycle times. The difference due to shorter cycle times could be reduced in adjusting parameters of the stop density model on link corridors (see section 4.3.3) and thus in adjusting ride times.
Table 21  Network properties of the design cases for the PT network of Winterthur. Ca. 15'000 CHF are estimated for the operation of the rail lines, ca. 110'000 minutes are estimated for the in-vehicle travel times on the rail network and ca. 280’000 minutes are estimated for perceived waiting times for trains and perceived transfer time from or towards rail.
Figure 40  Case 0: The PT network of Winterthur, realized in 2008.

<table>
<thead>
<tr>
<th>Line name</th>
<th>Headway [min]</th>
<th>Max headway complying with vehicle capacities</th>
<th>Cycle time [min]</th>
<th>Terminal times [min]</th>
<th>Number of vehicle</th>
<th>Smallest possible vehicle capacity if headways given</th>
<th>Regional line, not changed during optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>6.0</td>
<td>10.0</td>
<td>60.0</td>
<td>4.33</td>
<td>10</td>
<td>53</td>
<td>-</td>
</tr>
<tr>
<td>B2</td>
<td>6.0</td>
<td>10.0</td>
<td>72.0</td>
<td>4.93</td>
<td>12</td>
<td>79</td>
<td>-</td>
</tr>
<tr>
<td>B3</td>
<td>7.5</td>
<td>15.0</td>
<td>67.5</td>
<td>3.58</td>
<td>9</td>
<td>53</td>
<td>-</td>
</tr>
<tr>
<td>B4</td>
<td>10.0</td>
<td>20.0</td>
<td>30.0</td>
<td>5.42</td>
<td>3</td>
<td>53</td>
<td>-</td>
</tr>
<tr>
<td>B5</td>
<td>15.0</td>
<td>30.0</td>
<td>90.0</td>
<td>9.85</td>
<td>6</td>
<td>53</td>
<td>-</td>
</tr>
<tr>
<td>B7</td>
<td>15.0</td>
<td>30.0</td>
<td>45.0</td>
<td>4.43</td>
<td>3</td>
<td>53</td>
<td>-</td>
</tr>
<tr>
<td>B9</td>
<td>30.0</td>
<td>60.0</td>
<td>60.0</td>
<td>13.75</td>
<td>2</td>
<td>28</td>
<td>true</td>
</tr>
<tr>
<td>B10</td>
<td>15.0</td>
<td>30.0</td>
<td>45.0</td>
<td>4.82</td>
<td>3</td>
<td>28</td>
<td>-</td>
</tr>
<tr>
<td>B14</td>
<td>15.0</td>
<td>30.0</td>
<td>45.0</td>
<td>6.63</td>
<td>3</td>
<td>53</td>
<td>-</td>
</tr>
<tr>
<td>B610</td>
<td>30.0</td>
<td>60.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>28</td>
<td>true</td>
</tr>
<tr>
<td>B611</td>
<td>120.0</td>
<td>120.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>28</td>
<td>true</td>
</tr>
<tr>
<td>B615</td>
<td>120.0</td>
<td>120.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>28</td>
<td>true</td>
</tr>
<tr>
<td>B660</td>
<td>30.0</td>
<td>60.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>53</td>
<td>true</td>
</tr>
<tr>
<td>B665</td>
<td>60.0</td>
<td>60.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>28</td>
<td>true</td>
</tr>
<tr>
<td>B670</td>
<td>60.0</td>
<td>30.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>53</td>
<td>true</td>
</tr>
<tr>
<td>B676</td>
<td>30.0</td>
<td>60.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>53</td>
<td>true</td>
</tr>
<tr>
<td>B680</td>
<td>60.0</td>
<td>60.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>79</td>
<td>true</td>
</tr>
<tr>
<td>Mean B1-B14 (without B9)</td>
<td>11.2</td>
<td>21.9</td>
<td>56.8</td>
<td>5.5</td>
<td>6.1</td>
<td>53.1</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 22  Case 0: Line properties of the PT network of Winterthur realized in 2008 (Figure 40).
Figure 41 Case 2: The PT network of Winterthur, realized in 2008 after headway optimization (see also Table 23). Assignment is based on line nodes and link corridors.

<table>
<thead>
<tr>
<th>Line name</th>
<th>Headway [min]</th>
<th>Max headway complying with vehicle capacities</th>
<th>Cycle time [min]</th>
<th>Terminal times [min]</th>
<th>Number of vehicle</th>
<th>Smallest possible vehicle capacity if headways given</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>7.5</td>
<td>12.0</td>
<td>52.5</td>
<td>6.38</td>
<td>7</td>
<td>79</td>
</tr>
<tr>
<td>B2</td>
<td>3.0</td>
<td>7.5</td>
<td>55.0</td>
<td>3.30</td>
<td>22</td>
<td>28</td>
</tr>
<tr>
<td>B3</td>
<td>12.0</td>
<td>20.0</td>
<td>60.0</td>
<td>6.05</td>
<td>5</td>
<td>53</td>
</tr>
<tr>
<td>B4</td>
<td>15.0</td>
<td>30.0</td>
<td>30.0</td>
<td>8.77</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>B5</td>
<td>7.5</td>
<td>20.0</td>
<td>75.0</td>
<td>6.33</td>
<td>10</td>
<td>53</td>
</tr>
<tr>
<td>B7</td>
<td>15.0</td>
<td>30.0</td>
<td>45.0</td>
<td>7.83</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>B10</td>
<td>20.0</td>
<td>60.0</td>
<td>40.0</td>
<td>6.45</td>
<td>2</td>
<td>53</td>
</tr>
<tr>
<td>B14</td>
<td>30.0</td>
<td>60.0</td>
<td>30.0</td>
<td>3.55</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>Mean</td>
<td>13.8</td>
<td>29.9</td>
<td>48.4</td>
<td>6.1</td>
<td>6.5</td>
<td>43.8</td>
</tr>
</tbody>
</table>

Table 23 Case 2: Line properties of the PT network of Winterthur realized in 2008 after headway optimization, for the optimized lines only (see as well Figure 41). Assignment is based on line nodes and link corridors.
5.3.3 New Design of Winterthurs PT Network

For the PT network of Winterthur 4’414 NSR (nearly shortest routes) and out of them around 350 shortest lines are built. Together with the given 17 lines a set of 369 lines result. 100 lines are added to an initial PT network. According to the assignment process (see section 3.4.3) and equation (16) the 50 as “worst” evaluated lines are removed and replaced by other lines out of the set, as long as all line nodes and link corridors are served. This is repeated till all of the 369 lines have been evaluated, included or deleted. In some steps (50, 45, 40, 38, 36, 35) the number of lines is further reduced, till only 35 lines remain. With those lines the reduction process (see section 4.3.5) is started.

During the network reduction process (see section 4.3.5) Winterthurs new PT network is evaluated around 5'000 times. The reduced evaluations during shortening or merging with ACO (ant colony optimization) are performed about one million times. This leads to a total computing time of about two days on a 2GHz computer. Case 3 in Table 21 displays the results.

In a longer calculation, the reduction process was started with 45 lines. That resulted in around 10’000 evaluations, 2 million reduced evaluations and approximately 4 days computing time. Case 4 in Table 21 displays the results. The corresponding network is shown with line properties in Figure 42 and Table 24.
In Table 24 and Table 25 it can be seen that cycle times are often 60 minutes, the basic period of the network. Lines with cycle times of the basic period have terminal times independent from headways. Such lines with short terminal times have a cost advantage during the reduction process and in adapting headways to changes of the passenger load during the daytime. Beside cycle times of the basic period also cycle times of a multiple of the basic period or cycle times evenly divisible into the basic period by small divisors are advantageous.

<table>
<thead>
<tr>
<th>Line name</th>
<th>Headway [min]</th>
<th>Max headway [min], complying with vehicle capacities</th>
<th>Cycle time [min]</th>
<th>Terminal times [min]</th>
<th>Number of vehicle</th>
<th>Smallest possible vehicle capacity if headways given</th>
</tr>
</thead>
<tbody>
<tr>
<td>B2</td>
<td>30.0</td>
<td>60.0</td>
<td>60.0</td>
<td>11.0</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>B3</td>
<td>5.0</td>
<td>15.0</td>
<td>45.0</td>
<td>3.1</td>
<td>9</td>
<td>28</td>
</tr>
<tr>
<td>B4</td>
<td>10.0</td>
<td>30.0</td>
<td>20.0</td>
<td>4.6</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>B7</td>
<td>20.0</td>
<td>60.0</td>
<td>40.0</td>
<td>8.4</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>B10</td>
<td>30.0</td>
<td>60.0</td>
<td>60.0</td>
<td>16.8</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>B14</td>
<td>20.0</td>
<td>60.0</td>
<td>40.0</td>
<td>7.6</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>B56</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>5.3</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B75</td>
<td>15.0</td>
<td>60.0</td>
<td>30.0</td>
<td>6.9</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>B76</td>
<td>30.0</td>
<td>60.0</td>
<td>60.0</td>
<td>7.4</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>B89</td>
<td>7.5</td>
<td>20.0</td>
<td>52.5</td>
<td>3.3</td>
<td>7</td>
<td>28</td>
</tr>
<tr>
<td>B104</td>
<td>5.0</td>
<td>12.0</td>
<td>55.0</td>
<td>3.9</td>
<td>11</td>
<td>53</td>
</tr>
<tr>
<td>B116</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>6.7</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B117</td>
<td>20.0</td>
<td>60.0</td>
<td>60.0</td>
<td>7.0</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>B121</td>
<td>20.0</td>
<td>60.0</td>
<td>60.0</td>
<td>6.0</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>B158</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>7.0</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B287</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>11.65</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B305</td>
<td>30.0</td>
<td>60.0</td>
<td>60.0</td>
<td>6.4</td>
<td>2</td>
<td>53</td>
</tr>
<tr>
<td>B314</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>5.4</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B320</td>
<td>5.0</td>
<td>15.0</td>
<td>50.0</td>
<td>4.0</td>
<td>10</td>
<td>28</td>
</tr>
<tr>
<td>B325</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>9.3</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B332</td>
<td>30.0</td>
<td>60.0</td>
<td>60.0</td>
<td>4.3</td>
<td>2</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 24 Case 4: Line properties of the PT network, designed with existing lines after long calculation (see Figure 42). Assignment is based on line nodes and link corridors.
Figure 43  Case 6: Network design without existing lines (see also Table 25).

<table>
<thead>
<tr>
<th>Line name</th>
<th>Headway [min]</th>
<th>Max headway [min] complying with vehicle capacities</th>
<th>Cycle time [min]</th>
<th>Terminal times [min]</th>
<th>Number of vehicle</th>
<th>Smallest possible vehicle capacity if headways given</th>
</tr>
</thead>
<tbody>
<tr>
<td>B5</td>
<td>60.0</td>
<td>UpB.</td>
<td>120.0</td>
<td>33.0</td>
<td>2</td>
<td>53</td>
</tr>
<tr>
<td>B37</td>
<td>10.0</td>
<td>20.0</td>
<td>60.0</td>
<td>5.3</td>
<td>6</td>
<td>53</td>
</tr>
<tr>
<td>B69</td>
<td>15.0</td>
<td>30.0</td>
<td>60.0</td>
<td>4.0</td>
<td>4</td>
<td>53</td>
</tr>
<tr>
<td>B90</td>
<td>20.0</td>
<td>30.0</td>
<td>60.0</td>
<td>7.0</td>
<td>3</td>
<td>53</td>
</tr>
<tr>
<td>B102</td>
<td>15.0</td>
<td>60.0</td>
<td>45.0</td>
<td>4.4</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>B113</td>
<td>20.0</td>
<td>60.0</td>
<td>60.0</td>
<td>9.2</td>
<td>3</td>
<td>53</td>
</tr>
<tr>
<td>B151</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>5.0</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>B158</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>15.9</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B163</td>
<td>10.0</td>
<td>15.0</td>
<td>60.0</td>
<td>4.8</td>
<td>6</td>
<td>53</td>
</tr>
<tr>
<td>B241</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>13.1</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B267</td>
<td>6.0</td>
<td>20.0</td>
<td>30.0</td>
<td>3.1</td>
<td>5</td>
<td>28</td>
</tr>
<tr>
<td>B306</td>
<td>30.0</td>
<td>60.0</td>
<td>60.0</td>
<td>17.4</td>
<td>2</td>
<td>53</td>
</tr>
<tr>
<td>B326</td>
<td>5.0</td>
<td>7.5</td>
<td>50.0</td>
<td>4.2</td>
<td>10</td>
<td>53</td>
</tr>
</tbody>
</table>

Mean 28.5 41.7 60.4 9.7 3.6 43.4

Table 25  Case 6: Line properties of the PT network design without existing lines (see Figure 43).
Winterthurs PT network has been updated in 2007 and 2008 using a planner’s experience. Radial lines were combined into two additional through lines (through lines B3 and B5).

The presented approach allows to reduce total costs by 1.1 % (see Case 4 in Table 21). Considering that one third of total costs are related to the bus network, total costs of Winterthurs PT network can be improved by around 3%. To ensure that such small changes can bring sound improvements, the accuracy of both, the OD matrix and the assignment model, has to be sufficient. The relative error of passenger flows at maximum load sections of PT lines should not be higher than around 5% for each speed level. With data e.g. from electronic ticketing systems, it should be possible to validate and improve assignment models and to achieve the quality required. With the given transport model, as can be seen in Table 17, this accuracy is not given though.

Network design without existing lines (Case 6 in Table 21) causes results with 0.1% higher total costs than with the existing lines (Case 3 in Table 21). Consequently the nearly shortest lines as start network, is not the best start network for the reduction process. Other lines, such as line B3, B4 or B10 also could be good lines for the start network. Besides creating nearly shortest lines as start lines, it could be a successful strategy to additionally connect important transfer nodes with each other. This could be done using forms similar to the magnetic streamlines of a field that connects several magnetic poles (see Figure 44), covering the entire area as well.

Figure 44 Magnetical streamlines of the field displayed with small iron particles (source: Practical Physics, publ. 1914 by Macmillan and Company).
5.3.4 Sequential PT Network Design Using two Overlapping Planning Areas

Total computing times for optimizing the PT network of Winterthur are still too long, hence, no larger area networks or faster speed levels were optimized. In order to test the multiple area approach (Figure 2, loop 2), the area of Winterthur was divided into two overlapping planning areas (see Figure 45 and Figure 46). First, a PT network on the left side (Figure 45, Planning area 1) was designed, then in a second step (Figure 46, Planning area 2), a network for almost the entire area. The given lines resulting from the network optimization of Planning area 1 were fixed on the shaded area on the left side during the optimizing of Planning area 2. If line headways were basic headways (60 minutes), lines with fixed parts that touched the shaded area could be deleted. If lines with fixed parts were merged with other lines, the fixed part on the shaded area must not be changed.

Figure 45 Case 5: Sequential network design using two overlapping planning areas, with lines from network 2008. Firstly a PT network was designed in Planning area 1 on the left side of the entire area.

The given lines resulting from the network optimization of the first planning area (see Figure 45) are partly fixed on the shaded part at the left side during optimizing the second planning area (Figure 46). Only if line headways were basic headways (60 minutes), line parts of fixed lines, touching the shaded part, can be deleted. Otherwise one line touching the shaded part can still be merged with other lines. Just the shaded part must not be changed.
Figure 46  Case 5: Sequential network design using two overlapping planning areas, with lines from network 2008 (see also Table 26). In a second step a network for nearly the entire area was designed in Planning area 2. Lines on the shaded part at the left side are partly fixed.

The computing time of 1.5 days on a 2 GHz machine was slightly shorter compared to the single planning area calculation of two days. The network was evaluated around 10'000 times, with half a million reduced evaluations. The faster computing times for network evaluations are due to the smaller planning areas. Computing times could be significantly reduced with the help of the multiple area approach in the case of larger areas (using more planning areas).

Total network costs (see Case 5 and Case 3 in Table 21) are 0.1% higher than for network design of one single area. This makes 0.3%, considering that the cost share of the optimized bus network is 1/3 of total costs only. On one hand the planning area size is too small. On the other hand it should also be possible to merge two partly fixed lines. In creating the presented results it was only possible to merge a partly fixed line with a not fixed line.
<table>
<thead>
<tr>
<th>Line name</th>
<th>Headway [min]</th>
<th>Max headway complying with vehicle capacities</th>
<th>Cycle time [min]</th>
<th>Terminal times [min]</th>
<th>Number of vehicle</th>
<th>Smallest possible vehicle capacity if headways given</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>2.5</td>
<td>7.5</td>
<td>42.5</td>
<td>3.8</td>
<td>17</td>
<td>28</td>
</tr>
<tr>
<td>B2</td>
<td>60.0</td>
<td>UpB</td>
<td>60.0</td>
<td>11.0</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B3</td>
<td>6.0</td>
<td>20.0</td>
<td>48.0</td>
<td>4.6</td>
<td>8</td>
<td>28</td>
</tr>
<tr>
<td>B4</td>
<td>12.0</td>
<td>30.0</td>
<td>24.0</td>
<td>6.5</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>B5</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>4.7</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B7</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>18.4</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B10</td>
<td>12.0</td>
<td>60.0</td>
<td>36.0</td>
<td>4.8</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>B11</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>12.3</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B14</td>
<td>30.0</td>
<td>60.0</td>
<td>60.0</td>
<td>17.6</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>B46</td>
<td>15.0</td>
<td>30.0</td>
<td>45.0</td>
<td>3.9</td>
<td>3</td>
<td>53</td>
</tr>
<tr>
<td>B55</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>9.9</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B67</td>
<td>10.0</td>
<td>30.0</td>
<td>50.0</td>
<td>7.0</td>
<td>5</td>
<td>28</td>
</tr>
<tr>
<td>B81</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>16.8</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B85</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>17.3</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B88</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>14.8</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B89</td>
<td>20.0</td>
<td>60.0</td>
<td>40.0</td>
<td>4.8</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>B92</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>6.6</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B126</td>
<td>5.0</td>
<td>10.0</td>
<td>45.0</td>
<td>3.1</td>
<td>9</td>
<td>53</td>
</tr>
<tr>
<td>B128</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>10.7</td>
<td>1</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 26 Case 5: Line properties of the sequential PT network design of two overlapping areas with existing lines (see Figure 46).

5.3.5 Final Stop Placement

In the final stop placement, the line node placement method from section 4.2 has been adapted (see section 4.3.8). Further adjustments are recommended, to be able to successfully use it. However the results obtained with the adapted version show the following interesting effects.

In Table 27 various cases are compared with each other (see also Figure 34 and Figure 35). In a case, were passenger time is evaluated as high (cost balance factor \( f_{\text{cost}} = 1 \)), the existing stop positions lead to 0.6% decrease of the total cost function (equation (5)). Considering that the cost share of the optimized bus network is 1/3 of total costs only this makes 1.8%. If the cost balance factor \( f_{\text{cost}} = 20 \) (travel time costs are evaluated as low), the new stop positions in Figure 34 lead to 2.6% lower values of the total cost function than the existing stop positions because 29 stops less allow shorter travel times. This makes 7.8%, considering that the cost share of the optimized bus network is 1/3 of total costs only. After the final stop placement, access and egress times are slightly shorter than for the existing stop placement, but the other travel time components are longer. More transfers are required.
These results show that the placement of stops still holds potential for improvements. Instead of shortest walking times only, access/egress to stops can be performed additionally according to shortest travel times between link corridors and line nodes. The linear relationship (see equation (14)) between the density of the potential demand and the on-link stop distances could be improved to a more realistic nonlinear relationship. Such nonlinear relationships can be found in the existing stop positions, where stop distances are less regular.

### Table 27

<table>
<thead>
<tr>
<th>Stop case</th>
<th>Cost balance factor</th>
<th>Total perceived passenger travel time [min]</th>
<th>Total fleet size</th>
<th>Sum of Terminal Times (minimum, if all terminal times are 3 minutes) [min]</th>
<th>Access/Egress time [min]</th>
<th>Number of stops</th>
<th>0-transfer trips [%]</th>
<th>1-transfer trips [%]</th>
<th>Total passenger travel time costs [CHF]</th>
<th>Operating costs [CHF]</th>
<th>Total costs [CHF]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Existing</td>
<td>1</td>
<td>609’660</td>
<td>71</td>
<td>870 (420)</td>
<td>88’299</td>
<td>165</td>
<td>41.4</td>
<td>56.9</td>
<td>143’270</td>
<td>23’769</td>
<td>167’039</td>
</tr>
<tr>
<td>New</td>
<td>1</td>
<td>620’381</td>
<td>52</td>
<td>753 (318)</td>
<td>87’624</td>
<td>136</td>
<td>40.2</td>
<td>57.1</td>
<td>145’789</td>
<td>22’322</td>
<td>168’111</td>
</tr>
<tr>
<td>Existing</td>
<td>20</td>
<td>618’611</td>
<td>46</td>
<td>807 (252)</td>
<td>89’598</td>
<td>165</td>
<td>40.6</td>
<td>57.3</td>
<td>7’269</td>
<td>21’522</td>
<td>28’791</td>
</tr>
<tr>
<td>New</td>
<td>20</td>
<td>639’546</td>
<td>40</td>
<td>685 (240)</td>
<td>88’342</td>
<td>136</td>
<td>39.0</td>
<td>57.8</td>
<td>7’515</td>
<td>20’809</td>
<td>28’324</td>
</tr>
</tbody>
</table>

Table 27  
Network properties. Stop placement comparison of four cases. In two cases the cost balance factor has been changed.

### 5.4 Design Parameter Variation

#### 5.4.1 Variation of the Cost Balance Factor and of Maximum Bus Capacities

In Switzerland, PT is subsidized by politics. Ticket prices do not cover the full costs for traveling. This is the reason why the value of travel time savings, derived from stated preference surveys, only gives a hint how much a PT system should cost. In Table 28 it can be seen, that a value of travel time savings \( \nu_{\text{Time}} = 14.10 \text{ CHF/hour} \) (a cost balance factor \( f_{\text{cost}} = 1 \)) leads to 6 times higher user costs in equation (5) compared to operating costs.

While adjusting the cost balance factor \( f_{\text{cost}} \), it is possible to shift user costs and operating costs. With \( f_{\text{cost}} = 6 \) they have the same magnitude. In this case travel time costs are evaluated as significantly lower. In Table 28 the effects of adjusting the cost balance factor is shown. Additional, in one case the maximum vehicle capacity was lowered to one standing passengers/m² instead of two (compare Table 20), what leads to an increase of total costs by 1.4%.
Table 28  Network properties. Variation of the cost balance factor for the PT network of Winterthur and of vehicle capacities.

<table>
<thead>
<tr>
<th>Line name</th>
<th>Headway [min]</th>
<th>Max headway [min] complying with vehicle capacities</th>
<th>Cycle time [min]</th>
<th>Terminal times [min]</th>
<th>Number of vehicle</th>
<th>Smallest possible vehicle capacity if headways given</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>20.0</td>
<td>30.0</td>
<td>60.0</td>
<td>12.6</td>
<td>3</td>
<td>79</td>
</tr>
<tr>
<td>B3</td>
<td>20.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>10.7</td>
<td>3</td>
<td>79</td>
</tr>
<tr>
<td>B4</td>
<td>20.0</td>
<td>30.0</td>
<td>20.0</td>
<td>4.6</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>B7</td>
<td>30.0</td>
<td>60.0</td>
<td>30.0</td>
<td>3.4</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>B10</td>
<td>20.0</td>
<td>UpB.</td>
<td>40.0</td>
<td>6.8</td>
<td>2</td>
<td>53</td>
</tr>
<tr>
<td>B13</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>3.0</td>
<td>1</td>
<td>79</td>
</tr>
<tr>
<td>B14</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>17.6</td>
<td>1</td>
<td>79</td>
</tr>
<tr>
<td>B44</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>5.2</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>B56</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>5.3</td>
<td>1</td>
<td>79</td>
</tr>
<tr>
<td>B75</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>10.0</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>B86</td>
<td>30.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>6.5</td>
<td>2</td>
<td>79</td>
</tr>
<tr>
<td>B89</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>7.0</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>B153</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>5.5</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>B162</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>4.9</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>B164</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>5.4</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>B321</td>
<td>7.5</td>
<td>UpB.</td>
<td>52.5</td>
<td>5.5</td>
<td>7</td>
<td>79</td>
</tr>
<tr>
<td>B324</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>9.1</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>Mean</td>
<td>44.0</td>
<td>46.9</td>
<td>54.3</td>
<td>7.2</td>
<td>1.7</td>
<td>63.7</td>
</tr>
</tbody>
</table>

Table 29  Case 9: Line properties of a new PT network designed with existing lines (maximum vehicle load is two standing passengers/ m²); $f_{cov} = 6.$
Figure 47  Case 10: Network design with existing lines (see also Table 30). Cost balance factor \( f_{\text{cost}} = 6 \). Maximum vehicle load is one standing passenger/ \( m^2 \).

<table>
<thead>
<tr>
<th>Line name</th>
<th>Headway [min]</th>
<th>Max headway [min] complying with capacities</th>
<th>Cycle time [min]</th>
<th>Terminal times [min]</th>
<th>Number of vehicle</th>
<th>Smallest possible vehicle capacity if headways given</th>
</tr>
</thead>
<tbody>
<tr>
<td>B2</td>
<td>7.5</td>
<td>UpB.</td>
<td>60.0</td>
<td>6.5</td>
<td>8</td>
<td>93</td>
</tr>
<tr>
<td>B3</td>
<td>12.0</td>
<td>15.0</td>
<td>60.0</td>
<td>6.0</td>
<td>5</td>
<td>79</td>
</tr>
<tr>
<td>B4</td>
<td>20.0</td>
<td>30.0</td>
<td>20.0</td>
<td>4.6</td>
<td>1</td>
<td>79</td>
</tr>
<tr>
<td>B5</td>
<td>60.0</td>
<td>UpB.</td>
<td>120.0</td>
<td>30.0</td>
<td>2</td>
<td>53</td>
</tr>
<tr>
<td>B10</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>16.8</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>B14</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>17.6</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>B117</td>
<td>30.0</td>
<td>UpB.</td>
<td>90.0</td>
<td>17.9</td>
<td>3</td>
<td>79</td>
</tr>
<tr>
<td>B153</td>
<td>20.0</td>
<td>30.0</td>
<td>40.0</td>
<td>5.2</td>
<td>2</td>
<td>53</td>
</tr>
<tr>
<td>B164</td>
<td>15.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>5.1</td>
<td>4</td>
<td>79</td>
</tr>
<tr>
<td>B325</td>
<td>20.0</td>
<td>UpB.</td>
<td>40.0</td>
<td>5.1</td>
<td>2</td>
<td>79</td>
</tr>
<tr>
<td>Mean</td>
<td>30.5</td>
<td>32.8</td>
<td>61.0</td>
<td>11.5</td>
<td>2.9</td>
<td>70.0</td>
</tr>
</tbody>
</table>

Table 30  Case 10: Line properties of the new PT network, designed with existing lines (maximum vehicle load is one standing passenger/ \( m^2 \), see Figure 47); \( f_{\text{cost}} = 6 \).
Table 28 shows that with a cost balance factor of \( f_{\text{cost}} = 6 \) the fleet size becomes only 2/3 of the in 2008 necessary fleet size (compare Case 0 in Table 21). Compared to Case 4 in Table 21, it’s half the fleet size. For most lines, headways are just long enough to cope with passenger flows. The smaller fleet size requires larger vehicle capacities. In the conservative case, assuming one standing passengers/m² in the morning peak hour, only for the line B2 double articulated buses are necessary (see Table 30).

### 5.4.2 Variation of Maximum Bus Speed and Artificial Infrastructure Network

In the beginning of chapter 4 it is mentioned that it can be helpful to have a kind of “ideal network” as comparison to evaluate PT networks on given infrastructures. Here an example for such infrastructure network is presented (see Figure 48).

The existing infrastructure with the respective line nodes was considered as fixed for rail and for PT lines crossing the borders of the planning area. Since the planning area is quite small, the given regional lines serve the center of the area and many line nodes are already fixed. With those line nodes given, new line nodes were placed around. Minimum line node distances were set according to equation (14) to \( d_{\text{min},ib} = 350 \text{m} \) \((DP_{ib} \geq 300'000)\) and to \( d_{\text{min},ab} = 800 \text{m} \) \((DP_{ib} \leq 3'000)\).

<table>
<thead>
<tr>
<th>Design case</th>
<th>Cost balance factor</th>
<th>Maximum bus speed [km/h]</th>
<th>Total perceived passenger travel time [min]</th>
<th>Total fleet size</th>
<th>Sum of Terminal Times (minimum, if all terminal times are 3 minutes) [min]</th>
<th>0-transfer trips [%]</th>
<th>1-transfer trips [%]</th>
<th>2-transfer trips [%]</th>
<th>Total passenger travel time costs [CHF]</th>
<th>Operating costs [CHF]</th>
<th>Total costs [CHF]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 3: New Network With existing lines Stops not given.</td>
<td>1</td>
<td>F</td>
<td>586’499</td>
<td>57</td>
<td>945 (420)</td>
<td>37.2</td>
<td>61.4</td>
<td>1.4</td>
<td>137’827</td>
<td>22’841</td>
<td>160’668</td>
</tr>
<tr>
<td>Case 7: New Network With existing lines Stops not given.</td>
<td>1</td>
<td>50</td>
<td>573’909</td>
<td>54</td>
<td>958 (438)</td>
<td>38.0</td>
<td>60.9</td>
<td>1.2</td>
<td>134’868</td>
<td>22’951</td>
<td>157’820</td>
</tr>
<tr>
<td>Case 8: Artificial Network Stops not given.</td>
<td>1</td>
<td>50</td>
<td>560’787</td>
<td>77</td>
<td>1’116 (600)</td>
<td>38.4</td>
<td>61.1</td>
<td>0.5</td>
<td>131’785</td>
<td>25’177</td>
<td>156’962</td>
</tr>
</tbody>
</table>

Table 31 Network properties. Variation of maximum bus speed and use of an artificial infrastructure network; F: 2/3 of average car speed.
Investigation of space-time structures in public transport networks and their optimization

<table>
<thead>
<tr>
<th>Line name</th>
<th>Headway [min]</th>
<th>Max headway [min] complying with vehicle capacities</th>
<th>Cycle time [min]</th>
<th>Terminal times [min]</th>
<th>Number of vehicle</th>
<th>Smallest possible vehicle capacity if headways given</th>
</tr>
</thead>
<tbody>
<tr>
<td>B2</td>
<td>20.0</td>
<td>30.0</td>
<td>40.0</td>
<td>4.3</td>
<td>2</td>
<td>53</td>
</tr>
<tr>
<td>B3</td>
<td>5.0</td>
<td>12.0</td>
<td>40.0</td>
<td>3.8</td>
<td>8</td>
<td>53</td>
</tr>
<tr>
<td>B4</td>
<td>30.0</td>
<td>UpB.</td>
<td>40.0</td>
<td>9.6</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>B7</td>
<td>20.0</td>
<td>60.0</td>
<td>40.0</td>
<td>8.7</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>B10</td>
<td>15.0</td>
<td>60.0</td>
<td>30.0</td>
<td>3.9</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>B14</td>
<td>30.0</td>
<td>60.0</td>
<td>60.0</td>
<td>17.6</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>B72</td>
<td>15.0</td>
<td>30.0</td>
<td>45.0</td>
<td>3.8</td>
<td>3</td>
<td>53</td>
</tr>
<tr>
<td>B86</td>
<td>12.0</td>
<td>20.0</td>
<td>48.0</td>
<td>4.8</td>
<td>4</td>
<td>53</td>
</tr>
<tr>
<td>B106</td>
<td>15.0</td>
<td>20.0</td>
<td>45.0</td>
<td>4.2</td>
<td>3</td>
<td>79</td>
</tr>
<tr>
<td>B117</td>
<td>15.0</td>
<td>30.0</td>
<td>45.0</td>
<td>3.2</td>
<td>3</td>
<td>53</td>
</tr>
<tr>
<td>B122</td>
<td>30.0</td>
<td>60.0</td>
<td>30.0</td>
<td>7.8</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B150</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>11.4</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>B304</td>
<td>5.0</td>
<td>12.0</td>
<td>45.0</td>
<td>4.6</td>
<td>9</td>
<td>53</td>
</tr>
<tr>
<td>B325</td>
<td>60.0</td>
<td>UpB.</td>
<td>60.0</td>
<td>12.8</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>Mean</td>
<td>23.7</td>
<td>37.2</td>
<td>44.1</td>
<td>7.2</td>
<td>3.0</td>
<td>45.9</td>
</tr>
</tbody>
</table>

Table 32  Case 7: Line properties of the PT network, designed with existing lines (Maximum bus speed is 50km/h).

Figure 48  Case 8: Network design on an artificial infrastructure (see also Table 33); Maximum bus speed is 50km/h.
For the PT network used as comparison (see Table 31) the maximum bus speeds have been considered as fixed to 50km/h on all roads, to make the artificial infrastructure network comparable to the existing one. The faster maximum bus speed (50km/h instead of 2/3 of average car speed) leads to a smaller fleet size, to shorter travel times and consequently to lower total costs (-5.5% considering that one third of total costs are related to the bus network; see Table 31).

Nevertheless the PT network, built on the artificial infrastructure network, is slightly better (around 2.5%; bus network costs are 1/3 of total costs) than a network built on the existing infrastructure (see Table 31).

### 5.5 Summary of the Case Study Winterthur

Table 21, Table 27 and Table 28 are summarized in Figure 49 and Figure 50. Hereby travel times and costs concerning the bus network of Winterthur are displayed. Cost estimations for the train network are subtracted already. In Figure 49 it can be seen that the total travel time
on the bus network and the number of vehicles are inversely proportional. The number of vehicles varies by around 25%, including cases 7 – 10 by 50%.

---

**Figure 49** Total travel time on the bus network versus number of vehicles ($v_{max} = 2/3 \, v_{car}$; $f_{cost} = 1$). The design cases are listed in Table 34.

---

**Figure 50** Cost comparison ($v_{max} = 2/3 \, v_{car}$; $f_{cost} = 1$). The design cases are listed in Table 34.
Table 34 Design cases. Assignment is based on line nodes and link corridors \( (v_{\text{max}} = 2/3 v_{\text{car}}; f_{\text{cost}} = 1; \text{max. capacity} = 2 \text{ passenger/m}^2) \).

In Figure 50 it can be seen that the total cost are quite similar but the time costs and the operating costs are inversely proportional and vary considerably. The operating costs vary by around 20%, including cases 7 – 10 by 40%. This is 10% less than the number of vehicles in Figure 49. The fewer vehicles are used the larger is the capacity of each vehicle and the more costly the corresponding vehicle kilometer and vehicle hours become (compare Table 20).

While keeping total costs at one level (cases 2 – 6), the variability of time costs (ca. 7%) and operating costs (ca. 20%) could be used by politics to decide how the different costs of PT are covered and by whom. Hereby, the effects on modal split should be considered. The cases 7 and 8 in Figure 50 show that it is possible to reduce the systems operating costs together with passenger’s time costs by increasing vehicle speeds.
5.6 Comparison with Results of Other Approaches

Simplifications in transport models make it necessary to use benchmark problems in order to compare the results of different network design methods. The basis for all benchmark problems should be an accurate transport model.

Since a benchmark problem like this one did not exist, a modified version of Mandl’s benchmark problem [Mandl79] was used to compare the results of the guided stochastic search heuristic (GSSH) proposed in this research with results of other approaches. Only for two reviewed network design methods of chapter 2 a solution for Mandl’s benchmark problem exist. That’s why also other results are used for the comparison. The modifications to Mandl’s benchmark problem consist of an expansion developed by Baaj and Mahmassani [Baaj91] and other small changes. The largest of the changes introduced in this study was the use of the HBSMR assignment (see section 3.4.3 and [PTV07]). Figure 51 illustrates the benchmark problem.

The benchmark problem contains the following assumptions and conditions, which were applied to all compared line alignments. The OD matrix (originally for one day; see Mandl 1979) was assigned to the network as one hour (compare [Baaj91] and later literature). Therefore, many vehicles with a capacity of 50 passengers are required and resulting headways are quite short. During network evaluations, the HBSMR assignment [PTV07] was used and no schedule optimization was performed. The headway optimization of this study was applied to all given line alignments (results are shown in Table 37 to Table 43). All headways must be less than or equal to 30 minutes (introduced in this study). To make the results of the GSSH comparable to results of other approaches, headways were made continuous. The 60-minute basic period does not need to be evenly divisible by the headways. Instead of a maximum cycle time of 80 minutes [Baaj91], cycle times were not restricted here. Influence on results is small as well. For access and transfers, average waiting times were considered (compare equation (2)); with $a_{acc} = 0$, $a_{aw} = 1$, $a_{bw} = 1$, $a_{TranfWalk} = 0$, $a_{TranfW} = 1$, $b_{TranfPen} = 5\text{min}$). In Figure 51, run times on links are given and distances are not. Vehicle costs were set to $C_v = 130 \text{CHF/h}$. All stops have to be served. Terminal times (to turn vehicles plus buffer times) are not considered. A cost balance factor of $f_{cost} = 3.525$ was used, which means the ratio of passenger costs and service provider costs were adjusted. Consequently, the time cost values of passengers became $\nu_{time} = 4.0 \text{CHF/h}$ instead of $14.1 \text{CHF/h}$ (compare section 5.2.4).

With this value for time (or slightly lower values), the proposed method returned results (in terms of number of buses needed to operate the system) similar to those found optimal for network design in the literature. An explanation for that could be that service provider costs are restricted, e.g., to limit subsidies.
Investigation of space-time structures in public transport networks and their optimization

Figure 51 Mandl’s Swiss benchmark problem [Mandl79]. a) Network layout with node numbers and run times on links in minutes. b) OD matrix (Table 35) assigned to network links on shortest routes. If routes have the same travel time, demand is assigned to one of the routes.

During network design with the proposed GSSH approach, the network was evaluated between 5’000 and 8’000 times. The reduced evaluations during shortening or merging with ACO were performed more than 100’000 times. This made a total computing time of about one hour on a 2 GHz computer. Due to the more realistic transport modeling and the variability of line alignments, the proposed PT network design method is expected to develop PT networks close to the global optimum. Due to the stochastic nature of the GSSH approach and the sequence problem during reduction, the results differ slightly (local optima) in most runs of the algorithm.

Applying the GSSH approach to Mandl’s small benchmark problem (comparison 1 and 2 in Table 36) shows good results concerning the generalized costs as well as total travel times in combination with fleet sizes. With fewer vehicles, similar or lower total travel times were
achieved compared to results of other approaches. However, as no schedule optimization was performed, these results are limited.

Only one result of Zhao (Case S&M2 in [Zhao06]) and one of Zhao and Zeng (Case S&M2 in [Zhao08]) yield better results. The first study used a headway-based all-or-nothing assignment, the second a schedule-based assignment. Unfortunately, the resulting line alignments (network layouts) and schedule information were not published. A fleet size of 68 vehicles could not be achieved by the GSSH in combination with the HBSMR assignment. 70 vehicles was the smallest possible fleet size. The results of Zhao and Gan [Zhao03] Case S&M2 and Shih and Mahmassani [Shih94] Case S&M2 originally needed only 68 buses as well. With the HBSMR assignment, they needed 73 and 77 buses respectively.

In Zhao and Gan’s results [Zhao03] for Case S&M2 (see Figure 58), the lines B3 and B4 are not really necessary. Without these two lines, total costs would be smaller and even closer to total costs of the network designed with the GSSH (comparison 2; Figure 57). In Case S&M2 [Zhao03], the use of eight lines was a constraint for network design.

<table>
<thead>
<tr>
<th>Node</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>400</td>
<td>200</td>
<td>60</td>
<td>80</td>
<td>150</td>
<td>75</td>
<td>75</td>
<td>30</td>
<td>160</td>
<td>30</td>
<td>25</td>
<td>35</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>120</td>
<td>20</td>
<td>180</td>
<td>90</td>
<td>90</td>
<td>15</td>
<td>130</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>60</td>
<td>180</td>
<td>90</td>
<td>90</td>
<td>15</td>
<td>45</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>15</td>
<td>240</td>
<td>40</td>
<td>25</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>10</td>
<td>120</td>
<td>20</td>
<td>15</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>100</td>
<td>30</td>
<td>880</td>
<td>60</td>
<td>15</td>
<td>15</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>440</td>
<td>35</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>50</td>
<td>440</td>
<td>35</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>140</td>
<td>20</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>600</td>
<td>250</td>
<td>500</td>
<td>200</td>
<td>0</td>
<td>75</td>
<td>95</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>45</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

Table 35 Mandl’s Swiss benchmark problem [Mandl79] (see Figure 51): Symmetric OD matrix for an average day.
### Table 36 Comparison of line network layouts (compare Table 37 to Table 43) created by different search- or optimization methods for Mandl’s benchmark problem (see Figure 51) as expanded by Baaj and Mahmassani. Comparison 1 compares networks with 75 or more vehicles, comparison 2 with 68 or more.

* In Zhao and Zeng’s results (Case S&M2) the line network layout hasn’t been given. So, the values from Zhao and Zeng [Zhao08] using a schedule-based all-or-nothing assignment method are displayed here directly.

** In Zhao’s results (Case S&M2) the line network layout hasn’t been given. So, the values from Zhao [Zhao06] using a different headway-based all-or-nothing assignment method are displayed here directly.

*** Mandl’s solution should be compared assuming less demand.

<table>
<thead>
<tr>
<th>Total fleet size</th>
<th>Comparison 1</th>
<th></th>
<th>Comparison 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>≥ 75</td>
<td></td>
<td>≥ 68</td>
<td></td>
</tr>
<tr>
<td>Total passenger travel time [min]</td>
<td>188'314</td>
<td>189'043</td>
<td>202'487</td>
<td>208'602</td>
</tr>
<tr>
<td>Total fleet size</td>
<td>76</td>
<td>82</td>
<td>77</td>
<td>95</td>
</tr>
<tr>
<td>0-transfer trips [%]</td>
<td>92.7</td>
<td>92.9</td>
<td>87.2</td>
<td>76.5</td>
</tr>
<tr>
<td>1-transfer trips [%]</td>
<td>7.1</td>
<td>6.6</td>
<td>12.1</td>
<td>23.0</td>
</tr>
<tr>
<td>2-transfer trips [%]</td>
<td>0.2</td>
<td>0.5</td>
<td>0.7</td>
<td>0.6</td>
</tr>
<tr>
<td>Total passenger travel time costs [CHF]</td>
<td>12'554</td>
<td>12'602</td>
<td>13'499</td>
<td>13'906</td>
</tr>
<tr>
<td>Operating costs [CHF]</td>
<td>9'880</td>
<td>10'660</td>
<td>10'010</td>
<td>12'350</td>
</tr>
<tr>
<td>Total costs [CHF]</td>
<td>22'434</td>
<td>23'262</td>
<td>23'509</td>
<td>26'256</td>
</tr>
<tr>
<td>Network illustrated in</td>
<td>Figure 52</td>
<td>Figure 53</td>
<td>Figure 54</td>
<td>Figure 55</td>
</tr>
</tbody>
</table>
Figure 52  GSSH Comparison 1; network design solution for Mandl’s sample network (see Figure 51).

Table 37  Line properties of GSSH comparison 1 network design solution displayed in Figure 52. The “maximum headway” is the longest headway still complying with vehicle capacities.
Figure 53  Lee’s [Lee98] network design solution for Mandl’s sample network (see Figure 51).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>4.4</td>
<td>4.5</td>
<td>44</td>
<td>0 10</td>
<td>50</td>
<td>[9, 15, 7, 10, 11], [11, 10, 7, 15, 9]</td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td>2.3 UpB.</td>
<td>46</td>
<td>0 20</td>
<td>50</td>
<td>[5, 4, 6, 8, 10, 11], [11, 10, 8, 6, 4, 5]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B3</td>
<td>3.4 UpB.</td>
<td>96</td>
<td>0 28</td>
<td>50</td>
<td>[1, 2, 3, 6, 8, 10, 14, 13, 12], [12, 11, 13, 14, 10, 8, 6, 3, 2, 1]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B4</td>
<td>6.9 UpB.</td>
<td>76</td>
<td>0 11</td>
<td>50</td>
<td>[1, 2, 3, 6, 8, 10, 11, 12], [12, 11, 10, 8, 6, 3, 2, 1]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B5</td>
<td>8.4</td>
<td>9.3</td>
<td>42</td>
<td>0 5</td>
<td>50</td>
<td>[1, 2, 4, 12], [12, 4, 2, 1]</td>
<td></td>
</tr>
<tr>
<td>B6</td>
<td>4.5 UpB.</td>
<td>36</td>
<td>0 8</td>
<td>50</td>
<td>[1, 2, 3, 6, 15, 7], [7, 15, 6, 3, 2, 1]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 38  Line properties of Lee’s [Lee98] network design solution displayed in Figure 53.
Table 39 Line properties of Case 2 of Shih and Mahmassani's [Shih94] network design solution displayed in Figure 54.
Table 40  Line properties of Baaj and Mahmassani’s [Baaj91] Case 1 network design solution displayed in Figure 55.
Figure 56 Mandl [Mandl79]; network design solution for Mandl’s sample network (see Figure 51).

Table 41 Line properties of Mandl’s [Mandl79] network design solution displayed in Figure 56.
Figure 57  GSSH Comparison 2; network design solution for Mandl’s sample network (see Figure 51).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>B2</td>
<td>6.5</td>
<td>UpB. 26</td>
<td>0</td>
<td>4</td>
<td>50</td>
<td></td>
<td>[7, 15, 6, 3, 2, 4], [4, 2, 3, 6, 15, 7]</td>
</tr>
<tr>
<td>B5</td>
<td>21.0</td>
<td>24.2</td>
<td>42</td>
<td>0</td>
<td>2</td>
<td>50</td>
<td>[1, 2, 4, 12], [12, 4, 2, 1]</td>
</tr>
<tr>
<td>B8</td>
<td>2.0</td>
<td>2.1</td>
<td>46</td>
<td>0</td>
<td>23</td>
<td>50</td>
<td>[10, 8, 6, 3, 2, 1], [1, 2, 3, 6, 8, 10]</td>
</tr>
<tr>
<td>B16</td>
<td>2.3</td>
<td>UpB. 56</td>
<td>0</td>
<td>24</td>
<td>50</td>
<td></td>
<td>[5, 4, 6, 8, 10, 11, 13], [13, 11, 10, 8, 6, 4, 5]</td>
</tr>
<tr>
<td>B23</td>
<td>4.4</td>
<td>UpB. 62</td>
<td>0</td>
<td>14</td>
<td>50</td>
<td></td>
<td>[13, 14, 10, 7, 15, 8, 15, 9], [9, 15, 8, 15, 7, 10, 14, 13]</td>
</tr>
<tr>
<td>B27</td>
<td>6.0</td>
<td>7.1</td>
<td>30</td>
<td>0</td>
<td>5</td>
<td>50</td>
<td>[10, 11, 12], [12, 11, 10]</td>
</tr>
</tbody>
</table>

Table 42  Line properties of GSSH comparison 2 network design solution displayed in Figure 57.
Figure 58 Zhao and Gan’s [Zhao03] Case S&M2; network design solution for Mandl’s sample network (see Figure 51).

Table 43 Line properties of Zhao and Gan’s [Zhao03] Case S&M2 network design solution displayed in Figure 58.
6 Conclusions

6.1 Summary of the Work

A new approach for optimizing PT network designs was introduced in this doctoral thesis. The approach enables network designers to reduce some of the simplifications used in earlier methods and therefore improve the quality of the results. A PT-related total cost approach is applied that considers the entire PT chain from door-to-door. As a result, a software prototype for planning and improving PT network design was created and tested.

The approach first divides transport networks into speed-based levels, bus and tram, regional train, etc., and planning areas and then applies a guided stochastic search heuristic (GSSH). This procedure allows a sequential optimization of PT networks for each planning area and for each speed level. The small number of network evaluations enables the use of comparable realistic transport models in.

The GSSH was compared to other PT network design approaches using Mandl’s Swiss benchmark problem. It showed promising results, although these results are limited since no schedule optimization was performed.

Due to the stochastic nature of the GSSH approach and the sequence problem during reduction, in most runs of the algorithm slightly different results were achieved (local optima). It is possible to use the GSSH approach to develop an optimum network using a specified total number of vehicles (a measure of operating costs) by, e.g., adjusting the cost balance factor. However, this is an iterative process.

The multiple area approach and the GSSH were tested on a case study in the city of Winterthur, Switzerland by optimizing the slowest speed level (bus). Optimizing several speed levels could not be tested due to the slow computing times for network evaluations. Results from optimizing one single planning area are only slightly better than from just sequentially optimizing two planning areas. The radial PT network was updated in 2007 by experienced planners. In addition, the GSSH was able to improve total costs by about 3%.

To ensure that such small changes can bring sound improvements, the accuracy of the OD matrix and the assignment model has to be sufficient. The relative error of passenger flows at maximum load sections of PT lines should not be higher than 5% for each speed level. In the future, data available from electronic ticketing systems will make it possible to validate and improve assignment models and achieve a higher level of assignment quality.
While keeping total costs at one level, the variability of time costs (ca. 7%) and operating costs (ca. 20%) could be used by politics to decide how the different costs of PT are covered and by whom (see Figure 49). Hereby, the effects on modal split should be considered. The cases 7 and 8 in Figure 49 show that it is possible to reduce the systems operating costs together with passenger’s time costs by increasing vehicle speeds.

*Cycle times* are often equivalent to the basic period of the network. Lines with *cycle times* of the basic period (a multiple of the basic period or *cycle times* evenly divisible into the basic period by small divisors) have terminal times independent from headways. Such lines with short terminal times have an advantage during the reduction process and if headways are adapted, e.g., to changes of the passenger load during the daytime.

It was shown (see Table 24 and Table 25) that *cycle times* are often 60 minutes, the basic period of the network. Beside *cycle times* of the basic period also *cycle times* of a multiple of the basic period or *cycle times* evenly divisible into the basic period by small divisors are advantageous. Lines with *cycle times* of the basic period have terminal times independent from headways. Lines with short terminal times have an advantage during the reduction process and if headways are adapted to changes of the passenger load during the daytime.

The multiple area approach was tested by optimizing the slowest speed level. Several speed levels could not be tested due to the slow computing times for network evaluations. Results from optimizing one single planning area are slightly better than from just sequentially optimizing two planning areas. Furthermore, it could be seen that it is helpful to have a kind of “ideal infrastructure network” to compare and evaluate PT networks on given infrastructures.

### 6.2 Further Research

The main potential for further improvements of the approach lies in the following:

- The approach integrates stop placement into line planning. Scheduling is considered partly in terms of headway optimization combined with flexible line alignments. Schedule optimization is an important part of PT network design and should be fully integrated. For this case, the integration of stop placement can show its full effects. Travel times can be adjusted through the placement of stops. The integration of schedule optimization allows using schedule-based assignment to model passenger behavior. First, the assignment becomes more accurate. Second, travel times can be reduced while offering coordinated transfers, especially for less loaded lines with longer headways. Schedule-based assignments allow the consideration of more frequent service close to the most loaded section of a line. One important condition for realistic schedule-based assignments is that delay distributions should be considered, at least on an average level.
• Modal split could be included, e.g., according to Lee’s approach [Lee98]. This requires a small amount of additional computing time.

• To create a new start network, it is useful to concentrate on lines with cycle times a little bit shorter than the basic period (compare [Micha09]) or other important periods of multiple headway lengths. This reduces the number of lines in a different way than by deleting the lines evaluated as worst.

• The GSSH only partly allows finding new line alignments, especially in case of heavily loaded lines. The GSSH mainly changes routes evaluated as bad. The start network should already contain the most efficient line alignments. Besides the nearly shortest routes, this can also be lines developed with a different design strategy. Such a strategy could be connecting important transfer nodes with each other, similar to the magnetic streamlines of the field (Figure 44), covering the entire area as well.

• During the final stop placement, access/egress to stops could be modeled to minimize the travel times, not only to minimize walking times. In that case, travel times are calculated of walking time and riding time. The linear relationship (see equation (14)) between the density of the potential demand and the on-link stop distances could be changed to a more realistic non-linear relationship.

• The proposed approach still requires long computing times when used to optimize network designs for large-scale networks. One possibility to accelerate the process is parallelization. Ant colony optimization and genetic algorithms can easily be parallelized. Another possibility is to use faster programming languages, such as Java, and a faster assignment algorithm adapted to periodic event schedules. If computing times for planning areas are shortened enough, computing times for network design in entire areas can be handled as well since they are only linearly dependent on total area size.

6.3 Final Remarks

Numerous PT network and schedule design studies have been conducted to date. Due to the complexity of PT network design and the variety of approaches, it is difficult to define benchmark problem sets and standard transport models that allow different solutions to be compared. However, simplifications in network design methods and in the transport models used during network design require the application of benchmark problems in order to compare the results of different network design methods. Therefore, the basis for all benchmark problems should be an accurate transport model. Benchmark problem sets and standard transport models should be simple enough to keep the amount of required data low, but
complex enough to allow a comparison of PT network design methods under realistic conditions. In the future, the development of new benchmark problems will be of great interest. In particular, larger benchmark problems and benchmark problems that cover the entire PT chain from door-to-door should be considered.

The reduction approach in combination with the ant colony optimization shows promising results, as can be seen from the results for Mandl’s benchmark problem, as well as from the solutions for the more realistic study in the city of Winterthur. However, in the case of Winterthur, network evaluations today are too slow by about a factor 10 to be able to get satisfying results. In combination with the accuracy of the transport model, this still prevents being able to answer questions concerning, for example, optimal parameters for a given demand pattern such as the number of speed levels, the line distances and the transfer node distances.
7 References


[Biel02] Bielli, Maurizio; Caramia, Massimiliano; Carotenuto, Pasquale (2002) Genetic algorithms in Bus Network Optimization, Transportation Research Part C, 10 (1) 19–34.


Investigation of space-time structures in public transport networks and their optimization


[Ceder02] Ceder, Avishai; Gonzalez, Oneximo; Gonzalez, Hugo (2002) Design of Bus Routes – Methodology and the Santo Domingo Case, Transportation Research Record, 1791 35–43


[Vrtic05] Vrtic, Milenko; Fröhlich, Philipp; Schüssler, Nadine; Axhausen, Kay W.; et al. (2005) Verkehrsmodell für den öffentlichen Verkehr des Kanton Zürich, Endbericht, Amt für Verkehr des Kantons Zürich.


A 1 Curriculum Vitae

Bernhard Cornelius Alt

Born 11th January, 1975 in Zorneding (Germany)
Citizen of Germany

Practical Experience

2004 – 2010

Research Assistant at ETH Zurich, Institute for Transport Planning and Systems (IVT),
Chair for Transport Systems

2002 – 2003

2. Student Research Project: 26 weeks at Bombardier Transportation in Netphen.
Title: “Semiaktive Querkomfortregelung eines Hochgeschwindigkeitszuges”

1995 – 1996

Zivildienst Lebensgemeinschaft Bingenheim (near Frankfurt). Work with handicapped
in a gardening and a seeds production centre

Higher Education

1998 – 2003

Studies of Mechanical Engineering, University Siegen, focusing on Mechanics, Control
Systems and Mechatronics

1996 – 1997

Studies of Physics, University Stuttgart

1995

Abitur (A level), Freie Waldorfschule Filstal (Göppingen)