Line density and headway optimisation of urban public transport

Ulrich Schäffeler, IVT ETH

Conference paper STRC 2002
Session Infrastructure and Services

STRC | 2nd Swiss Transport Research Conference
Monte Verità / Ascona, March 20-22, 2002

Line density and headway optimisation of urban public transport

Ulrich Schäffeler
IVT
ETH
Zürich

Phone: 01-633 66 88
Fax: 01-633 10 57
e-Mail: schaeffeler@ivt.baug.ethz.ch

Abstract

Public transport networks are determined by their design parameters. The values of these parameters depend predominantly on the infrastructure and operation cost budget of a local authority. However, an outstanding knowledge of the network design for public transport is needed for determining these values with this budget restriction.

In this paper special attention is given on the two design variables line density and headway. Line density and frequency are closely connected. They influence directly the two quality criteria spatial accessibility and temporal disposability of the service to the public.

The relation of these two design parameters to each other is investigated with aid of a model for the optimal link distance on parallel public transport corridors and a model for the need of tangential lines in public transport networks.

Out of these models and their results it can be assumed that simple networks with central nodes in place of networks with high line density and numerous tangential services have to be promoted. This is particularly the case, if a high level in frequency and operational performance of the services (reliability) can be guaranteed.

Keywords

headway – line density – network design problem – network modelling – public transport – supply modelling – 2nd Swiss Transport Research Conference – STRC 2002 – Monte Verità

1 Introduction

Public transport networks are basically determined by 4 design parameters:

a. Stop density

b. Line density

c. Link density

d. Headway

These four parameters are closely interconnected. Their values depend predominantly on the infrastructure and operation cost budget of a local authority. Therefore, the four parameters can be varied within this budget. For determining these values, an outstanding knowledge of the network design for public transport is needed.

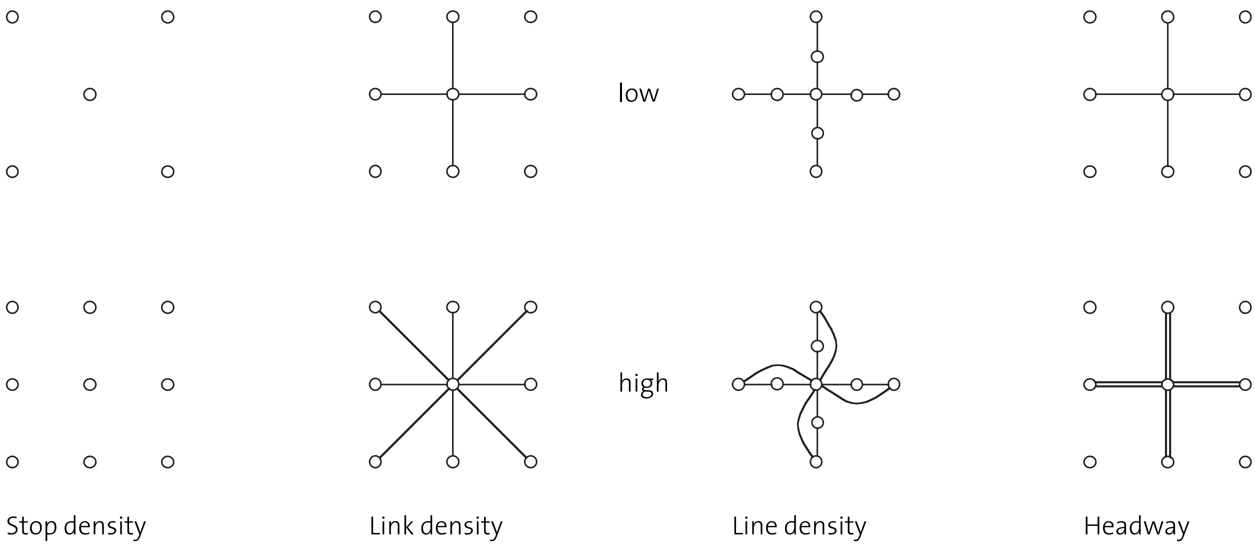

Figure 1: Network parameters for public transport

1.1 Optimisation in Network Design

For the determination of the network design parameters, goals to be achieved have to be formulated.

a. How does an optimal network for public transport look like and

b. How can “optimal” be defined in this context

“Optimal” can be defined in various ways. While in many cases the main purpose of transport policy is to guarantee a maximum of accessibility by a minimal consumption of resources, an optimisation under the following criteria shall be proposed:

Maximum use of public transport *for destinations, which are alternatively not covered by non-motorised means of transport*	⟺	*Minimum of operation* *and amortisation costs of the infrastructure*

A main challenge is to translate these optimisation criteria into mathematical functions for adequate models. Especially modelling the use of public transport appears difficult. In comparison with cost minimisation, this criterion can’t be modelled exclusively by analytical functions. Empirical functions with their uncertainty due to the difficulties of determining the input values have to be used. The question is, whether it is possible to gain anyway interesting and persistent outcomes.

Expressing cost minimisation in a mathematical function is much less complex. If operation has to be made with a fixed cost budget, this budget could also be expressed in a specific number of vehicles in question. Further, the vehicles can be dispatched on a certain number of public transport lines. Hence, every line has its own number, of vehicles, which is a function of the line density within a network. However, the number of vehicles influences the frequency on a line.

1.2 Line density and headway

In this paper special attention will be given on the two design variables line density and headway. Line density and frequency are closely connected. They influence directly the two quality criteria spatial accessibility and temporal disposability of the service to the public.

When it is assumed, that the number of vehicles in service is fixed, quality can be achieved either by high spatial accessibility (dense line network) or high temporal disposability (low headway) or a mix of these two.

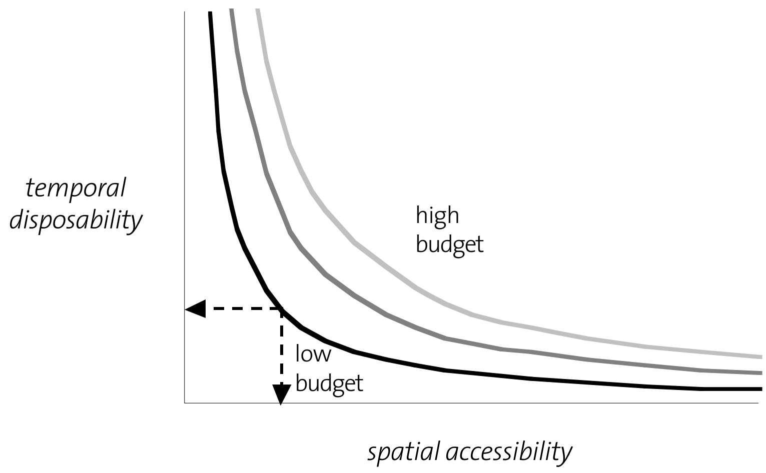

Figure 2: Relationship between temporal disposability and spatial accessibility (Iso-buget lines)

One can see in

Figure 2 that a high spatial accessibility can only be accomplished at the expense of a low temporal disposability and vice versa. Thus we need to know, which of the two quality criteria is more favourable for the entire quality of the public transport system with regard to travellers' behaviour.

The interrelationship between line density and frequency can be modelled by various different situations in network design. Furthermore, this problem will be considered not only by analytical models but also by an analysis of typical situations in existing transit networks.

In an earlier stage, research in this subject has been completed about the benefit of express lines [1].

Due to the limited scale of this paper, only the following two problems will be considered in the remainder sequel.

a. Optimal link distance on parallel corridors

b. Usefulness of tangential lines

2 Approach

The aim of our research is to find fundamental principles for solving the network design problem. Hence, it is necessary to generate models, which establish basic correlations between design parameters. Such models need to have a certain simplicity in order to be comprehensible. But this is in conflict with the aim of a precise replication of reality.

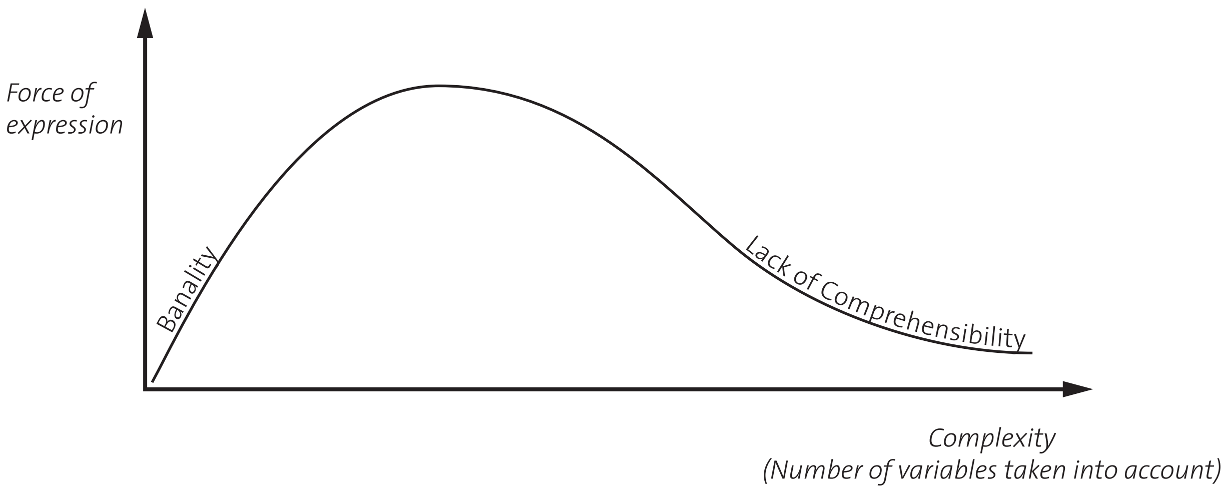

Figure 3: Dilemma of modelling detail

Figure 3 shows this dilemma of modelling, finding the right approach between banality and lack of comprehensibility.

Two main possibilities are considered for modelling network structures and getting statements on the correlation of design parameters:

a. Discussion of formulas

b. Discussion on graphs

The discussion of formulas requires a model for which an analytical solution can be found. This is mostly the case when input variables are linear.

If it is no more possible to find an analytical solution, a discussion of graphs is the only way for describing correlations between parameters. But then it is difficult to discuss more dimensional problems (more than two dimensions).

Thus, the two main approaches for the optimisation are:

a. linear → analytical solution

b. exponential → numerical solution

In this paper, only the linear approach will be presented.

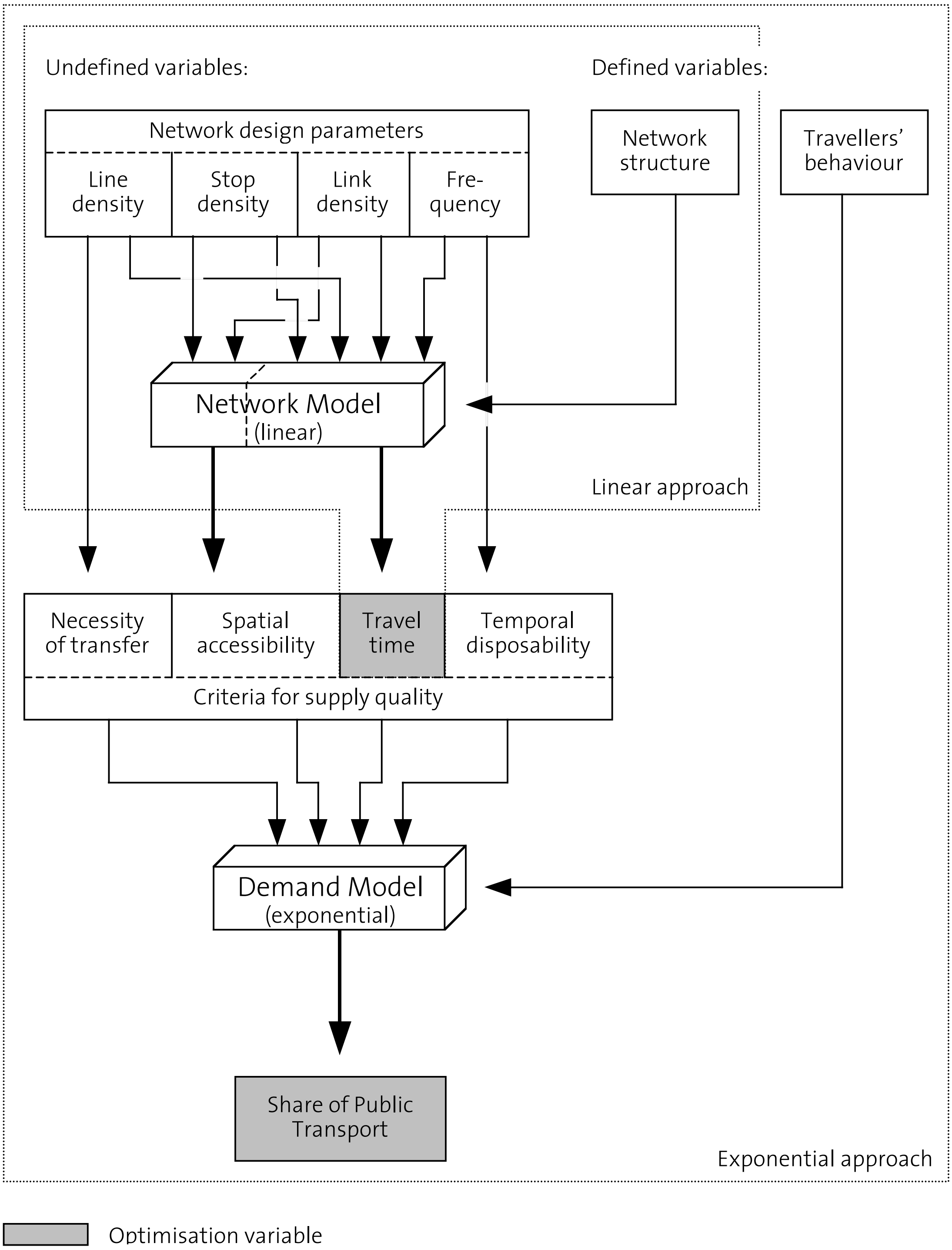

Figure 4: Flow of the optimisation

In figureFigure 4 one can see the optimisation process in network design. The planer has the four design parameters for defining a network structure. From this network structure a certain operation service results, which leads to a specific supply quality, expressed by the 4 quality criteria. This level of quality has its effect on travellers' behaviour, which can influences the share of public transport.

2.1 Linear Approach

Linear modelling implies a linearisation of all input variables. For the model of the supply quality of public transport, this requirement can be achieved without neglecting basic relations. Difficulties occur with the modelling of travellers' behaviour, the reaction of demand on a specific supply quality. A possibility for linearisation therefore is the reduction of the optimisation to one element of the supply quality. Due to the non-linear influence of the quality criteria on travellers' behaviour, the use of more than one criterion is delicate, because relations between the criteria and their specific weights have to considered in the right way.

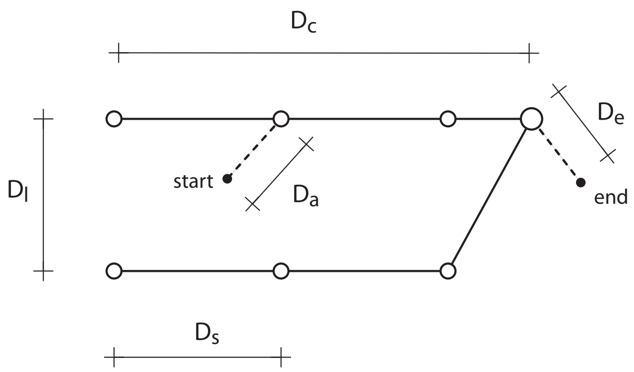

Figure 5: Variables for the linear approach

Here, the following elements of supply quality are of importance:

a. Door-to-door travel time

b. Spatial accessibility (access/egress distance)

c. Temporal disposability (frequency)

d. Necessity of transfer

The criterion "reliability" will not be considered in this step. For reasons of simplification, it is assumed, that the actual state does align with the target state.

When regarding the above listed quality criteria, the more or less direct link between the last three supply criteria and the network design parameters is striking. Only the quality criterion "travel time" or better "travel speed" cannot be derived directly from the network design parameters. For the determination of travel time a model is necessary. A main task therefore is the specification of the dependency of travel time on all other network design variables, which implicates, that the last three criteria in the list are not independent from the first criterion. This fact generates a basic difficulty in the optimisation process. Figure 4 shows this linear process within public transport network optimisation.

While travel time includes all network design parameters, it is advantageous to use as optimisation criterion for a linear approach.

2.1.1 Access Distance

The linear modelling of access distance depends on the chosen network structure (see chapter 3). However, access distance must be proportional to the sum of the two spatial network design parameters, stop distance and route path distance.

$$D_a = f_a \cdot (D_s + D_l) \tag{1}$$

For parallel route paths, the network geometry coefficient f_a can be specified with one quart (two times the half of stop and route path distance).

$$f_a = 0.25 \tag{2}$$

2.1.2 Travel Time

The model for door-to-door travel time is made up of:

a. model for access time

b. egress time

c. transportation time between two stops

d. waiting time before boarding

e. transfers

Access Time

Access time can be calculated with the model for the access distance and the agrees velocity:

$$T_a = \frac{f_a \cdot (D_s + D_l)}{v_a} \tag{3}$$

Egress Time

For trips to a city centre, the egress time T_e can be assumed as a constant value.

Transportation Time

The transportation time between two stops is determined by the acceleration time of a vehicle, the transportation time with the transportation velocity *v* and the deceleration time. Furthermore, the devell time at a stop T_{ba} has to be added.

$$T_s = \frac{D_s}{v} + \frac{v}{a} + T_{ba} \tag{4}$$

with:
$$\beta = \frac{v}{a} + T_{ba} \tag{5}$$

$$T_s = \frac{D_s}{v} + \beta \tag{6}$$

Waiting time

In urban networks, the waiting time T_w for a service can be approximated as half of the service alternating time. The alternating time itself it the reciprocal value of frequency.

$$T_w = \frac{1}{2 \cdot F} \tag{7}$$

Necessity to Transfer

Transfers can be included directly into the behaviour model. Thus a modelling of transfers for this purpose isn't necessary.

However, transfers have also an impact on travel time and therefore a model, which transforms it into time (T_u) is needed. For this model, two main situations can be distinguished:

a. transfers with connection guarantee

b. transfers without connection guarantee

In the first case, only the transfer time *U* has to be taken into consideration:

$$T_u = N_u \cdot U \tag{8}$$

In the second case (b) time for transfer is also influenced by the average waiting time for the connection service (half of frequency):

$$T_u = N_u \cdot \left(\frac{1}{2 \cdot F} \cdot U \right) \tag{9}$$

Total Travel Time

The non-weighted travel time is consequently the sum of the durations of above listed parts of a trip:

$$T_t = T_a + T_e + T_s + T_w + T_u \tag{10}$$

or:

$$T_t = \frac{f_a \cdot (D_s + D_l)}{v_a} + T_e + \frac{D_c}{D_s} \cdot \left(\frac{D_s}{v} + \beta \right) + \frac{1}{2 \cdot F} + N_u \cdot \left(\frac{1}{2 \cdot F} \cdot U \right) \tag{11}$$

2.1.3 Cost budget

The cost budget K_o is expressed by the constant number of vehicles in service. Therefore, the number of vehicles within a determined area Z_o has to be fixed. k_o expresses the operation costs of one vehicle within a certain time period.

$$K_o = k_o \cdot Z_o \tag{12}$$

The vehicles are spread on the route paths in this area (route path distance D_l). On a route path, if there is only one line going along on each route path), the number of vehicles is determined by the vehicle following distance D_o.

$$Z_o = \frac{1000\,m}{D_l} \cdot \frac{1000\,m}{D_o} = fix \tag{13}$$

The vehicle following distance is calculated as follows:

$$D_o = \frac{D_s}{F \cdot T_s} = \frac{D_s}{F \cdot \left(\frac{D_s}{v} + \beta \right)} \tag{14}$$

The cost budget can be formulated:

$$K_o = k_o \cdot \frac{(1000\,m)^2}{D_l \cdot D_s} \cdot F \cdot \left(\frac{D_s}{v} + \beta \right) \tag{15}$$

or

$$F = \frac{K_o}{k_o} \cdot \frac{D_l \cdot D_s}{(1000\,m)^2} \cdot \frac{1}{\frac{D_s}{v} + \beta} \tag{16}$$

Equation (16) shows, that frequency and link distance are direct proportional to each other. Thus, equation (16) can be simplified to the following expression:

$$F = \gamma \cdot D_l \tag{17}$$

3 Theoretical Examples

As mentioned in section 1.2, the main focus is on the two design parameters route path distance and frequency. Therefore, three network cases will be considered in which the relationsship of the two parameters can be investigated.

a. Optimal route path distance

b. Networks with tangential lines

c. Multilevel networks

Multilevel networks have already been investigated in an earlier stage of this project [1]. Here, the main results of this research will be included in the conclusions. Therefore, presentation in this paper is limited on the two first items.

The values for the input parameters in the following models will be as listed below:

Stop distance:	D_s	=	500 m
Average travel distance:	D_c	=	4000 m
Transfer time:	D_u	=	2 min

3.1 Optimal Route path distance

With the aid of the function for total travel time and the formula for the cost budget, a formula for the optimisation of the route path distance can be generated by substituting *F*.

$$T_t = \left(\frac{f_a \cdot (D_s + D_l)}{v_a} + T_e + \frac{D_c}{D_s} \cdot \left(\frac{D_s}{v} + \beta \right) \right) \cdot \left(\frac{D_s}{v} + \beta \right) + \frac{K_o \cdot D_l \cdot D_s}{2 \cdot k_o \cdot (1000\,m)^2} \tag{18}$$

The optimisation, which is minimisation of travel time, can be obtained by the derivation on route path distance (D_c)of the function above.

$$\frac{d\,T_t}{d\,D_l} = 0 \tag{19}$$

$$D_l = \sqrt{\frac{v_a \cdot k_o \cdot (1000\,m)^2}{2 \cdot f_a \cdot K_o \cdot D_s} \cdot \left(\frac{D_s}{v} + \beta \right)} \tag{20}$$

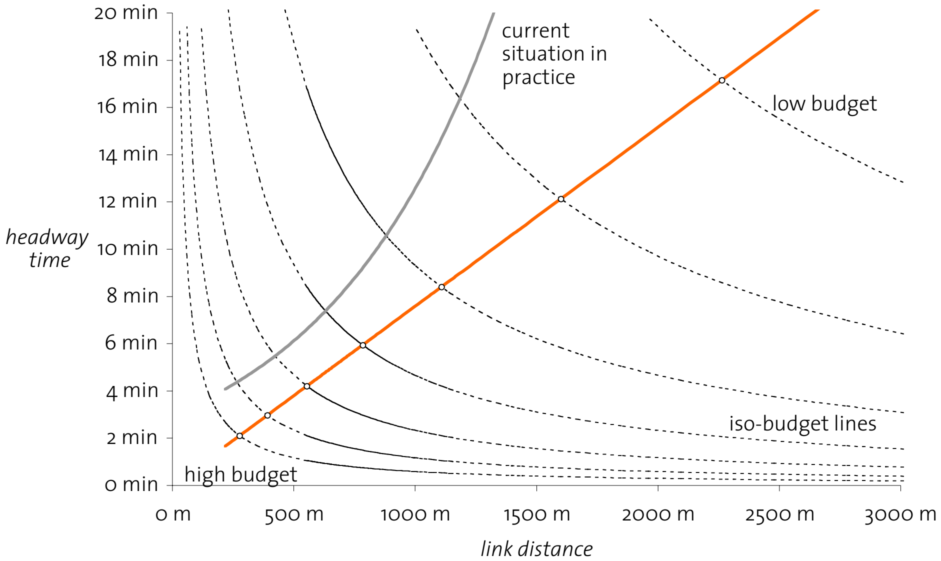

Figure 6: Optimal route path distances and service headway times for various operation costs (the indicated points show the doubling points of the operation costs)

Already in equation (15) one can see the linear correlation between route path distance and frequency. Of course, this can also be seen from Figure 6. Interesting there is how the curve for the optima tends to high frequencies and large link distances. Values in public transport networks are fare above this line, for example in Zurich, where rout path distance is in most cases lower than 1000 m and headway times about 7,5 minutes. Obviously, because of operational reasons it would be not clever to go with the service time headway below 6 minutes.

However, many cities don't even reach this operational limit for frequency. Often, headway times in public transport networks tend to be 10 or 15 minutes due to reasons of memorizing of the public.

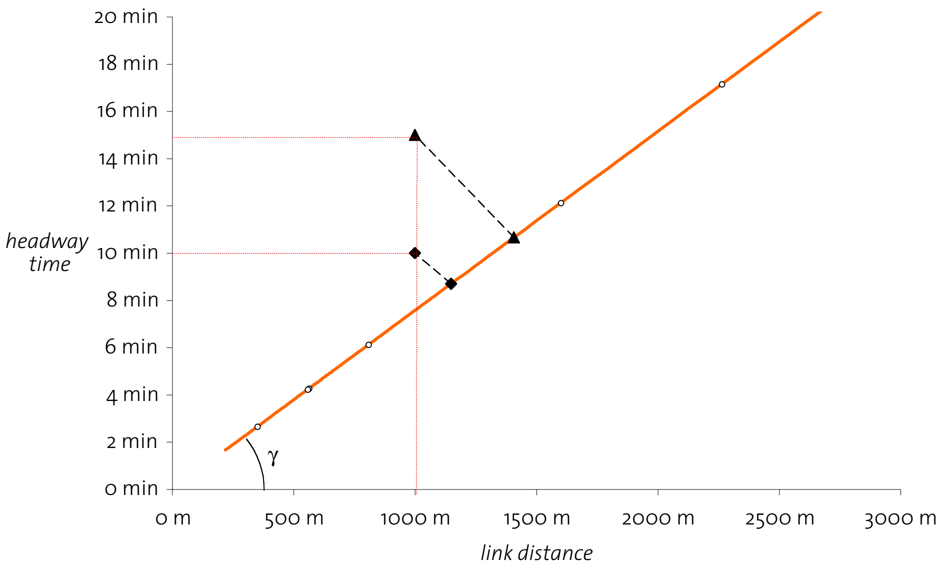

Figure 7: Figure 6 with typical network parameters of today's public transport networks (10 resp. 15 min, 1000 m)

In Figure 7 one can see the typical values of today's networks, 10 rep. 15 minutes headway with about 1000 m route path distance. The points of the optimisation curve which are linked to those values are based on the same operation cost. Thus, it is interesting to see, how frequency can be augmented when approaching the optima for total travel time.

3.2 Networks with Tangential Lines

The network model for tangential lines is based on the structure of Figure 8. The idea of this approach is to investigate, whether the operation of a tangential line is advantageous for the population of all travellers travelling form A to B. Travellers from stops in the surrounding of A resp. B would have to change for using the tangential line. These travellers will not be taken into account in this step, first it has to be proved, that a tangential line is advantageous for those, who could economise a transfer by using the tangential service.

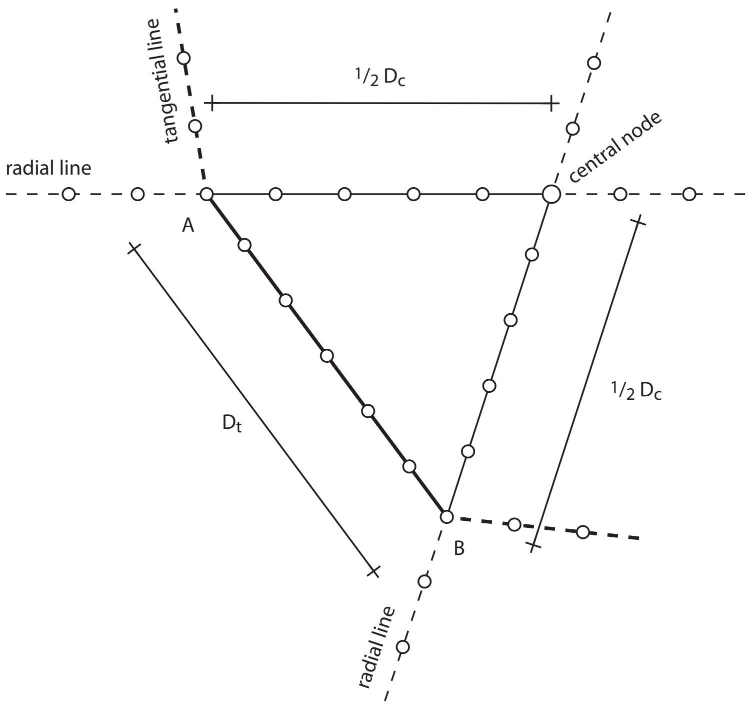

Figure 8: Network structure for tangential lines

With regard to Figure 8 one can find an answer on the question, whether it is more advantageous to travel from A to B via the central node C or whether the operation of a tangential line is beneficial.

The deviation coefficient ω shows the difference in distance between going via the central node and going on the direct tangential line.

$$\omega = \frac{D_c}{D_t} \tag{21}$$

Furthermore, the travellers with destination city centre have to be taken into account, because they have a loss in supply quality due to the operation of a tangential line (since the total budget is fixed, the number of vehicles in service in the radial line has to be reduced for the benefit of the tangential line).

3.2.1 Approach

The model is build up on a comparison between the situation with a tangential line and the situation without a tangential line. For both cases travellers to the central node C and travellers from A to B have to be considered. A tangential travel demand coefficient will be brought into the model:

$$\tau = \frac{\text{travellers between A and B}}{\text{travellers between A and C}} \tag{22}$$

The main question to be answered with this tangential model can be formulated as follows:

Which deviation must at least exist and how big has the share of tangential travellers to be, so that a tangential line is advantageous? Thus, the equilibrium, where average travel time in a network with a tangential line is lower for the population of all travellers' than in a network without one.

$$(1-\tau)\cdot T_{radial(F_2)} + \tau \cdot T_{tangential(F_2)} \leq (1-\tau)\cdot T_{radial(F_1)} + \tau \cdot T_{A-C-B(F_1)} \tag{23}$$

F_1 and F_2 are the service frequencies, that differ between the two situations with and without a tangential line (see in the beginning of this section).

In the model, the situation with a tangential line will be compared with the situation without a tangential line. The number of vehicles in service is for both situations the same.

Thus, the frequency within the network has to be determined for both situations. Of course, frequency for networks without tangential (situation 1) lines is higher than with tangential lines (situation 2).

$$F_{Situation\,1} > F_{Situation\,2}$$

The relation between the frequencies of both situations can be calculated with the aid of the operation costs (formula (15)). Because F_2 doesn't occur any more in formula (23) after solving it, a determination of F_2 is not necessary.

3.2.2 Results

For the results, two cases can be taken into consideration. First case considers transfers without connection guarantee (see chapter 2.1.2). Therefore the following formula has been derived:

$$\omega > \frac{\chi + \dfrac{1}{2\cdot F_1 \cdot \tau}}{\chi + \left(\dfrac{1}{2\cdot F_1} + U\right)} \tag{24}$$

with: $$\chi = \frac{D_c}{D_s} \cdot \left(\frac{D_s}{v} + \beta \right) \tag{25}$$

The outcomes can be seen in Figure 9.

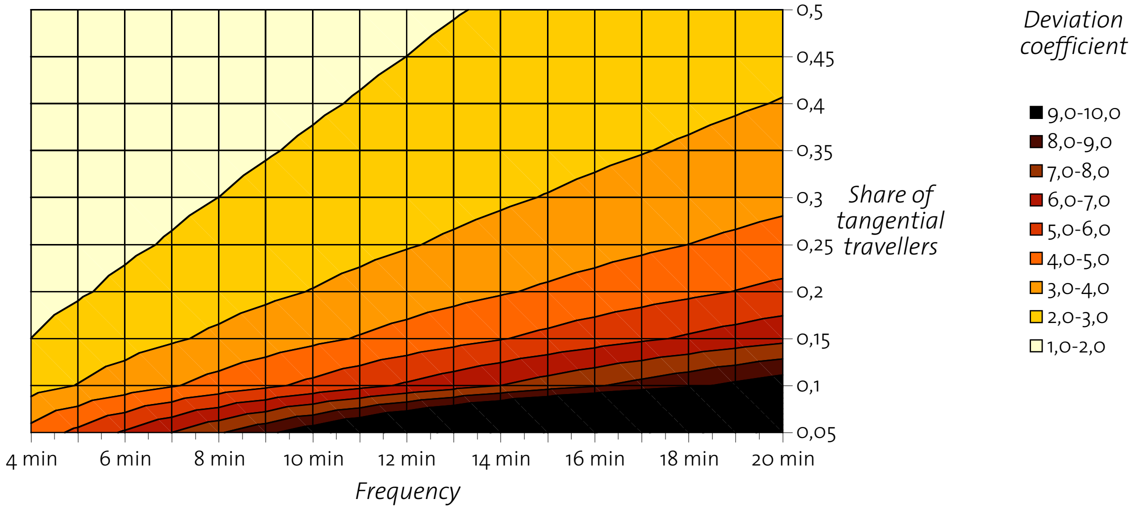

Figure 9: Deviation coefficient values for different shares of tangential travellers <u>without</u> connection guarantee

For central nodes with guaranteed connections, the formula looks a bit simpler:

$$\omega > \frac{\chi + \frac{1}{2 \cdot F_1 \cdot \tau}}{\chi + U} \tag{26}$$

The results with the same input values as already used above can be seen in Figure 10.

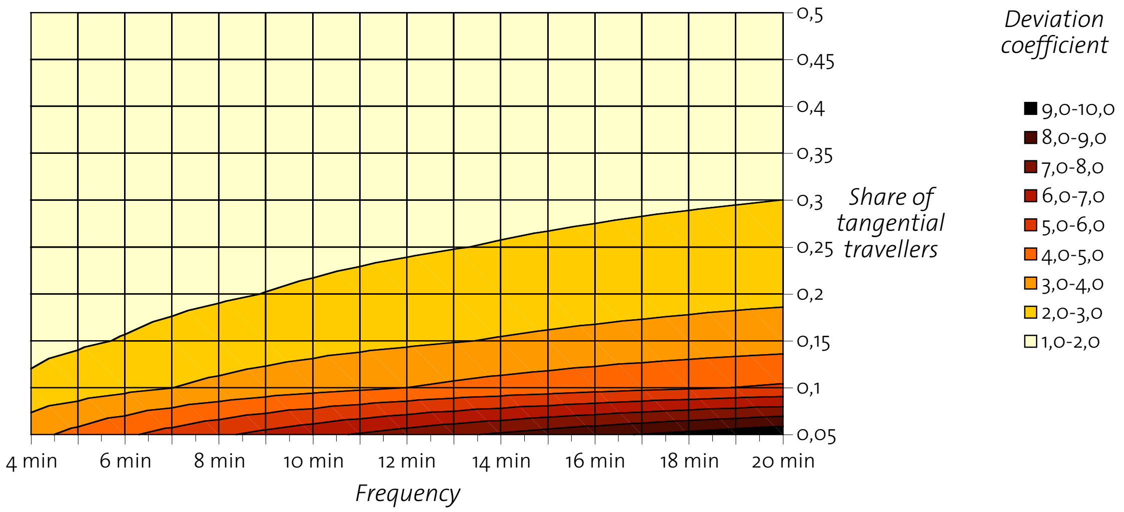

Figure 10: Deviation coefficient values for different shares of tangential travellers <u>with</u> connection guarantee

Figure 9 and Figure 10 show quiet different results. Whereas for networks without guaranteed connections the range for advantageous tangential lines services is very limited, is it for networks with guaranteed connections in a wide range favourable. Favourable are values, where the necessary deviation coefficient is lower than 3,0. Obviously, also networks with longer deviations via the central node are imaginable. But anyway, the level of service has to be high within a network for offering tangential lines. A requirement for tangential services are high frequencies or not guaranteed connections services.

For travellers starting or ending at a surrounding stop of A or B, tangential services are with regard to these outcomes not favourable. They have to make in every case one transfer. Furthermore, because of statically over determination of the network, transfers at A or B normally can't be with a guaranteed connection service. Thus, even if there is no connection service at the central node, a transfer at A or B isn't advantageous for those travellers.

An interesting approach is the situation with no deviation, $\omega = 1$ which means that the tangential line is on the same route path as the radial lines but without transfers requirement.

The following relationship can be found:

without connection guarantee:
$$F_1 \geq \frac{1}{2 \cdot U} \cdot \left(\frac{1}{\tau} - 1 \right) \qquad (27)$$

with connection guarantee:
$$F_1 \geq \frac{1}{2 \cdot \tau \cdot U} \qquad (28)$$

If τ is at least lower than 0,5, the frequency has to be higher than (two times) the reciprocal value of the transfer time. Consequently, direct lines are not advantageous in the context of this model, because transfer times are seldom higher than 3 minutes and frequencies can't be above 10 services per hour due to operation constraints.

It is obvious, that the influence of transfer time is decisive for the upset of additional lines. When taking a transfer penalty into consideration, the weighted transfer time could easily exceed 3 minutes. Hence, the equilibrium for an optimal operation of additional tangential lines can move quickly with the upset of a transfer penalty.

4 Conclusions and Outlook

Augmentation of the number of lines is quiet controversial. Though in the tangential model (chapter 3.2) tangential lines can be more favourable than transfers at a central node. But the optimum therefore is dependent on the importance of transfers and how travellers behave.

This leads to the main difference between the tangential model and an express line model [1]. Whereas in the tangential model additional lines lead to less transfers, express networks are build up on more interchange points. That means, a major penalty on transfers would lead to more tangential services but undermine the use of a multilevel network.

Therefore it is necessary to keep the transfer penalty as low as possible. Secondly, it is to be advised to promote simple networks with central nodes in place of networks with high line density and numerous tangential services.

However it is not an easy task to decrease the disadvantages of transfers for example via interchanges with connection services.

Therefore, a network has to be highly reliable, that services can keep their actual state in timetable with the target state.

The better the timetable and the better the operational performance of the services, the more are transfers acceptable and no tangential lines needed.

But another difficulty is the node structure of a network. Due to statically over determination, a network can have only a few interchanges with connection guarantee.

Without taking transfers into consideration (chapter 3.1), the network optimum on the side of low line density for the benefit of a high frequency is clear without any ambiguity. It could have been shown, that a shortening of the service internal time and an increase of route path distance is much more favourable for the total travel time than vice versa.

That’s why a main conclusion of this research is to prefer high temporal disposability instead of a high spatial accessibility.

The main uncertainty is the question of travellers’ behaviour. Till now, the models are based on objective criteria, mainly travel time. But it is known that travellers’ behaviour is not objective in a physical sense [3]. Thus it is necessary to develop a choice model for travellers’ behaviour, which is adapted to the needs of the here descript optimisation process.

Furthermore, the outcomes of this approach shall be tested in case studies of existing networks.

5 Literature

[1] Schäffeler, U.:
Verknüpfung öffentlicher Verkehrsnetze mit unterschiedlichen verkehrsplanerischen Funktionen,
Zwischenbericht zur Dissertation"Netzgestaltungsgrundsätze im öffentlichen Verkehr", Zürich 2001

[2] Schäffeler, U.:
Oft zu geringe Abstände: Haltestellen im ÖPNV
in: Der Nahverkehr, Issue 3/2001, Düsseldorf 2001

[3] Schäfer, F.:
Nachfrage effizient ermitteln
in: Der Nahverkehr, Issue 6/2001, Düsseldorf 2001

[4] König, A.:
Entscheidungsmodelle in der Verkehrsplanung
in: Materialien zur Vorlesung Verkehrsplanung, Issue 13, Zurich 2001