# Design of multimodal transport networks
## *A hierarchical approach*

Rob van Nes

# Design of multimodal transport networks
## *A hierarchical approach*

**Proefschrift**

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft
op gezag van de Rector Magnificus prof.dr.ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op woensdag 25 september 2002 te 16:00 uur
door

Robertus VAN NES

civiel ingenieur
geboren te Voorburg

Dit proefschrift is goedgekeurd door de promotor:
Prof.dr.ir. P.H.L. Bovy

*Samenstelling Promotiecommissie*

| | |
|---|---|
| Rector Magnificus | Voorzitter |
| Prof.dr.ir. P.H.L. Bovy | Technische Universiteit Delft, promotor |
| Prof.dr. N.H.M. Wilson | Massachusetts Institute of Technology, USA |
| Prof.Dr-Ing. K.W. Axhausen | Eidgenössische Technische Hochschule Zürich, Zwitserland |
| Prof.ir. L.H. Immers | Katholieke Universiteit Leuven, België |
| Prof.dr. P. Rietveld | Vrije Universiteit Amsterdam |
| Prof.dr. H.B. Roos | Erasmus Universiteit Rotterdam |
| Prof.ir. F.M. Sanders | Technische Universiteit Delft |

*"It is more important to have a clear understanding of general principles without, however, thinking of them as fixed laws, than to load the mind with a mass of detailed technical information which can readily be found in reference books or card indexes."*

*W.I.B. Beveridge*
*The Art of Scientific Investigations*
*Preparation (p. 4)*

In memory of Marga Knoester

# Preface

Multimodal transport, that is using two or more transport modes for a trip between which a transfer is necessary, seems an interesting approach to solving today's transportation problems with respect to the deteriorating accessibility of city centres, recurrent congestion, and environmental impact. Combining private transport and public transport in a truly multimodal transport system offers opportunities to capitalise on the strengths of the various systems while avoiding their weaknesses. The requirements for such a multimodal transport system, however, are high. Travellers have to be aware of the possibilities to change modes and the related benefit. Thus high quality travel information is crucial. Transfers between transport modes and services should be seamless, setting new standards for the design of transfer nodes and for the synchronisation of time-tabled transport services. Multimodal transport requires new organisational and financial arrangements between all actors involved. The most fundamental component of a multimodal transport system, however, is the multimodal transport network that consists of networks for private transport, public transport, and other transport services that are part of the multimodal transport system, including of course the transfer possibilities between these networks.

This thesis investigates the consequences of multimodal travelling for designing multimodal transport networks. It describes the characteristics of multimodal travel today and assesses its future potential. The analysis focuses on the way transport networks are organised in hierarchical network structures and determines the main mechanisms leading to these hierarchical network structures. Furthermore, an analysis is made of the role in a multimodal transport system of transport services other than private transport or public transport. The results provide new insights into the mechanisms determining hierarchical transport network structures. They show the potential impact of multimodal transport especially on the capacity requirements for public transport, and they show the possible roles of the various transport services that may be part of a multimodal transport system.

The work has been conducted at the Transportation Planning and Traffic Engineering Section of the Faculty of Civil Engineering and Geosciences of Delft University of Technology. The study is part of the DIOC research programme Seamless Multimodal Transport (SMM), which is initiated and financed by Delft University of Technology and is carried out within The Netherlands TRAIL Research School for Transport, Infrastructure and Logistics. This research programme studies all kinds of components of a multimodal transport system, such as travel demand modelling, design of transfer nodes, design of robust time-tables, and operational control of line-bound public transport services.

This thesis is the result of my research on this subject in the past five years. Some findings of this research have already been published as conference papers or journal articles. During these five years I had many inspiring discussions with my colleagues of the Transportation Planning and Traffic Engineering Section. Furthermore, parts of this thesis are influenced by previous work in my professional career at, again, the

# Contents

## REFERENCES

## APPENDIX A:    Potential for multimodal transport based on trip
                  purpose

# Chapter 1

# INTRODUCTION

## 1.1  Background

It cannot be denied that at the start of the twenty-first century the transport system in the wealthy developed countries has serious problems. The road networks suffer from recurrent congestion, the accessibility of economically important centres is deteriorating, and the negative impact on the environment is considered to be too high. The public transport systems, on the other hand, seem to be unable to cope with the changes in the transportation market. The requirements that have to be met increase, and the demand patterns are becoming even more dispersed. As a result public transport market share declines and its deficits increase. It is still unclear how these negative developments can be changed to improve the prospects for the transport system as a whole.

Multimodal transport might be an interesting approach to solve part of these mobility problems. Combining private and public transport in a multimodal transport system offers opportunities to capitalise on the strengths of the various systems while avoiding their weaknesses, and might therefore be an interesting alternative to the traditional strictly dichotomous choice between private car or public transport. The Dutch Advisory Council for Transport, Public Works and Watermanagement (Raad voor Verkeer en Waterstaat (2001)), for instance, introduces the concept of modal merge as opposed to the traditional notion of modal split. In that sense private transport could be an access or egress mode for public transport, especially for longer distances, leading to a larger market for public transport and to a reduction of car kilometres and congestion. And public transport might play an important role in the economically important centres, such as city centres, thus improving the accessibility and at the same time improving the environmental quality within these centres.

Of course, multimodal transport does already exist. Park & Ride facilities have been developed in an effort to attract car drivers to public transport. The car is used to access the train (or bus), which is then used for the main part of the trip. The Traintaxi, a dedicated shared taxi system in the Netherlands, is another example of a transport service

dedicated to collecting and distributing railway travellers. A concept focussed on the quality and accessibility of city centres is the Transferium, that is, a transfer facility at the city borders where drivers park their cars and use high quality urban public transport to access the city centre. An illustration of successful examples of multimodal transport in the Netherlands and abroad can be found in &Samhoud (2001). All these concepts and illustrations, however, are rather fragmented. They have been developed in different periods and by different actors in the transportation field. In other words, there is no real strategy for a truly coherent multimodal transport system.

A truly multimodal transport system is more comprehensive than it might seem at first sight. Given the complexity of all the transport service networks involved, especially the phenomena of timetables and transfers, it is clear that information is essential for a multimodal transport system. New information technology concepts are necessary for personal travel information before and during a trip, especially for trips that are made infrequently. Transport services should be punctual and should be synchronised at transfer nodes. The transfer itself should be as short as possible, in distance as well as in time, and preferably take place in an attractive environment. New services are possible too, for instance arranging tailor-made door-to-door trips such as the present transport services for the disabled.

It is clear that a substantial research effort is needed to determine how such a multimodal transport system should look like and how it can be realised. Which are the crucial elements from the viewpoint of the traveller, what kind of organisational arrangements are required, what is the best way to design a robust timetable, and how can a comfortable seamless transfer be arranged? Many of these questions have one point in common: they assume that there is something as a multimodal transport network. But how should a truly multimodal transport network look like and how should such a network be designed? This is exactly the subject addressed in this thesis.

Most literature on transport network design considers unimodal transport networks only: an urban network, a regional network, or a national network. Multimodal transport, however, implies combinations of different modes or transport services, and implies different hierarchical network levels, for instance for access, for the main part of the trip, and for egress, which are by definition strongly interrelated. A key question, for instance, is what the impact of increasing multimodal mobility will be on the characteristics of the public transport system. Will this lead to denser or to coarser networks? Should the location of access nodes be changed? In order to answer such questions an integrated multimodal network design methodology is needed that considers explicitly the relationships between modes or transport services and between network levels.

## 1.2   Main research questions

This thesis studies multimodal transport and its impact on transport network design, and it develops guidelines for multimodal network design. The concept of multimodal transport is used as the opposite of unimodal transport. It relates to personal trips consisting of combinations of modes, that is vehicle modes or service modes. Examples

are combinations of private transport and public transport as well as combinations of functionally different public transport services. Multimodal transport always requires transfers between modes. The difference with unimodal transfers, such as between urban bus services, is that multimodal transfers involve switching between different network levels, different modes, and different organisations. Furthermore, unimodal transport networks are often designed to minimise the number of transfers, while possible negative impact of the unavoidable transfers will be minimised. In multimodal transport networks, however, transfers are the key characteristic of multimodal transport. Minimising transfers would then ultimately lead to unimodal transport networks! These differences with unimodal transport make the multimodal transport network design problem a challenge.

The working hypothesis in this thesis is that the notion of multimodal transport will influence transport network structures. Combining modes will change the requirements for transport network design. Multimodal transport thus requires a new design methodology.

The analysis of the multimodal transport network design problem is based on the principle of hierarchical network structures in which different network levels are distinguished each having its own transport function as well as offering access to higher level networks. The concept of hierarchy does not imply that a given network level is more important than another, but is based on the notion that a transport system functions better if for different transport markets, especially by trip distance, different network levels are provided. Given the ongoing drive of people to travel faster and further, such hierarchical network structures are inevitable. As such, hierarchy can be seen as a universal phenomenon.

The main questions dealt with in this thesis are:

- What is the role of multimodal mobility today and what is it's potential in the future? Does, or may, it increase public transport usage and improve the accessibility of city centres?

- Which transport services should be considered in a multimodal transport system? Do demand-oriented transport services play an important role in a multimodal transport system?

- Which factors and which mechanisms determine the need to adopt different network levels both for private transport networks and for line-bound public transport networks, and how does the concept of multimodal transport affect these mechanisms?

- Does multimodal transport influence transport network design? Will it lead to coarser or denser public transport networks for instance, will it lead to a new hierarchy in network levels, will it introduce new network levels, or will it only influence the capacity requirements?

In order to answer these questions new analytical models are developed in this thesis for designing the different transport systems. The analytical framework is primarily based on economic objectives such as minimising total costs and maximising social welfare. These

models are applied to idealised situations, focusing on general network types and their main characteristics. In this way generic rules for the relationships within and between network levels can be derived. Given these generic rules, the impacts of combining modes or transport services can be established. Apart from the main modes private transport and line-bound public transport, an analysis is also made of the influence on multimodal transport of demand-oriented transport systems, that is, transport systems that are demand-oriented at an operational level such as rent-a-car and demand-responsive transport. The findings for all these types of transport systems finally lead to a set of guidelines for the design of multimodal transport networks.

## 1.3   Scientific and societal relevance

This thesis analyses multimodal mobility and its consequences on transport network design. It presents an assessment of multimodal mobility and its future potential. It shows that although multimodal mobility accounts only for 3% of all trips, it is important for accessing city centres and for interurban trips. Multimodal transport thus already fills roles that were expected and can be more important in the future.

The multimodal transport network design problem will be analysed using the framework of hierarchical transport networks. A uniform framework, based on well-known economic principles and a sound description of travel behaviour, will be developed for describing and analysing single-level and multilevel transport network design problems. This framework is suitable for private transport as well as for public transport services. Furthermore, it allows making a systematic comparison of different hierarchical network structures in public transport service networks.

This thesis will develop new perspectives on hierarchy in transport networks. It will be shown that hierarchy in transport networks is inevitable, due to the natural efficiency principle in travel behaviour and economic decision making. For private transport networks three different perspectives of hierarchy will be presented, each of which yielding similar results with respect to the relationships between network levels. This analysis will show that private transport systems have fundamental characteristics that lead to hierarchical transport network structures having simple relationships between network levels.

In the case of public transport services new analytical network design models will be developed using the uniform framework formulated in this thesis. These models allow a systematic comparison of single-level networks and hierarchical network structures in urban public transport networks such as express-services, trunk-feeder systems and zone systems. Special attention is given to the influence of the demand pattern and the relationship with hierarchy in spatial structures. Furthermore, the analytical models for transport network design will be extended to account for the possibility of alternative access modes for urban public transport services, and to determine the relationship between urban public transport services and interurban public transport services. In the latter case game theory will be used to account for the fact that urban and interurban transport services usually have different operators pursuing different objectives. The

analysis will establish that hierarchical public transport service network structures depend primarily on hierarchy in spatial structures.

Furthermore, an analysis will be made of the potential for multimodal transport of demand-oriented transport systems, such as rental services and demand responsive transport systems. For demand responsive transport services an analytical transport service design model will be presented to illustrate the limitations in cost efficiency that are inherent for this type of transport services.

The synthesis of the findings for hierarchical multilevel networks for both private transport and public transport services will show that the working hypothesis with which this thesis started, namely that the notion of multimodal transport will influence the way transport networks should be structured, is not true. The mechanisms that determine hierarchical transport network structures are robust with respect to multimodal transport. Multimodal transport, however, is shown to have a significant influence on the demand for interurban public transport services and thus on the required capacity.

The finding that multimodal transport does not require a new approach for transport network design limits the need for experimenting from a multimodal point of view with new concepts for infrastructure and transport services. Devoting special attention to the facilities at transfer nodes, and offering high quality transport services are more important.

The findings on optimal transport network characteristics and hierarchical relationship between network levels provide guidelines for the design of efficient and coherent transport system. Furthermore, the strong relationship between the spatial structure and public transport offers opportunities to improve the position of public transport through urban development. Finally, it is shown that demand responsive transport systems will have a limited impact on multimodal travelling. All these results limit the sometimes too optimistic view of new concepts presented in the context of multimodal transport and enable a more efficient approach to the mobility problems of today.

## 1.4  Thesis contents

The structure of this thesis is as follows. Chapter 2 introduces the subject of multimodal transport. It discusses the concepts of multimodal transport, the characteristics of multimodal mobility today and its future potential. The multimodal transport network design problem is introduced in Chapter 3. It presents literature review on transport networks and transport network design and it concludes by presenting typical characteristics of the multimodal transport network design problem. Since hierarchy is the basic principle used in this study, it is discussed in Chapter 4 for transport networks in general as well as for spatial structures.

The next two chapters analyse the main transport systems, that is private transport and line-bound transport services, and focus especially on the concept of hierarchy. The main principle used is network optimisation using economic objectives such as minimising

total costs or maximising social welfare. The private transport network is dealt with in Chapter 5. Existing guidelines and design methodologies are discussed and different mechanisms for distinguishing network levels investigated. A similar approach is followed in Chapter 6 for line-bound public transport service networks. Special attention is paid to the different concepts of hierarchy in public transport networks, and to the distinction between urban and interurban public transport.

Chapter 7 returns to the main subject of this thesis: multimodal transport network design. The findings of Chapters 5 and 6 are integrated in a multimodal context, and the remaining questions are analysed, that is the impact of alternative access modes on urban public transport network design and the potential for multimodal transport of rental services and demand responsive transport services. The combination of the previous conclusions and the results of these last analyses lead to the final conclusions with respect to the multimodal transport network design problem. Finally Chapter 8 summarises the main findings and conclusions of these thesis. It elaborates on the design guidelines for multimodal transport networks and presents recommendations for further research.

# Chapter 2

# MULTIMODAL TRANSPORT

## 2.1 Introduction

The first step in this thesis is an analysis of the phenomenon of multimodal transport itself. This chapter focuses on what multimodal transport networks are designed for, that is multimodal transportation of people. Questions that are answered are: what is meant by multimodal transport, what are the characteristics of multimodal mobility today, and what is the future potential of multimodal travelling?

The concept of multimodal transport is discussed in the following section using the layer model developed at Delft University of Technology (Schoemaker et al. (1999), Van Binsbergen & Visser (2001), Schaafsma (2001)), which will be extended to define and illustrate typical characteristics of multimodal transport in general. Furthermore, a definition of multimodal transport is given, which will be used in this thesis.

The characteristics of multimodal mobility today are analysed in Section 2.3 using an empirical analysis of the Dutch National Travel Survey. It is shown that multimodal transport is a niche market in transportation, which nevertheless plays a substantial role for specific trip types.

Furthermore, the main characteristics determining the share of multimodal travel will be established. These characteristics will then be used for a quantitative assessment of the future potential of multimodal mobility (Section 2.4). Two approaches will be developed: one based on trip purpose and the other on trip type. Both will show that multimodal travel might increase substantially but will remain a niche market.

Finally, this chapter concludes with a summary of the main findings on the characteristics of multimodal transport, the future potential of multimodal mobility, and the implications for multimodal transport network design.

## 2.2   Multimodal transport

### 2.2.1   Multimodal transport and the layer model

The layer model provides a framework to analyse the transportation system. The basic model (Schoemaker et al. (1999)) consists of three layers, Activities, Transport services and Traffic services, and two markets between them (Figure 2-1):

1. Transport market between activities and transport services;

2. Traffic market between transport services and traffic services.
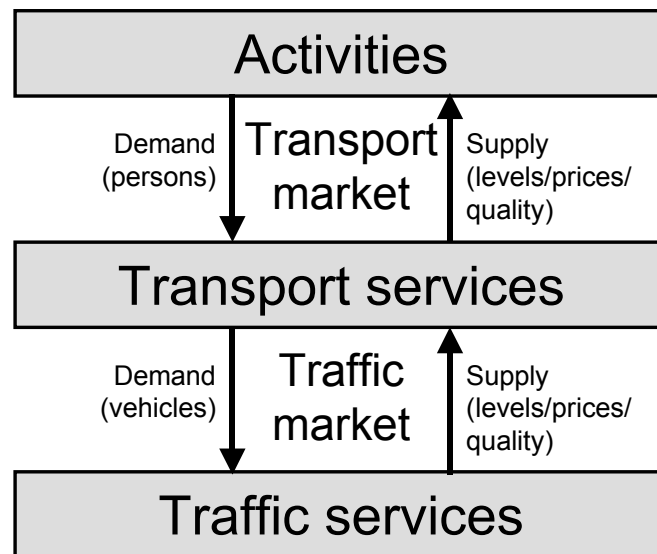
*Figure 2-1: Layer model of the transportation system*

Multimodal transport is related to the second layer: transport services. Transport services determine the quality of the whole trip from door to door, which is influenced by the vehicle, the network, and all service attributes. Transport services include private transport as well as public transport. The differences between the various transport services depend on the characteristics of all the three components and on who is responsible for the quality of those components.

In the case of public transport the concept of a transport service is quite clear. The public transport company determines nearly all characteristics of the transport service: the vehicle type, the service network, that is lines and timetables, and all service attributes, such as availability of travel information, travel costs, and the quality of the services offered. Only the infrastructure network that is available for the service network is not primarily determined by the public transport company. In the case of private transport such as private car, however, the concept of a transport service is less clear. The main point is that the driver provides transport for himself: the driver as service provider and the passenger are the same person. Just as the public transport company, the car-driver determines the quality of the vehicle and of the service during the trip, while the authorities determine the quality of the network used.

Multimodal transport implies that more than one transport service is used for making a trip, being combinations of private transport and public transport services or combinations of public transport services. This can be illustrated if the layer for transport services is split into the following elements (see Figure 2-2):
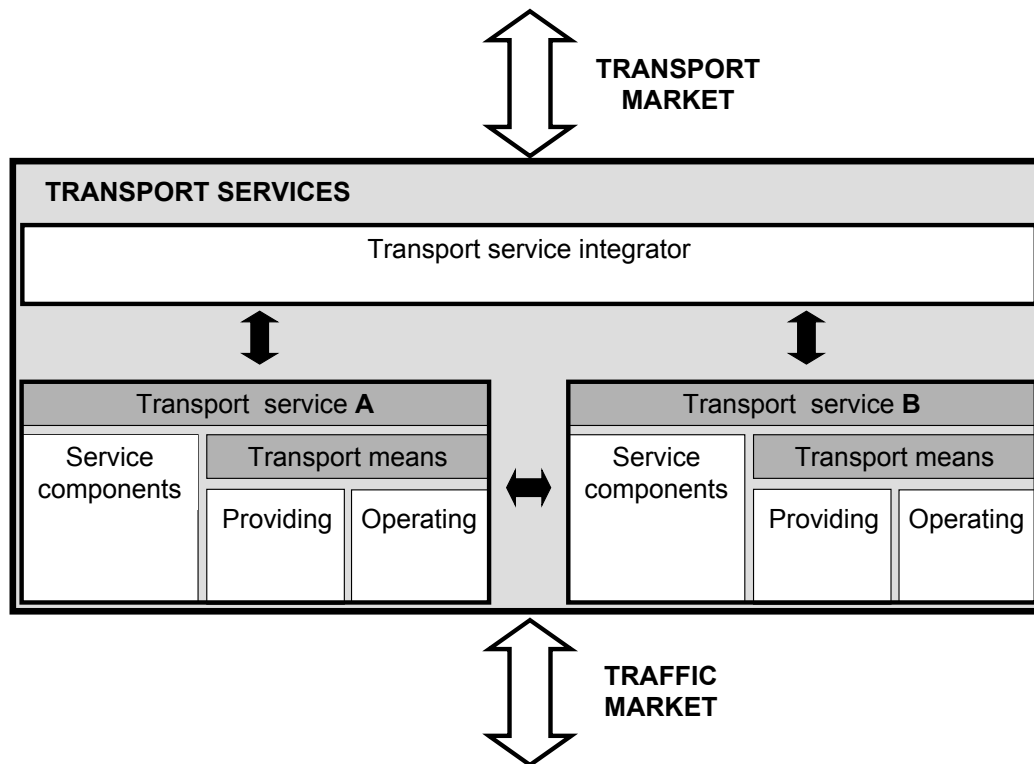


*Figure 2-2: Multimodal transport and the layer model*

- Transport service integrator, which decides or helps to decide which transport services are used for a specific trip. This might be a single transport service resulting in an unimodal trip or a combination of transport services leading to multimodal trip. Transport services are thus competing as well as working together. The role of transport service integrator is usually performed by the traveller himself, but can also be performed by a third party such as a travel agency or the Transvision company in the case of the Odessey-card. Upon request of the card-holder, Transvision arranges the whole trip using services such as rent-a-car (with or without a chauffeur), train, taxi, and Traintaxi, including all financial aspects;

- Transport services, the single or unimodal transport service such as urban, regional, or national public transport services, and private transport services as private car and bicycle, which determine travel time and travel costs. A transport service consists of service components and transport means, which are provided and operated;

- Service components, which include all components not related to the transport means, such as in the case of public transport services: the service network, ticketing, and providing information. In the case of private transport the traveller himself takes care of these aspects;

- Transport means, the vehicles used to provide transportation. They should be provided for and should be operated for the specific services;

- Operating transport means, which is taken care of by the driver and might be performed by either the traveller himself, a fellow traveller, or by a professional driver;

- Providing transport means, which can be done by the traveller himself, by a rental service, or by a public transport company. For a specific part of the trip providing transport means might include parking in the case of private vehicles, or so-called empty trips as in the case of a taxi-service.

Key elements for the transport service layer are thus transport service integrators, service components, providing transport means, and operating transport means. For both unimodal and multimodal trips, different parties might perform all these elements (see Table 2-1 for some examples). It might vary between one party taking care of all elements and a different party for each element.

*Table 2-1: Examples of parties involved in providing a transport service*

| | Transport service integrator | Service components | Providing transport means | Operating transport means |
|---|---|---|---|---|
| *Unimodal transport* | | | | |
| **Private car** | Traveller | Traveller | Traveller | Traveller |
| **Car passenger** | Traveller / (Car driver) | Car driver | Car driver | Car driver |
| **Public transport service** | Traveller | Public transport company / Information services / Selling points | Public transport company | Public transport company |
| *Multimodal transport* | | | | |
| **Private car / Bicycle*** | Traveller | Traveller | Traveller | Traveller |
| **Odessey Card** | Transvision | Railway company / Car driver / Traveller | Railway company / Car driver / Traveller | Railway company / Taxi company / Rental service |

*For instance a bicycle that is transported with the private car*

This more detailed description of the layer Transport services clearly shows three typical features of multimodal transport. First, the role of the Transport service integrator becomes more apparent. Combination of transport services requires at least good information on the transport services that are available. Furthermore, dedicated arrangements might be needed to assure a comfortable trip. Secondly, multimodal transport requires for the traveller to transfer between different transport services, being either private or public transport services. This implies a change of trip characteristics, which might have a negative influence on the attractiveness of multimodal transport. Furthermore, since each transport service involves different parties co-ordination might become a critical element. Thirdly, the service of providing transport means might be an interesting element in case the traveller hasn't any vehicle available, especially at the non-home-based part of the trip.

### 2.2.2   Definitions

The definition of multimodal transport in this thesis is that two or more different modes are used for a single trip between which the traveller has to make a transfer (see Figure 2-3). A mode might be defined by vehicle type or by transport function. The part of the trip where a single mode is used is called a leg. Typical examples are a trip in which a bicycle is used to access the railway system, or a trip is which an urban bus is used for the leg between railway station and the final destination. The opposite of a multimodal trip, that is a unimodal trip, thus are trips in which only a single mode is used, that is by private car or by a regional train service.
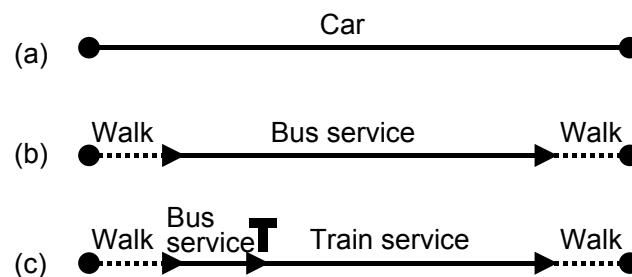


*Figure 2-3: Examples of unimodal trips (a, b) and a multimodal trip (c)*
*(the transfer point is denoted by the bold T)*

Although this definition seems to be quite simple, it deserves more discussion with regard to the following four aspects:

1.    Transfers;

2.    Modes and transport services;

3.    Trips instead of tours;

4.    The role of walking.

*Transfers* are an essential part of a multimodal trip. In order to use two or more modes travellers have to change modes at transfer nodes. However, since transfers are also a common phenomenon in unimodal public transport networks, the definition of transfers

needs to be more specific. In this thesis the term transfer is used for intermodal transfers, that is transfers where travellers change transport service networks or modes. The inclusion of transport services is essential since it implies that a transfer from one transport service network to another transport service network having other characteristics, is also an intermodal transfer. A typical example is the transfer from a regional bus to an urban bus. A transfer within a transport service network, between urban buses for instance, is then defined as an intramodal transfer. Intermodal transfers are special because they deal with different network types, which are designed separately by different operators or authorities, while for intramodal transfers usually one organisation is responsible for all these aspects.

Since a transfer implies extra travel time and/or travel costs while no distance is covered, the transfer itself has serious consequences for the transport services included in a multimodal trip. In order to be attractive compared to a unimodal transport service, the speed or the costs of a transport service in a multimodal trip should compensate for the delay and inconvenience of the transfer, as is illustrated in Figure 2-4. Multimodal transport requires fast or cheap transport services.
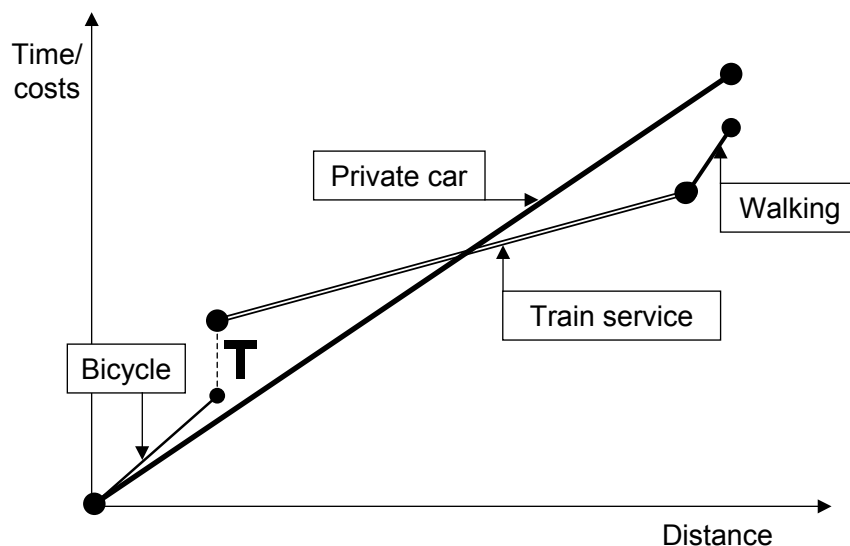


*Figure 2-4: Time/costs-distance diagram of an unimodal and a multimodal trip*

In unimodal transport network design, the transfer penalty in terms of additional time and costs usually leads to a focus on maximising direct trips, or on minimising transfers. Furthermore, the operators will also try to minimise the negative impact of the remaining transfers in their networks. Such an approach is clearly not suitable in a multimodal context, since it would ultimately lead to unimodal networks only.

*Modes and transport services* are terms that are closely related and at the same time have different meanings. A typical example of the usage of the term mode is in the mode-choice model in which the traveller's choice between for example, cycling, car, and public transport is modelled. In this context the term mode is usually associated with the vehicle used. In the case of public transport, however, the term mode is related to the service characteristics and not specifically to vehicle types such as bus, tram, metro, and

train. Thus a distinction can be made between service modes, namely private modes and public modes, and vehicle modes, which might be private vehicles such as private car and bicycle, and public vehicles such as bus and tram. Since multimodal transport is strongly related to transport services, the term mode in this thesis is usually related to service modes. Vehicle modes are thus of secondary importance. Furthermore, it should be noted that in the case of public transport services, different types of transport services can be distinguished having different characteristics with respect to accessibility, speed, frequency, fares, and vehicles used. These characteristics are often strongly related to functionally different network levels, that is urban, regional, and national public transport networks. Multimodal transport thus concerns transfers between private transport modes, between private transport and public transport services, and between functionally different types of public transport services. This distinction in mode types is illustrated in. Figure 2-5. Please note that for public transport services the vehicle mode might be ambiguous: bus, for example, might be a vehicle mode for urban public transport services as well as for regional and national public transport services.
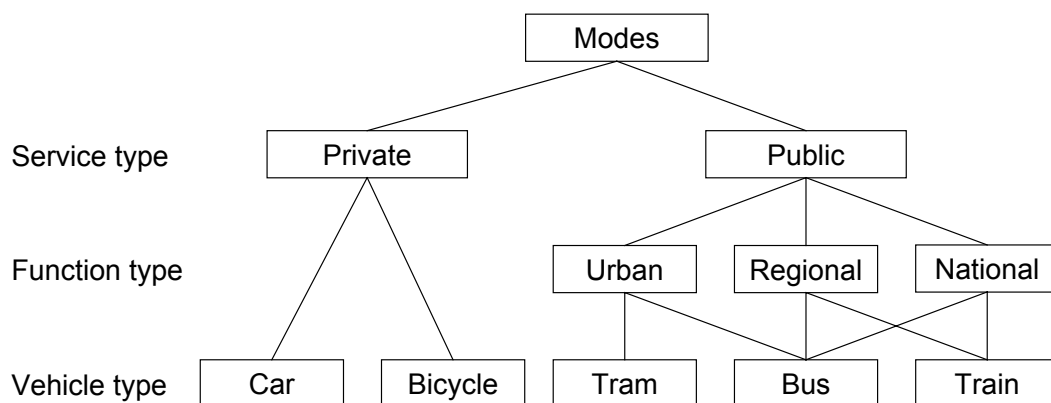


*Figure 2-5: Distinction in different types of modes*

The definition of multimodal transport should ideally be based on *tours* (see Figure 2-6) as suggested by Egeter et al. (1994) and not on trips. There is, however, a difference between a multimodal trip and a multimodal tour. A tour in which bus is used in the first trip, and in which the return trip is made as car passenger, consists of two unimodal trips. Although such a tour might be called a multimodal tour (see Figure 2-7), it is clear from this example that a multimodal tour is unlike a multimodal trip where two or more modes are used in a single trip. Both concepts, multimodal tours and multimodal trips, have their own characteristics. Multimodal transport network design is strongly related to multimodal trips. Tours become essential in describing travellers' behaviour.
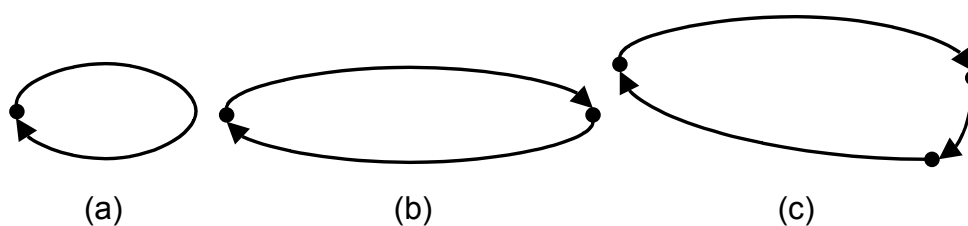


*Figure 2-6: Examples of tours consisting of 1 (a), 2 (b) or 3 trips (c)*
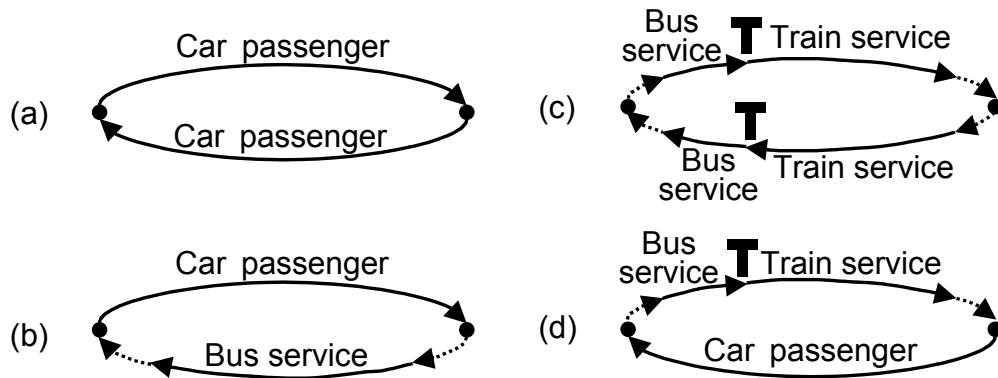
*Figure 2-7: Examples of unimodal and multimodal tours:*
*(a) unimodal tour consisting of 2 unimodal trips*
*(b) multimodal tour consisting of 2 unimodal trips*
*(c) multimodal tour consisting of 2 multimodal trips*
*(d) multimodal tour consisting of a multimodal and a unimodal trip*

On the other hand, this example illustrates that the concept of tours is relevant for the assignment of trips on a multimodal network. It is not very logical, for instance, that the return trip in this example could be made as a car driver. As concluded in the previous section, the availability of transport means plays an important role. For trips starting at home there may be more vehicle modes available than for trips starting at other locations. Furthermore, it must be taken into account that a private mode used for the first trip of the tour, should be returned to the home address. A typical example is that a traveller, who used a car as an access mode for the train, should return at the end of his tour to the railway station where the car was parked.

*Walking* is nearly always part of a trip. This is obvious in the case travellers have to walk to and from the stops of the public transport system, but using the car also requires walking to and from the parking place, although these distances might be short. Walking can thus be considered as a universal component at the start and the end of any trip, and is therefore not considered as a separate mode in the definition of a multimodal trip. Travellers who walk to the bus stop, ride the bus, and walk from the stop to their destination thus make a unimodal trip. In the case that a bicycle is used to access the bus system, however, the trip will be defined as a multimodal trip in which two services and two modes are used. The only exception to this rule is when walking can be seen as the main mode of the trip, that is, when walking is the mode used to cover the largest distance of the trip. This might occur if travellers use a bicycle or a car to access a shopping centre or a recreational area such as a park. The distance walked might then exceed the total distance covered by bicycle or car.

## 2.3   Multimodal mobility in practice

It was already stated that multimodal transport already exists, but what are the characteristics of multimodal mobility today? Characteristics such as the share of multimodal transport compared to unimodal transport and the modes used in a

multimodal trip are discussed in the first part of this section. The second part focuses on the factors that determine multimodal transport usage. These factors explain some of the characteristics found in the first part, and can be used to make an assessment of the future potential of multimodal mobility, which is presented in Section 2.4.

In order to answer these questions, an analysis was made of the Dutch National Travel Survey (CBS (1996)). This survey collects travel-data for more than 70,000 households annually, leading to data on about 600,000 trips. Since multimodal trips might have rare sequences of services and modes a combination is made of the surveys for 1995, 1996, and 1997. In this way there are more observations available for some particular combinations of modes, for instance, car driver as a main mode. The main mode is defined as the mode that is used to cover the largest distance of the trip, while the other modes used in the trip are classified as access and egress modes.

As opposed to the definitions discussed in the previous section, it was necessary in this analysis to make a distinction for public transport between the vehicle modes train, bus, and tram/metro. The database of the National Travel Survey does not allow making a distinction between service modes such as short-distance and long-distance train services, or between local bus, regional bus, and interurban bus services (Interliner), nor is it possible to distinguish intramodal and intermodal transfers. It is expected, however, that intramodal transfers are usually not reported explicitly in the survey. There are, for instance, very few trips consisting of two legs using train services. Finally, the analysis is focused on trips made by persons older than 12 years. This limitation allows a comparison with National Travel Surveys for the period 1985 to 1987, which do not include trips made by children.

### 2.3.1  Descriptive characteristics of multimodal trips

Table 2-2 shows the modal split for all trips with a distinction between unimodal and multimodal trips. The first observation that can be made is that the share of multimodal trips is small: 2.9% of all trips are multimodal. Compared to the period ten years earlier, this share has increased by 25%. This is mainly due to the fact that in 1990 the Students Public Transport Card was introduced, leading to a substantial increase in public transport usage for this part of the population. If students were not included in the comparison, the share of multimodal transport would be 2.1%, an increase of 10% in ten years.

Most multimodal trips (72%) consist of two legs, that is, two vehicle modes are used. 26% of multimodal trips contain three legs, and only 2% of multimodal trips consist of four or more legs. When looking at the main mode, that is, the mode used to cover the largest distance, train is the most important mode accounting for 59.2% of all multimodal trips. The second mode is bus, having 14.5%, followed by a group having a share of 6 to 7%: car passenger (7.3%), tram/metro (6.4%), and car driver (6.2%). Interestingly, walking is the main mode for 3.7% of all multimodal trips. Private modes are the main mode for 17.2% of all multimodal trips.

*Table 2-2: Modal split with distinction between unimodal and multimodal trips (NTS 1995-1997)*

| Main mode | All trips [%] | Unimodal [%] | Multimodal [%] | Percentage multimodal |
|---|---|---|---|---|
| Car driver | 36.2 | 36.0 | 0.2 | 0.5 |
| Car passenger | 13.1 | 12.9 | 0.2 | 1.6 |
| Train | 2.1 | 0.4 | 1.7 | 80.5 |
| Tram/Metro | 0.9 | 0.7 | 0.2 | 20.4 |
| Bus | 2.0 | 1.6 | 0.4 | 21.2 |
| Bicycle | 27.6 | 27.5 | 0.0 | 0.1 |
| Walking | 16.0 | 15.9 | 0.1 | 0.7 |
| Other | 2.1 | 2.1 | 0.0 | 1.7 |
| All modes | 100.0 | 97.1 | 2.9 | 2.9 |

Multimodal travel is dominant for the train, accounting for 80% of all train trips. The share of multimodal trips drops to about 20% for bus and tram/metro each. For the private modes it ranges between 1.5% for car passenger and 0.6% for car driver and bicycle.
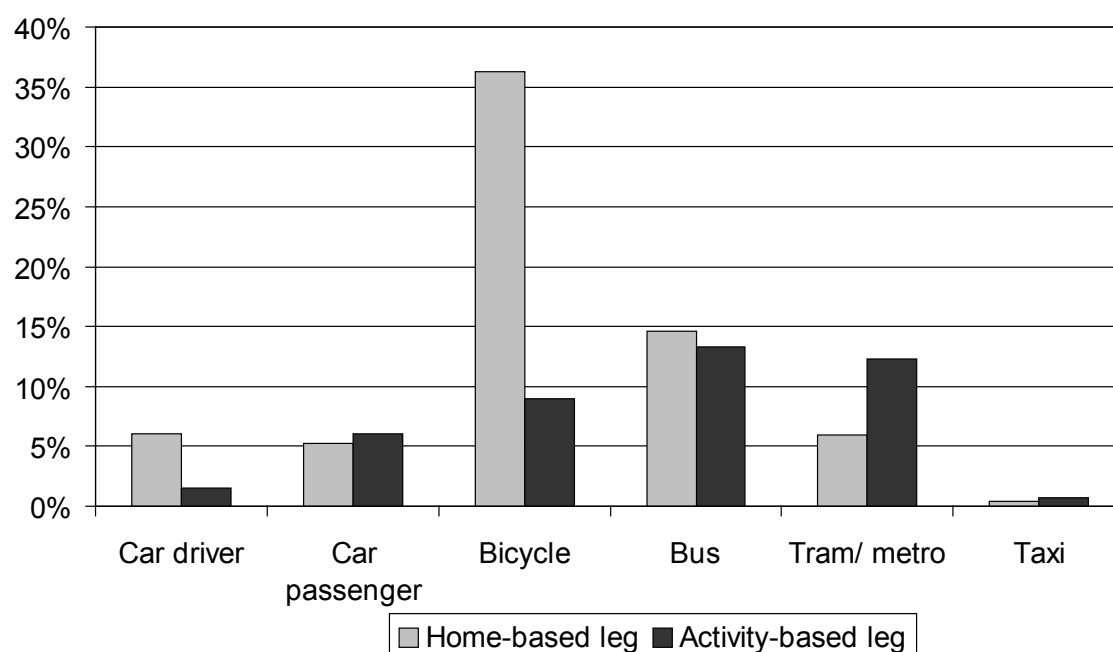


*Figure 2-8: Modal share of the home-based and the activity-based legs of multimodal trips, excluding walking (NTS 1995-1997)*

A characteristic mentioned in the discussion on trips or tours, is the difference in vehicle availability between the home-based part and the non-home-based or activity-based part of the trip. Figure 2-8 shows the modal split for the access and egress legs of the multimodal trips, excluding the percentages for walking which are 24.5% and 51.3% respectively. As expected, private modes, especially cycling, play an important role for the home-based part of the trip, accounting for 49% of the access/egress trips. For the activity-based part of the trip, this share drops to 17%. It is interesting to note that for the activity-based leg the importance of the car passenger and tram/metro modes is larger, while that for bus is equal. The share of taxi is limited to 1% for the activity-based part of the trip. Of course, these percentages are typical for the situation in the Netherlands, especially with respect to the role of the bicycle. For commuters to Chicago, for instance, Davidson & Yang (1999) reported 50% car drivers and 15% car passengers as access modes for rail transport services, while walking accounts for 83% of the egress modes.

This description of multimodal trips gives some insight into the main characteristics, but does not provide an explanation. That is the focus of the next section.

### 2.3.2    Factors determining the share of multimodal travel

Since most explanatory variables included in the National Travel Survey are ordinal variables, discriminant analysis was used to assess the main factors that determine multimodal travel (SPSS (1998)). In order to account for the contribution of the availability of vehicle modes at the home-based part of the trip a selection was made of trips starting at home. The explanatory variables included variables related to the traveller, such as age, education, and vehicle availability, and variables related to the trip itself, type of origin and destination area, trip purpose, and trip length. Three factors proved to be dominant in discriminating between unimodal and multimodal trips (in order of importance):

- Trip distance: longer travel distances have more multimodal trips;

- Type of destination area: multimodal trips are oriented to the main cities and especially the city centres;

- Trip purpose: the main trip purposes for multimodal trips are work and education.

Using these three factors 83% of the multimodal trips could correctly be classified. Trip distance is the dominant variable: directly accounting for classifying 78% of all multimodal trips. Inclusion of additional variables led to very small improvements in the classification results. Personal characteristics had only a minor influence, with the exception of the availability of the Students Public Transport Card. As already stated in the previous section, the availability of such a card has a strong positive relationship with multimodal trip making. These three main factors will be discussed in more detail, followed by some other factors such as car availability and the impact of tours.

### 2.3.2.1   Trip length

The importance of trip length can clearly be seen from Figure 2-9 and Figure 2-10, which show the trip length distribution and the differences in trip lengths with respect to the number of legs per trip.
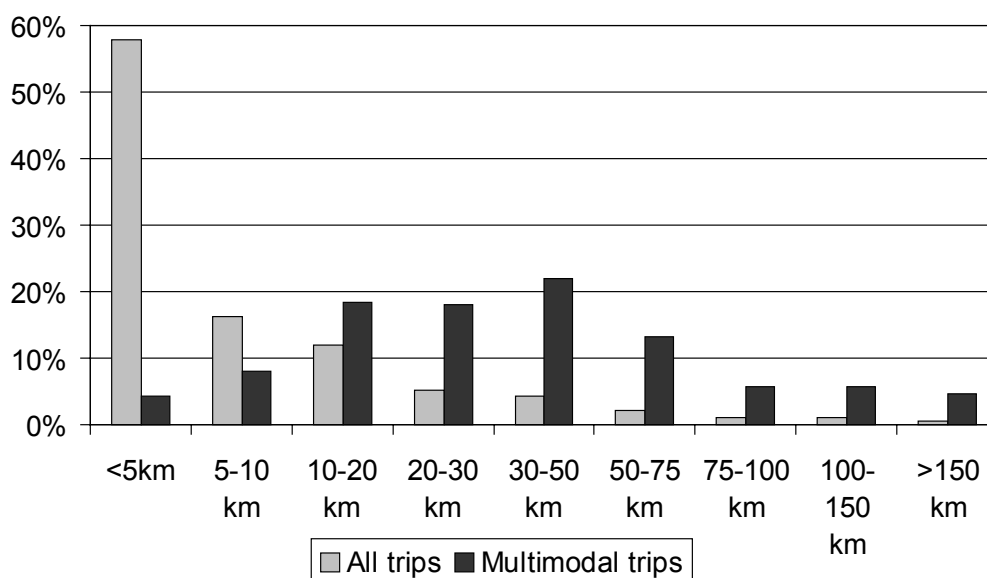


*Figure 2-9: Trip length distribution for all trips and for multimodal trips
(each trip types adds up to 100%)*
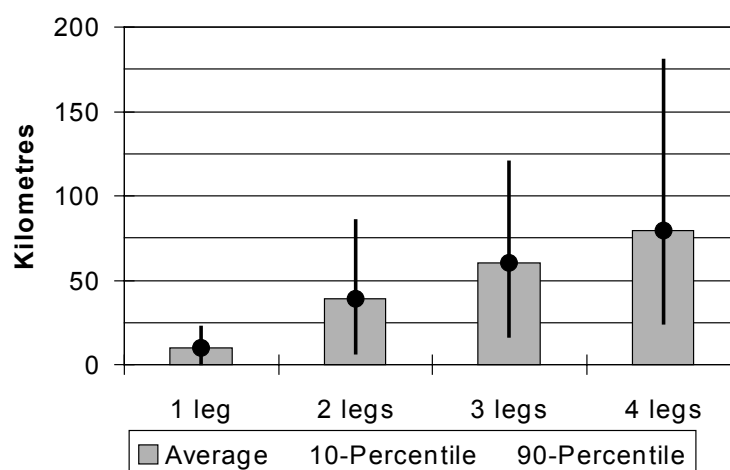


*Figure 2-10: Trip length characteristics in relation to the number of legs per trip
(NTS 1995-1997)*

The average total trip length of multimodal trips is 45 kilometres, more than 4.5 times the average unimodal trip length. Multimodal transport thus accounts for more than 12% of all kilometres travelled. Multimodal transport appears to be viable for trips longer than 10 kilometres and becomes an interesting alternative for trips longer than 30 kilometres,

having a modal share of approximately 15%. There is, however, a large difference in the distances with respect to the main modes used in the multimodal trip. Short trip lengths are found for the main mode walking (6 kilometres) and tram/metro (14 kilometres). For these main modes intra-urban trips are dominant. Bus and both car modes have medium trip lengths, varying between 26 kilometres (bus) and 42 kilometres (car passenger). It is interesting to note that for the main mode of car driver the trip length distribution includes short trips (intra-urban) as well as very long trips. The average trip length of a multimodal trip by train is 58 kilometres.
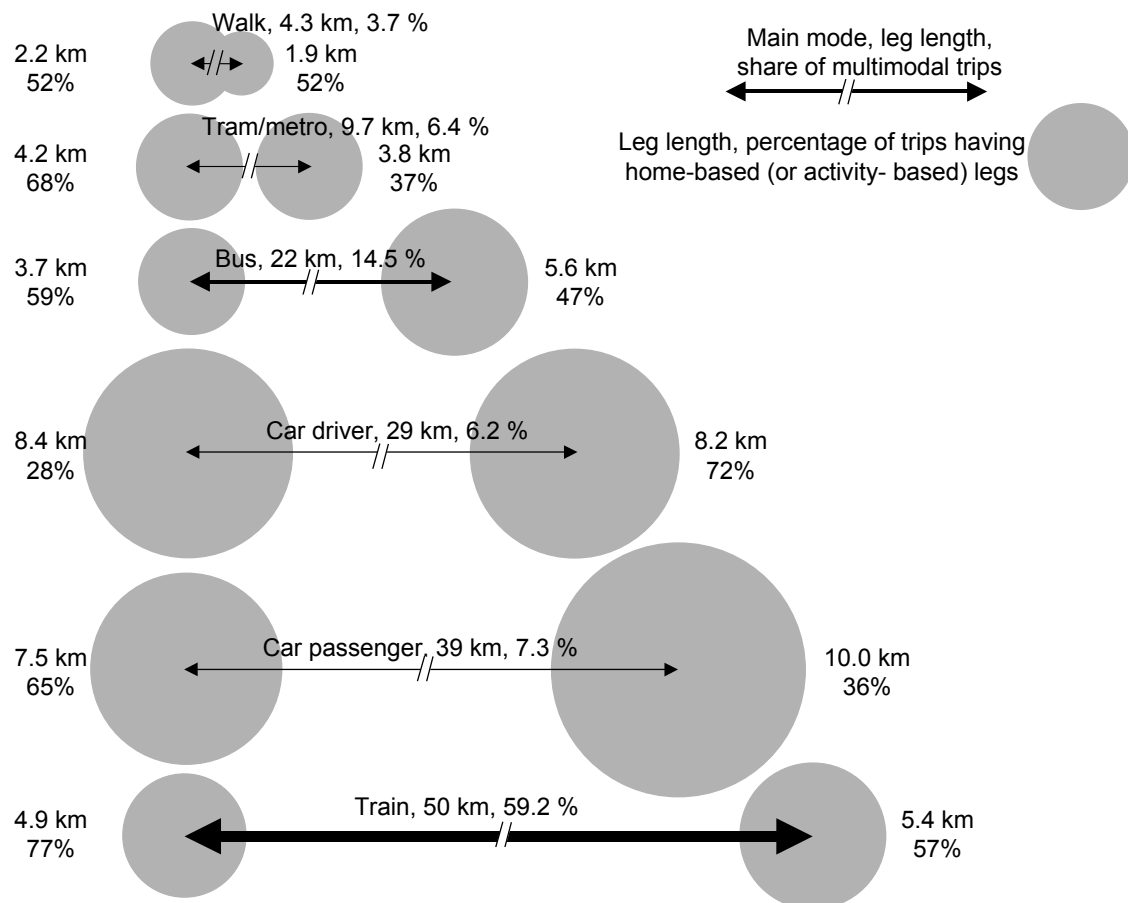


*Figure 2-11: Average leg lengths of multimodal trips by main mode for the home-based leg (left-hand side), the main mode, and the activity-based leg (right-hand side) (NTS 1995-1997)*

If multimodal trips are split into legs, there are three interesting characteristics that can be observed (see Figure 2-11). First, the differences in leg lengths between the main mode and access and egress modes suggest a hierarchical relationship between the networks of these modes. The network of the main mode is suited for long distance travel, while the networks for the access and egress modes are used for short distances. Second, the average length of the home-based leg, shown on the left, is slightly smaller than the length of the activity-based leg shown on the right, with the exception of the main modes walking, tram/metro, and car driver. Third, the access and egress legs for the main modes car driver and car passenger are substantially larger than those for public transport. In fact, these trips can be distinguished in trips having relatively short access distances,

especially by bicycle, and trips having very large access distances, for instance car passenger and car driver respectively in more rural areas. These latter trips are responsible for the relatively high average access distances. Furthermore, the percentage of multimodal trips having a home-based leg, is relatively low for the main mode car driver (28%), while for all other main modes the percentage of home-based legs is larger than that for the activity-based legs. In the case of the main mode train the percentages for both leg types add up to 134%, indicating a relatively high share of trips having three legs or more.

### 2.3.2.2   Type of destination area

The second discriminating factor is the destination area type. Figure 2-12 shows the distribution with respect to area type for the origin, that is the home-based part, and the destination, the activity-based part. For both ends of the trip, multimodal trips have a lower share for rural areas and a higher share for the main cities and the city centre. It is clear, however, that the difference between unimodal trips and multimodal trips is more distinct for the destination part.
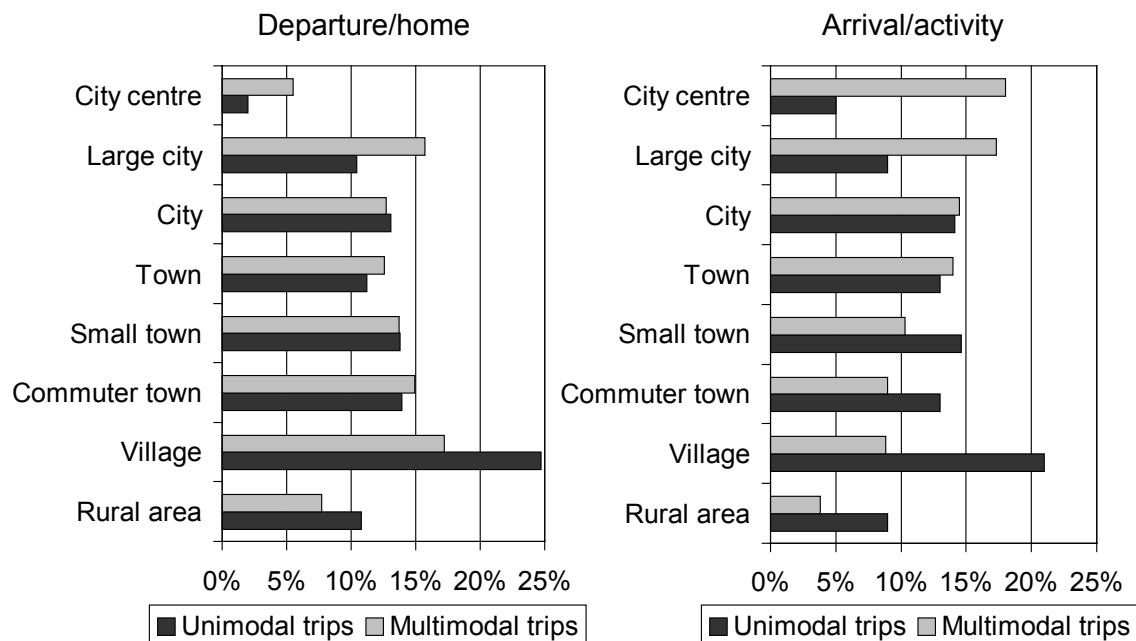


*Figure 2-12: Distribution of unimodal and multimodal trips with respect to area type for departure (home-based) and arrival (activity-based) (NTS 1995-1997)*

The importance of the main cities as the destination of multimodal trips explains the large share of tram/metro for the activity-based leg, as shown in Figure 2-8. The fact that the share for bus was indifferent with respect to the home-based leg and the activity-based leg is due to the fact that no distinction could be made between urban bus and regional bus. A rough classification of bus trips indicates that urban bus is more important for the activity-based leg too, just as is tram/metro.

Given the importance of trip length it is interesting to give special attention to the consequences of focussing on longer trips combined with area type. This is done by making a selection of interurban trips only. The shares of multimodal trips in interurban trips and in for trips to the four main cities in the Netherlands are 7% and 8% respectively. Combination of both trip types, i.e. interurban trips to the four main cities, shows a modal split of more than 20% for multimodal transport.

### 2.3.2.3   Trip purpose

The third discriminating factor for multimodal travelling is trip purpose. As can be seen in Table 2-3, multimodal transport plays an important role for work and especially for education trips. The latter illustrates, again, the strong relationship between the Student Public Transport Card and multimodal transport. Multimodal transport appears less interesting for trip purposes such as shopping, touring and picking-up or dropping off passengers.

*Table 2-3: Trip purposes with distinction between unimodal and multimodal trips (NTS 1995-1997)*

|  | All trips [%] | Multimodal [%] | Percentage multimodal |
|---|---|---|---|
| Work | 17.7 | 31.4 | 5.3 |
| Social | 15.6 | 14.8 | 2.8 |
| Education | 4.6 | 21.4 | 14.0 |
| Shopping | 24.5 | 9.4 | 1.1 |
| Business, private | 2.2 | 1.0 | 1.3 |
| Business, work | 3.1 | 2.4 | 2.3 |
| Recreation | 12.4 | 11.2 | 2.7 |
| Touring | 4.3 | 1.4 | 10 |
| Personal care | 3.1 | 1.3 | 1.2 |
| Pick-up/drop-off | 6.9 | 0.9 | 0.4 |
| Other | 5.6 | 4.8 | 2.6 |
| Total | 100.0 | 100.0 | 3.0 |

The importance for the work trips is to be expected. This trip purpose has a strong orientation to the centres of the main cities. Given the availability of the Students Public Transport Card the high share of multimodal trips for the trip purpose education is also not surprising. Similarly, the short distances related to shopping explain the low share of multimodal transport for this trip purpose. Such explanations, however, can not easily be found for the results for social and recreation trip purposes.

The fact that the share of multimodal transport for the work and education trips is higher than average leads to four observations. First, it is interesting to note that even though these trips have at least one transfer, they have a high trip-frequency, which implies that the penalty of a transfer seems to be acceptable. At least the overall benefits prevail over the discomfort of the transfer. Second, these trips are usually made in peak periods, periods in which the quality of public transport in terms of time-accessibility is usually the best. Third, the trip-frequency related to these trip purposes indicates that sufficient knowledge of the transport system may be expected to be available. Finally, it should also be noted that during the periods that most of these trips are made, the quality of the car system is worst due to congestion and that parking is often difficult and expensive.

### 2.3.2.4   Other factors

The discriminant analysis not only showed the main factors determining multimodal mobility, it also showed the factors that are not decisive for multimodal travelling. The fact that personal characteristics such as age and income have only a minor influence was already discussed. There are, however, two characteristics that need further discussion, namely car availability and the availability of a railway station. Furthermore, given the theoretical importance of tours, special attention is given to multimodal travel in complex tours.

Car availability, defined as having a driver's license and having a car in the household, has no substantial impact on multimodal mobility. If a distinction is made between travellers who have a car available and those who do not, the same discriminating factors resulted. It was also found, however, that more than 50% of the travellers making a multimodal trip had no car available, which is nearly twice the share for all travellers. Interestingly, this percentage is not influenced by the availability of Students Public Transport Card. A comparison for car availability between the characteristics of unimodal trip makers and multimodal trip makers, showed that in the case of multimodal trips car availability dropped substantially for the following characteristics:

- Travellers between 25 and 65 years, and especially for travellers between 30 and 50 years;

- Travellers who are participating in the labour market;

- Travellers having a relatively high level of education;

- Travellers having a relatively high personal income;

For all these groups the percentage of travellers having no car available is two or three times larger in the case of multimodal trip makers. This is exactly the opposite of what might be expected for these groups, since they generally have a relatively high level of car availability. Apparently, there is a group of travellers who make an explicit choice to have no driver's license or no car, and thus choose public transport. This public transport oriented group of travellers, however, provides no discriminating characteristics for classifying multimodal trips.

It might have been expected that, due to the importance of train in multimodal trip making, the availability of a railway station should also have been a discriminating factor. It is found, however, that railway station availability is of secondary importance if area type is also considered. Two remarks can be made with respect to this finding. First, there is a correlation between railway station availability and area type. Second, the variable area type is more detailed than a binary variable indicating whether a railway station is available or not. Furthermore, area type also stands for other characteristics, such as urban densities with respect to workplaces and urban facilities, and the availability of service modes to reach destinations that are further away from the railway station.

The distinction between trips and tours was already discussed in Section 2.2.2. It is usually assumed that unimodal transport is perfectly suited for more complex tours, that is, tours having two or more activities: unimodal transport offers the flexibility needed for complex tours. Table 2-4, however, shows that there is certainly no evidence to support this assumption. The share of tours consisting of at least one multimodal trip is even larger for more complex tours than for common tours consisting of two trips. In a way it seems to work the other way around. Common trips are easily made using unimodal transport and if travellers have to plan their trips, either due to the high trip frequency such as in the case of commuter trips, or in the case of complex tours, multimodal transport becomes a more interesting alternative. Another possible explanation might be that the complexity of multimodal trips makes it interesting to combine activities in a complex tour, which is made more attractive by the concentration of activities around major stations.

*Table 2-4: Share of multimodal trips per tour type (NTS 1995)*

|  | **Percentage of all tours** | **Percentage having multimodal trips** |
|---|---|---|
| **1 trip tours** | 10.7 | 3.0 |
| **Simple tours (2 trips)** | 68.0 | 2.4 |
| **Complex tours (> 2 trips)** | 21.3 | 3.0 |
| **All tours** | 100.0 | 2.6 |

### 2.3.3  Conclusion

This section described and analysed the characteristics of current multimodal trips. It was found that the share of multimodal transport is small: 2.9% off all trips are multimodal trips. Public modes, and especially train services, play an important role as the main modes used in multimodal trips. They account for 83% of all multimodal trips. Private modes, on the other hand, play an important role in the access and egress to and from the main modes. For the home-based leg they account for 65% of the trips and for the activity based leg for 35% (excluding walking).

The main factors associated with multimodal transport are trip length, destination area type and trip purpose. Multimodal transport is interesting for longer trips, that is more

than 10 and preferably more than 30 kilometres, is focused on larger cities and especially city centres, and is mostly used for work and education trips. Furthermore, it is found that multimodal transport has been stimulated by the introduction of the Students Public Transport Card.

This analysis of the Dutch National Travel Survey thus confirms the expected benefits of multimodal transport formulated in Chapter 1. Multimodal transport increases the patronage of, especially, long distance public transport. Given the share of 80% multimodal trips for train services, multimodal transport might even be seen as crucial for train usage. Second, multimodal transport is already an interesting alternative for long-distance trips, and therefore plays an important role in reducing the number of car kilometres. Finally, multimodal transport has already a substantial modal split, up to 20%, for trips to and from the main cities, and thus improves accessibility for these cities.


## 2.4   Future potential of multimodal transport

Given the fact that the share of multimodal mobility today is small, the next question is what is the potential of multimodal travel in the future. Will all possible improvements in the transport system lead to a doubling of the percentage of 2.9%, or more? What are the limitations to the potential of multimodal mobility? In order to answer these questions both sides of the transport market should be considered, that is the supply side and the demand side. Given the characteristics of current multimodal trips (see Section 2.3) and the possible changes on the supply side, the transport system, two assessments are made. The first is based on trip length and trip purpose, and the second on trip length and trip type.


### 2.4.1   Supply side considerations

Focusing on multimodal transport might lead to several improvements in the transportation system. Five important possible improvements are:

- Transfers become less uncomfortable due to better design of transfer nodes and better synchronisation of public transport services;

- Accessibility of public transport services for private modes might be improved, especially in rural areas having low densities;

- Availability of transport modes at the destination end might be improved by for instance rental services, especially for areas having a low quality of public transport services;

- Information might be more easily available, before and during the trip, thus making it easier to plan and complete multimodal trips;

- Financial aspects of multimodal trip making might be simplified, either by electronic payment facilities or by transport service integrators.

Improvements in the transport services themselves are not included in this selection, because these improvements are independent of a focus on multimodal transport.


### 2.4.2    Demand side considerations


When discussing the potential of multimodal mobility it is essential to make a distinction between niche markets and the transport market in general. A typical niche market is long-distance travel. In this market many travel agencies exist who arrange tailor-made trips either for recreation or for business purposes. An example is the earlier mentioned concept of the Odessey-Card of Transvision in the Netherlands, mainly for business trips. Upon request, Transvision arranges the whole trip using services such as rent-a-car, with or without a chauffeur, train, taxi and Traintaxi, including all financial aspects. Typical characteristics of these niche markets are that they are fully demand driven, and that they focus on additional services given the existing transport system. The fact that these niche markets are demand driven implies continual prospects for change in suppliers and in services. The market for the Odessey-Card, for instance, proved to be too small to make it profitable for the operator, and the service ceased in 2000.

The transport market in general, on the other hand, is characterised by a large number of trips having diverse characteristics, a strong relationship with infrastructure and spatial development, and a substantial involvement of the authorities. The transport market in general thus determines transport network design. There is, of course, still a strong relationship between supply and demand, however, a typical characteristic of this relationship is that it is strongly determined by basic activities such as work, education and shopping. Furthermore, given the common nature of these trips and the resulting trip frequencies, the focus is more on fundamental characteristics of travelling, that is travel time and costs, than on all related service aspects.

Finally, in assessing the potential of multimodal mobility the results of the analysis of current multimodal mobility should be taken into account, especially the critical factors of trip length, destination area, and trip purpose.


### 2.4.3    Assessment based on trip purpose


An estimate of the potential of multimodal mobility by simply doubling the current share clearly ignores typical characteristics of trips and of multimodal transport. For some trip purposes, for instance, multimodal transport might not be a credible alternative. A more educated guess can be made using the fact that multimodal mobility has a relatively large share for the work trip purpose. These trips are made frequently, implying a relatively good knowledge of the transport system, and are made during peak hours, in which the quality of the public transport system with respect to time-accessibility is best, while the quality for the alternative of private car is worst. It should be noted that the share of multimodal trips is higher for the trip purpose education, but this is a result of a strong bias towards public transport in general due to the Students Public Transport Card.

In this first assessment of the potential for multimodal mobility the usage of public transport for the work trip purpose is seen as a standard for all other trip purposes. For some trip purposes the potential of multimodal travel will be minimal, either because of relatively high usage of public transport (education) or because multimodal transport is not suited for these trips (touring, picking-up/dropping-off). These trips are excluded from the analysis. Furthermore, the analysis is focused on trips longer than 10 kilometres, the trip length where multimodal transport starts to play a role in modal split. Two classes are distinguished that account for nearly 95% of these trips:

1.   10 - 30 kilometres having an average share of multimodal trips of 6%;

2.   30 -100 kilometres having an average share of multimodal trips of 14%.

Substantial improvements in the transport system, focussed on multimodal travel, will increase the attractiveness of multimodal transport. Improvement in transfers alone might lead to a reduction of the travel time of a multimodal trip by 10%. Given the elasticities of the Dutch National Transport Model (AVV (1990)), a corresponding increase in demand of 10 to 15% might be expected. A more optimistic estimate might be that patronage might increase by 50% due to the combination of all possible improvements mentioned before.

In the most optimistic scenario (see Table 2-5) a maximum increase of the public transport share for work related trips is assumed to be 50% for each class, and the public transport share for all other trip purposes set equal to these values. The result is that the share of multimodal travel increases by more than 70% (see Table 2-5). An analysis of the trip purposes that account for this increase, shows that visiting accounts for 25% and 33% respectively, followed by work and recreation (19/22%, 19%). For the medium distances, the share of shopping in this increase is 18%, while for the long distances business trips account for 11% of the growth.

*Table 2-5: Potential for multimodal transport for three scenarios based on trip purpose*

|  | 1: Maximum | 2: Medium | 3: Minimum |
|---|---|---|---|
| **Scenario characteristics** |  |  |  |
| Growth factor for work trips | 50% | 50% | 15% |
| Correction for car passenger, time of day, value of time | No | Yes | Yes |
| **Results** |  |  |  |
| Growth class 1: 10-30 km | 81% | 30% | 14% |
| Growth class 2: 30-100 km | 76% | 30% | 13% |
| Growth multimodal transport | 72% | 28% | 13% |
| Resulting multimodal share | 5.0% | 3.7% | 3.3% |

This approach however, does not take into account the specific characteristics of the trip purposes. For some trip purposes, for instance visiting and recreation, the share of car passengers is relatively high, making private transport a relatively efficient mode.

Similarly, there are trip purposes for which a substantial share of trips is made in the evening, a period in which the service level of public transport is lower. Finally, since business trips have a high value of time, private transport might be, again, the most efficient mode. The second scenario takes these constraints into account pragmatically, limiting the growth of multimodal trips to 28% or a share of 3.7% (see Appendix A for a more detailed description of this analysis). The impact of the different trip purposes on this increase changes substantially. Work accounts for 50% or more. Other trip purposes for the medium distances are shopping and visiting with shares of 15% and 12% respectively. For the longer distances shopping (16%) and recreation (9%) are other important trip purposes.

The third and most conservative scenario is based on a general increase in work trips of 15% and takes into account all the constraints mentioned above. In this scenario the growth in multimodal trips is limited to 13%. Work is again the dominant trip purpose, having a share of 32% and 38%. Other important trip purposes are visiting (16% and 23%) and recreation (12% and 13%). Just as in the maximum scenario shopping is an important trip purpose for the medium distances (19%) and business for the longer distances (11%).

### 2.4.4    Assessment based on trip type

The second assessment focuses on the distinction between different trip types that will be influenced by the concept of multimodal transport. Five trip types are distinguished which differ with respect to the quality of the public transport system offered:

> A:    Between city centre and another agglomeration or town: usually train and one transfer to an urban public transport system;
>
> B:    Between agglomerations: train and two transfers to urban public transport systems;
>
> C:    Between city centre and other origins and destinations; train and one transfer to a regional public transport system;
>
> D:    Between agglomeration or towns and other origins and destinations; train and one transfer to an urban and one transfer to a regional public transport system;
>
> E:    Between all other origins and destinations; bus and regional trains and two transfers to regional public transport systems.

Figure 2-13 gives a schematic representation of these trip types and includes for each trip type the public transport share for trips between 10 and 50 kilometres and for trips longer than 50 kilometres plus the share of all trips (Van Goeverden & Schoemaker (2000)). If it is assumed that the differences in modal split between these trip types can be explained by the main characteristics of the transport system a simple model can be made to determine the share of public transport for each trip type (see Appendix B). The difference in speed for the main mode leads to a reduction in public transport usage of 45%, the addition of a transfer to an urban public transport system to a reduction of 51%, and the difference between urban and regional transport system to a reduction of 26%.

Urban agglomeration          Urban agglomeration

City centre                                A: 44/50                         City centre

B: 21/25 (2.5)
A: (0.6)

A: 44/50

C: 33/36 (0.6)                                                     Town

D: 11/15

D: (7.7)

D: 11/15

E: 4/6 (6.5)            Other origins
                        or destinations

Trip type, Percentage public transport 10-50 km /more than 50 km, (Share of all trips)
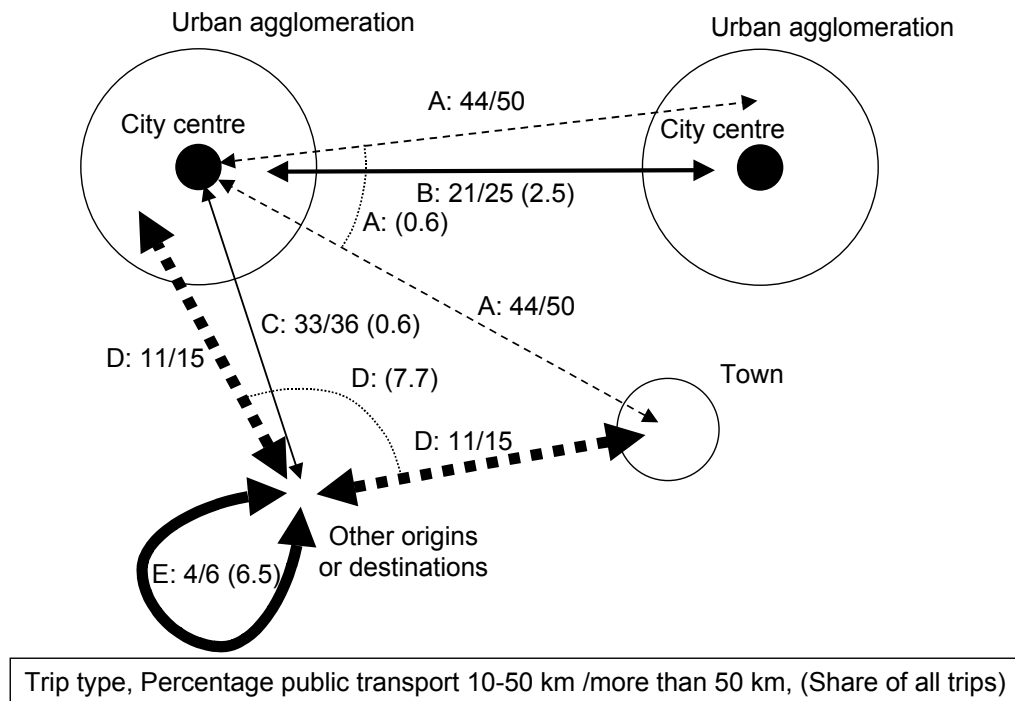
*Figure 2-13: Percentage public transport use for different trip types and share of all trips
(NTS 1995, Van Goeverden & Schoemaker (2000))*

This model is used to estimate the impact of the concept of multimodal transport. The difference in speed for the main mode will probably remain the same, but the transfer to an urban public transport system might be improved due to better synchronisation and information supply. Since access to the public transport system is a particular problem in regional areas, it might be expected that multimodal transport will reduce or even eliminate the difference between access by urban public transport and by regional public transport.

Given this line of reasoning two scenarios are defined. For the maximum scenario the reduction due to urban transfers is limited to 43% instead of 51% (an increase of 15%, just as in the assessment based on trip purpose) and it is assumed that there are no differences between regional and urban access. In the minimum scenario it is only assumed that the difference between regional and urban public transport will be halved. The results of this analysis are shown in Table 2-6. The maximum scenario leads to a growth of multimodal trips of 66% for the maximum scenario and 24% for the minimum scenario.

*Table 2-6: Potential for multimodal transport for two scenarios based on trip types*

| | Maximum scenario | Minimum scenario |
|---|---|---|
| **Scenario characteristics** | | |
| Improvement in urban transfers | Yes | No |
| Difference between urban and regional access | Eliminated | Halved |
| **Results** | | |
| **Growth of share of public transport per trip type** | | |
| A: City centre and another agglomeration or town | 0% | 0% |
| B: Between agglomerations | 4% | 0% |
| C: City centre and other origins and destinations | 12% | 6% |
| D: Agglomeration or towns and other origins and destinations | 14% | 5% |
| E: Between all other origins and destinations | 10% | 4% |
| **Total growth of share of public transport** | 10% | 4% |
| **Growth multimodal transport** | 66% | 24% |
| **Resulting share of multimodal transport** | 4.8% | 3.6% |

## 2.4.5   Conclusions

Comparison of both scenario approaches, based on trip purposes and based on trip types, to estimate the potential of multimodal mobility shows similar results. Three main scenarios can be distinguished. The minimum scenario leads to an increase of multimodal mobility of 15%, a medium scenario leads to an increase of about 25%, and a maximum scenario results in 5% multimodal travel, an increase of 70%.

The new multimodal trips are characterised by trip lengths between 10 and 100 kilometres and having an origin outside the main agglomerations or towns and a destination within agglomerations or towns. Since the share of public transport is lowest for trips between 10 and 30 kilometres, the impact of multimodal transport improvements will be the greatest for public transport systems serving these trip lengths. If it is assumed that train services are most suited to accommodate this growth, for instance due to the fact that train has a dedicated infrastructure, the maximum scenario leads to tripling the train trips for the medium trip lengths and to doubling the long distance train trips. Finally, given the importance of work and education trips in multimodal transport today, other trip purposes such as visiting and recreation will make up for the largest part of this increase.

Given the trip length of 10 kilometres and more, the main competing mode is private car. An increase of multimodal transport will thus lead to a reduction of trips by private car. Given a share of public transport of about 12% the estimated increases in public transport

usage will lead to a reduction of car trips of 2.3% for the minimum scenario, 4.6% for the medium scenario, and 12% for the maximum scenario, that is, for trips between 10 kilometres and 100 kilometres. Since multimodal transport is especially important for trips to the main cities and given the fact that car trips between 10 and 100 kilometres are dominant in congestion today, these percentages are high enough to expect a reduction of congestion for the medium and maximum scenarios. It should be noted, however, that given the importance of trip purposes such as visiting and recreation in the increase of multimodal transport, the impact for the peak hours might be less than suggested by these estimates.

## 2.5   Conclusions

This chapter focused on multimodal transport itself: the concept of multimodal transport, multimodal mobility today, and the future potential of multimodal mobility. The concept of multimodal transport is discussed using the layer-model for the transportation system. Since multimodal transport is mostly related with transport services, the description of this layer is extended by distinguishing the following elements:

- Transport service integrator;

- Transport service, consisting of service components and transport means;

- Transport means, which have to be provided for and which have to be operated.

The main characteristic of multimodal transport is that more than one transport service is used for a trip and that transfers between transport services are thus an essential element of multimodal transport. The presence of one or more transfers sets standards for the transport services used in a trip: the disutility of a transfer should be compensated for by the characteristics of main transport service used.

The current share of multimodal mobility in the Netherlands is small. Only 2.9% of all trips are made by multimodal transport. Most trips consist of one transfer only (72%). Train services are mostly used as the main transport service in a multimodal trip (59%). Private modes account for 17%. For the access and egress parts of the trip, private modes become more important: 65% for the home-based part of the trip and 35% for the activity-based leg, excluding walking. Finally, the differences in leg lengths between the main mode and the access and egress modes suggest a hierarchical relationship between the networks used. The higher level network, that is the network of the main mode, is suited for long distance travel, while the lower level network is used for access and egress having short leg lengths.

The main factors that determine multimodal trip making are:

- Trip length, multimodal transport is more suited for longer trips;

- Destination area type, multimodal trips are oriented to the four main cities and their centres;

- Trip purpose, more than 50% of the trips have work or education as a trip purpose.

For trips in which these factors are combined the share of multimodal mobility increases substantially. For interurban trips to and from the main cities the share of multimodal transport is about 20%. Multimodal transport thus already fills the roles that were expected. Multimodal transport concerns long distance trips between 10 and 100 kilometres and is especially important for trips to the main cities and their agglomerations.

Personal characteristics proved to have a limited effect on multimodal trip making, although the Students Public Transport Card substantially increased multimodal trip making. Furthermore, it was found that for multimodal travellers car availability drops from about 75% to less than 50%. However, car availability has no substantial influence on the choice between unimodal and multimodal transport either. Finally, an analysis of multimodal transport and tours showed an unexpected increase in multimodal trip making for more complex tours. It might be concluded that multimodal transport becomes a more relevant alternative if travellers have to plan their trip, either because of the trip frequency such as for trip purposes work and education or because of the complexity of the tour.

Two approaches are developed to assess the future potential of multimodal mobility, both leading to similar results. It is estimated that a policy aimed at facilitating multimodal transport might lead to an increase in multimodal trip making ranging between 13% and 70%. If this increase is accommodated by train services alone, it might lead to doubling or even tripling train usage. The expected reduction of car trips between 10 kilometres and 100 kilometres varies between 2.3% and 12%. Given the focus on trips to the main cities and on trips longer than 10 kilometres, this might lead to a reduction of congestion on the roads serving the main cities.

The findings on the characteristics determining multimodal mobility and the future potential of multimodal mobility influence the multimodal transport network design problem. The importance of trip length focuses the study on interurban networks, especially private transport networks and line-bound public transport service networks, while urban public transport service networks are clearly important access and egress modes. Demand responsive transport services such as taxi services seem less important as they have a minor share as access or egress modes. Furthermore, the small share of multimodal travel indicates that impact of multimodal transport on the use of private cars will be small, thus limiting the possible impact of multimodal transport on private transport network design. For line-bound public transport, however, substantial increases in usage might be expected, implying a substantial influence of multimodal transport on public transport service networks.

The consequence of these findings is that this thesis will adopt the notion of hierarchical transport networks in the analysis of the multimodal transport network design problem. The general notion of hierarchy and transport networks will be discussed in Chapter 4, while hierarchical transport network structures will be studied for interurban private transport networks (Chapter 5) and for both urban and interurban line-bound public

transport service networks (Chapter 6). Furthermore, attention will be given to the potential in a multimodal transport system of demand oriented transport services such as rental services and demand responsive transport (Chapter 7). First, however, the main characteristics of the multimodal transport network design problem will be discussed in the next chapter.

# Chapter 3

# MULTIMODAL TRANSPORT NETWORK DESIGN PROBLEM

## 3.1 Introduction

After having defined multimodal transport and assessing the future potential of multimodal mobility in Chapter 2, the second step in this thesis is the discussion of the multimodal transport network design problem itself. The purpose of this chapter is to introduce the network design problem in general and to establish a theoretical framework for the multimodal transport network design problem, which will be used in the remainder of this thesis.

The chapter consists of two parts. First, the topic of unimodal transport networks as such is discussed, including the main characteristics of the related transport network design problem. Topics examined are network types, network description, network design objectives, primary and secondary network characteristics, and transport network design approaches. The first part is concluded by the choice of network characteristics that will be used throughout this thesis, followed by a general formulation of the single-level transport network design problem. This general formulation is based on an economic perspective of transport network design and on a sound description of travel behaviour.

The second part focuses on multimodal transport and introduces two principles used in this thesis. First, it shows that the traditional network design problem discussed in the first part can be extended in two directions: hierarchical multilevel networks instead of single-level networks and multimodal networks instead of unimodal networks. Second, it is hypothesised that from the traveller's point of view there is a distinct relationship between the plausibility of switching modes and shifts between network levels. It seems more logical to switch modes if a switch is also made between network levels.

Following this hypothesis the formulation of the traditional network design problem is extended first to a multilevel network design problem, that is a unimodal network

consisting of hierarchically related network levels. Then, the implications for extending the unimodal multilevel network design problem to the multimodal multilevel transport network design problem are discussed. Finally, an overview of multimodal transport network design methods is given, which shows that the line of reasoning used in this thesis is a new approach.

The main characteristics of the analytical framework used in this thesis are summarised in Section 3.6.

## 3.2 Transport networks

### 3.2.1 Network categories

A transport network facilitates making a trip from an origin to a destination for a specific mode, and thus determines the characteristics of that trip. Following the discussion of the layer model presented in Section 2.2.1 transport networks can be found in both the transport service layer and the traffic service layer. There are thus two categories of transport networks that should be considered:

- Transport service networks, such as a bus service network or a train service network;

- Traffic service networks or physical networks, such as a road network or a railroad network.

A transport service network is, of course, always related to a physical network. A bus service network is based on a road network and a train service network on a rail network. From the traveller's point of view, however, it will only matter whether a train service can be used for his trip. A railway without train services is useless for the traveller. All constraints and possibilities arising from the rail network are already incorporated in the train services available. For private modes, on the other hand, the physical network will often determine whether a private mode is feasible for a specific trip. Both network categories determine all network characteristics that are relevant in transport network design.

### 3.2.2 Network characteristics

Network characteristics can be seen from two points of view: that of the network user (travellers) and that of the network investor or network operator. The main characteristics of any transportation network from the traveller's point of view are travel costs and travel time, with the latter determined by network characteristics such as space accessibility, time accessibility and network speed. These network characteristics can be described using the following definitions:

- *Space accessibility*: the number and distribution of access points where the traveller can enter and leave the network. Typical examples are bus stops, motorway ramps, and airports;

- *Time accessibility*: the distribution of opportunities per unit of time for the traveller to use the network. This characteristic is very common for public transport or airline services and can be described by timetables or service frequencies. For private transport the time accessibility is usually unlimited;

- *Network speed*: the average speed while travelling on the network, which is determined by the network structure and the design speed. Since speed is independent of the distance travelled, it is preferred to the perhaps more obvious alternative of travel time.

In this thesis travel quality will usually be described by travel time, thereby neglecting other characteristics such as travel costs or all other service related characteristics. Furthermore, transport networks determine primarily travel times, while the other characteristics have only an indirect relationship with the transport network itself.

It is evident that from the investor's or operator's perspective costs are the main network characteristic. The following costs can be distinguished:

- *Investment costs*, especially the costs of building the physical network, which is related to type of infrastructure and total length of the network;

- *Maintenance costs*, that is the costs for maintaining the quality of the infrastructure, which is, again, related to the length of the network;

- *Operating costs*, these costs are especially related to transport service networks and include costs such as operating the vehicles. These costs are determined by the length of the transport service network and the frequency with which these services are offered.

*Table 3-1: Relationship between network category and network characteristics*

| | Traveller | | | Investor or operator | | |
|---|---|---|---|---|---|---|
| | **Space accessibility** | **Time accessibility** | **Network speed** | **Invest-ment** | **Mainte-nance** | **Operation** |
| **Transport service network** | ▓ | ▓ | ▓ | (Vehicles) | (Vehicles) | ▓ |
| **Traffic service network** | ▓ | | ▓ | ▓ | ▓ | |

The characteristic that all these investor or operator costs have in common is the network density, which can be defined as the total network length per unit of area. Table 3-1 summarises the discussion of the relationship between the two main network categories and the network characteristics from the traveller's and the investor's or operator's point

of view respectively. It clearly shows that time accessibility and operating costs are strongly related to transport service networks and that investment and maintenance costs are mostly related to traffic service networks. It should be noted that in the case of transport service networks investment and maintenance costs may also exist, but these costs mostly relate to the fleet used to provide the services. In the case of public transport services having dedicated infrastructure, however, investment and maintenance costs of infrastructure also becomes an important phenomenon.

### 3.2.3   Network description

There are two main approaches to describing a transport network. Most common is to define a network as a set of nodes together with a set of links. Each link connects a pair of nodes. Usually there is a subset of nodes that are entry and exit nodes or access nodes, whereas the other nodes represent crossings where no access or egress is possible (see Figure 3-1). For public transport networks the representation includes public transport lines, that is a set of connected links and their nodes, and associated frequencies. This type of description is especially suited for transportation modelling and to describe all kinds of transportation networks found in practice.



9 access nodes          5 access nodes
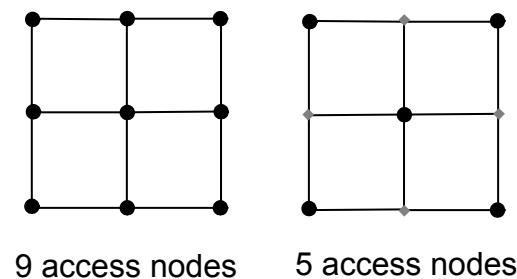
*Figure 3-1: Example of a network having different access densities*

The second approach looks at a network generically using specific network forms such as grid networks or radial networks, and aggregate network variables such as road spacing, line spacing, stop spacing, and frequency. These aggregate variables determine the main characteristics from both the traveller's and the investor's point of view. The level of detail with respect to topology of this second approach is clearly less than in the first approach, but it allows an analysis of the fundamental characteristics of networks. A typical example of such an extensive analysis can be found in Vaughan (1987).

The definition of these aggregate variables is strongly related to the network forms or morphological network structures that can be distinguished (Figure 3-2). Road spacing, for instance, is a suitable variable for network density in the case of a grid network or a triangular network, but not for a radial network. In that case the number of radials would be more appropriate. In the case of a radial/arc network or a circular network, arc spacing is a relevant variable. For line-bound transport networks, line spacing is more appropriate than road spacing. The relationships between these aggregate network variables and the main network characteristics for some common network types are given in Appendix C.
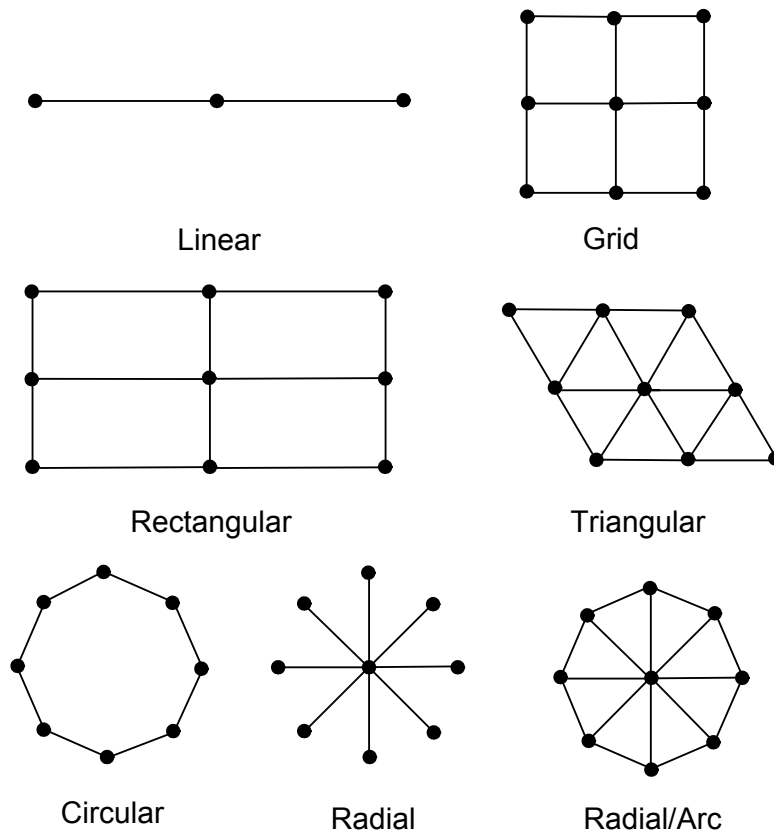
*Figure 3-2: Examples of network forms*

### 3.2.4   Second order network characteristics

As stated earlier, the first function of a transport network is facilitating trip making between origins and destinations. The main network characteristic for considering a transport service for a specific trip is the (expected) travel time, which is primarily determined by network characteristics such as space accessibility, time accessibility, and network speed. Apart from this line of reasoning, two other characteristics are often used as decision variables in transport network design: capacity and vehicle type or transport technology.

Capacity becomes an important issue when the transport network is not able to accommodate the demand while guaranteeing the expected quality. Usually quality is related to the expected travel time, but it is also related to crowding in buses or trains. The keywords here are 'expected quality' or 'expected travel time'. A transport network should facilitate trip making given a pre-defined quality of travelling or, more specifically related to the subject of transport network design, a pre-defined network speed. The capacity of the network or of its elements should be chosen such as to guarantee this quality and this network speed. Following this line of reasoning, capacity is in fact a second order characteristic of a transport network.

Vehicle type, or more generally transport technology, is the other prominent feature of transport systems. In many discussions, especially on public transport services, the

possibility of a new technologies dominates attention to the fundamental characteristics of a transport system, that is offering transport in order to participate in all kind of activities. The problem is similar to that of the red and blue buses in transportation mode choice modelling: adding a new alternative that is nearly identical except for the colour of the bus, or the transport technology that is used, does not influence mode choice substantially. Mode choice is primarily determined by travel time, and thus by the network characteristics space accessibility, time accessibility, and network speed. Transport technology can therefore also be seen as a second order characteristic.

## 3.3   Network design problem

### 3.3.1   Main characteristics

The question in transport network design is to determine a network that has an optimal performance given a specific design objective. Thus on the one hand there is a set of decision variables that determine the characteristics of the network, while on the other hand there is an objective against which the performance of the network is evaluated. Furthermore, there might be a set of constraints that limit the set of possible solutions.

The type of decision variables, and thus the characteristics of the design problem, depends on the method used to describe the network. The most common case found in literature is a description using nodes and links, while for public transport networks lines are also used as decision variables. The resulting transport network design problem can be illustrated using the following example. Given a set of access nodes, for instance, a set of 4 nodes, which allows 6 candidate links between these nodes. The question then is, which links should be included in the network. The decision variable is thus a binary variable determining whether or not a particular candidate link is included in the network. In order to connect all 4 nodes a minimum of 3 links is required. Therefore all networks consisting of 3, 4, 5, or 6 links are possible, which results in 42 possible networks. A similar line of reasoning for a set of 6 access nodes already results in 15 possible links and in more than 30.000 possible networks. From this numerical example it can easily be seen that the number of possible solutions increases more than exponentially with the size of the problem, which makes it a hard problem to solve. It has been shown that the network design problem in its simplest form is NP-complete, that is, no algorithm exists that can solve the network design problem in acceptable computation time, except of course for small networks (see e.g. Johnson et al. (1978)).

It is clear that if the decision variables may have more values than just being included in the network or not, for instance, accounting for the number of lanes that are available, then the combinatorial nature strongly increases. In the case of line-bound public transport services the impact on the combinatorial nature of the design problem is even larger: a new combinatorial aspect is introduced, namely that of all possible routes for each possible link network. Both extensions of the simple design problem discussed above substantially increase the complexity of the transport network design problem.

Using a more aggregate network description, however, greatly simplifies the network design problem. The decision variables then are limited to a few parameters such as network density, access density, network speed, and frequency, given a specific network type. The disadvantage is that due to the restriction to specific network types and the simplifying assumptions with respect to the demand pattern, the relationship with actual transport networks is limited. This drawback is perhaps the main reason that the discussion on transport network design in literature is dominated by the node- and link-based method to describe transport networks.

Before a general formulation of the network design problem to be used in this thesis is presented, three other aspects of the transport network design problem are discussed:

- The distinction in problem types;

- The additional complexity of the transport network design problem;

- The transport network design problem as discussed in the literature.

### 3.3.2   Problem types

Basically, there are two main types of the network design problem:

- *Designing a new network from scratch*. In this case a set of access nodes is given and the question is how these nodes should be connected with either transport services, such as bus lines, or with physical links, such as a road network;

- *Improving an existing network*. In this case not only is a set of access nodes given but also the existing network. The question then is whether transport services or links should be added to or removed from the network, or whether the capacity of existing services or links should be increased.

The main difference is the level of detail that is considered: transport services and network links only allowing the use of binary decision variables, or second order network characteristics such as capacity too, resulting in integer or real decision variables.

### 3.3.3   Additional complexity

Apart from the combinatorial nature, there are two other aspects that increase the complexity of the transport network design problem. The first follows from the conflict between the viewpoints of the traveller and the investor or operator. The traveller prefers direct connections between any origin and destination, and, if time accessibility also plays a role, at any time. The investor or operator favours a minimal network in space, and in time, thus reducing all cost factors. This dilemma is illustrated for road networks in Figure 3-3.

Travellers' optimum                           Investors' optimum
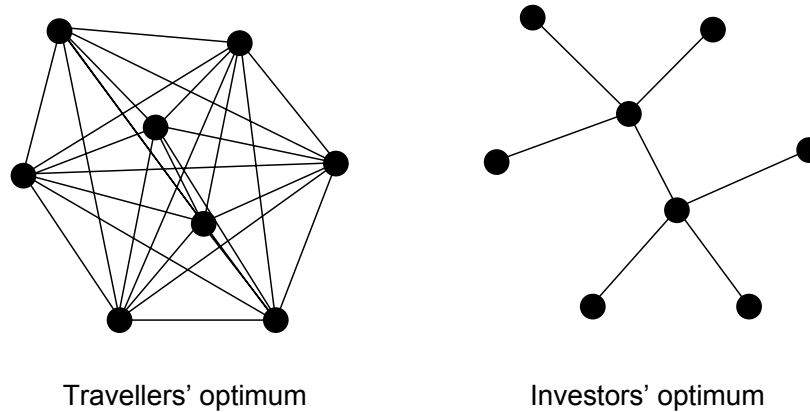
*Figure 3-3: Illustration of the difference in optimal network structures between the traveller's and the investor's point of view*

The design problem then is to compose a network that balances these two opposing objectives. There are three main approaches to find such a solution. The first approach formulates an objective that combines both opposing points of view. A common way is to choose the perspective of the authorities, who balance the interests of the travellers against those of the investors or the operators, for instance using economic principles such as social welfare or total societal costs. The second approach focuses on a single objective, usually the traveller's perspective, and includes the second perspective as a constraint, usually the available budget. The problem with this approach, however, is that it is unclear what this budget should be. The use of the budget as a constraint gives no insight into the trade-offs between investments costs and traveller's benefits. The third approach also chooses a specific objective, usually the investor's or operator's objective, and incorporates the behaviour of the other party involved, usually the travellers. A typical example of this approach is a public transport operator maximising profit while taking into account that offering inadequate services will reduce patronage and thus revenues. This also incorporates a trade-off between the operator's interests and those of the travellers. These three approaches are summarised in Table 3-2.

*Table 3-2: Three approaches to deal with opposing objectives*

| Approach | Formulation |
|---|---|
| **Combining objectives** | $O = g(P,T) + h(C_n)$ |
| **Focus on the traveller given a budget** | $O = g(P,T)$ <br> *sub* $C_n$ = given |
| **Focus on the investor or operator given travel behaviour** | $O = h(C_n(g(P,T)))$ |

*$O$: objective, $P$: demand, $T$: travel quality, $C_n$: network costs*

The second aspect that makes the transport network design problem a complex one, is the fact that travel behaviour and transport networks are strongly interrelated. Changes in the transport network lead to changes in travel behaviour. As such, the network design problem can be seen as a Stackelberg game in which one decision maker, that is the

network designer, has full knowledge of the decisions of the second decision maker, that is the traveller, and uses this knowledge to achieve his own objectives (Gibbons (1992), Cascetta (2001)). This principle is illustrated Figure 3-4.
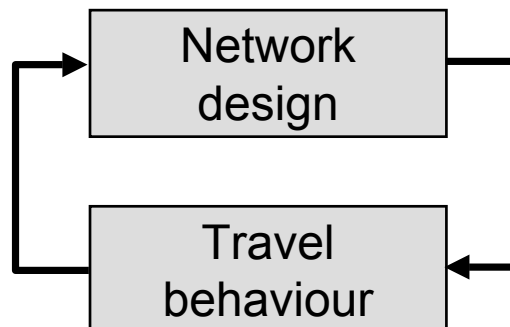


*Figure 3-4: Network design problem as a Stackelberg game*

In a Stackelberg game two problem types can be distinguished. The upper problem is the actual design objective in which the optimal network characteristics are determined given usage of the network by the travellers, while the lower problem describes travellers' behaviour given the network that is supplied. In this approach the network design problem is in fact the upper level problem. The lower problem usually deals with route choice while assuming a fixed level of demand, but it might also include other choices in travel behaviour such as mode choice and destination choice.

A typical example of this interaction between network design and travel behaviour is already shown in the third method for dealing with the opposing objectives of the traveller and the investor or operator. The transport network determines the quality of the transport services and thus the usage of the network. At the same time, if capacity is an issue the number of travellers using the network also influences the quality of the services. Travel demand can thus be assumed to be either fixed or dependent on travel quality (usually travel time), while travel quality (travel time) can be assumed to be dependent on the network only or on both the network and the level of demand. These possible relationships for the lower problem lead to four basic categories of the network design problem as shown in Table 3-3.

From a mathematical point of view there are several complications. First, the most realistic case is also the most complicated case, that is, demand depends on travel quality while travel quality depends on the level of demand. This is a fixed-point problem which is difficult to solve. Furthermore, for the formulation of travel times usually an equilibrium approach is used which for a detailed description lead to variational inequalities or using a slightly simplified approach to an optimisation problem. In the latter case, in which both the upper and the lower problem are written as optimisation problems, the network design problem can be classified as a bi-level optimisation problem. (Cascetta (2001)).

*Table 3-3: Basic categories of the network design problem*

|  | **Demand is fixed** | **Demand depends on travel quality** |
|---|---|---|
| **Travel quality depends on network only** | $P$=fixed $T$=g(*network*) | $P$=p($T$) $T$=g(*network*) |
| **Travel quality depends on network and level of demand** | $P$=fixed $T$=g(*network,P*) | $P$=p($T$) $T$=g(*network,P*) |

*P: demand, T: travel quality*

### 3.3.4  Literature on the network design problem

Given the complexity of the network design problem there is a huge amount of literature on this subject, from the fields of transportation science, mathematics, from an economic perspective, as well as from a societal perspective. The complexity is limited in various ways, for instance by adding constraints and thus limiting the number of possible solutions, or by limiting the relation between network supply and demand.

Furthermore, the fact that the transport network design problem is NP-complete leads to a special focus on efficient procedures to find (near-) optimal solutions. These procedures range from mathematically based methods such as branch-and-bound techniques to heuristic search approaches using for instance genetic algorithms (see for instance Magnanti & Wong (1984), Yang & Bell (1998), and Van Nes (1999)). In all these cases the design problem is formulated as a mathematical optimisation problem.

An alternative approach, also found in literature, is to develop design methodologies, that is, systematic stepwise procedures, thus avoiding the limitations of optimisation techniques in the case of NP-complete problems. A third approach focuses on the designer himself and uses artificial intelligence or knowledge-based techniques for support.

The following four tables give an overview of literature with respect to (for a more extensive discussion see Van Nes (1999)):

- Optimisation models for road network design (Table 3-4);

- Optimisation models for public transport service network design (Table 3-5);

- Design methodologies for transport networks (Table 3-6);

- Decision support systems (Table 3-7).

All tables cover the same characteristics of the network design problem:

- Network type: Service network (S) or physical network (P);

- Development type: New network (N) or network improvement (I);

- Objective type: Combined objectives (C), traveller oriented (T), or investor/ operator oriented (IO);

- Demand modelling: Dependent on network quality (Y) or fixed (N);

- Congestion modelling: Travel time dependent on level of demand (Y) or on network only (N).

Furthermore, these tables give a short description of the network that is designed, the objective, the design variables and the method used.

Optimisation models developed for road network design (Table 3-4) show a strong focus on network improvement while capacity constraints are modelled explicitly. This is mostly due to the fact that most of the road network already exists. Demand is usually kept fixed. The dilemma between traveller's and investor's interests is solved by combining both objectives in an objective from the authority's point of view, such as minimising total costs or maximising social welfare. Finally, the survey shows that a range of techniques is used to solve the optimisation problem. These optimisation models generally focus on single-level networks only, although Solanki et al. (1998) in order to tackle large-sized transport networks, divide the network into predefined network levels, which are optimised separately.

Optimisation models developed for public transport service networks (Table 3-5), on the other hand, focus entirely on designing new networks while capacity is rarely considered. Most of the design objectives are traveller oriented, with the exception of the design of an airline network. Furthermore, the design problem is often split into sub-problems, which are solved in a sequential order leading to stepwise procedures. Again, most models focus on single-level networks only. An exception can be found in Bouma & Oltrogge (1994) who distinguish different network levels a priori, and in models for airline networks (Aykin (1995), see also Section 3.5.2.2).

Compared to the optimisation models, design methodologies have less clear definitions of the objectives and the design variables (Table 3-6). Design methodologies are used for physical networks as well as for transport service networks. The approach followed by the Dutch Ministry of Public Works and Watermanagement (1998) even considers both network types, however, still as separate modes. Most methodologies focus on new networks and are traveller oriented. The level of demand and capacity are rarely considered. Another interesting difference compared to optimisation models is that many design methodologies distinguish different network levels. Furthermore, they use the concept of hierarchy for national and regional networks, following a top-down approach in the design process.

Decision Support Systems (Table 3-7) are all related to the design of single-level (urban) public transport service networks. In this case no explicit objective is defined. It is up to the planner to decide whether to focus on the traveller's or on the operator's interests. The main issue in this approach to the network design problem is to provide feedback on a specific network design, based on an assignment of the demand and general design rules.

A common finding in this survey is that most of the literature considers single-level networks only. In some cases different network levels are considered, but these are introduced as a decision variable. Solanki et al. (1998) use network levels to simplify their large-scale network design problem, while Bouma & Oltrogge (1994) make an a-priori distinction between network levels. Design methodologies on the other hand explicitly distinguish different network levels, however, these network levels are, again, an a-priori assumption.

*Table 3-4: Overview of optimisation models for road network design*

| Author(s) | Network type | Design problem type | Objective type | Demand modelling | Congestion modelling | Subject | Objective | Design variable | Method |
|---|---|---|---|---|---|---|---|---|---|
| **Steenbrink (1974)** | P | I | C | N | Y | Road network | Minimisation total costs | Link capacity | Capacity restraint |
| **LeBlanc & Boyce (1986)** | P | I | C | N | Y | Road network | Minimisation total costs? | Link capacity | Frank-Wolfe algorithm |
| **Boyce et al. (1990)** | P | I | C | Y | Y | Road network | Various objectives | Link tolls | Linear Programming Branch & Bound Simulated annealing |
| **Xiong & Schneider (1992, 1995)** | P | I | C | N | Y | Road network | Maximising fitness value (travel times and investment costs) | New links | Genetic algorithm |
| **Yang & Bell (1998)** | P | I | T C | Y | Y | Road network | Minimisation travel costs/ Maximisation consumer surplus/ multi-objective | Link capacity | Various techniques |
| **Solanki et al. (1998)** | P | I | T | N | N | Road network | Minimisation travel time given a budget | New links | Network partitioning and Branch and bound |
| **Lo & Tung (2001)** | P | I | T | N | Y | Road network | Minimisation unreliability | Link capacity | |

*Table 3-5: Overview of optimisation models for transport service networks*

| Author(s) | Network type | Design problem type | Objective type | Demand modelling | Congestion modelling | Subject | Objective | Design variable | Method |
|---|---|---|---|---|---|---|---|---|---|
| **Lampkin & Saalmans (1967)** | S | N | T | N | Y | Public transport | Minimisation travel time given fleetsize and vehicle size | Lines and frequencies | 4-step procedure<br>• Skeletons<br>• Lines<br>• Line selection<br>• Frequencies |
| **Hasselström (1979)** | S | N | T | Y | N | Public transport | Maximisation eliminated transfers given a budget | Lines and frequencies | 3-step procedure<br>• Link network<br>• Line generation<br>• Selection and frequencies |
| **Van Nes (1988)**<br>**Van Nes et al. (1988)** | S | N | T | Y | N | Public transport | Maximising travellers having no transfer given a budget | Lines and frequencies | 2-step procedure<br>• Line generation<br>• Selection and frequencies |
| **Ma & Schneider (1991)** | S P | N | C | N | N | Personal rapid transport | Minimisation total costs given a budget | Links | Heuristic procedure of adding and deleting links |
| **Bouma & Oltrogge (1994)** | S | N | T | N | N | Public transport | Maximisation travellers having no transfer | Lines and frequencies | Tree-search per sub-network |
| **Aykin (1995)** | S | N | IO | N | N | Air line network | Minimisation operator costs | Links and hubs | Branch & Bound Simulated annealing |
| **Ceder & Israeli (1998)** | S | N | C | N | N | Public transport | Minimisation travel time plus empty seat hours / minimisation of fleetsize | Lines and frequencies | 7-step procedure<br>• Line generation<br>• Path generation<br>• Line selection<br>• Assignment<br>• Frequencies<br>• Interchange<br>• Evaluation |
| **Bielli et al (2002)** | S | N | C | N | N | Public transport | Multi-criteria objective | Lines and frequencies | Genetic algorithm |

*Table 3-6: Overview of design methodologies for transport network design*

| Author(s) | Network type | Design problem type | Objective type | Demand modelling | Congestion modelling | Subject | Objective | Design variable | Method |
|---|---|---|---|---|---|---|---|---|---|
| **Schönharting & Pischner (1983)** | P | N | T | N | N | Road network | Maximum travel time constraints | Links and network levels | Hierarchical levels based on hierarchy of cities |
| **Immers et al. (1994), Immers & Egeter (1996)** | S | N | T | N | N | Public transport | Minimisation travel time | Links | Stepwise procedure<br>• Criteria<br>• Access nodes<br>• Link network<br>• Evaluation |
| **Kirchhoff et al. (1994, 1996)** | S | N | T | N | N | Public transport plus DRT | | Lines and transport service types | Hierarchical structure<br>• Radial lines<br>• Flexible routes<br>• Demand responsive transport system |
| **Franck et al. (1994)** | S | N | IO | N | N | Public transport | Cost efficiency constraint | Minimum settlement size<br><br>Network type | Network based approach |
| **Ministry of Transport, Public Works and Water-management (1998)** | P | N | T | N | N | Transport networks in general | Macro objectives<br>• Safety<br>• Welfare<br>• Economy<br>• Environ-ment | Links | Stepwise procedure<br>• Areas<br>• Relations/ trips<br>• Transport services<br>• Infrastructure |

*Table 3-7: Overview of Decision Support Systems for transport network design*

| Author(s) | Network type | Design problem type | Objective type | Demand modelling | Congestion modelling | Subject | Objective | Design variable | Method |
|---|---|---|---|---|---|---|---|---|---|
| **Janarthanan & Schneider (1988)** | S | N | C | N | N | Public transport | Multiple | Lines and frequencies | • Network design (manual)<br>• Assignment<br>• Feedback |
| **Baaj & Mahmassani (1991)** | S | N | C | N | N | Public transport | Multiple | Lines and frequencies | • Network design (computer aided)<br>• Assignment<br>• Line improvement |
| **Shih et al. (1998)** | S | N | C | N | N | Public transport | Multiple | Lines and frequencies | • Network design (computer aided)<br>• Transfer nodes<br>• Assignment<br>• Line improvement |

## 3.4  General formulation of the transport network design problem

The focus of this thesis is on the development of guidelines for multimodal transport network design. This section defines the approach for the transportation network design problem that will be used in the remainder of this thesis. Choices will be made with respect to problem type, design objectives, network description, and modelling travel behaviour. This section will conclude with a general formulation of the resulting design problem.

### 3.4.1  Design problem type

The topic of multimodal transport implies that different transport services should be considered in a similar systematic way, and thus that the approach should be suited for both service networks and physical networks. Since transport service networks are an essential part in multimodal transport the methodology should focus on the design of new networks.

### 3.4.2   Design objectives

In contrast to some of the design methodologies found in the literature (e.g. Immers et al. (1994)), an explicit objective in which the opposing objectives of both the traveller and the investor or operator are balanced is preferred. In that way an identical objective can be used for physical networks and for transport service networks. Typical examples of such objectives found in the literature, that are suitable for both types of networks are:

- Minimising total costs: that is minimising the sum of the costs involved in travelling, that is the total door-to-door travel time monetised using the value of time, plus the investments, maintenance and operating costs;

- Maximising social welfare: that is maximising the sum of consumer surplus and producer surplus. Consumer surplus consists of the benefits of all traveller who are able to travel at lower costs than they are willing to pay, while producer surplus is equivalent to profit (see e.g. Jansson (1996)).

The objective of maximising social welfare gives the most comprehensive description of the balance between the traveller's and the investor's objectives from an economic point of view (Berechman (1993), Yang & Bell (1997)). It incorporates the sensitivity of the demand for the changes in the service level that is supplied. This relationship between supply and demand, however, makes it also more complicated than the objective of minimising total cost in which a fixed level of demand is assumed. It can even be shown that combining a demand model with the objective of minimising total costs might lead to the trivial solution of offering no services at all, resulting in no travel costs and no investment, maintenance, or operational costs. In the case of urban public transport network design, it has been shown that both objectives yield similar outcomes for the resulting optimal designs (Van Nes (2000)). Finally, it should be noted that Berechman (1993) states that for long distance line-bound public transport profit maximisation might also be a suitable design objective. This option is considered in the analysis of the organisational aspects of hierarchical line-bound public transport networks (Section 6.5.3).

### 3.4.3   Network description

The type of network description used in this thesis is based on specific network forms and their aggregate network variables. This approach allows the development of analytical models which are suitable for acquiring insights into the main mechanisms of the network design problem and thus for developing design guidelines. The fact that such analytical models are based on simplifying assumptions, however, strengthens the importance of the interpretation of the results. Newell (1979), for instance, showed that even in a seemingly optimal situation for a grid network of public transport lines, the grid network was certainly not the optimal network configuration. For realistic networks such a theoretical network would surely be unsuitable. In this thesis though, the objective is not to advise on specific network types or optimal routings, but to analyse the impact of multimodal transport on transport network design. For that purpose, analytical models are certainly suitable to analyse and illustrate the main mechanisms involved.

The alternative approach of describing a network using a set of nodes, a set of links, and in the case of a public transport service network, a set of lines, obviously has the advantage that it can fit any topological situation found in reality. Furthermore, such a description allows a detailed description of travel behaviour. The crucial disadvantage, however, is that such an approach does not provide general insights that can be used in establishing guidelines for network design, unless a large number of cases are analysed.

The design variables in this study are thus the basic network variables for network density, access density, time accessibility, and speed (Section 3.2.3):

- Road spacing: the distance between parallel roads in a linear, grid or triangular network;

- Line spacing: the distance between parallel lines in a linear, grid or triangular network;

- Access spacing: the distance between access nodes for a road;

- Stop spacing: the distance between access nodes for a public transport line;

- Frequency: the number of runs per hour for a public transport line or network;

- Speed: the average speed in the network.

An explicit choice is made to include the service frequency as a decision variable in the transport network design problem. Cascetta (2001), in contrast, categorised frequency as a network performance variable comparable to capacity in road networks. Time accessibility, however, is a typical characteristic of many transport service networks, which, as will be shown in Chapter 6, has clear relationships with other network characteristics such as stop spacing and line spacing. Frequency should therefore be included from the start of the network design process.

The choice to focus the analysis on network forms and their basic network variables might also limit the influence of the demand pattern in space and time. In many network optimisation models the resulting network is strongly determined by the demand pattern used in the analysis. Although such a match between network and demand pattern is clearly desirable, it also raises the danger that the resulting network may not be robust to changes in the demand pattern. Especially in the case of physical networks, this might be a critical point in network design.

An illustration of this relationship between network forms and robustness can be found in Van Nes (1991). The optimisation model for public transport service networks developed by Van Nes (1988) was applied to the region of Utrecht in the Netherlands (AGV, 1989). In this study two optimal public transport service networks were designed. The first design was a tailor-made network for an assumed demand pattern, resulting in a diffuse network pattern. The second design used specific network forms, radial networks mostly, and a hierarchical network concept. For the overall network characteristics such as operational costs or total travel time, both network designs proved to be more or less equivalent, even though there were clear differences with respect to the number of transfers. A test of the robustness of these network designs with respect to the OD-matrix used in the design process showed that the performance of the first network design, that

is the diffuse network, dropped seriously, while the performance of the second network proved to be quite stable. The use of network forms thus led to a more robust network design. Another conclusion following from this analysis is that a network design model should explicitly account for the sensitivity of the demand for the quality of the network offered.

### 3.4.4 Description of travel behaviour

The transport network that is designed determines the travel times and influences travellers with respect to route choice, mode choice, destination choice, and even location choice. In the case of the design objectives maximising social welfare or maximising profit, the level of demand should depend on the quality of the services offered. In this thesis this will be modelled by incorporating mode-choice only, thus neglecting the influence of transport network quality on other transport related choices such as location choice and destination choice. In some cases when analysing line-bound public transport networks route choice will also be included in the analysis (Figure 3-5), for instance the choice between traditional public transport services and express services. For modelling these choice processes the framework of random utility theory is used (e.g. Ben-Akiva & Lerman (1985), Ortuzar & Willumsen (1994)). Incorporating both route choice and mode choice is modelled using the hierarchical logit model, or nested logit model, using the logsum of the route alternatives in the mode choice model. The details of this approach will be discussed in Section 6.4.3.1 for the choice between normal and express transport service, and in Section 7.3.1 for the case of alternative access modes.
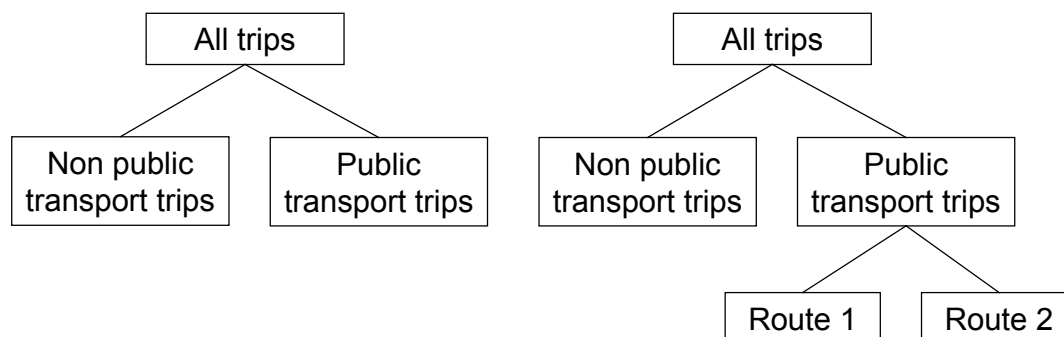


*Figure 3-5: Modelling travel demand using mode choice only (left-hand side) or including route choice as well (right-hand side)*

The choice between alternatives depends on the attractiveness of these alternatives, which is a function of travel time components, travel costs, and more subjective characteristics such as comfort. In this thesis the focus will be primarily on travel time, assuming all other factors to remain equal. Travel time is fully determined by the characteristics of the network and appears to be the primary descriptive variable in many transportation models (see. e.g. Manheim (1984), Hague Consulting Group (1997)). Using a normative approach with respect to travel behaviour characteristics Van Goeverden et al. (1998) established that travel time is an essential attribute for more than 80% of all trips. Travel times are determined excluding the influence of capacity. As will be shown in the following section, the multimodal transport network design problem is so

complex that this simplification is necessary. Furthermore, this choice is also in line with the statement that capacity is considered a second order network characteristic (Section 3.2.4): the main issue is to determine the required quality of the transport networks, while capacities should be chosen to guarantee that quality.

The analytical transport network design models developed and used in this thesis can thus be classified as the categories in the first row of Table 3-3: in the case of minimising total costs the first category results, while the case of welfare maximisation yields the second category.

### 3.4.5    General formulation

Following the classification of the network design approaches, the requirements for the proposed design methodology can be defined as:

- Suited for both service networks and physical networks;

- Focusing on the design of new networks;

- Using a combined objective of traveller's, investor's, and operator's interests;

- Modelling the level of demand using mode choice (maximising social welfare) or assuming a fixed level of demand (minimising total costs);

- No modelling of capacity constraints.

Given these choices the following design problem for the unimodal case can be formulated (the multimodal case is considered in the next section). The objective function considered is maximising social welfare ($SW$). The design variables for a specific network type ($N_t$) are space accessibility ($A_s$), time accessibility ($A_t$), network speed ($V_n$), and network density ($D_n$). In order to formulate the objective function a set of intermediate variables has to be defined. The design variables determine access and egress time ($T_a$), waiting time ($T_w$), and time spent on the network for a trip having length $L$ ($T_n$) (including transfers if relevant), which make up total travel time ($T_t$). Furthermore, these variables determine investment costs, maintenance costs and operational costs, which make up the total network costs ($C_n$). Since travel time is considered to be the main determinant in mode choice, travel time determines patronage ($P$), which determines the revenues ($R$). Please note that fares are no decision variables, and are assumed to be fixed. Given the level of demand the benefits for the traveller (consumer surplus ($CS$)) and those for the investor or operator can be determined (producer surplus ($PS$)), which together define the level of social welfare ($SW$). These relationships are illustrated in Figure 3-6.
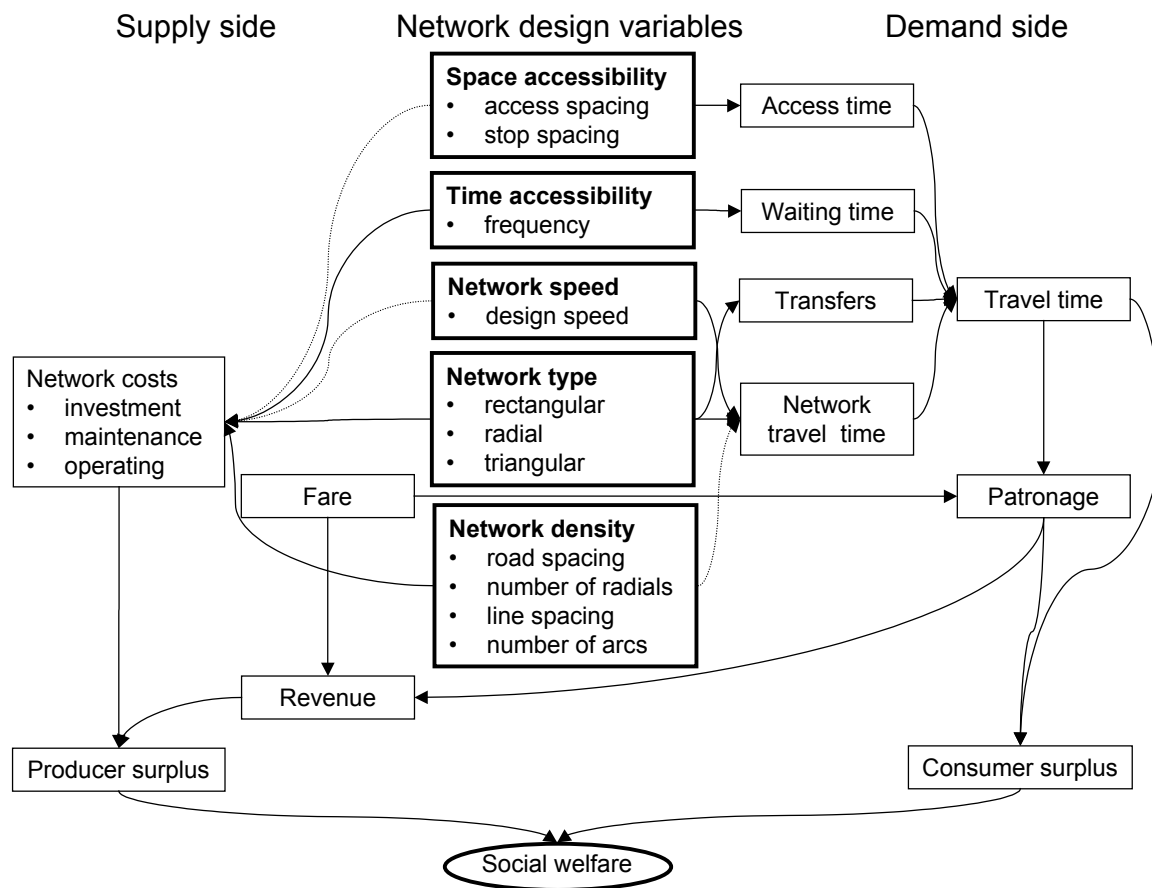
*Figure 3-6: Conceptual model of the relationships between the basic network variables and the objective of maximising social welfare*

Four remarks need to be made with respect to this scheme:

- Waiting time is only relevant if time accessibility is an issue such as for public transport service networks;

- Space accessibility is considered to be especially relevant for the network costs, that is mainly investment costs, if substantial costs are involved with providing access nodes, for instance, in the case of metro systems;

- Network speed has an influence on network costs, in the case that the required network speed implies grade-separated crossings, or in the case of transport services having operational costs only;

- Network density has an implicit relationship with travel time, depending on the way network density is defined for a specific network type.

For the objective of minimising total costs given a fixed level of demand, Figure 3-6 is slightly altered. The relationship between travel time and patronage is skipped and the items revenue and producer surplus are dropped. Consumer surplus is replaced by traveller's costs ($C_t$), and social welfare, of course, by total costs ($C$).

The relationships shown in Figure 3-6 are formulated mathematically as functions of the design variables and intermediate variables as in Table 3-8 (for the symbols used in this thesis see the List of symbols). The analyses in the following chapters will discuss these functions in more detail. Section 6.3.1, for example, gives an extensive presentation of the formulation for the urban public transport network design problem.

*Table 3-8: Mathematical representation of the single-level network design problem*

| Item | Type | Function |
|------|------|----------|
| *Objective maximising social welfare* | | |
| Access time | LOS | $T_a = f_a(A_s)$ |
| Waiting time | LOS | $T_w = f_w(A_t)$ |
| Time spend on the network | LOS | $T_n = f_n(L, N_t, D_n, V_n)$ |
| Total travel time | LOS | $T_t = f_t(T_a, T_w, T_n)$ |
| Patronage | Demand | $P = f_p(T_t)$ |
| Revenues | Demand | $R = f_r(P)$ |
| Consumer surplus | Demand | $CS = f_{cs}(P, T_t)$ |
| Network costs | Supply | $C_n = f_n(D_n, N_t, A_t, A_s, V_n)$ |
| Producer surplus | Supply | $PS = f_{ps}(R, C_n)$ |
| Social welfare | Objective | $SW = f_{sw}(CS, PS)$ |
| *Objective minimising total costs* | | |
| Patronage | Demand | $P$ = constant |
| Travel costs | Demand | $C_t = f_{ct}(P, T_t)$ |
| Total costs | Objective | $C = f_c(C_n, C_t)$ |

*LOS: Level of service*

## 3.5   Multimodal transport network design problem

### 3.5.1   Main characteristics

The discussion of the transport network design problem in the previous section mostly concerned unimodal transport network design. The main part of the literature on transport network design focuses on single modes or transport services and on single-level networks. Typical examples are urban or national road networks, and urban or national public transport service networks. A multimodal transport system, however, consists of several transport services and different network levels, which might also have different operators or authorities. A multimodal trip might consist of an access leg in which private car is used in an urban road network, and a main leg using an interregional train service on the national railroad network. The multimodal transport network problem thus adds at least two extra dimensions:

- Network levels;

- Combinations of transport services.

The third dimension of different operators or authorities is analysed separately in Section 6.5.3.

The common multilevel network configuration is illustrated in Figure 3-7. The highest network level, level 3, is characterised by a coarse network, limited accessibility, and high speeds, and is especially suited for long distance trips. The lowest network level, on the other hand, is fine-grained, has high accessibility and low speeds, making is suitable for short distance trips and for accessing the higher network levels.
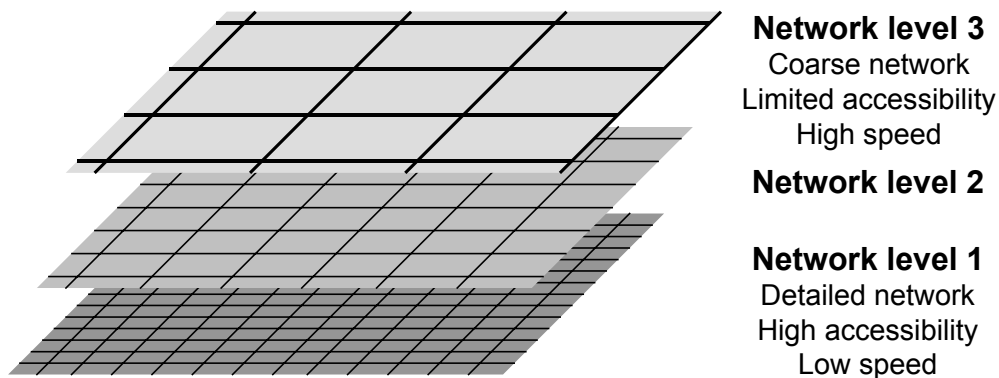


*Figure 3-7: Illustration of multilevel network*

The extension of the transport network design problem to the multimodal, multilevel network design problem is illustrated in Figure 3-8. Horizontally there are the transport services with a distinction between private and public modes. Examples of the private modes are private car (motorised vehicles), bicycle (human powered vehicles), and walking (no vehicles are used), while for the public modes a distinction is made between line-bound services and demand oriented services such as share-a-ride concepts. On the vertical axis are the various network levels that might be distinguished, which are strongly related to the distance travelled. For each transport service the grey area shows the distance range for which it might be suited. For many transport services different network levels might be distinguished. The dashed lines and the vertical arrows illustrate the boundaries between network levels. Typical examples for car networks are motorways, regional roads, urban motorways, and streets. The traditional transport network design problem usually deals with a single rectangle in this figure, for instance, an urban public transport service network.
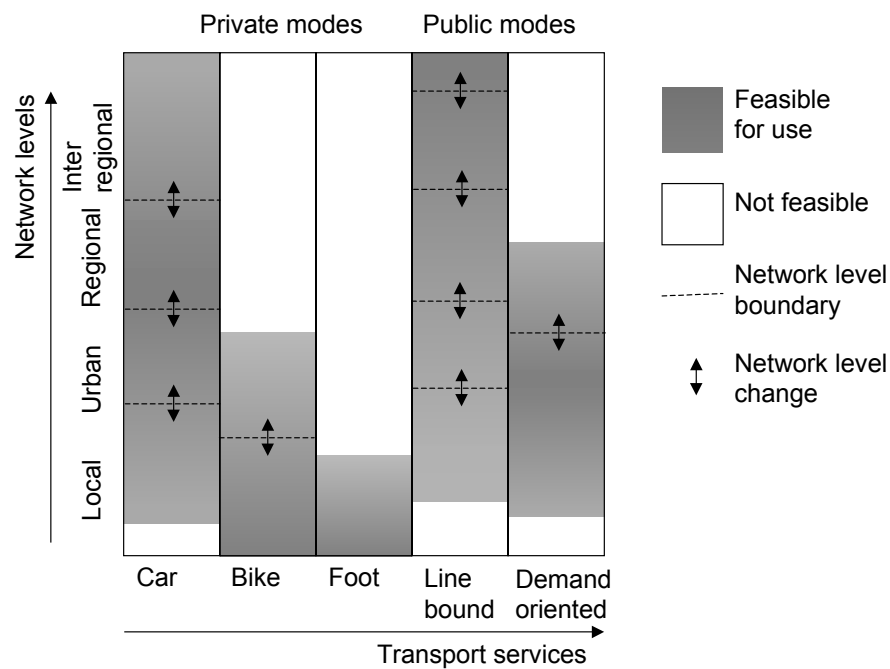
*Figure 3-8: Multimodal transport system consisting of modes and network levels*

Multilevel transport networks and multimodal transport networks are strongly related. A mode used to access another mode which is suited for the specific trip introduces a hierarchical relationship between these two modes and thus between the network levels that are used. The notion of hierarchy implies that a transport network, apart from having its own function, also provides access to higher level networks. Lower level networks support higher level networks. Consequently, multimodal transport implies multilevel transport networks. Multilevel networks on the other hand, are not necessarily multimodal networks. The road network, for instance, is clearly a hierarchical network having different network levels that are suited for specific trip lengths, while it is unimodal transport network because there is no need for the traveller to make a transfer: the same mode is used for all network levels.
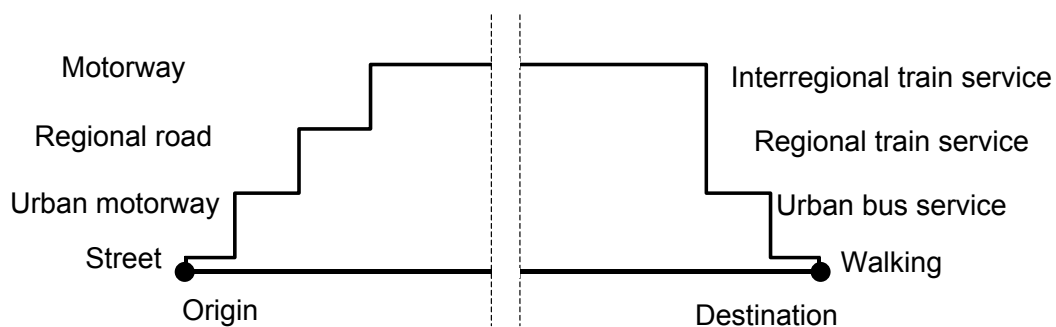


*Figure 3-9: Example of pyramidal route choice for private car (left-hand side) and line-bound public transport services (right-hand side)*

The notion of multilevel transport networks is also suitable for describing travel behaviour. Analyses by Bovy (1981, 1985) showed that the hierarchy in private transport networks enables a good description of route choice for cyclists and car drivers in cities.

The explicit use of the concept of hierarchical transport systems in route choice is called pyramidal route choice. Figure 3-9 gives an example of the network levels used for a long distance trip by private car and for line-bound public transport services. In a multimodal transport system, of course, private car might be used to access the interregional train service, yielding a similar representation!

The multilevel network concept introduces a second relationship between network levels. The lower-level network is used to access the higher-level network. The quality of the lower-level network thus determines the quality of the higher-level network. Furthermore, since there are travellers using both network levels, the quality of the higher-level network influences the patronage of the lower network. These additional relationships make the multilevel transport network design problem even more complicated than the single-level transport network design problem.

The focus of the literature on single-level unimodal transport network does not mean that no attention is paid to other transport services. In project PI (Dutch Ministry of Public Works and Watermanagement (1998)), for instance, both private and public modes are considered simultaneously, however, still as separate modes. In many studies in which the demand is assumed to be dependent on the quality of the services offered, mode-choice is used to describe this relationship, usually the choice between private car and public transport services. The various transport services are, again, analysed as separate modes.

In this more traditional unimodal context it is reasonable to assume that the relationship between network levels are transport service specific. The multimodal transport network design problem, however, introduces a dependency between different transport services. It is no longer necessary to assume that relationships for multilevel networks apply only for a specific transport service. On the contrary, other transport services might be essential to access higher-level transport networks. The quality of a network level of another transport service might thus be decisive for the characteristics of the higher-level network.

A typical question then is what will happen with the hierarchy of the transport service: will the characteristics of the network levels change or will they remain the same? What is, for instance, the impact on the hierarchy of line-bound public transport service networks if cycling replaces the role of walking? Will it change the characteristics of the lower-level network only, or will it influence all network levels of line-bound public transport services? Figure 3-10 schematically illustrates these possibilities. The left-hand side displays the network levels as shown Figure 3-8. The right-hand side shows two alternatives. In the first case only the characteristics of the lower-level network change, namely it is less suitable for short travel distances, while those for all other network levels remain identical. In the second case, however, the characteristics of all network levels change.
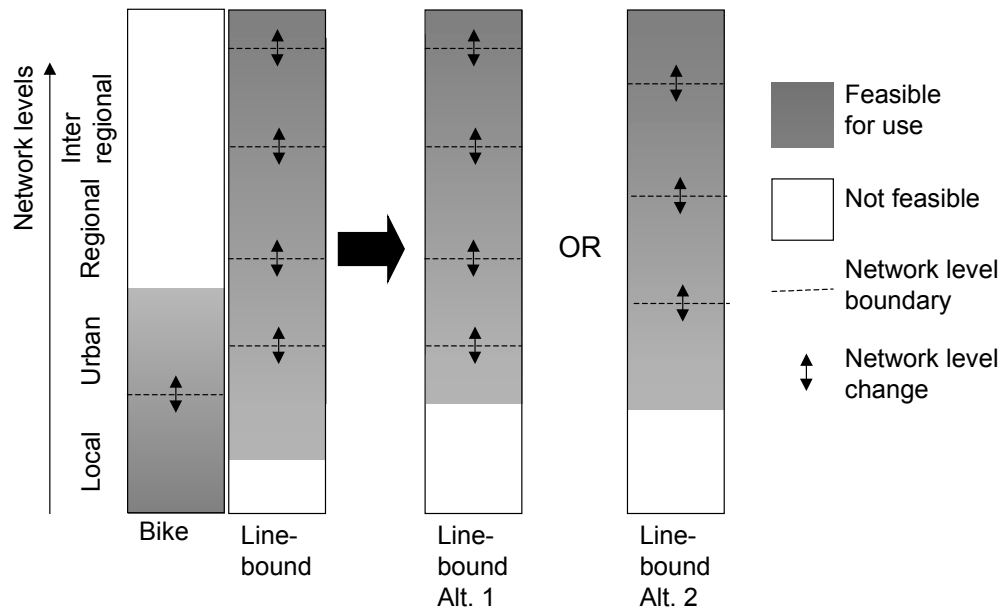
*Figure 3-10: Two possibilities for the hierarchy of line-bound public transport service networks because of combining bicycle and public transport:*
*Alternative 1: only the lower-level network characteristics change*
*Alternative 2: all network level characteristics change*

A third characteristic of the multimodal transport network design problem, which was already mentioned in Section 3.4, is that it deals with both transport service networks and physical transport networks. This is one of the reasons for choosing minimising total costs and maximising social welfare as design objectives. The combination of both network types, however, also adds to the complexity of the design problem. Service networks require physical networks. New links in the road network, for instance, enable new public transport services by bus, which in turn might even reduce the need for building those links. Combining transport services thus leads to additional combinatorial complexity.

It is clear that the complexity of the multimodal transport network design problem is substantially larger than that of the traditional network design problem. How to deal with this complexity? The key issue in this discussion is the interdependency between network levels for different transport services. The key to dealing with this interdependency can be found in travel behaviour. Multimodal transport will only be relevant if it is interesting for the traveller. Multimodal trips should be attractive with respect to travel time and travel costs. Furthermore, the concept of pyramidal route-choice requires that multimodal trips should have a logical pattern regarding the access leg, the main leg, and the egress leg. It might thus be stated that transfers between transport services are only plausible if they coincide with changes in network levels.

Given this assumption a two step procedure is used to develop a design methodology for multimodal transport networks. The first step is to analyse the relationships that determine the network level boundaries for a transport service. In this analysis a distinction is made between private transport networks (Chapter 5), in which physical networks are dominant, and line-bound public transport service networks (Chapter 6),

concentrating on service network design. The second step focuses on the consequences of combining transport services for these relationships between network levels (Chapter 7). In this way the multimodal transport network design problem is split up into two parts: unimodal multilevel networks and multimodal multilevel networks. Furthermore, the complexity following from detailed demand modelling, using mode- and route-choice in a multimodal transport network, is not considered in this thesis. The main goal is to determine basic relationships for transport network design, which are needed for the development of design guidelines.

### 3.5.2   The multilevel transport network design problem

#### 3.5.2.1   *Mathematical formulation of the multilevel network design problem*

The mathematical representation of the single-level transport network design problem presented in Section 3.4 can easily be extended to a multilevel network design problem by introducing two types of relationships between networks. The first relationship is that the characteristics of the lower-level network determine the access time to the higher-level network. If, for instance, two network levels are considered, having index 1 for the lower-level network and 2 for the upper level network, the mathematical formulation of the access and egress times changes for the upper level network. The space accessibility of network 2 ($A_{s,2}$) determines the trip length $L$ for network 1, which determines the final access time for network 2 (Figure 3-11).



Service area network 1
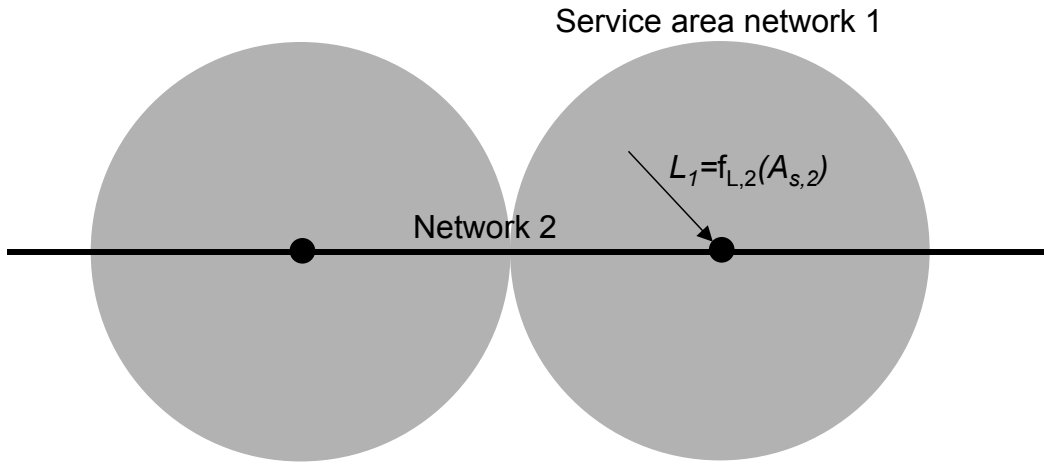
$L_1 = f_{L,2}(A_{s,2})$

Network 2

*Figure 3-11: Hierarchical concept in which network 1 provides access for network 2*

Mathematically, this relationship can be written as (see also Table 3-8):

$$T_{a,2} = f_{a,2}(A_{s,2}) = T_{t,1}\left(T_{a,1}, T_{w,1}, f_{n,1}\left(f_{L,2}(A_{s,2}), N_{t,1}, D_{n,1}, V_{n,1}\right)\right)$$   (3-1)

where:

$f_{L,2}$ = access distance as a function of the space accessibility of network 2

Furthermore, a distinction needs to be made between travellers using the lower-level network only, travellers who use the lower-level network to access the upper level network, and travellers using the upper level network only, for instance accessing it by other modes such as on foot. Each of these three categories has its own functions for the level of demand and the consumer surplus. The design objectives of maximising social welfare or minimising total costs are therefore functions of these three populations. This leads to changes in the formulae as shown in Table 3-9. All other equations remain identical to those in Table 3-8, except for adding an index for the network level. This description of the multilevel network design problem is used in Section 6.5.3.

*Table 3-9: Extended mathematical formulations for the two level hierarchical network design problem (see also Table 3-8)*

| Item | Type | Function |
|---|---|---|
| *Objective maximising social welfare* | | |
| Patronage population using level 1 | Demand | $P_1 = f_{p,1}(T_{t,1})$ |
| Patronage population using level 2 | Demand | $P_2 = f_{p,2}(T_{t,2})$ |
| Patronage population using both levels | Demand | $P_{12} = f_{p,12}(T_{t,1}, T_{t,1})$ |
| Revenues level 1 | Demand | $R_1 = f_{r,1}(P_1, P_{12})$ |
| Revenues level 2 | Demand | $R_2 = f_{r,2}(P_2, P_{12})$ |
| Consumer surplus | Demand | $CS = f_{cs}(P_1, P_2, P_{12}, T_{,t,1}, T_{t,2})$ |
| Producer surplus | Supply | $PS = f_{ps}(R_1, R_2, C_{n,1}, C_{n,2})$ |
| *Objective minimising total costs* | | |
| Patronage population using level 1 | Demand | $P_1 = \text{constant}_1$ |
| Patronage population using level 2 | Demand | $P_2 = \text{constant}_2$ |
| Patronage population using both levels | Demand | $P_{12} = \text{constant}_{12}$ |
| Travel costs | Demand | $C_t = f_{ct}(P_1, P_2, P_{12}, T_{t,1}, T_{t,2})$ |

### 3.5.2.2  *Multilevel network design in the literature*

While network design models focus on single-level transport networks, multilevel networks are often considered in design methodologies. Schönharting & Pischner (1983) distinguish different network levels for road networks based on a hierarchy of cities. Immers et al. (1994), Kirchhoff et al. (1994, 1996), and Immers & Egeter (1996) define different network levels for public transport service networks. There is, however, no specific argument for their classification. As a result other classifications, for instance as proposed by Van den Heuvel (1999) are just as reasonable.

Optimisation models usually focus on single-level networks. Exceptions can be found in literature on airline-network design. In these studies a two-layered network structure is assumed, in which the higher-level network connects the major airports or the hubs and

the lower-level network connects the other airports with those hubs (the spokes). A typical example of such a hub & spoke network is shown in Figure 3-12.
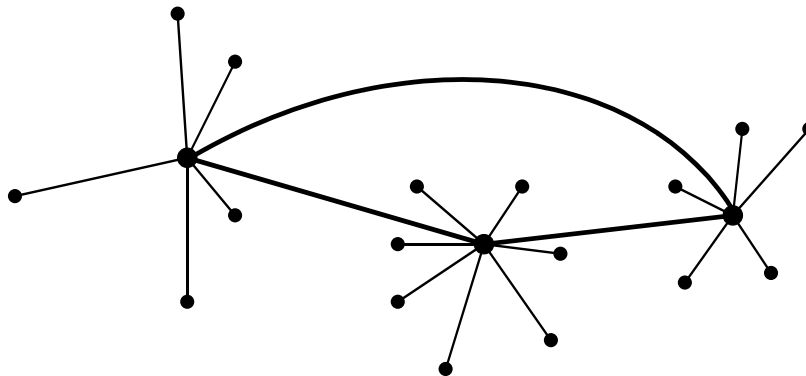


*Figure 3-12: Example of a hub & spoke network*

Airline networks are different from other transport service networks as the infrastructure needed consists only of nodes. Furthermore, the airline industry is a fully private market, which is characterised by fierce competition. As a result, the design objective is dominated by maximising profit or by minimising operational costs. It is clear that this approach does not fully account for the traveller's point of view. Given the fact that there are usually no alternative competing transport services except other airline services, the airline industry is able to focus on operational costs. In the long run, however, this might change. Aykin (1995), for instance, showed that airline networks offering more direct connections than a strict hub & spoke network might lead to lower operational costs while the percentage of travellers having a direct connections increases from about 15% to 57%, and in one case even to 74%! It is interesting to note that recently Boeing decided to develop a faster aeroplane, arguing that such an aeroplane would enable the airline industry to meet the traveller's wish to have more direct connections. This is exactly the opposite of the strategy of Airbus, which is developing a larger aeroplane in order to enable a cost reduction per seat for inter-hub connections.

Another point in airline network design is that when designing a hub & spoke network the difference in operational costs between the spokes and the inter-hub connections is usually assumed beforehand. O'Kelly & Brian (1998) state that this approach leads to an underestimation of the total operational costs.

It can be concluded that the multilevel concept in airline networks today is strongly dominated by operational costs. Furthermore, airlines are free to choose which connections they will offer. As a result it is questionable whether the concepts of airline networks are representative for multilevel service networks. The subject of multilevel transport service networks will be discussed further in Chapter 6.

Another exception found in literature is the model for road network design by Solanki et al. (1998). Given the NP-completeness of the network design problem, it is an interesting option to limit the size of the design problem. Solanki et al. divide the network in hierarchical clusters, each having its own budget. The networks for these clusters are designed in a top-down fashion in which virtual links are used for the lower-level

network clusters, and where the higher-level network clusters determine the entry and exit nodes. In this way the design problem for a large transport network is split into a set of smaller network design problems. Solanki et al, however, did not pay specific attention to the criteria to distinguish different network levels.

The overall conclusion is that usually a kind of hierarchy between network levels is assumed, but there is limited insight into the relationships that determine the hierarchy in network levels.

### 3.5.3   The multimodal multilevel transport network design problem

Given the formulation of the multilevel network design problem it seems easy to make the extension to the complete multimodal multilevel design problem, just add an additional index *m* for the transport service at each network level distinguished in Table 3-9. The introduction of extra transport services, however, also introduces a mode-choice problem: which of the travellers uses each transport service to access the higher-level network? A similar problem occurs if several higher-level networks are available. It is therefore no longer possible to distinguish different populations having their own relationship between supply and demand.

The literature on multimodal transport networks shows three ways to avoid this problem:

- Focusing on *specific combinations of modes* such as, for instance, car and public transport in the concept of transferia (Egeter et al. (1990), Van Binsbergen et al. (1992)), or the combination of public transport and bicycle (Nägele (1992), Van Goeverden & Egeter (1993));

- Focusing on *specific cases*, such as a corridor between two cities (Mu-consult (1996)) or a transferium for transferring from car to train (Mu-Consult (1998)). In these cases the existing networks were used as constraints and the focus was, again, on the potential of multimodal transport;

- Transforming the design problem into a *single-level network design problem* by defining a hypernetwork or supernetwork consisting of links for each mode plus transfer links for switching between modes (see Figure 3-13). A review of the hypernetwork or supernetwork approach can be found in Catalano et al. (2001). In a supernetwork the distinction between mode-choice and route-choice is replaced by a route-choice problem only. The transfer links thus not only stand for the physical transfer between modes, but also represent mode-choice characteristics. The supernetwork approach was suggested by Sheffi (1985). It was used for freight transport (Rutten (1995), Loreiro & Ralston (1996)), and recently for passenger transport (Ferrari (1999)). Given such a supernetwork formulation well-known network optimisation techniques are used to design the optimal network. These studies are usually limited to a single network level only: national or international freight transport or the urban transportation system.
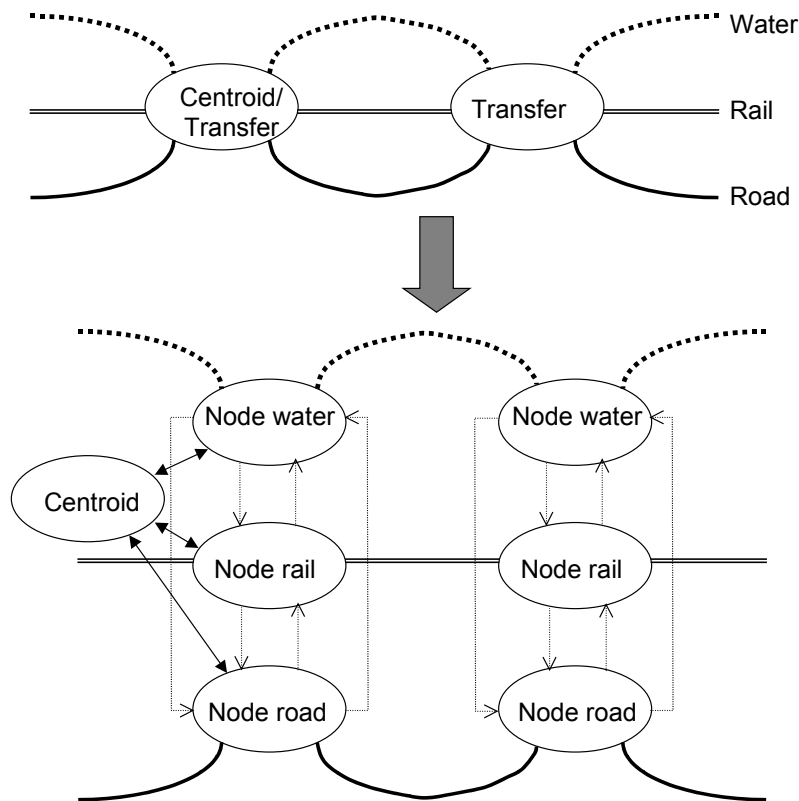
*Figure 3-13: Hypernetwork structure used by Loreiro & Ralston (1996)*

All these three approaches have in common that they do not fully consider the multimodal multilevel transport network design problem. The approach used by Van Binsbergen et al. (1992), however, gives some interesting insights. First, it is the only design methodology in this review, while all other studies are either optimisation models or use a pragmatic approach. Their design methodology is a top down approach consisting of four steps.

- Selection of destinations having good access by public transport and having sufficient transport demand;

- Selection of corridors having sufficient demand;

- Locating transferia and determining the expected number of travellers;

- Assessing the impact on accessibility and the environment.

Second, Van Binsbergen et al. not only determined where transferia should be located, but they also made some qualifications regarding existing infrastructure and transport services. Additional fast public transport services were necessary and some existing P&R-facilities were not suitable for transferia. This implicit evaluation is mainly thanks to the top-down approach in which the choice of the main destinations and the selections of corridors already lead, although implicitly, to the use of network types. It is the objective of this thesis to develop design guidelines that includes such design principles explicitly.

The literature on multimodal transport networks does not provide an approach that tackles the problem introduced by considering more transport services simultaneously. It might be questioned, however, whether it is necessary for a design methodology to include all dependencies. An important factor here is the type of relationships between network levels, which are analysed in the coming chapters. Given these relationships it will be possible to decide in which way considering more transport services simultaneously might be incorporated in the formulation of the multimodal transport network design problem. This will be discussed further in Chapter 7.

## 3.6  Conclusions

This chapter elaborated on the subject of the transport network design in general and the typical characteristics of multimodal transport network design in particular.

Simple transport network design was shown to be already a complicated problem, due to conflicts between the objectives of the traveller and those of the investor or operator, the dependency between supply and demand, and the combinatorial characteristics of the problem. The multimodal concept adds an additional complexity to this problem, namely the introduction of hierarchical levels within networks and the possibility of a functional replacement of a network level within a transport service by networks of other transport services. The interaction of network levels and transport services in a truly multimodal transport system suggests a dependency between networks that has not been analysed before.

This thesis will develop an approach to multimodal transport network design that takes these interactions into account. The conclusions following from this chapter provide the theoretical framework for the multimodal transport network design problem that will be used in the remainder of this thesis. This framework is suitable for both physical networks and transport service networks, and is based on an economic perspective on transport network design. General formulations have been presented for describing level of service, demand-side and supply-side elements, and objective functions for transport network design.

Objectives that will be used in the analyses are maximising social welfare and minimising total costs. These objectives are suitable for both physical networks and service networks, and they incorporate all relevant characteristics from the traveller's as well as the investor's or operator's point of view. Only in special cases will the objective of maximising profit be used.

In the case of the objective of maximising social welfare, the influence of the service quality of the transport network on the level of demand is described using mode-choice models based on random utility theory. In some cases will hierarchical logit or nested logit models will be used to account for lower level choices such as route-choice between different types of public transport services or access mode-choice for accessing public transport service networks. In the case of the objective of minimising total costs, it is assumed that the level of demand is fixed.

In order to deal adequately with the additional complexity inherent to the multimodal transport network design problem, it will be assumed that travel quality depends on network characteristics only. Characteristics such as capacity and transport technology are considered to be second order network characteristics that depend on the required quality and the actual situation with respect to level of demand and available infrastructure.

Furthermore, the analyses will adopt an aggregate description of networks, based on network types and aggregate network characteristics, such as network density, access density, network speed, and frequency or time accessibility. These characteristics will be defined using the design variables road spacing, line spacing, access spacing, stop spacing, speed, and frequency. This approach enables the development of analytical transport network design models, which consist of models describing the level of service, as well as demand and supply components. These network design models will be used to establish basic relationships for the network design variables.

Finally, a two step approach is proposed to analyse the multimodal transport network design problem. First, the impact of hierarchy within a network of a single transport service is analysed. Chapter 4 gives a general discussion on the subject of hierarchy, which is followed by detailed analyses for private transport networks (Chapter 5) and line-bound public transport service networks (Chapter 6). Second, the impact of combining different transport services is considered in Chapter 7, with respect to the characteristics of the networks that are combined as well as to the influence on the hierarchy of the transport services involved.

# Chapter 4

# HIERARCHY AS A BASIC PRINCIPLE

## 4.1 Introduction

Both the analysis of multimodal mobility today (Section 2.3) and the discussion on the multimodal transport network design problem (Section 3.5), suggest that the concept of hierarchical networks is a natural way to deal with typical characteristics of multimodal transport such as access, main modes, and egress. In multimodal travel changing modes or service types might be compared to changing network levels.

A hierarchical transport network in this context is a transport network or a transport service network in which functionally different network levels can be distinguished. Each network level is suited for specific trip types, especially with respect to trip length, while also providing access to higher level networks. Each level has its own characteristics regarding access density, network density and network speed. Higher level transport networks are suited for long-distance travel and have low access densities, low network densities, and high network speeds. Lower level networks are meant for short-distance travel, and thus have high access densities, high network densities, and low network speeds.

Interestingly, much of the literature on network design, such as discussed in the previous chapter, deals with single-level networks only. This lack of attention to the concept of hierarchy in transport networks is striking, since hierarchy is a very common phenomenon in nature. Hierarchy is a fundamental characteristic of resource distribution networks in biology (West et al. (1997, 1999)), such as can be seen in plants, for instance the branching of trees and sponges, and in animals and humans, for instance respiratory and cardiovascular systems. Similar suggestions for river networks and stream systems can be found in Haggett & Chorley (1969). The biological hierarchical network structures have proven to be optimal with respect to maximising metabolic capacity, that is maximising the surface area where resources are exchanged, while minimising internal transport distances. In their articles West et al. conclude that this network structure is so advantageous that it determines all kinds of fundamental relationships in biology.

The ideas of West et al., however, are not fully applicable to transport networks in general. The main difference with biological networks is that those networks have a one-to-many structure or vice versa, while in transport networks many-to-many patterns are dominant. Nevertheless, it is not difficult to define some kind of hierarchy for today's road networks or for public transport networks. This implies that the notion of hierarchical transport networks requires additional analysis.

This chapter presents a general discussion of hierarchy in transport networks: what are the main factors influencing hierarchy and what kinds of network structures are possible. The main purpose is make plausible that hierarchy in networks is a natural phenomenon, and thus is an appropriate approach for transport network design. The ideas presented in this chapter will be used in the following two chapters for a more detailed analysis of hierarchy in private transport networks (Chapter 5) and in public transport service networks (Chapter 6).

A second subject discussed in this chapter is the hierarchy in spatial structures or in settlements. As stated in Section 3.3.2, the demand pattern plays an important role in transport network design. Since the focus in this thesis is on generic relationships for transport networks, a minimum of assumptions about the demand patterns will be made if possible. For many analyses a uniform distribution demand pattern will be assumed, in some cases for a many-to-many pattern in others for a many-to-one pattern. These strong simplifying assumptions make it possible to focus on the mechanisms that are primarily determined by characteristics of transport networks, and not by the influence of the demand pattern. For urban networks the assumption of uniform demand patterns might be acceptable, but in the case of interurban transport networks this assumption is not always acceptable. The key function of transport networks is to offer transport facilities between origins and destinations, such as settlements of different sizes. The size and location of settlements thus influences trip lengths, which in turn are related to transport network levels. For analyses where the assumption of uniform demand patterns is not suitable, the framework of spatial hierarchy is used, which is discussed in Section 4.3.

## 4.2 Hierarchy in networks

### 4.2.1 Natural phenomenon

It can easily be demonstrated that hierarchy is a common phenomenon in transport networks. Let us assume a perfectly square grid network where all origins and destinations are located at the crossings, all links being equal in length and travel time. The demand pattern is uniformly distributed, that is, at every origin the same number of trips start in all directions having the same trip length, leading to the same number of arrivals at all destinations coming from al directions. Since it is a grid network the traveller may choose between a number of routes having the same length and travel time. In this hypothetical situation no hierarchy in demand or supply is assumed and at first sight no hierarchy in network usage results.

However, if small deviations to these assumptions occur, a process is started that leads at least to a hierarchical use of the network. Examples of such small changes are:

- Travellers might prefer specific routes, even though all routes are equal in time and length from an objective point of view. Such a preference might be due to habit, to the traveller's own perception of the routes or perception regarding the crossings, or to information provided by other travellers;

- Link characteristics might differ slightly leading to objective differences in route characteristics;

- Travellers might prefer to travel together, bringing in the stochastic element of travellers passing by and having an overlap with one of the possible routes;

- Some origins and destinations might be more attractive than others.

All of these deviations have the same effect regardless of the size of the change: namely some routes will become more attractive than others. This effect is mainly caused by the demand side of the transport system. The higher usage of some routes, however, also influences the supply side of the transport system. On the long run the most intensively used routes will receive better facilities and become more attractive, while the less used routes will be neglected. The supply side of the transport system thus strengthens the hierarchy started by the demand side. In fact, the process described here is an example from economics based on increasing returns (see e.g. Waldrop (1992), Arthur et al. (1987)), which is a fundamental characteristic in all kind of evolutionary processes, be they in economics or in biology. The final result in this case is a hierarchical network structure consisting of two link types, or put in other words, a higher-level network is superimposed on the original lower-level network.

This interaction between demand and supply can be illustrated using a simple grid network having a uniform demand pattern based on a direct demand model in which the attractions for all zones are equal. Travel demand is assigned to the network using a shortest path algorithm. The resulting flow pattern shows a concentration in the middle of the network, which is caused by the limited size of the network. A more essential phenomenon is that some links in the middle are used more often than others. In this case it is the result of the slightly higher demand for the zones in the middle of the network, instead of the various deviations from the assumptions mentioned earlier that everything is equal. If for the link having the largest flow the speed is slightly increased, for instance, by 10%, a process starts in which route patterns are changed by small improvements in the network and the network is influenced by the resulting usage of the network. The result of such a process for a 13 by 13 grid network, shown in Figure 4-1, clearly shows that such a process might lead to regular network structures which suggests a self-organising principle. It is interesting to note that a probit assignment instead of an all-or-nothing assignment yields similar results.
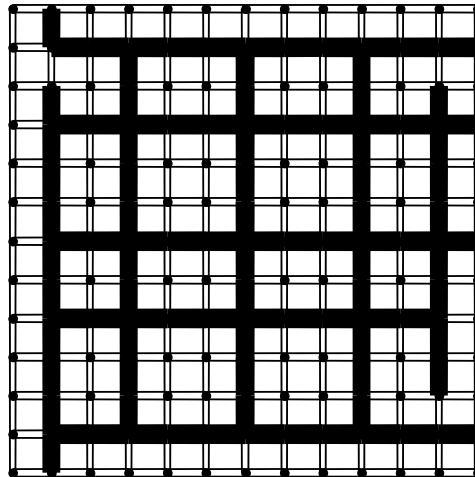
*Figure 4-1: Hierarchical network structure resulting from increasing the speed of the link having the largest flow by 10%*

This mechanism primarily applies to private transport networks. In public transport service networks, higher flows on specific routes might also lead to higher costs, since more vehicles may be required. It should be noted that a higher frequency might make the route more attractive as well. The analysis of Horner & O'Kelly (2001) for airline transport service networks, however, also showed that if there is just some mechanism leading to lower transport costs in the case of larger flows, it will automatically lead to concentration of flows on a limited number of services.

There is one important assumption in this argument that has not yet been discussed explicitly: that is travel distance. If, for instance, the trip lengths are limited to the length of single link, there will be no route choice and therefore no chance of a process leading to hierarchy. On the other hand, if longer trips do occur, such a concentration process is inevitable. Given human nature, longer trips will occur. Travelling might occur just for sake of travelling or probably more importantly for the possibility of trade.

The human drive to travel further if possible also raises the importance of speed. There is ample evidence that on average people adopt a fixed travel time budget for their out-of-home activities of 1 to 1.5 hours per day (Hupkes (1977), Zahavi (1981), Schafer (1998)). Given this constraint, the only way to increase one's interaction space is by increasing travel speed. Interestingly, the process leading to hierarchy described here also makes it possible to increase travel speed. This increase might be small due to the better facilities of the preferred routes such as better pathways or bridges instead of ferries, but might become substantial if higher speeds are possible or if faster transport modes also become an alternative. In the latter case the level of demand becomes an important factor as well, since it should be high enough to justify developing a higher level network. However, history shows that the demand level, just as trip length, has a strong tendency to increase continuously.

The introduction of faster modes speeds up the processes leading to hierarchical networks. Similarly, hierarchical transport networks lead to concentration of flows, and if these flows are large enough allowing for more efficient transport leading to lower travel costs per unit travelled (economies of scale), and for reducing negative impacts on the

environment, which also stimulates the development of hierarchical network structures. Hierarchical networks are thus a natural phenomenon resulting from the interaction between demand and supply that, due to technological developments and modern decision processes focussing on environmental impacts are becoming more common in transport networks (see Figure 4-2).
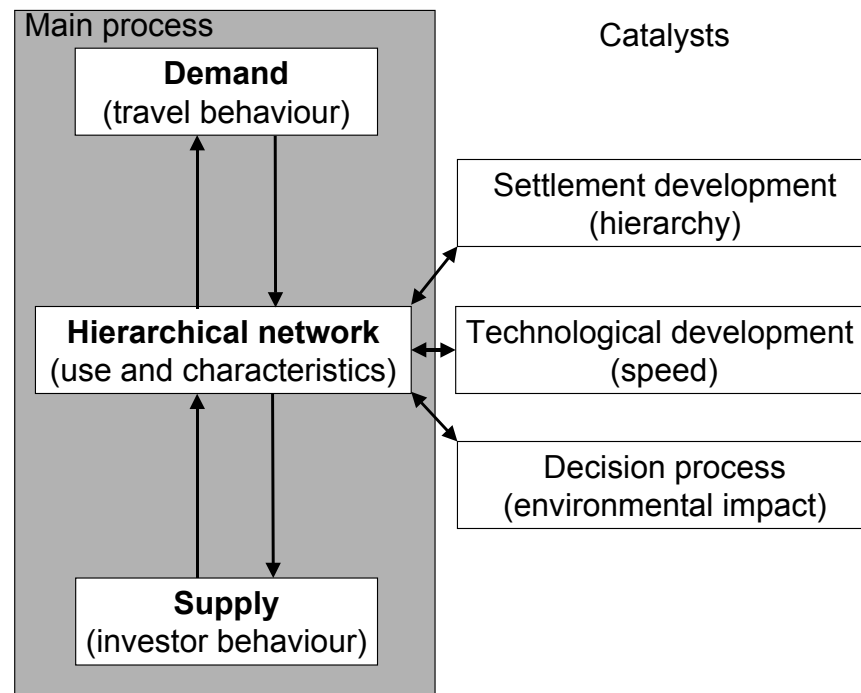


*Figure 4-2: Main factors leading to hierarchical networks*

The main process, that is the interaction between demand and supply, might have self-organising characteristics, such as is suggested by the results in Figure 4-1. Many networks, however, have been developed over a long period of time and are, therefore influenced by many factors. Hierarchy in spatial structure has always been such a factor. The importance of technology has substantially increased in the last two centuries. Rail networks were developed early in the 19th century and were a true accelerator for hierarchical network development in transport networks and spatial structures. The introduction of high-speed trains today will have a similar effect. The introduction of the private car in the beginning of the 20th century led to more ambiguous developments. Private car improved space accessibility and thus had a reverse effect with respect to spatial structure. At the same time, however, the private car allowed substantially higher speeds given the quality of the infrastructure, and can thus be seen as an accelerator for hierarchical road network development. In the second half of the 20th century a strong focus on planning processes especially with regard to environmental impacts, and the concept of bundling of transport and thus of infrastructure became dominant issues. Hierarchical networks can therefore be seen as a result of a continuous interaction process between demand and supply, which has a strong correlation with spatial development, and which is influenced over time by other developments such as technological developments and decision processes.

The variety of influences might lead to an unlimited number of hierarchical concepts. However, if transport networks do have self-organising characteristics a certain robustness in network characteristics can be expected. If such mechanisms do exist, they will limit the influence on network characteristics of all kind of external developments, and also the impact of multimodal transport.

The fact that hierarchy can be considered as a natural phenomenon however, does not explain the main relationships in a hierarchical network. For instance, which factors are most important and are there fundamental characteristics of transport networks themselves that determine the main relationships? The following two chapters focus on these questions for private transport (Chapter 5) and for line bound public transport services (Chapter 6) respectively. First, the next two sections will discuss some examples of configurations of hierarchical networks in theory and in practice.

### 4.2.2   Configurations of hierarchical networks

The analysis of hierarchical networks as a natural phenomenon leads to the conclusion that compared to lower-level networks higher-level networks have higher network speeds, lower network densities and lower access densities. Figure 4-3 shows some typical examples for a linear, grid, and a radial network, in which the higher-level network is superimposed on the lower-level network. It should be noted that in the case of the linear network the network densities of both network levels are identical.
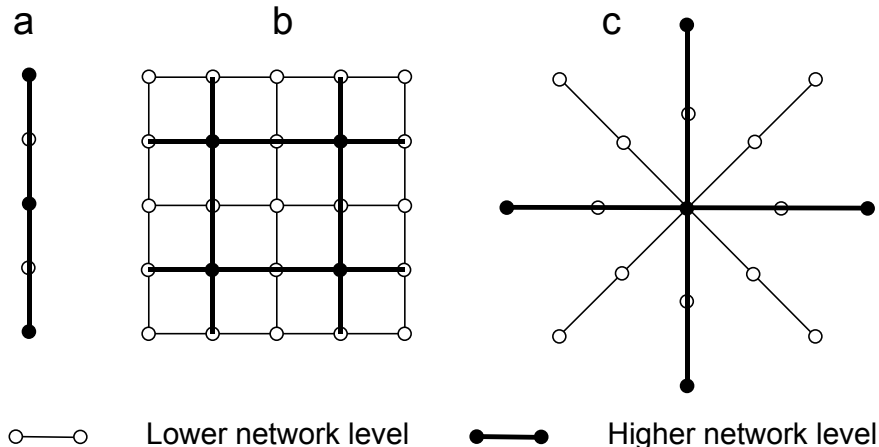


*Figure 4-3: Examples of hierarchical networks superimposed on lower-level networks*
*a: linear network, b: grid network, c: radial network*

As can be seen in Figure 4-3, it was implicitly assumed that the higher-level network has the same network type as the lower-level network. This does not always have to be the case: different network structure types can be combined in hierarchical structures. Given a specific network type there are two lines of thought. A new higher-level network should connect nodes that are already served, which limits the combinations of network types. Given a linear network, for instance, the only higher-level network that is possible is, again, a linear network. In the case of introducing a lower-level network, however, all

combinations are possible, as can be seen in Figure 4-4. Table 4-1 shows that many of these combinations are quite common.
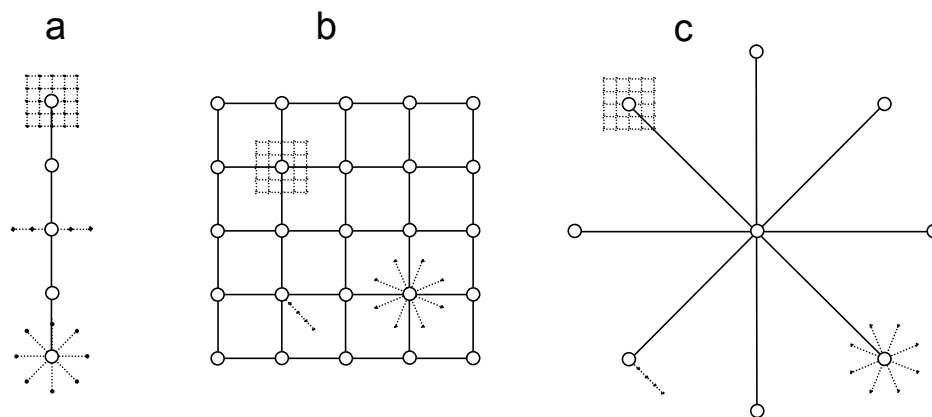


*Figure 4-4: Combinations of network types in the case of introducing lower-level networks for a: linear network, b: grid network, c: radial network*

*Table 4-1: Examples of hierarchical combinations of network types found in practice*

| Lower level | Upper level | | |
|---|---|---|---|
| | **Linear** | **Grid** | **Radial** |
| **Linear** | Feeder-line and public transport line | Driveway, especially in rural areas | Feeder-line and public transport network |
| **Grid** | Pedestrian network and public transport line | Local streets and urban motorway network | Pedestrian network and public transport network |
| **Radial** | Urban public transport network and interurban public transport line | Bicycle network in a neighbourhood and main urban bicycle network | Urban public transport network and interurban public transport network |

All examples presented so far only refer to two network levels. The number of network levels, however, can clearly be more than two. Choosing the optimal number of network levels is governed by the same dilemmas as the network design problem as discussed in Section 3.3. The main difference is that there are two mechanisms limiting the number of network levels: that of the investor or operator, because of the costs of building, maintaining and operating new network levels, and that of the traveller, because of the costs of transferring between network levels. If the transfers between network levels are seamless, such as the switches between network levels for private car, the difference might be minimal, but in the case of public transport an increase in the number of transfers is surely problematic from the traveller's point of view. This is the main reason for concepts that limit the number of transfers such as zone-systems in which a public transport line consists of an express-part, the higher-level network, and a feeder-part, the lower network, and central stations where several network levels connect at a single central node.

### 4.2.3   Hierarchy in practice

There are all kinds of hierarchical networks that can be found in practice, such as electricity, water supply, postal services, airlines, public transport, and road networks. These networks differ in the number of levels that exist. These differences are strongly determined by the characteristics of the related transport systems. Examples of such characteristics are:

- Demand pattern: many-to-one (or one-to-many) such as in electricity and water supply networks leading to tree-like network structures versus many-to-many such as in car traffic leading to hierarchical grid networks;

- Importance of travel time: time might be of virtually no importance such as in telecommunication networks allowing for, again, tree-like network structures, of limited importance such as for long distance travel or for freight transport leading to hub & spoke concepts, or time might be crucial as for frequently made trips in passenger transport and might thus limit the number of transfers;

- Costs of building and operating transfer-points between network levels: these costs may be low such as for road networks allowing many transfer points such as in hierarchical grid networks or may be high and thus limiting the number of transfer points and number of network levels as in for instance postal services or airlines;

- Scale benefits due to concentration of flows: these benefits might be substantial leading to more efficient techniques such as in the air-line industry, or may be small as in pedestrian networks.

These examples not only show characteristics that are related to hierarchy but these characteristics are also decisive for the network types used, especially with respect to grid networks and radial networks. Grid networks are especially suited for many-to-many patterns, while radial networks are especially suited in the case of an important central node or if travel time is less important, for instance in the case of short travel distances or high speeds.

Comparing these characteristics with those of multimodal transport shows some interesting features for the design problem. Given the focus on the general transport market, the demand pattern in general can be characterised as a combination of many-to-many patterns and, given the orientation to the main cities and the importance of public transport, many-to-one patterns. Travel time is the most decisive factor in all travel-related choices in passenger transport. Since transfers in public transport networks have a substantial penalty in time, the number of these transfers and thus the number of network levels per trip should be limited. The costs of building a transfer facility are substantial, but those for operating then might be limited because of the fact that travellers change modes themselves. Finally, the cost reductions due to concentration of flows are probably limited since higher-level networks usually require higher quality technologies.

## 4.3   Hierarchy in spatial structures

A look at any area, a region or a country, shows a variety of settlements, that can be divided into different categories or levels, for instance with respect to size, number of inhabitants, density, or specific functions. Ranking of these categories or levels leads to a hierarchy in spatial structures. Interestingly, such a ranking usually leads to similar results if different criteria are used for distinguishing categories. Apparently, there is a strong correlation between these criteria. If functions are used to categorise settlements, the definition of hierarchy is similar as for hierarchy in transport networks: each settlement level has its own functions and contributes to the market for specific functions located in higher level settlements.

The origin of such a hierarchy is therefore a complex subject on its own. There are many mechanisms that in one way or another determine the size and ranking of settlements. Safety, for instance, was an important factor in the Middle Ages. Transportation has influenced the maximum size of cities. Given an average time budget for travelling, cities should not exceed the distance that may be travelled in say 60 minutes (e.g. Morlok (1978)). Economic theories provide ample evidence for agglomeration processes (see for instance Christaller (1933,1963) and Lambooy (1980)). Furthermore, there are many urban functions, such as theatres, libraries and hospitals that require a minimum level of demand, for instance defined as a number of inhabitants within, again, a specific travel time.

These mechanisms vary between constraints and processes leading to equilibria. The factors that determine these equilibria, however, vary for industries, and for urban functions, and fluctuate over time. The net result of all these mechanisms is the hierarchy in settlements today. Given this complexity there is no universal model that explains spatial hierarchy. There are, therefore, many studies that try to describe or explain spatial hierarchy at an aggregate level. Descriptive studies use techniques such as ranking of cities (Zipf (1949), Van den Berg (1998), Salingaros & West (1999)), fractal analysis (Becker et al. (1994)), and morphological analysis (De Jong (1988b), De Jong & Paasman (1998)). Explanatory studies are mostly based on economic principles (Christaller (1933,1963), Lambooy (1980), Berechman & Small (1988), Medda et al. (1999)), while some other studies emphasise the importance of chance in spatial development (Arthur et al. (1987), Arthur (1988), Batty (1998), Page (1998)).

Given the purpose of this section, that is, defining a framework for spatial hierarchy as a tool for studying transport networks, the focus will be on descriptive studies. The explanatory studies, however, provide some interesting insights that explain and support the findings of the descriptive studies. Therefore, a selection of these studies will be discussed first.

### 4.3.1   Explanatory studies

Central place theory, in which one area, the central place, is more important than the surrounding areas, is one of the basic theories on spatial hierarchy. Christaller (1933,1963) defined three mechanisms leading to hierarchical structures (see Figure 4-5).
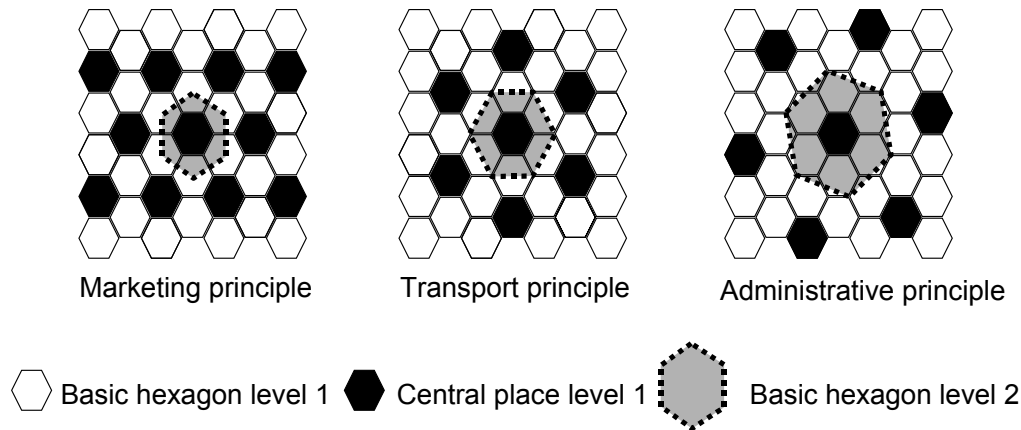


*Figure 4-5: Spatial structures according to the principles by Christaller (1933,1963)*

- Marketing principle, which is based on the assumption that competing central places will be located close together, but not next to each other. As a result the areas between central places are split up into three parts, leading to level 2 hexagons that are three times as large;

- Transportation principle, which is in fact the same principle as the marketing principle but with the additional constraint of a triangular transport network. The transport network limits the possible locations of competing central places. The areas between central places are now split up into two parts, leading to level 2 hexagons that are four times as large;

- Administrative principle, which is based on the philosophy of authorities. Neighbouring areas are not split up at all, leading to a higher-level of level 2 hexagons that are seven times as large.

The economies of agglomeration are another important mechanism for spatial hierarchy (Berechman & Small (1988)). Economies of agglomeration go a step further than economies of scale. They include positive externalities due to being located close together. Examples of such positive externalities are face-to-face contact and the possibility of information exchange. Furthermore, the process of agglomeration allows for new or larger facilities such as education or medical facilities, which in turn make agglomerations more attractive (Jansen (1980)). The transport system also plays a role. Being located close together reduces the transport costs between related industries and those for employees and customers. It should be noted, however, that due to the small share of transport costs in the total costs of the final product, transport costs are essential in the process of agglomeration but the actual value of the transport costs will be less important (Yegorov (1997)). Apart from these mechanisms leading to concentration there

are also some reverse mechanisms (Medda et al. (1999)). Agglomeration increases land use costs and leads to congestion, which increases transport costs. Jansen (1980) states that, although these reverse mechanisms will limit the growth of agglomerations, the agglomeration will generally maintain its attractiveness.

The last type of explanatory studies to be discussed focus on the location of agglomerations. The patterns defined by Christaller (Figure 4-5) have a limited resemblance with patterns found in reality. The main reason for this discrepancy is that agglomeration processes need specific circumstances, such as natural resources, location in the transport network, political relationships, and need sufficient time to achieve a kind of critical mass to ensure its survival (Arthur et al. (1987), Arthur (1988), Batty (1998)). Furthermore, the mechanisms influencing agglomeration development may vary over time, stimulating or slowing down the development of agglomerations (see for instance Lees & Hohenberg (1988)). It is therefore impossible to say beforehand where agglomerations will start and which will be successful.

### 4.3.2 Descriptive studies

Ranking cities is an intuitive way to describe spatial hierarchy. A study by Zipf (1949) showed a surprisingly strong relationship between population size and the reciprocal of city rank. Haggett (1965) states that the combinations of stochastic processes and the hierarchical concept of Christaller leads to a ranking as found by Zipf. Salingaros & West (1999) present similar multiplicity rules, which state that there is a balance between the size and number of cities of various sizes. They even state that deviations from these multiplicity rules, such as too many large cities or too many small cities, lead to dysfunctional regions.

Van den Berg (1998) discusses a study by Van Marrewijk et al. (1998) in which economic agglomeration models describing centrifugal and centripetal mechanisms were used to model spatial development. The study showed that the resulting ranking of settlements could be described using the relationship found by Zipf if an exponent for the city rank was introduced. A historical analysis for the Netherlands showed that this exponent varies over time. In 1600 the value of this exponent was 0.55 implying a limited variance in settlement size, which was explained by the poor quality of the transport system. At the end of the 19th century, when the rail and waterway networks were completed and industrial development was strong, the exponent equals 1 as in the original formulation by Zipf. In the 20th century, however, the value drops to 0.7 probably due to increasing congestion in the transport system. This analysis illustrates two things. First, there is a relationship between the mechanisms behind spatial hierarchy and the descriptive approach. Second, such a multiplicity rule only describes a relationship at a specific moment and provides no explanation. Its value is therefore limited.

The multiplicity rules of Salingaros & West (1999) are based on entropy as an organising principle, but they are also related to fractal geometry. Becker et al. (1994) also used fractals to study urban form. They found that city edges could be described using fractals, meaning that as cities grow city edges grow even faster. This implies that there is a strong

tendency to minimise the distance between any urban function and the city edge. Furthermore, they found that street patterns, that is, the lower-level networks, in historic cities are primarily determined by land use patterns and not by transport planning concepts.

The last descriptive approach discussed in this section is the morphological approach (De Jong (1988b), De Jong & Paasman (1998)). They defined a legend for analysing spatial planning proposals. The main principle used is that at each level new details should be visible. Objects, for instance, that are homogeneous at a certain level might be completely dispersed at a lower level. An experiment using black and white hexagonals shows that a reduction of the radius by a factor 3 is the minimal factor needed to achieve maximum diversity in homogeneity (see Figure 4-6). This factor 3 is defined as the scale-factor for settlements, that is the radius of a higher level settlement will always be thrice as large. It is interesting to note that this approach follows the administrative principle defined by Christaller.



Dispersion or                          Concentration or
maximal homogeneity                    minimal homogeneity
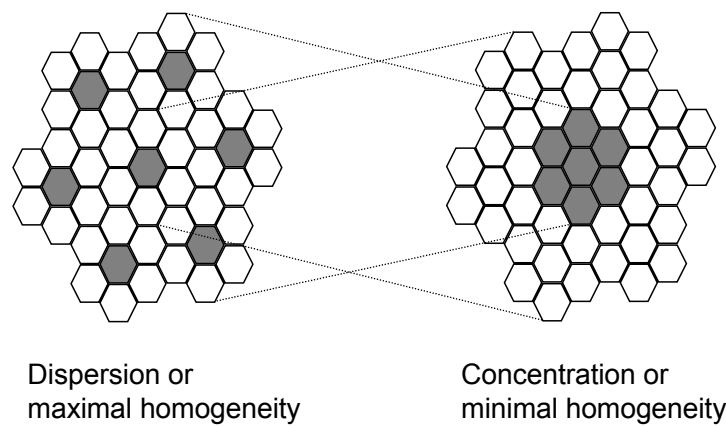
*Figure 4-6: Distinction between dispersion and concentration (De Jong (1988b))*

In combination with the fact that an urban area in the Netherlands with a radius of 1 kilometre has an average of 10,000 inhabitants, De Jong & Paasman used this scale-factor 3 to define a terminology for hierarchical spatial structures as shown in Table 4-2. It is interesting to note that while seven hexagons were used to illustrate the scale-factor 3, they also define a scale-factor 10 for the population of these settlements, which allows for varying the densities within a spatial unit.

Since the approach proposed by De Jong & Paasman is suitable for both theoretical purposes, for instance according to the concepts of Christaller, as well for realistic settlement structures, it will be used as the main framework to describe spatial hierarchy.

*Table 4-2: Characteristics of hierarchical levels in settlements*
*(De Jong & Paasman (1998))*

| Name | Radius [km] | Surface [km$^2$] | Inhabitants |
|---|---|---|---|
| Village – Neighbourhood | 0,3 | 0,3 | 1,000 |
| Town – District | 1 | 3 | 10,000 |
| City | 3 | 30 | 100,000 |
| Agglomeration | 10 | 300 | 1,000,000 |
| Metropolis | 30 | 3,000 | 10,000,000 |

### 4.3.3    Spatial structures and transport networks

The key function of transport networks is to offer transport facilities between origins and destinations such as between settlements of different sizes. The function, size and location of settlements thus determine trip characteristics such as the number of trips and trip lengths. The theoretical concepts for hierarchy in spatial structures lead to very rigid patterns, while in reality there is no clear pattern of settlements of specific sizes. This implies that the transport networks will not have rigid patterns either, but will be stretched to match the randomness in the locations of settlements. This argument, however, does not imply that specific network types such as grid networks of triangular networks are unrealistic, but emphasise once again, the importance of interpreting the results of analyses based on network types.

The hierarchy in settlement sizes can also be used to determine the range of travel distances related to trips between settlements of specific sizes. The hierarchy in settlements defined by De Jong & Paasman (1998) agrees with the administrative principle defined by Christaller (1933, 1963). According to the administrative principle the central city is surrounded by lower-level settlements, leading to service area of the central city having a radius of three times the radius of the central city (see Figure 4-5 and Figure 4-7). In that case the distance between two city centres is theoretically equal to 6 times the radius of the city.

Given the randomness in the location of settlements, however, it is also possible that the service areas of central cities overlap, which leads to spatial configurations that resemble the marketing principle as defined by Christaller (see Figure 4-5 and Figure 4-7). The distances between the central cities then are equal to four times the radius of the city. Such a spatial configuration might be representative for an urbanised region. In rural areas, on the other hand, it is also possible that the settlements are spaced further apart, having typical rural functions such as agriculture, parks or forests in between. In that case the distance between settlements might be 12 times the radius.
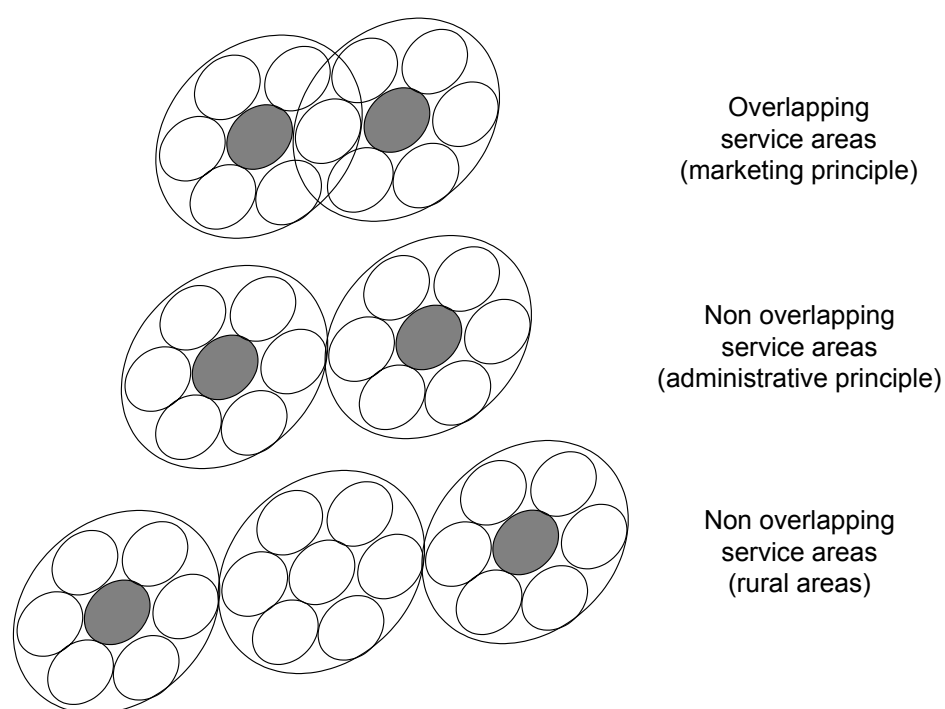
*Figure 4-7: Spatial configuration of settlements
with and without overlapping service areas*

In order to give an idea of the distances involved these principles are used to define ranges for the minimal distances between settlements of the same rank (Table 4-3). As can be seen, there is a large range of realistic distances between settlements of given rank. Thus the actual distances will depend on the actual situation.

*Table 4-3: Minimal distances between hierarchical settlements*

| Name | Radius [km] | Minimal distance between settlements [km] | | |
|---|---|---|---|---|
| | | **Urbanised** (Marketing principle) | **Basic** (Administrative principle) | **Rural** |
| **Village** | 0.3 | 1.2 | 1.8 | 3.6 |
| **Town** | 1 | 4 | 6 | 12 |
| **City** | 3 | 12 | 18 | 36 |
| **Agglomeration** | 10 | 40 | 60 | 120 |
| **Metropolis** | 30 | 120 | 180 | 360 |

## 4.4  Conclusions

This chapter discussed the subject of hierarchy in transport networks in general and the related subject of hierarchy in spatial structures. Network hierarchy is an important property of transport networks that is often neglected in literature on transport network design. It is argued in this chapter that hierarchy is a natural phenomenon in maximising performance while minimising the resources needed. The general principles explaining this phenomenon are that travellers try to maximise their utility, while investors in transport networks try to achieve economies of scale. The net result is that in transport networks different network levels will evolve, each of which is suited for specific trip types, while in spatial structures a hierarchy with respect to size and functions will emerge.

The finding that a hierarchy in transport networks can be considered a natural phenomenon does not explain the main relationships in a hierarchical network. These relationships will be analysed in more detail for private transport networks and line-bound public transport service networks in the following two chapters. A first look at existing hierarchical networks, however, already showed four important factors that influence hierarchical network patterns:

- Demand pattern;

- Importance of travel time;

- Costs of building and operating transfer facilities;

- Scale benefits due to concentration of flows.

These characteristics also have a strong influence on the network types used. For multimodal transport the first two characteristics seem to be most important. Given the focus of multimodal trips to and from the main cities a many-to-one pattern seems to be most relevant, although the more general notion of multimodal transport as an alternative for private transport also implies a many-to-many pattern. Travel time is the main factor influencing travel behaviour. Consequently, routes on the network should be as direct as possible, unless the speed compensates for the detour, and the number of transfers should be limited.

The hierarchy in spatial structures can be explained by many mechanisms, but till now, there is no grand theory explaining the development of a spatial hierarchy, and perhaps there never will be. Interestingly, there is a remarkable consistency in descriptive studies of hierarchy in settlements. Furthermore, the characteristics of these descriptions are in accord with the main mechanisms explaining the emergence of hierarchy. Since in this thesis a hierarchy in spatial structures is only used to account for the influence of the spatial pattern on transport systems, a choice is made for a distinction in hierarchical levels according to De Jong & Paasman (1998) as shown in Table 4-2. These levels will be used to determine the size of urban areas in the analysis of urban public transport services (Sections 6.3 and 6.4), and for the analysis of interurban transport networks in cases where distances between settlements are relevant.

# Chapter 5

# PRIVATE TRANSPORT NETWORKS

## 5.1 Introduction

This chapter elaborates the characteristics of hierarchical transport networks that are used by private modes, such as private car, bicycle and pedestrians. Hierarchical transport network structures are defined as networks having different network levels, each of which having their own transport function in terms of serving specific types of settlements or specific travel distances, while also providing access to higher network levels.

The focus on private transport eliminates service network issues such as transfers, time-accessibility and service components, and thus concerns infrastructure networks exclusively. The main network characteristics used in this analysis of hierarchy of private transport networks are thus space-accessibility, network density, and network speed.

In this chapter only grid and triangular networks will be discussed. Typical characteristics of private transport networks are that nearly every address has access to the network, and that the demand has a many-to-many pattern. Private transport networks should therefore have high space accessibility and should provide coverage of urban and rural areas. The dominant network types are then grid and triangular network structures, although in some cases radial network structures might also be relevant, for instance in case of short distances in local neighbourhoods, or in case of a dominant central node.

Given the importance of longer trip lengths for multimodal travel, the analysis will focus on higher-level networks. For lower-level networks, especially at the neighbourhood level, network structures are strongly influenced by factors other than transport. As Becker et al. (1994) concluded, land use characteristics, such as lot-size and function, determine the urban network structure. Furthermore, streets are used as an important structuring element in urban planning, and they provide space for other networks such as water, energy, and telecommunications. Due to this combination of factors there will be a

large variety of network characteristics. For the higher-level networks, however, the transport function becomes more dominant.

The purpose of this chapter is to show that hierarchical network structures exist and to establish the main relationships determining the emergence of hierarchy in private transport networks. Theoretically, such relationships might be the result of planning guidelines, of fundamental characteristics of transport networks in general, or might not exist at all. First, an overview is given of hierarchical properties of existing networks and planning guidelines, which will show that there is sufficient evidence of a hierarchical network structures. Second, the chapter focuses on the key question: what are the relationships that determine network hierarchy? Three different approaches are introduced. First, an analysis is presented of the fundamental characteristic of minimising travel costs that might determine this relationship. The next section presents an alternative analysis based on an economic approach using the objective of minimising total costs, which provides additional support for the relationship found. Finally, a simple network development strategy is presented which, surprisingly, again leads to similar hierarchical network structures. The findings on the relationships between network levels then is used to present a classification of road network levels, which is compared to the existing road networks in the Netherlands. Furthermore, some implications for future developments of road networks are discussed. The chapter concludes with the main conclusions with respect to hierarchy in private transport networks.

## 5.2   Private transport networks in retrospect

This section discusses characteristics such as number of network levels and network density, for the main private transport network in the Netherlands, the road network, and investigates how far planning guidelines might have influenced these characteristics. Some examples of planning guidelines in other countries are also presented.

### 5.2.1   Road network in the Netherlands

Paths are the natural result of people who are travelling for their daily needs or for trade. The earliest pathways found are wooden corduroy constructions in swamps. Some of these pathways developed into manufactured roads, for instance for military purposes (Lay (1990)). The best known example is the Roman road network, which was especially suited for military and administrative purposes. Interestingly, these roads had steep gradients, which made them unsuitable for freight transport, a phenomenon also found in Persia, Peru and China (Morlok (1978)). At the end of the Roman Empire, the quality of these roads deteriorated. Thus the development of the road network started again in the Middle Ages leading to haphazard networks. Roads were found parallel to the rivers (along the dikes), along the coast and to and from the main cities. They were used for postal services, for military purposes, and for trade, especially the trade routes to the east (Hessenroads). In 1800 the Netherlands had only 165 km of paved roads (Van der Woud (1987)).

In 1811 the French authorities made a classification of the road network in the French Empire consisting of five levels:

- Imperial roads class I: Amsterdam-Utrecht-Gorkum,-Breda-Antwerpen-Parijs;

- Imperial roads class II: Breda-'s Hertogenbosch-Nijmegen-Zwolle-Groningen;

- Imperial roads class III: Antwerpen-Rotterdam-'s Gravenhage-Haarlem, Brussel-'s Hertogenbosch-Venlo-Nijmegen, and Luik-'s Hertogenbosch-Utrecht;

- Regional roads (Routes Departemental);

- Local roads.

This categorisation should not be seen as a plan for a road network, but as a classification for administrative purposes for financing and maintenance, especially focussed on the interests of France. The roads from the Netherlands to Germany, for instance, are not included in the categorisation (Van der Woud (1987)).

The development of the road network in the Netherlands really started in the first half of the 19th century. The quality of the roads between the main cities was substantially improved as a first step towards a national road network (Van der Woud (1987)). Originally, many roads were toll roads, but the revenues dropped substantially because of the introduction of the train and later on also because of the tramways. Financing roads using private funds became a problem, thus making road network development primarily a concern of the local administrations.

An interesting development was the introduction of the rubber tyre, making the bicycle the main transport mode for short distance trips. The popularity of the bicycle led to the development of road maps and bicycle associations, the forebears of the national car associations (Lay (1990), Knippenberg & De Pater (1988)). Thus, the importance of the quality of especially local roads increased substantially. The introduction of the private car in the beginning of the 20th century influenced road network development only slowly. Road counts on national roads in 1923 showed an average of 9 cars, 70 bicycles, and 12 other carriages per hour (Knippenberg & De Pater (1988)).

The development of motorways started at the end of the first half of the 20th century. The national policy plan for the spatial structure in the Netherlands in 2000 (Staatsuitgeverij (1966)) presented a grid like motorway network having a road spacing of 15 to 20 km. The costs needed for realising this plan were clearly too high, and furthermore, the attention to environmental impacts and other land-use characteristics such as landscape and natural and cultural heritage increased. As a result the plans for the motorway network presented in more recent national transport plans (SVV (Second Structure Scheme for Traffic and Transport) (Ministry of Transport, Public Works and Watermanagement (1989)), NVVP (National Traffic and Transport Plan) (Ministry of Transport, Public Works and Watermanagement (2000)) have a substantially lower network density. The SVV focuses on a national road network that offers direct

connections between the 40 main urban areas in the Netherlands. This network should have a detour-factor of 1.4 at most, and the maximum travel distance to the access nodes is 10 kilometres, or 15 minutes. On top of that, special attention is given to the routes that connect the mainports Rotterdam and Schiphol Amsterdam to the neighbouring countries. Table 5-1 gives an overview of the characteristics of the road network in the Netherlands in 1996, while Figure 5-1 shows the national road network in the Netherlands according to the NVVP. Please note that not all national roads are motorways.
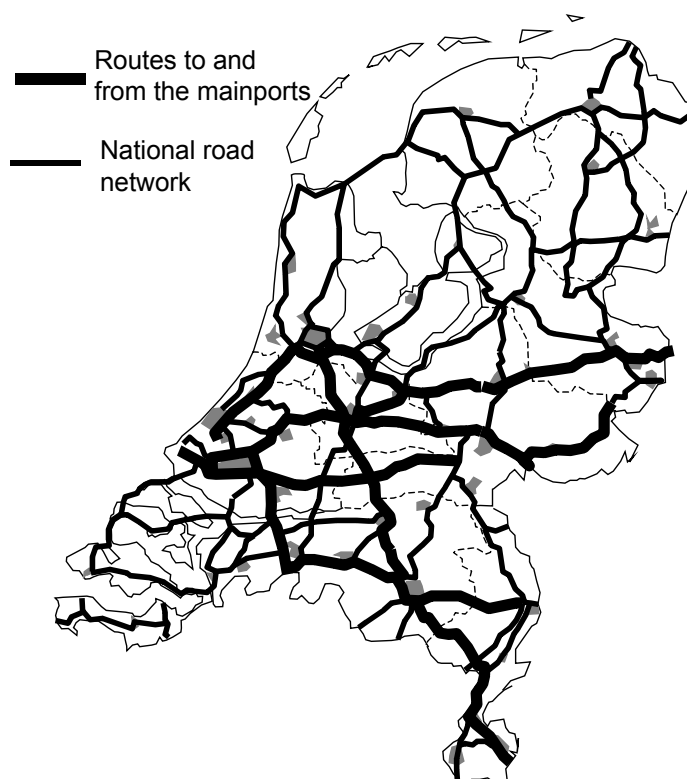


*Figure 5-1: National road network in the Netherlands according to the NVVP*
*(Ministry of Transport, Public Works and Watermanagement (2000))*

De Jong & Paasman (1998) present a morphological description of the road network in the Netherlands (Figure 5-2). They make a distinction between different levels:

- Continental level, consisting of transport axes following the coast line of Europe, that is Amsterdam - Paris and Amsterdam - Hamburg, and a second one at a distance of 500 kilometres;

- Fluvial level, consisting of transport axes along the main rivers in Europe, i.e. Rhine, Elbe, Seine and Somme having a spacing of 300 kilometres;

- National transport system, which is an orthogonal network with respect to the continental and fluvial system. Due to the shape of the continental system, or more precisely the shape of the coastline, there is a difference in orientation between the northern part and the southern part of the network.

*Table 5-1: Road length by type for the Netherlands in 1996*
*(classification based on the responsible authorities)(CBS)*

| Road type | Road length [km] | Percentage | Network density [km/km$^2$] |
|---|---|---|---|
| *All roads* | | | |
| **Motorways** | 2.208 | 2% | 0,06 |
| **Other national roads** | 998 | 1% | 0,03 |
| **Regional roads** | 6.910 | 6% | 0,20 |
| **Other road types** | 103.304 | 91% | 2,95 |
| **Total** | 113.419 | 100% | 3,24 |
| *Outside built-up areas* | | | |
| **Motorways** | 2.207 | 4% | 0,06 |
| **Other national roads** | 936 | 2% | 0,03 |
| **Regional roads** | 6.360 | 11% | 0,18 |
| **Other** | 48.699 | 84% | 1,39 |
| **Total** | 58.202 | 51% | 1,66 |
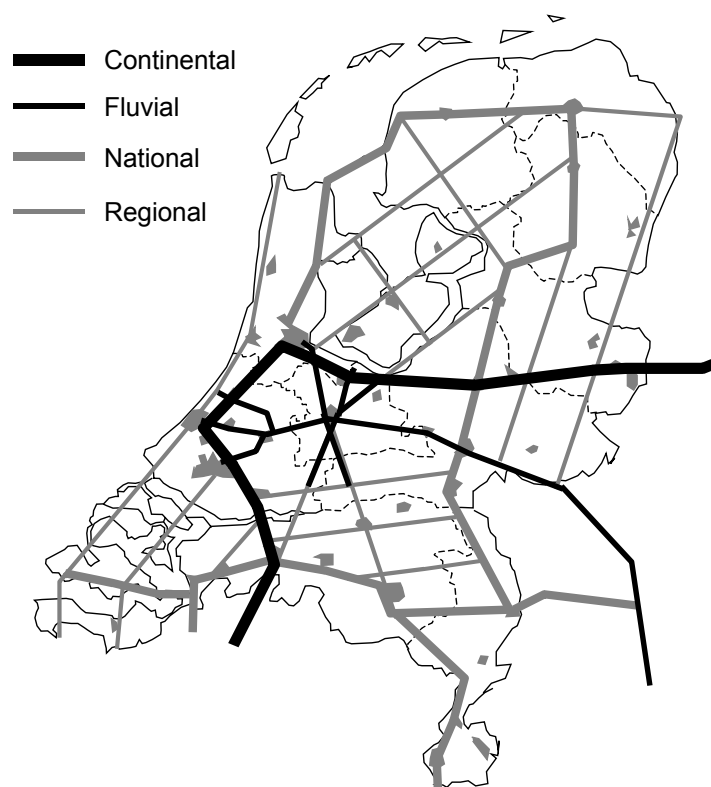


*Figure 5-2: Network levels according to De Jong & Paasman (1998)*

The national road transport system itself also consists of different network levels:

- National road network with a road spacing of 100 km. However, since the urban development in the Netherlands is located in the west the eastern part of this network has not yet been developed;

- Regional road network is rectangular network having a cell length of 50 km and a cell width of 20 km, which is equivalent to an average road spacing of 30 km. This network clearly shows the difference in orientation mentioned before;

- Local road system, having a road spacing of 10 km;

- Urban road system, which is again a rectangular network of 5 by 2 km or an average road spacing of 3 km.

Table 5-2 gives an overview of the main network characteristics for each level. Compared to the network densities found in the Netherlands as shown in Table 5-1 it can be seen that what is defined in the morphological approach as the regional road system is equivalent to the national motorway system. Furthermore, the density of the local road network is similar to the regional road network in the Netherlands.

*Table 5-2: Road network level characteristics according to De Jong & Paasman (1998)*

| Network level | Network density [km/km$^2$] | Average road spacing [km] | Square or rectangular |
|---|---|---|---|
| National roads | 0,02 | 100 | Square |
| Regional roads | 0,07 | 30 | 20 x 50 |
| Local roads | 0,20 | 10 | Square |
| Urban roads | 0,70 | 3 | 2 x 5 |

The morphological analysis shows a hierarchical network structure having a decreasing network density per network level with a scale-factor 3, which is also applicable to road spacing. Furthermore, De Jong & Paasman state that the network structure alternates per network level between a square grid and a rectangular network. De Jong (1988a) concluded that given the long history of the road network the scale-factor 3 for road spacing and road density can be seen as a kind of universal constant. It is interesting to note that the scale-factor 3 fits with the concepts for hierarchy in spatial structures as defined by De Jong (1988b) which were discussed in Section 4.3.2.

This morphological analysis might suggest that every network level should exist everywhere, however, that is certainly not necessary. First, the spatial structure should match the network structure and the demand should be high enough to justify all network levels. There is a positive relationship between population densities, and thus expected traffic intensities, and road investment costs (Puu (1978)). For the Netherlands it was already noted that, for instance, the eastern part of the national road network was less developed due to the limited urban development in that part of the Netherlands. Second, a

higher network level might replace roads of the lower network level. As a result the network density of the lower network level is reduced, leading to lower network costs. In the development of the national motorway system in the Randstad, for instance, the development of a regional network was skipped in favour of developing a motorway network (Hilbers et al. (1997), Immers et al. (2001)). The functions of the regional roads and the national roads are thus combined on the same infrastructure.

The morphological analysis by De Jong & Paasman strongly suggests a hierarchical network structure having a scale-factor 3 for road spacing and network density. The disadvantage of this morphological analysis however, is that it lacks an explanation for this phenomenon. Section 5.3 will analyse fundamental characteristics of transport networks that might offer such an explanation.

### 5.2.2   Guidelines for road network design

The Dutch guidelines for road design are primarily determined by a road design perspective (Commissie RONA, 1992). The guidelines roughly distinguish three network types:

- Higher network level, serving the main economic centres and having an average speed of 90 km/h.;

- Middle level network level, oriented at regional trips and having an average speed of 70 km/h.;

- Lower level network, consisting of the remaining roads having an average speed of 50 km/h.

Furthermore, these guidelines define 8 road categories, thus allowing the designer to choose a road category that fits local requirements. Four of these categories are even applicable to different network levels. From a traffic safety's perspective, however, it is recommended to reduce the number of road categories by making a clear distinction between the three main functions: flow, distribution or access (Wegman & Wouters (2002)). For the roads having a flow function a distinction might be made between motorways and motor roads, while for the other two functions a distinction is made between rural and urban areas.

An extensive analysis of the functions of a road network can be found in Schönharting & Pischner (1983). This analysis is based on a hierarchy in cities using the central place concept of Christaller (see Section 4.3.1). The road network offers transportation between cities of all kind of levels. Each road network level connects cities of a specific type and connects these cities with cities of the next higher level. Thus, different transport functions can be distinguished according to the city types that are served, as can be seen in Figure 5-3. Per function or trip type the average distance differs. Given politically determined values for the accepted travel time for these trip types, minimum speeds per road type can be derived. The results of this analysis are incorporated in the official guidelines for road network design in Germany (FGSV (1988)). Table 5-3 shows the main characteristics for each network level according to these guidelines. As can be seen,

each network level has a range of network speeds, thus allowing for subdivision in road types.
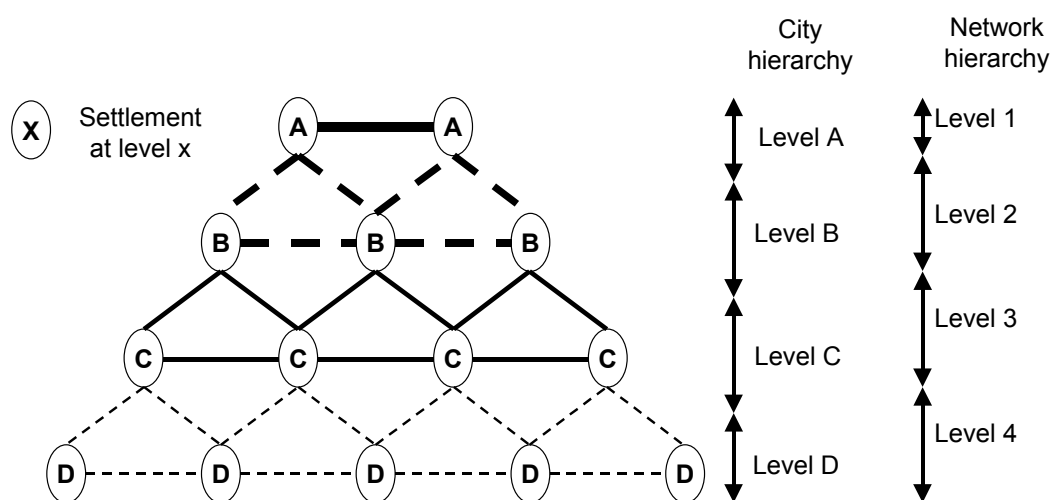


*Figure 5-3: Road network structure according to Schönharting & Pischner (1983)*

*Table 5-3: Road network level characteristics according to RAS-N (FGSV (1988))*

| Network level | Trip type | Trip distance [km] | Network speed [km/h] |
|---|---|---|---|
| **1 National** | A - A | 100 - 200 | 70 - 100 |
| **2 Interregional** | A- B and B - B | 50 - 100 | 60 - 90 |
| **3 Regional** | B - C and C - C | 25 - 50 | 50 - 80 |
| **4 Local** | C - D and D - D | 0 - 25 | 40 -60 |

*Number of network levels and trip types refer to Figure 5-3*

Bovy et al. (1994) also chose a functional approach to transport networks. Their analysis follows the same line of reasoning as Schönharting & Pischner (1983). The main difference is that Bovy et al. explicitly divide each network level into two types: offering access to cities of a higher-level and offering connections between cities. An aspect Bovy et al. focus on is that a hierarchical network leads to a strong concentration of flows, thus allowing the adoption of higher quality technologies and limiting environmental damage. Thus they define an explicit relationship between the hierarchy from a functional point of view and the technological characteristics of the network levels.

The impact of concentration of flows can be illustrated by the Dutch national road network, which has 3% of the total road length, and accommodates 40% of the vehicle-kilometres by car. However, it should also be noted that the motorway system might be too attractive: while from a functional point of view the motorway system is meant for long distance trips (network level 1) it is also used by medium and short distance trips (network levels 2 to 4). These trips experience the relatively high quality of the motorway system, which influences travel behaviour: location choice, destination choice, mode choice, and route choice. The net result is a relatively large increase of these medium and

short distance trips using the motorway network in quantity as well as in trip length, and thus leading to congestion earlier than expected. The implicit combination of functions thus might in the long run lead to a loss of quality for the long distance trips (Van Nes (1998), Bovy (2001)).

Both analyses clearly focus only on the functional characteristics of the road network. No specific attention is paid to network characteristics that might also determine network levels. It is interesting to note that there is a strong resemblance with the results of the morphological analysis discussed in the previous section. The central place concept used by Schönharting & Pischner (1983) agrees with the hierarchy concept for spatial structures defined by the De Jong (1988b). Furthermore, the hierarchy in network levels described by De Jong & Paasman (1998) matches the hierarchical concepts presented by Schönharting & Pischner (1983) and Bovy et al. (1994).

Finally, it interesting to note that guidelines for bicycle networks make a distinction between normal bicycle paths and main bicycle routes, having a path spacing of 150 to 200 metres and 450 to 600 metres respectively (Bach (1999)). The resulting scale-factor for path spacing matches the scale-factor for road networks found by De Jong & Paasman (1998).

## 5.3   Fundamental network characteristics for hierarchy

If there is a network characteristic that determines network hierarchy it is plausible that, given the interactions between network usage and network development described in Section 4.2.1, this characteristic should be related to the use of the network. A hierarchical network is only successful if it is used in a hierarchical fashion, that is that each network level is predominantly used by the categories of travellers it was meant to serve. From this point of view there are two criteria that might be used to characterise hierarchical networks. The first is based on reducing travel time or more generally travel costs, the general notion in travel behaviour. The use of this criterion focuses on the notion that the higher-level network should be used by a certain group of travellers. The second criterion deals with the elimination of shortcuts, which is equivalent to the idea that a higher-level network should be used as much as possible. Both criteria will be discussed further in the following sections.

### 5.3.1   Reducing travel costs

The basic principle of utility maximisation, which is often similar to reducing travel costs, implies that a higher-level network is only interesting to travellers if for a specific trip the use of the higher-level network leads to higher utility or lower travel costs. Since travel costs are primarily determined by travel times, especially for private transport, and since travel times are intuitively more related to transport networks than transport costs, the analysis will focus on travel times only.

It is assumed that there is a lower-level grid network *i* having road spacing $S_{r,i}$, and travel speed $V_i$ (see Figure 5-4). The higher-level network *i+1* has road spacing $S_{r,i+1}$, which is a multiple of $S_{r,i}$ ($sf_r \cdot S_{r,i}$), and travel speed $V_{i+1}$. Since non-integer values of the scale-factor for road spacing $sf_r$ often lead to inefficient land use, only integer values will be considered. Furthermore it is assumed that $sf_r$ is independent of the network level used in the analysis. Theoretically $sf_r$ might vary between different network levels, but given the findings of De Jong & Paasman (1998) that does not seem very plausible. Finally, it should be noted that in this analysis the higher-level network roads do not replace the lower-level network roads.
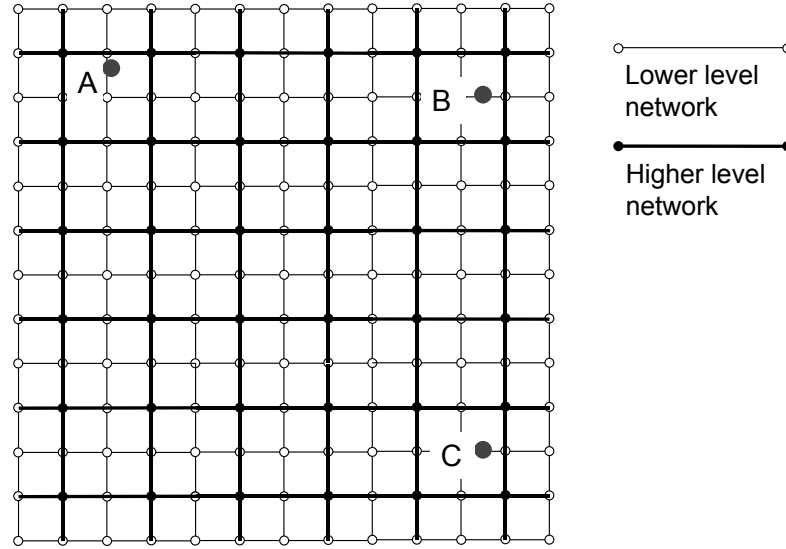


*Figure 5-4: OD-pairs in a hierarchical grid network*

In general, when using the higher-level network the trip travel times need to be less than when using the lower-level network only (see Van Nes (1998)). Let us look at a specific trip in a hierarchical grid network, for instance from A to B or from A to C (see Figure 5-4). The trip lengths for both the lower-level network and the higher-level network will be more or less equal, except for some short trips. The trip using the higher-level network, however, will also include access time and egress time using the lower-level network since the origin and destinations are often not located on the higher-level network. Please note that the access time and egress time for the lower network are excluded from the analysis. The difference in speeds between the two levels will have to compensate for the travel time that is lost with access and egress using the lower-level network. It can be shown that the travel speed for the higher-level network should fulfil inequality (5-1) (see Van Nes (1998)).

$$V_{i+1} > \frac{V_i}{1 - \frac{2 \cdot f_a \cdot sf_r}{f_d} \cdot \frac{S_{r,i+1}}{L}}$$

(5-1)

with:

$L$     = airline distance between origin and destination

$f_d$    = factor for the detour between origin and destination due to

        the network structure, $1 < f_d < \sqrt{2}$

$f_a$    = factor for the average access and egress distance for using the

        higher level network, $0.19 < f_a < 0.33$

$V_i$    = network speed in network $i$, i.e. travel speed between access and egress points

$S_{r,i}$  = road spacing of network $i$

$sf_r$   = scale factor for road spacing, $sf_r \in \{1,2,3,4,.......\}$

or formulated the other way around:

$$sf_r > \frac{f_d \cdot L}{2 \cdot f_a \cdot S_{r,i+1}} \cdot \left(1 - \frac{V_i}{V_{i+1}}\right) \qquad (5\text{-}2)$$

These inequalities show that the required difference in travel speeds is strongly dependent on the trip lengths in the spatial system that is considered. For long trips, a small difference in speeds will be sufficient to make using the higher-level network attractive. The choice of a specific trip length, however, is arbitrary.

### 5.3.2    Elimination of shortcuts

The second criterion for network hierarchy from the traveller's point of view is the elimination of shortcuts. In a way this is a stronger formulation of the former criterion of reducing travel time, which aims at accommodating a minimum set of trips only, namely having an arbitrary minimum trip length. Eliminating short cuts, on the other hand, focuses at maximising the use of the higher-level network.

A possible approach in this case is to look at the maximum detour for a shortcut in a single grid. This detour determines the necessary difference in travel speed between the network levels. The most realistic scale-factor for road spacing can be found by calculating this travel speed ratio for a set of scale-factors, and selecting the scale-factor resulting in the lowest travel speed ratio. The choice for the lowest value is based on the intuitive notion that the lower the travel speed ratio, the easier it will be to develop a higher-level network.

In this approach, the only assumption is that the trip length is equal to or longer than the road spacing of the higher-level network. Within a grid network the trip having the maximum detour using the higher-level network can be defined as the trip between two nodes that are located at the middle of two opposing sides of the grid of the higher-level network (see Figure 5-5 (a)). In the case that $sf_r$ is uneven, this trip is located between two nodes as close to the middle as possible (Figure 5-5 (b)). The trip distance using the lower-level network is always $S_{r,i+1}$.
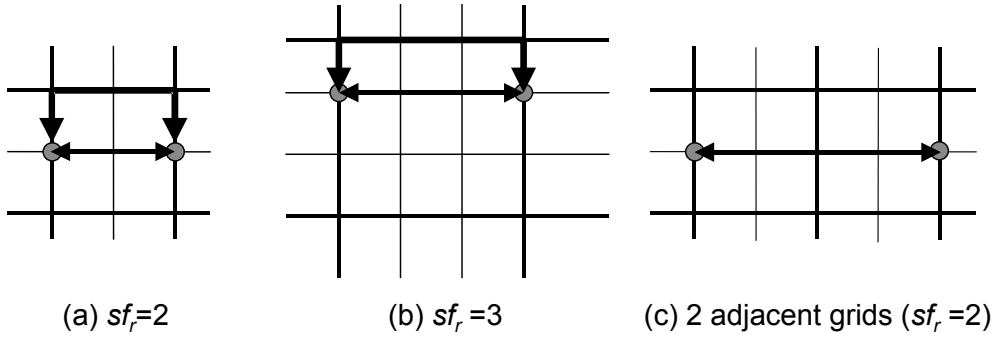
*Figure 5-5: Maximum detours in a hierarchical grid network*

If $sf_r$ is even, the trip distance using the higher-level network is twice as large, which implies that the travel speed for the higher-level network should be at least twice as high in order to have a shorter travel time using the higher-level network. This implies that in this case no choice for the most realistic scale-factor can be made. In case $sf_r$ is uneven, the trip distance for the higher-level network becomes $(2 \cdot sf_r - 1)/sf_r$ as large. In order to have a shorter travel time using the higher-level network, the travel speed should increase accordingly. It can easily be shown that the smallest increase of travel speed is found if $sf_r$ equals 3: $V_{i+1}$ then is $1.67 \cdot V_i$. As $sf_r$ increases the necessary increase in travel speed converges to a factor 2. In both cases the maximum travel speed ratio is 2. Apparently, it is not necessary to have larger travel speed ratios to avoid short cuts.

It is possible, of course to relax the criterion that determines whether the higher-level network is interesting or not, for instance, by using the trip between the middle of the opposing sides of two, or more, adjacent grids (see Figure 5-5 (c)). In that case the travel speed ratio will decrease. Again, the choice of the number of adjacent grids is arbitrary.
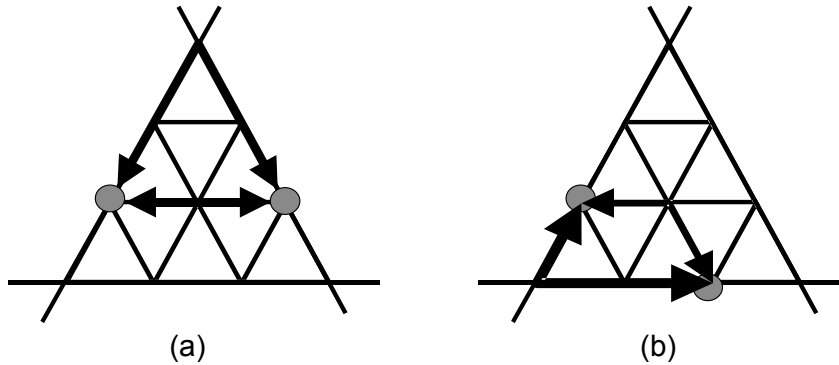


*Figure 5-6: Maximum detour in triangular networks ($sf_r$=3)*

When triangle networks are considered, the situation becomes slightly more complicated. For any trip along a straight line, the travel speed ratio equals 2 (see Figure 5-6 (a)), which makes it impossible to choose a scale-factor. For any other trip using the lower-level network within the higher-level network triangle, the maximum trip length is $sf_r$-1 and the trip length using the higher-level network $2 \cdot sf_r$-3 (Figure 5-6 (b)). This implies that if $sf_r$ equals 3, the travel speed ratio will have the lowest value: 1.5. If the criterion for the trip length is relaxed, the problem becomes identical to the problem for grid networks.

This analysis clearly shows that the existence of a scale-factor 3 for the road spacing of hierarchical road networks can be explained using a simple and plausible mechanism based only on network characteristics. The corresponding scale-factor for network speed $sf_v$ is 1.67 and should not be larger than 2. An interesting feature of this analysis is that it requires an absolute minimum of assumptions.

## 5.4   Consistency with economic approach

Although travel time is the main factor in network usage, there are obviously other factors that also influence the development of transport networks. A typical example is an economic approach using objectives such as minimising total costs or maximising social welfare. The question then is whether such an economic approach leads to completely different results for hierarchical networks or yields similar results.
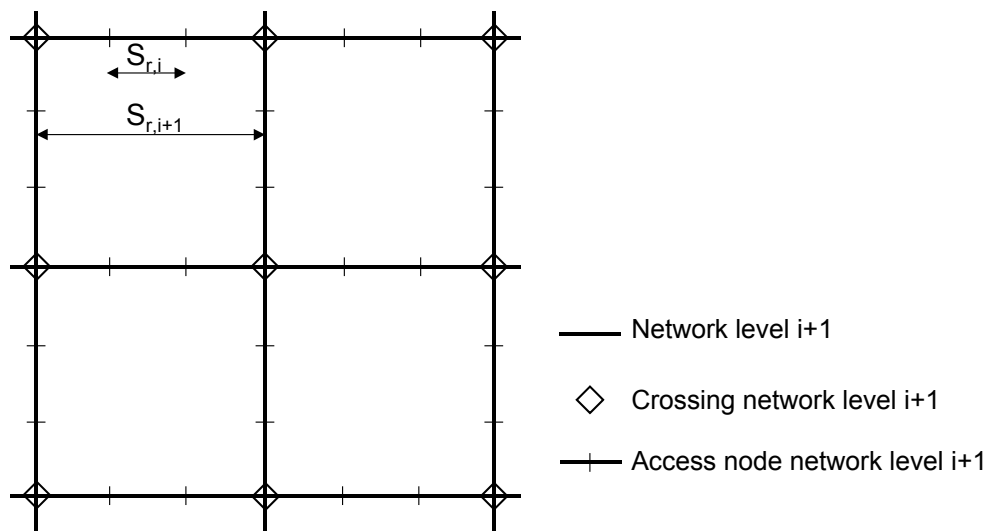


*Figure 5-7: Main characteristics of network level i+1*

From an economic perspective the problem of hierarchical networks can be analysed using the following situation. Given a grid network, for instance of regional roads, having road spacing $S_{r,i}$ and travel speed $V_i$, the question is whether a higher-level network, for instance a motorway network (Figure 5-7), having a road spacing $S_{r,i+1}$, which is equal to or larger than $S_{r,i}$, and a travel speed $V_{i+1}$, which is higher than $V_i$, should be built or not. In order to answer this question the objective of minimising total costs can be used. Since the lower-level network is given, only the changes in costs have to be taken into account, which are the investor's costs for building and maintaining the higher-level network, and the travellers' benefits due to the reduction in travel times as a result of using the higher-level network. If the benefits outweigh the investment and maintenance costs, the higher-level network should be built, and in all other cases it should not. The value of $Sr,i+1$ for which the net costs are minimal is of course the optimal road spacing for the objective of minimising total costs.

This approach to determine optimal road spacing has earlier been suggested for urban areas. Creighton (1970) determined optimal values for expressways using a graphical approach, since no suitable analytical approach was available to describe the usage of the expressway network. Black (1976) also studied the optimal spacing for urban expressways, using analytical estimates of the flows for each network level given an average trip length. These equations, however, do not account for the difference in quality between the network levels. The following sections present an analytical model for determining the optimal scale-factor for the road spacing of interurban networks, which accounts for the quality of both network levels and explicitly considers the trip length distribution.

### 5.4.1   Investment and maintenance costs

For a unit area the investment costs are determined by the road length or road density ($D_{r,i+1}$), the number of crossings of the higher-level network or crossing density ($D_{c,i+1}$), and the access density ($D_{a,i+1}$). These densities can be calculated as (see Figure 5-7 and Appendix C):

$$D_{r,i+1} = \frac{2 \cdot S_{r,i+1}}{S_{r,i+1}{}^2} = \frac{2}{S_{r,i+1}} \tag{5-3}$$

$$D_{c,i+1} = \frac{1}{S_{r,i+1}{}^2} \tag{5-4}$$

$$D_{a,i+1} = \frac{D_{r,i+1}}{S_{r,i}} - D_{c,i+1} = \frac{2}{S_{r,i} \cdot S_{r,i+1}} - \frac{1}{S_{r,i+1}{}^2} \tag{5-5}$$

The investment costs per square kilometre can then be calculated by multiplying these densities by the costs for a kilometre road of the higher-level network, the costs for junctions, and those for crossings respectively. Table 5-4 gives an overview of cost factors recommended for the Netherlands. The investment costs should be amortised over a period of 30 years using a discount rate of 4%, yielding an annual payment ($c_a$) of 5.8%. The annual maintenance costs are estimated as 1.5% of the investment costs ($c_m$).

*Table 5-4: Cost factors for regional roads and motorways (Ministry of Transport, Public Works and Watermanagement (1996))*

|                          | Regional road | Motorway |
|--------------------------|---------------|----------|
| Cost per kilometre [k€]  | 3,750         | 7,500    |
| Cost per junction [k€]   | 2,500         | 15,000   |
| Cost per crossing [k€]   | -             | 7,500    |

*Cost factors for 1995, increased by 10%*

The total annual costs for investment and maintenance of the higher-level network ($C_{im,i+1}$) can then be calculated using:

$$C_{im,i+1} = (c_a + c_m) \cdot (c_{dr,i+1} \cdot D_{r,i+1} + c_{dc,i+1} \cdot D_{c,i+1} + c_{da,i+1} \cdot D_{a,i+1}) \qquad (5\text{-}6)$$

with:

$C_{im,i+1}$ = investment and maintenance costs per year for level $i+1$

$c_{dr,i+1}$ = cost factor for road density for level $i+1$

$c_{dc,i+1}$ = cost factor for crossing density for level $i+1$

$c_{da,i+1}$ = cost factor for access density for level $i+1$

### 5.4.2  Traveller benefits

On the other hand there are the benefits of the new higher-level network. Travel speed increases leading to shorter travel times and thus lower travel costs. Given a trip length $L$ the travel time consists of an access and egress time on a lower-level network, and the travel time on the network itself. The access and egress time on a lower-level network varies depending on the direction of the trip in relation to the orientation of the grid. Holroyd (1967) gives an extensive analysis of this. In this analysis it is assumed that the sum of the access and egress distance ($L_a$) is a function of the road spacing of the network level to be used, for instance in the case of using network level $i$ only:

$$L_{a,i-1} = 0.5 \cdot S_{r,i} \qquad (5\text{-}7)$$

The trip length on network level $i$ ($L_{t,i}$) is then:

$$L_{t,i} = L - L_{a,i-1} \qquad (5\text{-}8)$$

The total travel time in case of using network level $i$ only ($T_{n,i}(L)$) then becomes:

$$T_{n,i}(L) = \frac{0.5 \cdot S_{r,}}{V_{i-1}} + \frac{L - 0.5 \cdot S_{r,i}}{V_i} \qquad (5\text{-}9)$$

For simplicity sake, the check whether the resulting trip length on the higher-level network is negative has been omitted in this equation.

In the case of a new higher-level network the access and egress time consists of two parts: access and egress on network level *i-1* and on network level *i*. The total travel time can then be calculated as:

$$T_{n,i+1}(L) = \frac{0.5 \cdot S_{r,i}}{V_{i-1}} + \frac{0.5 \cdot S_{r,i+1}}{V_i} + \frac{L - 0.5 \cdot S_{r,i} - 0.5 \cdot S_{r,i+1}}{V_{i+1}} \qquad (5\text{-}10)$$

Given a trip length distribution, consisting of different classes having trip length $L_k$ and $P_k$ trips (see Table 5-5), the total travel time for each class with and without the higher-level network can be calculated using Equations (5-9) and (5-10). Multiplying the difference in travel time by the number of trips per year and by the travellers' value of time ($c_t$) gives the travellers benefits for that class. Summation of all classes finally yields the total benefit of the higher-level network:

$$C_{t,i+1} - C_{t,i} = \sum_k \left( c_t \cdot \left( T_{n,i+1}(L_k) - T_{n,i}(L_k) \right) \cdot P_k \right) \qquad (5\text{-}11)$$

*Table 5-5: Trip length distribution for all trips in the Netherlands longer than 20 kilometres (NTS, 1995)*

| Distance class [km] | Average trip length [km] | Number of trips per person per year |
|---|---|---|
| 20-30 | 23 | 58 |
| 30-40 | 33 | 30 |
| 40-50 | 44 | 17 |
| 50-75 | 57 | 23 |
| 75-100 | 82 | 11 |
| 100-150 | 125 | 11 |
| 150-200 | 175 | 4 |
| 200-300 | 225 | 2 |
| >300 | 350 | 0 |

In order to be complete, the costs paid by the travellers should be subtracted from the total benefit, especially the operating costs of the vehicles. There are, however, two reasons to omit these costs from this analysis:

- Since the trip lengths for the alternatives considered, that is using the lower-level network only or also using the higher-level network, will be more or less equal, the difference in the operating costs will be small;

- The operating costs of the vehicles are to a large degree determined by taxes, which are used among other matters to finance the road infrastructure. The objective of minimising total costs thus requires that, if operational costs are considered, the share used to finance road infrastructure should be excluded. The resulting operational costs will probably be small compared to the travel costs based on travel time.

### 5.4.3   Numerical analysis

The equations presented in the previous sections are used to determine the net benefits for a number of cases varying the road spacing of the higher-level network. All parameter

values adopted are representative for the Netherlands. Apart from the cost factors and the trip length distribution presented in Table 5-4 and Table 5-5, it is assumed that:

- The number of inhabitants per square kilometre equals 429 persons, which is representative for the Netherlands;

- The average value of time for trips by car is € 7,75 per hour;

- The road spacing of network level *i* is 10 kilometres (see also Section 5.2);

- The ratio for the travel speed is 1.67, following the line of reasoning presented in Section 5.3.2;

- The average travel speed on the higher-level network is 120 km/h.

The values for the investment and maintenance costs and for the travellers' benefit are calculated for different values of the scale-factor for road spacing. The result is illustrated in Figure 5-8, which is similar to the graphs developed by Creighton (1970) for assessing the expressway spacing in urban areas. As the road spacing of the higher-level network increases, the costs for investment and maintenance decreases. At the same time the savings in travel costs decreases. An optimum for the total costs is found in the case that the road spacing equals 30 kilometres, or put another way, the scale-factor equals 3! Furthermore, it should be noted that in this case the travellers' benefit outweighs the investment and maintenance costs: there is certainly an economic reason to build the higher-level network.
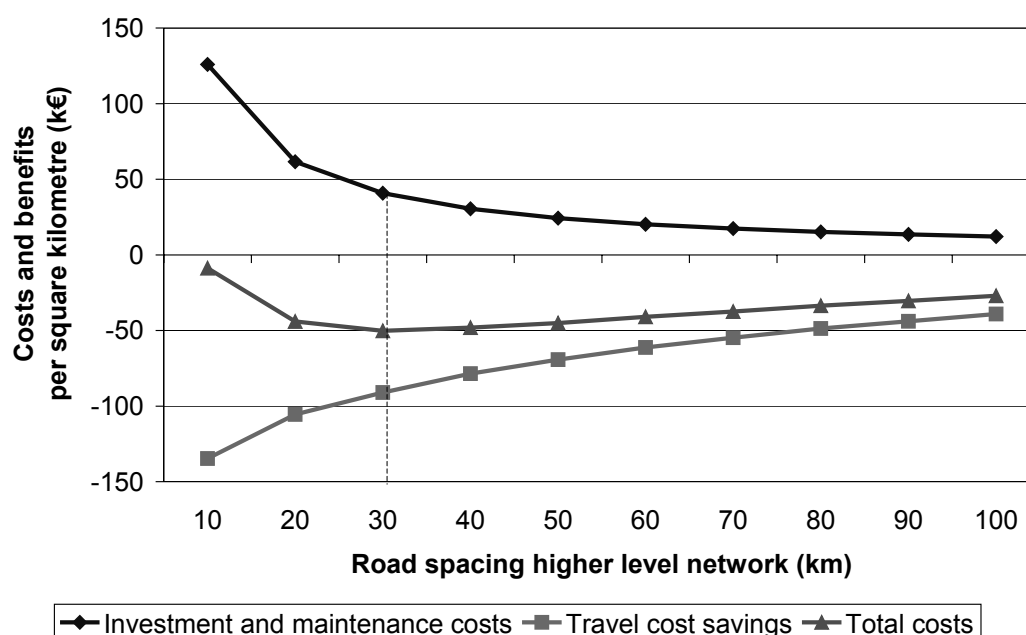


*Figure 5-8: Relationship between various costs and the road spacing of the higher-level network*

In order to illustrate the sensitivity to the assumptions, alternative cases have been analysed varying the travellers' value of time and the travel speed ratio. Changes in cost factors or in population density can be translated into a change of the value of time. The

results of this analysis are shown in Table 5-6. The finding of the scale-factor 3 for road spacing seems to be quite robust. A larger travel speed ratio results in the theoretically expected value for the scale-factor of 2. A lower travel speed ratio, on the other hand, shows in that case that the benefits are too small leading to larger values for the scale-factor.

*Table 5-6: Optimal values for road spacing for network level i+1 under different assumptions*

| Travel speed ratio | Road spacing level *i+1* for different values of time [€/h] [km] | | | Scale-factor for road spacing for different values of time [€/h] | | |
|:---:|:---:|:---:|:---:|:---:|:---:|:---:|
| | **6.50** | **7.75** | **9.00** | **6.50** | **7.75** | **9.00** |
| **1.5** | 50 | 40 | 30 | 5 | 4 | 3 |
| **1.67** | 30 | 30 | 30 | 3 | 3 | 3 |
| **2.0** | 30 | 20 | 20 | 3 | 2 | 2 |

This analysis is, given the subject of this thesis, focused only on passenger transport. Inclusion of freight transport will increase the possible benefits since freight transport has a significantly higher value of time. On the other hand, the reduction of travel time will be less because of speed limitations for freight traffic. It is estimated that given an average level of 10% freight transport for motorways that the extra benefit due to freight transport is 20% to 25%, which yields similar results as the sensitivity analysis above for the highest value of time.

This analysis clearly shows that an alternative approach for road network design that might be used today leads to results that are consistent with those based on the fundamental network characteristic based on minimising shortcuts. This illustrates that alternative approaches that might be decisive at a certain point of time do not have to conflict with the fundamental objective of minimising detours.

## 5.5   Network development strategy

The economic approach discussed in the previous section considers the situation in which it is decided to build a complete network. As such it is clearly a hypothetical situation. It is more likely that higher-level network development is a stepwise process. The question is then whether such a stepwise process leads to different results with respect to the hierarchical network structure.

To answer this question an analysis is made of how a higher-level network might evolve (Van Nes & Van der Zijpp (2000)). The network used is identical to that in Section 4.2.1. The lower-level network has a grid structure. The demand pattern is derived using a direct demand model having equal attractions for all origins and destinations. The network development process is identical: at every step the speed of a single link is

increased. However, instead of selecting the link having the largest flow, the link yielding the largest reduction in the total travel time after its increase in speed is chosen. This approach implies a more sophisticated analysis for which link has to be upgraded to the higher-level network, but the method is still based on a greedy algorithm. Given the many years needed to develop a higher-level network, this assumption is not as unrealistic as it might seem at first sight.

The lower-level network consists of a set of links $X$. The links of the higher-level network can found in the set $Y$. At the beginning the set $Y$ is empty. At every step $j$ an analysis is made of which link $x$ of the set $X$ would yield the largest reduction of the total travel time in the network ($T_n$) if it would be included in the set $Y$:

$$Y_{j+1} = Y_j \cup \mathop{\text{argmin}}_{x \in X} T_n\big(Y_j \cup x\big) \tag{5-12}$$



28 steps

84 steps

104 steps

144 steps
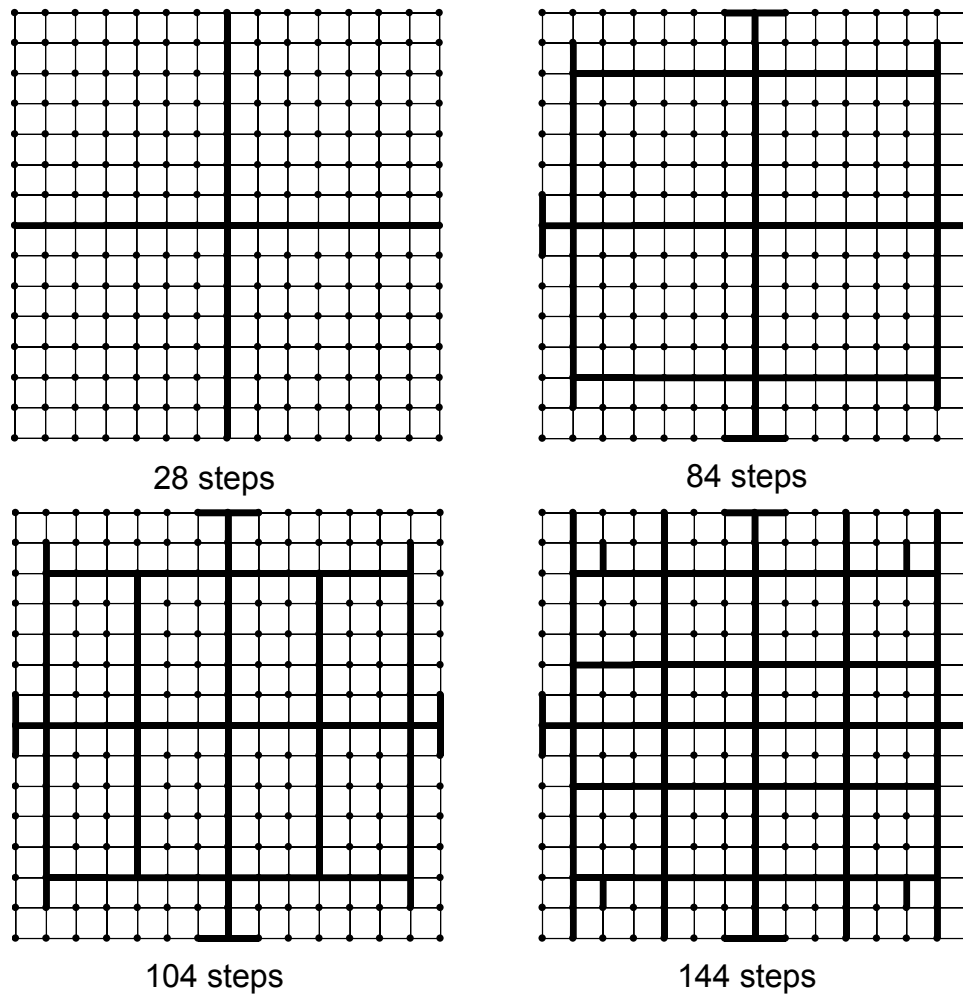
*Figure 5-9: Higher-level network structure for different stages in a stepwise process*

This algorithm is applied to a grid network consisting of 15 by 15 nodes. The network improvement consists of doubling the speed of a link. The results are shown in Figure 5-9. The first higher-level network that appears after 28 steps can be described as a cross,

or, put another way, as a part of grid network having a very large scale-factor for the road spacing. The next structure that emerges (84 steps) is a higher-level network having a grid of 5 by 6 units. The algorithm then constructs short cuts resulting in a grid of 5 by 3 units (104 steps). The next network structure that is found after 144 steps is a grid consisting of two grid types: 3 by 3 and 2 by 3. In the next steps the algorithm develops access roads to the grid structure only, as illustrated in Figure 5-10. The final result of this approach is, of course, the situation that the speed of all links has been doubled. The process of starting with coarse networks, which are refined later on, fits with the graph shown in Figure 5-8 in the previous section: the optimal situation is approached from the right hand-side of the graph.
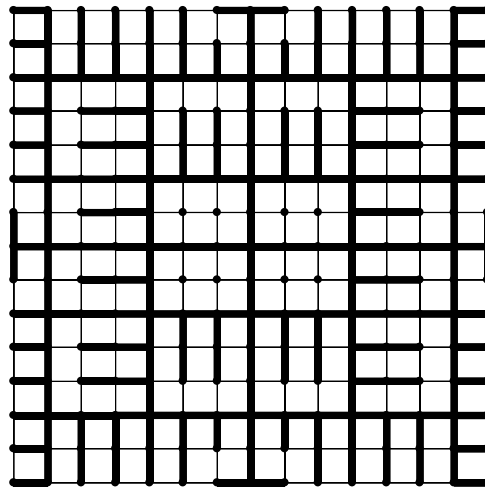


*Figure 5-10: Network structure after 232 steps*

This is, of course, a simple analysis which might be influenced by the assumptions made, especially the increase in speed, the size of the network, and the choice to focus on travel time only.

In case the speed is upgraded using a factor 1.5 instead of a factor 2 a similar pattern evolves, resulting in a smaller higher-level network grid of 2 by 2 units and 2 by 3 units. The smaller increase in speed allows for more shortcuts in the larger coarser grid structures. Other differences with the results presented in Figure 5-9 are:

- Due to the smaller increase in speed it takes slightly longer to develop higher-level grid patterns: more links in the border are upgraded. The effect of bundling on the higher-level network is less strong;

- The first grid pattern of 5 by 6 units is divided asymmetrically into sections of 2 by 5 and 4 by 5 units.

The analysis shows some irregularities in the border of the network. Such irregularities can only be excluded if an infinite network is used. Due to computational limitations of this approach it is not possible to analyse larger networks. The analysis of smaller networks, however, showed identical network structures, which implies that the border effects do not seriously influence the results.

In this approach no attention is given to the investment costs related to upgrading the links from set *X* to *Y*. A possibility is to introduce a minimum criterion for selecting a link, for instance the ratio between the reduction in travel time and the investment costs. Using such a criterion will stop the network development process at a certain point. In the long run, however, demand will grow, either due to the improved network or due to other influences, and the development process will continue again. Neglecting investment costs is therefore not a serious limitation in this analysis.

The main point in this analysis is that a simple strategy for network improvement results in distinct network structures for the higher-level network. Furthermore, it shows that an incremental network development strategy is also consistent with the fundamental network characteristic of minimising shortcuts.

## 5.6   A hierarchy for private transport networks

The analysis of mechanisms that determine hierarchy in private transport networks has shown that simple rules exist that define the relationships between network levels. In this section these rules will be used to present a classification of road network levels. Next, this classification will be compared with existing hierarchies in Dutch road network. Furthermore, some implications for future developments of private transport will be discussed.

### 5.6.1   Classification of private transport network levels

Using the principles found in this chapter a new classification for road network levels can be defined (Table 5-7). The road spacing is based on the findings of De Jong (1988a) and De Jong & Paasman (1998), since they are in agreement with the range of the minimal distances between settlements for each level defined in Table 4-3, and match existing networks. The access spacing is based on the scale-factor 3 for road spacing. The speed is determined using the maximum speed for the national motorway network and the scale-factor 1.67 for speed. Finally, it should be noted that the numerical values in this classification should be interpreted as average values and not as rigid standards.

This classification has two interesting consequences:

- The lack of higher-level networks;
- The two functions of each network level.

Given the principles established in this chapter the national network level should have a road spacing of 100 kilometres, an access spacing of 30 kilometres, and a network speed of approximately 170 to 200 kilometres per hour. It might be questioned whether such high speeds are realistic without substantial changes in the private transport system, such as the introduction of automated vehicles.

*Table 5-7: Classification of road network levels*

| Network level | Spatial level | Road spacing [km] | Access spacing [km] | Speed [km/h] |
|---|---|---|---|---|
| **Urban** | | | | |
| **Street** | Neighbourhood | 1 | 0,3 | 20 |
| **Arterial** | District | 3 | 1 | 35 |
| **Expressway** | 'City' | 10 | 3 | 55 |
| **Interurban** | | | | |
| **Local** | Village | 3 | 1 | 35-40 |
| **Regional** | Town | 10 | 3 | 60-70 |
| **Interregional** | City | 30 | 10 | 100-120 |
| **National** | Agglomeration | - | - | - |
| **International** | Metropolis | - | - | - |

Another consequence of the proposed classification is that the interregional network fulfils two functions. It connects the main cities, but due to the connections with the lower network it also provides access to the main cities. This implies that the mechanism described in Section 5.2.2 is relevant. If motorways are used to access cities that are located closer than 10 to 20 kilometres, the relatively high quality enjoyed by these travellers will influence their travel behaviour, leading to a relatively large growth of these trip types in quantity as well as trip lengths. These changes in travel behaviour might ultimately lead to congestion, and thus to a loss of quality for the trips between cities, which were the main trips to be accommodated.

### 5.6.2 Comparison with existing network levels in the Netherlands

The highest existing network level in the Netherlands is the interregional network level, which serves the 40 main urban areas, that is cities having more than 70.000 inhabitants. As stated earlier, road networks serving agglomerations or metropolises do not yet exist. This implies that there might be a need to develop a new higher network level to connect the main agglomerations. It is interesting to note that in practice the motorway network connecting the main agglomerations in the Randstad has, due to the high level of congestion, a lower quality than the motorway network outside the Randstad. This is exactly the opposite of what is required according to the classification in this thesis.

Please note that in this discussion capacity plays an important role. In the analysis in this chapter it was explicitly assumed that sufficient capacity would be provided. In practice this requirement was met when the motorways were built, however, the increasing demand with respect to volume and trip length requires larger capacities, or, which is more in line with the findings in this chapter, the development of a higher level network. The high costs involved with these measures are obviously too high to avoid any congestion at all. This implies that there will always be a discrepancy between the

theoretical classification and the actual situation. On the long run, however, the fundamental mechanisms underlying this classification will generally prevail.

An example of a suggestion for higher level networks in the Netherlands are the high-speed roads between the main agglomerations suggested by Werkgroep '2duizend' (1999). In fact, the distinction between the main routes to and from the mainports and the other roads also suggests a need for a higher network level. Another suggestion in that direction is the proposal to develop a 'doorstroomroute' along the A4-motorway in the Netherlands: a dedicated route for long-distance traffic having an average access spacing of 10 kilometres meant to avoid delay due to congestion on the A4-motorway (NEI & DHV (2000)).

Such a new network level, however, should have an access spacing of 30 kilometres, and a network speed of approximately 170 to 200 kilometres per hour. These requirements certainly exceed the qualities suggested in some of the proposals mentioned. The 'doorstroomroute', for instance, should be seen as restoring the original function of the interregional network and not as the introduction of a new network level. If the economic approach described in Section 5.4 is used to assess the potential of such a new network, assuming that the costs involved are twice as high as those for the national road network, it is found that no net benefit results. Apparently, there is no justification yet for such a network.

The consequences of combining several network levels on the motorway network can clearly be seen near the main agglomerations, where congestion is common phenomenon. The A4-motorway near The Hague, for instance, has an average access spacing of 4 kilometres, which is nearly equivalent to the access spacing of the local network. Such an access spacing is clearly too short for the interregional network level.

Other observations that can be made given this classification are:

- Higher-level road networks having a rectangular grid, such as the regional network in the Netherlands are very susceptible to shortcuts (see Figure 5-2), which might lead to higher network densities than required. Grid network structures are therefore preferable.

- It has already been concluded that the Randstad lacks a regional network (Hilbers et al. (1997)). The hierarchy concept established in this chapter support the proposal of Immers et al. (2001) to invest in a regional road network for the Randstad to reduce congestion on the interregional network.

- The speed differences between the subsequent network levels as defined by the German design guidelines (FGSV (1988), Section 5.2.2) are too small. Larger differences based on scale-factors for speed, which range between 1.67 and 2, are required for a more efficient use of the network levels. This implies that for a specific situation a subset of network levels should be used, which given the characteristics of a specific level should match the scale-factors defined in this chapter.

- The scale-factor 3 for road spacing is also applicable for bicycle paths, as can be seen in Bach (1999). If the distinction in spatial levels is used, however,

the values for the path spacing are increased. For the lower-level network it becomes 300 metres instead of 150 to 200 metres and for the higher-level network 1 kilometres instead of 450 to 600 metres.


## 5.7  Conclusions

This chapter discussed the concept of hierarchy in private transport networks, that is transport networks in which different network levels can be distinguished each of which is suited for specific trip types and having its own network characteristics in terms of space accessibility, network density, and network speed. Each network level has its own transport function and provides access to higher network levels. The main question studied, is whether there are fundamental characteristics that determine the hierarchy of network levels. Private transport networks have developed over time and are influenced by many factors. A leading motive found in network design methodologies today, is the function of road networks, which is determined by the hierarchy in spatial structures. The influence of characteristics of transport networks themselves, however, is rarely considered.

Even though road network development has been influenced by many factors, a morphological analysis of the Dutch road network showed that there is a simple rule that describes network hierarchy. Every new network level has a road spacing three times as large and a network density that is three times smaller.

The analysis in this chapter clearly shows that this scale-factor 3 for road spacing and network density can be explained by a simple but fundamental network characteristic: minimising shortcuts. This explanation is based on fundamental characteristics of travel behaviour and economic decision making. Travellers maximise their utility, which in the case of route-choice can be seen as minimising travel costs, while basic economic principles require that a higher-level network is only justified when it minimises the use of the lower-level network. The role of the lower-level network then is limited to accommodating short travel distances and providing access and egress to the higher-level network.

The concept of minimising shortcuts also provides a scale-factor for the network speed. The difference in speed should be larger than 50% for triangular networks and larger than 67% for grid networks. The maximum difference required is 100% for both triangular and grid networks. These differences are larger those found in the German design guidelines (FGSV (1988)).

This chapter also shows that two alternative approaches to network development, such as an economic approach based on minimising total costs, or a network development strategy in which that link is improved which yields the largest reduction in travel time, lead to identical scale-factors for road spacing. This finding might seem surprising at first sight, but both approaches are based on travel behaviour characteristics, in which minimising travel time is an important objective. Given the robustness of the derived scale-factor for road spacing it is concluded that private transport networks have self-

organising properties, which should be taken into account in transport network design. This implies that private transport network structures are robust with respect to multimodal transport.

Finally, it is interesting to note that the scale-factor 3 for road spacing fits perfectly with the hierarchy in spatial structures as defined by De Jong & Paasman (1998). The scale-factor 3 for the radius of the settlements per level (see Section 4.3.2) also leads to a scale-factor 3 for the distances between settlements of subsequent levels (see Figure 5-11). A functional approach will thus also lead to a scale-factor 3. The only discrepancy is that given the theoretical distribution in space this approach would lead to a triangular network instead of a grid network. Such a theoretical allocation of settlements, however, is very unlikely in reality, given the stochastic element in settlement locations and the high costs of triangular networks.
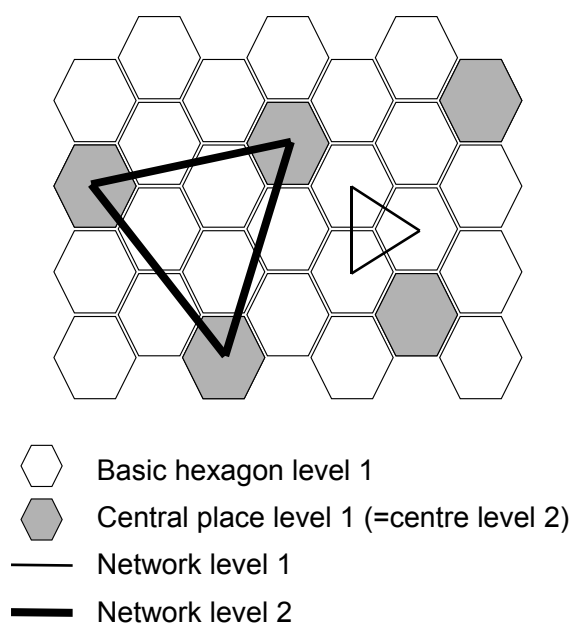


*Figure 5-11: Relationship between hierarchy in spatial structures and hierarchy in road networks*

# Chapter 6

# LINE-BOUND PUBLIC TRANSPORT SERVICE NETWORKS

## 6.1 Introduction

This chapter analyses the characteristics of hierarchical transport networks of line-bound public transport services. In contrast to private transport, issues such as transfers and time-accessibility are essential elements in transport network design. Transport network design thus concerns all network characteristics: that is space-accessibility, time-accessibility, network density, and network speed.

Another difference compared to private transport is the introduction in the analysis of operational costs. When public transport services use the same infrastructure as private modes, infrastructure costs are ignored, making operational costs the critical factor from the operator's point of view. In the case of dedicated infrastructure such as for rail services, however, both infrastructure and operational costs are included in the analysis.

The lower-level private transport networks are primarily determined by functions other than transport, thus allowing the analysis to focus on higher-level transport networks. For public transport service networks, however, transportation is always the primary function. All network levels are therefore relevant in the analysis, although the influence of trip length on multimodal trip making still justifies special attention to higher-level networks. A second argument for considering lower-level public transport networks, is the importance of these networks as access and egress modes, especially for the activity based end of the trip. This function implies that a discussion of lower-level public transport networks is also included in this chapter.

For line-bound public transport service networks a distinction is made between urban and interurban networks. In the case of urban networks the demand pattern is more or less continuously distributed while for interurban networks the demand pattern is determined by settlement structures, and is therefore discretely distributed. A second characteristic

following from this distinction is the network structure usual for each type. Urban networks generally have radial network structures or consist of a set of urban corridors. For interurban networks, however, area covering network structures such as grid and triangular networks are often more suitable, even though the TGV-network in France also has a radial structure.

The purpose of this chapter is similar to that of the previous chapter: to establish relationships that determine the emergence of hierarchy in line-bound public transport service networks. First a description is given of existing public transport networks and planning guidelines. This overview will show that hierarchy is a common characteristic of public transport service networks. Next, the analysis focuses on the emergence of hierarchy. This analysis is divided into three parts.

The first part deals with urban public transport networks (Section 6.3). An analytical approach for single-level transport network design is introduced, which will be extended later to analyse multilevel urban public transport networks. This analytical model consists of models describing the level of service, the demand and supply components, as well as the design objectives. Furthermore, the results of an application of the analytical model will be used to discuss the common objections against adopting findings following from analytical models. This discussion will show that although analytical models are based on important simplifications, the results are still useful and worth considering.

The second part of the analysis deals with multilevel urban public transport networks (Section 6.4). The analytical model presented for the design single-level networks will be extended to multilevel network design for analysing hierarchical public transport network structures under different assumptions with respect to city size and demand pattern. Since all these analytical transport network design models are based on the theoretical framework presented in Chapter 3, it is possible to make a systematic comparison of commonly found hierarchical network structures. Furthermore, an approach similar to the economic approach for private transport networks will be used. These analyses will show that public transport systems as such do not possess characteristics that lead automatically to hierarchical network structures. Hierarchy in urban public transport networks only makes sense if there is a hierarchy in spatial structures.

The third part focuses on interurban public transport networks (Section 6.5). Special attention will be given to the (dis-) similarities with the mechanisms for hierarchy in private transport networks. This part will also show that the hierarchy in spatial structures is decisive for the hierarchy in public transport networks. Another important aspect analysed is the dependency between consecutive network levels, especially in the case of different operators for these network levels. Combining game theory and analytical transport network design models, it will be shown that the phenomenon of different operators has interesting consequences for the financial relationships between network levels.

The findings on the mechanisms determining hierarchy will be used to propose a classification for public transport networks. This classification is compared to the public transport system in the Netherlands, which leads to suggestions for future developments of the Dutch public transport system. Finally, the chapter concludes with a summary of

the main conclusions with respect to hierarchy in line-bound public transport service networks.


## 6.2   Public transport networks in retrospect

This section presents the characteristics of public transport networks and their hierarchy, with emphasis on the Netherlands. Furthermore, an analysis is made of existing planning guidelines for the Netherlands and for other countries, giving some insights into the way the design guidelines may have influenced existing network structures.

The focus on the Netherlands is of course not fully representative for the development of public transport networks worldwide. Comparison with historical developments in Europe and the United States (White (1995), Black (1995)), for instance, shows that the Netherlands were relatively slow in adopting new, but expensive, transport technologies such as railways and metro systems. However, the main characteristics and relationships discussed are certainly valid for public transport networks in general.


### 6.2.1   Public transport in the Netherlands

One of the oldest kinds of public transport is the stagecoach. In the Netherlands, however, the track boat has been more important for many centuries. The quality of the roads was poor, the western and more densely populated part of the Netherlands had a reasonable network of waterways, and water-borne public transport was much cheaper than by stagecoach. In the 17$^{th}$ century there was already an extensive network of water-borne public transport services. In 1830 there were about 800 departures per week from Amsterdam serving from 120 to 180 destinations (Knippenberg & de Pater (1988)).

In the beginning of the 19$^{th}$ century people started to consider public transport by water as old-fashioned, while at the same time the possibilities of public transport by road improved. It was the railways that really became an alternative to public transport by water. Although the first railways in the Netherlands were developed with a focus on freight transport, it was soon discovered that passenger transport could be very profitable. Railways allowed substantially higher travel speeds than public transport by water: 20 to 30 km/h instead of 7 to 8 km/h. It is interesting to note that due to the focus on freight transport, and of course the lower investment costs, many railway stations were located outside the cities (Van der Woud (1987)). A phenomenon that can still be seen in our cities today: many railway stations are located adjacent to the city centres.

At the end of the 19$^{th}$ century, the tram became an interesting transport mode for urban and regional transport, at first drawn by horses, then propelled by steam and later by electricity. The popularity of train and tram had a large impact on road transport, as can be seen in the reduction of revenues from toll roads. From 1930 on, however, buses became more popular. They were more flexible and were considered to be a more modern mode of transportation. Furthermore, private car became a realistic alternative to public transport. More people could afford a car and the quality of the road network

improved. As a result the tram network was reduced to urban networks and many regional tramways were closed. The railway network today has therefore a large resemblance to the railway network at the end of the 19th century. The total length of the railway network in 1980 was 2.500 kilometres (passenger transport only) which is equivalent to a network density of 0.07 km/km$^2$, comparable to the density of the motorway network in the Netherlands. Figure 6-1 shows the railway network for the Netherlands, with a distinction between lines used for long distance travel services (Intercity network) and by other rail services.
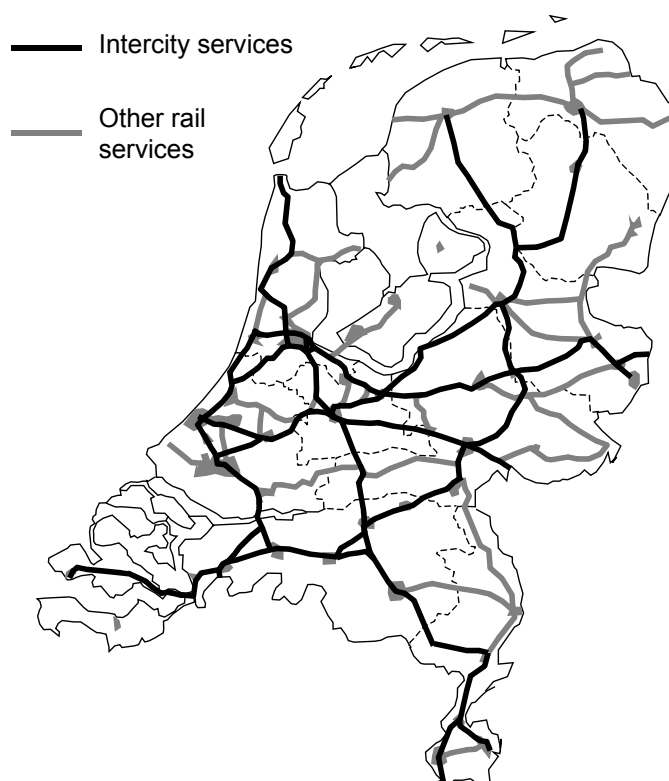


*Figure 6-1: Railroad network in the Netherlands and their use by train services*

Public transport services today are offered at different network levels which differ with respect to trip length but also with respect to organisation: namely national railways, regional bus, and urban public transport. The national Dutch Railway Company (NS) distinguishes three network levels:

- Intercity services offering connections between the 25 main cities in the Netherlands. The stop spacing is about 30-40 kilometres;

- Interregional train services (Express trains) for cities having 50.000 inhabitants or more. The stop spacing is approximately 20 to 30 kilometres. Outside the Randstad area the Interregional train services coincide with Intercity services;

- Local train services especially suited for trips with lengths between 5 and 30 kilometres, which operate in urban agglomerations or in rural areas. The stop spacing is 2 to 10 kilometres.

For regional bus three service levels can also be distinguished:

- Interliner, high quality express services which are comparable with local train services, and sometimes also with interregional train services. The stop spacing is about 5 kilometres and in urbanised areas 1 to 2 kilometres;

- Express services offering direct services to and from cities, often limited to peak hours;

- Local services, having a stop spacing of 1 to 2 kilometres and in urbanised areas about 400 metres.

In urban public transport networks above a certain city size, two network levels can be distinguished (see Section 6.4.1 for hierarchical network structures in urban agglomerations worldwide):

- Express services such as metro or light rail systems, having a stop spacing of 600 to 800 metres;

- Local services, usually bus and tram systems, having a stop spacing of 400 metres.

From a functional point of view it appears that urban agglomerations are supported by three service levels, while interurban transport shows four levels (see Table 6-1). Since the difference between Interregional train services and Intercity services is small, it is arguable whether Intercity services are really an additional network level.

*Table 6-1: Functional classification of the public transport system in the Netherlands (based on various schedules of transport service operators)*

| Travel distance | Urban | | | Interurban | | |
|---|---|---|---|---|---|---|
| | Network | Stop spacing [km] | Freq. [veh/h] | Network | Stop spacing | Freq. [veh/h] |
| 0 - 8 km | Bus/tram | 0.4 | 4 - 8 | Local bus | 1 – 2 (0.4) | 1 - 2 |
| 3 - 10 km | Metro/ light rail | 0.6 - 0.8 | 6 - 12 | | | |
| 5 - 30 km | Local train/ light rail | 3.0- 5.0 | 4 - 8 | Local train/ light rail/ Interliner | 3 - 5 (1 – 2) | 2 - 4 |
| 30 - 100 km | | | | Interregional/ Intercity | 20 – 30 | 2 |
| 30 - 300 km | | | | Intercity (Randstad only) | 30 – 40 | 2 |

Van Goeverden & Schoemaker (2000) present an overview of public transport service network characteristics for the main network types in the Netherlands. Due to the focus of their study, however, the distinction between network levels is primarily based on

technology and organisation and not on functional characteristics. Nevertheless, their overview provides some interesting findings. Using theoretical network structures, such as triangular and radial networks, the main network characteristics were determined that match more aggregate data found in maps and annual reports of various transport operators. Table 6-2 shows the main network characteristics derived from their study. A distinction is made between public transport in urban agglomerations and interurban public transport services. For both categories, scale-factors are derived for the higher network levels.

*Table 6-2: Network characteristics for public transport service networks in the Netherlands (based on Van Goeverden & Schoemaker (2000))*

| | Network type | Stop spacing [km] | Line spacing / radius [km] | Opera-tional speed [km/h] | Average frequency [veh/h] | Network density [km/km$^2$] | Access density [/km$^2$] |
|---|---|---|---|---|---|---|---|
| **Urban agglomerations** | | | | | | | |
| Urban bus | Grid | 0.4 | 1.5 | 23 | 4.3 | 1.733 | 3.889 |
| Urban tram | Radial | 0.4 | 5.0 | 18 | 9.7 | 0.659 | 1.660 |
| Metro | Radial | 0.9 | 7.5 | 35 | 9.8 | 0.223 | 0.253 |
| Urban rail | Radial | 3.8 | 14.0 | 65 | 4.2 | 0.150 | 0.041 |
| **Scale-factors** | | | | | | | |
| Metro/urban bus | | 2.3 | | 1.5 | 2.3 | 7.8 | 15.4 |
| Metro/urban tram | | 2.3 | 1.5 | 1.9 | 1.0 | 3.0 | 6.6 |
| Urban rail/metro | | 4.2 | 1.9 | 1.9 | 0.4 | 1.5 | 6.2 |
| **Interurban services** | | | | | | | |
| Regional bus | Triangular | 0.8 | 10.0 | 38 | 2.0 | 0.450 | 0.577 |
| Regional rail | Triangular | 6.3 | 50.0 | 70 | 2.2 | 0.076 | 0.011 |
| National rail | Triangular | 24.0 | 80.0 | 95 | 2.2 | 0.045 | 0.002 |
| **Scale-factors** | | | | | | | |
| Regional rail/regional bus | | 8.3 | 5.0 | 1.8 | 1.1 | 5.9 | 51.2 |
| National rail/regional rail | | 3.8 | 1.6 | 1.4 | 1.0 | 1.7 | 7.3 |

As expected, in all cases the higher-level networks have larger stop spacing, higher operational speed and lower densities. It is interesting that the difference in densities between bus systems and rail systems is relatively large. This might be due to the influence of infrastructure costs for rail services. Furthermore, the changes in operational speed are similar: the scale-factors vary between 1.4 and 1.9. The values for the frequencies on the other hand, do not change very much. For interurban transport the average frequency is 2 services per hour, while for urban transport services the frequency is 4 or 10 services per hour. Finally, it is interesting to note that the scale-factor for the

network density for regional and national rail is relatively low. This is probably caused by the reduction of the rail infrastructure in the middle of the 20th century, leading to a heavily used infrastructure.

### 6.2.2    Guidelines for public transport service network design

The hierarchical structure network levels present in practice can also be found in planning guidelines. The following discussion on guidelines for public transport service network design is divided into three parts:

- Hierarchy in public transport service networks;
- Urban public transport networks;
- Interurban public transport networks.

#### 6.2.2.1  *Hierarchy in public transport service networks*

Hierarchy in public transport networks is a common phenomenon in Dutch recommendations for network planning as can be seen in Table 6-3, which shows travel distances for each network level.

*Table 6-3: Travel distances per network level for public transport networks according to Dutch planning recommendations*

| Network level / Spatial level | Immers & Egeter (1996) | Raad voor Verkeer en Waterstaat (1996) | Van den Heuvel (1997) |
|---|---|---|---|
| **Local** | - | 1 – 3 | - |
| **Agglomeration** | - | 3 – 10 | 5 - 25 |
| **Regional** | 8 - 30 | 10 – 30 | 25 - 40 |
| **Interregional** | 30 - 80 | 30 – 100 | 40 -100 |
| **National** | 80 - 300 | 100 – 300 | 100 - 300 |
| **International** | > 300 | 300 - 1,000 | > 300 |

The names used to identify the network levels are based on the spatial level they are related to. The classification of Immers & Egeter (1996) is very similar to that of the Dutch Advisory Council for Transport, Public Works and Watermanagement (Raad voor Verkeer en Waterstaat (1996)), while the classification of Van den Heuvel (1997) differs in distinguishing two network levels for trips between 30 and 100 kilometres. The arguments supporting these categorisations, however, are limited. Only Van den Heuvel (1997) provides an analysis leading to his classification, which is based on minimising weighted travel time given a fixed operations budget. It should be noted, however, that using his results a classification in three levels, in which the National level replaces both the Interregional and International level, might be just as good.

Finally, it is interesting to note that the classification of the Raad voor Verkeer en Waterstaat bears a strong resemblance to the hierarchy in road networks as defined by De Jong & Paasman (1998) (see Section 5.2.1). As shown in Section 5.7 the classification of De Jong & Paasman fits the hierarchy in spatial structures. As such, it is possible to use the hierarchy in spatial structures to determine the hierarchy in public transport services, an approach to public transport network planning advocated in Germany (VÖV (1981), Köhler (1989)). Bierschenk & Keppeler (2000)) analysed transport networks in Germany and found that public transport service networks should have a similar hierarchy as proposed for road networks (FGSV (1988)), that is, a network hierarchy based on the hierarchy in spatial structures.

### 6.2.2.2  *Urban public transport service networks*

The guidelines for urban public transport network design focus mostly on values for stop spacing or maximum access distances. Table 6-4 shows general guidelines published for the Netherlands. Just as with the guidelines on hierarchy in public transport networks, arguments supporting these values are missing. Generally, these values seem to be based on common practice. The finding of Egeter (1993) that the stop spacing in urban public transport networks should be about 600 metres, for instance, is still not accepted in planning practice, even though it was included in advice on public transport network planning (Projectbureau IVVS (1995)). Some explanations for this phenomenon will be discussed later on in Section 6.3.5.

*Table 6-4: Guidelines for urban public transport networks (Bach (1999))*

|  | Stop spacing [m] | Maximum access distance [m] | Average speed [km/h] | Trip lengths [km] |
|---|---|---|---|---|
| **Urban bus** | 300 - 500 | 400 | 12 - 20 | 2- 6 |
| **Urban tram** | 300 - 500 | 400 | 12 - 20 | 2 - 6 |
| **Light rail tram** | 400 - 700 | 600 | 18 - 25 | 4 - 10 |
| **Metro** | 700 - 1,400 | 800 | 30 - 40 | 5 - 14 |
| **Urban rail** | 1,500 - 2,000 | 1,000 | 40 - 50 | 7 - 20 |

An overview of guidelines for the maximum access distance to bus stops is given by Mensebach (1994). The values vary between 160 metres to 1,000 and even 1,500 metres, the most common values being 400 to 600 metres. The arguments for these values are unclear. The only exception is the lowest value of 160 metres, which is based on a study on the reduction in patronage as a function of the access distance for the city of Bielefeld (Walther (1973)). This study illustrates a general belief that access distance is a decisive factor in choosing public transport, independent of other characteristics. Many studies, however, have shown that the door-to-door travel time is the key variable in explaining traveller behaviour (Van der Waard (1988), Central Transportation Planning Staff (1997)). The value of 160 metres based only on access distance seems too short-sighted, and therefore not realistic.

Vuchic & Musso (1991) give corresponding guidelines for metro systems. Metro stations should be located at large squares, at railway stations and at intersections with other public transport services. Following common practice in other cities, the stop spacing should be between 500 and 800 metres in city centres (see Paris, Hamburg and Philadelphia) and 1,000 to 3,000 metres for outlying suburbs (London, Moscow, and San Francisco). For the two metro-systems in the Netherlands, stop spacing is about 1,000 metres in Rotterdam and 750 to 800 metres in Amsterdam.

### 6.2.2.3  Interurban public transport networks

The guidelines for interurban public transport service networks discussed in this section focus on the design of a network given a specific network level. Two approaches can be distinguished.

1.    Network design given a specific geographic situation;

2.    Feasibility analysis of specific network structures.

Immers et al. (1994) and Immers & Egeter (1996) present a design methodology for regional and national public transport networks, consisting of four-steps:

- Definition of parameters and performance criteria;

- Selection of access nodes;

- Design of a minimal network that connects all access nodes and fulfils the constraints with respect to the circuity;

- Final adjustment of the network.

Table 6-5 shows the criteria used in the design of the national and regional public transport service networks in the Netherlands. The optimal stop spacing is derived using a simple model for the trade-off between access time and in-vehicle time. Given the importance of walking and cycling for accessing the train network, however, this approach is questionable (see also Section 6.5). There are several options for the criterion to select and sort settlements: city size, number of inhabitants, number and type of urban functions, number of trips to and from the city having a minimal trip length, et cetera. For the national network, for instance, the number of trips longer than 80 kilometres was used. All other criteria are mostly based on planning practice.

The network design procedure itself can be characterised as developing a kind of minimum spanning tree, which is adjusted to satisfy the circuity criterion for the main movements that have to be served by the network. As a result the network density will be rather low, thus allowing for higher frequencies. Two aspects are explicitly not dealt with in this design methodology, that is the budget for operating the services and the existing infrastructure for public transport services. The final design is considered to be the optimal network from the traveller's point of view, which then is used to confront such constraints from the operator's or authority's perspective.

*Table 6-5: Criteria for interurban public transport service networks defined by Immers et al. (1994) and Immers & Egeter (1996)*

|  | **National network** | **Regional network** |
|---|---|---|
| **Travel distance** | 80 - 300 km | 8 – 30 km |
| **Optimal stop spacing** | 30 - 40 km | 3 – 6 km |
| **Number of access nodes** | 25 to 45 | Depends on the area |
| **Size of service areas** | 6 to 8 km | 4 to 5 km |
| **Selection/sorting criterion** | Service level | > 10,000 inh. |
| **Maximum access distance** | 40 km | 5,000 inh. < 20 km |
| **Maximum circuity** | 1.5 | 1.5 to 2.5 |

Frank et al. (1994) describe an alternative design approach. They start with a specific network structure for a regional public transport service network, for instance, a grid network with or without diagonals. Furthermore, they assume that the demand is concentrated at the nodes of the network. For each network structure considered the minimum level of demand per node is determined that is required to meet the official German standards with respect to cost efficiency. Their analysis for Germany showed that for nearly half of Germany other means of public transport were necessary.

### 6.2.2.4   Conclusion

A general conclusion from this survey is that current guidelines on public transport service network planning are primarily based on past practice. The arguments supporting these guidelines are limited, and design objectives are rarely mentioned. Exceptions are studies by Egeter (1993), Immers et al. (1994) and Van den Heuvel (1997), who focus on minimising travel time on a fixed budget. The budget itself, however, is not related to the performance of the public transport network. The hierarchical concepts advocated in Germany (VÖV (1981), Köhler (1989), Schönharting (1997), Bierschenk & Keppeler (2000)) are based on hierarchy in spatial structures.

In the next sections, the arguments for hierarchical network structures for both urban and interurban public transport networks will be analysed using clear design objectives. First, however, the single-level urban public transport network design problem is discussed, presenting among other matters the analytical framework used in the analyses in this chapter.

## 6.3   Single-level urban public transport networks

The design of single-level public transport network systems is not the main topic in the thesis but is included for three reasons. First, it introduces an analytical approach to the public transport service network design problem, which will be extended later to hierarchical public transport networks. Second, it illustrates the gap between theory and

practice, or in other words, the robustness of planning practice with respect to theoretically based planning guidelines (see also Section 6.2.2.2). And third, urban public transport networks play an important role in multimodal transport as access modes to higher level networks.

The analytical model for urban public transport network design is based on the perspective of transport network design as a Stackelberg game: that is the network design influences the demand, while the demand influences the network design. The analytical model consists of the following components (see Table 3-8):

- Level of service: travel time elements;

- Demand side: patronage, revenues, consumer surplus, and travel costs;

- Supply side: operational costs and producer surplus;

- Design objectives: maximising social welfare, and minimising total costs.

The following sections will give a detailed description of these components and their application. Most of this section is based on Van Nes & Bovy (2000) and on Van Nes (2000) which gives an extensive review of the subject.

### 6.3.1   Analytical model

An analytical model is formulated for an urban corridor in which parallel public transport lines offer transport services to the city centre (Figure 6-2). Given the radial structure of many urban public transport networks it might seem unrealistic to consider a linear network instead. However, radial networks can often be seen as a set of corridors and it can be shown that the results for a corridor are also valid for a radial network structure (Van Nes (2000)).
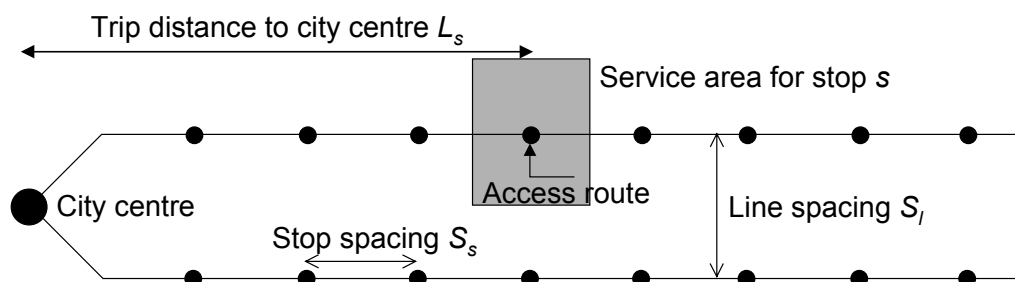


*Figure 6-2: Layout of the public transport network in an urban corridor*

The analytical model is used to determine optimal values for the decision variables stop spacing $S_s$, line spacing $S_l$, and frequency $F$. The design objectives used are maximising social welfare and minimising total costs. As such the analytical model presented in the following sections fully meets the description of the network design problem given in Section 3.4.5 and Table 3-8. The decision variables determine the quality offered in terms of travel times to the city centre. Similarly, the decision variables determine operational costs, and if relevant also investment costs. In the case of maximising social welfare, travel times influence the number of travellers, which determines the revenues. Please

note that the fares are assumed to be fixed. Revenues and operational costs determine producer surplus or profit, which together with consumer surplus yields social welfare. In the case of the objective of minimising total costs, however, the patronage is assumed to be fixed, that is independent of the quality of the services offered.

Theoretically, the influence of the decision variables on all travel characteristics involved varies with the location of the traveller in the corridor. As such the travel time should be determined for a small area within the service area of a given stop, having a specific level of demand. The total travel time following from a specific network design is then the aggregation of travel times for all those small areas per stop for all stops, multiplied by the corresponding levels of demand. Such a detailed approach, however, is beyond the scope of this thesis. It would introduce a level of detail that is inconsistent with arguments supporting the use of analytical models, while making the design problem unnecessary complicated. Therefore, the following assumptions are introduced:

- The stop spacing is assumed to be constant;

- The service area per stop is assumed to be a rectangle having a uniformly distributed demand. In that case the average access time to the stop can easily be calculated using the central point of the service area (see Equation 6-2). It can be shown (see Appendix D) that in this case the access time is slightly overestimated since shorter access times, that is areas close to the stop, will yield a higher patronage than long access times, locations farther away. This effect is partially compensated by the fact that there are more small areas farther away than close to the stop. The overestimation of the access time will lead to a slight underestimation of the impact on the patronage;

- The relationship between travel times and demand is modelled for the average perceived travel time in the corridor, instead of modelling it for each stop separately. It can be shown (see Appendix D) that this assumption will lead to a slight overestimation of the patronage, as short trips by public transport are generally less attractive than longer trip lengths. The impact of this assumption will decrease if the average travel time is based on trips having a minimal trip length.

Since the impacts of these assumptions are small, and some assumptions even have opposite effects with respect to the estimated level of demand, the net effect on the findings in this chapter will be negligible. Introducing this level of detail thus makes the analysis unnecessary complex. Finally, it can be noted that these assumptions are similar to those used in other analytical models for urban public transport networks (e.g. Kocur & Hendrickson (1982), Tsao & Schonfeld (1984), Chang & Schonfeld (1993b), Spasovic et al. (1994)).

### 6.3.1.1   *Travel time*

The perceived travel time of a trip to the city centre from a service area of a stop $s$ at a distance $L_s$ consists of access time, waiting time, in-vehicle time, and egress time. Transfers are not relevant for a trip to the city centre. Weights are used to account for the fact that travellers have different valuations for the different parts of the trip.

$$T_c(L_s) = w_a \cdot T_a + w_w \cdot T_w + T_i(L_s) + w_e \cdot T_e \tag{6-1}$$

where:

$T_c$ = total weighted travel time for a trip to the city centre

$T_a$ = access time

$T_w$ = waiting time

$T_i$ = in - vehicle time

$T_e$ = egress time

$w_k$ = weight for time element $k$

Stop spacing and line spacing jointly determine the access time:

$$T_a = \frac{f_a \cdot f_d \cdot (S_s + S_l)}{V_a} \tag{6-2}$$

where:

$f_a$ = access routing factor

$f_d$ = demand distribution factor

$V_a$ = access speed

In case of access routes parallel and perpendicular to the lines $f_a$ equals 0.25, while a uniformly distributed demand yields a value for $f_d$ of 1.

The waiting time is determined by the frequency:

$$T_w = \frac{f_w}{F} \tag{6-3}$$

where:

$f_w$ = factor for the waiting time

A value of 1,800 seconds for $f_w$, for instance, results in a waiting time of half the headway, which is suitable for frequent services as in urban public transport networks.

The in-vehicle time is determined by the average travel distance, the maximum speed and the time lost at each stop:

$$T_i = \frac{L_s}{V} + \frac{L_s}{S_s} \cdot T_s = \frac{L_s}{S_s} \cdot \left( \frac{S_s}{V} + T_s \right) \tag{6-4}$$

where:

$L_s$ = average travel distance from stop $s$ to city centre

$V$ = maximum speed

$T_s$ = time lost at a stop

Since the focus is on trips to the city centre, it is assumed that the egress time $T_e$ is fixed. Substitution of Equations (6-2) to (6-4) into Equation (6-1) then yields:

$$T_c(L_s) = w_a \cdot \frac{f_a \cdot f_d \cdot (S_s + S_l)}{V_a} + w_w \cdot \frac{f_w}{F} + \frac{L_s}{S_s} \cdot \left(\frac{S_s}{V} + T_s\right) + w_e \cdot T_e \qquad (6\text{-}5)$$

The average perceived travel time to the city centre then is the weighted average of the travel time of the passengers for all stops:

$$T_c = \frac{\sum_{s=1}^{n_s} T_c(L_s) \cdot P(L_s)}{\sum_{s=1}^{n_s} P(L_s)} \qquad (6\text{-}6)$$

where:

$P(L_s) = $ passengers boarding at stop $s$ at distance $L_s$

$n_s \quad = $ number of stops

This equation can easily be simplified to:

$$T_c = w_a \cdot \frac{f_a \cdot f_d \cdot (S_s + S_l)}{V_a} + w_w \cdot \frac{f_w}{F} + \frac{L_c}{S_s} \cdot \left(\frac{S_s}{V} + T_s\right) + w_e \cdot T_e \qquad (6\text{-}7)$$

where:

$$L_c = \text{average travel distance to the city centre}: \left(\frac{\sum_{s=1}^{n_s} L_s \cdot P(L_s)}{\sum_{s=1}^{n_s} P(L_s)}\right)$$

Please note that this derivation is independent of the level of demand that is assumed for each stop. Equation (6-7) can thus be used for any distribution of the demand along the line.

### 6.3.1.2  *Patronage*

In general, travellers are sensitive to the quality of the services offered. This applies to captives as well as to choice travellers. High quality results in high patronage and vice versa: higher patronage leads to higher frequencies which make the service more attractive. Higher patronage might be due to shifts in destination choice, longer trips due to people travelling to the city centre instead of to a local shopping centre, or in mode choice, choosing public transport instead of private car or bicycle. In the analysis it is assumed that modal split has the largest influence on the patronage. Only in the case of large changes in the quality of the services offered this assumption should be reconsidered. The following analyses, however, will show that the changes found for the public transport network structures do not require a more detailed analysis including destination choice models.

From a theoretical point of view, mode choice is influenced by three components: weighted door-to-door travel time, travel costs, and travel comfort. Since the focus is on the design variables stop spacing and line spacing, it is assumed that there will be no differences between trip alternatives with respect to fares or comfort. As a result, the description of the relationship between supply and demand will be limited to weighted travel time only. Furthermore, the sensitivity for changes in the transport system will vary between populations. Some populations might be considered to be captive to public transport, that is, they will chose public transport independent of the quality offered, while other travellers are very sensitive to changes in the public transport network. In this analysis, however, an average sensitivity will be used which accounts for different types of travellers.

A logit-choice model is chosen to describe public transport choice as a function of perceived travel time:

$$P(T) = P_0 \cdot \frac{\exp(-\alpha \cdot T)}{\exp(-\alpha \cdot T) + \sum_{m=1}^{n_m} \exp(-\alpha_m \cdot T_m)} \qquad (6\text{-}8)$$

where:

$P_0$ = total travel demand for transport per square kilometer in trips

$\alpha$ = demand sensitivity to public transport quality

$n_m$ = number of modes excluding public transport

$\alpha_m$ = demand sensitivity for mode $m$

$T_m$ = weighted travel time for mode $m$
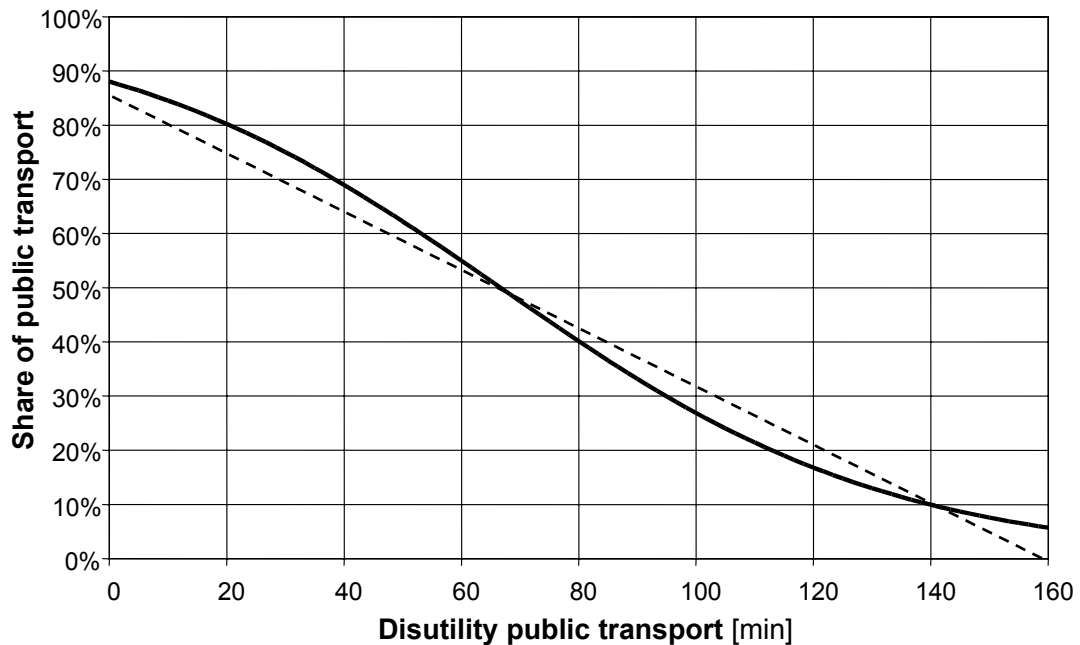


Figure 6-3: Share of public transport as function of the weighted travel time using a logit-mode-choice model given a fixed quality for the other modes

Figure 6-3 shows the share of public transport, that is $P(T_c)/P_o$, as a function of the weighted travel time using a logit-choice model. The dashed line shows the simplification to a linear relationship that is often used instead for analytical purposes (see for instance Kocur & Hendrickson (1982), Chang & Schonfeld (1993b), Spasovic et al. (1994), and Chang & Yu (1996)).

### 6.3.1.3   Revenues

The revenues for a public transport operator are determined by the fares paid by the travellers and the subsidy. The patronage determines the revenues from the traveller $R_t$:

$$R_t = r_t \cdot P(T_c) \text{ or } R = r_{tk} \cdot L_c \cdot P(T_c) \qquad (6\text{-}9)$$

where:

$\quad r_t \quad$ = fare paid by the traveller

$\quad r_{tk} \quad$ = fare per kilometre paid by the traveller

The subsidy $R_s$ might be a fixed amount or might be related to the patronage:

$$R_s = s \text{ or } R_s = s_t \cdot P(T_c) \text{ or } R_s = s_{tk} \cdot L_c \cdot P(T_c) \qquad (6\text{-}10)$$

where:

$\quad s \quad$ = fixed subsidy

$\quad s_t \quad$ = subsidy per traveller

$\quad s_{tk} \quad$ = subsidy per kilometre per traveller

### 6.3.1.4   Consumer surplus

The consumer surplus is the traveller's component in the objective of maximising social welfare. It represents the benefits of travellers who can make their trip with lower costs or shorter travel times compared to their maximum acceptable travel costs or travel time (Jansson (1996)). The consumer surplus is defined as the integral of the demand function from $T_c$ to infinity, but this leads to complicated formulations. A commonly used alternative is to use a linear approximation to the logit-mode-choice model, which reduces the integral to the area of a triangle. The disadvantage, however, is that the non-linear characteristics of the demand function are no longer accounted for.

Therefore, a new approximation approach is used in which the demand is modelled using the logit-mode-choice model, while the consumer surpluses are calculated as the surface of the grey triangles in Figure 6-4. Due to economic conventions the axes have been switched in comparison with the graph for the logit-mode-choice model as depicted in Figure 6-3.
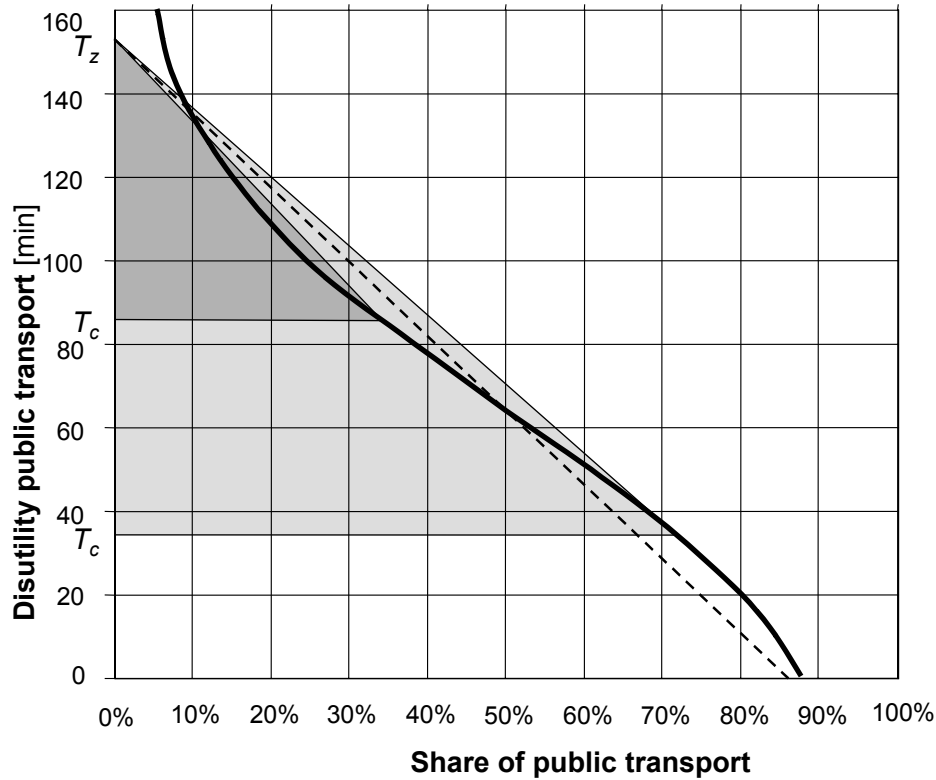
*Figure 6-4: Consumer surplus assuming a linear relationship between supply and demand*

Since the relationship between supply and demand is assumed to depend only on perceived travel time, the vertical axis represents weighted travel time. In order to monetise travel time, it should be multiplied by the average value of time. Again, the value of time might vary between populations. In fact, it might be expected that new travellers opting for public transport have a higher value of time. In this analysis, however, an average value of time is used for all travellers. Theoretically, fares should also be included in the definition of consumer surplus. However, since fares are assumed to be constant they are dropped in these analyses.

The formula for the consumer surplus $CS$ for a given perceived door-to-door travel time $T_c$ becomes:

$$CS(T_c) = 0.5 \cdot P(T_c) \cdot (T_z - T_c) \cdot c_t \qquad (6\text{-}11)$$

where:

$T_z$ = travel time where the demand for public transport vanishes

$c_t$ = value of time for passengers

The value of $T_z$ in these analyses is defined as the travel time for which the linear approximation of the logit-mode-choice model equals zero.

### *6.3.1.5   Travel costs*

The traveller's component in the objective of minimising total costs are the travel costs, that is the product of travel time and the number of passengers monetised using the passenger's value of time, plus the fares paid by the passengers:

$$C_t = c_t \cdot T_t \cdot P + R_t \tag{6-12}$$

As a matter of fact the travel costs can be regarded as being complementary to consumer surplus, since the travel costs are equivalent to the surface of the rectangle under the grey triangles in Figure 6-4.

### *6.3.1.6   Operational costs*

Finally, the operational costs $C_o$ have to be determined. Since, the demand is expressed per square kilometre, the operational costs should also be related to a unit area of a square kilometre. The operational costs are determined by total driving time of the vehicles within the unit area. This time depends on the frequency, the number of lines per unit area, the driving time of one vehicle within the unit area, and the fact that the line operates in two directions:

$$C_o = (c_o + c_{im}) \cdot F \cdot \frac{1000}{S_l} \cdot \frac{1000}{S_s} \cdot \left( \frac{S_s}{V} + T_s \right) \cdot 2 \tag{6-13}$$

where:

   $c_o$   = operational costs, that is, driver and vehicle costs per vehicle hour

   $c_{im}$   = investment and maintenance costs per vehicle hour

Please note that in case of a bus system the factor $c_{im}$ will be equal to zero, while for rail services the investment and maintenance costs will have a substantial influence.

### *6.3.1.7   Objective functions*

Maximising social welfare can be written as the sum of consumer surplus (Equation (6-11)) and producer surplus, that is revenues minus operational costs (Equations (6-9) and (6-13)):

$$MAX\{SW\} = MAX\{CS(T_c) + R_t - C_o\} =$$
$$MAX \left\{ \begin{array}{l} 0.5 \cdot P(T_c) \cdot (T_z - T_c) \cdot c_t + \\[2mm] r_t \cdot P(T_c) - (c_o + c_{im}) \cdot F \cdot \dfrac{1000}{S_l} \cdot \dfrac{1000}{S_s} \cdot \left( \dfrac{S_s}{V} + T_s \right) \cdot 2 \end{array} \right\} \tag{6-14}$$

Please note that subsidies are not included in this formulation, since these should be regarded as additional costs for the authorities.

The alternative objective of minimising total costs can be formulated as the sum of the traveller's costs and the operational costs. Since the fares paid by the traveller reduce the costs for the operator, they can be excluded from the total costs for travelling, which then becomes:

$$MIN\{C\} = \min\{C_t + C_o\} =$$

$$MIN\left\{c_t \cdot T_c \cdot P + (c_o + c_{im}) \cdot F \cdot \frac{1000}{S_l} \cdot \frac{1000}{S_s} \cdot \left(\frac{S_s}{V} + T_s\right) \cdot 2\right\} \qquad (6\text{-}15)$$

Please note that in the latter objective a fixed level of demand is used instead of using a function to describe the patronage. This is a more robust formulation for the objective of minimising total costs, since it is possible to minimise the total costs by reducing patronage. In fact, the trivial solution for this objective is to offer no public transport services at all, thus having no operational costs and reducing the patronage and consequently the travel costs to zero.

### 6.3.2   Solving the model

The objective function of maximising social welfare is difficult to solve analytically but can easily be solved numerically. The shape of the objective function given a fixed frequency is shown in Figure 6-5.
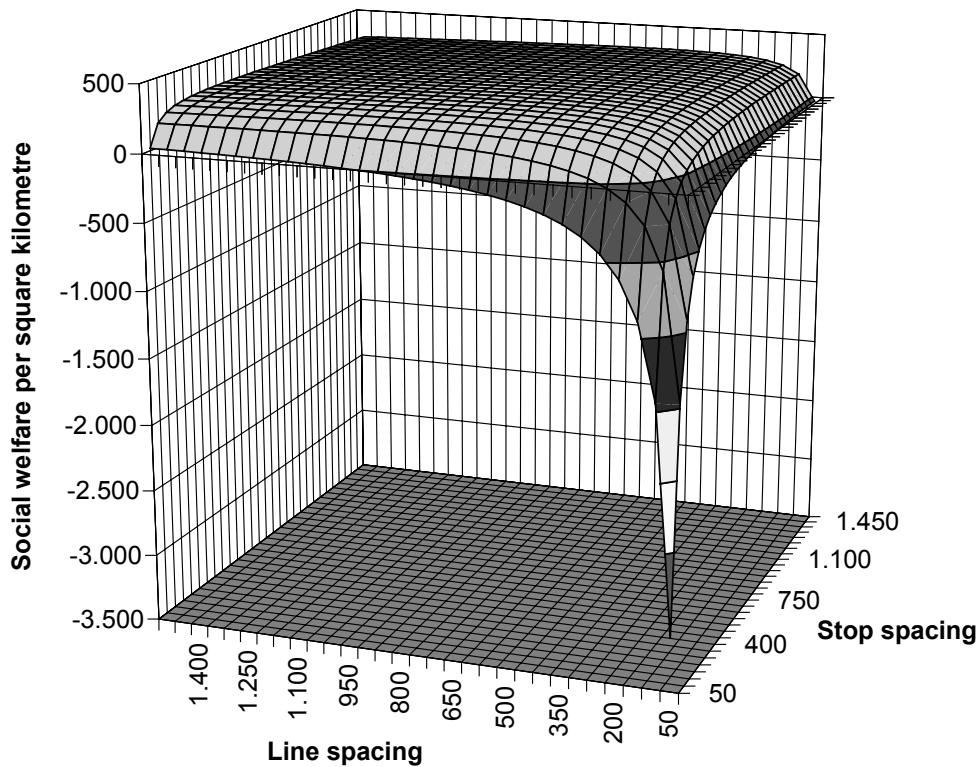


*Figure 6-5: Maximising social welfare*

It clearly shows that there is a large range of values for the stop and line spacing where the values of the objective function are similar. This is the same phenomenon as shown in the analysis of the optimal road spacing for a new road network level in Section 5.4.3. It interesting to note that, given the values used in this example, there is a net benefit for urban public transport.

The alternative objective function of minimising total costs can easily be solved analytically by setting the derivatives with respect to the decision variables stop spacing, line spacing and frequency equal to zero. This leads to the set of equations given in Equation (6-16), which can be solved using a Gauss-Seidel iteration scheme.

$$S_s^* = \sqrt{\frac{\lambda}{\kappa} \cdot \left( \frac{\rho \cdot F^*}{S_l^*} + L_c \right)} \qquad (6\text{-}16a)$$

$$S_l^* = \sqrt{\frac{\lambda}{\kappa} \cdot \left( \frac{\rho}{S_s^*} + \frac{\rho}{\lambda \cdot V} \right) \cdot F^*} \qquad (6\text{-}16b)$$

$$F^* = \sqrt{\frac{\tau}{\lambda} \cdot \left( \frac{1}{\dfrac{\rho}{S_s^*} + \dfrac{\rho}{\lambda \cdot V}} \right) \cdot S_l^*} \qquad (6\text{-}16c)$$

where:

$$\kappa = \frac{w_a \cdot f_a \cdot f_d}{V_a}$$

$$\lambda = T_s$$

$$\rho = \frac{1000^2 \cdot (c_o + c_{im}) \cdot 2}{P \cdot c_t}$$

$$\tau = w_w \cdot f_w$$

The square roots in the formulae and the fact that the decision variables influence each other show, again, that the sensitivity of the optimal values is limited. The optimal stop spacing increases with access speed, time lost per stop, the operational and investment costs, and trip length respectively, while it decreases with the access factor and the weight for access time. Similar relationships apply to the optimal line spacing, with the distinction that the trip length does not directly influence the optimal line spacing, and that the influence of the time lost per stop is less. Furthermore, the optimal line spacing decreases if the maximum speed increases. The optimal frequency increases as the weight of waiting time, the maximum speed and the time lost at stops increases, and decreases if the operational and investment costs increases.

The intermediate variables in Equation (6-16) can be seen as the main factors with respect to spatial access ($\kappa$), time lost at stops ($\lambda$), costs ratio between supply and demand ($\rho$), and time accessibility ($\tau$). These factors will discussed in more detail in the analysis of the sensitivity of this approach (Section 6.3.4).

### 6.3.3   Application of the model

The derived models are used to determine realistic optimal values for the decision variables stop spacing, line spacing and frequency. The results will be compared to the characteristics of urban public transport networks in the Netherlands today, leading to recommendations for planning practice. Since analytical models are clearly a gross simplification of reality, this application of an analytical model for the single-level urban public transport networks is also used to discuss the sensitivity of the results (Section 6.3.4) and to discuss commonly used objections against adopting these results in practice (Section 6.3.5).

Two typical situations have been analysed. The first is a bus network comparable to that in the neighbourhoods Zuilen and Overvecht in the city of Utrecht in the Netherlands, while the second is a tram network based on characteristics of the southern part of The Hague in the Netherlands. The network characteristics of these networks are presented in Table 6-6.

*Table 6-6: Network characteristics of traditional bus and tram networks*

|  | **Bus** <br> e.g. Utrecht Zuilen/Overvecht | **Tram** <br> e.g. Southern part of The Hague |
|---|---|---|
| **Corridor length** [km] | 5 | 7 |
| **Stop spacing** [m] | 350 | 400 |
| **Line spacing** [m] | 550 | 1,000 |
| **Peak frequency** [veh/h] | 5 | 8 |

The parameters values used in this analysis are given in Appendix E. The level of demand is chosen such that the occupancies of bus and tram near the city centre are more than 80%. The linear approximation of the logit mode choice model (Equation 6-8) for the objective of maximising social welfare is based on the 10th and 90th percentile. The traveller's weighting of the travel time elements is based on Van der Waard (1988), even though these values are based only on route choice analysis. Comparison with other studies on travel behaviour (e.g. DHV (1994), Hague Consulting Group (1997), Yai et al. (1997), Central Transportation Planning Staff (1997), Nielsen (2000)), and with other studies on single-level urban public transport design (Van Nes (2000)), shows that these values are representative for describing average travel behaviour. The derivation of the cost factors can be found in Appendix F. These values are based on a detailed analysis of operational and infrastructure costs based on Van Goeverden & Schoemaker (2000), and are slightly lower than those used in earlier studies (Van Nes (2000)).

The results of these applications are shown in Table 6-7. Included are the optimal values for the decision variables stop spacing, line spacing, and frequency, the related densities, the values of the objective functions, and the consequences for travel time, patronage, operational costs, and profit. Please note that for the objective of minimising total costs the level of demand is assumed to be fixed, that is, equivalent to the level of demand in the traditional situation.

*Table 6-7: Optimal network characteristics for the selected objectives for a bus network and a tram network respectively*

| | Bus (trip length 3 km) | | | Tram (trip length 5 km) | | |
|---|---|---|---|---|---|---|
| | Traditional | Max. Social welfare | Min. Total costs | Traditional | Max. Social welfare | Min. Total costs |
| **Stop spacing** [m] | 350 | 640 | 641 | 400 | 757 | 762 |
| **Line spacing** [m] | 550 | 752 | 759 | 1,000 | 797 | 820 |
| **Frequency** [veh/h] | 5 | 8 | 7 | 8 | 7 | 7 |
| **Stop density** [/km$^2$] | 5.2 | 2.1 | 2.1 | 2.5 | 1.7 | 1.6 |
| **Line density** [km/km$^2$] | 1.8 | 1.3 | 1.3 | 1.0 | 1.3 | 1.2 |
| **Social welfare** [€/km$^2$] | 411 | 434 | 430 | 744 | 779 | 772 |
| **Total costs** [€/km$^2$] | 299 | 281 | 276 | 515 | 485 | 481 |
| **Travel time** [min] | 20.8 | 18.2 | 18.8 | 25.1 | 22.8 | 22.9 |
| **Weighted travel time** [min] | 28.2 | 26.7 | 27.5 | 33.5 | 32.3 | 32.5 |
| **Patronage** [/km$^2$] | **100** | 102 | **100** | **150** | 151 | **150** |
| **Revenues** [€/km$^2$] | 34 | 35 | 34 | 85 | 86 | 85 |
| **Operational costs** [€/km$^2$] | 68 | 59 | 51 | 104 | 85 | 83 |
| **Profit** [€/km$^2$] | -34 | -24 | -17 | -19 | 0 | 2 |
| **Total costs per trip** [€] | 2.98 | 2.76 | 2.76 | 3.44 | 3.20 | 3.21 |
| **Total costs per kilometre travelled** [€/km] | 0.99 | 0.92 | 0.92 | 0.69 | 0.64 | 0.64 |

*Note: Exogenous input values are shown in bold print. Adopted parameter values are given in Appendix E.*

Compared to the traditional network structure, the network densities for the optimal bus network are much lower. Stop spacing is twice as large and line spacing is nearly 50% larger. Furthermore, the frequency is substantially higher. These differences imply

shorter travel times, and thus a higher-level of demand, and lower operational costs. Reductions in operational costs range between 13 and 25%.

For the tram network the optimal stop density is significantly lower but the optimal line density is higher than in the traditional network. Stop spacing is, again, nearly twice as large. Line spacing, however, is 20% lower, and the frequency 12%. Improvement of the space accessibility is apparently more important than improvement of time accessibility. Travel times are lower and patronage slightly higher. The reduction of operational costs is substantial: up to 20%.

These results clearly show that traditional network structures can certainly be improved from the viewpoint of maximising social welfare as well as minimising total costs. Both objectives yield nearly identical optimal network structures. The optimal network for the objective of minimising total costs is slightly coarser than the network for the objective of maximising social welfare. With respect to the results for the frequencies, it should be noted traditional networks might focus on matching capacities, and thus frequencies, with the level of demand, instead of considering frequencies as an important network design variable with respect to time accessibility.

Interestingly, the differences between the traditional and the optimal designs vary between bus and tram. For the bus corridor the differences for line density and frequency are clearly larger than for the tram corridor. Apparently, the traditional tram network is already more in line with the design objectives adopted in this analysis, which might be due to the higher costs involved in providing tram services. The finding that the stop spacing should be significantly increased, however, is identical for both corridor types.

These findings suggest that in the design of urban public transport networks stop spacing of 650 to 7500 metres, depending on the average trip length, and line spacing of 750 to 800 metres are appropriate. Frequencies should preferably be in the range of 7 or 8 vehicles per hour. These recommendations differ substantially from the guidelines discussed in Section 6.2.2.2, although some of these recommendations are similar to those from earlier studies, such as the study by Egeter (1993). Since, the fact that these recommendation follow from an analytical approach might be a barrier for them to be accepted in planning practice, the next section discusses the sensitivity for the underlying assumptions. This is followed by an assessment of commonly used objections against these coarser networks.

### 6.3.4   Sensitivity for the underlying assumptions

Given the flatness of the objective functions involved (Section 6.3.2), the sensitivity of these optimal design values to the values of the parameters used in the analysis will be limited. The analytical approach presented in Section 6.3.2 showed six key input factors for the objective of minimising total costs (see the factors $\kappa$, $\lambda$, $\rho$, $\tau$, $V$, and $L_c$ in Equation (6-16)). These are:

- Access factor ($\kappa$), which depends on the traveller's weight for access time, the routing pattern around the stop, the distribution of demand around the stop, and which is inversely proportional to access speed;

- Stop loss factor ($\lambda$), which is equivalent to the time lost per stop;

- Costs factor ($\rho$), depending on the ratio between operational costs, that is the operational costs per vehicle, and traveller costs, patronage and traveller's value of time;

- Waiting factor ($\tau$), which depends on the traveller's weight for waiting time, and the factor for calculating the waiting time;

- Maximum speed for public transport ($V$);

- Travel distance ($L_c$).

Generally speaking the optimal values for stop spacing, line spacing, and frequency are proportional to these key input factors. Exceptions are the access factor for all decision variables, and the maximum speed for line spacing only. Thus a larger stop loss factor implies larger values for stop and line spacing, and a higher weight for access time will lead to lower values for the optimal stop and line spacing.
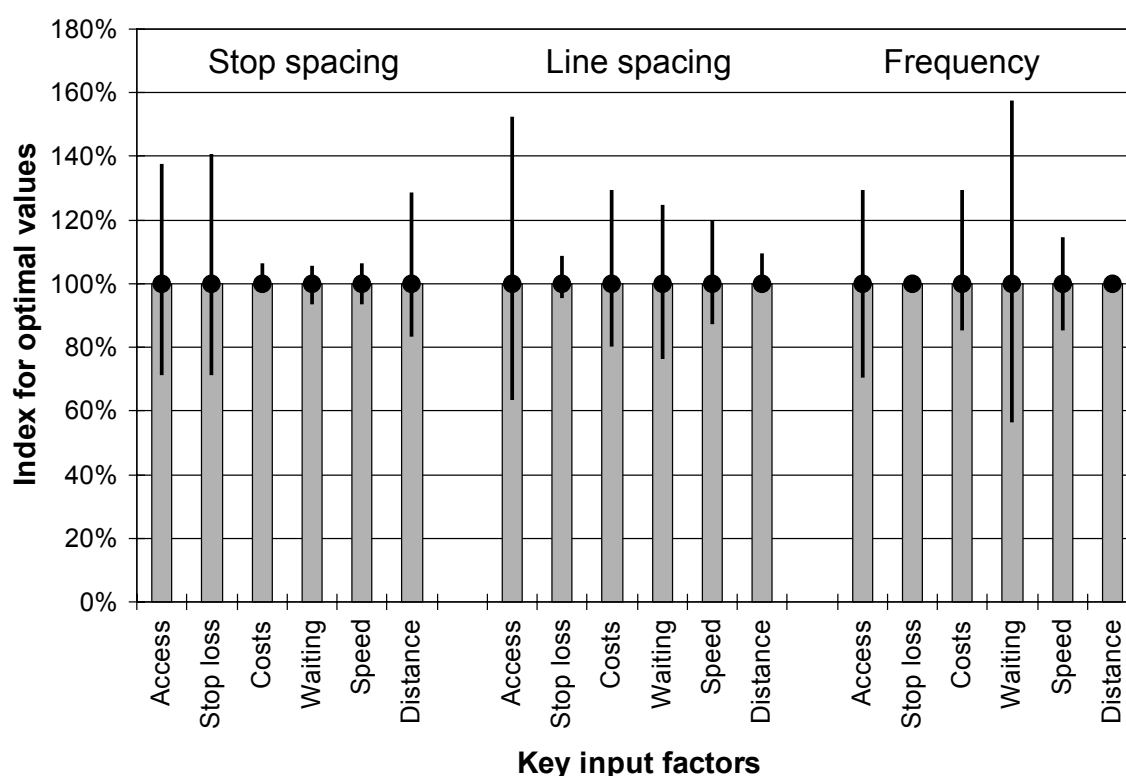


*Figure 6-6: Impact on optimal values for stop spacing, line spacing and frequency of doubling and halving key input factors for the objective of maximising social welfare*

Figure 6-6 shows the impact on the optimal values for stop spacing and line spacing for maximising social welfare in the cases that these key input factors are doubled or halved.

Given the square root relationships found for the decision variables the maximum impact would be an increase of the optimal values of 41% or a decrease of 29%. For some key input factors, however, the relationships are not strictly proportional, and, furthermore, the decision variables interact with each other. As a results the deviations are sometimes larger such as for line spacing and the access factor, and for frequency and the waiting factor, but in most case the deviations are smaller than expected given the square-root relationships.

The access factor is clearly most important, since it influences all three decision variables. Time lost at stops and trip length only influence the optimal stop spacing. The cost factor and especially the waiting factor have a substantial influence on line spacing and frequency. The last key factor, maximum speed, has the smallest impact on the optimal values for line spacing and frequency.

The values of the parameters used in the analysis are generally quite conservative (see also Van Nes (2000)). There are three parameters that deserve particular attention. First, the level of demand that is assumed is representative for a peak period. For off-peak periods the demand is lower leading to lower values for all decision variables. Second, the maximum speed of 50 km/h might be too optimistic for urban public transport. Delays at intersections might lead to lower values. This would imply larger line spacing and lower frequencies. Finally, the most arbitrary assumption is about the trip length. The values chosen focus on longer trip lengths, that is, on trips from the outer areas to the city centre. It might be that decision makers consider short trips just as important, leading to a 10% to 15% lower value for the optimal stop spacing.

### 6.3.5   Objections against adopting results from analytical models

The analysis in this chapter clearly suggests that the values for stop spacing that are currently used and also recommended in guidelines are too low. This is also true for the design variables line spacing and frequency in bus networks. The proposed changes in urban public transport network structures lead to shorter travel times and especially to lower operational costs. The conclusion seems obvious: reduce the number of stops, and in the case of bus networks increase the line spacing and the frequencies. Some of these recommendations were even proposed earlier, such as a stop spacing of 600 metres by Egeter (1993). Public transport planners, however, seem reluctant to adopt these new guidelines. The question is why? There are several kinds of objections that might be applicable:

- The analytical model is too theoretical;

- Current guidelines have such a long history, they must be right;

- Reduction of accessibility is critical for public transport usage;

- The situational context is more important than planning guidelines;

- Increasing stop spacing implies eliminating stops, leading to objections from current public transport users.

The theoretical nature of the analytical model has already been discussed in Section 6.3.1. It has been shown, however, that optimising a radial network structure instead of a corridor leads to similar results (Van Nes (2000)). Van Nes furthermore showed that the inclusion of trip types other than trips to the city centre only, might lead to even lower network densities. Trips from the city centre deal with egress time instead of access time, which has a lower weight and makes accessibility less important. Transverse trips lead to longer travel distances, making in-vehicle time more important. Finally, transfers increase the importance of waiting time, and thus of frequency, also leading to lower densities. Tangential trips are the only exception, having short travel distances and where access and egress time are both influenced by stop spacing. In all other cases, however, the results of the modelling analysis are certainly representative for urban public transport networks that are oriented to the city centre.

The long history of current guidelines is a difficult argument to deal with, especially since it is not clear which objectives have been used in the past. What is clear, however, is that the situation was certainly different in the past: smaller cities and thus shorter travel distances, higher demand densities, lower operational costs especially for personnel, lower speeds, and other competing modes. In such a situation higher network densities with respect to stop and line spacing might be justified, suggesting a relatively strong focus on space accessibility. If, for instance, it is assumed that the level of demand is twice as high, while trip lengths, speeds, operational costs and fares are 50% lower, the objective of maximising social welfare yields for bus networks a stop spacing of 390 metres, a line spacing of 550 metres and a frequency of 10 vehicles per hour. Although it is not likely that such an approach has been used in the past, it interesting to note that these optimal values are in line with the currently used values for stop and line spacing. The situation today is obviously different: cities have grown substantially as have travel distances, densities have decreased, personnel costs have increased, and public transport has to compete with other modes such as the private car. Such large changes require a reconsideration of public transport planning practice.

Focusing on stop accessibility only ignores the fact that travellers consider the door-to-door travel time in choosing public transport. Access time to stops is an important travel time element, but not the only one. Offering high accessibility naturally leads to longer in-vehicle times and longer waiting times, or to higher operational costs, all of which should also be considered in public transport network design. The argument that larger access distances would have an unacceptable impact on public transport usage is therefore too great a simplification of the design problem. Furthermore, there is no evidence that travellers have a maximum access distance. The only exception here is a study by Walther (1973). In his analysis, however, it is not clear how far the existing network influences the relationships found: if it is not necessary to walk more than 400 metres to find a bus stop, nobody will walk longer distances. On the other hand, the city of Almere in the Netherlands has a bus network having a stop spacing of 600 metres and a line spacing of 800 metres (De Heij & Maassen (1995)). These network characteristics are clearly an exception but the Almere network has the highest cost efficiency ratio in the Netherlands. Thus the case of Almere shows that the access distances following from the results in Table 6-7 can clearly be acceptable.

The fourth objection that might be raised against reducing network densities is based on the structure of cities today. The location of streets and urban facilities might limit the possibilities to adopt optimal values for stop and line spacing. This might certainly be true with respect to the line spacing. If the scheme for the road spacing of De Jong & Paasman (1998) is extended to urban road networks (see also Table 5-2) the road spacing would be 1,000 metres for arterial streets and 300 metres for collector streets. Line spacing would thus be limited to multiple values of 300 metres only, making it difficult to adopt a value of 750 metres. A similar line of reasoning, primarily based on the assumptions that stops should be located at crossings of collector streets, might be applied to stop spacing. The location of urban facilities, however, should be accounted for as well. Several design studies showed that higher values for stop spacing, and sometimes line spacing too, are possible in existing cities and that the new networks offer the benefits expected (Koot & Govers (1995), Roedoe (1995), Schäffeler (1999)). All three studies concluded that an average stop spacing of 600 metres is possible. Schäffeler also made a design for an average stop spacing of 750 metres, yielding a slightly longer travel time and a further reduction of operational costs. The resulting pattern of stops was irregular compared to the design for 600 metres. The differences for the objectives of minimising total costs were minimal. It can thus be concluded that, although urban structure limits the possibilities for adopting optimal values for stop and line spacing, it is not an argument for persisting in using currently used values.

The last objection is based on the simple fact that an elimination of stops leads to complaints from current public transport users, while the benefits are meant for potential public transport users. Furthermore, local authorities are sensitive for such complaints, and thus limit public transport network changes. This is probably the main reason that larger values for stop and line spacing are only found in new networks, such as in Almere, and for new lines, for instance in The Hague and Rotterdam (Projectbureau TramPlus (1992), Stadsgewest Haaglanden (1999)).

This discussion of the possible objections against adopting larger values for stop and line spacing shows that the environment in which public transport networks are developed is quite conservative. Although most arguments can be countered, the general opinion is that space accessibility is essential and thus that eliminating stops will certainly reduce the number of travellers while the possible benefits are viewed as uncertain. The fact that most objections can be countered, however, implies that even though an analytical approach requires strong simplifications, the results are certainly realistic and useful for planning practice. Analytical models are thus a fruitful approach.

## 6.4  Multilevel urban public transport networks

This section is the second part of analysis in this chapter. It uses the analytical framework presented in the previous section to analyse the possible emergence of hierarchical public transport networks in large urban areas. In order to give some insight into what is meant by hierarchical urban public transport networks, an overview is given of multilevel networks today. Then, the main question for this chapter is formulated, followed by the

analysis of all kind of hierarchical network structures, while accounting for city size, demand patterns, and hierarchy in spatial structures.

### 6.4.1 Multilevel networks today

In many large cities different network levels can be distinguished. If transport technology is used to distinguish different network levels, the lower-level network usually consists of bus or tram systems, while for the higher-level network metro or light rail systems are used. Furthermore, rail systems provide a third level offering access to suburbs and surrounding towns. A typical example of such a hierarchical network structure is shown in Figure 6-7. For simplicity sake, the bus lines in the city centre and in the suburbs and surrounding towns have been omitted.
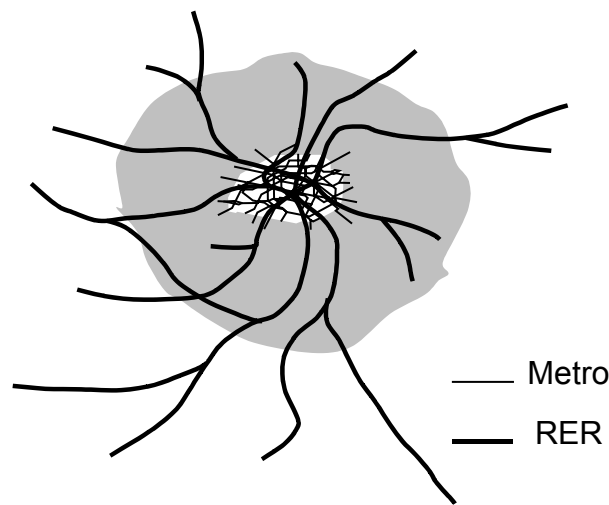


*Figure 6-7: Hierarchy of metro and rail-network for the agglomeration of Paris*

A survey of 11 cities in the world (Table 6-8) shows that a technology based distinction already reveals many of the characteristics expected of multilevel networks:

- Increasing stop spacing, showing a scale-factor of 2.4;

- Increasing line length (defined as network length divided by number of lines), having a scale-factor of 2.2;

- Increasing network speed, having a scale-factor of 1.7 to 2.6.

Interestingly, there is no clear pattern for frequencies: those for level 2 are the highest, while those for level 3 are the lowest, especially in off peak periods. Furthermore, there is a clear distinction between the two scale-factors for the number of lines. Apparently, the infrastructure costs involved with levels 2 and 3 are so high that the number of lines remains nearly constant.

*Table 6-8: Main characteristics for multilevel networks in urban agglomerations around the world based on transport technology (Kwakernaak & Van Nes (2001))*

| | Level 1 | Level 2 | Level 3 | Level 2/ Level 1 | Level 3/ Level 2 |
|---|---|---|---|---|---|
| **Stop spacing** [km] | 0.4 | 0.9 | 2.2 | 2.5 | 2.4 |
| **Network length/line** [km] | 9 | 20 | 43 | 2.3 | 2.1 |
| **Speed** [km/h] | 17 | 30 | 77 | 1.7 | 2.6 |
| **Number of lines** | 30-800 | 9 - 13 | 8 - 11 | 0.1 | 0.8 |
| **Frequency in peak** [veh/h] | 13 | 29 | 10 | 2.2 | 0.3 |
| **Frequency off-peak** [veh/h] | 4 | 12 | 2 | 3.1 | 0.2 |

Furthermore, some typical types of transport services are found in practice (see Vuchic & Musso (1991), Bruijn (1994), Bouman et al (2001)): skip stop services, express services, trunk and feeder services, and zone services (Figure 6-8). Similar concepts can be found in freight transport (see for instance Kreutzberger (1999)).
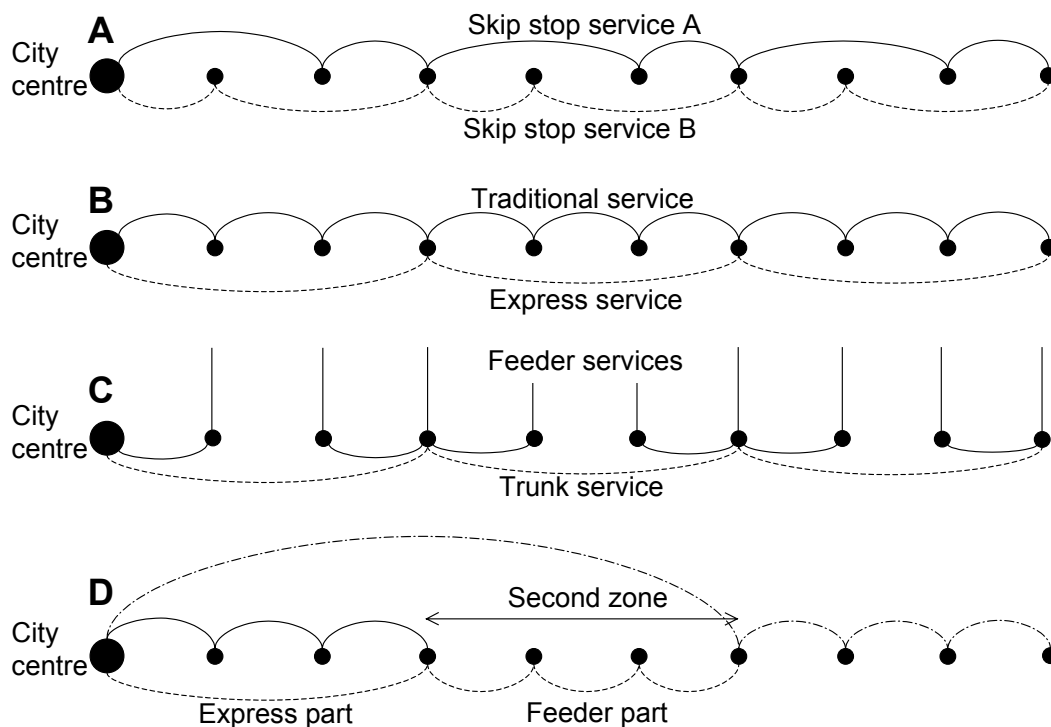


*Figure 6-8: Main types of multilevel urban networks:*
*A: Skip-stop system, B: Express system, C: Trunk-Feeder system, D: Zone system*

Skip-stop service is a service type in which two lines have parallel routes (Vuchic (1973)). Each line serves the main stops whereas the other stops are served by one line only. This concept is especially useful if the line length is too long to offer acceptable travel times if al stops are served by all vehicles. Skipping stops reduces travel time and leads to more balanced occupancy of the vehicles. However, the transport function of

both lines is identical, having an identical line length, frequencies, number of stops, and average speed. The only difference is that the frequency for the main stops is twice as high as the frequency for the other stops. From a functional point of view therefore, skip-stop services are not really a new network level in itself.

Express services are usually found in parallel to more traditional lines. They have a larger stop spacing serving the main stops only, and have a higher average speed. Express lines too, are found in cases where the length of traditional lines become too long to provide attractive travel times. Travellers might use the traditional service to access the express line or they might access the express line directly. In the latter case no transfers are required.

Trunk and feeder services usually consist of a high quality transport service to the city centre (trunk), whereas more traditional transport services offer transport to that trunk line (feeder). This concept is in fact a hierarchical solution to the dilemma of offering short access distances while guaranteeing fast transport to the city centre. The disadvantage, however, is that an obligatory transfer between feeder and trunk service is introduced.

Zone services can be seen as a special case of trunk and feeder services. The term zone is related to the service area that is served by the feeder part on the line. Zone lines also make a distinction between feeder function and transport function but eliminate the transfer. The transport parts of the lines can be bundled, for instance on dedicated infrastructure, and allowing higher frequencies on the trunk. A limitation is that this concept does not allow the introduction of new transport technologies, apart from dedicated infrastructure, which offer higher quality or have lower operational costs per traveller.

### 6.4.2    Research question: emergence of hierarchy

The overview of multilevel networks shows that they are quite common and that there are several alternatives to make such a hierarchy operational. Two studies tried to derive optimal characteristics for hierarchical transport networks. Lesley (1973) developed an analytical model to determine the optimum number of multi-modal interchange stations for a city. In his approach he focused on minimising travel time only, assuming a fixed ratio between network levels for speeds. His results tend to favour a minimum of network levels having a scale-factor for speed of about 3. His analysis, however, does not give any attention to the interests of other parties involved in public transport network design such as authorities, investors and operators. Van den Heuvel (1997), on the other hand, used the operations budget as a constraint, while minimising travel time to determine the number of network levels having fixed qualities. The level of the operations budget is independent of the quality offered. Furthermore, the result of his study is ambiguous with respect to the optimal number of network levels, which might be 3 or 5.

Thus neither study answers the key question in this chapter: which factors determine the emergence of hierarchy in urban public transport networks. This analysis focuses on an urban agglomeration having an average radius of 10 kilometres and maximum radius of

17 kilometres (see Table 4-2). The following sections will explore this question in more detail, using the analytical framework presented in Section 6.3. Basically, there are three differences compared to the urban corridor analysed earlier that might be responsible for the hierarchical structure:

- Longer corridor length;

- Different demand pattern: no longer oriented just towards the city centre;

- Different demand distribution: no longer uniformly distributed, but concentrated at specific locations.

Just as in the analysis of the emergence of hierarchy in private transport networks a minimum of assumptions on demand patterns and spatial structures will be used to determine whether urban public transport networks have some kind of self-organising property or not. Since assumptions on concentrations of demand might easily lead to a kind of self-fulfilling prophecy with respect to hierarchical structures, the analysis starts off considering a uniform demand pattern in a longer corridor (Section 6.4.3). The analysis focuses on the question whether an optimally designed multilevel network is preferable to a single-level network. Next, the impact of different demand patterns is considered (Section 6.4.4), an analysis using the concept of super-positioning higher-level networks presented in Section 5.4. Finally, the consequences of hierarchy in urban spatial structures are analysed (Section 6.4.5). This analysis will show that only a hierarchy in spatial structures provides a rationale for multilevel urban public transport networks.

## 6.4.3   Longer corridor length

In this case, the situation is assumed to be identical to that described in Section 6.3.1: an urban corridor in which parallel lines offer transport to the city centre. The demand pattern is uniformly distributed in the corridor, while only trips to the city centre are considered. Three hierarchical network structures found in current practice are analysed: express services, trunk and feeder services, and zone services.

For each of these network structures an optimal network is designed, which is then compared to a single-level bus network that is optimised for the average trip length to the city centre using the objective of maximising social welfare. The main characteristics of these reference situations are shown in Table 6-9. Compared to the optimal characteristics for a bus network in a 5-kilometre corridor, the travel distances are much larger leading to larger stop spacing, a slightly lower frequency, higher revenues, and higher social welfare. The impact of the trip length is especially strong for the 17-kilometre corridor.

*Table 6-9: Characteristics of the reference single-level bus network for three corridors*

| Corridor length [km] | 5 | 10 | 17 |
|---|---|---|---|
| Average trip length [km] | 3[*] | 5 | 8.5 |
| Stop spacing [m] | 640 | 738 | 905 |
| Line spacing [m] | 752 | 717 | 770 |
| Frequency [veh/h] | 8 | 7 | 7 |
| Weighted travel time [min] | 26.7 | 31.6 | 39.1 |
| Operational costs [€/km$^2$] | 59.1 | 51.2 | 44.3 |
| Profit [€/km$^2$] | -24.5 | 5.5 | 52.2 |
| Total costs [€/km$^2$] | 281 | 309 | 363 |
| Social welfare [€/km$^2$] | 434 | 523 | 691 |

*\* In the analysis in Section 6.3.3 trips shorter than 1 kilometre were ignored*

For the higher-level network it is generally assumed that the maximum speed is 50% higher than for the lower network, with the lowest scale-factor for speeds found for road networks, and yields the highest speed that is realistic for an urban area. Such an increase in speed might be possible if the higher-level network has dedicated infrastructure. Theoretically, there are therefore two possibilities that might be considered: buses having separate bus lanes or light rail systems. However, since the first possibility is clearly the cheapest (see also Appendix F) while the benefit in travel speed is equal, the analysis focuses only on bus systems having separate bus lanes.

The frequencies on the higher-level network need particular attention. The most natural approach would be to make it an additional design variable. However, if under the assumptions used hierarchical network structures are less efficient than single-level networks, the higher-level network should not exist, which is equivalent to a frequency of zero vehicles per hour. Apart from some numerical problems this possibility might lead to, such a result does not provide any insight into the differences between single-level and multilevel public transport networks. In order to achieve that insight, it is assumed that there is a relationship between the frequencies of both network levels. Frequencies of existing higher-level networks are substantially higher than those for the lower networks. However, it is not a priori clear what the explanation for these higher frequencies is. It is possible for instance that this is caused by the high level of demand following from concentrations of demand around the stops of the higher-level systems. This is what will be analysed in Section 6.4.5. In this analysis based on a uniform demand pattern, therefore, it will be assumed that the frequency of the higher-level network is equal to that of the lower-level network, thus assuring that both systems are realistic alternatives. Of course, the impact of both assumptions will be discussed in the following sections.

### 6.4.3.1  Express services

In the case of express services, two service types are available. The lower network is a bus service having many stops, while the higher-level network consists of a parallel bus service having dedicated bus lanes and stopping at a smaller number of stops. These stops coincide with stops of the lower-level network.

It is assumed that travellers choose which service is most suitable for their trip to the city centre, and travel directly to the relevant stop. Since it is expected that the stop spacing of the express system will be too small to make the option of using the lower-level network to access the express service really attractive, the option of transferring between both service types is not considered in the analysis (see also Section 6.4.3.2). The fact that traveller may choose between alternative implies that route choice should be included in the analysis as well, as was discussed in Section 3.4.4. In that case the choice for public transport versus other transport modes is determined by the aggregate quality of the routes that are available. The main extension of the analytical model presented in Section 6.3.1 thus is the incorporation of the route choice model to determine the aggregate quality of the services offered, which then will be used to determine the patronage. As discussed earlier, nested logit-models will be used for this analysis.

In the case of an express service a second transport service is introduced having a larger stop spacing $S_{s,2}$, frequency $F_2$, and maximum speed $V_2$. All other characteristics such as access factor, access speed, and weights for travel time elements, are assumed to be equal. The decision variables are thus the stop spacing, the line spacing, and the frequency for each network level. The line spacing and frequency for both network levels, however, are assumed to be equal.

It is obvious that for short travel distances the lower-level network will be most attractive while for long distance trips the higher-level network will be most attractive. Therefore, it is no longer possible to use a unit area of one square kilometre having an average travel distance $L_c$ such as in Section 6.3.1, but it is necessary to analyse the full corridor.

For each unit area at a distance $L_x$ from the city centre, travellers are assumed to choose between either using the lower-level network, i.e. the 'slow' service, or the higher-level network: the express service. At any distance $L_x$ the travel time using network $j$ is (see equation (6-5)):

$$T_{c,j}(L_x) = w_a \cdot \frac{f_a \cdot f_d \cdot (S_{s,j} + S_l)}{V_a} + w_w \cdot \frac{f_w}{F_j} +$$

$$\frac{L_x}{S_{s,j}} \cdot \left( \frac{S_{s,j}}{V_j} + T_s \right) + w_e \cdot T_e$$

(6-17)

For the route choice between the two service types a logit model is assumed, similar to the model used for mode choice given in equation (6-8). Following Ben Akiva & Lerman

(1985) the aggregate travel time used to determine the level of demand, can be derived using the logsum over the utilities of the route choice model. This logsum is given by:

$$\widetilde{U}_c(L_x) = \overline{U}_c(L_x) + \frac{1}{\mu} \cdot \ln\left(\frac{\sum_{j=1}^{n_a} \exp\left(\mu \cdot \left(U_{c,j}(L_x) - \overline{U}_c(L_x)\right)\right)}{n_a}\right) + $$

$$\frac{1}{\mu} \cdot \ln(n_a) \qquad\qquad (6\text{-}18)$$

with:

$U_{c,j}(L_x)$ = utility in route choice model for alternative $j$: $U_{c,j}(L_x) = \alpha_r \cdot T_{c,j}(L_x)$

$\overline{U}_c(L_x)$ = average utitlity in route choice model for services $1$ to $n_a$

$\widetilde{U}_c(L_x)$ = resulting utility using the logsum

$\mu$ = scale parameter

$n_a$ = number of alternatives

$\alpha_r$ = sensitivity of route choice for travel time

The logsum automatically accounts for the case where an alternative service has a positive benefit. If there is no positive benefit, for instance in the case of short distances or long distances, the logsum equals the maximum utility of both services, i.e. the lower-level network and the higher-level network respectively. In the case that both service types are realistic options, the utility determined using the logsum is higher than the maximum utility of both service types. It should be noted that in the case of route choice the utility is usually negative, that is a maximum utility is equivalent to a minimum travel time.

Since the travel demand model is based on travel times, the utility based on the logsum is transformed in time again by dividing it by $\alpha_r$, the sensitivity to travel time in the route choice model. Theoretically all parameters used should be estimated simultaneously. In this case, however, no suitable data was available. Therefore, commonly used parameters for the individual models are used, thus requiring this pragmatic transformation. The average travel time in the corridor becomes:

$$T_c = \frac{\displaystyle\int_{L_x} \frac{\widetilde{U}_c(L_x)}{\alpha_r} \cdot P(L_x)}{\displaystyle\int_{L_x} P(L_x)} \qquad\qquad (6\text{-}19)$$

Using equation (6-19) for $T_c$ the social welfare can be calculated using the formulas presented in Section 6.3. The only exception is the formula for the operational costs (6-13) which is replaced by the summation of the operational costs for each network. These costs are written as:

$$C_{o,j} = \left( c_{o,j} + c_{im,j} \right) \cdot F_j \cdot \frac{1000}{S_l} \cdot \frac{L_l}{S_{s,j}} \cdot \left( \frac{S_{s,j}}{V_j} + T_s \right) \cdot 2 \tag{6-20}$$

where:

$L_l$  = line length

Given this description of the express system it is possible to determine the optimal network variables for the objective of maximising social welfare. For the numerical analysis the integral in equation (6-19) is replaced by a summation over a set of corridor segments. Values assumed for $\alpha_r$ and $\mu$ are 0.3 and 1.0 respectively. The results of this analysis are presented in Table 6-10. Please note that it is assumed that the line spacing and frequency for both networks are equal and that $S_{s,2}$ is a multiple of the stop spacing of the lower-level network using a scale-factor $sf_s$:

$$S_{s,2} = sf_s \cdot S_{s,1} \tag{6-21}$$

*Table 6-10: Optimal network characteristics for an express system*

|  | 10 km corridor | | 17 km corridor | |
|---|---|---|---|---|
|  | Reference (single-level) | Express system | Reference (single-level) | Express system |
| **Stop spacing level 1** [m] | 738 | 579 | 905 | 626 |
| **Line spacing** [m] | 717 | 929 | 770 | 965 |
| **Frequency** [veh/h] | 7 | 6 | 7 | 5 |
| **Scale-factor for stop spacing** | - | 2 | - | 2 |
| **Stop spacing level 2** [m] | - | 1159 | - | 1252 |
| **Weighted travel time** [min] | 31.6 | 31.8 | 39.1 | 39.3 |
| **Operational costs** [€/km²] | 51.2 | 68.0 | 44.3 | 52.9 |
| **Profit** [€/km²] | 5.5 | -11.4 | 52.2 | 43.5 |
| **Social welfare** [€/km²] | 523 | 505 | 691 | 681 |

The optimal stop spacing for level 1 is clearly shorter than the reference situation, while line spacing is larger and the frequency is lower. The stop spacing for level 2 is twice as large as that for level 1. The net result of an express system is that the weighted travel time is higher, as are the operational costs. The profits and thus the values for social welfare are therefore lower: -5.6% for the 10-kilometre corridor and -2.3% for the 17-kilometre corridor.

It will be obvious that because of the higher investment costs the assumption for a light rail system will lead to even larger reductions of social welfare. Dropping the assumption that both frequencies are equal, indeed results in a frequency of zero as suggested earlier, which results in a single-level network having the highest value for social welfare.

Similarly, dropping the constraint of equal line spacing results in a very large line spacing for one the networks, which is equivalent with a single-level network. In this specific case, assuming a uniform demand pattern and considering only trips to the city centre, there is no net benefit from introducing an express system compared to the single-level network. The possible mechanisms leading to this finding will be discussed in Section 6.4.3.4.

### 6.4.3.2 *Trunk and feeder services*

A trunk-feeder system is a truly multilevel network as discussed in Section 3.5. The trunk line is the higher-level network having dedicated bus lanes, and few stops. The lower-level network is the bus feeder network offering access to the stops of the trunk network. Trips to the city centre thus always require a transfer from a feeder service to the trunk service. As discussed in the beginning of Section 6.4.3 the frequencies of both network levels are equal. For trips to the city centre a higher frequency on the trunk line might seem more logical, however, that would ignore the fact that for a return trip a higher frequency of the feeder network would be desirable. Furthermore, it is assumed that the frequencies are high enough to eliminate the need to synchronise both service types. In this analysis only trips to the city centre are analysed. There are thus no travellers using only the feeder network, nor are there travellers using only the trunk network. The analysis focuses purely on travellers using both networks.

Feeder systems have been studied in the literature, however, these studies focused mostly on the feeder part only (Kuah & Perl (1988), Chang & Schonfeld (1991)). Wirasinghe (1980) also included the stop spacing for the trunk network as a decision variable, but he assumed that the trunk system itself is given including the operational costs involved. Furthermore, none of these studies make a comparison with a single-level system.

For trunk-feeder services two typical lay-outs of the feeder network are distinguished (Figure 6-9):

- Perpendicular feeders where the stop spacing of the trunk line is related to the line spacing of the feeder lines;

- Radial feeder networks whereas the stop spacing of the trunk line is twice the radius of the feeder network.

For the perpendicular feeder network an analytical model is formulated to assess the potential of a trunk-feeder system in an urban corridor having a uniformly distributed demand pattern. The results of this analysis are used to discuss the case of radial feeder networks. Finally, the assumption that all travellers use the feeder network to access the trunk line is reconsidered.
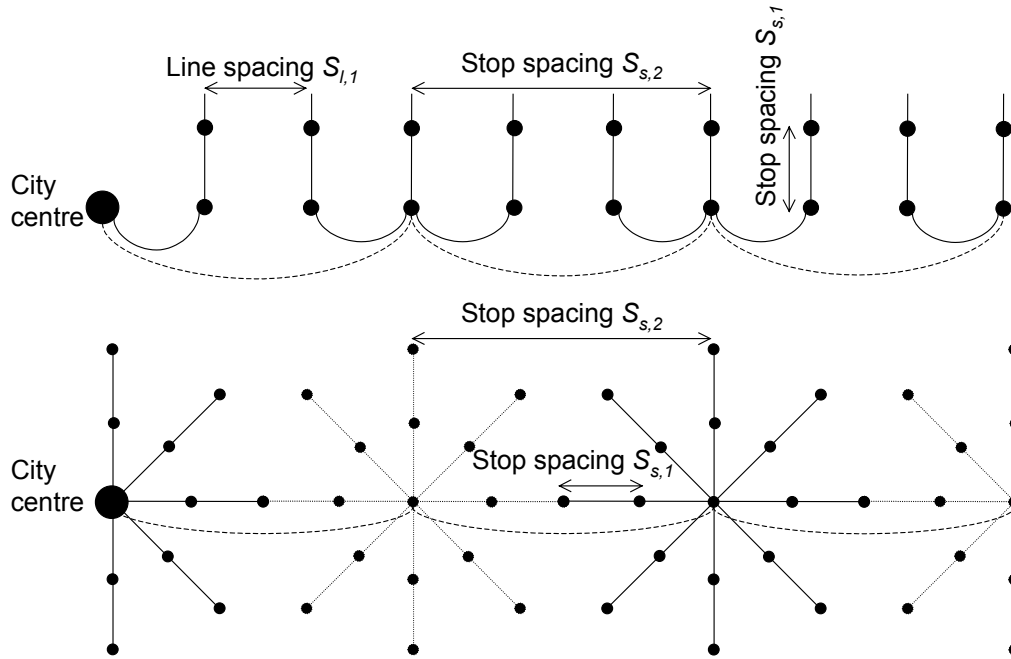
*Figure 6-9: Typical configurations of a trunk-feeder network:*
*feeder lines perpendicular to the trunk line and radial feeder networks*

*Perpendicular feeder lines*

In the case of feeder lines perpendicular to the trunk line, an analytical model is formulated for maximising social welfare. It is assumed that the stop spacing of the trunk line is related to the line spacing of the feeder network using a scale-factor $sf_{sl}$ (see Figure 6-9):

$$S_{s,2} = sf_{sl} \cdot S_{l,1} \tag{6-22}$$

The decision variables are stop and line spacing for the feeder network ($S_{s,1}$ and $S_{l,1}$), the frequency $F_1$ (which is equal to $F_2$), the scale-factor for the stop spacing $sf_{sl}$, and the line spacing for the trunk network $S_{l,2}$.

The access time to the trunk network is determined by the characteristics of the feeder network, a transfer penalty, and the space accessibility of the trunk line service, that is, its stop and line spacing. The access time to the trunk network is assumed to be independent of the distance to the city centre, and is formulated as:

$$
T_{a,2} = w_a \cdot \frac{f_a \cdot f_d \cdot (S_{s,1} + S_{l,1})}{V_a} + w_w \cdot \frac{f_w}{F_1} +
$$
$$
\frac{f_a \cdot f_d \cdot (S_{s,2} + S_{l,2})}{S_{s,1}} \cdot \left( \frac{S_{s,1}}{V_1} + T_s \right) + w_t \tag{6-23}
$$

where:

$\quad w_t \quad$ = transfer penalty

The average total travel time to the city centre then becomes:

$$T_c = T_{a,2} + w_w \cdot \frac{f_w}{F_2} + \frac{L_c}{S_{s,2}} \cdot \left( \frac{S_{s,2}}{V_2} + T_s \right) + w_e \cdot T_e \qquad (6\text{-}24)$$

The operational costs are, just as in the case of the express system, the summation of those for the feeder and the trunk network. The operational costs of the feeder network consist of a component perpendicular and a component parallel to the trunk line. The component parallel to the trunk line is determined by the stop spacing of the trunk line, the line spacing of the trunk line and the number of feeder lines per stop of the trunk line. All costs are determined per square kilometre. The equation for the operational costs thus becomes (see equation (6-13)):

$$C_{o,1} = \left( c_{o,1} + c_{im,1} \right) \cdot F_1 \cdot \frac{1000}{S_{l,1}} \cdot \frac{1000 + \dfrac{1}{sf_{sl}} \cdot \dfrac{f_a \cdot f_d \cdot S_{s,2}}{0.5 \cdot S_{l,2}}}{S_{s,1}} \cdot \left( \frac{S_{s,1}}{V_1} + T_s \right) \qquad (6\text{-}25)$$

with:

  $sf_{sl}$  = number of feeder lines per stop on the trunk line

     (scale factor for the stop spacing)

Using the formulas described in Section 6.3 the objective function for maximising social welfare is formulated, and then solved numerically for the two corridor lengths. Furthermore, two alternatives are considered with respect to the transfer to the trunk line. The first alternative considers only waiting times while ignoring any transfer penalties, and is thus an optimistic scenario. The second alternative explicitly accounts for the fact that the waiting time at a transfer has a higher weight and includes a transfer penalty of 5.7 minutes (see e.g. Van der Waard (1988)). The results of this analysis are shown in Table 6-11 in which the four trunk-feeder designs are compared to the corresponding single-level system.

The optimal stop spacing of the feeder network varies around 600 metres. Its line spacing is relatively large, 900 to 1100 metres, enabling higher frequencies. This is due to the fact that waiting time occurs twice in the total travel time. The stop spacing of the trunk network is three quarters of the line spacing and varies between 2,700 and 3,100 metres. The line spacing is about 4 kilometres. The fact that the line spacing is longer than the stop spacing reduces the impact of the component of the operational costs for the part of the feeder network parallel to the trunk line (see equation (6-25)).

If transfer penalties are ignored, a trunk-feeder system leads to lower travel times in the case of a 17-kilometre corridor. In all other cases the weighted travel times are higher. Furthermore, the operational costs of a trunk feeder system are always higher than those of a single-level system. As a result the values for social welfare are also lower: -8.3% to -2.7% if transfer penalties are ignored, and -22.6% to -11.7% if transfer penalties are included in the analysis. In this case, having assumed a uniform demand pattern and

considering trips to the city centre only, there is, again, no positive benefit from introducing a higher-level network.

*Table 6-11: Optimal network characteristics for a trunk-feeder system compared to the single-level reference*

| | 10 km corridor | | | 17 km corridor | | |
|---|---|---|---|---|---|---|
| | Reference (single-level) | Trunk feeder system | | Reference (single-level) | Trunk feeder system | |
| | | No transfer penalties | Transfer penalties | | No transfer penalties | Transfer penalties |
| *Feeder network* | | | | | | |
| **Stop spacing** [m] | - | 616 | 655 | - | 630 | 664 |
| **Line spacing** [m] | - | 889 | 955 | - | 969 | 1030 |
| **Frequency** [veh/h] | - | 10 | 12 | - | 9 | 11 |
| **Scale-factor** $sf_{sl}$ | - | 3 | 3 | - | 3 | 3 |
| *Trunk network* | | | | | | |
| **Stop spacing** [m] | 738 | 2,667 | 2,865 | 905 | 2,908 | 3,091 |
| **Line spacing** [m] | 717 | 3,748 | 4,212 | 770 | 4,032 | 4,457 |
| **Frequency** [veh/h] | 7 | 10 | 12 | 7 | 9 | 11 |
| **Weighted travel time** [min] | 31.6 | 33.2 | 41.2 | 39.1 | 38.6 | 46.4 |
| **Operational costs** [€/km$^2$] | 51.2 | 71.1 | 87.5 | 44.3 | 66.4 | 74.1 |
| **Profit** [€/km$^2$] | 5.5 | -25.0 | -35.2 | 52.2 | 30.3 | 19.4 |
| **Social welfare** [€/km$^2$] | 523 | 479 | 405 | 691 | 673 | 611 |

*Radial feeder lines*

In the case of a radial feeder network, the stop spacing and line spacing of the trunk line network should be equal in order to have a systematic coverage of the corridor. Furthermore, the access time to the trunk network will be shorter due to the direct routing to the trunk line stops. It might therefore be expected that the optimal values for stop spacing and line spacing for the trunk line network will be larger than those found in the case of perpendicular feeder lines. An analysis for the perpendicular network using the additional constraint that the stop spacing should equal the line spacing, shows a scale-factor of 4 and optimal values for the stop and line spacing ranging between 3,800 and 4,600 metres. The operational costs of the radial network will be higher due to the high line density near the trunk line stop (see Van Nes (2000)), while those for the trunk line

will be lower. It is expected that the net results will be slightly better than those for a perpendicular feeder network, however, the net effect compared to the reference situation will still be negative.

*Usage of the feeder network*

Given the access distances to the trunk network an analysis can be made of the impact of the assumption that all travellers use the feeder network to access the trunk network. After all, travellers may choose to walk to the trunk line stop instead of using the feeder network. Basically, a feeder network is only interesting for travellers if the total access time for the higher-level network using the feeder network is lower than the weighted access time for the access mode walking.

This line of reasoning can best be illustrated by focussing on trips parallel to the trunk line. In this way there is no need to assume an arbitrary value for the line spacing: the access distance perpendicular to the line is equal for both network types. If the origin of the trip is located between two stops of the trunk line, travellers can either walk to the stop of the trunk line or use the feeder line. The average access distance to the trunk line stop can be written as:

$$L_a = f_a \cdot f_d \cdot S_{s,2} \tag{6-26}$$

In the case of walking, the weighted access time becomes:

$$T_{aw,2} = w_a \cdot \frac{L_a}{v_{a,2}} = \frac{w_a \cdot f_a \cdot f_d \cdot S_{s,2}}{v_{a,2}} \tag{6-27}$$

Using the feeder network, the total weighted travel time to the trunk line stop consists of access time, waiting time, in-vehicle time for the feeder network, and a transfer penalty (compare Equation (6-7)):

$$T_{af,2} = w_a \cdot \frac{f_a \cdot f_d \cdot S_{s,1}}{v_{a,1}} + w_w \cdot \frac{f_w}{F_1} + \frac{L_a - f_a \cdot f_d \cdot S_{s,1}}{S_{s,1}} \cdot \left( \frac{S_{s,1}}{V_1} + T_s \right) + w_t \tag{6-28}$$

Please note that the in-vehicle distance is reduced by the access distance for the feeder network. The time components for the trunk service, i.e. waiting time, in-vehicle time, and egress time, are assumed to be equal for both cases. Furthermore, it is assumed that walking is the only access mode and that the factors for access are equal for both service types. Finally, it is assumed that the stop spacing of the trunk line is a multiple of the stop spacing of the feeder network. If access using the feeder network is more attractive than walking to the trunk line stop, the scale-factor for the stop spacing $sf_s$ should fulfil the following inequality:

$$sf_s \geq \dfrac{\kappa \cdot S_{s,1} + \dfrac{\tau}{F_1} + w_t - f_a \cdot f_d \cdot \left(\dfrac{S_{s,1}}{V_1} + \lambda\right)}{\kappa \cdot S_{s,1} - f_a \cdot f_d \cdot \left(\dfrac{S_{s,1}}{V_1} + \lambda\right)}$$                          (6-29)

where:

$$\kappa = \dfrac{w_a \cdot f_a \cdot f_d}{V_{a,1}}$$

$$\lambda = T_s$$

$$\tau = w_w \cdot f_w$$

Substitution of commonly used parameter values (valid for the Netherlands) results in minimum values for the scale-factor for stop spacing, which are shown in Table 6-12. An analysis, where, instead of the average access distance, the maximum access distance is used as the main criterion ($L_a=0.5 \cdot S_{s,2}$), is also included. The results show that the optimal stop spacing of the trunk network should be larger than 1 kilometres at least and preferably larger than 2 kilometres. For realistic values of stop spacing for the feeder network, 600 to 800 metres, the scale-factor for stop spacing is about 3.

*Table 6-12: Minimum values of the scale-factor for stop spacing for a trunk-feeder network*

| Stop spacing feeder network | $L_a$ is average access distance | | $L_a$ is maximum access distance | |
| :---: | :---: | :---: | :---: | :---: |
| | Minimum value scale-factor stop spacing | Minimum stop spacing trunk network | Minimum value scale-factor stop spacing | Minimum stop spacing trunk network |
| 300 | 6.0 | 1,800 | 3.0 | 900 |
| 400 | 4.7 | 1,900 | 2.4 | 950 |
| 500 | 4.0 | 2,000 | 2.0 | 1,000 |
| 600 | 3.4 | 2,050 | 1.7 | 1,050 |
| 700 | 3.1 | 2,150 | 1.5 | 1,100 |
| 800 | 2.8 | 2,250 | 1.4 | 1,150 |
| 900 | 2.6 | 2,350 | 1.3 | 1,200 |
| 1,000 | 2.5 | 2,450 | 1.2 | 1,250 |

All values are lower than those found for the optimal stop spacing of the trunk line in Table 6-11. This shows that the assumption that all travellers use the feeder network is reasonable, although, of course, travellers near the trunk line stop will usually walk to the trunk line stop. Finally, it is interesting to note that the minimal stop spacing for the trunk network according to this line of reasoning is larger than the stop spacing found for the express network (Section 6.4.3.1). The assumption that the lower-level network is not used to access the express network is thus also justified.

### 6.4.3.3 Zone system

In a zone system a line combines two functions: collecting (or distributing) passengers within a zone and offering an express service from that zone to the city centre (and vice versa) (Figure 6-10). For the express part dedicated bus lanes are available, thus allowing higher speeds. The essential difference with the trunk-feeder concept is that in this case it is not necessary to make a transfer between both network levels. Since each zone is served by its own line, the number of lines is equal to the number of zones. Travellers walk to the nearest stop and travel to the city centre. The possibility to walk in the opposite direction of the line in order to board at the last stop of the next zone, which is served by the express part, is not considered in the analysis.
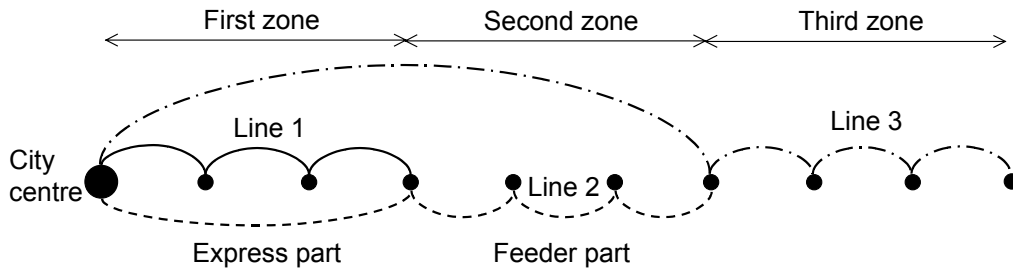


*Figure 6-10: Zone-system consisting of three zones*

An analytical model is developed for maximising social welfare. The decision variables are the number of zones, the stop spacing of the feeder part, the line spacing and the frequency.

From the travellers point of view the average travel time to the city centre can be written as in equation (6-7) except for the in-vehicle time which now consists of two parts. The in-vehicle time for the feeder part ($T_{i,1}$) is determined by the average trip length within the zone and can be described by:

$$T_{i,1} = \frac{0.5 \cdot \dfrac{L_r}{Z}}{S_{s,1}} \cdot \left( \frac{S_{s,1}}{V_1} + T_s \right) \tag{6-30}$$

where:

$L_r$ = corridor length or radius of the city

$Z$ = number of zones

The average in-vehicle time on the express part ($T_{i,2}$) is described by:

$$T_{i,2} = \frac{\dfrac{1}{Z} \cdot \sum_{j=1}^{Z-1} \left( \dfrac{j}{Z} \cdot L_r \right)}{V_2} = \frac{L_r \cdot (Z-1)}{2 \cdot Z \cdot V_2} \tag{6-31}$$

Please note that in the case of a single zone, the in-vehicle time on the express part equals zero.

The operational costs per square kilometre ($C_o$) are determined by the frequency, the number of lines, the total travel time of the vehicles within the corridor for both the feeder part and the express part of the line (in two directions):

$$C_o = F \cdot \frac{1000}{S_{l,1}} \cdot 2 \cdot \left( \begin{array}{l} \left(c_{o,1} + c_{im,1}\right) \cdot \frac{1000}{S_{s,1}} \cdot \left( \frac{S_{s,1}}{V_1} + T_s \right) + \\[2mm] \left(c_{o,2} + c_{im,2}\right) \cdot 1000 \cdot \frac{(Z-1)}{2 \cdot V_2} \end{array} \right) \tag{6-32}$$

Combining these formulations for in-vehicle time and operational costs and those of Section 6.3.1 leads to the objective for maximising social welfare, which then is optimised for the design variables stop spacing, line spacing, frequency, and number of zones. The results of this analysis are shown in Table 6-13 and are compared to the corresponding single-level references. A zone system proves to be an interesting alternative. For a 10-kilometre corridor a 2-zone system leads to an increase of the social welfare of 2.5%, and for a 17-kilometre corridor a 3-zone system gives an improvement of 3.3% compared to the reference situation. Even though the operational costs for a zone system are higher due to the larger number of lines, the reduction in travel time is large enough to compensate for it. Interestingly, the stop spacing for the feeder part is nearly constant: ± 600 metres.

*Table 6-13: Optimal network characteristics for a zone system*

|  | 10 km corridor | | 17 km corridor | |
|---|---|---|---|---|
|  | Reference (single-level) | Zone system | Reference (single-level) | Zone system |
| **Number of zones** | 1 | 2 | 1 | 3 |
| **Stop spacing** [m] | 738 | 584 | 905 | 603 |
| **Line spacing** [m] | 717 | 800 | 770 | 969 |
| **Frequency** [veh/h] | 7 | 6 | 7 | 6 |
| **Weighted travel time** [min] | 31.6 | 29.6 | 39.1 | 34.4 |
| **Operational costs** [€/km$^2$] | 51.2 | 54.4 | 44.3 | 53.5 |
| **Profit** [€/km$^2$] | 5.5 | 3.2 | 52.2 | 44.6 |
| **Social welfare** [€/km$^2$] | 523 | 536 | 691 | 714 |

The fact that a zone system has a clear positive benefit on social welfare leads to two questions:

- Is a zone system also interesting for corridors shorter than 10 kilometre?

- What will happen if no bus lanes are assumed, i.e. the only benefit from a zone system is that stops are skipped on the express part?

An analysis for a 6-kilometre corridor still shows a positive benefit of 1.5% if a 2-zone system is introduced. Without bus lanes the maximum speed is not increased and there is no increase in operational costs. The results for three corridor lengths are shown in Table 6-14. It clearly shows that in this case a zone system is interesting for longer corridors only, and that the benefit is less: 0.8% for the 10-kilometre corridor and 1,1% for the 17-kilometre corridor.

*Table 6-14: Optimal network characteristics for a zone system in which the express part and the feeder part have identical characteristics*

|  | 6 km corridor | | 10 km corridor | | 17 km corridor | |
|---|---|---|---|---|---|---|
|  | Reference (single-level) | Zone system | Reference (single-level) | Zone system | Reference (single-level) | Zone system |
| **Number of zones** | 1 | 1 | 1 | 2 | 1 | 2 |
| **Stop spacing** [m] | 631 | 631 | 738 | 589 | 905 | 693 |
| **Line spacing** [m] | 731 | 731 | 717 | 856 | 770 | 868 |
| **Frequency** [veh/h] | 7 | 7 | 7 | 7 | 7 | 6 |
| **Weighted travel time** [min] | 27.6 | 27.6 | 31.6 | 30.1 | 39.1 | 37.4 |
| **Operational costs** [€/km$^2$] | 55.0 | 55.0 | 51.2 | 60.3 | 44.3 | 48.3 |
| **Profit** [€/km$^2$] | -21.1 | -21.1 | 5.5 | -2.8 | 52.2 | 48.8 |
| **Social welfare** [€/km$^2$] | 425 | 425 | 523 | 527 | 691 | 699 |

In order to gain some insight into the relationships for the optimal number of zones it is interesting to analyse the objective function of minimising total costs (Equation (6-15)), again using equations (6-30) to (6-32) for in-vehicle time and operational costs. At the optimum the derivative with respect to the number of zones should equal zero. The optimal number of zones is then given by (see also Equation (6-16)):

$$Z^* = \sqrt{\left( \frac{L_r}{\rho_2} \cdot \left( \frac{V_2 \cdot \lambda}{S_{s,1}^*} + \frac{V_2}{V_1} - 1 \right) \cdot \frac{S_{l,1}^*}{F_1^*} \right)} \qquad (6\text{-}33)$$

with:

$$\rho_2 = \frac{1000^2 \cdot (c_{o,2} + c_{im,2}) \cdot 2}{P \cdot c_t}$$

This equation shows that the optimal number of zones increases with corridor length $L_r$, speed on the express-part $V_2$, the line spacing $S_{l,1}$ and the number of travellers $P$ (the

denominator of $\rho_2$), and decreases with the operational costs per vehicle per hour $c_{o2}+c_{im,2}$ and the frequency $F_1$.

These results are in line with those found in literature (Tsao & Schonfeld (1983), Furth (1986), Chang & Schonfeld (1993a)). In their analyses the vehicle capacity is introduced as an additional constraint, which also makes a zone system more interesting from an operator's point of view. For longer corridors the combination of optimal frequency and vehicle capacity might not be sufficient to accommodate the total demand, leading to higher frequencies, and thus to higher operational costs than assumed for the reference situation. A zone system allows more efficient operation of the vehicles that are available.

### 6.4.3.4  Discussion of findings

The analysis in the previous three sections focussed on hierarchical systems in long urban corridors having a uniform demand pattern of trips oriented to the city centre. In general the introduction of a higher-level network such as an express network or a trunk-feeder network leads to a lower value for the design objective of maximising social welfare. In order to be attractive, especially in terms of speed, the higher-level network will be more expensive, thus increasing the operational costs. Although the higher-level network is interesting for trips where origins are located near stops of the higher-level network, a trunk-feeder network introduces a transfer for origins located farther away from those stops, while in the case of an express network the choice between two service types does not always lead to shorter travel times. Given the assumption of a uniformly distributed demand pattern, these effects lead to a net increase of travel time. A single network is therefore a better alternative. This conclusion implies that hierarchical structures such as those found in practice, can not adequately be explained using the assumptions used in this analysis. The following sections therefore will discuss alternative assumptions with respect to the orientation of the demand pattern (Section 6.4.4) and the distribution of the demand (Section 6.4.5).

A zone system, however, proves to be an interesting option for longer corridors. It leads to shorter travel times, while operational costs may also decrease if vehicle capacity is included in the analysis. The main benefit compared to the trunk-feeder network is that there is no transfer between the two components of the system, thus eliminating the transfer penalty and the additional waiting time. Compared to the express system, the zone system integrates the characteristics of both service types instead of forcing a choice between both service types, while the costs required for providing a higher quality are lower. It is intriguing that design guidelines do not pay attention to zone systems and the practical applications appear to be limited, for example in Paris or Chicago. A possible reason for this might be that zone systems focus explicitly on trips to the city centre. For more diffuse demand patterns for instance, a zone system might be less beneficial than an express system.

Finally, some remarks can be made regarding the assumption of dedicated bus lanes for the higher-level network and the equality of the frequencies of both systems. Given the finding of a negative net effect of introducing a higher-level network under the

assumption of bus lanes having low additional costs while offering a substantial increase in speed, it is clear that assuming other systems for the higher-level network will always lead to larger reductions in social welfare. The operational costs increase while the quality with respect to travel time remains the same. The finding that the net benefit of a higher-level network is negative implies that an optimisation in which both frequencies are independent design variables will lead to a solution in which one of the frequencies will be zero, thus resulting in a single-level network. Such a solution might lead to some computational difficulties, but more importantly, it does not provide insight into the characteristics of a hierarchical network structure. Both assumptions might seem arbitrary at first sight, but they allow a more complete view of hierarchical network structures.

### 6.4.4   Non-centre oriented demand patterns

The previous section dealt with demand patterns focussed only on the city centre. For larger cities this assumption might no longer hold. This section therefore deals with large urban areas having a varied pattern of origins and destinations. Two cases are considered:

- A linear corridor in which two service networks can be distinguished: one for short distance trips and the other for long distance trips. An analytical model will be developed to determine optimal network characteristics for these networks;

- A higher-level grid network that is superimposed on a lower-level grid network, a situation comparable to the analysis for road networks in Section 5.4.

The key question in this section is whether under the assumption of a uniform demand pattern that is no longer oriented to the city centre, mechanisms can be found that determine the emergence of hierarchical network structures.

#### *6.4.4.1   Linear network (corridor)*

This analysis still deals with an urban corridor, however, the trips are no longer oriented to the city centre. Within the corridor two parallel public transport services are offered. The first service is a bus service that focuses on short trips, while the second service has dedicated bus lanes and is designed to accommodate long distance trips within the corridor. For the same reasons as in the analysis of the express system in Section 6.4.3.1, it is assumed that the line spacing and frequencies of both services are equal. Each service, however, has its own stop spacing. It is expected that the differences between trip lengths will automatically lead to different values. No transfers between both service types are assumed.

Travellers choose a service type depending on their trip length. In this case it is assumed that all kinds of trips are made within the corridor and that there are no clear concentrations of origins or destinations. Given a trip length distribution a distinction can be made between short-distance trips and long-distance trips, each of which are served by their own network. These networks are optimised simultaneously as single-level

networks for their own specific trip types using the objective of maximising social welfare. Thus, for each network level the objective function is comparable to equation (6-14). The design variables are then the stop spacing for each network ($S_{s,1}$, $S_{s,2}$), line spacing ($S_{l,1}=S_{l,2}$), frequency ($F_1=F_2$), and the distance that distinguishes between short-distance and long-distance trips, that is the distance criterion $L_d$.

A key point in this analysis is the way the distinction between trip types is incorporated. Given a trip length distribution for a specific corridor length it is possible to calculate for every possible value of the distance criterion the resulting average distance for short-distance trips and long-distance trips, as well as the share of short distance trips. The total demand per square kilometre, which is based on the reference situation, remains equal in all analyses. The demand is split into short-distance and long-distance trips according to the share of short distance trips, which depends on the value of the distance criterion used.

Figure 6-11 shows the three trip length characteristics for trips larger than 2 kilometres in the 17-kilometre corridor based on the trip length distribution from the Dutch National Travel Survey. For each characteristic a function can be estimated that describes the relationship with the distance criterion $L_d$. Since trip length distributions can be approximated by exponential distributions, a logarithmic function is estimated:

$$f_l(x) = \gamma_1 \cdot \ln(L_d) + \gamma_2 \qquad\qquad (6\text{-}34)$$
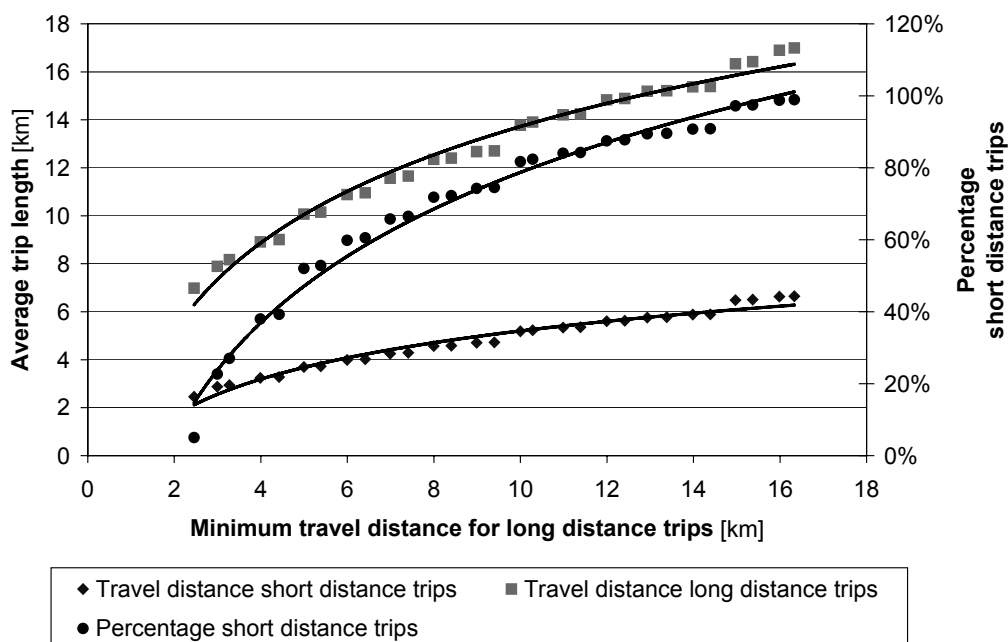


*Figure 6-11: Average trip distances for short-distance and long-distance trips and the share of short distance trips as a function of distance criterion $L_d$ for the 17-kilometre corridor excluding short trips (NTS 95)*

These functions are used to define the relationship between the distance criterion and the average trip length for each network, and to determine the share of each trip type class.

Since short distance trips are dominant in the trip length distribution a distinction is made between trip length distributions including and excluding short trips. The total demand per square kilometre, however, is identical in all analyses. Excluding short trips thus leads to longer trips and thus higher revenues. All functions used in the analyses for the 10 kilometre and the 17-kilometre corridor that are based on equation (6-34) have an R-square larger than 0.9. It should be noted that in the case where short trips are included, a linear function might be more appropriate. Such a function, however, does not lead to a feasible optimum.

The resulting optimal network characteristics for the two corridor lengths, with and without short trips, are shown in Table 6-15. The results are compared to the corresponding single-level references. Please note that the values for the reference situation differ from those presented in Table 6-9 and used in earlier analyses. This is due to the different trip length distribution that is assumed here.

*Table 6-15: Optimal network characteristics for a two-level network in an urban corridor*

|  | 0 - 10 km corridor | | 1 - 10 km corridor | | 0 - 17 km corridor | | 2 - 17 km corridor | |
|---|---|---|---|---|---|---|---|---|
|  | Ref. single level | Two levels | Ref. single level | Two levels | Ref. single level | Two levels | Ref. single level | Two levels |
| **Distance criterion** [km] | - | 2.76 | - | 4.03 | - | 3.80 | - | 6.52 |
| **Stop spacing level 1** [m] | 738 | 557 | 738 | 636 | 905 | 576 | 905 | 733 |
| **Line spacing** [m] | 717 | 963 | 717 | 964 | 770 | 979 | 770 | 1,001 |
| **Frequency** [veh/h] | 7 | 6 | 7 | 6 | 7 | 6 | 7 | 6 |
| **Stop spacing level 2** [m] |  | 1,043 |  | 1,063 |  | 1,211 |  | 1,259 |
| **Weighted travel time** [min] | 29.0 | 31.1 | 31.0 | 33.4 | 33.4 | 34.2 | 37.6 | 39.0 |
| **Operational costs** [€/km$^2$] | 51.2 | 69.9 | 51.2 | 67.5 | 44.2 | 66.2 | 44.2 | 61.0 |
| **Profit** [€/km$^2$] | -9.4 | -28.9 | 2.5 | -14.8 | 17.0 | -5.2 | 43.4 | 26.1 |
| **Social welfare** [€/km$^2$] | 461 | 423 | 510 | 473 | 544 | 516 | 654 | 626 |

The stop spacing for the lower-level network in both cases including very short trips is similar: 557 and 576 metres. If very short trips are excluded, however, the difference between the two corridors is clearly larger, nearly 100 metres. The stop spacing of the higher-level network is generally a factor of 2 larger than the stop spacing of the lower-level network. As expected the alternatives excluding short trips have a larger distance criterion, higher profits and higher social welfare. The distance criterion is 1 kilometre larger for the 10-kilometre corridor, and 2.5 kilometres larger for the 17-kilometre corridor. A sensitivity analysis for the assumptions with respect to line spacing and

frequency showed only minor deviations. The maximum difference for social welfare, for instance, was less than 0.5 %.

Just as was found in the analysis of the express system (Section 6.4.3.1), the introduction of a higher-level network leads to an increase of operational costs, which is limited by an increase in line spacing and a reduction in frequency. Furthermore, due to larger access distances and lower frequencies the average travel times are also higher, yielding lower values for the objective of maximising social welfare. The reduction in welfare varies between 7% and 8% for the 10-kilometre corridor and 4% to 5% for the 17-kilometre corridor. Adding an additional network level, again, proves to have a negative benefit.

### 6.4.4.2   Grid network

In this section the restriction of a linear corridor is relaxed, and a grid network is assumed that serves a many-to-many travel pattern in a square city. This situation is highly similar to that analysed for private transport in Chapter 5 in which a higher-level network is super-positioned on a lower-level network. There are however two important differences between private transport and line-bound public transport that should be taken into account in this analysis: time accessibility and transfers. Since the analysis focuses on higher-level networks, the third difference between private transport networks and public transport service networks, that is the limited space accessibility of the lower public transport network, is not relevant here.

Private transport modes can easily cross between network levels, while in public transport travellers have to change vehicles at transfer nodes and have to wait till the next service arrives. As a result the difference between using the lower-level network or using the higher-level network as well, is no longer a function of distance and speed only, but also includes transfer penalties and additional waiting times. The higher-level should not only compensate for the detour but also for the additional transfers and waiting times, which makes the concept of elimination of shortcuts dependent on the trip lengths considered.

The economic line of reasoning (Section 5.4), however, might be applied to line public transport services as well. The area studied is a large city, for instance 20 by 20 kilometres wide, having a diffuse demand pattern. In the analysis the trip length distribution for trips up to 30 kilometres will be used. The city is served by a lower-level bus network having a line spacing of 500 metres, a frequency of 6 vehicles per hour, and a speed of 20 kilometres per hour. Compared to the analyses for the national road network, the demand densities in urban areas are clearly higher. A density of 4,300 inhabitants per square kilometre is realistic for West European urban areas. Since public transport might not be an option for all inhabitants a reduction of 33% is applied.

For the higher-level network several options are considered, ranging between low cost options such as a network of bus lanes and high cost options such as an underground system. The main characteristics of these systems are presented in Table 6-16.

*Table 6-16: Characteristics of higher-level urban networks*

| | Bus lane | | Light Rail | Metro | |
|---|---|---|---|---|---|
| | Simple | Exclusive | | Elevated | Tunnel |
| **Average speed** [km/h] | 30 | 40 | 40 | 40 | 40 |
| **Frequency** [veh/h] | 6 | 6 | 4 | 9 | 9 |
| **Operating cost** [€/veh/km] | 3.48 | 3.48 | 7.68 | 12.74 | 12.74 |
| **Investment costs** [€/m] | 1,250 | 2,500 | 5,500 | 15,000 | 40,000 |
| **Stations** [% of metro stations] | - | - | - | 50% | 100% |

Just as in the case of private networks the benefits of introducing a higher-level network in terms of travel time savings are calculated and are reduced by a payment for the investment costs and the operational costs. For the calculation of travel time it is assumed that nearly all travellers using the lower network transfer only once. Given the diffuse demand pattern and the size of the city, only 5% of the trips can be made without a transfer. This percentage is based on the percentage of the city surface that is covered by two perpendicular lines. Similarly, most travellers using the higher-level network will use the lower-level network for access or egress. Only 10% of the travellers access the higher-level network on foot. The higher-level network has a low line density, implying that it is likely that more travellers can reach their destinations without transferring to another higher-level network line. This percentage is assumed to be 35%, which is based on a line spacing of the higher-level network of 4 kilometres.
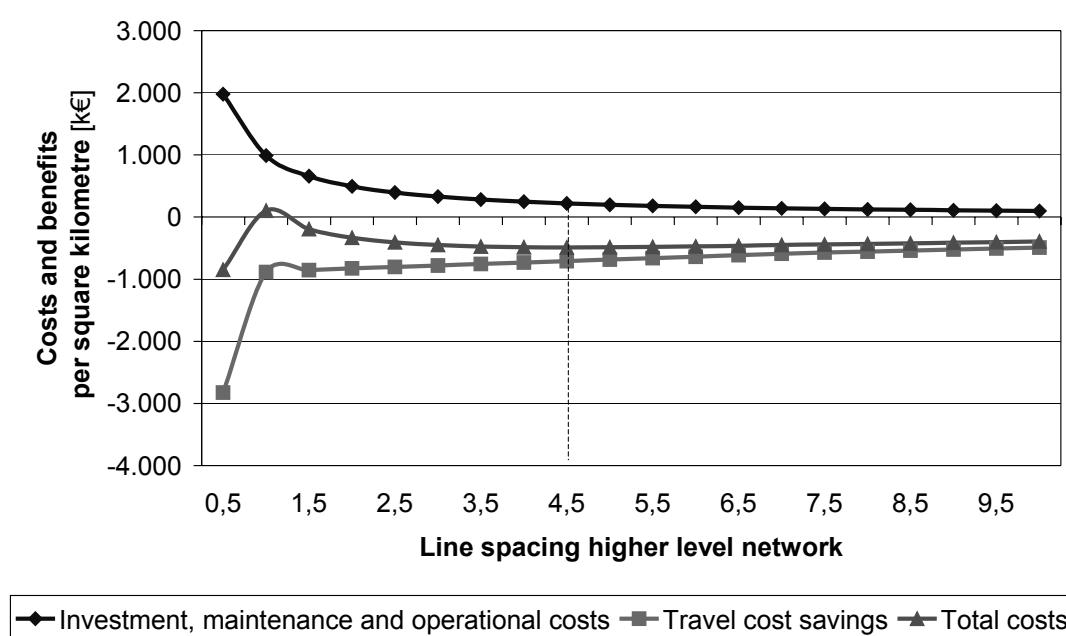


*Figure 6-12: Relationship between various costs and line spacing for the higher-level public transport network (exclusive bus lane)*

An interesting option that appeared in the analysis is that of replacing the lower-level network completely. In that case the additional costs of transferring to the higher-level network are eliminated completely while all travellers benefit from the higher average speed. For the bus lane systems this proved to be a realistic solution as shown in Figure 6-12. However, if the possibility of replacement is excluded, another optimum can be found. Table 6-17 shows the optimal line spacing for all options for the higher-level network, and also indicates whether replacement is an option. The scale-factor for the line spacing is very large, the minimum value being 9. The metro system is apparently not suitable under the assumptions used, since it leads to an increase of total costs.

*Table 6-17: Optimal line spacing for an urban higher-level grid network*

| | Bus lane | | Light Rail | Metro | |
| --- | --- | --- | --- | --- | --- |
| | **Simple** | **Exclusive** | | **Elevated** | **Tunnel** |
| **Replacement** | Yes | Yes | No | No | No |
| **Line spacing level 2** [km] | 6 | 4,5 | 7,5 | 10[*] | 10[*] |
| **Scale-factor for line spacing** | 12 | 9 | 15 | 20 | 20 |
| **Change in total costs** [k€/km²] | -171 | -489 | -229 | 186 | 814 |

[*] *Maximum value considered for a city of 20 by 20 kilometres*

The scale-factors for line spacing are substantially higher than those found in practice (see for instance Table 6-2), even though a relatively high level of demand is assumed and that in the case of bus lanes only a small increase in costs is considered. The results for the metro option even lead to the conclusion that metro systems are not economically justifiable. Just as in the analysis of a linear corridor the additional costs involved in providing a higher-level network are too high to be compensated by the benefits of higher speeds. There is apparently, under the assumption of a uniformly distributed demand pattern, no strong rationale to develop higher-level public transport networks.

### 6.4.4.3   Conclusions

The analysis of the corridor leads to the conclusion that having a multilevel network does not lead to an improvement of social welfare compared to single-level systems. For a grid network a multilevel network might be an option, but the resulting scale-factors for the line spacing do not resemble scale-factors found in practice. It can therefore be concluded that relaxing the constraint that public transport is only focussed on trips to the city centre, does not lead to the emergence of a hierarchical network structure. The next section focuses on the third assumption in these analyses that might explain the emergence of hierarchical public transport service networks: the concentration of the demand pattern.

### 6.4.5   Concentration in demand patterns

Sections 6.4.3 and 6.4.4 analysed hierarchical network structures under the assumption of a uniformly distributed travel demand pattern, which is oriented to the city centre or has a diffuse demand pattern. In both cases it was found that the introduction of a higher-level network might reduce travel times for travellers whose origins are located near stops of the higher-level network, but that for all other travellers the travel time increases due to the longer access distances or the transfer to the higher-level network. Combined with the increase in operational costs due to introducing an additional service network this leads to lower values for social welfare compared to the single-level network.

If instead of a uniformly distributed travel demand pattern concentrations in the travel demand pattern are assumed, these conclusions might change. In the case of higher densities around the stops of the higher-level network, in terms of inhabitants, workplaces or facilities, the number of travellers that benefit from the service quality of the higher-level network increases substantially. This might lead to a net positive benefit from introducing a higher-level network.

Considering such concentration of demand near stops of the higher-level network, however, introduces a special kind of dependency in the transport network design problem. If the demand densities are given beforehand, the pattern of the stops for the higher-level network will be determined by the pattern of locations having higher densities. In that case, hierarchy in urban public transport networks is not a characteristic of public transport networks themselves but represents a characteristic of the hierarchy of the urban system. Alternatively, when the density of the demand is considered a result of the public transport network design process, that is of the resulting access density of the higher-level network, transport network design becomes similar to urban design. It is even possible that a system results in which an increase in access density leads to an increase in demand, which allows a further increase in density, and so on. Another possibility might be that demand will be located only around the higher-level stops, thus allowing elimination of the lower-level network. These are typical examples of a fixed point problem as discussed in Section 3.3.3. In this case the description of travel behaviour not only relates to route choice, mode choice, and destination choice, but also incorporates location choice.

In the context of the topic of this chapter, determining mechanisms that lead to emergence of hierarchical network structures, such a fixed point approach is clearly too detailed. Therefore, the impact of higher demand densities around stops of the higher-level network is illustrated using a descriptive model for an express system in an urban corridor. The analyses only deals with trips oriented to the city centre.

The layout of the lower-level network is assumed to be fixed having a stop spacing of 400 or 600 metres and a line spacing of 1,000 metres (Figure 6-13). The stop spacing of the higher-level network is a multiple of that of the lower-level network. The higher-level has dedicated bus lanes, allowing higher speeds. The line spacing for both networks is equal. The difference with the previous approaches presented in this chapter is that, instead of using a unit area of a square kilometre, the unit area is defined as the service

area of a stop of the lower-level network. For service areas where the express line stops, different demand levels will be assumed using a scale-factor $sf_d$.
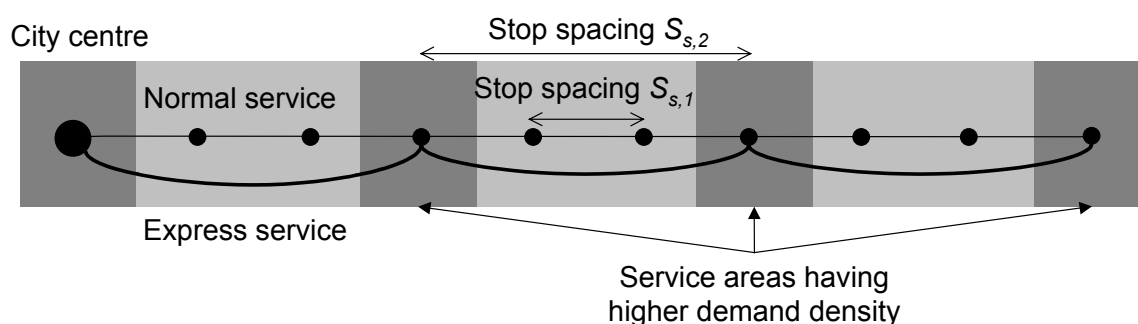


*Figure 6-13: Layout of an express system in an urban corridor in which the express line stops have higher demand levels*

Travellers in a unit area may choose to use the lower-level network or to use the higher-level network to travel to the city centre. The average travel time for each service area follows from the logsum of the travel times for each network level (see Section 6.4.3.1). If both systems stop in the unit area the choice is based on differences in frequencies and in-vehicle times only. For unit areas in between the stops of the higher-level network, travellers may choose to walk along the line, either in the forward or backward direction, to one of the stops of the higher-level network. The stop spacing of the express system is assumed to be small enough to ignore the possibility of using the lower-level network to access the higher-level network (see also Table 6-12).

In order to avoid the problems related to the fixed-point character of the design problem including variable demand densities, the analysis follows a two-step procedure. First the value of the scale-factor for the stop spacing ($sf_s$) is determined simultaneously with the frequencies for both network levels ($F_1, F_2$), while assuming an equal demand level at each stop. For simplicity sake, the objective of minimising total costs is used. Second, given the optimal stop spacing of the higher-level network determined in the first step, the level of demand for the service areas of the express line stops is then doubled and tripled. For both scenarios the optimal values for the frequencies are determined using the objective of minimising total costs. Instead of assuming equal frequencies for both services as in earlier approaches, a minimum frequency of 4 vehicles per hour is used in this analysis.

The results are compared to those for the single-level references as shown in Table 6-9. Since the different situations have different numbers of travellers, the total costs are scaled to the number of travellers in the reference situation. It should be noted that due to the fact that the descriptive model is based on service areas per stop instead of a fixed line length, the results may differ slightly from previous analyses.

The results for the 10-kilometre corridor are shown in Table 6-18. The optimal scale-factor for the stop spacing in the case of a uniformly distributed demand pattern is 3, which is higher than that found in Section 6.4.3.1. For the cases having a lower-level stop spacing of 400 metres, this seems realistic since the stop spacing of the higher-level

network is in both cases about 1,100 - 1,200 metres. Furthermore, the frequency on the lower network equals the minimum frequency used as an additional constraint. As expected the total costs of the alternatives having equal demand is higher than for the single-level reference (7%). If the demand for the higher-level network is doubled, however, an express system leads to a reduction of total costs (2%), even though the frequency of the higher-level network is increased. In the case of tripling the demand the reduction in total costs is 7%.

*Table 6-18: Network characteristics of an express network having different demand levels at the stops of the higher-level network for the 10-kilometre corridor*

| | Ref. (single level) | Stop spacing lower-level network is 400 m | | | Stop spacing lower-level network is 600 m | | |
|---|---|---|---|---|---|---|---|
| **Scale-factor for demand level at express line stops** | | 1 | 2 | 3 | 1 | 2 | 3 |
| **Stop spacing level 1** [m] | 738 | 400 | 400 | 400 | 600 | 600 | 600 |
| **Line spacing** [m] | 717 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| **Frequency level 1** [veh/h] | 7 | **4** | **4** | **4** | **7** | **7** | **7** |
| **Scale-factor for stop spacing** | - | **3** | 3 | 3 | **3** | 3 | 3 |
| **Stop spacing level 2** [m] | - | 1,200 | 1,200 | 1,200 | 1,800 | 1,800 | 1,800 |
| **Frequency level 2** [veh/h] | - | **8** | **9** | **11** | **5** | **8** | **9** |
| **Weighted travel time** [min] | 31.6 | 32.7 | 30.9 | 29.3 | 33.1 | 31.2 | 30.2 |
| **Operational costs** [€/km$^2$] | 51.2 | 64.3 | 68.8 | 77.7 | 60.4 | 71.5 | 75.1 |
| **Total costs** [€/km$^2$] | 309 | 331 | 304 | 286 | 336 | 315 | 299 |

*The decision variables used to minimise total costs are shown in bold print*

In the case of a lower-level stop spacing of 600 metres, the scale-factor 3 yielded slightly lower total costs than the expected value of 2. This is due to the fact that the frequencies between both networks are allowed to differ. The result is a delicate balance between lower-level and higher-level networks, while in most cases the lower-level network has the minimum frequency implying that the higher-level network is favoured over the lower-level network. Increasing the demand level does not directly lead to lower total costs. Only in the case of tripling the demand at the higher-level stops, a reduction of 3% is found.

The results for the 17-kilometre corridor appear to be similar (Table 6-19). In the case of the 400-metre stop spacing, the optimal network configurations are identical to those for the 10-kilometre corridor. The only difference is that the reduction of the total costs increases to 9%. In the 600-metre case the scale-factor for the stop spacing is 2, leading to the stop spacing for the higher-level network of 1,200 metres found in previous analyses. Furthermore, there is a tendency to favour the higher-level system, especially if

the demand for the higher-level network stops is increased. In this case an express system is already beneficial if the demand is doubled (3%).

*Table 6-19: Network characteristics of an express network having different demand levels at the stops of the higher-level network for the 17-kilometre corridor*

| | Ref. (single level) | Stop spacing lower-level network is 400 m | | | Stop spacing lower-level network is 600 m | | |
|---|---|---|---|---|---|---|---|
| **Scale-factor for demand at higher-level network stops** | | 1 | 2 | 3 | 1 | 2 | 3 |
| **Stop spacing level 1** [m] | 905 | 400 | 400 | 400 | 600 | 600 | 600 |
| **Line spacing** [m] | 770 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| **Frequency level 1** [veh/h] | 7 | **4** | **4** | **4** | **5** | **4** | **4** |
| **Scale-factor for stop spacing** | - | **3** | 3 | 3 | **2** | 2 | 2 |
| **Stop spacing level 2** [m] | - | 1,200 | 1,200 | 1,200 | 1,200 | 1,200 | 1,200 |
| **Frequency level 2** [veh/h] | - | **8** | **10** | **11** | **7** | **10** | **12** |
| **Weighted travel time** [min] | 39.1 | 37.7 | 35.2 | 34.0 | 38.9 | 36.4 | 34.7 |
| **Operational costs** [€/km$^2$] | 44.3 | 65.4 | 74.7 | 79.3 | 62.0 | 69.9 | 79.1 |
| **Total costs** [€/km$^2$] | 363 | 377 | 348 | 329 | 387 | 351 | 331 |

*The decision variables used to minimise total costs are shown in bold print*

This analysis illustrates clearly that higher demand levels around stops of the higher-level network make a hierarchical network system advantageous, while a uniform demand pattern does not. It shows furthermore, that there is a general tendency to reduce the frequency of the lower-level network to the minimum frequency, thus enabling higher frequencies for the higher-level network, especially if the stop spacing of the higher-level network is about 1,200 metres. If the stop spacing would be larger, for instance 1,800 metres, there are more benefits from using the lower-level network, leading to a delicate balance between both systems. In all cases, however, it is obvious that increasing demand levels around stops of the higher-level network reduces the importance of the lower-level network, and leads to higher frequencies for the higher-level network. This mechanism might also explain the higher frequencies for the second level, which was noted in Table 6-8.

### 6.4.6   Conclusions

This part of Chapter 6, Section 6.4, focused on the emergence of hierarchical urban public transport networks. Several types of hierarchical network structures were studied under different assumptions to find out whether hierarchical network structures have self-organising properties, just as private transport networks.

The analysis of hierarchical networks in urban areas, however, shows that public transport networks by themselves do not naturally lead to a hierarchy such as an express system or a trunk-feeder system. This conclusion holds for a uniformly distributed travel demand pattern and for the corridor lengths analysed. The differences between the results for the 10-kilometre corridor and the 17-kilometre corridor suggest that a break-even point might be found for still longer corridor lengths. In fact, an exploratory analysis shows that for a 30-kilometre corridor a hierarchical network leads to a minor positive effect for the objective of social welfare maximisation (+0.1% for an express system and +0.6% for a trunk feeder system). However, the realism of the assumption of a uniformly distributed demand pattern becomes even more questionable as the area size increases. Corridors longer than 17 kilometre can no longer be looked at as urban corridors, but should be regarded as interurban public transport, which is discussed in the following section (Section 6.5).

The finding that scenarios considering concentrations of demand around the stops of the higher-level network are the only situations in which hierarchical network structures appear advantageous, leads to the conclusion that spatial patterns determine the hierarchy in an urban public transport system. This finding clearly distinguishes public transport networks from private transport networks, which have their own mechanisms leading to a hierarchy, which happen to be consistent with the hierarchy for spatial structures (see Chapter 5).

Zone systems prove to be an exception. A zone system is an interesting alternative for corridors longer than 6 kilometres if dedicated bus lanes are available for the express part, and for corridors longer than 10 kilometres if the characteristics for the express part are similar to those of the feeder part. Surprisingly, zone systems are seldom found in practice. An important limitation of zone systems might be that they focus too much on trips to the city centre only, thus being less suitable for other trip types. Furthermore, they require identical transport technologies for both parts of the line. This requirement can be met by using buses that use bus lanes on the express part, but this is not consistent with the general perception that rail services have a higher quality than road-bound public transport. It should be noted that this positive bias towards rail services has not been confirmed in practice (e.g. Axhausen et al. (2001)). Finally, the urban structure itself might make zone systems less attractive. The location of areas having higher densities might not match the preferred allocation of the zones, or the demand level related to these higher densities might require a higher capacity for the express part.

## 6.5   Multilevel regional and national public transport networks

The third part of analysis in this chapter deals with hierarchy in interurban public transport networks. The first two parts dealt with single-level urban public transport networks and multilevel urban public transport networks respectively. The main difference between urban public transport networks on the one hand and regional and national public transport networks in the other is that the demand is clearly no longer continuously distributed but is concentrated in settlements of different sizes. Furthermore, regional and national networks tend to have area covering network structures instead of

networks that are primarily focussed on a single destination. Existing regional and national public transport networks have more network levels than urban public transport networks. It is shown in Section 6.2.2.1 that at least four different network levels may be distinguished for the Netherlands: namely regional, interregional, national, and international. Finally, having multiple levels in a public transport service system also has organisational consequences. For urban public transport networks it is realistic to assume a single viewpoint only for designing a network, while for interurban networks different operators and different authorities need to be considered in analysing the designs.

The main topic in this section is, again, the emergence of hierarchical structures. Given the findings on urban multilevel public transport networks it seems plausible that spatial hierarchy is decisive for the hierarchy in public transport networks as well. This assumption is supported by the modal split of access and egress modes for long distance public transport services such as train services. The modal split for the access and egress modes for trips made by train in the Netherlands show that lower-level public transport systems such as bus, tram, and metro only account for 20% to 25% of all trips. Other modes such as walking (nearly 50%) and cycling are dominant. In the case of the main cities in the Netherlands these percentages become 40% walking, 40% bus-, tram-, and metro-services, and 20% private modes. The demand for higher-level networks thus depends strongly on the densities around the stops within walking or cycling distance. Spatial structures, again, seem to determine the hierarchy for regional and national public transport networks.

Ending the analysis of interurban public transport networks by accepting this conclusion, however, might be premature since this conclusion does not state that interurban public transport network do not have mechanisms that might lead to hierarchy. It only states that it is plausible that spatial structures determine interurban public transport structures. In order to find out whether interurban networks do have their own of mechanism that might lead to hierarchy two different types of analysis are pursued.

First, the fact that regional and national networks have area covering structures suggests an analysis comparable to the analysis of private networks in Chapter 5, in which a higher-level network is superimposed on a lower-level network (Section 6.5.1). The question in this analysis is whether a higher-level network is worthwhile, and if so, what the scale-factor for the line spacing should be. The second type of analysis focuses on the dependency between network levels, that is the relationship between the network characteristics of subsequent network levels (Section 6.5.2). This analysis uses the concept of a multilevel network as discussed in Section 3.5, and is comparable to the analysis of the trunk-feeder network in Section 6.4.3.2.

Finally, the fact that interurban public transport networks interact with urban public transport networks, while for both networks different authorities and operators might be responsible, requires an additional extension of the multilevel network design problem. The model presented in Section 6.5.3 uses game theoretical concepts to account for these organisational aspects.

### 6.5.1   Super-positioning a higher-level network

In the analysis of hierarchy in private transport networks two approaches were used for the analysis of super-positioning a higher-level network on an existing lower-level network. The first focused on elimination of shortcuts, while the second used an economic approach. This section analyses the possibilities to use these two approaches for interurban public transport networks. There are, however, two important differences compared to the analysis for private networks that should be considered for this analysis. The first is the fact that in public transport services changing between network levels requires a transfer. The second difference is that regional and national public transport services tend to have a triangular network structure instead of a grid network structure (see Table 6-2). This is especially relevant for the economic approach.

#### *6.5.1.1   Elimination of shortcuts*

For private transport networks a simple relationship between network levels could be derived, without specific assumptions on the demand pattern (Section 5.3.2). The need to transfer, however, changes the analysis, since the higher-level network should compensate not only for the detour but also the transfer. The necessary differences in speeds will therefore be higher for the lower-level networks, that is networks suited for short travel distances, than for higher-level networks which deal with longer travel distances.

For a grid network, the impact of transfers is even larger, because it is not only necessary to transfer to the higher-level network but also to transfer while using the higher-level network only (Figure 6-14). In that case the barriers for using the higher-level network become very large indeed, unless of course the destinations are such that no additional transfers are necessary. This option, however, requires additional assumptions with respect to the demand pattern, and thus emphasises the importance of spatial structures.
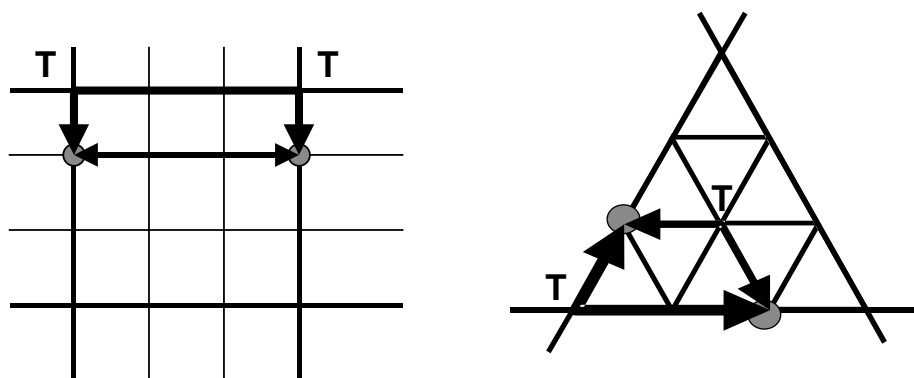


*Figure 6-14: Transfers for trips using the lower-level or the higher-level public transport network for a grid network and a triangular network*

The argument for triangular network structures is slightly different. If it is assumed that travellers have origins and destinations located close to the stops of the higher-level network, again an assumption on spatial structures, it is still possible to derive a generic

relationship for the differences in speeds. In that case there is no transfer from the lower-level network to the higher-level network, and if the detour shown in the right-hand side of Figure 6-14 is considered, travellers have to transfer using the lower-level network as well as using the higher-level network. The criterion to eliminate detours may then be simplified to differences in in-vehicle time only, implying a minimum scale-factor for speed of 1.5 and a scale-factor for line spacing of 3, just as for triangular private transport networks.

### 6.5.1.2  Economic approach

The main question in the economic approach is: given a lower-level public transport network, for instance regional buses, what are then the benefits and costs of introducing a higher-level network such as regional train or express train services? In this analysis the benefits consist of the saving in travel time, while the costs are determined by the necessary infrastructure and the operational costs of the transport services.

Details of this approach can be found in Section 5.4, while some of the consequences of applying this approach to public transport are discussed in Section 6.4.4.2. In the case of interurban public transport the following differences should be accounted for:

- Triangular network structures instead of grid networks;

- The level of demand;

- The characteristics of the transport services offered at each network level, especially with respect to frequencies;

- Transfers between network levels and within network levels;

The triangular network structure changes the formulas for the costs for private networks presented in Section 5.4, as can be seen in Appendix C. The difference in network structure also reduces the necessary number of transfers. From each transfer node it is possible to travel in six directions instead of four.

The demand densities in urban areas are assumed to be equal to those in the analysis for the private networks. Since public transport might not be an option for all travellers, however, two scenarios are analysed having reductions of 33%, and 60%. In the following analysis the trip length distribution for trips longer than 10 kilometres will be used. The value of time is based on the value for train users: € 5.50.

The lower-level network is a regional bus network having a line spacing of 10 kilometres, a frequency of 2 vehicles per hour, and a speed of 35 kilometres per hour. For the higher-level public transport network two options are considered, that is a low cost option as a network of regional train services and a high cost option such as an express train service network. The main characteristics of these systems are presented in Table 6-20 (see also Appendix F). It should be noted that the differences in speed that are assumed are quite large. Furthermore, the frequencies of the transport services are low compared to those assumed in the analysis of urban public transport networks. These low frequencies might require some kind of synchronisation between services. Therefore, instead of accounting

for both transfer waiting time and transfer penalty, the description of the transfer is limited to a fixed transfer penalty of 13 minutes.

*Table 6-20: Characteristics of higher-level regional networks*

|  | Regional train services | Express train services |
|---|---|---|
| **Average speed** [km/h] | 70 | 105 |
| **Frequency** [veh/h] | 2 | 2 |
| **Operating cost** [€/veh/km] | 5.47 | 7.69 |
| **Investment costs** [€/m] | 5,510 | 8,220 |

For the calculation of travel time it is assumed that nearly all travellers using the lower network transfer only once: only 10% of the trips can be made without transfers. Similarly, most travellers using the higher-level network will use the lower-level network for access or egress. Only 25% of the travellers access the higher-level network using other modes. The triangular structure of the higher-level network also influences the percentage of trips that have to transfer within the higher-level network. For a triangular network it is likely that more travellers can reach their destination without transferring to another higher-level network line. Furthermore, most travellers will try to minimise the number of transfers and thus choose destinations that require fewer transfers. Therefore this percentage is assumed to be 67%.

*Table 6-21: Optimal line spacing for a regional higher-level network*

|  | Regional train | | Express train | |
|---|---|---|---|---|
| **Demand for public transport** | 40% | 67% | 40% | 67% |
| **Line spacing level 2** [km] | 100 | 70 | 100 | 70 |
| **Scale-factor for line spacing** | 10 | 7 | 10 | 7 |
| **Change in total costs** [k€/km²] | -8.3 | -40.0 | -12.4 | -56.1 |

The calculation of the costs for each possible value of higher-level network line spacing is identical to that for private transport networks. Table 6-21 shows the optimal line spacing for all options for the higher-level network, and the net effect on the total costs. The scale-factor for the line spacing is quite large, the minimum value being 7. The express train system appears to be preferable, although the differences are small. The higher costs, investment and operational, are more than compensated for by the higher speed. The optimal values for line spacing are in the same range as those found in practice (see Table 6-2), however, those for the regional train services are equal to those for the express train services. Interestingly, the option of replacing the lower-level network completely is not realistic in this case. Apparently, the costs of building a higher-level network are too high. Figure 6-15 shows the cost functions for the alternative of an express train service network and a demand level of 67%. It is clear that the objective function is very shallow, allowing for many near optimal solutions.
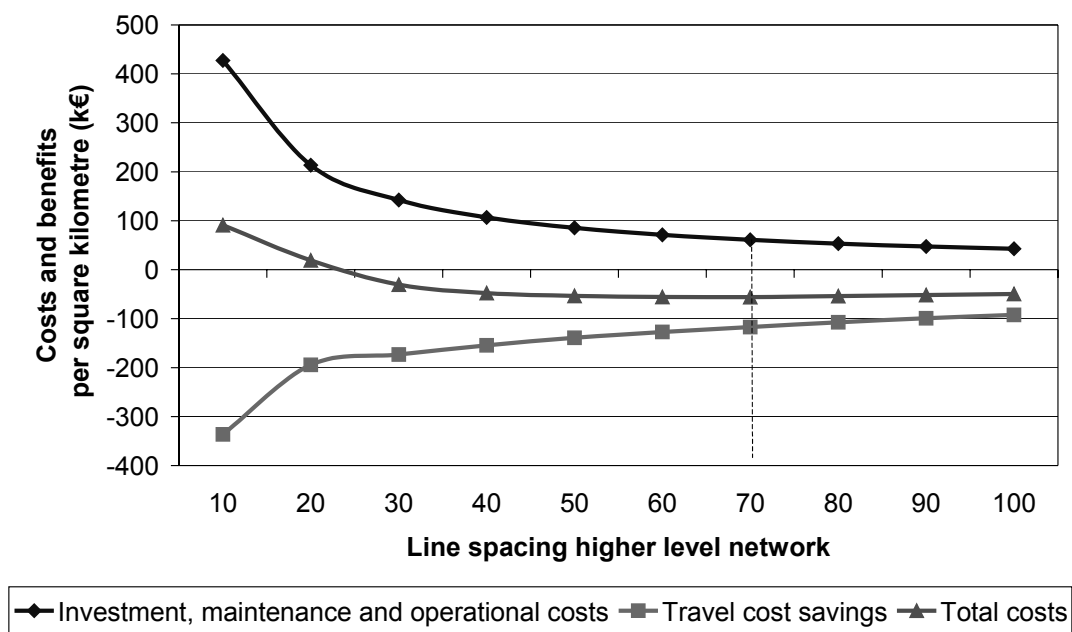
*Figure 6-15: Relationship between various costs and line spacing for the higher-level public transport network (express train, 67% of the demand)*

It should be noted that the assumptions with respect to the number of transfers strongly influence these results. If, for instance, it is assumed that most origins and destinations are located close to stops of the higher-level network, the percentage of travellers that have to transfer from the regional bus system to the higher-level network will be lower. Table 6-22 shows the results if only 25% of the travellers have to transfer between network levels. The optimal line spacing is clearly lower, especially for the express train service network. Furthermore, the reduction in total costs per square kilometre is larger.

*Table 6-22: Optimal line spacing for a regional higher-level network assuming that the demand is concentrated around the stops of the higher-level network*

|  | Regional train | | Express train | |
|---|---|---|---|---|
| **Demand for public transport** | **40%** | **67%** | **40%** | **67%** |
| **Line spacing level 2** [km] | 80 | 40 | 80 | 40 |
| **Scale-factor for line spacing** | 8 | 4 | 8 | 4 |
| **Change in total costs** [k€/km$^2$] | -16.7 | -66.1 | -20.8 | -81.6 |

It is interesting to note that the values for line spacing for regional train and express train services are always identical. This is due to the fact that benefits following from the higher speed are nearly proportional to the increase in investment and operational costs. The net differences are too small to be noted using integer decision variables.

The picture that emerges from this analysis is not very clear. The analysis based on eliminating detours suggests a scale-factor of 3 for line spacing and a scale-factor of 1.5 for speed, however, this is based on a very specific case. The economic approach

supports the line spacing found in practice, but these results are also sensitive to the assumptions used. The general tendency is that the high investment costs involved make higher-level networks very costly, leading to large-scale-factors for line spacing as well as for speed. In contrast to the analysis for private networks, this line of reasoning does not lead to generic rules for hierarchy in interurban public transport service networks.

### 6.5.2   Hierarchy as a network characteristic

The concept of hierarchy as a network characteristic is based on the fact that each network level has two functions: providing transport for it's own trips and offering access and egress to higher-level networks. It is typical for urban public transport networks that the access nodes for the higher-level networks are often located close to the city centre. The average access distance for the higher-level network is thus equivalent to the average trip length to the city centre.

Following the suggestion by Immers & Egeter (1996) (see Section 6.2.2.3), it can be assumed that for the higher-level network a similar line of reasoning can be used as for single-level urban public transport networks (see Section 6.3.1). In their analysis of the optimal access density, a simple relationship between both network levels is derived. The access distance to the higher-level network is a function of both its stop spacing and line spacing, which are a multiple of the stop and line spacing of the lower-level network:

$$L_{a,2} = f_{a,2} \cdot f_{d,2} \cdot \left(S_{s,2} + S_{l,2}\right) = f_{a,2} \cdot f_{d,2} \cdot \left(sf_s \cdot S_{s,1} + sf_l \cdot S_{l,1}\right) \tag{6-35}$$

If the access distance to the higher-level network equals the average trip length to the city centre, $L_{a,2}$ equals $L_c$. For the cases analysed in Section 6.3.3 the average trip length was 3 and 5 kilometres respectively. If for simplicity sake it is assumed that that the scale-factor for line spacing equals that for stop spacing, it is simple to formulate a relationship for the scale-factor:

$$sf_s = \frac{L_c}{f_{a,2} \cdot f_{d,2} \cdot \left(S_{s,1} + S_{l,1}\right)} \tag{6-36}$$

In the case of a homogeneous distribution of the demand the factor $f_{d,2}$ would equal 1, while for an access network having a grid structure $f_{a,2}$ would be 0.25. For a higher-level network, however, it is more likely that the demand is concentrated around the stops, implying that the demand factor $f_{d,2}$ should be smaller than 1. Following the scale-factor 3 for spatial structures as discussed in Section 4.3, a realistic assumption would be that the demand factor $f_{d,2}$ is reduced to one-third.

Substitution of these factors, the average trip lengths, and the corresponding values for the stop spacing and line spacing for urban public transport networks (Table 6-7) in equation (6-36) yields the scale-factors for stop spacing and line spacing shown in Table 6-23. It is clear that the scale-factors derived on the assumption of a strict hierarchy between networks are very large, and do not resemble any of the values found in practice. This finding leads to the conclusion that the assumption that the lower-level network

determines the network characteristics of the higher-level network is not realistic. Interurban public transport networks also do not appear to have a mechanism leading to the emergence of hierarchy.

*Table 6-23: Scale-factors for higher-level networks assuming a strict dependency between network levels*

| | Bus | | Tram | |
|---|---|---|---|---|
| | Reference | Max. Social welfare | Reference | Max. Social welfare |
| **Average trip length** [m] | 3,000 | 3,000 | 5,000 | 5,000 |
| **Stop spacing level 1** [m] | 350 | 640 | 400 | 757 |
| **Line spacing level 1** [m] | 500 | 752 | 1,000 | 797 |
| **Scale-factor** | 42.4 | 25.9 | 42.9 | 38.6 |
| **Stop spacing level 2** [m] | 14,824 | 16,552 | 17,143 | 29,228 |
| **Line spacing level 2** [m] | 21,176 | 19,448 | 42,857 | 30,772 |

### 6.5.3   Organisational aspects and hierarchy

#### 6.5.3.1   Introduction

In the analysis of hierarchy in public transport networks, there is one issue that has not yet been analysed: the impact of the organisation of public transport. For urban public transport networks there is usually a single organisation responsible for the network design, either the public transport company itself or the local authorities. The combination of urban and regional or national public transport networks, however, includes other parties as well. Each of these parties is likely to be focused on its own transport market and have its own objectives. There is thus no guarantee that all network levels are designed consistently, which is an implicit assumption in all previous analyses in this thesis. The question in this section is therefore to assess the impact of having different operators or authorities in a multilevel public transport system (see also Van Nes (2002)).

The situation of multiple parties who are responsible for public transport network design, that is operators or authorities, can be analysed using game theory (see e.g. Gibbons (1992)). If there are two operators, each of them optimising their own objective, two classic cases can be distinguished. The first case is called a Cournot-game in which both operators optimise their own network simultaneously, while assuming an optimised network for the other operator. In the second case a distinction can be made between a leader and a follower, who optimise their objectives in that order. Such a sequential game is called a Stackelberg-game. A typical characteristic of a Stackelberg-game is that the leader can take advantage of his knowledge of the expected behaviour of the other party

involved. In a Cournot-game, however, in which both operators are assumed to have full and identical information, and in which they optimise their networks simultaneously, it is not possible for one operator to take advantage of the other.

As described in Section 3.3 the transport network design problem is an example of a Stackelberg-game in which the network builder, or operator, is the leader while the traveller is the follower. In the case of multilevel multi-operator public transport network design a Stackelberg seems less realistic. It might be stated that in the past the location of higher-level nodes, especially railway stations, were determined by the higher-level network operators only, forcing the lower-level network operators to adapt their networks. However, since the city centres of today are mostly located close to those railway stations, there is no specific reason to assume that one operator or network level should be seen as a leader and the other as a follower. Therefore the analysis will be based on a Cournot-game.

Another game theoretical concept is also relevant for this analysis. A Cournot-game may be simulated using a stepwise iterative procedure in which each player, or actor, optimises its objective using the knowledge of the decision of the other actors made in the preceding steps. Such a procedure can be stopped if no player is able to improve its objective further. In that case equilibrium has been achieved which is defined as a Nash-equilibrium.

In the case of public transport network design the operators, or authorities, might pursue their own objectives, either profit maximisation or welfare maximisation. All parties may have identical objectives, or, which is more likely according to the analysis of public transport economies by Berechman (1993), have different objectives, namely welfare maximisation for urban networks and profit maximisation for interurban networks. In all cases a comparison is made with the situation of a single operator, or authority, maximising the multilevel public transport service network. The following section describes the characteristics of the analytical model used. Next the results of the analysis will be discussed.

### 6.5.3.2   *Formulation of the model*

The system used to analyse this situation consists of an urban corridor having a network of parallel lines offering transport services to the city centre and an interurban network connecting two city centres (Figure 6-16). The design variables for this network are stop spacing $S_{s,1}$, line spacing $S_{l,1}$, and frequency $F_1$. The inter-urban network is a single line offering transport from centre 1 to centre 2. Since in interurban transport the location of the stops is primarily determined by the spatial organisation of cities, the set of design variables is limited to the frequency $F_2$.
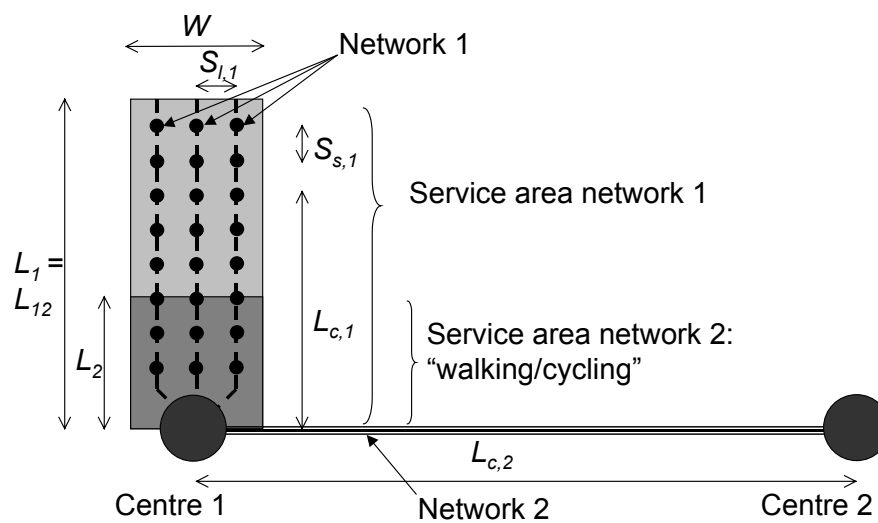
*Figure 6-16: Layout of the multilevel public transport network*

At first sight, this schematic lay out might seem unrealistic. However, since many cities have radial public transport networks, it is often possible to distinguish a set of corridors that make up the radial structure. Furthermore, there are usually several city centres that are an alternative for the local city centre. The realistic situation of several urban corridors and a number of alternative city centres is thus for simplicity sake reduced to a single corridor and a single alternative city centre. This simplification might affect the realism of a numerical analysis, but does not influence the main mechanisms resulting from the dependency between the networks.

The demand for the urban network consists of two populations: travellers to the city centre and travellers who use the urban network to access the interurban network. Similarly, the demand for network 2 consists of travellers who access the network using the urban public transport network, and travellers who use other modes, such as walking and cycling, to access the interurban network. Given the characteristics of these other modes, the service area of the latter population is smaller than that of the first. These travel choices are illustrated in Figure 6-17.

This nested choice model is simplified by making a distinction between three populations, each having its own travel quality (i.e. the grey boxes in Figure 6-17):

1.  Travellers $P_1$ using the urban public transport network to travel to the city centre 1;

2.  Travellers $P_2$ accessing the interurban public transport network on foot or by bicycle for travelling to city centre 2;

3.  Travellers $P_{12}$ using both networks to travel to city centre 2. This group has a transfer at the station in centre 1.
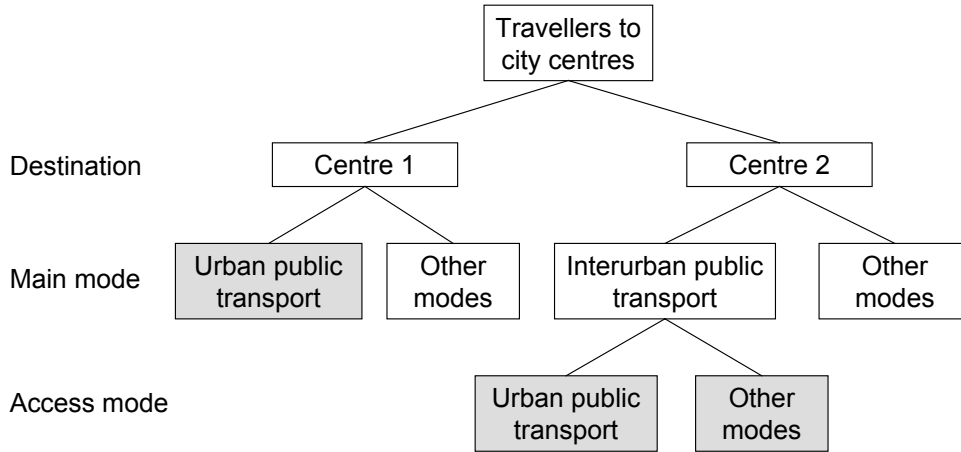
*Figure 6-17: Travel choices in a multilevel transport network*

The formulation of the analytical model is similar to those used in the previous sections (Section 6.3). The main differences are that indices for the network and the population are introduced, while all equations are related to the total system instead of a unit area of one square kilometre.

The formulation of the travel time for population $P_1$ is similar to Equation (6-7):

$$T_{c,1} = w_{a,1} \cdot \frac{f_{a,1} \cdot f_{d,1} \cdot (S_{s,1} + S_{l,1})}{V_{a,1}} + w_{w,1} \cdot \frac{f_w}{F_1} +$$
$$\frac{L_{c,1}}{S_{s,1}} \cdot \left( \frac{S_{s,1}}{V_1} + T_{s,1} \right) + w_{e,1} \cdot T_{e,1}$$

(6-37)

with:
$L_{c,1}$ = average travel distance to city centre 1

Travellers to city centre 2 who walk or cycle to the stop of the interurban network have their origin in a service area defined by width $W$ and length $L_2$. Theoretically, the value of $L_2$ depends on the characteristics of the urban public transport system and those of the alternative access modes. In order to analyse the impact of a population which uses the interurban network only, however, it suffices to assume that the value of $L_2$ is fixed. The travel time can be written as:

$$T_{c,2} = w_{a,2} \cdot \frac{f_{a,2} \cdot f_{d,2} \cdot \left( \frac{W}{2} + L_2 \right)}{V_{a,2}} + w_{w,2} \cdot \frac{f_w}{F_2} + \frac{L_{c,2}}{V_2} + w_{e,2} \cdot T_{e,2}$$

(6-38)

with:
$f_{a,2}$ = factor for the access distance as a function of $W$ and $L_2$

The population of travellers $P_{12}$ first use the urban network, yielding access time, waiting time and in-vehicle time, and then transfer to the interurban network, yielding a second waiting time and a second in-vehicle time:

$$T_{c,12} = w_{a,12} \cdot \frac{f_{a,1} \cdot f_{d,1} \cdot (S_{s,1} + S_{l,1})}{V_{a,1}} + w_{w,12} \cdot \frac{f_w}{F_1} + \frac{L_{c,1}}{S_{s,1}} \cdot \left( \frac{S_{s,1}}{V_1} + T_{s,1} \right) +$$

$$w_{w,12} \cdot \frac{f_w}{F_2} + \frac{L_{c,2}}{V_2} + w_{e,12} \cdot T_{e,2}$$

(6-39)

Since there is little knowledge about the behavioural differences between the three populations, it is assumed that all parameters are equal, that is $w_{a,1} = w_{a,2} = w_{a,12}$ et cetera, with one exception. Research on travel behaviour shows that the weight for the waiting time at transfer nodes is higher than the weight for the waiting time at the first stop (Van der Waard (1988)). Therefore, a higher value is used for the weight $w_{w,12}$ for waiting time for population $P_{12}$. All other assumptions are similar to those used in previous analyses (see Appendix E). The egress times in both city centres are assumed to be equal. Finally, it should be noted that synchronisation of timetables is not taken into account.

The demand is no longer related to a unit area of a square kilometre but is related to total service area of the urban network. Using Equation (6-8) a logit-model is used to determine the total demand for each population:

$$P_p(T_{c,p}) = \frac{\exp(-\alpha \cdot T_{c,p})}{\exp(-\alpha \cdot T_{c,p}) + \sum_m \exp(-\beta_m \cdot T_{m,p})} \cdot P_{0,p} \cdot W \cdot L_p$$

(6-40)

The revenues $R_i$ consist of the fares paid by the travellers and the related subsidies. Please note that the population using both networks is always included in the equation.

$$R_i = R_{t,i} + R_{s,i}$$
$$= r_{t,i} \cdot (P_i(T_{c,i}) + P_{12}(T_{c,12})) + s_{t,i} \cdot (P_i(T_{c,i}) + P_{12}(T_{c,12}))$$

(6-41)

with:

$r_{t,i}$ = fare paid by the traveller for network $i$

$s_{t,i}$ = subsidy per traveller for network $i$

The operational costs per system depend on the number of lines, the frequency and the total turnaround time of the lines. The operational costs for network 1 are specified as:

$$C_{o,1} = (c_{o,1} + c_{im,1}) \cdot F_1 \cdot \frac{W}{S_{l,1}} \cdot 2 \cdot \frac{L_1}{S_{s,1}} \cdot \left( \frac{S_{s,1}}{V_1} + T_{s,1} \right)$$

(6-42)

The operational costs for network 2 are defined by:

$$C_{o,2} = (c_{o,2} + c_{im,2}) \cdot F_2 \cdot 2 \cdot \left( \frac{L_{c2}}{V_2} + T_b \right)$$

(6-43)

where:

$T_b$ = buffer time

The objective of profit maximisation is simply the producer surplus, either per system or as the summation of both systems.

$$Max(PS_i) \tag{6-44a}$$

or

$$Max\left(\sum_i PS_i\right) \tag{6-44b}$$

The objective of welfare maximisation needs special attention. The producer surplus includes the subsidies, which should be regarded as additional costs for the objective of social welfare. Therefore the social welfare should be calculated as the summation of consumer surplus and producer surplus minus the subsidies, again per system or for both systems together. Furthermore, the total welfare is not equivalent to the sum of the welfare of both systems, because in that case the consumer surplus for population $P_{12}$ would be accounted for twice. The formulae thus become:

$$Max(SW_i) = Max(CS_i + CS_{12} + PS_i - R_{s,i}) \tag{6-45a}$$

or

$$Max\left(\sum_i SW_i - CS_{12}\right) = Max\left(\sum_p CS_p + \sum_i PS_i - \sum_i R_{s,i}\right) \tag{6-45b}$$

### 6.5.3.3  *Analysis of operating strategies*

The developed analytical model can be used to analyse the single operator case as well as the multi-operator case. In the latter case the analysis is based on a Cournot-game. This implies an iterative approach in which each network is optimised individually while assuming optimal characteristics for the other network determined in the previous step, until a Nash-equilibrium is achieved.  In all cases such an optimum has been found. Table 6-24 summarises the scenarios that are analysed.

*Table 6-24: Overview of scenarios with respect to number of operators and objectives*

| Scenario | Number of operators | Strategy or design objective | |
|---|---|---|---|
| | | Urban network | Interurban network |
| 1 | 1 | Profit | |
| 2 | 1 | Social welfare | |
| 3 | 2 | Profit | Profit |
| 4 | 2 | Social welfare | Social welfare |
| 5 | 2 | Social welfare | Profit |

The numerical example is based on an urban corridor 5 kilometres long and 2 kilometres wide. The length of the service area for the interurban network is assumed to be fixed at 2 kilometres. The distance between the cities is 20 kilometres. For the urban network the subsidy equals the fare paid by the traveller. The interurban system operates without

subsidy. The operational costs of the interurban network include a payment for the infrastructure. As a reference an urban network is assumed having a stop spacing of 400 metres, a line spacing of 500 metres, and a frequency of 4 vehicles per hour. The frequency of the interurban network is also 4 vehicles per hour.

Since the level of demand depends on the quality of the services offered, the total level is chosen such that in the reference situation $P_{12}$ is 40% of all travellers using the urban network, an estimate based on an analysis for the four main cities in the Netherlands using the Dutch National Travel Survey. $P_{12}$ is also 40% of all travellers using the interurban system, an assumption based on the share of public transport as an access/egress mode for train trips in urbanised areas. The parameters describing travel behaviour are assumed to be equal for all three populations, with two exceptions. The access speed for population $P_2$ is based on walking and cycling access modes, while the weight for the waiting time for population $P_{12}$ is higher than for the other populations, thus accounting for the higher weight related to transfers.

The results of the various scenarios are given in Table 6-25. Each column represents the network characteristics for a scenario, the values for the objective functions, the travel time per population and the operational costs for each network. The results for the single operator cases have similar characteristics as those for single-level networks. Compared to welfare maximisation, profit maximisation leads to lower network densities and longer travel times, even though in this analyses a level of subsidy for the urban network is assumed that is representative to the level proposed by the Dutch authorities. Welfare maximisation leads to a slightly coarser urban network than the reference situation and to higher frequency. A typical difference with the analysis for single-level networks is that the frequencies for the urban network are higher, which is due to the higher weight for the waiting time for population $P_{12}$, while in all cases the frequency for the interurban network is lower than in the reference situation. This reduction of the interurban frequency reduces the travel quality on the interurban network, especially for population $P_2$ who only use the interurban network. The higher quality of the urban network, however, reduces the travel time for the other two populations.

Figure 6-18 shows the comparison of the case of two operators having identical objectives to the single operator cases. In this comparison there is an interesting difference between the two design objectives. In the case of profit maximisation the quality of the urban network is reduced, leading to a higher profit for the urban network. The total profit for the public transport system, however, is reduced. In the case of welfare maximisation the differences in network characteristics are small, only a slight reduction of the frequency in the urban network, while the values for welfare are virtually identical. What happens here, is that due to the focus on single network optimisation, the effects of changes in the network on the usage of the other network are not accounted for. This mechanism fully applies to the case of profit maximisation, while in the case of welfare maximisation the concept of consumer surplus also includes the consumer surplus for population of travellers $P_{12}$ who use both networks.

*Table 6-25: Optimal values for network design variables and corresponding network characteristics for different scenarios with respect to number of operators and objectives*

| | Reference | Single operator | | Two operators | | |
|---|---|---|---|---|---|---|
| | | Profit total | Welfare total | Game profit | Game welfare | Game profit / welfare |
| **Stop spacing level 1** [m] | 400 | 833 | 717 | 950 | 728 | 732 |
| **Line spacing level 1** [m] | 500 | 1,397 | 847 | 1,996 | 894 | 916 |
| **Frequency level 1** [veh/h] | 4.0 | 6.8 | 10.2 | 4.2 | 9.4 | 9.0 |
| **Frequency level 2** [veh/h] | 4.0 | 1.8 | 3.0 | 1.7 | 3.0 | 1.7 |
| **Profit level 1** [€] | 113 | 346 | 60 | 408 | 139 | 109 |
| **Profit level 2** [€] | 210 | 1,271 | 964 | 1,095 | 960 | 1,384 |
| **Total profit** [€] | 323 | 1,616 | 1,024 | 1,502 | 1,099 | 1,493 |
| **Welfare level 1** [€] | 3,874 | 3,443 | 4,149 | 2,824 | 4,150 | 3,569 |
| **Welfare level 2** [€] | 4,218 | 3,928 | 4,860 | 3,401 | 4,807 | 4,084 |
| **Total welfare** [€] | 6,629 | 6,535 | 7,459 | 5,676 | 7,450 | 6,709 |
| **Travel time $P_1$** [min] | 30 | 31 | 25 | 39 | 26 | 26 |
| **Travel time $P_2$** [min] | 53 | 67 | 57 | 68 | 57 | 68 |
| **Travel time $P_{12}$** [min] | 68 | 86 | 65 | 98 | 66 | 83 |
| **Operational costs level 1** [€] | 557 | 243 | 641 | 101 | 554 | 517 |
| **Operational costs level 2** [€] | 3,290 | 1,474 | 2,489 | 1,398 | 2,468 | 1,398 |
| **Total operational costs** [€] | 3,847 | 1,717 | 3,130 | 1,499 | 3,022 | 1,915 |

The last scenario deals with the case of two operators having different objectives. For the urban network the objective is welfare maximisation and for the interurban network it is profit maximisation. This configuration is in line with the ideas of Berechman (1993) and is representative for the situation in the Netherlands. The results of the Cournot-game in this case differ, of course, from the single operator cases, as can be seen in Figure 6-19.

The characteristics of the urban network are similar to those for the single operator case maximising social welfare, while those for the interurban network resemble those for the single operator case maximising profit. The interesting thing that happens here, is that due to the quality of the urban network the travel time of population $P_{12}$, travellers using both networks, is lower than those for all other analyses using the objective of profit maximisation. As a result, the interurban network has more travellers without investing in its own quality, thus yielding the highest profit of all scenarios.
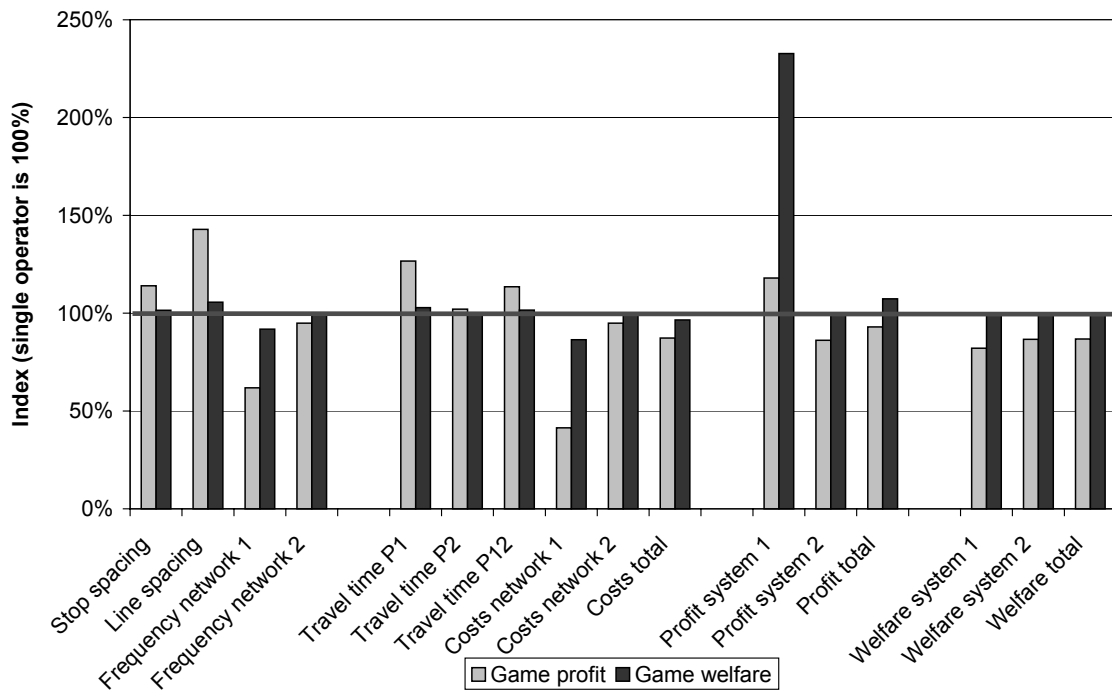
*Figure 6-18: Optimal network characteristics in the case of two operators relative to those for the case of a single operator*
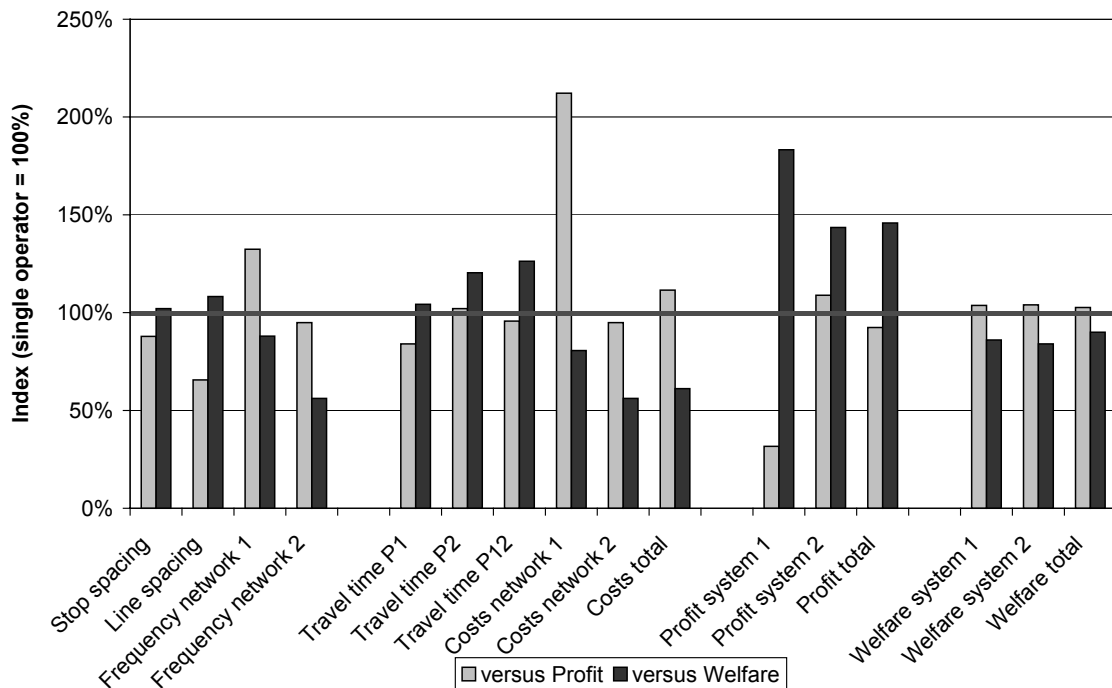


*Figure 6-19: Optimal characteristics in the case of two operators having different design objectives relative to the case of a single operator for two objectives*

### *6.5.3.4  Conclusion*

This analysis showed that organisational aspects influence multilevel public transport network design significantly. It is found that in the case of profit maximisation there is a rationale to subsidise the lower-level network compensating the lower profits or losses that are the result of offering the transport quality that leads to larger benefits of the higher-level network. This rationale is based on the fact that there is a reasonably sized population of travellers using both network levels. This finding is exactly the mechanism that is used in airline networks using a hub & spoke structure (see Section 3.5.2.2, Aykin (1995), O'Kelly & Brian (1998)).

If all network levels are designed using the objective of welfare maximisation, however, there is no need for subsidies between network levels. Welfare maximisation thus seems the most robust design objective. The impact of a population using both network levels is then limited to the higher weight for waiting time that is assumed for that population, leading to slightly coarser network and higher frequencies.
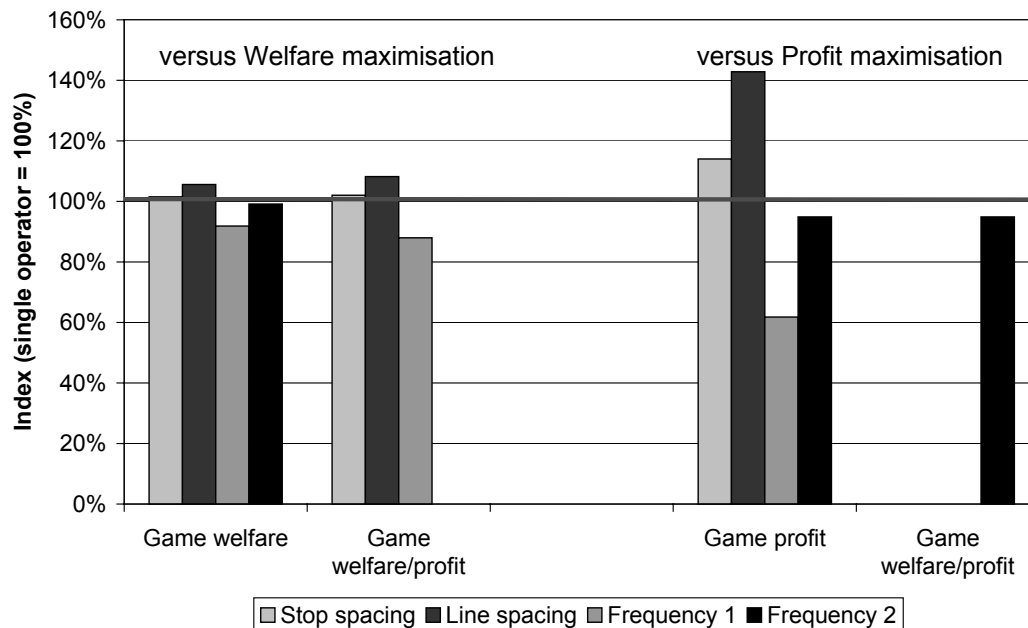


*Figure 6-20: Impact of organisational arrangements on the decision variables*

Interestingly, the main findings from this analysis deal with financial aspects such as direct or indirect subsidies between network levels. With respect to the decision variables in public transport network design, the impact of the organisation on the optimal values is limited for the objective of welfare maximisation and obviously larger in the case of profit maximisation. Figure 6-20 shows the differences from the single operator case for the decision variables only. The largest difference is found when comparing the decision variables for the urban network in the case of profit maximisation. In this case the design dilemma between space accessibility and time accessibility is judged differently depending on the impact of population $P_{12}$ having a higher weight for the waiting time. For the interurban network, however, the differences are small.

**6.5.4    Conclusions**

The analysis of regional and national line-bound public transport services does not lead to clear conclusions with respect to the mechanisms determining hierarchy, even though it is obvious that in practice several levels can be distinguished. Some evidence has been found implying a scale-factor 3 for the line spacing and a scale-factor 1.5 for speed might be applicable, but this is based on very specific assumptions. An economic analysis of superimposing a higher-level train service network on a regional bus service network, supports line spacing values found today. However, it also showed that this analysis is very sensitive to the parameter values used.

Since walking and cycling are important access and egress modes, it is likely that demand densities located within walking and cycling distance determine the allocation of stops at the higher-level. This is supported by the finding that focussing on a strict dependency between network levels, that is the lower-level network only provides access to the higher-level, leads to unrealistic scale-factors for the stop spacing. Spatial structures, again, determine the network structures in public transport.

Finally, it is found that organisational arrangements in multilevel networks significantly influence the network characteristics if design objectives other than social welfare are used. The important element in this case is the size of the population that uses several network levels compared to the populations using a single network level only. These consequences, however, are larger for the financial characteristics than for the decision variables in network design themselves.

## 6.6    A hierarchy for public transport service networks

The search for a mechanism determining hierarchy in public transport networks led to the conclusion that the hierarchy in spatial structures determines public transport hierarchy for both urban and interurban networks. In this section a classification of network levels is presented that is based on the hierarchical levels for spatial structures defined by De Jong & Paasman (1998) (see Section 4.4), and the insights gained in this chapter. First, the classification is presented. Then, it is compared with existing hierarchies in Dutch public transport service networks such as discussed in Section 6.2, which is followed by a discussion of the possibilities for new network levels.

**6.6.1    Classification of public transport network levels**

The classification of the hierarchical levels in line-bound public transport networks is based on the spatial hierarchy defined by De Jong & Paasman (1998). This classification has three network levels for urban areas and five network levels for interurban public transport networks.

The line spacing for urban networks is based on a linear network structure, an urban corridor, although in some cases a radial structure might be more appropriate. For the

higher urban network levels a radial network structure is assumed, in which case no values for line spacing are presented. The stop spacing equals twice the radius of the settlement size that is served. An example of the proposed urban network structure is given in Figure 6-21. An urban network for a city offers transport services at the level of neighbourhoods, while an express-system in an agglomeration offers transport services at the level of districts. The size of the network is thus two levels higher than the spatial level that is served. Furthermore, is should be noted that this approach to public transport service network design also leads to the possibility of an internal ring.
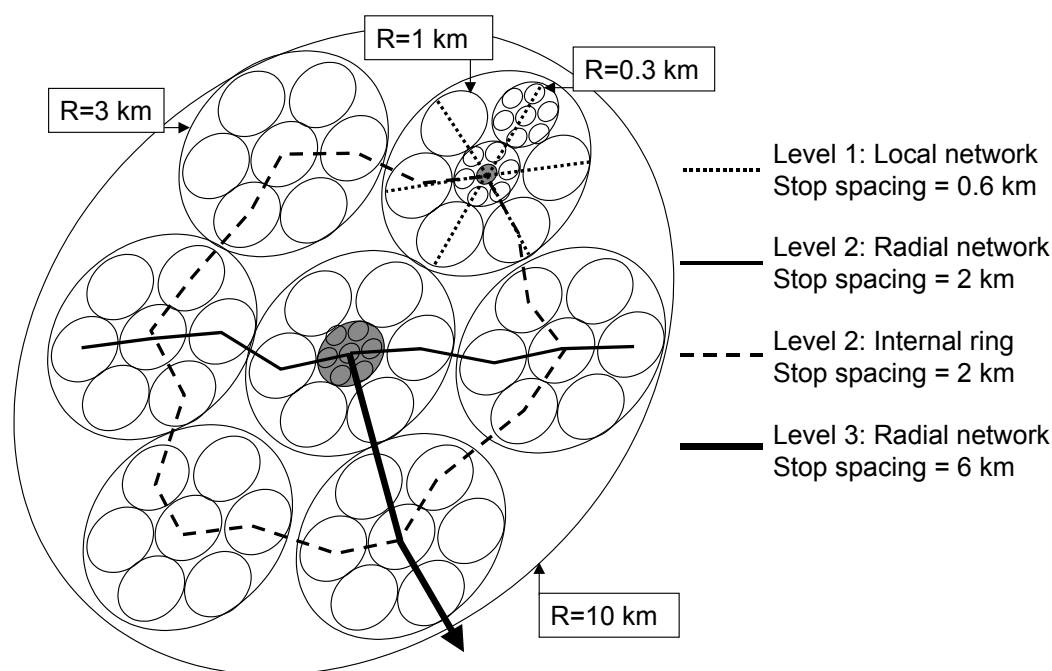


*Figure 6-21: Layout of a multilevel urban public transport network based on the hierarchy in spatial structures*

The line spacing for the interurban networks is based on a triangular network structure having a scale-factor of about 3. An important difference with the classification for road networks is that the access spacing equals the line spacing, instead of being one-third of the line spacing. This exception is due to the fact that an additional stop in a public transport service network leads to a time loss and thus leads to longer travel times, while for private transport networks an additional access node has no consequences for the travel time. Since no clear principles for the scale-factor for speeds have been found, two options are presented. The first is based on the concept of eliminating shortcuts in a triangular network as discussed in Section 6.5.1.1, while the second follows from the scale-factor for speeds for private transport networks.

These principles lead to the classification shown in Table 6-26, which is comparable to that for private transport networks shown in Table 5-7. Again, it should be noted that the numerical values represent global averages and not rigid standards.

*Table 6-26: Classification of public transport network levels*

| Network level | Spatial level | Line spacing [km] | Stop spacing [km] | Speed [km/h] Scale-factor 1.5 | Speed [km/h] Scale-factor 1.67 |
|---|---|---|---|---|---|
| **Urban (linear / radial networks)** | | | | | |
| **Urban** | Neighbourhood | 0,6 – 0,8 | 0,6 | 20 | 20 |
| **Express-services** | Districts | Radial | 2 | 30 | 35 |
| **Agglomeration-services** | 'City' | Radial | 6 | 45 | 55 |
| **Interurban (triangular networks)** | | | | | |
| **Local** | Village | 3 | 3 | 30 | 30 |
| **Regional** | Town | 10 | 10 | 45 | 50 |
| **Interregional** | City | 30 | 30 | 70 | 85 |
| **National** | Agglomeration | 100 | 100 | 105 | 140 |
| **International** | Metropolis | 300 | 300 | 160 | 235 |

### 6.6.2  Comparison with existing network hierarchies in the Netherlands

When this concept of hierarchy is compared with current network characteristics in the Netherlands (see Table 6-2 and Table 6-8), some interesting conclusions can be drawn. For urban networks the characteristics with respect to speed resemble those found in practice. However, the values for stop spacing following from this approach are, again, significantly larger than the values found in practice. It seems as if the location of stops is not only determined by the centres of the spatial units that are served, but also by the borders between adjacent spatial units, especially for the second network level of express services. Interestingly, the values proposed for the urban public transport network are reasonably in line with the results found using the analytical models used in this chapter.

For the interurban networks there is less similarity with existing network levels in the Netherlands. A typical phenomenon is that there are several types of transport services available at each network level. Demand responsive transport systems and regional buses serve the first level, focusing on transport to and from villages. The second level, oriented to towns, is also served by regional buses, but also by express bus services (Interliner) and local train services. Express and Intercity train services are the main public transport service for the third network level. Since there is currently no public transport service having an average speed of 140 kilometres per hour, the fourth network level seems to be missing, while high-speed train services are a typical example for the fifth network level.

The fact that different types of transport services are operational within a specific network level is no contradiction to the concept of a hierarchical public transport system. As long as the service quality required for a specific network level is guaranteed, it is not

relevant whether the service is provided by taxi, bus or train. In fact, the distinction of different transport services offers opportunities to provide cost efficient transport while providing adequate capacity.

Transport services operating at two network levels, however, conflict with the hierarchical network structure. The common description used in this thesis is that each network level serves settlements related to that level and at the same time is a feeder for the next higher-level network. This interpretation of a hierarchical network for private transport networks was illustrated in Figure 5-3.

If a transport service operates at two network levels, however, it becomes unclear which network level determines the network characteristics. Are they determined by the highest network level, while providing access to the highest network level as well, or are they determined by the lowest network level and are thus insufficient to provide the quality required for the highest network level? The disadvantage of the latter possibility is obvious: the quality of the higher network level is too low. In the first case, in which the transport is also a feeder for its own network level, the hierarchy can be as shown in Figure 6-22. The disadvantage of this approach is that in many cases the feeder function to its own network is mixed with the primary function of the lower-level network. Furthermore, the costs for providing transport service on the lower network level might be too high due to the requirements for the higher network level. Finally, the necessary stops at the lower level might reduce the quality for the higher-level network. This might apply to the average speed, but also to the level of comfort. Crowding due to facilitating short distance trips might reduce the comfort for the long distance trips. This is in fact a similar argument as that for private transport networks as discussed in Sections 5.2.2 and 5.7. Combining different network levels within single transport services thus lead to an ambiguous network concept, while the common definition used in this thesis results in clear network structures.
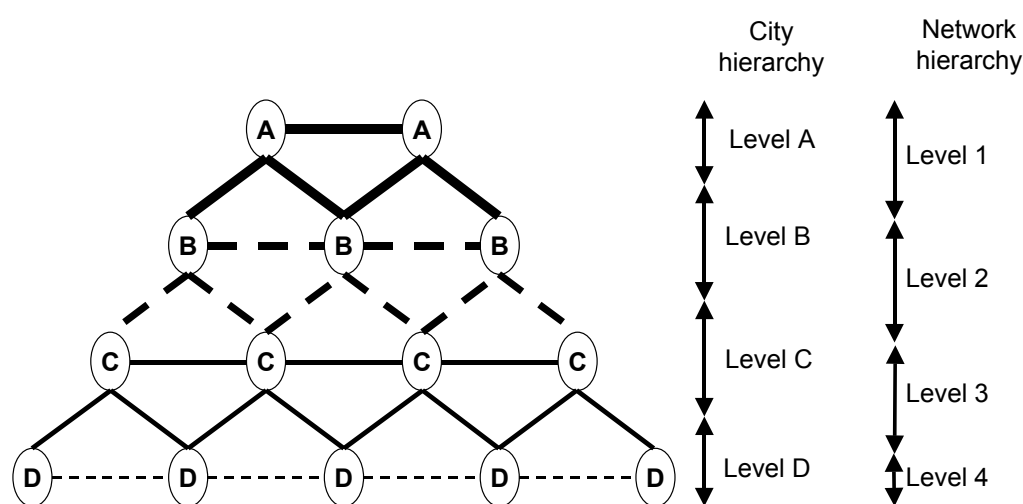


*Figure 6-22: Hierarchical public transport service network concept in which each level serves settlements of its own level and provides access to its own level (compare Figure 5-3)*

This argument may be illustrated by the Intercity network in the Netherlands. According to the classification proposed in this thesis, a separate network level can be distinguished for providing public transport services between agglomerations. In the Netherlands there are only three agglomerations that are large enough to be called agglomerations, that is Amsterdam, Rotterdam and The Hague. In the near future Utrecht and Eindhoven might also qualify.

Since the Intercity network is stated to be a higher network level than express-train services, it might me assumed that the Intercity network should serve these main agglomerations. In the Randstad area this is partly true. The Intercity network offers direct connections between these agglomerations. Only between The Hague and Amsterdam additional stops are included, which relate to a lower network level. Outside the Randstad, however, many other cities are included in the Intercity network, and in some cases even stops that might be classified as towns which are related to the regional network level. Furthermore, the travel speed offered, that is 95 kilometres per hour in the Randstad, is relatively low compared to the required speed for the national network, and is only slightly higher that the speed of the express-train network, being 80 kilometres per hour. Intercity train services today seem more a marketing concept than a separate network level in a hierarchical public transport service network from a functional point of view.

A similar line of reasoning can be applied to the Dutch Interliner concept, which stands for medium distance bus services having high quality buses and high quality stops. The characteristics of the Interliner services, however, differ substantially. The line length ranges between 24 and 137 kilometres, stop spacing between 1.8 and 19 kilometres, and the average speed between 24 and 64 kilometres per hour. Given these characteristics Interliner services might be classified as public transport services for the local as well as regional network level.

Finally, it is interesting to compare the average speeds for each network level with those found for private transport. If the speeds for the interregional level, 70 to 85 km/h, are compared to the average speed for the motorway network, which serves the same category of settlements, they are clearly lower. At first sight this gives public transport a weak position as an alternative to private transport. However, it should be noted that the public transport stops of the Interregional network are located in or near the city centre, while the access nodes of the motorway system are located at the periphery of cities. This emphasises once again the importance of public transport for access to city centres. Finally, public transport services allow for higher-level networks, even up to international networks, which do not exist for private transport networks.

### 6.6.3   Possibilities for new network levels

The finding that there is currently no transport service available at the national level in the Netherlands, raises the question whether such a network level should be developed. Examples of such public transport services can be found in other countries such as France and Germany having high-speed trains serving agglomerations and metropolises. In the

last decade two proposals for a new public transport network level in the Netherlands have been made.

Schoemaker et al. (1993) proposed the concept ARGUS, a high speed public transport service between the agglomerations of Amsterdam, Rotterdam, The Hague, and Utrecht, plus Schiphol Airport. The main objective in their design is to drastically reduce the travel time between the city centres to 20 minutes, thus enabling the Randstad to function as a true metropolis. They conclude that the existing railway network should be improved in order to facilitate higher maximum speeds, up to 200 kilometres per hour, and that a high-speed train service for the Amsterdam – Schiphol – Rotterdam corridor is required. Please note that such a network would eliminate the double function noted for the Intercity network.

The average speed for the latter train service proposed today as a part of the high speed train connection Amsterdam - Brussels - Paris is 120 to 140 kilometres per hour, a speed that matches the requirement for the fourth level proposed in this thesis. The proposed high-speed shuttle service between Amsterdam and Breda, however, does not fit with the proposed hierarchical concept. Since Breda can be classified as a city rather than an agglomeration, public transport services at the interregional level would be sufficient.

An alternative proposal for the Randstad area is that of Siemens Nederland N.V. (2000) for a very high-speed magnetic levitation train, the MAGLEV system. This system is also meant to connect the main agglomerations in the Randstad, but the stops are located in the periphery of the agglomerations and not in the city centres. For these peripheral locations the level of demand will be substantially lower due to the lower urban densities. The role of access modes other than walking and cycling will be more important, especially private car. The main idea is thus to compensate for the access costs by higher speed. The MAGLEV technology allows high average speeds in combination with relatively short stop spacing. The average speed for the Randstad system is 200 kilometres per hour, a speed that matches the requirements of the fifth network level.

Given the finding in this thesis that public transport service networks are strongly dependent on the hierarchy in settlements, it is questionable whether the peripheral locations are suitable for such a system. The demand that is generated by the high density of urban facilities in the city centre is not available. In order to have a demand level that is somewhat comparable, a very large number of travellers need to access the system by car. Since the accessibility of agglomerations is already a significant problem, the same will apply to the accessibility of the stops of the MAGLEV-system. Furthermore, the size of the parking facilities must not be underestimated. It might be argued that city centres will migrate towards these new locations, just as happened in the past in relation to railway stations. It should be noted, however, that the distances between city centres and railway stations were much shorter, the structure of the cities was different, and that in many cities today there is still a distinction between the city centre and the railway station. Such a development would anyway be a long-term outcome that would probably take much too long to make such a system profitable.

## 6.7  Conclusions

### 6.7.1  Summary

This chapter has focused on line-bound public transport networks, and especially on the relationships that determine hierarchy in such networks. Just as with private transport networks hierarchy is a common phenomenon in practice as well as in many design guidelines. The analysis in this thesis shows that such a hierarchy can not be explained just by the characteristics of public transport networks, as is possible for private transport networks (see Section 5.7). There are two characteristics in which public transport services differ from private transport that partly explain this finding. The first is the fact that in public transport services time accessibility plays a dominant role. Changing network levels implies transfers, which have a negative impact on the total travel quality. Higher-level networks should therefore not only compensate for the possible detour but also the necessary transfers. Second, there is an important difference in costs involved. In private transport networks the analysis deals only with roads, for which the investment costs per kilometre increases as the level of the network increases. In public transport service networks, however, operational cost need to be added as well, while there is also a clear difference in levels of investment cost between network levels. The lower-level public transport networks are generally road based networks, which have low investment costs since they use large parts of the private transport network, while the higher public transport level networks are based on dedicated rail networks, leading to high investment costs. In the analysis in this chapter, the investment costs are explicitly included in the analysis for public transport services having dedicated infrastructure. These factors make it plausible that the differences in design between subsequent public transport service network levels are larger than in private transport networks. This conclusion is, however, not in line with the network levels found in practice. There are thus other factors that influence hierarchy in public transport service networks.

This section discusses the main conclusion following from this chapter, with respect to single-level and multilevel public transport networks. Other subjects that will be discussed are the relationship between network levels and the consequences of special network configurations.

### 6.7.2  Single-level public transport service networks

For single-level urban public transport networks is has been found that currently used values for stop spacing are too low from a welfare perspective. Stop spacing should vary between 600 and 800 metres, depending on the city size considered. Line spacing should be about 800 metres, while peak frequencies should be 7 to 8 vehicles per hour. These coarser networks will yield shorter travel times, except for very short trip lengths, and especially lower operational costs.

The currently used values for stop spacing and line spacing might be explained from a historical perspective having smaller cities, higher levels of demand, and lower

operational costs, while actual frequencies might also be influenced by vehicle capacity, a characteristic that is not included in these analyses.

### 6.7.3   Multilevel public transport networks

An extensive analysis of hierarchy in urban networks shows that unless a hierarchy in demand densities is assumed, a single-level network is optimal from a social welfare perspective. The addition of a faster higher-level network has two effects resulting in a net negative impact on social welfare. The operational costs are increased leading to a potentially lower profit. Part of this increase is compensated for by a lower quality for the lower-level network. The resulting network consists of a lower-level network having lower frequencies, and a higher-level network, which does not have high frequencies either, and which has limited space accessibility. For travellers having an origin close to the stops of the higher-level network, the introduction of the higher-level network leads to shorter travel times, but for all other travellers travel times are increased. The net result is a reduction in social welfare. Only when the number of travellers who benefit from the higher-level network is increased substantially, does a hierarchical network structure lead to an increase in social welfare. The analysis showed that this higher demand should be twice or preferably three times as high as the uniformly distributed demand level assumed in the analysis. Hierarchical urban public transport services are therefore only justified in cities or agglomerations having a clear hierarchical spatial structure.

For train services, which are usually related to higher-level interurban networks, it has been found that walking and cycling are dominant access and egress mode. This suggests that spatial structures influence the hierarchy in public transport networks. Furthermore, the analysis of interurban public transport service networks shows no evidence of relationships explaining hierarchy as a natural characteristic of public transport networks themselves. Only for a very specific case a scale-factor of 3 for line spacing and a scale-factor of 1.5 for speed can be found, while other analyses show that the network characteristics tend to very large scale-factors for line spacing. These findings lead to the conclusion that for hierarchical interurban networks the main explanation can again be found in the hierarchical organisation of settlements. As such the findings in this chapter support the ideas in Germany on network design (Köhler (1989), Bierschenk & Keppeler (2000)).

The finding that spatial structures are decisive when it comes to explaining a hierarchy for public transport networks, leads to a classification for public transport network levels that consists of three levels for urban areas, while for interurban public transport five levels are distinguished (Table 6-26).

### 6.7.4   Relationships between network levels

The system of network levels proposed in Section 6.6.1 is based on the hierarchy in spatial structures, since no relationship between network levels could be established based only on the network characteristics of line-bound public services. The combination of spatial hierarchy and hierarchical public transport service networks, however,

determines a relationship between network levels with respect to the location of the transfer nodes between network levels. Furthermore, as has been found in this chapter, financial relationships between network levels do exist.

The analysis of existing multimodal transport in Section 2.3 shows that private modes, and especially walking and cycling play an important role in multimodal transport. The importance of public transport as an access mode, however, is larger at the activity-based part of the trip. Given the fact that city centres and access nodes of higher-level networks are usually located close together, it is possible to transfer between various network levels at the same location (see also Section 4.2.3). Such a network configuration reduces the number of transfers and increases the service area of the higher-level networks. It might thus be concluded that although there is no clear relationship except spatial concentrations that determines hierarchy for public transport service networks, it is important to combine the central nodes of the different network levels at the same location. If these transfer nodes are located in or next to city centres, this type of relationship is beneficial for all network levels involved.

Another important finding in this chapter is that, depending on the organisational structure, financial relationships between network levels exist, especially if the objective of profit maximisation is used for one or more network levels. If profit maximisation is used for all network levels, it is interesting for the operators of the higher-level networks to subsidise the lower-level network. Due to the resulting higher quality of the lower-level network, the higher-level networks are more attractive for travellers who use the lower-level network to access their networks, leading to a higher total profit. An interesting situation appears if different objectives are used. The most likely situation then is that welfare maximisation is used for the lower-level network, while profit maximisation is used for the higher-level networks (Berechman (1993)). In that case, the operators of the higher-level networks benefit from the resulting higher quality of the lower networks too, but without investing in the lower-level network themselves.

### 6.7.5   Specific network configurations

Part of the controversy between the theoretical concept of hierarchy as proposed in this thesis and current public transport networks as discussed in Section 6.6.2 can be explained by typical network configurations based on an operational perspective, capacity constraints, and the desire to eliminate transfers. Connecting two radial lines to an agglomeration, and thus creating a transversal line, does eliminate a transfer for travellers who have an origin on one radial line and a destination on the other. Furthermore, the fact that transversal lines do not require all kinds of operational facilities at the main stop in the city might also have operational advantages. This does not mean, however, that the line is designed for trips from one end of the line to the other end.

Connecting radial lines, either at the main stop agglomerations or, the opposite, in the middle of two agglomerations, also explains the existence of stops on the periphery of the agglomeration. These stops are natural if a radial line serving the city centre is considered, but might not be appropriate if the whole transversal line is examined.

Another example is the combination of transport services of different network levels, such as a train service between agglomerations, for instance between The Hague and Utrecht, and an express train service between cities, e.g. Utrecht – Groningen. The resulting train service is a typical example of the zone-system, which is analysed for an urban network in Section 6.4.3.3. As was shown then, a zone-system has some interesting characteristics from the traveller's perspective as well as from the operator's point of view. For interurban networks, however, trips are less clearly oriented to a single destination, which might reduce the possible benefits. Bouman et al. (2000) found that for rail services a strict hierarchy leads to the best performance for current travellers as well as for future travellers. It should be noted that the zone-system was the best of the other service concepts that were analysed.

Combinations of urban network levels with interurban network levels might also have positive benefits (Bruijn (1994)). But again this raises the question, which network level is decisive for the system requirements. Furthermore, since the main advantage is that transfers are eliminated, it is essential to provide competitive travel times by limiting the number of stops in the urban area. The difference in network levels should not be too large.

Finally, systematic usage of a zone system might lead to an ambiguous concept of network hierarchy in which a network level serves settlements related to that level, is a feeder for its own network level, while also providing transport services at a lower-level. Such a combination of transport functions might be unnecessarily expensive and might lead to a reduction of the quality offered for the higher-level network trips. Furthermore, it limits the possibilities for developing coherent lower level networks. Zone-systems are therefore an interesting concept to consider, but it is important to be aware of their possible disadvantages.

# Chapter 7

# MULTIMODAL TRANSPORT NETWORK DESIGN

## 7.1 Introduction

After defining the multimodal transport network design problem in Section 3.5, the concepts of multilevel transport networks were extensively analysed for private transport networks (Chapter 5) and for line-bound public transport service networks (Chapter 6). This chapter returns to the main topic of this thesis, that is, multimodal transport network design. The analysis in this chapter is divided into two parts.

The first part summarises the findings on both transport network types and on the estimated potential of multimodal transport presented in earlier chapters. Then it establishes the implications for multimodal transport networks. The elaboration focuses on the hypothesis formulated in Section 3.5, namely that the combination of modes or transport services influences the network characteristics of the related transport services and also the hierarchy in transport networks. The discussion will show that unimodal transport networks are robust in the sense that multimodal travel has a limited impact on the transport network in general. Furthermore, the consequences of considering alternative access modes for urban public transport network design will be analysed. The analytical transport network design model for urban public transport networks will be extended to incorporate access mode-choice. The analysis will show that walking remains the most important access mode for urban public transport services.

The second part of the analysis deals with typical transport services that are often mentioned as important elements in multimodal transportation, that is rental services and demand responsive transport systems. Rental services for cars or bicycles might eliminate the differences in vehicle availability between the home-based and the activity-based part of the trip (see Section 2.3.1) and might thus improve the attractiveness of activity oriented transfer points in a multimodal transport system. This is especially relevant in

cases where the quality of public transport services is not high enough to compensate for the limitation in the availability of private modes.

Demand responsive transport systems are an alternative for the limited accessibility in space and time of traditional line-bound public transport services. Demand responsive transport might eliminate or reduce these limitations. The fact that these transport services are fully demand oriented makes them very suitable for tailor-made multimodal trips. The Traintaxi in the Netherlands, a share-a-cab service to and from railway stations, has already shown that such transport services have a clear potential in multimodal transportation, especially for the activity -based part of the trip (De Bruijn (1998)).

It is interesting to note that both types of transport services lack a distinct network, since they use the road network. Rental services introduce a new limitation of the accessibility of the network, which is determined by the location of the rental services. In the case of demand responsive transport services the accessibility of the network is ideally comparable to that of private transport services. In practice, however, the need to make a reservation still leads to a substantial difference with private transport services. Furthermore, the quality of the transport service itself will vary with actual demand patterns in space and time.

The analysis of these demand oriented transport services will show that while they can improve the attractiveness of multimodal transport, they have no impacts on multimodal transport network design.

Finally, the consequences of these findings for multimodal transport network design are discussed, resulting in a focus on the location of transfer nodes and the role and characteristics of more detailed network design models.

## 7.2   Synthesis

In formulating the multimodal transport network design problem in Section 3.5 it was hypothesised that combining unimodal transport services would influence the characteristics of the transport service networks themselves, especially the characteristics of the hierarchy in transport networks. In order to verify this hypothesis, the characteristics of the two main transport network types, that is private transport networks and line-bound public transport services, were analysed in Chapters 5 and 6 respectively. The findings of these analyses have clear consequences for multimodal transport network design.

### 7.2.1   Private transport

Private transport networks were found to have self-organising characteristics that determine the hierarchy of higher-level networks (Chapter 5). This mechanism leads to scale-factors for road spacing and network speed, implying that the road spacing of each consecutive level is three times larger, while the corresponding network speed should be

1.67 times larger. Interestingly, the scale-factor 3 for road spacing matches the concept of hierarchy for spatial structures defined by De Jong (1988a, 1998b), which has a scale-factor 3 for the radius of settlements at consecutive levels (Section 4.3).

Given the fact that fundamental mechanisms, following from travel behaviour and economic principles, govern these self-organising properties the concept of multimodal transport will not have any consequences for the main characteristics of private transport networks. Furthermore, the limited potential of multimodal transport, as assessed in Section 2.4, also resulted in a limited influence on second order network characteristics such as capacity. There are, however, two exceptions to this finding. Since multimodal transport is especially relevant for trips to and from city centres, multimodal transport might reduce the use of roads to and from the city centre and might slightly reduce the level of congestion on the highways to and from the main cities.

The lower-level private transport networks are mainly determined by factors such as land use, urban planning, and the requirements of urban systems for water, energy and telecommunications. These factors for instance explain the differences in characteristics between urbanised and rural areas. Urbanised areas have fine-grained networks and low network speeds, while in rural areas the network density is clearly less and network speeds are higher. The main mechanisms determining network hierarchy, however, should not be ignored. These mechanisms determine whether the network will be used as intended or not.

## 7.2.2   Public transport

For line-bound public transport networks no evidence has been found for self-organising mechanisms determining network hierarchy. This finding might easily lead to the conclusion that combining alternative modes and line-bound public transport results in different network characteristics. It was also found, however, that hierarchy in line-bound public transport networks is mainly determined by hierarchy in spatial structures. This finding is supported by the fact that for train services, which are to a large degree equivalent to higher-level line-bound public transport services, many other modes are used for access and egress. Put into other words, higher-level line-bound public transport networks are in fact already multimodal transport systems. Excluding other modes such as private car or bicycles for access and egress would seriously reduce the patronage for these higher-level networks.

Since the hierarchy in spatial structures changes slowly over time, the hierarchy in public transport networks seems just as robust as the hierarchy in private transport networks. Thus the initial conclusion that combining alternative modes and line-bound public transport influences public transport network characteristics no longer holds for higher network levels. The question with respect to the lower-level public transport network, however, is still unanswered and is analysed further in Section 7.3.

The finding of a strong relationship between hierarchy in spatial structures and public transport network levels might seem old-fashioned. There are, for instance, scenarios in which sub-urbanisation is dominant leading to lower densities which make it impossible

to provide efficient lower-level public transport networks (Goudappel Coffeng & Beleidsonderzoek en -advies (1998)). The consequence is that other modes are called for to fulfil the function of lower public transport networks, such as private car, bicycles and demand responsive transport systems. Another finding in their study, however, is that even in a sub-urbanisation scenario, there will still be concentrations of urban functions that provide good possibilities for higher-level public transport networks, although the level of patronage might be lower, leading to less emphasis on capacities. The hierarchy concept for public transport networks itself, however, is also suitable for scenarios assuming lower densities and less hierarchy in spatial structures.

### 7.2.3 Multimodal transport

The overall conclusion following from the analyses of private transport networks and public transport networks is that multimodal transport does not require a new approach for designing transport networks. Properly designed unimodal transport networks facilitate multimodal transport as well as unimodal transport. The proposed network levels for private transport and line-bound public transport networks are summarised in Table 7-1.

*Table 7-1: Network levels for private transport and line-bound public transport (in case of different characteristics: Private transport / Public transport)*

| Network levels | Related spatial level | Network size | Road and line spacing [km] | Network speed [km/h] |
|---|---|---|---|---|
| **Urban** | | | | |
| Street / Urban | Neighbourhood | City | 1 / 0,6 – 0,8 | 20 |
| Arterial / Express service | District | Agglomeration | 3 / radial | 35 |
| Expressway / Agglomeration service | 'City' | Metropolis | 10 / radial | 55 |
| **Interurban** | | | | |
| Local | Village | Small region | 3 | 35-40 / 30 |
| Regional | Town | Region | 10 | 60 - 70 / 50 |
| Interregional | City | Province | 30 | 100 - 120 / 85 |
| National | Agglomeration | Country | 100 | - / 140 |
| International | Metropolis | Part of continent | 300 | - / 235 |

The proposed network levels can be regarded as guidelines for transport network design. Please note that the values for road spacing, line spacing, and network speed represent global averages and should not be interpreted as rigid standards. While these guidelines are not always different from existing guidelines, the type of support for these guidelines is quite different. The guidelines developed here follow from extensive analyses of the characteristics of transport networks. Some of the existing guidelines are based on current practice, which given the self-organising properties of private transport networks might be in line with these new guidelines, such as the scale-factor for the path spacing in bicycle networks (Bach (1999)). The analysis of urban public transport networks, however, shows that guidelines on for instance stop spacing might be outdated due to changes in urban structures and related costs. Other guidelines, such as those advocated in Germany for interurban transport networks (e.g. FGSV (1988)), are based on the a priori distinction in spatial structures. The analyses in this thesis, focusing primarily on network characteristics, show that this distinction is indeed proper. As such, this thesis provides additional theoretical support for the German guidelines for private transport and public transport network design (VÖV (1981), FGSV (1988), Köhler (1989), Schönharting (1997), Bierschenk & Keppeler (2000)). This theoretical support, however, also identifies the need to make a clear distinction between network levels. While the German guidelines allow the possibility of having small differences in speed between network levels, these guidelines clearly recommend substantial differences in quality in order to exploit fully the benefits of hierarchical network structures.

The estimate of the future potential for multimodal transport given in Section 2.4 showed that improvements in multimodal services might lead to a substantial increase of public transport usage. If this growth were to be accommodated by train services alone, train trips between 10 and 30 kilometres would be tripled, while trips between 3 and 100 kilometres would be doubled. Improved multimodal transport thus implies a capacity problem for line bound public transport. It is more likely, however, that due to the limited rail infrastructure the existing train service network is not suited to accommodate the extra demand, and that additional transport services are necessary. Such transport services probably need dedicated infrastructure in order to meet the standards for the associated network levels, which makes them difficult to develop.

### 7.2.4   Sensitivity

The basic principle determining hierarchy in private transport networks, that is eliminating short cuts, and the simple rule that a hierarchy in spatial structures is fundamental for hierarchy in public transport networks, are so elementary that this classification is very robust. The values for the associated qualities, however, are more likely to differ. The stochastic nature of the location of settlements influences the actual road and line spacing of the networks. Furthermore, the main principle used in the analysis is the economic concept of minimising total costs or maximising social welfare. The findings of the numerical examples are thus sensitive to the parameter values. Although all values used in the analyses are based on reasonable assumptions, it is obvious that the parameter values are subject to discussion. This is especially true for the parameters related to infrastructure investments. There is a tendency for the costs of infrastructure to increase, for instance due to the increase in required quality and

environmental concerns, making it more difficult to develop higher-level networks. This increase in costs might be countered by higher demand densities or by including other benefits in the analysis, such as the benefits of viable and attractive city centres. Overall, it is expected that building higher-level infrastructure will become more difficult, thus hampering the development of balanced network structures and leading to heavily used networks.

The speeds at each level depend on the scale-factor for speeds, which is again quite robust, and on the assumption of the speed at a specific network level. It must be noted that travel time is used as an approximation of the more appropriate concept of generalised costs. The differences in speed between network levels may thus be less if there is a decrease in costs and vice versa: Introducing a toll for higher-level network usage requires a larger difference in speeds.

These restrictions on the analyses emphasise that the classification of network levels and the associated qualities should be seen as general guidelines and not as strict rules that have to be obeyed at all times.

### 7.2.5   Transfer nodes

Since higher-level line-bound public transport networks are already a multimodal transport system, the concept of multimodal transport will not substantially change in public transport network design. However, it strengthens the importance of the accessibility of access nodes for higher-level public transport networks, including parking facilities for private car and bicycles. The dependency between hierarchy in spatial structures and public transport networks, however, determines the location of these access nodes.

The importance of accessibility and the location of these access nodes lead to an interesting design dilemma. From a multimodal transport point of view the accessibility by private car is important, while the higher densities related to hierarchy in spatial structures make it very expensive to provide adequate facilities for parking private cars in city centres. Furthermore, these higher densities set high standards for the environmental quality in city centres, which limit the possibilities to provide high quality access by private car.

Alternatively, the opposite approach of introducing new mostly peripheral access nodes to higher-level line-bound public transport networks which are primarily based on access by private car, lacks all the benefits provided by the concentration of activities found in urban city centres. Egeter et al. (1990) estimated that such a transfer node would generate 2,000 travellers per day at most, which is very small compared to an average access node of the regional public transport network. If all travellers needed for an economically sound operating transport system would arrive by private car, the size of the parking space would be enormous. On the other hand, an analysis of the potential of a transfer point in the national train service network in the Netherlands showed that given very high standards for the public transport system, the number of car drivers willing to transfer to

public transport was still very limited (Mu-Consult (1998)). New access nodes based on car access alone do not seem to be a very fruitful solution.

The accessibility of access nodes of higher-level public transport networks for bicycles is easier. Bicycles require less space for parking facilities and bicycles are an accepted environmentally benign transport mode in city centres.

## 7.3   Impact of alternative access modes on urban public transport networks

Although combining alternative modes and higher-level public transport does not influence the network hierarchy of line bound public transport networks, it increases the patronage of these networks. For the lower-level public transport network, however, the question remains whether combining alternative access modes may influence urban public network design. At first sight, it should. If, for instance, bicycles are used to access urban public transport, the access speed is four times higher than the speed for walking. If this access speed is used to determine the optimal network characteristics for the tram network analysed in Section 6.3.3, the impact on the network and its performance is substantial. Stop spacing is nearly twice as large, line spacing more than twice as large, while the optimal frequency becomes 12 vehicles per hour. The weighted travel time is about 40% lower, as is the operational cost. Social welfare is 26% higher. This approach, however, is clearly too simplistic. Travellers may choose between using a bicycle to access public transport and walking to the stop. Given the resulting values for stop and line spacing, the maximum access distance becomes about 850 metres, which makes it likely that many travellers will prefer to walk. Therefore, a more detailed analysis is needed to assess the impact of alternative access modes.

An option would be to follow the suggestion by Jung (1996) who used the average access time for all access modes to determine the optimal network characteristics. In his study, however, Jung used scenarios with respect to the modal split for the access modes, and did not consider a more detailed description of travel behaviour such as access mode choice. Furthermore, Jung focussed on minimising travel time only, and did not use the preferred objective of maximising social welfare. The following section presents an analytical model that explicitly accounts for access mode choice, using the objective of maximising social welfare. This approach is an improvement of the method used by Van Nes (2001).

### 7.3.1   Incorporating access mode choice

In this section the analytical model developed in Section 6.3.1 is extended to account for alternative access modes. Travellers thus make two different choices: they choose to travel by public transport or not and to walk or to cycle to the access node (Figure 7-1). The approach used in this analysis is similar that used to describe travel behaviour or an express-system (Section 6.4.3.1). The difference is that for an express system travellers may choose between two different routes between origin and destination, while in the

case of alternative access modes travellers may only choose how to access the public transport network. The trip time elements within the public transport system, that is waiting time, in-vehicle time and egress time, are equal for both alternatives. Thus, in this analysis the logsum, which in the case of express services was used to determine the aggregate travel time for express services, should only apply to the access part of the trip.
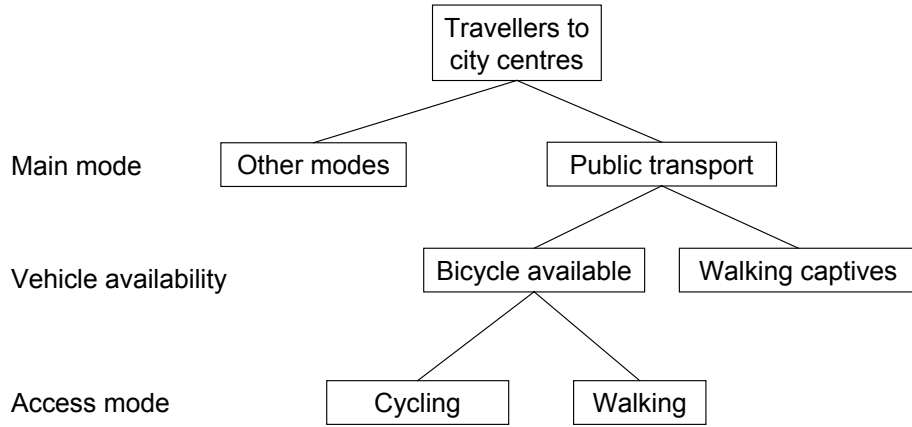


*Figure 7-1: Travel choices in the case of cycling as an alternative access mode for urban public transport*

The choice between walking and cycling can be described using a logit-mode choice model as shown in Equation (6-8). It is assumed that the choice between walking and cycling is primarily determined by the travel times for each mode having a sensitivity $\alpha_w$ for walking and $\alpha_b$ for cycling. Furthermore, a mode specific constant for cycling $\beta$ is introduced to account for the different perception of both modes. Using data from the Dutch National Travel Survey an estimate of the coefficients $\alpha_b$, $\beta$, and $\alpha_w$ can be made, yielding $\alpha_b$=0.07, $\beta$=-1.25 and $\alpha_w$=0.11.

These values, however, describe only the choice between walking and cycling as separate modes, and not the choice between two access modes to urban public transport. It seems reasonable to assume that the sensitivity to the travel times will not be influenced by whether or not it is an access mode. Therefore a time penalty for using a bicycle as an access mode is introduced. Such a penalty might represent just the time needed for getting and parking a bicycle or the subjective time penalty related to leaving the bicycle unattended at the stop. Furthermore, the model assumed that all travellers may choose between both modes. In practice, however, there is a population who are not able to use a bicycle, either because there are physical constraints or simply because there is no bicycle available. Thus the percentage of travellers using a bicycle for accessing urban public transport can be written as:

$$p_b(L_a) = \frac{\exp\left(-\alpha_b \cdot \left(\dfrac{L_a}{V_b} + T_{pb}\right) + \beta\right)}{\exp\left(-\alpha_b \cdot \left(\dfrac{L_a}{V_b} + T_{pb}\right) + \beta\right) + \exp\left(-\alpha_w \cdot \dfrac{L_a}{V_w}\right)} \cdot (1 - p_{cw}) \qquad (7\text{-}1)$$

where:

$p_b$ = share of public transport travellers using a bicycle

$L_a$ = access distance which is a function of stop spacing and line spacing

$\alpha_b$ = demand sensitivity to cycling time

$\beta$ = mode specific constant for cycling

$\alpha_w$ = demand sensitivity to walking time

$V_x$ = average speed for mode $x$

$p_{cw}$ = share of public transport travellers that is captive to walking

$T_{pb}$ = time penalty for using a bicycle as an access mode

The fact that a percentage of the travellers might use a bicycle reduces the average access time for trips to the city centre. The impact of the introduction of an alternative access mode can be described using the logsum (see also Section 6.4.3.1). The (dis-) utility for these two alternatives is given by:

$$\tilde{U} = \overline{U} + \frac{1}{\mu} \cdot \ln\left( \frac{\exp(\mu \cdot (U_b - \overline{U})) + \exp(\mu \cdot (U_w - \overline{U}))}{2} \right) + \frac{1}{\mu} \cdot \ln(2) \tag{7-2}$$

where:

$\overline{U}$ = average utility for walking and cycling

$\tilde{U}$ = resulting utility using the logsum

$\mu$ = scale parameter

$U_b$ = utility for cycling = $-\alpha_b \cdot \left( \dfrac{L_a}{V_b} + T_{pb} \right) + \beta$

$U_w$ = utility for walking = $-\alpha_w \cdot \left( \dfrac{L_a}{V_w} \right)$

In the formulation of the objective functions for determining the optimal values for the stop spacing, line spacing and frequency decision variables (see Section 6.3.1), however, access time is used instead of the utility from the access mode choice-model. This utility can be transformed into access time by dividing it by $\alpha_w$. Thus for travellers able to choose between walking and cycling the access time is given by:

$$\tilde{T}_a = \frac{\overline{U}}{\alpha_w} + \frac{1}{\mu \cdot \alpha_w} \cdot \ln\left( \frac{\exp(\mu \cdot (U_b - \overline{U})) + \exp(\mu \cdot (U_w - \overline{U}))}{2} \right) + \frac{\ln(2)}{\mu \cdot \alpha_w} \tag{7-3}$$

where:

$\tilde{T}_a$ = aggregate access time for the access modes walking and cycling

In this analysis it is further assumed that bicycles are used for public transport access only. They are parked at the stop, and are not taken along in the bus or tram to be used for egress as well. Thus, all equations for travel time, patronage, and revenues defined earlier in Section 6.3.1 can be used to formulate the components for the objective function for

maximising social welfare for travellers who can (in which Equation (7-3) replaces Equation (6-2)) and who cannot choose between walking and cycling (using Equation (6-2)). The formulation of the operational costs depends solely on the decision variables themselves (Equation (6-13)).

Using this formulation of the network design problem, the optimal values for the decision variables can be calculated numerically. The problem, however, is that little is known of the values of the new parameters that were introduced, that is, the time penalty for using a bicycle to access public transport ($T_{pb}$), and the percentage of travellers not able to use a bicycle ($p_{cw}$). A rough estimate of these values is made based on two observations while assuming that $\mu = 1$:

- Van der Waard (1988) found that nearly 40% of travellers using urban public transport claimed not to have a bicycle available for that specific trip. This finding results in a value for $p_{cw}$ of 0.40;

- In urban bus networks in the Netherlands the use of bicycles as an access mode is insignificant. Given the stop and line spacing for the reference situation used in the application of the single-level network optimisation model and the access mode choice model defined in Equation (7-1), this leads to a time penalty of 35 minutes.

It is interesting to note that in a study of penalties in intermodal transfers in the USA (Liu et al. (1998)), the transfer penalty from car to train was thrice as high as the traditional transfer penalty. Given a transfer penalty of 13 minutes (Central Transportation Planning Staff (1997)) the intermodal transfer penalty would then be 39 minutes, which is quite close to the proposed value of 35 minutes.

### 7.3.2 Application of the access-mode choice model

Using the model and the values for the new parameters described above, the optimal values for stop and line spacing are determined for an urban corridor served by a tram service and having an average trip length of 5 kilometres. Since the values for the parameters used in the base scenario are clearly rough estimates, the analysis also includes scenarios in which the values of these parameters are reduced by 50%. In order to find out what the maximum impact of combining cycling and urban public transport could be a maximum scenario is analysed having a time penalty of only 2 minutes and a share of captives of 20%, which is nearly equivalent to the share of elderly. The results of this analysis are shown in Table 7-2. The optimal network characteristics for the case of walking only are included as a reference (see Table 6-7).

*Table 7-2: Optimal network characteristics for a tram network in the case of cycling as an additional access mode (trip length is 5 kilometres)*

| | Reference | Walking and cycling as access modes | | | | |
|---|---|---|---|---|---|---|
| | Optimal for walking only | Base | Penalty -50% | Captives –50% | Penalty and captives –50% | Penalty 2 min captives –50% |
| **Penalty for cycling** [min] | - | **35** | **17.5** | **35** | **17.5** | **2** |
| **Percentage walking captives** [%] | - | **40%** | **40%** | **20%** | **20%** | **20%** |
| **Stop spacing** [m] | 757 | 764 | 778 | 766 | 785 | 849 |
| **Line spacing** [m] | 797 | 804 | 820 | 806 | 829 | 942 |
| **Frequency** [veh/h] | 7 | 7 | 7 | 7 | 7 | 8 |
| **Social welfare** [€/km$^2$] | 779 | 786 | 801 | 788 | 808 | 855 |
| **Total costs** [€/km$^2$] | 485 | 480 | 468 | 478 | 463 | 425 |
| **Percentage walking** | 100 | 98 | 92 | 97 | 90 | 75 |
| **Weighted travel time walking only** [min] | 32.3 | 32.4 | 32.5 | 32.4 | 32.6 | 33.0 |
| **Weighted travel time walking and cycling** [min] | - | 31.5 | 29.8 | 31.6 | 29.9 | 25.4 |
| **Patronage** [/km$^2$] | 151 | 152 | 153 | 152 | 154 | 158 |
| **Revenues** [€/km$^2$] | 86 | 86 | 87 | 86 | 87 | 89 |
| **Operational costs** [€/km$^2$] | 85 | 84 | 82 | 84 | 81 | 79 |
| **Profit** [€/km$^2$] | 0 | 2 | 5 | 2 | 6 | 10 |
| **Total costs per kilometre travelled** [€] | 0.64 | 0.63 | 0.61 | 0.63 | 0.60 | 0.54 |

*Note: Exogenous input values are shown in bold print*

The first conclusion from this analysis is that in the base case, the optimal network characteristics are nearly identical to those for the case of walking only. Given the values assumed for the time penalty for using a bicycle and the percentage walking captives, cycling is apparently not an interesting alternative to walking. Only 2% of the travellers use a bicycle to access the tram system, even though the access distance is clearly larger than that assumed for the traditional bus network used to determine the penalty for using a bicycle as an access mode. Due to the shorter travel times of travellers using a bicycle, patronage is slightly higher. The operational costs are a little lower, and social welfare is slightly higher.

The next two scenario's show that reducing the time penalty has a larger impact on the network characteristics than reducing the share of walking captives, leading to 8%

cycling and 3% cycling respectively. The combination of both assumptions still leads to small differences compared to the reference situation. The optimal values for stop and line spacing are 4% higher, leading to slightly longer travel times for travellers who can only walk. The travel times for travellers who are free to choose are 7% lower, even though only 10% of the travellers use a bicycle. Operational costs are 5% lower. The maximum scenario shows that 75% of the travellers still walk to the stop. The differences with the reference situation, however, are clearly larger. The optimal values for the decision variables, including the frequency, are 12% to 18% higher. Travel time is reduced by more than 20%, that is for travellers who can chose between walking and cycling, while operational costs are 7% lower.

Given realistic parameter estimates for the access-mode choice model the introduction of cycling as an alternative access mode has a minimal influence on the optimal network decision variables. The scenarios in which these parameters are reduced by 50%, clearly a generous assumption, still result in only small changes. Only in the obviously unrealistic case of a very small time penalty, is the resulting network less dense and the frequency higher. The resulting network, however, is closer to the reference network of walking only, than in the case that all travellers would use a bicycle, as discussed in the beginning of Section 7.3. The main explanation for this finding is that the access distance will always remain too short to make cycling an attractive alternative to walking as an access mode. The share of travellers that are able to cycle and that are willing to use a bicycle to access urban public transport, thus is too small to make a difference in urban public transport network design.

### 7.3.3   Conclusion

The impact on urban public transport network design of introducing cycling as an alternative access mode is small. This finding also implies that the impact of other alternative access modes or transport services will be even less. Transport services such as peoplemovers or demand responsive transport systems introduce limitations in space and time, which lead to penalties that are probably larger than for cycling. In the case of peoplemovers, the fixed routes might also lead to an increase of the percentage walking captives. The concept of multimodal transport thus has no influence on the design of urban public transport networks.

This conclusion does not mean that no attention should be given to the combination of cycling and urban public transport. As was shown in the base scenario including alternative access modes reduces travel times slightly. Providing parking facilities for bicycles at specific locations, for instance at stops farther away from the city centre or at stops having large access distances, will certainly make urban public transport more attractive. The impact, however, is too small to justify different network structures. Walking is and will be the dominant access mode for the lower-level urban public transport network.

Finally, it should be emphasised that these conclusion are only relevant for single-level urban public transport network, or in the case of multilevel public transport networks also

for the lower-level network. For higher-level public transport networks, however, cycling is an important mode for access as well as for egress.

## 7.4 Rental services

Rental services of cars and bicycles increase the vehicle availability for the home-based part and especially at the activity-based part of the trip. It is therefore an interesting component of a multimodal transport system. The only network characteristics related to rental services are access density, that is the density of rental facilities, and time accessibility in terms of opening hours and perhaps a waiting time if there is no vehicle available. Once a traveller has a vehicle he uses available the infrastructure network, such as the bicycle network or the car network. As such, rental services have no real relationship with network design in general. There is, however, the question of the possible impact of this type of transport services on multimodal transport.

### 7.4.1 Expectations

In their study on the impact of integrating cycling and public transport, for example, Van Goeverden & Egeter (1993) concluded that such an integration might lead to 14% more kilometres travelled by public transport. The largest growth is found for trips of between 10 and 40 kilometres long. The impact of cycling at the activity-based part of the trip is twice that of the home-based part. A limitation of this analysis is that it focused primarily on travel time, while the price of renting bicycles is also an important, if not dominant, characteristic (Molin & Hoogendoorn-Lanser (2001)).

In literature on rental services for cars there is a strong focus on car-sharing, especially from an environmental point of view. The main idea is that car-sharing needs less vehicles to accommodate the travel demand, and that less polluting vehicles such as electric vehicles might be used. An overview of car-sharing and related mobility services can be found in Wagner (2000) and Bernard & Collins (2002). Many projects have been started but only a few have achieved some level of success, and their impact on mobility is limited. The Stadtauto- in Berlin, for instance, has 5,000 participants and is thus a large scale car-sharing project (Cousins (1998)). Compared to the total number of 3.5 million inhabitants, however, its share in total mobility is negligible.

### 7.4.2 Operational aspects

Car-sharing projects can be classified into three main categories (Barth & Sasheen (2002)). Figure 7-2 gives an illustration of these concepts. Single-station systems usually operate from small depots in the neighbourhood (A). These systems are usually not oriented to be used in combination with other transport modes. Systems that are designed to provide access and egress from transfer nodes are station cars, which usually have larger depots or rental facilities (B). Multiple station systems have several locations where vehicles can be rented or returned (C). Of course, combinations of these categories

exist as well. The difference between depots and rental facilities is that depots have few vehicles and no staff, while a rental facility have more vehicles and can provide personal service.

Home-based end                                                    Activity-based end



━━━━ Main mode
─────── Access or egress using a rented vehicle
- - - - - - Access or egress to a depot

● Rental facility
• Depot
◦ Activity location
○ Home
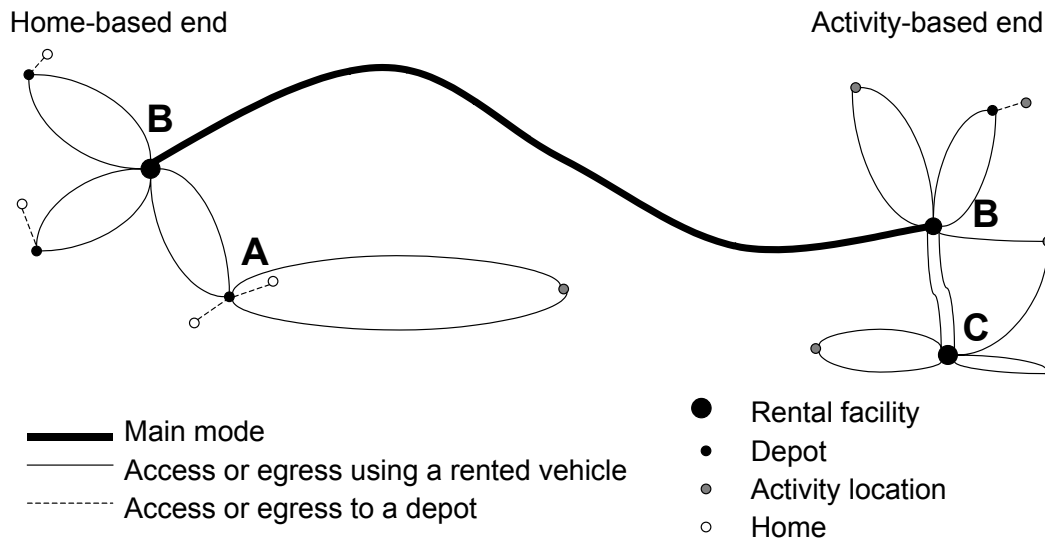
*Figure 7-2: Typical examples of car-sharing operations*

Given the focus on multimodal transport, there are three components of car sharing systems that have to be considered:

- The location of the depots or rental facilities;

- The number of vehicles needed;

- The price of renting vehicles.

With respect to the location of rental services in combination with multimodal travelling a distinction can be made between the home-based part and the activity-based part of the trip. At the home-based part rental services are interesting for travellers who have no vehicle available to access public transport. In that case, the residential density determines the location of the rental services. For the activity-based part of the trip, the transfer node is decisive for the location. The difference between trip parts also leads to another important difference. Given a density of households and assumption on the willingness to use such a service and the average trip rate, it is possible to determine optimal values for the number of vehicles that should be available (see e.g. Cousins (1998) and Zou (1999)). The general idea is that the facilities should be relatively small, that is 4 to 6 vehicles, in order to have a balance between number of vehicles and easy access to the vehicles themselves.

At the activity-end of the trip, demand for rental services might fluctuate substantially, especially at transfer nodes for the higher-level networks, leading either to large number of vehicles or to a large percentage of denied requests. This phenomenon introduces an interesting dilemma. Since the vehicle availability is more limited at the activity-end of the trip, it seems attractive to provide rental services for that part of the trip, while from an operational point of view, that is guaranteeing good service at low costs, it is more

interesting to provide rental services for the home-based part of the trip. Since bicycles cost less and require less space than cars, it is clear that the disadvantage of activity based rental services is smaller for bicycles than for cars. Furthermore, rental services for cars at transfer nodes have similar problems with parking at transfer nodes as discussed for private cars in Section 7.2.

An important aspect related to the price of renting a vehicle is the rental period itself. In the case of single location at the activity-based part of the trip, the vehicle is rented not only for the trip itself, but also for the time needed for the activity and the return trip. For home based rental services the time related to the main part of the trip is also included. The rental period is thus longer, often much longer, than the time the vehicle is needed for the trip itself, making the costs unnecessary high and increasing the number of vehicles needed. A possible solution to this problem is to change the concept of rental services from a single location service into a multiple station service (see e.g. Massot et al. (1999) and Barth & Todd (1999)). Suitable locations for additional rental facilities are office areas or large recreational facilities. Vehicles can then be used for trips starting at those locations as well. Extension to a multiple station rental facility, however, introduces the problem of balancing the number of vehicles and the demand for each location.

A typical form of rental services and multimodal transport is the station car demonstration project in San Francisco (Spiekerman & Weinstein (1998), Nerenberg et al (1999)). A multiple station rental facility was developed for a selected set of stops on the BART system and for several companies. The combination with offices allows the vehicles to be used in two directions, that is from home to the BART station and vice versa, and from the station to the office and vice versa. Furthermore they could be used for short tours from home or work. Special parking lots for parking and recharging the vehicles were located at the BART-stations. The experiment, in which 94 people participated using 40 vehicles, showed that such a system is technologically feasible, however, there was no analysis of the economic feasibility. The ratio between the number of participants and number of vehicles suggests that such services would be quite expensive.

### 7.4.3  Conclusions

From this discussion of the characteristics of rental services it can be concluded that they might have benefits for a multimodal transport system, especially for transfer nodes having no or low quality public transport as an egress mode. The price of rental services, however, will be relatively high, thus limiting the impact on multimodal transportation. Apart from the elements discussed here it is possible to increase the quality of the service component related to the rental service, for instance automatic payments, possibility of reservations, accessible locations at the transfer node, et cetera. Such improvements, however, will usually lead to higher prices.

Given the conclusion of the organisational relationship between network levels in Section 6.5.3 it is possible for the higher-level network operator to subsidise rental services in order to increase the total profit, but given the small numbers of users that are expected the size of this subsidy is limited. Rental services are thus a useful element in a

multimodal transport system, but its usage will be limited due to the high prices associated with such services. It should be noted, however, that there are obviously more possibilities for small and relatively cheap vehicles such as bicycles than for large and expensive types such as cars.


## 7.5   Demand responsive transport systems

Demand responsive transport is another type of service that might play an important role in a multimodal transport system. An important limitation of line-bound public transport services is the reduction in accessibility in time and space. Demand responsive transport service are provided on demand and eliminate resistances such as access, egress, and transfers, while waiting at the stop or station is replaced by waiting at home or at the activity location. The fact that demand responsive transport services are fully demand oriented make them very suitable for tailor-made multimodal trips.

The Traintaxi in the Netherlands, a share-a-cab service to and from railway stations, has already shown that such transport services have clear potential in multimodal transport, especially for the activity-based part of the trip (De Bruijn (1998)). In 1997 3.8 million times the Traintaxi was used to travel to and from the railway station. Marketing research showed that 16% of these trips were new train users, usually making long distance trips while travelling 1$^{st}$ class.

The cost-effectiveness of the Traintaxi, however, is still a point of concern. The Traintaxi is subsidised by the Dutch National Railway Company (NS), which is in line with the conclusion about the organisational relationship between network levels (Section 6.5.3). After 9 years of growing patronage, the number of users has decreased in the last two years to 3 million trips per year. Furthermore, the users were less satisfied with the quality offered, especially with waiting times, which were considered to be too high. Recently, it was decided to limit the service areas of the Traintaxi, and some locations have been closed because of the small number of users (Zwijgers (2001)).

The question to be answered is whether the dilemmas shown in the diagnosis of the Traintaxi systems are representative for demand responsive transport in a multimodal transport system. Therefore, this section focuses on the following two phenomena:

- The quality offered to the traveller in terms of price, travel times (waiting time and in-vehicle time) and reliability;

- The costs of demand responsive transport services, which are determined by the demand density and the size of the service area.

### 7.5.1   Quality for the traveller

In the context of multimodal transport, demand responsive transport is used for access and egress to and from transfer nodes.

For accessing transfer nodes travellers have to contact the operator before making a trip. Presently, most demand responsive transport systems in the Netherlands have a minimum period for making a reservation of 1 hour before the desired departure time. They might use the system individually, such as the traditional taxi service, or together with other travellers in a share-a-cab service. In both cases a waiting time might apply if there is a shortage of vehicles at the requested moment of departure. In a share-a-cab system, however, waiting times might be longer in order combine more trips. The traveller is brought to the transfer node, but in the case of a share-a-cab service a detour is possible in order to facilitate other travellers using the same vehicle at the same time.

For travellers arriving at a transfer node and using a demand responsive transport system to travel to their destination, the process is similar. The main distinction is that, depending on the systems, the requests are not made beforehand, making it more difficult to provide the service promptly. On the other hand, concentration of demand at a specific location enables the operator to have some vehicles waiting for the travellers. For a share-a-cab service, the possibility of immediate requests makes it more difficult to allocate the vehicles efficiently, which might lead to larger detours for the travellers.

Compared to the traditional taxi service, a share-a-cab service thus has several disadvantages, which are compensated by the lower price of such a service: longer waiting times, longer in-vehicle times, and thus a less reliable transport service. In fact, Wilson & Hendrickson (1980) claim that reliability is the Achilles' heel of demand responsive transport systems. In the case of trips starting at home, or at other places where no vehicles are waiting for travellers, the fact that travellers have to make reservations first also has to be taken into account. Experience with demand responsive transport systems shows that this is a serious barrier to their use.

### 7.5.2   Costs of demand responsive transport systems

The cost effectiveness of demand responsive transport systems seems to be limited. The study of Adebisi & Hurdle (1982) shows that in contrast to traditional public transport services, demand responsive transport systems have no economics of scale, that if the systems become larger the costs involved increase correspondingly. Demand responsive transport is thus only an alternative for line-bound public transport if the demand levels for public transport are too low for economically sound operations. Analysis of combinations of line bound public transport and flexible route strategies near the origin or destination by Chang & Schonfeld (1991) and Chang & Yu (1996) shows that the ratio between traveller costs and operator costs differs greatly between public transport and demand oriented transport services. In the latter case operational costs make the largest part of the total costs, which makes demand responsive transport interesting only for relatively low levels of demand.

### 7.5.2.1   Analytical model for a one-to-many service

The characteristics of the cost-effectiveness of a one-to many demand responsive transport system can be illustrated with a small analytical model. Given are a transfer node, a service area $B$ and a number of travellers $P$ arriving at the transfer node to travel to their destination using demand responsive transport. The assumption of travellers arriving at the transfer node can easily be replaced by travellers travelling to the transfer node. Combining trips in both directions will not affect the analysis either, unless the functions of distributing and collecting passengers are dealt with separately as in Daganzo et al. (1977).

The operator of the demand responsive transport system, tries to combine trips in space and time by dividing the service area into $n_b$ sectors and $n_t$ time periods, which is in fact a kind of frequency of the transport service (Figure 7-3). For each sector of the service area and in each time period a travelling salesman tour is made to distribute the travellers. The design problem that can be formulated is to determine the optimal values for the decision variables $n_b$ and $n_t$. Just as in earlier analyses the objective of minimising total costs is used. The number of travellers is thus assumed to be independent of the quality of the services offered.
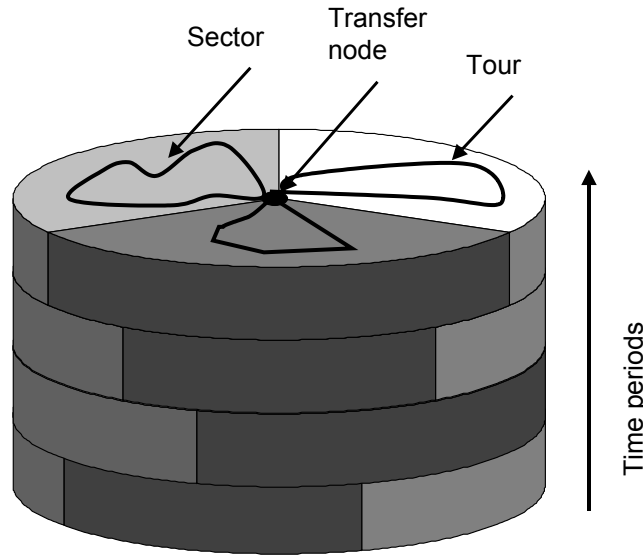


*Figure 7-3: Concept of sectors and time periods for one-to-many demand responsive transport systems*

Given the findings of Stein (1978) the length of the tour ($L_{tr}$) is determined by the size of the service area and the number of addresses to be served:

$$L_{tr} = b \cdot \sqrt{\frac{B}{n_b} \cdot \frac{P}{n_b \cdot n_t}} \qquad (7\text{-}4)$$

Assuming that the number of addresses per tour is relatively small and that the network used is a grid network, the value of the constant $b$ used in the analysis is 1.15 (see also

Chang & Schonfeld (1991) and Chang & Yu (1996)). For large numbers of addresses and Euclidean distances a value of 0.765 would be appropriate.

The arrival pattern of the travellers at the transfer node is assumed to be uniformly distributed. The number of time periods thus determines the waiting time (compare Equation (6-3)):

$$T_w = \frac{f_w}{n_t} \qquad (7\text{-}5)$$

If it is assumed that all travellers are travelling from the transfer node to their destination, the average in-vehicle time is given by:

$$T_i = \frac{0.5 \cdot L_{tr}}{V} \qquad (7\text{-}6)$$

The total travel time to the destination is thus the sum of the waiting time and the in-vehicle time, while accounting for the fact that travellers have a different value for the waiting time (compare Equation (6-1)):

$$T_t = w_w \cdot T_w + T_i \qquad (7\text{-}7)$$

The operational costs are determined by the number of vehicles needed, which depends on the riding time summed over all sectors and all time periods:

$$T_r = n_b \cdot n_t \cdot \frac{L_{tr}}{V} \qquad (7\text{-}8)$$

The minimum number of vehicles needed for one hour of operation follows directly from the total riding time. If the service area is divided into sectors, however, it is more realistic that each sector is served by at least one vehicle, since otherwise no vehicle would be available within a time period, leading to longer waiting times. The operational costs can thus be written as:

$$C_o = c_o \cdot MAX\left( n_b \cdot n_t \cdot \frac{L_{tr}}{V}, n_b \right) \qquad (7\text{-}9)$$

The most realistic assumption would be that the number of vehicles in each sector and each time period is an integer number. Such a formulation, however, leads to an objective function having many local optima. The formulation used is thus better suited to gain insight into the characteristics of a demand responsive transport system.

The objective function for the total costs can be written as (compare Equation 6-15):

$$C = P \cdot T_t \cdot c_t + C_o \qquad\qquad\qquad (7\text{-}10)$$

Since this formulation of the objective function includes a maximisation function, it is not possible to solve it analytically, but it might be solved numerically. Opting for a formulation based on either Equation (7-8) or on the number of sectors, however, leads to trivial solutions for the decision variables. In the optimal situation the number of sectors or the number of time periods would become very large, which is clearly not realistic.

### 7.5.2.2   *Application of the model*

In order to assess the impact of service area size and level of demand on the cost effectiveness, the model is applied for a set of hypothetical cases. The service areas considered have a radius of 3 kilometres, 5 kilometres and 7 kilometres, while for level of demand values of 10, 30, and 50 trips per hour are used. The lowest level of demand is representative for a peak period for an average traintaxi-system today.

The value of time for travellers is assumed to be equal to that for public transport travellers, that is € 4.90 per hour, while the operational costs for demand responsive transport are € 25.-- per vehicle per hour. For each combination of service area and demand density the optimal values for the number of sectors and number of time periods are determined numerically, and are then used to calculate the weighted travel time, and the operational costs. Figure 7-4 gives an impression of the shape of the objective function. In general the shape is comparable to those found for public transport, however, it should be noted that in this case the objective function is less shallow.

The results for these scenarios are shown in Table 7-3. From left to right it gives the impact of increasing demand, while from top to bottom the consequences of larger service areas are given. The diagonal from the upper left to the lower right gives an impression of the influence of a larger service area having a constant demand density. In order to compare the demand responsive transport system with line-bound public transport the total costs per traveller and per kilometre travelled are also included. In general, the total costs per kilometre are twice as high as those for urban public transport (see Table 6-7).

If the level of demand increases, the number of sectors and the number of time periods increase as well, leading to higher operational costs and lower travel times. The total costs per traveller are lower. However, since a higher demand level implies longer tours (see Equation (7-4)), trip length increases as well leading to relatively small reductions of the costs per kilometre travelled or even to increases. The impact of the changes in demand level also illustrates the impact of varying demand levels during a day. The scale-economies with respect to the level of demand are thus limited.
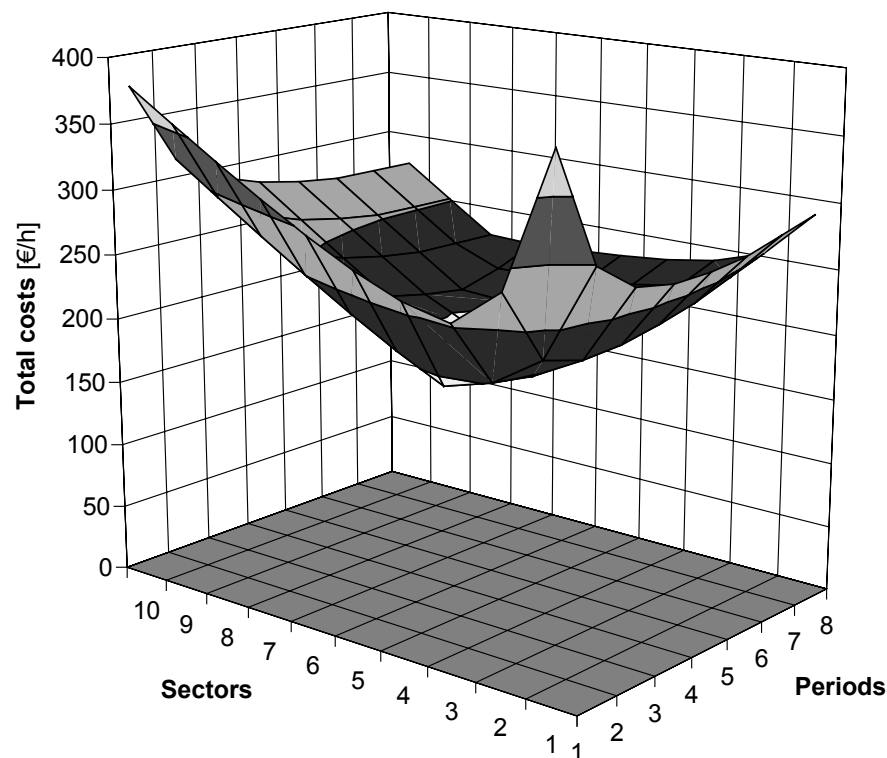
*Figure 7-4: Minimising total costs for a one-to-many demand responsive transport system (radius is 5 kilometres, demand is 30 travellers per hour)*

An increase in service area leads to an increase in nearly all characteristics. The only exception is the number of time periods. For larger service areas an increase in sectors has more impact than an increase in time periods. Interestingly, the consequences for the total costs per kilometre travelled are less clear. A reduction of the number of time periods leads to a larger demand per tour, to longer tours having more travellers and thus to lower costs per kilometre travelled. However, if the number of sectors also increases, the demand per tour will decrease, yielding shorter tours having less travellers and higher costs per kilometre. The net impact on the cost varies with the values of both decision variables. It should be noted that for larger service areas the number of time periods is rather low, leading to relatively large waiting time. This implies that in order to provide an acceptable service level, for instance having maximum waiting times of 15 minutes or providing synchronised services, suboptimal solutions having a minimal number of time periods should be implemented.

The third relationship shown in Table 7-3 is that of increasing the service area having a constant demand density, that is the diagonal of this table. In this case larger service areas lead to more sectors, longer travel times, and higher operational costs. The total costs per traveller are more or less constant, while the total costs per kilometre travelled are lower.

Finally, it should be noted that the average number of travellers per sector per time period is relatively low, about two travellers. This implies that taxis are a reasonable type of

vehicle for such a transport service. In the case of concentrations of specific destinations in time and space, of course, larger vehicles may be necessary.

*Table 7-3: Optimal characteristics of a one-to-many demand responsive transport system for different demand levels and various service area sizes*

| Demand level [/h] | 10 | 30 | 50 |
|---|---|---|---|
| *Radius = 3 kilometres* | | | |
| **Number of time periods** | 4 | 4 | 6 |
| **Number of sectors** | 2 | 3 | 5 |
| **Weighted travel time** [min] | 18.5 | 19.6 | 12.8 |
| **Operational costs** [€/h] | 50.0 | 83.7 | 132.4 |
| **Total costs per traveller** [€] | 6.51 | 4.39 | 3.69 |
| **Total costs per kilometre travelled** [€/km] | 2.69 | 1.57 | 2.09 |
| *Radius = 5 kilometres* | | | |
| **Number of time periods** | 2 | 3 | 4 |
| **Number of sectors** | 2 | 5 | 7 |
| **Weighted travel time** [min] | 39.6 | 24.7 | 19.0 |
| **Operational costs** [€/h] | 57.0 | 125.0 | 180.2 |
| **Total costs per traveller** [€] | 8.93 | 6.18 | 5.15 |
| **Total costs per kilometre travelled** [€/km] | 1.57 | 1.92 | 2.00 |
| *Radius = 7 kilometres* | | | |
| **Number of time periods** | 2 | 3 | 3 |
| **Number of sectors** | 3 | 6 | 8 |
| **Weighted travel time** [min] | 38.5 | 26.3 | 25.9 |
| **Operational costs** [€/h] | 79.8 | 169.2 | 218.4 |
| **Total costs per traveller** [€] | 11.12 | 7.79 | 6.49 |
| **Total costs per kilometre travelled** [€/km] | 2.09 | 2.07 | 1.78 |

### 7.5.3   Conclusions

The main conclusion that can be drawn from this analysis is that a one-to-many demand responsive transport system is expensive and will remain expensive. Increasing demand levels will reduce the costs per trip, but the total costs of the system will increase. The scale-economies that have been found are too small to become an alternative for line-

bound public transport, except in cases having a low level of demand (Heinzel (1996), Mageean & Nelson (2001)).

The possibility of extending the one-to-many system to a many-to-many system will not improve the cost efficiency either. On the contrary, a many-to-many pattern lacks the natural tendency for bundling of a one-to-many system, which makes it more difficult to combine trips efficiently. Only in the case of very low demand levels is it beneficial to expand the one-to-many system in order to reduce the time that vehicles and their drivers are idle.

From the travellers' point of view the assessment of demand responsive transport is less clear. Demand responsive transport systems might offer interesting travel times, but the price is clearly higher than that for line-bound public transport. Furthermore, there is less certainty on the arrival times, so transfers or appointments might be missed, or additional waiting time is introduced in order to minimise the risk of arriving late. Finally, the fact that for specific trip types a reservation is required, makes demand responsive transport systems less attractive. To date the experience with demand responsive transport systems suggests that the disadvantages outweigh the possible benefits.

## 7.6   Conclusions

This chapter focused on the consequences of multimodal transport on transport network design. A synthesis was made of the findings on hierarchical network structures for private transport and line-bound public transport. Furthermore, an analysis was made of the potential of two types of transport services that might improve the prospects for multimodal transport.

The most important conclusion from this discussion is that multimodal transport does not ask for new network structures. There is thus no need for a new approach to transport network design, which is the opposite of what was expected. The analysis in this thesis, however, emphasised the importance of hierarchy in transport networks and supports the design principles used in the German guidelines. Properly designed unimodal transport networks are robust with respect to their hierarchical network structures. The hierarchical structure of private transport networks is based on fundamental principles, while the impact of multimodal transport on network usage is expected to be small. For public transport networks the relationship with spatial hierarchy is essential for the hierarchical network structure. Changing this network structure only in favour of multimodal transportation eliminates the foundation of public transport networks. Furthermore, it was established that alternative access modes have no significant influence on urban public transport network design. Walking should be considered as the main access mode for lower-level urban public transport networks.

The analysis of demand oriented transport services such as rental services or demand responsive transport systems, showed that such services have small benefits for multimodal transport systems. The costs involved with such transport services, and thus

the prices for using the services, are generally too high to attract large numbers of travellers.

If multimodal transport does not require adapted network structures, the problem of multimodal transport network design is reduced to the allocation of transfer nodes. Again, since the hierarchy in spatial structures primarily determines the location of access nodes of line-bound public transport services, there are limited possibilities for specific transfer nodes between private transport and line-bound public transport:

1. Transfer nodes in city centres, which offer access to higher-level public transport networks, combined with access modes walking, cycling, and local public transport;

2. Transfer nodes in local centres offering access to medium distance public transport, accessible by bicycle and private car;

3. Transfer nodes at the edge of cities offering access to the city centre using high quality urban public transport, combined with access by private car, and preferably directly connected to higher-level road networks.

Given these location of transfer nodes, the design of these nodes, and especially the design of the parking facilities, and the quality of the transport services offered are then decisive for the prospects of multimodal transportation.

It should be noted that transfer nodes that are accessible by private car, that is types 2 and 3, are not only limited to access nodes of line-bound public transport. Similar types of transfer nodes might be offered at access nodes of higher-level private transport networks, allowing for combining the transport modes car driver and car passenger (carpool). Transfer nodes at the edge of cities might even be suited for rental services. In that case the transfer nodes operate in two directions:

• To the city centre: main mode private car and egress mode urban public transport;

• To destinations outside the city: access modes urban public transport and cycling and main mode rental car.

Once the typology of the transfer nodes has been chosen, the exact location and the associated characteristics can be determined using a multimodal network design model using a detailed network description based on zones, nodes, links, and lines. Examples of such models can be found in Carrese et al. (1996), Ferrari (1999) and García & Marín (2001). In fact such models are comparable with traditional network based network design models as described in Section 3.3.4. For a detailed modelling of the time accessibility of public transport services, especially for low frequency services, the concept of diachronic graphs provides interesting possibilities (Nuzzolo & Russo (1996), Nuzzolo et al. (2001)).

A typical characteristic of multimodal transport network design models having this level of detail is the way the combination of modes is modelled. Traditional transport models make an explicit distinction between modes and are therefore not able to model multimodal transport properly. If there are only a few relevant combinations of modes,

they might be defined as separate modes and then be included in a traditional approach. Multimodal transport, however, consists of a diverse set of various combinations of modes, which limits the possibilities of such an approach. An interesting alternative is to model all networks involved and the transfer links between the networks in a supernetwork (Sheffi (1985)), and to use new models that combine route-choice and mode-choice as route-choice in a supernetwork (Benjamins et al. (2002)). An extensive overview of such supernetwork, or hypernetwork, approaches can be found in Catalano et al. (2001).

Another important element in a multimodal transport network design model is the description of the costs involved. The numerical analyses on hierarchy in private transport networks and in line-bound public transport showed the importance of considering costs, and especially investment costs, in the analysis. A more detailed view of the costs involved and of the economical benefits of multimodal transport facilities will seriously influence the detailed design of transport services and transfer facilities.

# Chapter 8

# CONCLUSIONS

The main research theme in this thesis is whether the notion of multimodal transport requires a new approach to the design of transport networks. This final chapter summarises the main results of the thesis. First, a short summary is given of the problem studied in this thesis and the approach used to analyse the multimodal transport network design problem. This is followed by an extensive discussion of the findings in this thesis and the conclusions regarding multimodal transport network design. Next, the resulting guidelines for multimodal transport network design are presented, and finally recommendations for further research are given.

## 8.1  Short summary

Multimodal transport might be an interesting approach to solve today's mobility problems such as recurrent congestion, deteriorating accessibility, and negative environmental impact. Combining private and public transport in a multimodal transport system offers opportunities to capitalise on the strengths of the various systems while avoiding their weaknesses, and might therefore be an interesting alternative to the traditional strictly dichotomous choice between private car or public transport.

Multimodal transport involves many issues, such as information systems, financial systems, operational control of transport services, and organisation of services. A crucial prerequisite for multimodal transport, however, is the existence of a multimodal transport network. The question then is what the main characteristics of a multimodal transport network are. How are efficient multimodal networks structured, and what does that imply for transport network planning? This thesis provides answers to these questions. The main research question thus is: what is the influence of multimodal transport on transport network design?

In this context multimodal travel means using two or more modes for making a trip, that is different vehicle modes or functionally different service modes. Typical examples of

multimodal trips are: using the private car to travel to a city while transferring to an urban public transport service to access the city centre, or making a long distance trip by train while using private modes such as car or bicycle, or public transport services such as bus to access the train station. Multimodal travel thus implies an intermodal transfer between different modes or services. Trips in which only a single vehicle mode is used, or which have only a transfer within a specific public transport service network, such as between urban buses, are considered to be unimodal trips. The main distinction between the two types of transfers is that intermodal transfers, and thus multimodal transport, deal with different networks, different modes or service types, which are designed and operated by different authorities, investors or operators. Following the elaboration of the notion of multimodal transport, an extensive quantitative analysis has been made of the characteristics of multimodal mobility today and its future potential.

The specification of the multimodal transport network design start with a discussion on the network design problem and on the extension to the multimodal transport design problem. The network design problem is seen as a Stackelberg game in which an investor, or operator, takes decisions on the network structure, while having full knowledge of its impacts on the actions of the other party involved: the traveller. With respect to the approach adopted in the analysis of the consequences of multimodal travel on transport network design two important choices were made.

First, it was observed that multimodal transport nearly always implies that a particular mode or transport service is used for covering the main distance, while other modes are used to access and/or egress from the main mode used. Therefore, the concept of access and egress modes versus main modes suggests a hierarchical view of transport systems, and thus of transport networks. Hierarchy in transport networks implies that different network levels are distinguished, each suited for covering specific distances and for offering access to higher network levels. The highest network level is usually a coarse network having high speeds and limited spatial accessibility. The lowest network level serves short distance trips and provides access to higher-level networks. This network level thus has high network densities, slow speeds, and high space accessibility. Hierarchical network structures can easily be distinguished in transport networks today, but they seem to be considered a fact of life, instead of an important decision variable in transport network design. This thesis will therefore explore the issue of hierarchical transport networks first.

Second, two methods can be distinguished to describe transport networks, and thus to specify the transport network design problem. The first method focuses on abstract network types while using aggregate network characteristics such as average stop spacing, line spacing, road spacing, frequency and speed. This method has the advantage that analytical models can be derived specifying the main relationships for these network characteristics. On the other hand, however, such analyses are clearly a simplification of reality. Real networks show a nearly unlimited variety in their characteristics compared to the abstract network types, while many assumptions used to formulate analytical network design models are based on strongly simplifying assumptions, especially with respect to travel demand patterns. In contrast, the second method to specify transport networks accounts for all the details that can be found in transport networks today, such as local constraints and capacity constraints, as well as the varying demand pattern over

space and time. This method is therefore especially suited to analyse specific problems in transportation planning. However, since these specific models allow all kinds of details to be included in the analysis, they do not easily provide general knowledge on the main drivers of transport network structure. Analytical models of aggregate network characteristics are considered more suited for that purpose. This thesis develops and uses such analytical network design models to analyse the impact of multimodality on transport network design. The limitations of this abstract network approach, however, are acknowledged in the interpretation of the findings.

Given these choices of a hierarchical network concept and the use of analytical network design models describing aggregate network characteristics, an extensive analysis is made of hierarchy in transport networks in general, in private transport networks, and especially in line-bound public transport networks. The analytical network design models developed in this thesis consistently use the same framework for both types of transport networks, that is the economically based design objectives of minimising total costs or maximising social welfare, while random utility theory is used to describe traveller behaviour. The design models consist of level of service models, demand models and supply models. Innovative elements in these models are the multilevel nature of the network design problem, the incorporation of lower level travel choices such as route-choice or access mode-choice, and the possibility that different actors, such as authorities, investors or operators are responsible for transport network design. Furthermore, the importance of the models developed in this thesis is that in comparison with earlier studies on transport network design, they focus especially on the relationships for hierarchical network structures and that their mutual consistency allows a systematic comparison between different hierarchical network structures.

The concept of hierarchy in private transport networks is approached from three different directions. A new perspective on hierarchy in private transport networks is developed, based on the elimination of shortcuts in hierarchical transport networks. A model for super-positioning a higher-level network is developed to show the main mechanisms involved from an economic point of view. Furthermore, the issue of hierarchy is analysed from an evolutionary perspective on transport networks. All three types of analysis show a surprisingly consistent outcome on hierarchy in private transport networks.

Concerning line-bound public transport networks a new formulation is established for the network design problem of single-level urban networks, such as bus or tram networks. This analytical model is then extended to the multilevel network design problem. New models are developed to describe all kinds of possible hierarchical network structures for large urban areas, such as express services, trunk-feeder systems, and zone-systems. The impacts of trip length distribution and demand patterns are explicitly considered in the analysis. Just as for the case of private transport networks, design models are developed for the analysis of super-positioning a higher-level network for urban public transport networks as well as interurban public transport networks. Furthermore, the multilevel network design problem is extended further to include the fact that different network levels generally involve different actors. A game theoretical approach is used to account for the fact that these actors pursue their own objectives. This analysis approach provides interesting insights into the interaction between network levels.

A synthesis follows in which the findings from the analyses of hierarchy in private transport and public transport networks are used to redefine the possible impacts of multimodality on transport network design. This synthesis includes three additional analyses. First, the single-level urban public transport network model is extended to include alternative access modes to walking only. Second, an assessment is made of the potential impact of rental services as access and egress modes on multimodal transportation. Third, an analysis is made of the potential of demand responsive transport systems as access and egress modes on multimodal trip making. This analysis includes the development of an analytical design model describing a demand responsive transport system offering access to, and egress from, a multimodal transfer point.

The findings derived from these analyses are summarised in the next section, followed by a discussion on the conclusions that can be drawn for multimodal transport networks.


## 8.2   Findings on multimodality and its impacts on network design


### 8.2.1   Multimodal mobility: small but important

Multimodal mobility today represents 3% of all trips in the Netherlands. However, this does not imply that multimodal mobility is unimportant. Indeed multimodal travel does have substantial shares in certain market segments:

- 15 % of Dutch trips longer than 30 kilometres are multimodal;

- 20% of trips to and from the four main cities in the Netherlands are multimodal.

The main mode for nearly 60% of multimodal trips is train, or looked at in another way, 80% of all train trips are multimodal trips. Private modes are the main mode for 17% of multimodal trips. Multimodal mobility is thus an important phenomenon for long distance travel and for access to city centres.

For the assessment of the future potential of multimodal transport two demand estimations were made: one based on trip purpose and the other on trip type, the results of which however were similar. The medium scenarios yield an increase of the multimodal travel share of about 25%, leading to a multimodal share of 3.6% of all trips. The maximum scenarios result in an increase of 70%, yielding a modal share of 5.0% of all trips. This assessment shows that multimodal mobility will remain only a small fraction of all trips even though an increase of 25%, and certainly an increase of 70%, will require substantial increases in the capacity of public transport services. If it is assumed that the increase of 70% in multimodal travel is fully accommodated by train services, trips between 10 and 30 kilometres would triple, while trips between 30 and 100 kilometres would double. These values are obviously too optimistic, but they show that an increase in multimodal transport can have substantial impacts on the required public transport service capacities. Multimodal transport thus has a strong relationship with line-bound public transport services.

### 8.2.2    Hierarchy in transport networks is a natural phenomenon

Hierarchy in transport networks can be seen as a natural phenomenon. In fact, a single-level network is highly unstable, except in hypothetical cases. If there is just a small bias favouring some routes over others, or if there is some kind of interaction between supply and demand, the development of a second network level having other characteristics is inevitable. The resulting multilevel network consists of a higher level network that is suited for covering long distances, while the lower network is suited for short distances and to access the higher level network.

In reality, mechanisms leading to such a hierarchy are unavoidable. From the perspective of the demand for transport, not all possible routes will be equally attractive for all travellers. Travellers will thus favour some routes over others, leading to differences in usage. On the supply side, it is logical that from an economic point of view heavily used routes are more likely to become more attractive due to all kinds of investments for additional facilities or improvements of the routes themselves. On both demand side and supply side improving efficiency appears to be a leading principle. Furthermore, not all origins and destinations will have equal demand. Agglomeration tendencies lead to specific concentrations of demand in space, thus making routes to and from these locations more important. Technological developments make it possible to increase travel speed significantly making higher network levels more attractive, while modern decision processes stimulate the concentration of flows on a limited set of routes. The continuous interaction between demand and supply, the correlation with agglomeration processes, and the influence of technological developments and decision processes, all these mechanisms are constantly working together leading to hierarchical transport network structures: hierarchy in transport networks is a fact of life.

### 8.2.3    Hierarchical private transport networks exhibit self-organising properties

For private transport networks a fundamental mechanism, which can be qualified as a self-organising property, has been established in this thesis that dictates the relationship between network levels. For grid networks the road spacing of a higher-level network should be about three times the road spacing of the lower-level network, while the average speed should be about two-thirds higher. This fundamental relationship has been established using a number of different approaches based on travel behavioural principles, such as travel cost reduction and minimisation of shortcuts, and on economic principles such as minimising total costs. Interestingly, the scale-factor 3 for road spacing following from these analyses, matches the scale-factor for road networks that was found by a morphological analysis (De Jong (1998a)).

### 8.2.4    Spatial structures determine hierarchy in line-bound public transport service networks

For line-bound public transport networks no fundamental hierarchy rule based only on network characteristics could be established. Typical phenomena such as transfers and the high costs of infrastructure hamper the natural tendency for hierarchy in transport

service networks. Concentration of travel demand in space, however, appears to provide a perfect counterbalance for this mechanism. The hierarchy in spatial structures thus determines the hierarchy in line-bound public transport networks, leading to a scale-factor of about 3 for both line spacing and stop spacing.

The characteristics found for today's access and egress legs of long distance public transport services, such as train services, support the importance of spatial hierarchy in public transport network structures. For these access and egress legs three types of travel modes can be distinguished with corresponding shares found in the Netherlands:

- Walking, accounting for nearly 50% of all access and egress legs;

- Public transport services, making up more than 25%;

- Private transport, such as private car and especially bicycles, also representing more than 25% of all legs.

For train trips to and from the four main cities in the Netherlands these percentages change to 40%, 40% and 20% respectively. All these percentages, of course, concern the split as it exists today. However, they clearly illustrate the importance of the location of the access nodes, accounting for nearly half the patronage, as well as the importance of multimodal access for interurban public transport services, for which both private transport and public transport services play an important role.

## 8.2.5 Urban public transport service networks are too dense

For urban public transport networks it has been shown using a unimodal analysis, that current public transport networks are too dense. Given the size and characteristics of the cities today and the operational costs for urban public transport, an 80% larger stop spacing, and for bus networks also a 35% larger line spacing, is preferable from a social welfare point of view. Coarser networks yield shorter travel times, except of course for very short trips, and lower operational costs. Furthermore, it has been shown that many objections raised against such a coarsening in urban public transport network structure can easily be countered using basic principles of travel behaviour and economic analysis.

## 8.2.6 Alternative access modes do not influence public transport network design

On the basis of an in-depth multimodal access analysis it has been shown that walking remains the main access mode for urban public transport networks. This pertains to single-level urban networks as well as to the lower-level network of multilevel public transport systems. The introduction of alternative access modes such as cycling is shown to have only minor consequences for the values of the optimal network characteristics. A dedicated behavioural analysis shows this is partly due to a rather high penalty for using a bicycle to access urban public transport, and partly due to the fact that access distances will always be so short that walking will remain the preferred access mode. Note that alternative access modes, such as cycling, remain essential for higher-level public transport networks although this does not influence their optimal network design.

### 8.2.7    Financial relationships between network levels exist

A new type of analysis developed in this thesis incorporates the case that different public transport network levels may have different authorities or operators, who may have their own objectives. Using a game theoretical approach different organisational scenarios for multilevel network design have been analysed. With respect to the network design variables, the impact the design objective is more decisive than the case of different authorities or operators. Networks designed using the objective of maximising social welfare are denser and have higher service quality, while the objective of profit maximisation leads to coarse networks having long travel times. The main impact of the case of different authorities or operators is that frequencies become more important due to explicitly considering travellers transferring between both network levels.

Furthermore, the outcomes of these scenarios show that interesting financial relationships exist between network levels. Since long distance public transport is more likely to be profitable, it is in the interest of the higher-level network operators to ensure good accessibility by lower-level public transport services. This can be achieved by subsidising lower-level public transport networks, or other local transport networks, to improve the quality for travellers using those networks to access the higher-level network. An even more profitable situation occurs if other actors, such as local authorities, strive to improve the quality of their lower-level networks for their own objectives, such as in the case of an urban public transport network that is designed using the objective of maximising social welfare. In that case the interurban operator obtains high quality access to his transport services, without any financial input from his side.

### 8.2.8    Demand oriented transport services are useful but not decisive

Demand oriented transport services such as rental services and demand responsive transport services can make multimodal transport more attractive. The costs of providing such services, however, are so high that their impact on multimodal mobility will always be small, although they might be perfectly suited for specific population segments. In the case of rental services, the costs of the vehicles themselves as well as those for parking the vehicles are relatively high, thus limiting the benefits of scale economies. It is obvious that this finding is more critical for car rental services than for bicycle rental services. For demand responsive transport the share of personnel costs is too high to achieve benefits of scale that are in line with those found in traditional public transport services. This assessment might change, however, if the operational costs could be substantially reduced, for instance by using fully automated vehicles in a demand responsive transport system, thus eliminating the costs of a paid driver.

## 8.3    Conclusions for multimodal transport networks

Given the outcomes of a variety of analyses, the main conclusion of this thesis is that multimodal transport does **not** require significant restructuring of transport networks. The explanation for this conclusion is that properly structured transport networks, that is

private transport and multilevel line-bound public transport networks, on themselves are already well suited for serving multimodal travel demand. The emphasis in this conclusion is on the word *properly*. Clear rules have been established that determine hierarchy in both private transport and line-bound public transport networks, and thus also in multimodal transport networks. Ignoring these basic rules leads to poorer performance of all networks involved. Furthermore, the level of demand for multimodal transport is now, and is likely to remain, too small to justify changes in unimodal transport network structures.

The conclusion that an efficient multimodal transport system requires properly designed transport networks, reduces the multimodal network design problem to the allocation of transfer nodes. Given the strong relationship between hierarchy in spatial structures and in public transport networks, this leads to clear criteria for the location of intermodal transfer nodes:

- Within city centres, offering access to higher-level public transport networks (comparable to railway stations in cities);

- Within local centres, offering access to cities by public transport services (comparable to traditional Park&Ride facilities);

- At the edge of cities near motorways, offering access to city centres by urban public transport services, or even offering access to the motorway network using urban public transport or bicycles (comparable to Transferia).

The attractiveness of multimodal travel thus depends more on the quality of the transport services offered than on newly designed transport networks. The quality of the transfer nodes, the transport services themselves, the availability of information and all kinds of financial aspects are decisive. Stimulating multimodal mobility does not require a new grand design for the transport system, but benefits more from doing little things properly.

Multimodal transport is essential for the accessibility of city centres and for the profitability of higher-level public transport services. Actors who are responsible for these two issues, should take the lead in the development of multimodal transport services and facilities.

## 8.4   Implications for transport network planning

### 8.4.1   Planning guidelines

The findings and conclusions following from the analyses in this thesis lead to the following guidelines for transport network design:

- Hierarchy in spatial settlements determines the hierarchical levels in both private transport and public transport networks, leading to five possible network levels oriented to villages, towns, cities, agglomerations and metropolises respectively (Table 8-1). Each network level offers transport

facilities for trips between settlements of the same rank, and offers access to higher-level networks;

*Table 8-1: Hierarchy in interurban network levels*

| Network level | Related spatial level | Road and line spacing [km] | Speed [km/h] | |
|---|---|---|---|---|
| | | | Private transport | Public transport |
| **International** | Metropolis | 300 | - | 235 |
| **National** | Agglomeration | 100 | - | 140 |
| **Interregional** | City | 30 | 100 – 120 | 85 |
| **Regional** | Town | 10 | 60 – 70 | 50 |
| **Local** | Village | 3 | 35 – 40 | 30 |

- Private transport networks should ideally have a scale-factor of about 3 for road spacing and access spacing and a scale-factor of about 1.67 for speed;

- Line spacing and stop spacing in public transport networks have ideally a scale-factor of about 3, while a speed scale-factor of 1.5 to 1.67 is appropriate;

- Urban public transport networks should have stop spacing of 600 to 800 metres and line spacing of 750 to 800 metres;

- Providing rental services or offering demand responsive transport services at access nodes of higher-level public transport networks is sensible from the perspective of increasing patronage and providing transport services for specific niche markets.

Furthermore, three types of intermodal transfer nodes should be distinguished:

- Transfer nodes to higher-level networks, that is at the interregional level or higher, which should be located in city centres. Special attention is needed for accessibility by transport modes that are suitable for an urban environment, such as walking, cycling and urban public transport;

- Transfer nodes at the regional network level, which should be centrally located in towns and have good accessibility and good facilities especially for private modes such as car and bicycle. Similar types of transfer nodes are also possible at on and off-ramps of the motorway network, facilitating the opportunity to share a ride by private car;

- Transfer nodes at the city edges, offering access to city centres by urban public transport for travellers arriving by private car. In combination with rental services such transfer nodes might also be interesting for accessing the motorway network, using urban transport modes, such as public transport and bicycles as access modes.

Please note that these guidelines are not always different from existing guidelines. The type of support for these guidelines, however, is quite different. The guidelines developed here follow from extensive analyses of the characteristics of transport networks, while some of the existing guidelines are based on current practice, and other on the a priori distinction in spatial structures. The analyses in this thesis showed that guidelines based on current practice might be outdated to changes in urban structures and related costs. Furthermore, they showed that assuming a distinction in spatial structures is a proper approach. As such, this thesis provides additional theoretical support for the German guidelines for private transport and public transport network design (FGSV (1988), Köhler (1989), Schönharting (1997), Bierschenk & Keppeler (2000). This theoretical support, however, also identifies that, instead of allowing small differences in speed between network levels, a clear distinction between network levels is required to exploit the benefits of a hierarchical network structure. Finally, it should be noted that these new guidelines re-emphasise the importance of the relationship between hierarchy in spatial structures and transport networks, especially for public transport service networks. Higher-level public transport networks require high densities around the stops, which are by definition associated with the hierarchy in spatial structures.

### 8.4.2   Use of these guidelines

These guidelines are based on various conclusions with respect to network hierarchies, such as:

- The relationship with hierarchy in spatial structures;
- The scale-factors for road spacing, stop spacing, line spacing and speed;
- The optimal values for urban public transport networks;
- The typology of transfer nodes.

Given the background of these conclusions, however, these guidelines should not be interpreted as rigid standards. Actual situations will always lead to deviations from the ideal constructs used in the analyses. For most of these guidelines such deviations are no problem at all. The analytical models that have been used to determine these optimal values have relatively flat objective functions implying that small deviations have no great consequences for the overall performance of the design.

In the case of large deviations, however, the performance of the networks involved will deteriorate. Examples of such deviations are:

- Skipping a network level in private transport networks;
- Creating small differences in quality between network levels;
- Applying a very short stop spacing;
- Locate transfer nodes badly.

What might seem clever at first sight might be expensive in the long run. This is especially important to consider, as many decisions require new transport infrastructure,

such as roads, rail or parking facilities, which once they have been built are impossible or impractical to move to a better location.

## 8.5   Recommendations for further research

This thesis focused on the consequences of multimodal trip making for transport network design. The analyses using hierarchical network concepts and analytical models clearly showed that the multimodal perspective does not require a new approach to network design. On the contrary, properly designed unimodal transport networks are essential for the attractiveness of multimodal trip making.

Apart from this conclusion, and the findings and guidelines presented in this chapter, the analyses in this thesis show three major opportunities for further research. First, the possibilities for exploring the analytical approach further are discussed, including the possibility to make the design guidelines operational. Second, typical consequences following from this thesis for a more detailed approach of multimodal transport network design are considered. Finally, there is a need for more detailed knowledge on the costs involved and on multimodal travel behaviour.

### 8.5.1   Extending the analysis

From a theoretical point of view, it is interesting to carry the ideas in this thesis further. The thesis analysed basic configurations such as two-level transport networks and considering only two transport modes. The analytical approach of the multimodal multilevel network design problem, however, can be extended to many levels and many modes, while accounting for all the traveller categories involved. Such an analysis might provide more detailed insight into the relationships between network levels, transport modes, and traveller categories. Another extension might be to include additional transport system characteristics such as fares and capacity in the analysis.

Another promising opportunity is the extension of the multimodal, multilevel, multi-operator network design problem with the distinction between transport infrastructure provider and transport service operator. Since transport infrastructure is often seen as the domain of the authorities, while transport service operators might be classified as private companies, the differences in objectives and in time-horizons might show interesting relationships that can have implications for planning practice.

A further intriguing question is whether the concepts used in this thesis are also applicable to freight transportation. Are there similar generic rules for network hierarchy, or is the diversity in freight transportation so large that no such rules can be defined?

Finally, it is worthwhile to develop an interactive design tool for multimodal transport networks, that explicitly uses these guidelines to judge existing networks and new proposals, as well as suggesting improvements. Such a tool would be an easy way to

provide regional and national planners with insights into critical characteristics in their transport networks.

### 8.5.2  Multimodal transport network design models

The design guidelines are especially powerful in the first stages of the planning process: which options are realistic and if so how should the network structure be changed. However, if it comes to deciding where to locate specific facilities within the boundaries provided by these guidelines, more detailed tools are required. In such a case, local constraints, such as network topology, the allocation of urban functions, and more detailed network characteristics such as capacity become crucial. Detailed network models and equally detailed demand models are essential to assess the costs and benefits involved, and to determine the consequences of deviating from the guidelines. It should be noted that it is possible that due to all kinds of local constraints, only clearly sub-optimal solutions might be feasible! For such detailed network models special attention is needed for modelling multimodal transport trip making properly, and for accounting for all costs involved.

### 8.5.3  Additional knowledge on costs and travel behaviour

Design of transport networks requires adequate knowledge of all costs involved and of travel behaviour. In this thesis many assumptions have been made on the costs of infrastructure and operation as well as on travel behaviour, using existing data where possible.

This study focused on transport network design, using optimisation techniques to find a balance between traveller's interests and investors or operators interests. Multimodal transport, however, is shown to be important for the accessibility of economically important centres. The question is whether the benefits resulting from this improved accessibility should also be incorporated into the analysis, or whether the general economic concept of consumer surplus provides sufficient insight into the size and nature of those benefits.

For the costs involved in transport services and transport networks a broad range of values has been found, especially for the investment costs, while in public transport operational costs are often regarded as classified information. Since the models used in this thesis all have relatively flat objective functions, deviations from the values assumed would not seriously influence the conclusions presented here. For more detailed network analyses, however, more detailed knowledge of the costs involved is necessary. Good assessments of the costs of new infrastructure or operating public transport services might be critical for the decision whether or not to develop transport facilities.

The phenomenon of a broad range of possible values also applies to the parameters used to describe travel behaviour. Furthermore, it was found that there is virtually no knowledge of the willingness to transfer in a multimodal trip. Just as was noted when discussing the need for a better assessment of the costs involved, the relatively flatness of

the objective functions makes the main conclusions quite robust with respect to changes in the values that were assumed, while for actual decisions a more detailed description is essential. This is especially true if relatively rare or new combinations of modes are considered. Examples are the combination of bicycles and lower-level public transport networks, and the role of demand responsive transport systems in a multimodal transport system. After all, the more detailed the decision that is considered, the more detailed the model that is needed, and the greater the importance of more detailed data!

# References

&SAMHOUD (2001), Ketenmobiliteit in beeld, Wereldse successen, AVV, Rotterdam

ADEBISI O., V.F. HURDLE (1982), Comparing fixed-route and flexible-route strategies for intraurban bus transit, *Transportation Research Record 854, pp. 37-43*

AGV (1989), STRategische Netwerkstudie Openbaar Vervoer (STRENOV), AGV, Nieuwegein

ARTHUR W.B. (1988), Urban systems and historical path dependence, *In: Ausubel J.H., R. Herman, Cities and their vital systems: Infrastructure Past, Present & Future* National Academy Press, Washington

ARTHUR W.B., Y.M. ERMOLIEV, Y.M., KANIOVSKI (1987), Path dependent processes and the emergence of macro-structure, *European Journal of Operational Research, Vol. 30, pp. 294-303*

AVV (1990), Elasticiteiten Landelijk Modelsysteem, AVV, Rotterdam

AXHAUSEN K.W., T. HAUPT, B. FELL, U. HEIDL (2001), Searching for the rail bonus: Results from a panel SP/RP study, *Paper presented at the 80th Annual Meeting of the Transportation Research Board,* Transportation Research Board, Washington

AYKIN T. (1995), Networking policies for hub-and-spoke systems with application to the air transportation system, *Transportation Science, Vol. 29, No. 3, pp. 201-221*

BAAJ M.H., H.S. MAHMASSANI (1991), An AI-Based approach for transit route system planning and design, *Journal of Advanced Transportation, Vol. 25, No. 2, pp. 197-210*

BACH B. (1999), Ontsluitingsstructuren, *In: V.J.D. de Groot, E. Kanters, J.W.M. Korsmit, Handboek Verkeers- en Vervoerkunde,* VUGA, Den Haag

BARTH M., M. TODD (1999), Simulation model performance analysis of a multiple station shared vehicle system, *Transportation Research C, No. 7, pp. 237-259*

BARTH M., S. SASHEEN (2002), Classification of shared vehicle systems, *Paper 02-3854, CD-ROM of the 81st Annual meeting of the Transportation Research Board,* Transportation Research Board, Washington DC

BATTY M. (1998), Urban evolution on the desktop: simulation with the use of extended cellular automata, *Environment and Planning A, Vol. 30, pp. 1943-1967*

BECKER S., K. BRENNER, P. FRANKHAUSER, K. HUMPERT, U. KULL, E. SCHAUR (1994), Selbstorganisation urbaner Strukturen, *ARCH+, Vol. 27, Heft 121, pp. 57-68*

BEN-AKIVA M., S.R. LERMAN (1985), Discrete choice analysis: Theory and application to travel demand, The MIT Press, Cambridge

BENJAMINS M., CH. LINDVELD, R. VAN NES (2002), Multimodal travel choice modeling: supernetwork approach, *Paper 02-2948, CD-ROM of the 81st Annual Meeting of the Transportation Research Board,* Transportation Research Board, Washington DC

BERECHMAN J. (1993), Public transit economics and deregulation policy, North Holland, Amsterdam

BERECHMAN J., K.A. SMALL (1988), Modeling land use and transportation: an interpretive review for growth areas, *Environment and Planning A, Vol. 20, pp. 1285-1309*

BERNARD III M.J., N.E. COLLINS (2002), Shared Evs in Denver: Extending the station car and car sharing concepts, *Paper 02-3081, CD-ROM of the 81st Annual meeting of the Transportation Research Board,* Transportation Research Board, Washington DC

BIELLI M., M. CARAMIA, P. CAROTENUTO (2002), Genetic algorithms in bus network optimization, *Transportation Research C, Vol. 10, No. 1, pp. 19-34*

BIERSCHENK H., B. KEPPELER (2000), Verbindungsqualitäten für die Entwickelung integrierter Verkehrsnetze, *Internationales Verkehrswesen, Vol 52, No. 11, pp. 484-491*

BLACK A. (1976), Optimizing density of development with respect to transportation costs, *Highway Research Record 207, pp. 22-31*

BLACK A. (1995), Urban mass transportation planning, McGraw-Hill Inc, New York

BOUMA A., C. OLTROGGE (1994), Linienplanung und Simulation für öffentliche Verkehrswege in Praxis und Theorie, *ETR Eisenbahntechnische Rundschau, Vol. 43, pp.369-378*

BOUMAN P.A., A.A. BRUIJN, S.C. KIEFT (2001), Intercity stopt ook te Vleuten West: Alternatieve bedieningsmodellen door de bril van de reiziger, *In: Colloquium Vervoerplanologisch Speurwerk, pp. 869-888,* CVS, Delft

BOVY P.H.L. (1981), Het kortste tijd routekeuzeprincipe: een toetsing van de voorspellende kwaliteit, *Verkeerskunde, Vol. 32, No. 6, pp. 291-297*

BOVY P.H.L. (1985), Funktionele wegklassificatie in een fietsnetwerk onderzocht met routedata, *In: De Wit T., Verkeerskundige Werkdagen 1985,* KIVI/SVT, Driebergen-Rijsenburg

BOVY P.H.L. (2001), Traffic flooding the low Countries; how the Dutch cope with motorway congestion, *Transportation Reviews, Vol. 21, No. 1, pp. 89-116*

BOVY P.H.L., TH.J.H. SCHOEMAKER, A.J. VAN BINSBERGEN (1994), Vervoerkundige aspecten bij integrale prioriteitstelling hoofdinfrastructuur, Ministry of Transport, Public Works and Watermanagement, Den Haag

BOYCE D.E., L.J. LEBLANC, G.R.M. JANSEN (1990), Design of a toll system for a congested intercity road network: the Dutch road pricing system, *Paper prepared for the North American Meeetings, Regional Science Association International*, Boston

BRUIJN A.A. (1994), Combinatielijnen: verschillende stelselvormen op één lijn, Deelrapport I, Mogelijkheden van combinatielijnen, *Report VK 5101.704,* Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft

CARRESE S., S. GORI, T. PICANO (1996), Relationships between parking locations and traffic flows in urban areas, *In: L. Bianco, P. Toth, Advanced methods in transportation analysis,* Springer-Verlag, Berlin

CASCETTA E. (2001), Transportation systems engineering: Theory and methods, Kluwer Academic Publishers, Dordrecht

CATALANO S., C. LINDVELD, K. CARLIER (2001), A hypernetwork approach toward multimodal transportation analysis, Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft

CBS (1996), Onderzoek verplaatsingsgedrag 1995, Documentatie voor tape-gebruikers, CBS, Heerlen

CEDER, A. Y. ISRAELI (1998), User and operator perspectives in transit network design, *Transportation Research Record 1623, pp. 3-7*

CENTRAL TRANSPORTATION PLANNING STAFF (1997), Transfer penalties in urban mode choice modeling, *Travel model improvement program,* US Department of Transportation and US Environmental Protection Agency

CHANG S.K., P.M. SCHONFELD (1991), Optimization models for comparing conventional and subscription bus feeder services, *Transportation Science, Vol. 25, No. 4, pp. 281-298*

CHANG S.K., P.M. SCHONFELD (1993A), Optimal dimensions of bus service zones, *Journal of Transportation Engineering, Vol. 119, No. 4, pp. 567-585*

CHANG S.K., P.M. SCHONFELD (1993B), Welfare maximization with financial constraints for bus transit systems, *Transportation Research Record 1395, pp. 48-57*

CHANG S.K., W.J. YU (1996), Comparison of subsidized fixed- and flexible-route bus systems, *Transportation Research Record 1557, pp. 15-20*

CHRISTALLER W. (1933), Die zentralen Orte in Süddeutschland: eine ökonomisch-geographische Untersuchung über die Gesetzmässigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischen Funktionen, Jena

CHRISTALLER W. (1966), Central places in southern Germany; translated by C.W. Baskin, Prentice Hall, New York

COUSINS S.H. (1998), A model of car availability in car-sharing schemes of different sizes, *Traffic Engineering and Control, pp. 90-94*

COMMISSIE RONA (1992), Richtlijnen voor het ontwerp van niet autosnelwegen (RONA) (voorlopig), Hoofdstuk 1 Basiscriteria, SDU Uitgeverij, Den Haag

CREIGHTON R.L. (1970), Urban transportation planning, University of Ilinois Press, Urbana

DAGANZO C.F., C.T. HENDRICKSON, N.H.M. WILSON (1977), An approximate analytical model of many-to-one demand responsive transportation systems, *In: T. Sasaki, T. Yamakeo, Proc. of the 7th int. symposium on transportation and traffic theory,* The Institute of Systems Science research, Kyoto

DAVIDSON W.A., L. YANG (1999), Modeling of commuter rail station choice and access mode combinations, *Paper presented at the 78th Annual Meeting of the Transportation Research Board,* Transportation Research Board, Washington

DE BRUIJN T. (1998), Treintaxi hengelt naar steun gemeenten, *OV-magazine, Vol. 4, No. 6, pp. 16-18*

DE HEIJ H.T.A., J.H.M. MAASSEN (1995), Almere: hoogwaardig openbaar vervoer in de praktijk, *Verkeerskunde, Vol. 46, No. 7/8, pp. 44-48*

DE JONG T.M. (1988A), Morfologische typologie van netwerken, *Monografieën milieuplanning / SOM3,* Delft University of Technology, Faculty of Architecture, Delft

DE JONG T.M. (1988B), Milieudifferentiatie, *Monografieën milieuplanning / SOM 2,* Delft University Press, Delft

DE JONG T.M., M. PAASMAN (1998), Het metropolitane debat, Een vocabulaire voor besluitvorming over de kaart van Nederland, Stichting voor Milieu en Stedelijke Ontwikkeling, Zoetermeer

DHV (1994), Het Randstad model, DHV, Amersfoort

EGETER B. (1993), Systeemopbouw in stedelijke gebieden, theorievorming en netwerkoptimalisatie, *Report VK 5115.301,* Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft

EGETER B., TH.J.H. SCHOEMAKER, A VAN BINSBERGEN (1994), Multimodale opbouw van het personenvervoersysteem, Verkenning en theoretisch fundament, *Report VK 1111.301,* Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft

EGETER B., TH.J.H. SCHOEMAKER, M.J. NOBELEN, J.G. QUEE, J.H.M. SEERDEN (1990), Transferia, *Report VK 1501.301,* Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft

FERRARI P. (1999), A model of urban transport management, *Transportation Research B, Vol. 33, pp. 43-61*

FGSV (FORSCHUNGSGESELLSCHAFT FÜR STRASSEN- UND VERKEHRSWESEN) (1988), Richtlinien für die Anlage von Strassen (RAS), Teil: Leitfaden für die funktionale Gliederung des Strassennetzes (RAS-N), Kirschbaum Verlag, Bonn-Bad Godesberg

FRANK D., C. STORATH, J. SUMPF (1994), Mindestsiedlungsdichte für den ÖPNV, *Internationales Verkehrswesen, Vol. 46, No. 1+2, pp. 25-31*

FURTH P.G. (1986), Zonal route design for transit corridors, *Transportation Science, Vol. 20, No. 1, pp. 1-12*

GARCIA R., A. MARÍN (2001), Urban multimodal interchange design methodology, *In: Pursula M., Niitymäki J.: Mathematical methods on optimization in transportation systems,* Kluwer Academic Publishers, Dordrecht

GIBBONS R. (1992), Game theory for applied economists, Princeton University Press, Princeton

GOUDAPPEL COFFENG, BELEIDSONDERZOEK & -ADVIES (1998), Collectief vervoer in een posturbaan milieu, Eindrapport, Projectbureau Integrale Verkeers- en Vervoerstudies, Den Haag

HAGGETT P (1965), Locational analysis in human geography, Edward Arnold (Publishers), London

HAGGETT P., R.J. CHORLEY (1969), Network analysis in geography, Edward Arnold (Publishers), London

HAGUE CONSULTING GROUP (1997), Landelijk Model Systeem 5.0 Deel D, *HCG*, Den Haag

HASSELSTRÖM D. (1979), A method for optimization of urban bus route networks, VOLVO Bus Corporation, Göteborg

HEINZEL G. (1996), Individueller ÖPNV mit den 'Anrufbus' - ein Angebot zur rechten Zeit?, *Internationales Verkehrswesen, Vol. 48, No. 6, pp. 43-49*

HILBERS H.D., I. WILMINK, G.R.M. JANSEN (1997), Randstad, Ruhrgebied en de regio Antwerpen-Brussel-Gent, In: Egeter B., N. Kalfs, Colloquium Vervoerplanologisch Speurwerk 1997, pp. 739-758, CVS, Delft

HOLROYD E.M. (1967), The optimum bus service: a theoretical model for a large uniform urban area, *In: Vehicular traffic science, Proceedings of the 3rd Int. Symposium on the Theory of Traffic Flow,* Elsevier, New York

HORNER M.W., M.E. O'KELLY (2001), Embedding economies of scale concepts for hub network design, *Journal of Transport Geography, Vol. 9, pp. 255-265*

HUPKES G. (1977), Gasgeven of afremmen, toekomstscenario's voor ons vervoerssysteem, Kluwer, Deventer

IMMERS L.H., B. EGETER (1996), Visie regionaal openbaar vervoer, Advies aan IPO, Report VK 5101.307, TU Delft/TNO-INRO, Delft

IMMERS L.H., C.A. SMITS, B. EGETER, C.D. VAN GOEVERDEN, TH.J.H. SCHOEMAKER (1994), Ontwerpmethodiek nationale OV-dienstverlening, TNO-INRO, Delft

IMMERS L.H., I.R. WILMINK, J.E. STADA (2001), Bypasses voor bereikbaarheid, TNO-INRO, Delft

JANARTHANAN N., J.B. SCHNEIDER (1988), Development of an expert system to assist in the interactive graphic transit design process, *Transportation Research Record 1187, pp. 30-46*

JANSEN A.C.M. (1980), Aspekten van de moderne lokatietheorie, *In: J.G. Lambooy, Ekonomie en ruimte, Deel 1: lokatietheorie en regionale vraagstukken*Van Gorcum, Assen

JANSSON K. (1996), Welfare aspects on organization of passenger transport, *In: D. Hensher, J. King, T.H. Oum, World Transport Research Part 3 (7th WCTR, 1995),* Pergamon, Oxford

JOHNSON D.S., J.K. LENSTRA, A.H.G. RINNOOY KAN (1978), The complexity of the network design problem, *Networks, No. 8, pp. 279-285*

JUNG V. (1996), Lange anwege und hohe Geschwindigkeit oder: Kuze Anwege und niedrige Geschwindigkeiten, *Internationales Verkehrswesen, Vol. 48, No. 5, pp. 12-17*

KIRCHHOFF P., M. FRIEDRICH, M. HALLER (1994), ÖPNV-Planung in Landkreisen, *Der Nahverkehr, No. 9, pp. 10-16*

KIRCHHOFF P., M. FRIEDRICH, M. HALLER (1996), Rufbus-linien im Landkreis Erding, *Der Nahverkehr, No. 4, pp. 29-32*

KNIPPENBERG H., B. DE PATER (1988), De eenwording van Nederland, Schaalvergroting en integratie sinds 1800, SUN, Nijmegen

KOCUR G., C. HENDRICKSON (1982), Design of local bus service with demand equilibration, *Transportation Science, Vol. 16, No. 2, pp. 149-170*

KÖHLER U. (1989), Die Bedeutung der Zentralität von Orten für die Planung von Verkehrsnetzen, *Internationales Verkehrswesen, Vol. 41, No. 2, pp. 85-89*

KOOT A., B. GOVERS (1995), Systeemopbouw openbaar vervoer in de praktijk, *Verkeerskunde, Vol. 46, No. 7/8, pp. 26-30*

KREUTZBERGER E. (1999), Innovative networks and new-generation terminals for intermodal transport, *TRAIL 5th Annual Congress 1999, Five years " Crossroads of theory and practice", Part 3,* Delft University Press, Delft

KUAH G.K., J. PERL (1988), Optimization of feeder bus routes and bus-stop spacing, *Journal of Transportation Engineering, Vol. 114, No. 3, pp. 341-354*

KWAKERNAAK C.V., R. VAN NES (2001), Openbaar vervoerssystemen in wereldsteden, Een beschrijving van de hiërarchie in netwerkopbouw (concept), Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft

LAMBOOY J.G. (1980), Ekonomie en ruimte, Deel 1: lokatietheorie en regionale vraagstukken, Van Gorcum, Assen

LAMPKIN W., P.D. SAALMANS (1967), The design of routes, service frequencies and schedules for a municipal bus undertaking, a case study, *Operations Research Quarterly, Vol. 18, No. 4, pp. 375-397*

LAY M.G. (1990), Handbook of road technology, Vol. 1, Planning and pavements, 2nd edition, *Transportation Studies Vol. 8,* Gordon and Breach Science Publishers, New York

LEBLANC L.J., D.E. BOYCE (1986), A bilevel programming algorithm for exact solution of the network design problem with user-optimal flows, *Transportation Research B, Vol. 20, No. 3, pp. 259-265*

LEES L.H., P.M. HOHENBERG (1988), How cities grew in the western world: a systems approach, *In: Ausubel J.H., R. Herman, Cities and their vital systems: Infrastructure Past, Present & Future,* National Academy Press, Washington

LESLEY L.J.S. (1973), Determining the optimum number of multi-modal interchange stations for a city, *PTRC Winterconference,* Sussex University, Sussex

LIU R., R.M. PENDAYLA, S. POLZIN (1998), Simulation of the effects of intermodal transfer penalties on transit use, *Transportation Research Record 1623, pp. 88-95*

LO H.K., Y.K. TUNG (2001), Network design for improving trip time reliability, *Paper presented at the 80th Annual Meeting of the Transportation Research Board,* Transportation Research Board, Washington

LOUREIRO C.F.G., B. RALSTON (1996), Investment selection model for multicommodity multimodal transportation networks, *Transportation Research Record 1522, pp. 38-46*

MA J., J.B. SCHNEIDER (1991), Designing personal rapid transit (PRT) networks, *Journal of Advanced Transportation, Vol. 25, No. 3, pp. 247-268*

MAGEEAN J.F., J.D. NELSON (2001), Some observations on the organisation of demand responsive transport services, *Transporti Europei*, pp. 49-57

MANHEIM M.L. (1984), Fundamentals of transportation systems analysis, Volume 1: Basic concepts, The MIT Press, Cambridge

MAGNANTI T.L., R.T. WONG (1984), Network design and transportation planning: models and algorithms, *Transportation Science, Vol. 18, No. 1, pp. 1-55*

MASSOT M.H., J.F. ALLOUCHE, E. BÉNÉJAM, M. PARENT (1999), Praxitelle: Preliminary results from the Saint-Quentin experiment, *Paper presented at the 78th Annual Meeting of the Transportation Research Board,* Transportation Research Board, Washington

MEDDA F., P. NIJKAMP, P. RIETVELD (1999), Urban industrial relocation: the theory of edge cities, *Environment and Planning B, Vol. 26, pp. 751-761*

MENSEBACH W. (1994), Strassenverkehrstechnik, Werner-Verlag GmbH., Dusseldorf

MINISTRY OF TRANSPORT, PUBLIC WORKS AND WATERMANAGEMENT (1989), Second Transport Structure Plan, Ministry of Transport, Public Works and Watermanagement, Den Haag

MINISTRY OF TRANSPORT, PUBLIC WORKS AND WATERMANAGEMENT (1996), AGV, Vrije Universiteit Amsterdam: Handboek economische effecten infrastructuur, Ministry of Transport, Public Works and Watermanagement, Den Haag

MINISTRY OF TRANSPORT, PUBLIC WORKS AND WATERMANAGEMENT (1998), Raamwerk integrale beleidsvorming hoofdinfrastructuur (Raamwerk Pi), Ministry of Transport, Public Works and Watermanagement, Den Haag

MINISTRY OF TRANSPORT, PUBLIC WORKS AND WATERMANAGEMENT (2000), Van A naar Beter, Nationaal Verkeers- en Vervoerplan 2001-2020, Ministry of Transport, Public Works and Watermanagement, Den Haag

MOLIN E.J.E., S. HOOGENDOORN-LANSER (2001), De OV-fiets: het gebruik voorspeld op basis van 'stated choices', *In: Colloquium Vervoerplanologisch Speurwerk 2001,* CVS, Delft

MORLOK E.K. (1978), Introduction to transportation engineering and planning, McGraw-Hill Book Company, New York

MU-CONSULT (1996), Multimodaal  personenvervoer, uitwerking van kansrijke concepten
Projectbureau Integrale Verkeers- en Vervoerstudies, Den Haag

MU-CONSULT (1998), Multimodaal vervoer: Kwantificering potenties, Eindrapport, Projectbureau Integrale Verkeers- en Vervoerstudies, Den Haag

NÄGELE R.C., P.T. WILBERS, R.A. DE BRUIN (1992), Integratie fiets - openbaar vervoer, Traffic Test, Veenendaal

NEI, DHV (2000), Doorstroomroute A4, Zaandam/Almere - Belgische grens, NEI, Rotterdam

NERENBERG V., M.J.BERNARD III, N.E. COLLINS (1999), San Francisco Bay area station car demonstration evaluation results, *Paper presented at the 78th Annual meeting of the Transportation Research Board,* Transportation Research Board, Washington

NEWELL G.F. (1979), Some issues relating to the optimal design of bus routes, *Transportation Science, Vol. 13, No. 1, pp. 20-35*

NIELSEN O.A. (2000), A stochastic transit assignment model considering differences in passengers utility functions, *Transportation Research B., Vol. 34, pp. 377-402*

NUZZOLO A., F. RUSSO (1996), Stochastic assigment models for transit low frequency services: some theoretical and operative aspects, *In: L. Bianco, P. Toth, Advanced methods in transportation analysis,* Springer-Verlag, Berlin

NUZZOLO A., F. RUSSO, U. CRISALLI (2001), A doubly dynamic schedule-based assignment model for transit networks, *Transportation Science, Vol. 35, No. 3, pp. 268-285*

O'KELLY M.E., BRIAN D.L. (1998), Hub location with flow economies of scale, *Transportation Research B, Vol. 32, No. 8, pp. 605-616*

ORTÚZAR J. DE D., L.G. WILLUMSEN (1994), Modelling transport, 2nd edition, John Wiley & Son, Chichester

PAGE S. E. (1998), On the emergence of cities, *Working Papers Santa Fe Institute,* Santa Fe Institute, Santa Fe

PROJECTBUREAU IVVS (1995), Aspecten bij het ontwerpen van netwerken voor het stads- en streekvervoer, Ministry of Transport, Public Works and Watermanagement, Den Haag

PROJECTGROEP TRAMPLUS (1992), Handboek TramPlus, Projectbureau TramPlus, Rotterdam

PUU T. (1978), Towards a theory of optimal roads, *Regional Science and Urban Economics, Vol. 8, pp. 225-248*

RAAD VOOR VERKEER EN WATERSTAAT (1996), Advies knooppunten in openbaar vervoernetwerken, Dutch Advisory Council for Transport, Public Works and Watermanagement, Den Haag

RAAD VOOR VERKEER EN WATERSTAAT (2001), Van modal split naar modal merge, Advies over de toekomst van hetregionaal verkeer en vervoer, Dutch Advisory Council for Transport, Public Works and Watermanagement, Den Haag

ROEDOE A. (1995), Grofweg beter, een casestudie naar een nieuw stedelijk OV-netwerk, *Report VK 5115.701,* Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft

RUTTEN B.J.C.M. (1995), Intermodaal railvervoer: realiteit of utopie? Een ontwerpmethode voor een netwerk van intermodale inland weg-railterminals, *Tijdschrift voor Vervoerwetenschap, No. 4, pp. 323-337*

SALINGAROS N.A., B.J. WEST (1999), A universal rule for the distribution of sizes, *Environment and Planning B, Vol. 26, pp. 909-923*

SCHAAFSMA A. (2001), Dynamisch railverkeersmanagement, besturingsconcept voor railverkeer op basis van het Lagenmodel Verkeer en vervoer, *TRAIL Thesis Series, T2001/7,* Delft University Press, Delft

SCHAFER A. (1998), The global demand for motorized mobility, *Transportation Research A, Vol. 32, No. 6, pp. 455-477*

SCHÄFFELER U. (1999), Optimale Haltestellenabstände für den öffentlichen Nahverkehr in den Niederlanden unter spezieller Beachtung des multimodalen Verkehrs, *Masters Thesis,* Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft

SCHOEMAKER TH.J.H., B. EGETER, C.D. VAN GOEVERDEN (1993), ARGUS, een metropolitaan vervoerstelsel in de Randstad; ontwerp en vervoerprognose, *Report VK 5117.301,* Delft University of Technology, Delft

SCHOEMAKER TH.J.H., K. KOOLSTRA, P.H.L. BOVY (1999), Traffic in the 21st century - a scenario analysis for the traffic market in 2030, *In: Weijnen M.P.C., E.F. ten Heuvelhof, The infrastructure playing field in 2030, pp. 175-194,* Delft University Press, Delft

SCHÖNHARTING J. (1997), Integrierte Netzgestaltung: konzeptioneller Ansatz und Qualitätsziele, *Strassenverkehrstechnik, No. 4, pp. 161-167*

SCHÖNHARTING J., TH. PISCHNER (1983), Untersuchungen zur Ableitung van Strassenkategorieabhängigen Soll-Qualitäten aus Erreichbarkeitsanalysen, *Forschung Strassenbau und Strassenverkehrstechnik, Heft 399*

SHEFFI Y. (1985), Urban transportation networks, Equilibrium analysis with mathematical programming methods, Prentive Hall, Englewood Cliffs

SHIH M.C., H.S. MAHMASSANI, M.H. BAAJ (1998), Planning and design model for transit route networks with coordinated operations, *Transportation Research Record 1623, pp. 16-23*

SIEMENS NEDERLAND N.V. (2000), Zweven is vrijheid, Klaar voor vertrek, Siemens Nederland N.V., Den Haag

SOLANKI R.S., J.K. GORTI, F. SOUTHWORTH (1998), Using decomposition in large-scale network design with a quasi-optimization heuristic, *Transportation Research B, Vol. 32, pp. 127-140*

SPASOVIC L.N., M.P. BOILE, A.K. BLADIKAS (1994), Bus transit service coverage for maximum profit and social welfare, *Transportation Research Record 1451, pp. 12-22*

SPIEKERMAN T., A. WEINSTEIN (1998), Evaluating th epotential commuter market for station car usage at the Bay area rapid transit district, *Paper presented at the 77th Annual meeting of the Transportation Research Board,* Transportation Research Board, Washington

SPSS INC. (1998), SPSS for Windows, release 9.0.0, SPSS Inc., Chicago

STAATSUITGEVERIJ (1966), Tweede nota over de ruimtelijke ordening in Nederland, Verkorte uitgave, Staatsuitgeverij, Den Haag

STADSGEWEST HAAGLANDEN (1999), Agglonet bereikt meer, Stadsgewest Haaglanden, Den Haag

STEENBRINK P. (1974), Optimization of transport networks, John Wiley & Sons, New York

STEIN D.M. (1978), Scheduling dial-a-ride transportation systems, *Transportation Science, Vol. 12, No. 3, pp. 232-249*

TSAO S., P. SCHONFELD (1983), Optimization of zonal transit service, *Journal of Transportation Engineering, Vol. 109, No. 2, pp. 257-272*

TSAO S., P. SCHONFELD (1984), Branched transit services: An analysis, *Journal of Transportation Engineering, Vol. 110, No. 1, pp. 112-128*

VAN BINSBERGEN A., J. VISSER (2001), Innovation steps towards efficient good distribution systems for urban areas, *TRAIL Thesis Series, T2001/5,* Delft University Press, Delft

VAN BINSBERGEN A.J., SCHOEMAKER TH.H.J., B. EGETER, E. HERINGA (1992), Transferia, locatiekeuze en raming van gebruik, *Report VK 1510.302,* Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft

VAN DEN BERG (1998), Een onvoorspelbare geldhoop, *NRC Handelsblad, 28 november 1998,* Perscombinatie, Amsterdam

VAN DEN HEUVEL M.G. (1997), Openbaar vervoer in de Randstad, een systematische aanpak, Thesis Publishers Amsterdam, Amsterdam

VAN DER WAARD J. (1988), Onderzoek weging tijdelementen, Deelrapport 3, Analyse routekeuzegedrag van openbaar vervoerreizigers, *Report VK 5302.303,* Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft

VAN DER WAARD J. (1988), The relative importance of public transport trip-time attributes in route-choice, PTRC 1988

VAN DER WOUD A. (1987), Het lege land, De ruimtelijke orde van Nederland 1798-1848, Meulenhof Informatief, Amsterdam

VAN GOEVERDEN C.D., B. EGETER (1993), Gecombineerd gebruik van fiets en openbaar vervoer, Verwachte effecten op de vervoerwijzekeuze van optimale fietsbeschikbaarheid in voor- en natransport, *Report VK 5116.301,* Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft

VAN GOEVERDEN C.D., B. EGETER, H.D. HILBERS (1998), Nieuwe vervoersystemen, Staalkaart vraag, *Report VK 1109.302,* Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft

VAN GOEVERDEN C.D., TH. J.H. SCHOEMAKER (2000), Openbaar Vervoer op eigen benen, Vervoerkundige effecten, AVV, Rotterdam

VAN MARREWIJK C., S. BRAKMAN, H. GARRRETSEN (1998), Moderne handelstheorie en lokatietheorie, *Economisch en Sociaal Tijdschrift*, Vol. 52, No. 4, pp. 479-507

VAN NES R. (1988), Optimalisatie openbaar vervoernetwerken, Een optimalisatiemodel voor het ontwerpen van openbaar vervoernetwerken, *Report VK 5301.301,* Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft

VAN NES R. (1991), Strategische openbaar vervoerstudies in stedelijke agglomeraties, *In: De Wit T., Verkeerskundige Werkdagen 1991, pp.243-251,* CROW, Ede

VAN NES R. (1998), Schaalniveaus in wegennetwerken, *In: Colloquium Vervoerplanologisch Speurwerk 1998: Sturen met structuren, pp. 1507-1522,* CVS, Delft

VAN NES R. (1999), Design of multimodal transport systems, Setting the scene: Review of literature and basic concepts, *TRAIL Studies in Transportation Science Series, No. S99/3,* Delft University Press, Delft

VAN NES R. (2000), Optimal stop and line spacing for urban public transport networks, Analysis of objectives and implications for planning practice, *TRAIL Studies in Transportation Science S2000/01,* Delft University Press, Delft

VAN NES R. (2001), The impact of altrnative access modes on urban public transport network design, *European Journal of Transportation and Infrastructure Research, Vol. 1, No. 2, pp. 127-210*

VAN NES R. (2001), Multi-level optimization of public transport networks, *Paper 02-2493, CD-ROM of the 81st Annual Meeting of the Transportation Research Board,* Transportation Research Board, Washington

VAN NES R., N.J. VAN DER ZIJPP (2000), Scale-factor 3 for hierarchical road networks: a natural phenomenon?, *ProceedingsTRAIL 6th Annual Congress, Part 1, pp. 131-156,* Delft University Press, Delft

VAN NES R., P.H.L BOVY (2000), The importance of objectives in urban transit network design, *Transportation Research Record 1735, pp. 25-34*

VAN NES R., R. HAMERSLAG, L.H. IMMERS (1988), Design of public transport networks, *Transportation Research Record 1202, pp. 74-83*

VAUGHAN R. (1987), Urban spatial traffic patterns, Pion Limited, London

VÖV (1981), Emphelungen für einen Bedienungsstandard im öffentlichen Personennahverkehr, Reihe Technik, VÖV 1.41.1, Köln

VUCHIC V.R. (1973), Skip-stop operation as a method for transit speed increase, *Traffic Quarterly, Vol. 27, No. 2, pp. 307-327*

VUCHIC V.R., A. MUSSO (1991), Theory and practice of metro network design, *Public Transport International, Vol. 39, No. 3, pp. 298-325*

WAGNER C. (2000), Car sharing enterprises and mobility systems, service development and technology strategies in the US and Europe, *Paper presented at the 79th Annual Meeting of the Transportation Research Board,* Transportation Research Board, Washington

WALDROP M.M (1992), Complexity: the emerging science at the edge of order and chaos, Simon & Schuster, New York

WALTHER K. (1973), Nachfrageorientierte bewertung der Streckenführung in Öffentlichen Personennahverkehr, *Veröffentlichten des Verkehrswissenschaftlichen Institutes des RWTH Aachen,* Aachen

WEGMAN F., R. WOUTERS (2002), Road safety policy in the Netherlands: facing the future, Annales des Ponts et Chaussées, nouvelle série no. 101, janviers-mars 2002

WERKGROEP '2DUIZEND' (1999), De consument is koning, eindraport, Stichting werkgroep '2duizend, Amersfoort

WEST G.B., J.H. BROWN, B.J. ENQUIST (1997), A general model for the origin of allometric scaling laws in biology, *Science, Vol. 276, No. 5309, pp. 122-126*

WEST G.B., J.H. BROWN, B.J. ENQUIST (1999), The fourth dimension of life: Fractal geometry and allometric scaling of organisms, *Science, Vol. 284, 4 june 1999, pp. 1677-1679*

WHITE P. (1995), Public transport, Its planning, management and operation, 3rd edition, UCL Press Limited, London

WILSON N.H.M., C. HENDRICKSON (1980), Performance models of flexibly routes transportation services, *Transportation Research Part B, Vol. 14B, pp. 67-78*

WIRASINGHE S.C. (1980), Nearly optimal parameters for a rail/feeder-bus system on a rectangular grid, *Transportation Research A, Vol.14, No. 1, pp. 33-40*

XIONG Y., J.B. SCHNEIDER (1992), Transportation network design using a cumulative genetic algorithm and neural network, *Transportation Research Record 1364, pp. 37-44*

XIONG Y., J.B. SCHNEIDER (1995), Processing of constraints in transportation network design problem, *Journal of Computing in Civil Engineering, Vol. 9, No. 1, pp. 21-28*

YAI T., S, IWAKURA, S. MORICHI (1997), Multinomial probit with structured covariance for route choice behavior, *Transportation Research B, Vol. 31, No. 3, pp. 195-207*

YANG H., M.G. BELL (1997), Traffic restraint, road pricing and network equilibrium, *Transportation Research B, Vol. 31, No. 4, pp. 303-314*

YANG H., M.G. BELL (1998), Models and algorithms for road network design: a review and some new developments, *Transport Reviews, Vol. 18, No, 3, pp. 257-278*

YEGOROV Y. (1997), Self organization of spatial infrastructure, *Working papers Santa Fe Institute,* Santa Fe Institute, Santa Fe

ZAHAVI Y. (1981), The UMOT-Urban Interactions, DOT-RSPA-DPB, US Department of Transportation, Washington DC

ZIPF G.K. (1949), Human behavior and the principle of least effort, Addison-Wesley, Cambridge MA

ZOU P. (1999), Optimierung der Organisations-, Dispositions- und Finanzierungsmodelle van Carsharing, *Strassenverkehrstechnik, No. 8, pp. 376-380*

ZWIJGERS T. (2001), Treintaxi gaat het liefst op in een landelijk CVV-systeem, *OV-Magazine, Vol. 7, No. 8, pp. 11-13*

# Appendix A
## Potential for multimodal transport based on trip purpose

The assessment of the potential of multimodal transport based on trip purpose consists of two elements: input from the National Travel Survey regarding the modal split per trip purpose per distance class and a set of assumptions for general growth factors and specific correction factors. These growth and correction factors are applied per trip purpose per distance class to assess the potential share of public transport given the set of assumptions.

The input from the National Travel Survey (1995) is:

- Percentage of trips per trip purpose per distance class (10 - 30 kilometres, 30 - 100 kilometres);

- Share in modal split for car driver, car passenger and public transport.

The assumptions are:

- Growth for trip purpose work: 50% or 15%;

- Growth of share of public transport to match the (already increased!) share for work trips (excluding trip purposes education, touring, and picking-up and dropping-off passengers);

- Correction for time of day: 0.67 (social, recreation);

- Correction for car occupancy: share car passenger divided by share car passenger for work trips (maximum correction is 1);

- Correction for value of time: value of time business trips divided by value of time non-business trips (business only).

The following results can then be calculated:

- Growth of share of public transport multiplied by all correction factors;

- Resulting share for public transport per trip purpose and distance class.

Three scenarios are analysed:

- Maximum: Growth factor work is 50%, all trip purposes except education, touring and picking-up/dropping-off, match the increased share of public transport for the trip purpose work. No correction factors are used;

- Medium: The maximum scenario plus all correction factors;

- Minimum: Growth factor for work is 15% plus all correction factors.

This table shows an overview of input, assumptions and results for the medium scenario:

| | Work | Social | Shopping | Business personal | Business work | Recreation | Personal care | Other | Excluded | **Total** |
|---|---|---|---|---|---|---|---|---|---|---|
| **10-30 km** | | | | | | | | | | |
| *National Travel Survey* | | | | | | | | | | |
| percentage of all trips % | 4.1 | 2.9 | 2.4 | 0.3 | 0.7 | 2.1 | 0.4 | 0.8 | 2.4 | **16.2** |
| share car driver % | 66.1 | 46.7 | 49.6 | 65.0 | 83.1 | 40.0 | 51.1 | 47.9 | 34.7 | **51.1** |
| share car passenger % | 10.5 | 39.4 | 35.1 | 23.0 | 10.4 | 43.0 | 36.9 | 34.2 | 14.8 | **26.2** |
| share public transport % | 12.0 | 6.7 | 7.9 | 7.7 | 3.4 | 6.1 | 7.5 | 8.2 | 19.1 | **10.1** |
| *Assumptions* | | | | | | | | | | |
| growth factor work | 1.5 | 1.5 | 1.5 | 1.5 | 1.5 | 1.5 | 1.5 | 1.5 | | |
| growth to share trip purpose work | 1.0 | 1.7 | 1.3 | 1.3 | 4.3 | 2.0 | 1.4 | 1.2 | | |
| correction time of day | 1.00 | 0.67 | 1.00 | 1.00 | 1.00 | 0.67 | 1.00 | 1.00 | | |
| correction car passenger | 1.00 | 0.27 | 0.30 | 0.46 | 1.00 | 0.24 | 0.29 | 0.31 | | |
| correction value of time | 1.00 | 1.00 | 1.00 | 1.00 | 0.25 | 1.00 | 1.00 | 1.00 | | |
| *Results* | | | | | | | | | | |
| growth public transport | 1.5 | 1.3 | 1.4 | 1.6 | 1.1 | 1.3 | 1.4 | 1.4 | 0.0 | **1.30** |
| share public transport % | 18.0 | 8.7 | 10.9 | 12.4 | 7.0 | 8.0 | 10.5 | 11.2 | 19.1 | **13.1** |
| **30-100 km** | | | | | | | | | | |
| *National Travel Survey* | | | | | | | | | | |
| percentage of all trips % | 1.8 | 1.4 | 0.4 | 0.0 | 0.4 | 0.8 | 0.0 | 0.4 | 0.9 | **6.2** |
| share car driver % | 66.5 | 44.5 | 45.2 | 0.0 | 82.6 | 40.2 | 0.0 | 46.2 | 30.3 | **51.7** |
| share car passenger % | 12.6 | 43.8 | 39.4 | 0.0 | 10.4 | 43.9 | 0.0 | 35.7 | 18.8 | **27.9** |
| share public transport % | 18.9 | 10.4 | 14.2 | 0.0 | 6.4 | 10.4 | 0.0 | 13.9 | 35.2 | **16.3** |
| *Assumptions* | | | | | | | | | | |
| growth factor work | 1.5 | 1.5 | 1.5 | 0.0 | 1.5 | 1.5 | 0.0 | 1.5 | | |
| growth to share trip purpose work | 1.0 | 1.7 | 1.0 | 0.0 | 3.4 | 1.7 | 0.0 | 1.0 | | |
| correction time of day | 1.00 | 0.67 | 1.00 | 1.00 | 1.00 | 0.67 | 1.00 | 1.00 | | |
| correction car passenger | 1.00 | 0.29 | 0.32 | 1.00 | 1.00 | 0.29 | 1.00 | 0.35 | | |
| correction value of time | 1.00 | 1.00 | 1.00 | 1.00 | 0.25 | 1.00 | 1.00 | 1.00 | | |
| *Results* | | | | | | | | | | |
| growth public transport | 1.5 | 1.3 | 1.3 | 1.0 | 1.9 | 1.3 | 1.0 | 1.47 | 0.0 | **1.30** |
| share public transport % | 28.3 | 13.8 | 18.7 | 0.0 | 11.9 | 13.8 | 0.0 | 19.0 | 35.2 | **21.3** |
| **Total 10-100 km** | | | | | | | | | | |
| Percentage of all trips | 5.9 | 4.3 | 2.8 | 0.3 | 1.1 | 2.9 | 0.4 | 1.2 | 3.3 | **22.4** |
| growth public transport | 1.5 | 1.3 | 1.4 | 1.6 | 1.0 | 1.3 | 1.4 | 1.4 | 0.0 | **1.30** |
| share public transport % | 21.2 | 10.4 | 12.0 | 12.4 | 8.8 | 9.6 | 10.5 | 13.8 | 23.5 | **15.4** |

*Excluded trips are trip purposes education, touring and picking-up/dropping-off.*

The increase of the public transport share by 30% for a trip type that accounts for 22.4% of all trips implies that 0.8% of all trips switched to public transport. Given the current share of 2.9% for multimodal transport, this switch to public transport is equivalent to an increase of multimodal trips by 28% leading to a total of 3.7% of all trips.

# Appendix B
## Potential for multimodal transport based on trip type

The trip types used in the assessment of the potential of multimodal transport are:

A:  Between city centre and another agglomeration or town: usually train and one transfer to an urban public transport system;

B:  Between agglomerations: train and two transfers to urban public transport systems;

C:  Between city centre and other origins and destinations; train and one transfer to a regional public transport system;

D:  Between agglomeration or towns and other origins and destinations; train and one transfer to an urban and one transfer to a regional public transport system;

E:  Between all other origins and destinations; bus and regional trains and two transfers to regional public transport systems.

The following table shows the size of these trip types and the corresponding share of public transport in the modal split (National Travel Survey 1995, Van Goeverden et al. (1998)).

| Trip type | 10 - 50 km | | Longer than 50 km | |
|---|---|---|---|---|
| | Percentage. of all trips | Share public transport | Percentage of all trips | Share public transport |
| **A** Between city centre and another agglomeration or town | 0.4 | 44 | 0.2 | 50 |
| **B** Between agglomerations | 1.7 | 21 | 0.8 | 25 |
| **C** Between city centre and other origins and destinations | 0.4 | 33 | 0.2 | 36 |
| **D** Between agglomeration or towns and other origins and destinations | 6.2 | 10.5 | 1.5 | 15 |
| **E** Between all other origins and destinations | 5.9 | 4 | 0.9 | 6 |
| **Subtotal** | 14.6 | 10.6 | 3.6 | 18.1 |
| **All trips longer than 10 km** | | | 18.2 | 12.1 |

In the analysis it is assumed that the share of public transport for trip type A is the maximum that is possible. For all other trip types the share of public transport is lower due to the next three characteristics of the public transport system:

* Additional transfer (RT);

* Difference between regional and urban public transport (RR);

* Difference in speed between trains and buses and regional trains (RS).

A simple model is used in which three characteristics are represented by multipliers for the public transport share of trip type A. The additional transfer (RT) applies to trip types B and D, the difference between urban an d regional transport (RR) to trip types C, D, and E, and the difference between train and regional buses (RS) for trip type E only. The values for these reduction factors are estimated by minimising the sum of the squares of the differences for the public transport share. Given these reduction factors two scenarios are defined:

- Maximum scenario: factor for transfer is increased by 15% and the difference between urban and regional public transport is eliminated;

- Minimum scenario: differences between urban and regional public transport are halved.

The following table shows the values for these multipliers as estimated and as used for the two scenarios:

|                | Estimated values | Maximum scenario | Minimum scenario |
|----------------|------------------|------------------|------------------|
| RT (transfer)  | 0.49             | 0.57             | 0.49             |
| RR (regional)  | 0.74             | 1.00             | 0.87             |
| RS (speed)     | 0.54             | 0.54             | 0.54             |

The following table shows the results for these two scenarios:

| Trip type | 10 - 50 km | | | | Longer than 50 km | | | |
|-----------|------------|------|------|------|-------------------|------|------|------|
|           | Percentage. of all trips | Share public transport | | | Percentage of all trips | Share public transport | | |
|           |            | NTS  | Max. | Min. |                   | NTS  | Max. | Min. |
| A         | 0.4        | 44.0 | 44.0 | 44.0 | 0.2               | 50.0 | 50.0 | 50.0 |
| B         | 1.7        | 21.0 | 24.9 | 21.7 | 0.8               | 25.0 | 28.3 | 24.7 |
| C         | 0.4        | 33.0 | 44.0 | 38.2 | 0.2               | 36.0 | 50.0 | 43.4 |
| D         | 6.2        | 10.5 | 24.9 | 16.4 | 1.5               | 15.0 | 28.3 | 18.6 |
| E         | 5.9        | 4.0  | 13.6 | 7.7  | 0.9               | 6.0  | 15.4 | 8.8  |
| Subtotal  | 14.6       | 10.6 | 21.4 | 14.8 | 3.6               | 18.1 | 27.5 | 20.6 |
| All trips longer than 10 km | | | | | 18.2 | 12.1 | 22.6 | 16.0 |

The increase of the public transport share up to 16.0% (or 22.6%) for a trip type that accounts for 18.2% of all trips implies that 0.7% (or 1.9%) of all trips switched to public transport. Given the current share of 2.9% for multimodal transport, this switch to public transport is equivalent to an increase of multimodal trips by 24% (or 66%) leading to a total of 3.6% (or 4.8%) of all trips.

# Appendix C
**Basic variables and characteristics for the main network types**

**Network type: Linear**
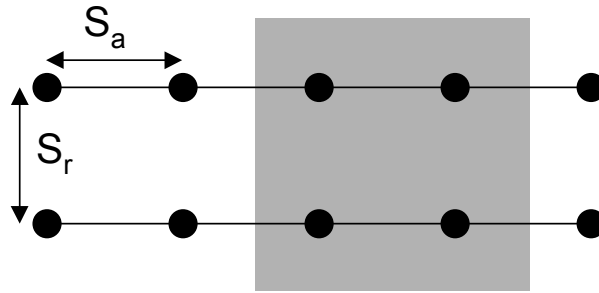


$S_a$ = Access spacing

$S_r$ = Road spacing

Network density: $$D_n = \frac{1}{S_r} \quad \frac{[km]}{[km^2]}$$

Crossing density:     Not relevant

Access density: $$D_a = \frac{1}{S_r \cdot S_a} \quad \frac{1}{[km^2]}$$

**Network type: Grid**



$S_a$ = Access spacing
$S_r$ = Road spacing
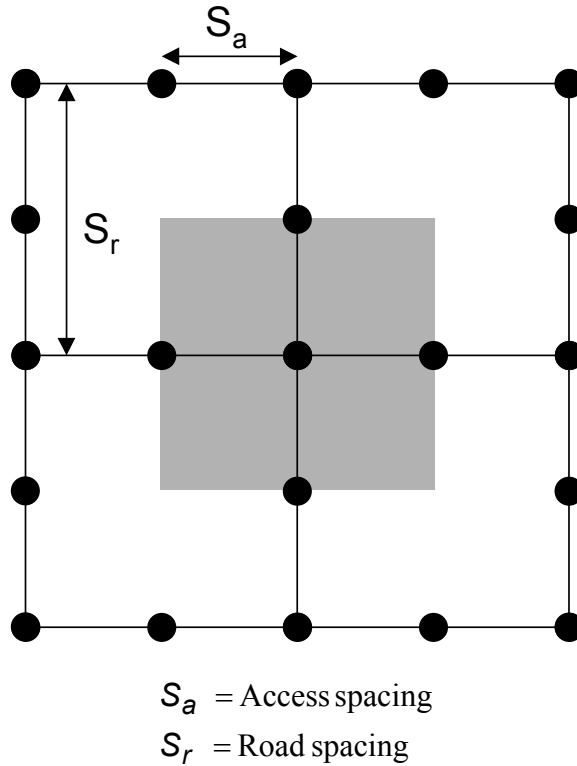
Network density:     $D_n = \dfrac{2}{S_r} \quad \dfrac{[km]}{[km^2]}$

Crossing density:     $D_c = \dfrac{1}{S_r{}^2} \quad \dfrac{1}{[km^2]}$

Access density:     $D_a = \dfrac{2}{S_r \cdot S_a} - \dfrac{1}{S_r{}^2} \quad \dfrac{1}{[km^2]}$

**Network type: Triangle**
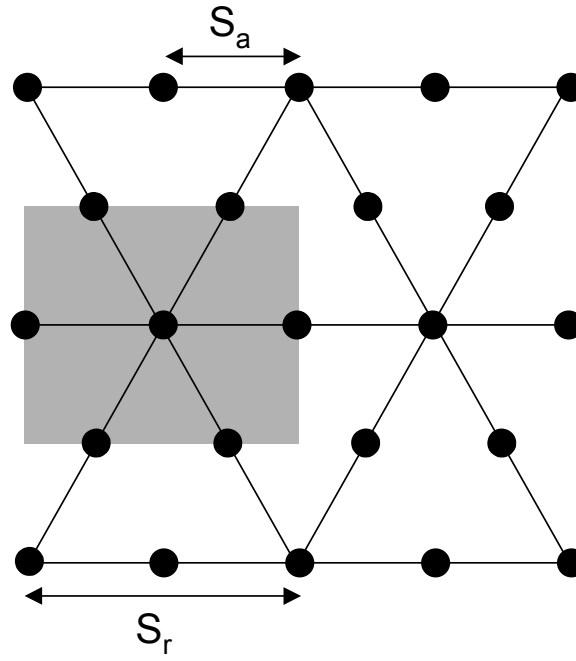


$$S_a = \text{Access spacing}$$
$$S_r = \text{Road spacing}$$

Network density:
$$D_n = \frac{3}{S_r \cdot \sin(\pi/3)} \quad \frac{[km]}{[km^2]}$$

Crossing density:
$$D_c = \frac{1}{S_r^2 \cdot \sin(\pi/3)} \quad \frac{1}{[km^2]}$$

Access density:
$$D_a = \frac{3}{S_r \cdot S_a \cdot \sin(\pi/3)} - \frac{2}{S_r^2 \cdot \sin(\pi/3)} \quad \frac{1}{[km^2]}$$

**Network type: Radial**

In this case there are 2 possible definitions for the densities, both of which are presented here:

1.  All criteria are related to the area served, that is the circle having radius $L_r$;
2.  All criteria are related to the square in which the network is located, having a width of $2L_r$.
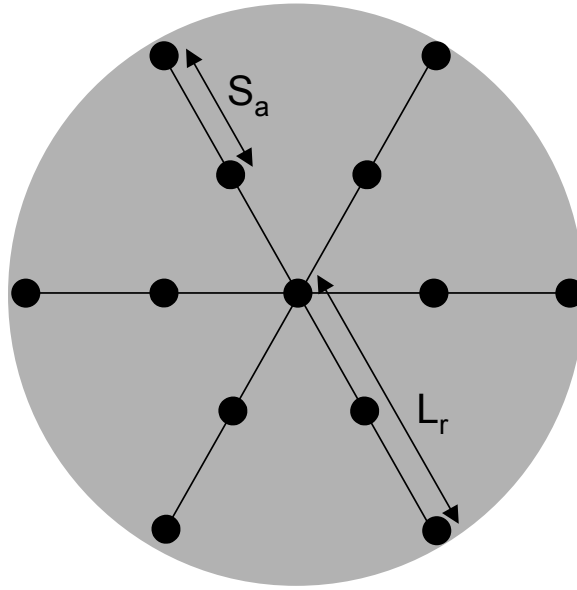


$$S_a = \text{Access spacing}$$
$$L_r = \text{Length of a radial}$$
$$NR = \text{Number of radials}$$

Network density:
$$D_n = \frac{NR}{\pi \cdot L_r} \quad \frac{[km]}{[km^2]} \quad \text{or} \quad D_n = \frac{NR}{4 \cdot L_r}$$

Crossing density:
$$D_c = \frac{1}{\pi \cdot L_r^2} \quad \frac{1}{[km^2]} \quad \text{or} \quad D_c = \frac{1}{4 \cdot L_r^2}$$

Access density:
$$D_a = \frac{NR}{\pi \cdot L_r \cdot S_a} + \frac{1}{\pi \cdot L_r^2} \quad \frac{1}{[km^2]} \quad \text{or} \quad D_a = \frac{NR}{4 \cdot L_r \cdot S_a} + \frac{1}{4 \cdot L_r^2}$$

# Appendix D
## Assessment of the consequences of using average distances in the analytical approach

The analytical model for urban public transport network design presented in Section 6.3.1 is based on two assumptions that might seem too simple at first sight. The first assumption is that the access time to the urban public transport system is determined using the average access distance, while assuming a uniformly distributed demand pattern around the stop. The second assumption is that the patronage is determined for the average trip length only. For both assumptions a more realistic approach is possible. The question, however, is whether a more detailed approach is necessary. First the assumption with respect to access time is analysed, and secondly the assumption for the trip length.

*Access time*

The service area of a stop can be divided into small squares having co-ordinates ($x,y$). Each square has the same potential for public transport to the city centre. Squares located close to the stop, however, have short access distances making public transport attractive, while squares further away have large access distances, which make public transport less attractive. Thus the actual demand for public transport varies over the service area.



Average access distance                    Detailed approach
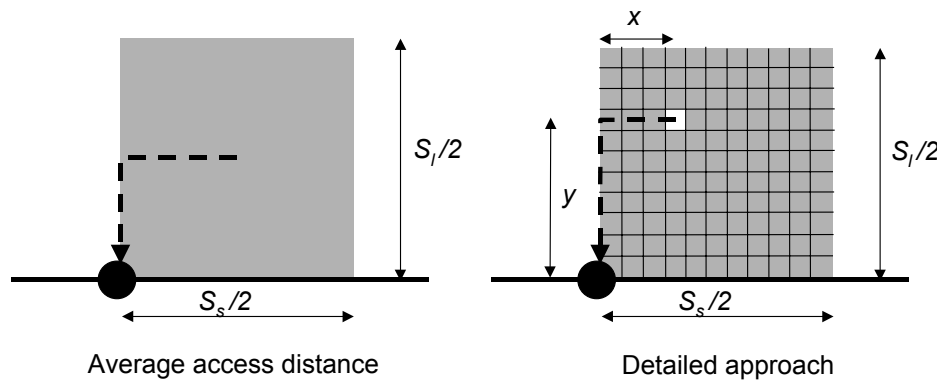
The total demand for the service area can be determined by summating the demand for all squares. If only the access time is considered while assuming that all other travel elements are fixed, the travel time for square ($x,y$) can be written as:

$$T_c(x,y) = \frac{w_a \cdot (x+y)}{v_a} + T_{pt}$$

with:

$T_c$ = weighted travel time to the city centre

$w_a$ = weight for access time

$v_a$ = access speed

$T_{pt}$ = weigthed travel time by public transport from the stop to the city centre

Please note that for simplicity sake the access routes are parallel and perpendicular to the public transport line. For each square the patronage can be determined using the logit-mode-choice model. The total patronage then is the sum of the patronage of all squares, or, if the squares are infinitely small, the integral over the service area defined by half the stop spacing and half the line spacing.

$$P_{sa} = \int_{x=1}^{S_s/2} \int_{y=1}^{S_l/2} P(T_c(x,y)) dxdy$$

with:

$P_{sa}$ = patronage for the service area

$S_s$ = stop spacing

$S_l$ = line spacing

$P(\ )$ = logit - mode - choice model for each square

The approach used in this thesis simply uses the average access distance to the stop. The total travel time then becomes:

$$T_c = \frac{w_a \cdot \frac{(S_s + S_l)}{4}}{v_a} + T_{pt}$$

If the patronage for the stop is calculated using both approaches and realistic values for the variables, it is found that the simple approach slightly overestimates the patronage. The maximum difference for service areas up to 750 by 750 metres (which is equivalent to stop spacing and line spacing of 1.500 metres), however, is 0.6%, which is clearly negligible.

In the previous analysis no average access distance is used, since the patronage was determined for each square separately. An alternative approach is to determine an average access time by weighing the access time of each square by its patronage. Thus short access times will have a larger impact than long access times. This effect is partly compensated by the fact that there are obviously more squares having long access times. The resulting average access time can then be used in the analytical model in the same way as described in Section 6.3.1. Using the same values for the variables it can be established that compared to this approach the simple approach slightly overestimates the average access time. If short trips are considered, trip length is 3 kilometres, the difference is 6%, while for longer distances the difference decreases. For 10-kilometre trips the difference is only 2%. The net effect on the patronage is even smaller, namely 2% and 0.2% respectively.

Both approaches clearly show that the error that is introduced by using the average access distance is very small indeed. Given the limited sensitivity of the objective functions, this simplification will have no significant effect on the results.

*Average trip length*

For the assessment of the patronage the average trip length is used. In Section 6.3.1.1 it was shown that this is appropriate if a fixed demand is assumed for all stops. In reality however, the demand depends on the quality of the services offered. Given the penalties related to using public transport, especially access time and waiting time, it is obvious that public transport is more attractive for long distance trips than for short distance trips. In this more realistic approach the total patronage becomes:

$$P = \sum_{s=1}^{n_s} P(T_c(L_s))$$

where:

$P(T_c(L_s))$ = patronage for stop $s$ at distance $L_s$

$n_s$ = number of stops

Again using realistic values for all variables, the following figure shows the ratio between the simple approach based on the average trip length and this more realistic approach:
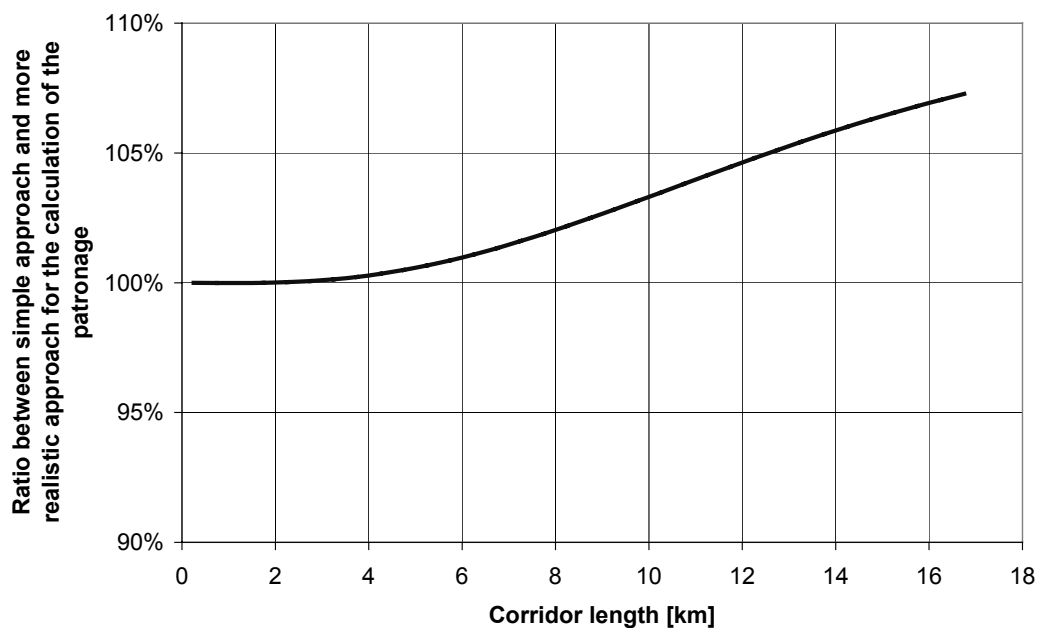


This figure clearly shows that the difference between both approaches is negligible for the urban corridors considered in the analysis of single-level networks (less than 2%), while for longer corridor lengths the patronage is slightly overestimated. The net effect, however, is still small. Since all analyses consider only comparisons of cases having identical corridor lengths, this overestimation of the demand will not influence the results. The only difference is that the values for social welfare will be too high.

# Appendix E

**Parameter values used for optimising urban public transport networks**

| Parameter | Symbol | Bus | Tram | Units |
|---|---|---|---|---|
| **Travel distance** | $L_c$ | 3 | 5 | km |
| **Access speed** | $v_a$ | 1.1 | 1.1 | m/s |
| **Factor access distance** | $f_a$ | 0.25 | 0.25 | |
| **Factor demand pattern** | $f_d$ | 1 | 1 | |
| **Maximum speed public transport** | $v$ | 13.9 | 13.9 | m/s |
| **Time lost at stops** | $T_s$ | 34 | 34 | s |
| **Factor waiting time** | $f_w$ | 1,800 | 1,800 | s |
| **Egress time** | $T_e$ | 180 | 180 | s |
| **Regular frequency** | $F$ | 5 | 8 | veh/h |
| **Fare** | $r_{tk}$ | 0.11 | 0.11 | €/km |
| **Subsidy** | $r_{sk}$ | 0 | 0 | €/km |
| **Operating costs per vehicle** | $c_o + c_{im}$ | 80 | 150 | €/h |
| **Weight access time** | $w_a$ | 2.2 | 2.2 | |
| **Weight waiting time** | $w_w$ | 1.5 | 1.5 | |
| **Weight egress time** | $w_e$ | 1.1 | 1.1 | |
| **Value of time travellers** | $c_t$ | 4.90 | 4.90 | €/h |
| **Travel demand per square kilometre (fixed)** | $P$ | 100 | 150 | Pas/km$^2$ |
| **Coefficient public transport** (logit-mode-choice) | $\alpha$ | 0.03 | 0.03 | min$^{-1}$ |
| **Coefficient private car** (logit-mode-choice) | $\alpha_m$ | 0.08 | 0.08 | min$^{-1}$ |
| **Average speed private car** (logit-mode-choice) | | 4.2 | 4.2 | m/s |
| **Parking penalty** (logit-mode-choice) | | 300 | 300 | s |

# Appendix F
## Derivation of the cost factors for public transport modes

The derivation of the cost factors for public transport modes consists of two components:

- Infrastructure costs for rail-bound public transport;
- Operational costs per public transport technique.

The investment and maintenance costs are based on data of the Ministry of Transport, Public Works and Watermanagement (1996). The investment costs are amortised over a period of 30 years using a discount rate of 4%, yielding an annual payment of 5.8%. The yearly maintenance costs are estimated as 3% of the investment costs. The resulting annual costs per kilometre are then transformed into costs per kilometre per day and given the average number of vehicles per day into costs per vehicle per kilometre per day. Multiplication with the average speed per vehicle type yields the costs per vehicle per hour.

|  | Bus lane | Tram | Metro elevated | Metro tunnel | Train low | Train high |
|---|---|---|---|---|---|---|
| **Investment costs** (price level 2,000) [m€/km] | 1.25 | 2.61 | 15.03 | 40.08 | 5.51 | 8.22 |
| **Annual payment factor** [%] | 5.8 | 5.8% | 5.8% | 5.8% | 5.8% | 5.8% |
| **Maintenance factor** [%] | 3.0 | 3.0% | 3.0% | 3.0% | 3.0% | 3.0% |
| **Total annual payments** [%] | 8.8 | 8.8% | 8.8% | 8.8% | 8.8% | 8.8% |
| **Annual costs** [m€/km] | 0.11 | 0.23 | 1.32 | 3.53 | 0.48 | 0.72 |
| **Daily costs** [€/km] | 302 | 629 | 3.624 | 9.663 | 1.329 | 1.981 |
| **Vehicles per day** | 324 | 486 | 342 | 342 | 144 | 144 |
| **Daily costs per vehicle** [€/km] | 0.93 | 1.29 | 10.60 | 28.26 | 9.23 | 13.76 |
| **Average speed** [km/h] | 35 | 19 | 35 | 35 | 70 | 70 |
| **Costs per vehicle per hour** [€/h] | 33 | 24 | 371 | 989 | 646 | 963 |

The main input for the operational costs per vehicle hour are the costs per trip offered per vehicle type as derived by Van Goeverden & Schoemaker (2000). These costs include personnel costs, that is driving staff and other service related personnel, and vehicle costs, consisting of investment, maintenance and operating costs. Although these values were determined for 1990, it is assumed that these cost factors are still representative for the costs in 2,000, or in other words, it is assumed that inflation is compensated for by efficiency improvement. Given the capacity per vehicle unit and the average number of units per vehicle the average costs per vehicle can easily be calculated. Multiplication by the average speed per vehicle yields the operational costs per vehicle.

Finally, the infrastructure and maintenance costs should be added to the operational costs to determine the total operational costs per vehicle hour.

| | Urban bus | Tram | Metro eleva-ted | Metro tunnel | Urban train | Natio-nal Train (Rand-stad) | Regio-nal bus | Regio-nal train | Natio-nal Train (rural) |
|---|---|---|---|---|---|---|---|---|---|
| **Operating hours** | | | | | | | | | |
| **Base** | 6 | 6 | 6 | 6 | 8 | 9 | 6 | 8 | 9 |
| **Peak** | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| **Low** | 9 | 9 | 9 | 9 | 7 | 6 | 8 | 7 | 5 |
| **Frequencies** [veh/h] | | | | | | | | | |
| **Base** | 4 | 10 | 10 | 10 | 4 | 3.5 | 2 | 1.8 | 1.9 |
| **Peak** | 7.5 | 14 | 12 | 12 | 5.5 | 3.6 | 3.6 | 2.2 | 2.2 |
| **Low** | 2.5 | 6.5 | 8 | 8 | 3.5 | 3.3 | 1 | 1.5 | 1.5 |
| **Vehicle units** | | | | | | | | | |
| **Base** | 1 | 1 | 2.4 | 2.4 | 4 | 5 | 1 | 3 | 6 |
| **Peak** | 1 | 1 | 3.7 | 3.7 | 7.5 | 6.5 | 1 | 5.8 | 8 |
| **Low** | 1 | 1 | 1.3 | 1.3 | 2.5 | 4 | 1 | 2 | 6 |
| **Average frequency** | 3.8 | 8.9 | 9.3 | 9.3 | 4.1 | 3.5 | 1.8 | 1.8 | 1.8 |
| **Average number of vehicle units** | 1.0 | 1.0 | 2.2 | 2.2 | 4.3 | 4.9 | 1.0 | 3.3 | 6.4 |
| **Cost per trip kilometre offered** [€ct/km] | 4.99 | 4.13 | 2.41 | 2.41 | 2.63 | 2.13 | 2.31 | 2.63 | 1.95 |
| **Capacity per unit** | 68 | 164 | 240 | 240 | 68 | 64 | 67 | 68 | 64 |
| **Speed** [km/h] | 23 | 18 | 35 | 35 | 65 | 93 | 38 | 83 | 98 |
| **Operational costs** [€/h] | 78 | 122 | 446 | 446 | 499 | 628 | 59 | 483 | 786 |
| **Infrastructure costs** [€/h] | - | 24 | 371 | 989 | 646 | 646 | - | 646 | 646 |
| **Percentage infrastructure costs** | 0% | 17% | 45% | 69% | 56% | 51% | 0% | 57% | 45% |
| **Total costs** [€/h] | 78 | 146 | 817 | 1,435 | 1,145 | 1,273 | 59 | 1,129 | 1,432 |
| **Rounded values** [€/h] | 80 | 150 | 800 | 1,400 | 1,100 | 1,300 | 60 | 1,100 | 1,400 |

For an urban bus using dedicated bus lanes the average costs per vehicle hour including infrastructure becomes € 110.--.

# Glossary

| Term | Explanation |
| --- | --- |
| Access node | Node used to enter a network (entry node), e.g. a railway station |
| Access time/distance | Time/distance to travel from an origin to an access node and to enter a network |
| Access density | Number of access nodes of a network per unit area |
| Base-bound mode | Mode of which the vehicle is only available at a specific location, usually the home address, and should finally be returned to that specific location |
| Carpool | Arrangement of two or more people to share the use, the costs or both of travelling in a private car for specific trips |
| Depot | Rental facility having few vehicles and no staff |
| Design speed | Maximum speed for which a network element is designed to be used comfortably |
| Egress node | Node used to leave a network (exit node), e.g. a bus-stop |
| Egress time/distance | Time/distance to leave a network and to travel from the egress node to a destination |
| Entry point | Node used to enter a network (access node) |
| Exit point | Node used to leave a network (egress node) |
| Grid network | Network of evenly spaced perpendicular lines, resulting into square cells |
| Hierarchical network | Composite network consisting of different layers (networks), which offer connections within a specific level as well as offer connections to and from the next higher-level |
| Hierarchy | Distinction in different levels in which the higher levels depend for their performance on the lower levels |
| Leg | Part of a trip in which one type of transport service is used |
| Level | Layer in a hierarchical system |
| Line | Route of a transport service defined by a string of subsequent links and nodes, and its service characteristics |
| Line density | Total line length per unit area |
| Link | Network element connecting two nodes |
| Mode | Combination of vehicle type and transport service type |
| Multilevel networks | Networks in which different network levels can be distinguished, each having its own transport function |

| Term | Explanation |
| --- | --- |
| Multimodal | Involving the use of more than one mode, that is vehicle modes or service modes |
| Multimodal mobility | Demand characteristics of multimodal transport |
| Multimodal tour | Tour in which different transport services (or modes) are used |
| Multimodal transport | General notion of multimodal travelling and multimodal transport services |
| Multimodal trip | Trip in which different transport services (or modes) are used between which a transfer is made |
| Network | System of nodes and links (and in public transport systems lines too) that describes a transportation system |
| Network density | Total link length of a network per unit area |
| Network speed | Average speed between an entry and an exit node |
| Node | Intersection of two or more links or an access/egress point |
| Rectangular | Network of evenly spaced perpendicular lines, resulting into rectangular, not necessarily square, zones |
| Scale-factor | Factor describing the systematic change of a variable or characteristic between two successive levels in a hierarchical system |
| Sector | Part of a service area of a demand responsive transport system that is served by a single tour |
| Service area | Area that is being served by a specific transport service or an access node of a transport service |
| Service network | Network that is primarily defined by service characteristics, e.g. public transport service network |
| Space accessibility | Number and spatial distribution of access and egress nodes of a network within a unit area |
| Time accessibility | Distribution of the moments per unit of time that travellers can travel on a network |
| Tour | Travel unit starting and ending at home and consisting of one or more successive trips |
| Traffic | Transportation of vehicles |
| Transport | Transportation of persons (or goods) |
| Transport service | Combination of vehicle type, network, and service attributes |
| Trip | Travel unit starting at an origin and ending at a destination |
| Unimodal transport | Transport without a transfer, that is between vehicle modes or between functionally different transport services |

# List of symbols

| Symbol | Unit | Explanation | Indices | Explanation |
|---|---|---|---|---|
| *A* | /km$^2$ | Accessibility | *s*<br>*t* | Space<br>Time |
| *B* | km$^2$ | Service area demand responsive transport system (DRT) | | |
| *C* | €<br>€/km$^2$ | Costs | *n*<br>*t*<br>*im*<br>*o* | Network<br>Travel<br>Investment and maintenance<br>Operating |
| *CS* | €/km$^2$ | Consumer surplus | | |
| *D* | /km$^2$<br>km/km$^2$ | Density | *a*<br>*c*<br>*r*<br>*s*<br>*l*<br>*n* | Access<br>Crossing<br>Road<br>Stop<br>Line<br>Network |
| *F* | /h | Frequency | | |
| *L* | m<br>km | Distance | *a*<br>*c*<br>*c1*<br>*c2*<br>*d*<br>*k*<br>*l*<br>*p*<br>*r*<br>*s*<br>*t*<br>*tr*<br>*2* | Access<br>Centre<br>Centre 1<br>Centre 2<br>Distance criterion<br>Distance class<br>Line length<br>Population<br>Radius<br>Stop *s*<br>Trip<br>Length DRT tour<br>Access system 2 |
| *M* | | Number of modes | | |
| *N* | | Network | *t* | Network type |
| *O* | | Objective | | |

| Symbol | Unit | Explanation | Indices | Explanation |
|---|---|---|---|---|
| $P$ | /km$^2$ | Patronage | 0 | Total demand per km$^2$ |
| | | | 1 | Demand system 1 |
| | | | 2 | Demand system 2 |
| | | | 12 | Demand systems 1 and 2 |
| $PS$ | €<br>€/km$^2$ | Producer surplus | | |
| $R$ | €<br>€/km$^2$ | Revenue | s<br>t | Subsidy<br>Fares |
| $S$ | m<br>km | Spacing | a<br>l<br>r<br>s | Access<br>Line<br>Road<br>Stop |
| $SW$ | €<br>€/km$^2$ | Social welfare | | |
| $T$ | min | Time | a<br>af<br>aw<br>b<br>c<br>e<br>i<br>m<br>n<br>pb<br><br>s<br>t<br>w<br>z | Access<br>Access feeder line<br>Access walking<br>Buffer time<br>Travel time to city centre<br>Egress<br>In-vehicle<br>Mode<br>Network<br>Penalty cycling as access mode<br>Stop loss<br>Travel<br>Wait<br>Demand for public transport vanishes |
| $\overline{T}$ | min | Average time | | |
| $\widetilde{T}$ | min | Resulting time using the logsum | | |
| $U$ | | Utility | b<br>w | Bicycle<br>Walking |
| $\overline{U}$ | | Average utility | | |

| Symbol | Unit | Explanation | Indices | Explanation |
|---|---|---|---|---|
| $\tilde{U}$ | | Resulting utility using the logsum | | |
| $W$ | km | Width of corridor | | |
| $V$ | km/h | Speed | $a$ | Access |
| | m/s | | $b$ | Bicycle |
| | | | $w$ | Walking |
| $X$ | | Set links lower-level network | | |
| $Y$ | | Set links higher-level network | | |
| $Z$ | | Number of zones | | |
| $b$ | | Factor for length travelling salesman tour DRT-system | | |
| $c$ | € | Cost factor | $a$ | Annual payment |
| | €/h | | $da$ | Access density |
| | €/km$^2$ | | $dc$ | Crossing density |
| | €/km$^2$/h | | $dr$ | Road density |
| | | | $im$ | Investment and maintenance |
| | | | $m$ | Maintenance |
| | | | $o$ | Operating |
| | | | $t$ | Travel time |
| $f$ | | Factor | $a$ | Access |
| | | | $d$ | Detour |
| | | | $w$ | Waiting |
| $i$ | | Network level | | |
| $j$ | | Index | | |
| $k$ | | Index | | |
| $m$ | | Mode | | |
| $n$ | | Number | $a$ | Alternatives |
| | | | $b$ | Sectors |
| | | | $m$ | Modes |
| | | | $s$ | Stops |
| | | | $t$ | Time periods |
| $p$ | % | Demand factor | $cw$ | Captive to walking |

| Symbol | Unit | Explanation | Indices | Explanation |
|--------|------|-------------|---------|-------------|
| $r$ | € | Revenue factor | $t$ | Fare per traveller |
|  | €/km |  | $tk$ | Fare per kilometre |
| $s$ | € | Subsidy factor | $t$ | Subsidy per traveller |
|  | €/km |  | $tk$ | Subsidy per kilometre |
| $sf$ |  | Scale-factor | $d$ | Demand |
|  |  |  | $l$ | Line spacing |
|  |  |  | $r$ | Road spacing |
|  |  |  | $s$ | Stop spacing |
|  |  |  | $sl$ | Stop spacing on line spacing |
|  |  |  | $v$ | Speed |
| $w$ |  | Weight | $a$ | Access |
|  |  |  | $e$ | Egress |
|  |  |  | $i$ | In-vehicle |
|  |  |  | $t$ | Transfer |
|  |  |  | $w$ | Wait |
| $x$ |  | Link lower-level network |  |  |
| $f()$ |  | Function | $a$ | Access |
|  |  |  | $c$ | Total costs |
|  |  |  | $cs$ | Consumer surplus |
|  |  |  | $ct$ | Traveller costs |
|  |  |  | $l$ | Trip length distributions |
|  |  |  | $n$ | Network |
|  |  |  | $p$ | Patronage |
|  |  |  | $ps$ | Producer surplus |
|  |  |  | $r$ | Revenues |
|  |  |  | $sw$ | Social welfare |
|  |  |  | $t$ | Travel time |
| $g()$ |  | Function traveller's objective |  |  |
| $h()$ |  | Function investor or operator's objective |  |  |
| $P()$ |  | Demand function |  |  |
| $\alpha$ | /min | Demand sensitivity | $b$ | Bicycle |
|  |  |  | $m$ | Mode |
|  |  |  | $w$ | Walking |

| Symbol | Unit | Explanation | Indices | Explanation |
|--------|------|-------------|---------|-------------|
| $\beta$ | | Mode specific constant cycling | | |
| $\gamma$ | | Parameter trip length distribution functions | | |
| $\kappa$ | s/m | Aggregate access parameter | | |
| $\lambda$ | s | Aggregate stop loss parameter | | |
| $\mu$ | | Scaling parameter | | |
| $\rho$ | | Aggregate costs parameter | | |
| $\tau$ | | Aggregate time accessibility parameter | | |

# SUMMARY

**Design of multimodal transport networks, A hierarchical approach**

**Rob van Nes**

*Introduction*

Multimodal transport is seen as an interesting approach to solve today's mobility problems such as recurrent congestion, deteriorating accessibility, and negative environmental impact. Combining private and public transport in a multimodal transport system offers opportunities to capitalise on the strengths of the various systems while avoiding their weaknesses, and might therefore be an interesting alternative to the traditional strictly dichotomous choice between private car or public transport.

Multimodal transport involves many issues, such as information systems, financial systems, operational control of transport services, and organisation of services. A crucial prerequisite for multimodal transport, however, is the existence of a multimodal transport network. The question then is what the main characteristics of a multimodal transport network are. How are efficient multimodal networks structured, and what does that imply for transport network planning? This thesis provides answers to these questions. The main research question thus is: what is the influence of multimodal transport on transport network design?

This summary is structured as follows. First, a brief description of multimodal transport is given, followed by an assessment made of multimodal travel today and its future potential. Next the multimodal transport network design problem is discussed. This discussion results in two choices for the analyses in this thesis: the notion of hierarchical transport networks and the use of analytical transport network design models. The actual analysis starts with a general discussion on hierarchy in transport network and spatial structures, followed by detailed analyses of the emergence of hierarchical network structures for private transport networks and especially for line-bound public transport networks. Finally, the implications of these analyses for the multimodal transport

network design problem are discussed. This discussion also includes an analysis of the potential of demand oriented transport systems, such as rental services and demand responsive transport systems.

*Multimodal transport*

In this context multimodal travel means using two or more modes for making a trip, that is different vehicle modes or functionally different service modes. Typical examples of multimodal trips are: using the private car to travel to a city while transferring to an urban public transport service to access the city centre, or, making a long distance trip by train while using private modes such as car or bicycle, or public transport services such as bus to access the train station. Multimodal transport thus implies an intermodal transfer between different modes or services. Trips in which only a single vehicle mode is used, or which have only transfers within a specific public transport service network, such as between urban buses, are considered to be unimodal trips. The main distinction between the two types of transfers is that intermodal transfers, and thus multimodal transport, deal with different networks, different transport modes or transport service types, which are designed and operated by different actors such as authorities, investors and operators.

Using data of the Dutch National Travel Survey, an extensive analysis is made of the characteristics of multimodal mobility today. It is found that the share of multimodal mobility today represents 3% of all trips. However, this does not imply that multimodal transport is unimportant. Indeed multimodal travel does have substantial shares in certain market segments:

- 15 % of Dutch trips longer than 30 kilometres are multimodal;

- 20% of trips to and from the four main cities in the Netherlands are multimodal.

Discriminant analysis shows that the main factors determining multimodal transport usage are trip length, destination area type, and trip purpose. Long distance trips to heavily urbanised areas having trip purpose work and education are most likely to be multimodal trips. The main mode for nearly 60% of multimodal trips is train, or looked at in another way, 80% of all train trips are multimodal trips. Private modes are the main mode for 17% of multimodal trips. Multimodal mobility is thus an important phenomenon for long distance travel and for access to city centres.

For the assessment of the future potential of multimodal transport two demand estimations are made: one based on trip purpose and the other on trip type, the results of which however are similar. The medium scenarios yield an increase of the multimodal travel share of about 25%, leading to a multimodal share of 3.6% of all trips. The maximum scenarios result in an increase of 70%, yielding a modal share of 5.0% of all trips. This assessment shows that multimodal transport will remain only a small fraction of all trips even though an increase of 25%, and certainly an increase of 70%, will require substantial increases in the capacity of public transport services. If it is assumed that the increase of 70% in multimodal transport is fully accommodated by train services, train trips between 10 and 30 kilometres would triple, while trips between 30 and 100 kilometres would double. These values are obviously too optimistic, but they show that

an increase in multimodal transport can have substantial impacts on the required public transport service capacities. Multimodal transport thus has a strong relationship with line-bound public transport services.

*Transport network design problem*

The transport network design problem is seen as a Stackelberg game in which an investor, or operator, makes decisions on the transport network structure, while having full knowledge of its impact on the actions of the other party involved: the traveller. The objective of the investor or operator might be maximising profit, maximising social welfare or minimising total costs, while the travellers' objective is usually minimising generalised travel costs. This perspective of the transport network design problem leads to bi-level formulation in which the upper problem is given by the investor's or operator's perspective, and the lower problem is determined by traveller behaviour.

For the analysis of the consequences of multimodal travel on transport network design two important choices were made.

First, a choice is made for adopting the concept of hierarchical network structures. Multimodal transport nearly always implies that a particular mode or transport service is used for covering the main distance, while other modes are used to access and/or egress from the main mode used. Therefore, the concept of access and egress modes versus main modes suggests a hierarchical view of transport systems, and thus of transport networks. Hierarchy in transport networks implies that different network levels are distinguished, each suited for covering specific distances and for offering access to higher network levels. The highest network level, usually a coarse network having high speeds and limited accessibility. The lowest network level serves short distance trips and provides access to higher-level networks. This network level thus has high network densities, slow speeds, and high space accessibility. Hierarchical network structures can easily be distinguished in transport networks today, but they seem to be considered a fact of life, instead of an important decision variable in transport network design. In fact most studies on transport network design deal only with unimodal single-level networks. For multimodal transport, however, multimodal multilevel transport networks are required. In the analysis a stepwise procedure is used, in which first unimodal multilevel transport network are analysed, and second the consequences of multimodal multilevel transport networks are considered.

Second, a choice is made for developing and using analytical transport network design models. Basically, there are two methods that can be distinguished to describe transport networks, and thus to describe the transport network design problem. The first method focuses on abstract network types while using aggregate network characteristics such as average stop spacing, line spacing, road spacing, frequency and speed. This method has the advantage that analytical models can be derived specifying the main relationships for these network characteristics. On the other hand, however, such analyses a clearly a simplification of reality. Real networks have a nearly unlimited variety in their characteristics compared to the abstract network types, while many assumptions used to formulate analytical network design models are based on strongly simplifying assumptions, especially with respect to travel demand patterns. In contrast, the second

method to specify transport networks accounts for all the details that can be found in transport networks today, such as local constraints and capacity constraints, as well as the varying demand pattern over space and time. This method is therefore especially suited to analyse specific problems in transportation planning. However, since these specific models allow all kinds of details to be included in the analysis, they do not easily provide general knowledge on the main drivers of transport network structure. Analytical models based on aggregate network characteristics are considered more suited for that purpose.

This thesis thus uses analytical network design models to analyse the impact of multimodality on transport network design. The analytical network design models developed in this thesis consistently use the same framework for both types of transport networks, that is the economically based objectives of minimising total costs or maximising social welfare, while random utility theory is used to describe traveller behaviour. The design models consist of level of service models, demand models and supply models. Innovative elements in these models are the multilevel nature of the network design problem, the incorporation of lower level travel choices such as route-choice or access mode-choice, and the possibility that different actors, such as authorities, investors or operators are involved. Furthermore, the importance of the models developed in this thesis is that in comparison with earlier studies on transport network design, they focus especially on the relationships for hierarchical network structures and that their mutual consistency allows a systematic comparison between different hierarchical network structures.

*Hierarchy*

Hierarchical network structures are a common phenomenon in nature, as can be seen in resource distribution networks in plant, animals, and humans. It has been shown for biological networks that hierarchical structures are optimal with respect to maximising metabolic capacity, that is maximising the surface where resources are exchanged, while minimising internal transport distances. The difference with transport networks, however, is that all these biological networks have one-to-many patterns, while in transport networks many-to-many patterns are dominant.

It can easily be shown that hierarchy in transport networks can also be seen as a natural phenomenon. In fact, a single-level network is highly unstable, except in hypothetical cases. If there is just a small bias favouring some routes over others, or if there is some kind of interaction between supply and demand, the development of a second network level is inevitable. The resulting multilevel network consists of a higher level network that is suited for covering long distances, while the lower network is suited for short distances and to access the higher level network.

In reality, mechanisms leading to such a hierarchy are unavoidable. From the perspective of the demand for transport, not all possible routes will be equally attractive for all travellers. Travellers will thus favour some routes over others, leading to differences in usage. On the supply side, it is logical that from an economic point of view heavily used routes are more likely to become more attractive due to all kinds of investments for additional facilities or improvements of the routes themselves. On both demand side and supply side improving efficiency appears to be a leading principle. Furthermore, not all

origins and destinations will have equal demand. Agglomeration tendencies lead to specific concentrations of demand in space, thus making routes to and from these locations more important. Technological developments make it possible to increase travel speed significantly making higher network levels more attractive, while modern decision processes stimulate the concentration of flows on a limited set of routes. The continuous interaction between demand and supply, the correlation with agglomeration processes, and the influence of technological developments and decision processes, all these mechanisms are constantly working together leading to hierarchical transport network structures: hierarchy in transport networks is a fact of life.

Hierarchy also exists in spatial structures and might stimulate the development of hierarchical network structures. In order to avoid the self-fulfilling prophecy that the spatial hierarchy that is assumed determines the hierarchical structure of transport networks, many analyses on hierarchical transport networks will be based on a uniformly distributed demand pattern. In the case where assumptions on spatial structures are needed, the classification defined by De Jong & Paasman (1998) will be used. In this classification the radius of settlements of the next higher level are always a scale-factor larger.

*Private transport networks*

Three types of analysis were used to establish relationships determining hierarchical network structures for private transport networks.

The first analysis develops a new perspective on hierarchy in private transport networks, based on the elimination of shortcuts in hierarchical transport networks. This analysis needs a minimum of assumptions on travel characteristics, while leading to a kind of self-organising principle for hierarchical network structures. For a grid network it is established that the road spacing of a higher-level network should be thrice as large as the road spacing for the lower-level network, while the average speed should be 67% higher.

The second analysis uses an economic point of view for super-positioning a higher-level network while minimising total costs. Interestingly, this analysis showed that using such an economically based approach while assuming parameter values that might be used today, also leads to the conclusion that the road spacing of a higher-level network should be thrice the road spacing of the lower-level network. It should be noted, that the shape of the objective function proves to be rather flat, implying that especially higher values for the road spacing of the higher-level network are near optimal. The flat objective function is a commonly encountered phenomenon when using analytical models for transport network design.

The third analysis is based on an evolutionary perspective on transport networks. A grid network is improved in a stepwise procedure by increasing the quality, that is the speed, of the link that would yield the largest reduction in travel times. This analysis showed, again, that a scale-factor three for road spacing is very plausible.

*Line-bound public transport service networks*

Since public transport is the main transport service for multimodal transport, single-level transport service networks are also analysed. The analytical network design model for single-level networks is later extended to describe all kinds of hierarchical public transport networks. Furthermore, the analysis of hierarchy in public transport service networks is split into two types, namely urban public transport networks and interurban public transport networks.

For line-bound public transport service networks the network design problem is newly formulated for single-level urban networks, such as bus or tram networks. The model describes an urban corridor in which parallel lines offer transport services to the city centre. The decision variables are stop spacing, line spacing and frequency. It has been found that from a social welfare point of view current public transport networks are too dense. Given the size of the cities today and the operational costs for urban public transport, an 80% larger stop spacing and for bus networks a 35% larger line spacing, is preferable. Coarser networks yield shorter travel times, except of course for very short trips, and lower operational costs. Furthermore, it is shown that many objections raised against such a change in network public transport network structure can easily be countered using basic principles of travel behaviour and economic analysis.

Just as for private transport networks, the analysis of the emergence of multilevel public transport networks starts from a minimum of assumptions, since these might influence the hierarchical structure for public transport networks. First, the analytical model for a single-level network is extended to describe hierarchical networks for a long urban corridor. Systems that are analysed are express services, trunk-feeder systems, and zone-systems. The first two systems, however, proved to have a lower performance compared to a single-level network. Only the zone-system proved to be interesting with respect to social welfare. Relaxation of the assumptions with respect to the demand pattern (many-to-many instead of many-to-one), network structure (grid network instead of a linear corridor), or demand densities, showed that multilevel network structures have only a better performance in the case of higher densities around stops of an express system. Hierarchy in urban public transport networks is thus determined by hierarchy in spatial structures.

Current travel patterns already suggest that interurban public transport strongly depends on access and egress on foot and by private modes. Public transport accounts only for 25% of all access and egress parts of trips using train services. In the main cities in the Netherlands this percentage becomes 40%. A similar analysis as for private transport networks, that is super-positioning a higher-level network, showed also that for interurban public transport networks hierarchy is not really a natural phenomenon. Only if higher demand densities are located around the stops of the higher-level network, or vice versa, a hierarchical structure is plausible.

The main mechanisms that explain why hierarchical network structures are apparently not a self-emerging phenomenon are transfers and the costs involved in providing the higher-level network. While private transport has seamless switches between network levels, public transport requires a transfer between network levels. The higher network

level should thus not only compensate for the possible detour, but also for these transfers. Second, the higher-level public transport networks usually require dedicated infrastructure to guarantee their quality. The costs of providing higher level transport service networks are thus very high.

Given the classification for hierarchical spatial structures, the conclusion that spatial hierarchy determines the hierarchy for public transport networks also implies a scale-factor of about 3 for both stop spacing and line spacing.

A new type of analysis is developed to determine whether the interaction between urban and interurban public transport networks influences the network characteristics. This extension of the multilevel network design problem incorporates the fact that different network levels generally involve different actors in designing and operating transport services. Game theory is used to account for the fact that these actors pursue their own objectives. This model provides interesting insights into the interaction between network levels. The consequences of different actors are primarily determined by the design objectives that are adopted. The objective of maximising social welfare leads to attractive transport service networks from the traveller's perspective, while the objective of profit maximisation yields coarse networks having long travel times. The main impact of the case of different authorities or operators is that frequencies become more important due to explicitly considering travellers transferring between both network levels. Furthermore, the analysis shows that financial relationships between network levels exist. Since long distance public transport is more likely to be profitable, it is in the interest of the higher-level network operators to ensure good accessibility by lower-level public transport services. This can be achieved by subsidising lower-level public transport networks, or other local transport networks, to improve the quality for travellers using those networks to access the higher-level network. An even more profitable situation occurs if other actors, such as local authorities, strive to improve the quality of their lower-level networks for their own objectives, such as in the case of an urban public transport network that is designed using the objective of maximising social welfare. In that case the interurban operator obtains high quality access to his transport services, without any financial input from his side.

*Multimodal transport networks*

Finally, the findings of the analysis of hierarchy in private transport and public transport networks are used to redefine the possible impact of multimodality on transport network design. It is concluded that the structure of both private transport and public transport service networks will not be influenced by multimodal travelling. Private transport networks have self-organising properties determining its hierarchical structure, while the share of multimodal transport is, and is likely to remain, too small to change private transport network structures. The strong relationship that is established between hierarchy in spatial structures and public transport network structures, also implies that hierarchical public transport networks are robust with respect to multimodal travelling. In contrast to private transport, however, the future potential for multimodal mobility is large enough to require substantial increases in capacity. These findings with respect to the hierarchy of both network types are used to define a classification of network levels that can be used as design guidelines.

Furthermore, this synthesis leads to three additional analyses. First, the single-level urban public transport network model is extended to account for alternative access modes to walking. Second, an assessment is made of the potential impact of rental services on multimodal transportation. Third, an analysis is made of the potential of demand responsive transport systems on multimodal trip making.

The analysis of multimodal access for urban public transport networks show that walking remains the main access mode for urban public transport networks. This pertains to single-level urban networks as well as to the lower-level network of multilevel public transport systems. The introduction of alternative access modes such as cycling is shown to have only minor consequences for the values of the optimal network characteristics. A dedicated behavioural analysis shows this is partly due to a rather high penalty for using a bicycle to access urban public transport, and partly due to the fact that access distances will always be so short that walking will remain the preferred access mode. Note that alternative access modes, such as cycling, remain essential for higher-level public transport networks although this does not influence their optimal network design.

Demand oriented transport services such as rental services and demand responsive transport services can make multimodal travelling more attractive. The costs of providing such services, however, are so high that their impact on multimodal mobility will always be small, although they might be perfectly suited for specific population segments. In the case of rental services, the costs of the vehicles themselves as well as those for parking the vehicles are relatively high, thus limiting the benefits of scale economies. It is obvious that this finding is more critical for car rental services than for bicycle rental services. For demand responsive transport the share of personnel costs is too high to achieve benefits of scale that are in line with those found in traditional public transport services. This assessment might change, however, if the operational costs could be substantially reduced, for instance by using fully automated vehicles in a demand responsive transport system, thus eliminating the costs of a paid driver.

*Intermodal transfer nodes*

The conclusion that an efficient multimodal transport system requires properly designed transport networks, reduces the multimodal network design problem to the allocation of transfer nodes. Given the strong relationship between hierarchy in spatial structures and in public transport networks, this leads to clear criteria for the location of intermodal transfer nodes:

- Within city centres, offering access to higher-level public transport networks (comparable to railway stations in cities);

- Within local centres, offering access to cities by public transport services (comparable to traditional Park&Ride facilities);

- At the edge of cities near motorways, offering access to city centres by urban public transport services, or even offering access to the motorway network using urban public transport or bicycles (comparable to Transferia).

The attractiveness of multimodal travel thus depends more on the quality of the transport services offered than on newly designed transport networks. The quality of the transfer nodes, the transport services themselves, the availability of information and all kinds of financial aspects are decisive. Stimulating multimodal mobility does not require a new grand design for the transport network, but benefits more from doing little things properly.

# SAMENVATTING

## Ontwerp van multimodale vervoernetwerken, een hiërarchische aanpak

## Rob van Nes

*Inleiding*

Multimodaal vervoer wordt vaak gezien als een interessante mogelijkheid om de huidige vervoer- en verkeersproblemen zoals congestie, bereikbaarheidsproblemen en negatieve milieueffecten, op te lossen. Het combineren van particulier vervoer en openbaar vervoer in een multimodaal vervoersysteem biedt kansen om de sterke punten van beide vervoersystemen maximaal te benutten en tegelijkertijd de nadelen van hun zwakke punten te beperken. Multimodaal vervoer is daarmee een alternatief voor de traditionele strakke scheiding tussen particulier vervoer en openbaar vervoer.

Multimodaal vervoer heeft betrekking op een groot aantal onderwerpen, zoals informatievoorziening, betalingssystemen, de operationele beheersing van vervoerdiensten en de organisatie van vervoerdiensten. Een essentieel onderdeel van een multimodaal vervoersysteem is echter het multimodale netwerk. De vraag is hoe zo'n multimodaal netwerk eruit ziet of eruit zou moeten zien. Is dat anders dan de netwerken die we nu kennen, en wat betekent dat voor het ontwerp van vervoernetwerken? Deze dissertatie geeft een antwoord op deze vragen. De centrale onderzoeksvraag is dan ook wat de invloed is van multimodaal vervoer op het netwerkontwerp.

Deze samenvatting heeft de volgende opzet. Eerst wordt een korte omschrijving gegeven van multimodaal vervoer, gevolgd door een beschrijving van multimodale verplaatsingen van vandaag en een raming van de potentie van multimodaal vervoer. Vervolgens wordt het multimodale netwerkprobleem besproken. Hierbij worden twee belangrijke keuzen voor de verdere analyse gemaakt: het gebruik van hiërarchische netwerkstructuren en het gebruik van analytische netwerkontwerpmodellen. De feitelijke analyse begint met een meer algemene discussie over hiërarchie in vervoernetwerken en in ruimtelijke structuren, gevolgd door gedetailleerde analyses van het ontstaan van hiërarchische

netwerkstructuren in particuliere vervoernetwerken en in het bijzonder in lijngebonden openbaar vervoernetwerken. Tot slot worden de implicaties van deze analyses voor het multimodale netwerkontwerpprobleem besproken. Hierbij komt tevens de potentie van vraag georiënteerde vervoersystemen zoals verhuurdiensten en vraagafhankelijke vervoersystemen aan de orde.

*Multimodaal vervoer*

Multimodaal vervoer is gedefinieerd als verplaatsingen waarbij gebruik wordt gemaakt van twee of meer vervoerwijzen, dat wil zeggen voertuigen of functioneel verschillende vervoerdiensten. Typische voorbeelden zijn het gebruik van de eigen auto om naar de stad te reizen en daar over te stappen op een metro om naar het centrum te gaan, of het gebruik van de trein voor een lange afstandsverplaatsing waarbij voor het voortransport gebruik wordt gemaakt van de fiets of de stadsbus. Multimodaal vervoer betekent dus een overstap tussen vervoerwijzen of vervoerdiensten. Verplaatsingen waarbij slechts één vervoermiddel of één type vervoerdienst wordt gebruikt zijn gedefinieerd als unimodale verplaatsingen. Verplaatsingen waarbij binnen een bepaalde vervoerdienst wordt overgestapt, bijvoorbeeld tussen twee stadsbuslijnen, zijn eveneens unimodale verplaatsingen. Een overstap in een multimodale verplaatsing onderscheidt zich doordat wordt overgestapt tussen verschillende netwerken, tussen verschillende vervoerwijzen en vervoerdiensten, die bovendien worden ontworpen en geëxploiteerd door verschillende partijen zoals overheden en vervoerbedrijven.

Het Onderzoek Verplaatsingsgedrag is gebruikt om een analyse te maken van de karakteristieken van multimodaal vervoer op dit moment. Het aandeel multimodaal vervoer is op zich beperkt tot 3% van alle verplaatsingen. Dit betekent echter niet dat multimodaal vervoer onbelangrijk is. Integendeel, multimodaal vervoer heeft een belangrijk aandeel in specifieke marktsegmenten:

- 15% van verplaatsingen langer dan 30 kilometer is multimodaal;
- 20% van de verplaatsingen van en naar de vier grote steden in Nederland is multimodaal.

Discriminant analyse laat zien dat de verplaatsingsafstand, het type bestemmingsgebied en het verplaatsingsmotief de belangrijkste invloedsfactoren voor multimodaal vervoer zijn. Lange afstandsverplaatsingen naar stedelijke gebieden met de motieven werk en studie hebben het hoogste aandeel multimodaal vervoer. De belangrijkste vervoerwijze voor meer dan 60% van alle multimodale verplaatsingen is de trein. Omgekeerd geldt dat 80% van de verplaatsingen per trein multimodale verplaatsingen zijn. Particuliere vervoerwijzen zijn de hoofdvervoerwijze voor 17% van de multimodale verplaatsingen. Multimodaal vervoer is dus een belangrijk fenomeen voor lange afstandsvervoer en de bereikbaarheid van stedelijke centra.

Er zijn twee ramingen gemaakt voor het toekomstige aandeel van multimodaal vervoer: één gebaseerd op het verplaatsingsmotief en de ander op basis van het verplaatsingstype. Beide ramingen geven gelijkwaardige resultaten. De middenscenario's laten een toename van het aandeel multimodaal vervoer van 25% zien, oftewel 3,6% van alle verplaatsingen. De maximale scenario's resulteren in een groei van 70%, of een aandeel

van 5,0% van alle verplaatsingen. Deze ramingen laten zien dat multimodaal vervoer een bescheiden aandeel in de verplaatsingsmarkt zal behouden. Anderzijds betekent een toename van 25%, en zeker een groei met 70%, wel dat een uitbreiding van de capaciteit van openbaar vervoerdiensten noodzakelijk is. Als wordt verondersteld dat de groei van 70% volledig wordt opgevangen door treindiensten, dan wordt het aantal reizigers tussen de 10 en 30 kilometer verdrievoudigd en tussen de 30 en 100 kilometer verdubbeld. Deze waarden zijn natuurlijk gebaseerd op een aantal vergaande veronderstellingen, maar laten duidelijk zien dat een toename van multimodaal vervoer grote consequenties heeft voor het openbaar vervoer. Multimodaal vervoer heeft dus een belangrijke relatie met openbaar vervoer.

*Netwerkontwerpprobleem*

Het netwerkontwerpprobleem is benaderd als een Stackelberg-spel waarin een investeerder of exploitant een netwerk ontwerpt met volledige kennis over het gedrag van de gebruikers van het netwerk. De doelstelling van de ontwerper kan zijn winstmaximalisatie of welvaartsmaximalisatie terwijl de reiziger meestal reiskosten of reistijden wil minimaliseren. Deze benadering van het ontwerpprobleem leidt tot een bi-level formulering waarin het bovenste probleem het ontwerpersprobleem is met de optiek van de investeerder of de exploitant en het onderste probleem betrekking heeft op het gebruik van het netwerk, dat wil zeggen de optiek van de reiziger.

Om inzicht te krijgen in de consequenties van multimodaal reizen op het ontwerp van vervoersnetwerken zijn twee belangrijke keuzes gemaakt.

Op de eerste plaats is gekozen voor het gebruik van het concept van hiërarchische netwerken. Multimodaal vervoer impliceert meestal dat een specifieke vervoerwijze, hetzij particulier vervoer of openbaar vervoer, primair wordt gebruikt voor het overbruggen van de verplaatsingsafstand. De andere vervoerwijzen worden dan gebruikt als voor- of natransport. Het concept van voor- en natransport en hoofdvervoerwijze suggereert een hiërarchische benadering voor een multimodaal vervoernetwerk. In een hiërarchisch netwerk kunnen verschillende deelnetwerken worden onderscheiden, die elk geschikt zijn voor het overbruggen van specifieke afstanden en tevens toegang geven tot hogere orde netwerkniveaus. Het hoogste netwerk niveau is geschikt voor lange afstandsverplaatsingen en is meestal een grofmazig netwerk met weinig toegangspunten en met hoge reissnelheden. Het laagste netwerkniveau daarentegen is geschikt voor korte afstanden en heeft een fijnmazige netwerkstructuur met veel toegangspunten maar met een lage snelheid. Hiërarchische netwerkstructuren zijn vaak duidelijk herkenbaar maar worden meestal als een natuurlijk gegeven beschouwd en niet als een ontwerpvariabele van het netwerkontwerpprobleem. In de praktijk hebben veel studies over netwerkontwerp het over enkellaags netwerken, bijvoorbeeld een stedelijk, een regionaal of een nationaal netwerk. Multimodaal vervoer betekent echter per definitie een meerlaagse netwerkstructuur. In de analyse in deze dissertatie is een stapsgewijze aanpak gebruikt. Eerst is het enkellaagse netwerkontwerpprobleem uitgebreid naar meerlaagse netwerken, zowel voor netwerken voor particulier vervoer als voor openbaar vervoernetwerken. Daarna zijn de consequenties van de uitbreiding naar multimodal meerlaagse netwerken onderzocht.

Op de tweede plaats is gekozen voor het gebruik van analytische netwerkontwerpmodellen. Er kunnen twee methoden worden onderscheiden om vervoernetwerken, en dus ook het netwerkontwerpprobleem, te beschrijven. De eerste methode gaat uit van abstracte netwerktypen en gebruikt globale karakteristieken zoals halteafstand, lijnafstand, maaswijdte (afstand tussen parallelle wegen), frequenties en snelheden. Deze methode heeft het voordeel dat analytische modellen kunnen worden opgesteld die de belangrijkste relaties tussen de netwerkkarakteristieken beschrijven. Aan de andere kant vraagt deze analytische benadering belangrijke vereenvoudigingen van de werkelijkheid. Bestaande netwerken hebben in vergelijking met de abstracte netwerktypen een bijna ongelimiteerde variatie in kenmerken, terwijl voor het formuleren van een analytisch netwerkontwerpmodel zijn sterk vereenvoudigende aannames noodzakelijk, met name voor de beschrijving van de vervoervraag. De tweede methode om netwerken te beschrijven daarentegen kan rekening houden met allerlei aspecten uit de praktijk zoals lokale randvoorwaarden, capaciteiten en variaties van de vervoervraag in ruimte en tijd. Deze methode is dan ook bij uitstek geschikt voor de analyse van specifieke vervoerproblemen in de praktijk. Aangezien deze tweede methode erg gedetailleerd is, is hij minder geschikt voor het verkrijgen van algemene kennis over de manier waarop vervoernetwerken in elkaar zitten. Analytische modellen gebaseerd op meer globale netwerkkenmerken zijn hiervoor beter geschikt.

In deze dissertatie worden analytische netwerkontwerpmodellen gebruikt om de invloed van multimodaal vervoer op het netwerkontwerpprobleem te analyseren. Deze analytische modellen voor netwerkontwerp gebruiken consistent een zelfde raamwerk. Dit raamwerk is gebaseerd op economische doelstellingen als welvaartsmaximalisatie en kostenminimalisatie, terwijl voor de beschrijving van het reizigergedrag gebruik wordt gemaakt van de random-nuts theorie. Het netwerkontwerpmodel bestaat uit deelmodellen voor de aanbodskwaliteit, het reizigersgedrag, en het aanbod van vervoerdiensten zelf. Innovatieve onderdelen in deze modellen zijn de uitbreiding naar meerlaagse netwerken, het verwerken van onderliggende keuzeprocessen zoals routekeuze en vervoerwijzekeuze voor voortransport, en het expliciet rekening houden met de mogelijkheid dat verschillende partijen verantwoordelijk zijn voor het ontwerp en de exploitatie van verschillende netwerkniveaus in een meerlaags vervoersysteem. Een andere belangrijke eigenschap van de ontwikkelde modellen is dat ze expliciet gericht zijn op het fenomeen hiërarchische meerlaagse netwerken, en dat ze het, door het consistente gebruik van hetzelfde analyse raamwerk, mogelijk maken verschillende hiërarchische netwerkvormen systematisch te vergelijken.

*Hiërarchie*

In de natuur zijn hiërarchische netwerken gebruikelijk, zoals in distributienetwerken in planten en dieren. Onderzoek heeft laten zien dat deze hiërarchische netwerken optimaal zijn voor het maximaliseren van hun metabolische capaciteit, dat wil zeggen het oppervlak waar stoffen worden uitgewisseld, terwijl tegelijkertijd de interne transportafstanden worden geminimaliseerd. Deze biologische netwerken hebben echter een duidelijk verschil met transportnetwerken, omdat deze netwerken alleen betrekking hebben op distributie of collectie vanuit of naar een centraal punt, terwijl transportnetwerken in het personenvervoer een veel meer diffuus patroon van verplaatsingen verzorgen.

Het kan echter eenvoudig aannemelijk worden gemaakt dat hiërarchie in vervoernetwerken ook een natuurlijk fenomeen is. Een enkellaags netwerk blijkt zeer instabiel, behalve in zeer hypothetische omstandigheden. Als er maar een kleine aanleiding is waardoor sommige routes interessanter zijn dan andere, of als er een wisselwerking is tussen vraag en aanbod, is de ontwikkeling van een tweede netwerkniveau onvermijdelijk. Het resulterende meerlaagse netwerk bestaat uit een hoger orde netwerk voor langere afstanden en een lager orde netwerk voor korte afstanden en toegang tot het hoger orde netwerk.

In werkelijkheid zijn de mechanismen die tot zo'n hiërarchie leiden onontkoombaar. Vanuit de optiek van de gebruiker, zullen nooit alle routes even aantrekkelijk zijn. Reizigers zullen dus sommige routes verkiezen boven andere. Aan de aanbodkant is het economisch gezien voor de hand liggend dat drukkere routes interessanter zijn voor investeringen in extra faciliteiten of het aantrekkelijker maken van de routes zelf. Het algemene principe dat aan deze processen ten grondslag ligt is efficiency verhoging. Bovendien zullen niet alle herkomsten en bestemmingen een gelijke vervoervraag hebben. Agglomeratieprocessen leiden tot concentraties van de vervoervraag in de ruimte, en maken dus de routes van en naar deze locaties belangrijker. Technologische ontwikkelingen maken het mogelijk de snelheden te verhogen, waardoor het gebruik van hogere orde netwerken aantrekkelijker wordt. De huidige beslissingsprocessen leggen in toenemende mate het accent op bundeling zodat allerlei negatieve milieueffecten worden geminimaliseerd. De continue interactie tussen vraag en aanbod, versterkt door agglomeratieprocessen, technologische ontwikkelingen en besluitvormingsprocessen, leiden eenduidig tot hiërarchische netwerkstructuren: hiërarchische vervoernetwerken zijn een gegeven.

Ook in de ruimtelijke structuur bestaat een hiërarchie van nederzettingen. Aangezien deze hiërarchie mede sturend kan zijn voor de ontwikkeling van hiërarchische netwerkstructuren, wordt in de analyses vaak gebruik gemaakt van een uniforme verdeling van de vervoervraag. Indien het noodzakelijk is om aannamen te doen over nederzettinggrootte en ruimtelijke spreidingspatronen, wordt gebruik gemaakt van de indeling van De Jong & Paasman (1998). Bij deze indeling is de straal van een nederzetting van een volgend schaalniveau altijd een schaalfactor 3 groter.

*Particulier vervoernetwerken*

Drie typen analyses zijn gebruikt om te analyseren welke netwerkeigenschappen bij particuliere vervoernetwerken, zoals bijvoorbeeld wegennetwerken, tot hiërarchie leiden.

De eerste analyse geeft een nieuw perspectief op hiërarchie in wegennetwerken en is gebaseerd op het elimineren van kortsluitroutes. Deze aanpak vraagt een minimum aan veronderstellingen omtrent vraagpatronen en reizigersgedrag en leidt tot een soort zelforganiserend principe. Voor een grid netwerk geeft deze benadering een schaalfactor 3 voor de maaswijdte, dat wil zeggen dat de maaswijdte een volgend netwerkniveau steeds een factor 3 groter is, terwijl voor de snelheid een schaalfactor van 1,67 van toepassing is.

De tweede analyse is gebaseerd op de economische benadering van minimalisatie van de totale kosten in het geval van superpositie van een hoger orde netwerk. Deze analyse laat zien dat zo'n economische benadering, met reële parameterwaarden, ook leidt tot een optimale schaalfactor voor de maaswijdte die gelijk is aan 3. Hierbij moet worden opgemerkt dat de doelstellingsfunctie relatief vlak is, waardoor er veel bijna optimale oplossingen zijn, met name voor hogere waarden voor deze schaalfactor. De platte doelstellingsfuncties zijn overigens een gebruikelijk fenomeen bij analytische netwerkontwerpmodellen.

De derde analyse gaat uit van een meer evolutionaire ontwikkeling van een wegennet. Een rasternetwerk is stapsgewijs verbeterd door steeds de kwaliteit, oftewel snelheid, van die link te verhogen die tot de grootste reductie van de totale reistijd zou leiden. Ook deze analyse laat zien dat een schaalfactor 3 voor de maaswijdte erg plausibel is.

*Lijngebonden openbaar vervoernetwerken*

Aangezien openbaar vervoer de meest sterke relatie heeft met multimodaal vervoer, is ook het enkellaags netwerkontwerpprobleem in beschouwing genomen. Het bijbehorende analytisch netwerkontwerpmodel is in tweede instantie uitgebreid naar meerlaagse netwerken. De analyse van hiërarchie in lijngebonden openbaar vervoernetwerken is verder in twee delen verdeeld, namelijk stedelijke openbaar vervoernetwerken en interlokale openbaar vervoernetwerken.

Voor enkellaagse stedelijke openbaar vervoernetwerken is een nieuw netwerkontwerpmodel geformuleerd dat geschikt is voor bijvoorbeeld bus- en tramnetwerken. Dit model beschrijft een stedelijke corridor waarin parallelle lijnen het vervoer naar het centrum verzorgen. De ontwerpvariabelen zijn de halteafstand, de lijnafstand en de frequentie. Toepassing van het model met reële parameterwaarden laat zien dat vanuit de optiek van welvaartsmaximalisatie de huidige stedelijke openbaar vervoernetwerken te fijnmazig zijn. Gegeven de grootte van de steden nu en de huidige exploitatiekosten zouden de halteafstanden circa 80% groter moeten zijn en bij busnetwerken zou ook de lijnafstand 35% groter moeten zijn. Deze grovere netwerken leiden tot kortere reistijden, behalve voor korte reisafstanden natuurlijk, en met name tot lagere exploitatiekosten. Bovendien laat deze analyse zien dat veel bezwaren tegen het streven naar grovere stedelijke openbaar vervoernetwerken kunnen worden tegengesproken met basisprincipes van reizigersgedrag en economische analyse.

Net als bij particuliere netwerken is de analyse naar verklaringen voor het ontstaan van hiërarchische netwerken gebaseerd op een minimum aan veronderstellingen. Is het analytische model voor een stedelijke corridor uitgebreid voor hiërarchische netwerkstructuren in een lange corridor met een uniform verdeelde vervoervraag. Onderzochte netwerkstructuren zijn sneldiensten, stamlijnen met aparte aanvoerlijnen, en een zonesysteem. De eerste twee netwerkstructuren leiden echter tot een lagere welvaart dan een enkellaags netwerk. Alleen een zonesysteem leidt tot een netto hogere welvaart. Andere aannamen voor het vraagpatroon (diffuus in plaats van centrum georiënteerd), netwerkstructuur (grid netwerk in plaats van een lineair netwerk), en concentraties van de vervoervraag, laten zien dat alleen hogere concentraties van de vervoervraag rond de

halten van de sneldienst netto tot een hogere welvaart leiden. Hiërarchie in stedelijk openbaar vervoer wordt dus bepaald door de hiërarchie in de stedelijke structuur.

Kenmerken van verplaatsingen met de trein suggereren al dat interlokaal openbaar vervoer erg afhankelijk is van voor- en natransport te voet en met de fiets. Openbaar vervoer verzorgt slechts 25% van alle voor- en natransport voor de trein. In de vier grote steden komt dit percentage op zo'n 40%. Ook de analyse van het superpositioneren van een hoger orde netwerk op een interlokaal openbaar vervoernetwerk laat zien dat openbaar vervoernetwerken geen eigenschapen hebben die eenduidig tot een vaste hiërarchie leiden. Alleen indien de vervoervraag is geconcentreerd op herkomsten en bestemmingen rond de toegangspunten van het hogere orde netwerk, blijkt een hiërarchische netwerkstructuur zinvol.

De verklaring waarom lijngebonden openbaar vervoernetwerken niet uit zich zelf een mechanisme hebben dat tot hiërarchische structuren leidt, is de overstap en de kosten van hogere orde netwerken. Voor particulier vervoer is de overgang tussen netwerkniveaus naadloos, terwijl bij openbaar vervoer altijd een overstap noodzakelijk is. Het hogere orde netwerk moet hierdoor niet alleen de eventuele omweg compenseren, maar ook de noodzakelijke overstappen. Verder hebben hogere orde openbaar vervoer netwerken meestal eigen infrastructuur nodig om hun gewenste kwaliteit te kunnen garanderen. De kosten die hiermee gepaard gaan zijn erg hoog.

Gegeven de indeling voor schaalniveaus van nederzettingen, leidt de conclusie dat ruimtelijke hiërarchie bepalend is voor de hiërarchie in openbaar vervoernetwerken tot een schaalfactor 3 voor de halteafstand en de lijnafstand.

Een nieuw soort analyse is ontwikkeld om na te gaan wat de interactie is tussen een stedelijk en interlokaal openbaar vervoernetwerk. Bij deze uitbreiding van het meerlaagse netwerkontwerpprobleem is met behulp van speltheorie rekening gehouden met de mogelijkheid dat voor elk netwerk een andere partij verantwoordelijk is voor het ontwerp en de exploitatie, en dus een eigen ontwerp maakt. De belangrijkste consequenties voor de ontworpen netwerken zijn het gevolg van de gebruikte ontwerpdoelstellingen. Welvaartsmaximalisatie leidt tot netwerken met hoge dichtheden en aantrekkelijke reistijden, terwijl winstmaximalisatie tot grove en laagfrequente netwerken leidt die niet aantrekkelijk zijn voor de reiziger. Wel is het zo dat, omdat expliciet rekening wordt gehouden met reizigers die beide netwerken gebruiken, het belang van de frequentie toeneemt. Deze reizigers moeten immers overstappen tussen beide netwerken.

Een tweede consequentie van aparte netwerkontwerpen in een meerlaagse netwerkstructuur is dat financiële relaties tussen netwerken zichtbaar zijn. Aangezien lange afstandsvervoer eerder rendabel is, is het in het belang van de interlokale vervoerder om goed toegankelijke toegangspunten te hebben. Dit kan onder andere worden bereikt door lokale vervoerders te subsidiëren om een hogere vervoerkwaliteit van en naar hun toegangspunten te verzorgen. Een nog gunstigere situatie ontstaat als voor het ontwerp van het stedelijke netwerk de doelstelling van welvaartsmaximalisatie wordt gehanteerd. Dan is een goede toegankelijkheid van het interlokale vervoer verzorgd, zonder dat de interlokale vervoerder hierin hoeft te investeren.

*Multimodale vervoernetwerken*

Tenslotte zijn de resultaten van de analyses van hiërarchische netwerkstructuren in particuliere vervoernetwerken en lijngebonden openbaar vervoernetwerken gebruikt om de consequenties van multimodal vervoer voor het ontwerp van multimodale vervoernetwerken opnieuw te formuleren. De conclusie is dat de structuur van zowel particuliere vervoernetwerken als van lijngebonden openbaar vervoernetwerken niet door multimodaal vervoer worden beïnvloed. Particuliere vervoernetwerken hebben zelfstructurerende eigenschappen die de hiërarchische netwerkstructuur bepalen. Verder is het aandeel van multimodaal vervoer te klein, en zal dat waarschijnlijk ook blijven, om een verandering in netwerkstructuren te rechtvaardigen. De sterke relatie tussen de hiërarchie van nederzettingen en de netwerkhiërarchie van openbaar vervoernetwerken, betekent ook dat openbaar vervoernetwerken robuust zijn ten aanzien van multimodaal vervoer. In dit geval is echter de mogelijke groei van multimodaal vervoer zo groot, dat wel een vergroting van de capaciteit vereist is. De bevindingen over de hiërarchie voor beide netwerktypen zijn gebruikt voor een indeling in schaalniveaus voor vervoernetwerken die gebruikt kan worden als ontwerprichtlijn.

Verder leidt deze synthese tot drie aanvullende analyses. Eerst is het enkellaags ontwerpmodel voor stedelijke openbaar vervoernetwerken uitgebreid met de vervoerwijzekeuze voor het voortransport. Hiermee kan worden nagegaan in welke mate andere vervoerwijzen in het voortransport het netwerkontwerp kunnen beïnvloeden. Ten tweede is een analyse gemaakt van de mogelijke rol in multimodaal vervoer van vervoermiddelverhuur. Ten derde is een analyse gemaakt van de potentie van vraagafhankelijke vervoersystemen in een multimodaal vervoersysteem.

De analyse van multimodaal voortransport voor stedelijke openbaar vervoernetwerken laat zien dat lopen altijd de belangrijkste voortransport vervoerwijze blijft. Dit geldt zowel voor enkellaagse openbaar vervoer netwerken als voor de laagste netwerkniveaus van meerlaagse openbaar vervoernetwerken in steden. Het introduceren van alternatieve voortransport vervoerwijzen heeft slechts een beperkte invloed op de optimale waarden van de ontwerpvariabelen van een stedelijk openbaar vervoernetwerk. Een analyse van reizigersgedrag laat zien dat dit zowel wordt veroorzaakt door een hoge penalty voor het gebruik van andere vervoerwijzen in het voortransport als doordat de voortransport afstanden zo kort blijven dat lopen vaak de meest voor de hand liggende vervoerwijze is. Overigens zijn andere voortransport vervoerwijzen zoals fietsen wel essentieel voor de bereikbaarheid van hogere orde openbaar vervoernetwerken, maar ze hebben geen invloed op het optimale netwerkontwerp.

Vraaggeoriënteerde vervoerdiensten zoals verhuurdiensten en vraagafhankelijke vervoersystemen kunnen multimodaal vervoer aantrekkelijker maken. De exploitatiekosten van dergelijke vervoerdiensten zijn echter zo hoog, dat hun aandeel in multimodaal vervoer altijd klein zal blijven. Dat neemt niet weg dat deze vervoersystemen voor specifieke reiziger segmenten een belangrijke rol kunnen vervullen. Bij verhuurdiensten zijn de kosten van de voertuigen en de stalling relatief hoog, en zijn de mogelijkheden om schaalvoordelen te behalen beperkt. Dit geldt natuurlijk meer voor autoverhuur dan voor de verhuur van fietsen. Bij vraagafhankelijke vervoersystemen is het aandeel van de personeelskosten te groot om een zelfde

schaalvoordeel te behalen zoals bij traditioneel openbaar vervoer. Indien dergelijke systemen volledig automatisch zouden kunnen functioneren, kan dit natuurlijk veranderen.

*Intermodale overstappunten*

De conclusie dat een efficiënt multimodaal vervoersysteem vereist dat particuliere vervoernetwerken en openbaar vervoernetwerken op een juiste manier zijn ontworpen, betekent dat het multimodale netwerkontwerpprobleem is gereduceerd tot de locatie van intermodale overstappunten. Op basis van de sterke relatie tussen hiërarchie in nederzettingen en de hiërarchie in openbaar vervoernetwerken kunnen voor de locatie van intermodale overstappunten duidelijke criteria worden afgeleid:

- Binnen stedelijke centra, toegangspunten voor hogere orde openbaar vervoernetwerken (vergelijk de treinstations in de grote steden);

- In centra van plaatsen, toegangspunten voor openbaar vervoer naar de steden toe (vergelijk de traditionele P&R faciliteiten);

- Aan de rand van steden bij snelwegen, toegangspunten voor stedelijk openbaar vervoer naar de stadscentra (vergelijk de transferia), of wellicht toegangspunten naar het autosnelwegennet met fiets of stedelijk openbaar vervoer als voortransport.

De attractiviteit van multimodaal vervoer is dus meer afhankelijk van de kwaliteit van de aangeboden vervoerdiensten dan van een nieuw multimodaal netwerkontwerp. De kwaliteit van overstappunten, de vervoerdiensten zelf, de beschikbaarheid van informatie en allerlei financiële aspecten zijn bepalend voor het succes van multimodaal vervoer. Het stimuleren van multimodaal vervoer vraagt niet om een grote maatregel als een nieuw netwerkontwerp, maar heeft meer baat bij het goed doen van eenvoudige dingen.

# CURRICULUM VITAE

Rob van Nes was born in 1958 in Voorburg, the Netherlands. He received his master's degree (with distinction) in Civil Engineering from the Delft University of Technology in 1982.

In 1984 Rob's career started at the Transportation Research Laboratory of Delft University of Technology. The first year he was working at traffic flow analysis for two-lane rural roads. After that he developed an interactive optimisation model for the design of public transport networks, a research supported by the Netherlands Technology Foundation (STW).

In 1988 Rob joined AGV consultants in Nieuwegein. Here, he was mainly involved in the development and application of various transportation models in all kinds of transport studies. In 1995 Rob switched to Arends & Samhoud Verkeers- en Vervoerkundige Diensten where he worked as a senior consultant.

His interest in the fundamental aspects of transportation science led him in 1997 to Delft University of Technology again, were he is working as assistant professor with the Transportation Planning and Traffic Engineering Section of the Faculty of Civil Engineering and Geosciences. He gives courses on transportation systems and traffic networks. During his research on multimodal transport network design, he has presented various papers in the Netherlands and at international conferences, and published articles in national and international journals.

**TRAIL Thesis Series**

A series of The Netherlands TRAIL Research School for theses on transport, infrastructure and logistics.

Nat, C.G.J.M., van der, *A Knowledge-based Concept Exploration Model for Submarine Design*, T99/1, March 1999, TRAIL Thesis Series, Delft University Press, The Netherlands

Westrenen, F.C., van, *The Maritime Pilot at Work*: E*valuation and Use of a Time-to-boundary Model of Mental Workload in Human-machine Systems*, T99/2, May 1999, TRAIL Thesis Series, Eburon, The Netherlands

Veenstra, A.W., *Quantitative Analysis of Shipping Markets,* T99/3, April 1999, TRAIL Thesis Series, Delft University Press, The Netherlands

Minderhoud, M.M., *Supported Driving: Impacts on Motorway Traffic Flow*, T99/4, July 1999, TRAIL Thesis Series, Delft University Press, The Netherlands

Hoogendoorn, S.P., *Multiclass Continuum Modelling of Multilane Traffic Flow*, T99/5, September 1999, TRAIL Thesis Series, Delft University Press, The Netherlands

Hoedemaeker, M., *Driving with Intelligent Vehicles: Driving Behaviour with Adaptive Cruise Control and the Acceptance by Individual Drivers*, T99/6, November 1999, TRAIL Thesis Series, Delft University Press, The Netherlands

Marchau, V.A.W.J., *Technology Assessment of Automated Vehicle Guidance - Prospects for Automated Driving Implementation*, T2000/1, January 2000, TRAIL Thesis Series, Delft University Press, The Netherlands

Subiono, *On Classes of Min-max-plus Systems and their Applications*, T2000/2, June 2000, TRAIL Thesis Series, Delft University Press, The Netherlands

Meer, J.R., van, *Operational Control of Internal Transport*, T2000/5, September 2000, TRAIL Thesis Series, Delft University Press, The Netherlands

Bliemer, M.C.J., *Analytical Dynamic Traffic Assignment with Interacting User-Classes: Theoretical Advances and Applications using a Variational Inequality Approach*, T2001/1, January 2001, TRAIL Thesis Series, Delft University Press, The Netherlands

Muilerman, G.J., *Time-based logistics: An analysis of the relevance, causes and impacts*, T2001/2, April 2001, TRAIL Thesis Series, Delft University Press, The Netherlands

Roodbergen, K.J., *Layout and Routing Methods for Warehouses*, T2001/3, May 2001, TRAIL Thesis Series, The Netherlands

Willems, J.K.C.A.S., *Bundeling van infrastructuur, theoretische en praktische waarde van een ruimtelijk inrichtingsconcept*, T2001/4, June 2001, TRAIL Thesis Series, Delft University Press, The Netherlands

Binsbergen, A.J., van, J.G.S.N. Visser, *Innovation Steps towards Efficient Goods Distribution Systems for Urban Areas,* T2001/5, May 2001, TRAIL Thesis Series, Delft University Press, The Netherlands

Rosmuller, N., *Safety analysis of Transport Corridors*, T2001/6, June 2001, TRAIL Thesis Series, Delft University Press, The Netherlands

Schaafsma, A., *Dynamisch Railverkeersmanagement, besturingsconcept voor railverkeer op basis van het Lagenmodel Verkeer en Vervoer*, T2001/7, October 2001, TRAIL Thesis Series, Delft University Press, The Netherlands

Bockstael-Blok, W., *Chains and Networks in Multimodal Passenger Transport. Exploring a design approach*, T2001/8, December 2001, TRAIL Thesis Series, Delft University Press, The Netherlands

Wolters, M.J.J., *The Business of Modularity and the Modularity of Business*, T2002/1, February 2002, TRAIL Thesis Series, The Netherlands

Vis, F.A., *Planning and Control Concepts for Material Handling Systems*, T2002/2, May 2002, TRAIL Thesis Series, The Netherlands

Koppius, O.R., *Information Architecture and Electronic Market Performance,* T2002/3, May 2002, TRAIL Thesis Series, The Netherlands

Veeneman, W.W., *Mind the Gap; Bridging Theories and Practice for the Organisation of Metropolitan Public Transport*, T2002/4, June 2002, TRAIL Thesis Series, Delft University Press, The Netherlands

Van Nes, R*., Design of multimodal transport networks, a hierarchical approach,* T2002/5, September 2002, TRAIL Thesis Series, Delft University Press, The Netherlands