Choice Set Generation in Multi-Modal Transportation Networks

Maria Stella FIORENZO-CATALANO

Delft University of Technology, 2007

This thesis is a result from a project
funded by the co-operative research programme T3 of the
Netherlands Organization of Applied Scientific Research TNO
and the Netherlands Research School for Transport, Infrastructure
and Logistics TRAIL.

Cover illustration: Maria Stella Fiorenzo-Catalano

Choice Set Generation in Multi-Modal Transportation Networks

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 dinsdag 5 juni 2007 om 12.30 uur

door

Maria Stella FIORENZO-CATALANO

Master of Science in Computer Science

University of Pisa (Italy)

geboren te Venezuela

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

Toegevoegd promotor: Dr. ir. R. van Nes

Samenstelling promotiecommissie :

Rector Magnificus	voorzitter
Prof. dr. ir. P.H.L. Bovy	Technische Universiteit Delft, promotor
Dr. ir. R. van Nes	Technische Universiteit Delft, toegevoegd promotor
Prof. ir. F.M. Sanders	Technische Universiteit Delft
Prof. dr. O.A. Nielsen	Technical University of Denmark
Prof. dr. H.J.P. Timmermans	Technische Universiteit Eindhoven
Prof. dr. H.J. Meurs	Radboud Universiteit Nijmegen
Prof. dr. ir. E.C. van Berkum	Universiteit Twente

This thesis is the result of a Ph.D. study carried out from 2000 to 2004 at Delft University of Technology, Faculty of Civil Engineering and Geosciences, Transport and Planning Section.

TRAIL Thesis Series no. T2007/6, The Netherlands TRAIL Research School

TRAIL
P.O. Box 5017
2600 GA Delft
The Netherlands
Phone: +31 (0) 15 278 6046
Fax: +31 (0) 15 278 4333
E-mail: info@rsTRAIL.nl

ISBN 978-90-5584-087-8

Printed in The Netherlands

”Ignoti nulla cupido”
- Ovidio (*Ars amatoria, III, 397)*

Contents

List of Figures

List of Tables

Notation

The main shorthand and symbols that are used in this thesis are presented as follows:

Shorthand

P&R : Park and Ride facilities
B&B : Branch and Bound technique
UM : Uni-Modal
MM : Multi-Modal
OD : OD pair
RSG : Route Set Generation
CSG : Choice Set Generation
SP : Shortest Path
KSP : *k* Shortest Paths approach
CKSP : Constrained *k* Shortest Paths approach
MC : Monte Carlo approach
AMC : Accelerated Monte Carlo approach
MCL : Monte Carlo Labelling combination approach
AMCL : Accelerated Monte Carlo Labelling combination approach
MM : Main modes of a multi-modal trip
HM : Home-based modes of a multi-modal trip
AM : Activity-based modes of a multi-modal trip
$MM-CSG$: Multi-Modal Choice Set Generation approach

Network and Choice set notation

N : Set of nodes
L : Set of links
M : Set of modes
$G = (N,L)$: Uni-modal graph with set of nodes and links
$G = (N,L,M)$: Multi-modal graph with set of nodes, links, and modes
o, d : Origin, destination indices
i, j, x : Node indices
a, b, $\hat{a}$, $\hat{b}$, l, z : Link indices
h, k, r, p : Route indices
c : Link cost
m : Mode index
u, v : PT stops index
r : Run index of PT line
cen : Centroid index
G_v : Service sub-graph for each PT line v
G_d : Demand sub-graph
G_{ae} : Access / egress sub-graph
G_{vr} : Service sub-graph for each run r of PT line v

Q : Group of individual travellers
q : Individual traveller index
s : User class index
T : Time period such as morning or evening peak
T_w : Time window such as [9:00, 10:00]
t : Departure time instant, for example, 9:15
n : Choice situation index representing $\{OD, t, q\}$
ms, MS : Master choice set for the traveller's, researcher's perspective
OS : Objective choice set for the researcher's perspective
ss, SS : Subjective choice set for the traveller's, researcher's perspective
cs, CS : Consideration choice set for the traveller's, researcher's perspective
ca, CA : Chosen alternative for the traveller's, researcher's perspective
$\mathbf{ss^a}$: Actual joint subjective choice set
$\mathbf{cs^a}$: Actual joint consideration choice set
$\mathbf{ca^a}$: Actual chosen alternative
$\mathbf{SS^o}$: Observed joint subjective choice set
$\mathbf{CS^o}$: Observed joint consideration choice set
$\mathbf{CA^o}$: Observed chosen alternative

Route choice and Choice set generation model notation

U_{kq} : Utility function of alterative k and individual q
V_{kq} : Systematic part of the utility function of alterative k and individual q
ε_{kq} : Random term capturing the various uncertainties
P_{kq} : Probability that the alterative k is chosen by individual q
$X_{kq}(y)$: Vector of y attributes of the route alterative k and the individual q
$\beta(y)$: Vector of y parameters to be estimated
y : Index of the Y attributes considered by individual q
μ : Model's scale parameter
δ : Weight coefficient of the route overlap
τ : Path travel time
CF_k : Commonality factor of path k
d_{kh} : Length (cost) of links in common to paths k and h
d_k : Overall path lengths (costs) of paths k
l_a : Length of link a
PS_k : Path size of route k
γ_0, γ_1 : Positive parameters to be estimated
γ : Size assignment parameter
Γ_k : Set of links belonging to route k
δ_{ak} : Binary variable (0/1) indicating whether link a is part of route k or not
$|CS|$: Cardinality of the consideration set, i.e. the total number of routes in the CS
η : Measure of similarity of routes k and h

F : Disutility measure such as time, distance, number of links
$F_O(x)$: Shortest time or distance from origin O to node x
$F_D(x)$: Shortest time or distance from node x to destination D
$d[i,j]$: Shortest connection (in distance, time, or cost) between any nodes i and j
α : Detour parameter $\alpha \geq 1$
Δ : Mutual route overlap percentage between two paths $0 \leq \Delta \leq 1$
θ_1, θ_2 : Route comp. thresholds $0 \leq \theta_1 \leq 1$, $0 \leq \theta_2 \leq 1$, and $\theta_1 < \theta_2$
ω_{max}, ω_{min} : Max and min route detour percentages $\omega_{max} \geq 1$, $\omega_{min} \geq 1$
S : Choice set size
$\bar{C}_{qk}$: (Stochastic) generalized cost of a route k perceived by the individual q
δ_{ma} : Binary variable (0/1) indicating whether link a belongs to mode m or not
R : Total number of alternative routes in a network
κ : Number of draws from a network with R alternative routes
$\mu(X)$, $\mu(\beta)$: Expectation associated with X and β
$\sigma(X)$, $\sigma(\beta)$: Variance associated with X and β
$\hat{\varepsilon}$: Random error term attached to parameters β
$\tilde{\varepsilon}$: Random error term attached to attributes X

Chapter 1

Introduction

1.1 Context and background

Multi-modal trips, i.e. trips using two or more vehicular modes between which a transfer is necessary, are a common travel phenomenon which are expected to become more important in the future. Although multi-modal trips overall only account for less than 3% of all total passenger transport in The Netherlands, as noted in Van Nes (2002), multi-modal transport merits attention as it occupies an important niche for longer distance inter-urban transport. For instance, over 20% of the inter-urban trips from and to the larger Dutch cities are multi-modal trips with usually train as the main transport mode. An increase in the market share of multi-modal transport may increase public transport occupancy rates and improve liveability of city centres. To improve the opportunities and conditions for multi-modal transport, enlarged insight is required into the possibilities for multi-modal trip making with respect to the availability of travel modes (supply side) as well as the preferences of individual travellers (demand side).

Possible benefits of multi-modal transport are reductions of long distance car trips by offering better access to long distance public transport and improving the accessibility of city centres, for instance, by introducing transfer points to high quality public transport services at the outskirts of the city. In order to achieve these potential benefits of multi-modal transport it is necessary to provide facilities for multi-modal transport, for instance, for transfers between modes. Such transfer points require stops or stations offering high quality public transport services and sufficient parking facilities for private cars and bicycles. It is therefore important to assess the performance of such transfer points in advance. In order to analyse multi-modal planning problems, such as the location and design of inter-modal transfer points, the travel demand modelling needs to be capable of analysing and predicting the use of multi-modal trips.

Given the growing importance of chain mobility, there is an increasing number of planning problems that require tools for analysing multi-modal travel choices, which adequately can deal with a simultaneous choice of routes, transport services, travel modes and transfer locations. In the past thirty years, respectable progress has been made towards quantitative modelling of transportation systems. For practical reasons, progress has centred on uni-modal rather than multi-modal systems. To handle multi-modal transport systems, the classical modelling approach might be generally appropriate, but it shows serious shortcomings when true multi-modal transport involving different modes and services within a trip needs to be modelled.

For example, currently available toolkits to assess transport investments are based on a rigid distinction between modes, especially cars and public transport. Slow modes such as bicycles and walking are often not fully accounted for. Consequently, the impact of combining modes in a trip is difficult to assess. A pragmatic approach is to limit multi-modal transport to a few promising combinations of modes for specific trip types, which are simulated in a uni-modal network, being either private car or public transport. However, the analysis in Van Nes (2002) shows that multi-modal transport covers a large variety of modal combinations with an even larger diversity of trip characteristics. Furthermore, in a number of countries such as Denmark and The Netherlands, bicycles play an important role in multi-modal transport. Therefore, the pragmatic approach falls short of covering the whole range of multi-modal transport possibilities.

Route-choice and mode-choice are the standard ingredients of classical transport models. Route-choice is usually modelled using (equilibrium) assignment, while mode-choice is typically done separately (before the assignment step). Although the choice of interchanges between public transport modes has been investigated in the literature, the choice of entry-exit points of the transit network has not been modelled in a general way so far. This can be seen as a lack of classical transport modelling techniques. This issue has been considered in the literature in which several proposals have been presented (e.g. Fernandez et al. (1994)). However, these approaches are essentially hardly applicable and recommendable when it comes to real-world networks and especially multi-modal networks.

On the basis of available literature it can be concluded that the poor handling of multi-modal trips is caused among other matters by the strict separation between mode choice and route choice that is maintained in classical modelling approaches.

To remedy the shortcomings of current approaches, this thesis proposes a multi-modal modelling approach that accounts for simultaneous choice of mode and route through a multi-modal transportation network. The modelling approach is based on a representation of the multi-modal network known as a supernetwork. In this modelling framework, mode choice is the result of route choice in the multi-modal supernetwork. As a result, the travel behavioural modelling focuses on route choice only, the quality of which is now the main determinant of the quality of the model system itself, which as a result receives much attention.

In order to better model route choice in the supernetwork, we assume that a traveller has a set of possible (multi-modal) route alternatives available to him or her for a specific trip (i.e. *choice set*) from which he or she chooses the alternative that is most suited for his/her travel need (see e.g. Bovy, P.H.L. and Stern, E. (1990) and Hoogendoorn-Lanser (2005)). Choice sets play a critical role in choice model estimation and demand prediction, significantly influencing the validity of parameter estimates and predicted demand levels. Profound research into choice set modelling makes progress in tackling choice set misspecification problems and in identifying the variables that determine the individual's choice set formation.

Furthermore, the explicit generation of choice alternatives *prior* to the choice prediction is preferred by assuming that a clear distinction in the modelling between a choice set generation step and, conditional on that set, the genuine choice modelling step, may significantly improve choice analysis and prediction quality. Therefore, the route choice modelling is split into the *choice set generation* and the consequent route choice modelling.
The idea of this choice set modelling approach is based on the same assumption of the two stages process introduced by Manski (1977), although Manski's approach differs from our proposed approach. Manski's approach is a probabilistic approach, which can be applied only with limited choice sets, and usually it is applied for estimation purposes, whereas our approach is proposed to be applied in the route choice context, with typically large choice sets, and for demand prediction purposes.

Choice set generation consists in finding all feasible routes that a traveller might consider for travelling from his origin to his destination. Choice set generation for route choice modelling is known as a difficult problem compared to other choice modelling problems such as mode choice or destination choice. It is well known that in a route choice context, choice set composition is a critical aspect because very many routes may be available whereas only a limited subset of those are actually perceived while even less are actually considered by trip makers. In a multi-modal transport network, route finding and generation is even more difficult because of the multiple different types of choices involved in a multi-modal trip. The specific theoretical challenge in modelling multi-modal trips is in the multi-dimensional character of these trips encompassing a multitude of choices with respect to routes, travel modes, transport service types, and interchange locations between public transport modes, access/egress locations from private to public transport modes and vice versa. This is the challenge that is addressed as subject of study in this thesis.

The research presented in this thesis is part of a joint research project, the so-called Hypernetwork project, carried out in the Transportation Research School TRAIL and in collaboration with the Organization for Applied Scientific Research (TNO). The Hypernetwork project aims to develop a practically applicable model approach to analyse proposed multi-modal passenger transport systems analytically. The project aims to:

- Develop the analysis methodology;
- Provide its theoretical foundations;

- Develop required travel demand prediction models;
- Establish required computational tools;
- Demonstrate the validity and practical applicability.

The developed multi-modal demand prediction model should be applicable to a much larger group of research issues, compared with the current methods of analysis for passenger traffic. It should mainly be appropriate for investigating effects on chain mobility of changes in public transport services. The model should support planning decisions like determining the location of a multi-modal interchange, perform market analysis of multi-modal services, and determine the level of competition between multi-modal alternatives. It should also be able to evaluate (both fixed and demand driven) public transport services or the influence of properties of the interchange facility like parking space, parking toll, transfer time, etc.

This thesis presents parts of the achievements of the multi-modal transport modelling approach developed within the Hypernetwork project, focusing on the route choice modelling part and more specifically on the choice set generation model for multi-modal trips, that is on generating a set of realistic multi-modal routes (transport modes, service types, transfer nodes, etc.) that a group of travellers might consider for making a multi-modal trip.

1.2 Research objectives and related research questions

This thesis deals with choice set generation in the case of multi-modal trips exhibiting multiple-choice dimensions (such as transport modes, transfer nodes, public transport service types, and routes), which are represented jointly as a case of route choice in a multi-modal network for predicting flows in the multi-modal network.

The research objective of this thesis thus is to establish a choice set generation model and algorithm, and demonstrate its validity and feasibility for demand prediction purposes.

Note that this thesis starts from the presumption of superiority of explicit a priori choice set generation. For prediction applications, explicit generation of route choice sets *prior* to route and link flow calculation instead of during the iterative flow calculation is strongly favoured for several reasons. A major advantage of a priori choice set generation for prediction purposes is that a priori given choice sets allow much more flexibility and realism in behavioural assumptions in the route choice models to be adopted. In that case, no restrictions exist on the type of choice model or utility functions specification to be adopted. Applicability of advanced analytical approaches, easy consideration of route overlap, non-linear utility functions, and route-based attributes are only a few advantages of applying a priori generated choice sets. A priori generation in a network context offers also implementation and computational advantages in iterative network assignment approaches since no repeated optimal route search is necessary. This has, for instance, been demonstrated in dynamic equilibrium modelling of a large road network (Bliemer, M.J.C. and Versteegt, E. and Castenmiller, R., 2004).

Obviously the consequence of a priori choice set generation poses some limitations. First of all, the generated routes need to be stored, and for large networks, the memory requirements could be huge, although with new technologies, problems related to memory requirements are decresing. Moreover, the advantages of a priori choice set generation are obtained at the expense of greater computational complexity related to the large dimension of the choice sets. Second, in a priori choice set generation special attention is required to adequately handle the impact of heavy congested situation. Finally, a priori choice set generation approach is especially convenient for demand prediction purposes, since large choice sets containing non-relevant routes are better than the ones that are too small and in which relevant routes are missing.

An important working hypothesis in this thesis is that the composition of individual choice sets is strongly determined by individual preferences for trip attributes. In fact, several studies (see Bovy, P.H.L. and Stern, E. (1990)) have shown that in order to obtain proper choice sets traveller's characteristics have to be taken into account. Different travellers may have different preferences (so-called taste variation) so that the population consists of a mix of several segments, each with their own attribute weights, and the determination of their choice sets depends on their preferences and the main factors that influence route choice behaviour.

The specific central research question of this thesis focusing on choice sets is the following: how to generate appropriate choice sets in multi-modal networks accounting for the variety of multi-modal travel behaviour?

The more detailed research questions derived from the central question are among others the following:

- Which are the main characteristics of adequate route choice sets?
- What is the size and composition of an appropriate choice set?
- How to determine the quality of the generated choice sets?
- Which (model-based) methods are appropriate to generate choice sets?
- How to generate appropriate choice sets in multi-modal networks?

In order to answer all these questions in this thesis, new choice set terminology is introduced, a new formulation of criteria for reasonable choice sets in uni-modal and multi-modal networks is developed, choice set generation approaches are analyzed and compared, and a new choice set generation approach for multi-modal networks is proposed and evaluated. All these topics are dealt with in the next chapters.

1.3 Thesis contributions

The specific contributions of this thesis can be grouped into the following five topics:

1. Multi-modal modelling approach and multi-modal transportation network representation.
2. Choice set notions and concepts.
3. Criteria definitions for adequate choice sets in uni-modal and multi-modal networks.
4. Analysis and comparison of choice set generation approaches.
5. New choice set generation approach for multi-modal networks.

1.3.1 Multi-modal modelling approach

In this thesis, a new approach for analyzing multi-modal travelling based on the use of a supernetwork methodology is introduced, in which the networks of all available travel modes are combined into a single supernetwork that includes transfer possibilities between modes. Multiple-choice dimensions in multi-modal trips are represented jointly as route choice in the supernetwork while by explicitly generating the individual choice alternatives prior to the choice modelling maximum flexibility is available in adopting the most suitable choice modelling approach. Moreover, a new multi-modal network representation is presented that takes into account several layers of uni-modal networks and links them via waiting and walking links.

1.3.2 Choice set notions and concepts

New choice set notations and related terminology is presented thereby explicitly accounting for different stages in the travel choice process (available, known, feasible, considered or chosen), viewpoint (travellers or researchers), number of travellers (individual traveller or group of travellers), origin (actual, observed or generated), and applications (analysis, estimation or prediction). Furthermore, conditions that apply to choice sets in the different stages of the decision-making process are regrouped and new clear definitions are given.

1.3.3 Criteria for adequate choice sets

An important topic dealt with in this thesis is the formulation of the main characteristics that adequate choice sets and appropriate generation processes should satisfy in uni-modal and multi-modal networks. The requirements for adequate route choice sets for prediction of route and link flows in uni-modal transport networks are introduced in this thesis. The new criteria concern requirements for a reasonable route, for adequate choice sets at individual level and at group level, and appropriate choice set generation for prediction purposes.

The new criteria established for the uni-modal case are extended to the multi-modal case, focusing on the differences between uni-modal and multi-modal networks, such as, *combination* of several continuous-type and discontinuous-type transport service systems; the *transfer* between transport modes; the *sequence of modes* within a multi-modal trip networks, the route overlap problem which is highly different from the overlap problem in uni-modal *continuous* networks.

1.3.4 Analysis and comparison of choice set generation approaches

Another important contribution of this thesis is the evaluation of the choice set generation approaches for uni-modal networks to be applied for prediction purposes.

A generic choice set generation procedure is specified for the purpose of characterising current route set generation procedures. Based on the generic generation scheme, a classification of the choice set generation methods known from literature is provided, after which each of the generation methods is described in a structured comparable way as a basis for our evaluation of their adequacy for our purposes, on the basis of which recommendations are given for the best approach to be adopted. Due to the generated choice sets size and compositions (many overlapping routes might be generated) a filter process in which some of the defined criteria are applied after the generation step is proposed in order to obtain more adequate choice sets.

1.3.5 Choice set generation approach for multi-modal networks

As the main contribution of this thesis, we present a newly developed route choice set generation algorithm, the so-called doubly stochastic approach, applicable to uni-modal and multi-modal networks, thereby taking into account travellers' origin and destinations, vehicle availability, and travellers groups. The basic hypothesis of this approach is that traveller's preferences with respect to route attributes highly determines the size and composition of the choice set considered in his choice decision. From this hypothesis we state that a traveller's trip utility function can be used as the basis for a generating function with which attractive alternatives can be generated through optimal path search in the given network by stochastically varying network attributes and attribute preferences. While the adopted variances in network attribute values (travel time, waiting time, travel cost, etc.) reflect, among other matters, the variation in perception of the attributes among the travellers and their differences in knowledge about the network, the adopted variances in the related parameter values reflect the variation in the preferences for these attributes within the population of travellers.

1.4 Outline of the thesis

This thesis can roughly be divided into three parts:

1. Multi-modal modelling approaches and uni-modal and multi-modal transportation network representations (Chapters 2 and 3).
2. Conceptual framework for choice set concepts and route choice modelling (Chapters 4 and 5).
3. Choice set generation approaches for uni-modal and multi-modal networks (Chapters 6 and 7).

Chapter 2 introduces the main characteristics and definitions of multi-modal transportation and modelling approaches for analysing multi-modal travel choices. The architecture for proposed new multi-modal demand prediction system is presented in Chapter 2, the noticeable characteristics of which are that a modal split module is absent and a route set generation module prior to the route choice is specified. Chapter 3 focuses on the network modelling concepts for uni-modal and multi-modal trip analysis of which the supernetwork concept will receive ample attention. This network concept has to facilitate the proposed approach in route set generation and route choice analysis. In Chapter 4 a start will be made with the route set generation subject by introducing the necessary choice set concepts. This chapter outlines a theoretical framework with conceptual notions about choice sets, distinguishing between traveller's and researcher's perspective, individual travellers and group of travellers, and actual, observed, generated behaviour. Chapter 5 then gives an overview of route choice modelling, constituting the behavioural context for the route set generation modelling in Chapters 6 and 7. In Chapter 6, we first deal with the route set generation for uni-modal networks. After formulating quality requirements for reasonable routes, adequate choice sets at individual level and at group level; available route set generation procedures will be presented and systematically assessed. Chapter 7 is devoted specifically to the multi-modal case. The specific quality requirements in the multi-modal case are elaborated. Current route set generation approaches for multi-modal networks are reviewed and assessed, after which a new doubly-stochastic generation approach is presented. The validity and effectiveness of this approach will be tested with observations. The findings and conclusions of this thesis and its prospects for future work are summarized in the final Chapter 8.

Chapter 2

Key issues in multi-modal transportation planning

2.1 Introduction

Having introduced the topic of this thesis, we first elaborate somewhat on the objectives of multi-modal transportation planning being the context of the new developments for multi-modal travel demand modelling dealt with in this thesis. Within this context, a presentation of the main characteristics and definitions of multi-modal transport is given answering the question what is meant by multi-modal transport. Subsequently, this chapter focuses on modelling approaches for analysing multi-modal travel choices in a planning context. Starting from a description of the current approaches for multi-modal transportation modelling, we present a new, more appropriate modelling approach specifically addressing the modelling of multi-modal trips.

Section 2.2 deals with the main planning problems typical for multi-modal transportation. The characteristics of current multi-modal mobility are introduced in Section 2.3 using the empirical analysis carried out by Van Nes (2002). It is shown that although multi-modal transport is a niche market in transportation at large, this class of trips nevertheless plays a substantial role in satisfying travel demand for specific trip types. Subsequently, a set of related notions is introduced and definitions of multi-modal transport and trip making are given that will be used in this thesis based on the terminology introduced by Van Nes (2002).

Current practice in static transportation modelling is shown in Section 2.4 including an overview of current approaches to handle multi-modal transportation modelling in Section 2.4.3. Three of these approaches of current practice are discussed in more detail, namely the joint uni-modal transportation modelling, the extended classical approaches, and the supernetwork approach. The main contribution of this chapter is the presentation of a new multi-modal transportation modelling approach, presented in Section 2.5, based on the supernetwork concept integrating the various uni-modal sub networks into a consistent single network. This approach allows treating the various travel choices such as modes, access/egress nodes, uni-modal routes, etc. as chosing a single route in a supernetwork. Essential part of the proposed approach is the a priori route choice set generation in the supernetwork being the main subject of this thesis.

2.2 The transportation planning context

In contrast to the earlier, now obsolete approaches of separate transportation planning procedures for private (car, bicycle) and public transport modes, modern transportation planning is characterized by a multi-modal approach in which the various modes are considered as co-operating elements to jointly provide the best possible efficient transport services to the public (see e.g. Cascetta (2001)). Multi-modal transportation planning is an effort to combine the strengths of the different modes and avoid their weaknesses as far as possible in order to optimise the use of resources and to improve travel services to the public in terms of travel time, cost, reliability, etc.. Specific objectives in multi-modal transport planning include strengthening the functions of cities by improving their accessibility and strengthening the public transport system by improving their access with private modes (see e.g. Bovy, P.H.L. (2003)). By offering better multi-modal transport services, planners' objectives are among other matters to reduce or alleviate road traffic congestion and increase the use of public transport. Typical instruments used in multi-modal transportation planning are the establishment of physical linkages between different modes such as Park & Ride facilities, inter-modal transfer stations, carpool points, and adapted public transport services, co-ordination of service schedules, multi-modal travel information systems, and the like. Park & Ride facilities are created mainly for transfers from private car to train or metro. Carpool locations might be defined as locations where transfers are made between car driver and car passenger transport modes. Van Nes (2002), for example, describes the following possibilities for specific transfer locations between private and public transport:

1. Transfer nodes in city centres offering access to higher-level inter-city public transport networks, combined with access modes walking, cycling, and local public transport;
2. Transfer nodes in local centres that provide access to medium distance public transport, accessible by bicycle and private car;
3. Transfer nodes at the edge of cities providing 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.

In the context of strategic transportation planning, the location of inter-modal transfer facilities is one of the questions of primordial importance.

As a means to assess the effectiveness of long-term multi-modal policies, current travel demand prediction methodologies do not provide the appropriate capabilities in dealing with the typical characteristics of multi-modal trip making, in particular they fail to capture the rich variety in modal composition of multi-modal trips such as the mix of private and public modes. Given the growing importance of multi-modal transportation, there is a need for new demand modelling methodologies that can sufficiently handle the typical multidimensional travel decision making of travellers in the case of multi-modal trips. These travel choice dimensions include among other matters: choices of transport services and vehicular modes for different parts of the trip, choices of transfer points where to switch modes, choices of routes in the modal networks between transfer points, etc. The new demand modelling methodology to be developed should facilitate not only the estimation of the use of planned multi-modal facilities (in order to assess their cost-effectiveness) but should equally well enable the estimation of the costs and benefits of travelling via the various singular and combined modes available. This thesis will contribute to the establishment of such a new demand estimation methodology for application in strategic multi-modal transportation planning. Its contribution first involves the establishment of a modelling architecture (see Section 2.5) that appropriately can deal with the complex and flexible structures of multi-modal trips. Basic to this architecture is the establishment of a so-called supernetwork being an integration of the various uni-modal sub networks such that it facilitates modelling the various mode choices, service choices, and transfer location choices involved in multi-modal trip making as a single route choice in the supernetwork. Part of this architecture is a separate new step in the demand modelling process, namely the generation of route choice sets in the supernetwork in advance of the choice modelling and network assignment steps. It is this generation step being the main topic of this thesis.

2.3 Multi-modal travelling

2.3.1 Characteristics of multi-modal trips

Main characteristic of a multi-modal transport is that more than one mechanized transport mode is used for a trip from origin to destination. One or more transfers between transport modes are thus an essential element of multi-modal transport. According to Van Nes (2002), based on mobility data from the Dutch National Travel Survey, the current share of multi-modal transport in The Netherlands is only about 3% if viewed nationwide. However, the share of multi-modal trips is more than 20% for inter-urban trips to and from the main Dutch cities. Of all the multi-modal trips, the largest share (75%) is with public transport as main mode (train and bus), while car accounts for 15%. Train alone accounts for nearly 60%. Of all train trips, more than 80% are multi-modal. The statistical analyses carried out by Van Nes (2002) show that the main factors determining multi-modal trip making are:

- Trip length: multi-modal transport is more used for longer trips;
- Destination area type: multi-modal trips are predominantly oriented to the main cities and their centres;
- Trip purpose: more than 50% of the multi-modal trips are for work or education purposes.

Thus, for trips in which these factors are combined the share of multi-modal transport appears to be substantially.

These figures from The Netherlands are similar for other Western European countries such as Germany or Switzerland. However, in some Eastern Asian countries (Singapore, Hong Kong) the share of multi-modal trips is much higher, attaining even 60% in Singapore (Lam, 2002). In USA multi-modal transport is suited only for very long distance trips, using air transport, or car-train and Park & Ride transfer locations. Usually there are few transfers from inter-urban to urban public transport (see Cirillo, C. and Axhausen, K.W. (2002)).

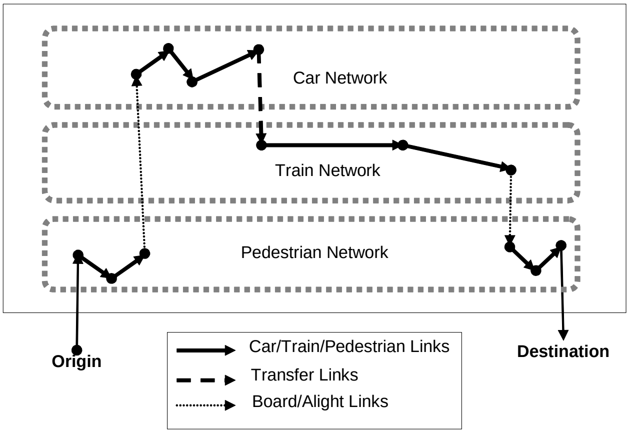


Figure 2.1: Representation of a multi-modal trip using two vehicular transport modes (car and train).

2.3.2 Definitions of multi-modal trips

In this thesis we use the same notions and terminology of multi-modal transport as defined in Van Nes (2002). According to those definitions, multi-modal trips are trips in which two or more different vehicular modes or transport services (excluding walking) are used for a single trip from origin to destination, between which at least one transfer is necessary. Figure 2.1 shows an example of a multi-modal trip in which a car is used to access the railway system; the dashed link represents the transfer from car to train. A uni-modal trip is a trip in which only a single vehicular mode or walking-only is used. Since the term *mode* might have different meanings, its definition needs to be more specific.

A *travel mode* might refer to service type, to transport function, or to vehicle type, or a combination of these. A distinction can be made between service modes, namely private and public transport modes, and vehicle modes, such as private car and bicycle (private vehicles) and bus, tram, metro and train (public vehicles). Since multi-modal 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.

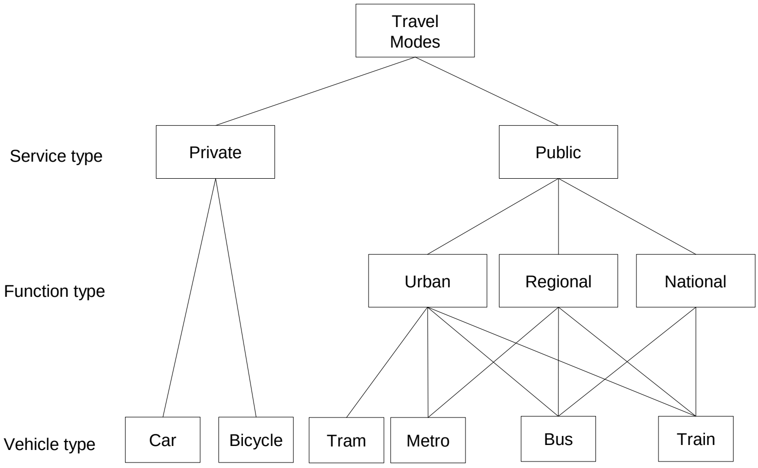


Figure 2.2: Distinction of different types of travel modes.

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 vehicle types used. These characteristics are often strongly related to functionally different network levels, that is urban, regional, national public transport networks. Multi-modal transport thus concerns transfers between private transport modes, between private and public transport services, and between functionally different types of public transport services. This distinction in mode types is illustrated in Figure 2.2. 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.

The part of a trip where a single mode is used without transferring is called a *leg*. Apart from the walking legs, multi-modal trips consist of multiple vehicular legs by definition, whereas uni-modal trips consist of a single mechanised mode leg that might also contain one or more walking legs.

Transfers are essential parts of a multi-modal trip. In order to use two or more transport modes travellers have to change mode at transfer nodes. Since transfers may also occur in uni-modal public transport networks, for example changing from one bus line to another, the definition of transfers needs to be more specific. Two types of transfers may be distinguished namely the *inter-modal transfers*, which are transfers where travellers change transport service network or modes, and the *intra-modal transfer*, which is a transfer within a transport service network, between urban buses for instance. A typical example of inter-modal transfer in a multi-modal trip is the transfer from an urban bus to a regional bus (see Figure 2.3). In this thesis, the term transfer is used for *inter-modal* as well as *intra-modal* transfers. Given these transfer definitions, a trip consisting of at least two mechanized modes between which at least one transfer is required, is then defined as a *multi-modal trip*.

Figure 2.3 shows a few examples of uni-modal and multi-modal trips. In the case of uni-modal trips, the same vehicle mode or transport service with the same function type (e.g. urban) is used. In the case of multi-modal trips, there are transfers between private and public transport services (car and train), and between functionally different types of public transport services (urban and regional bus services, or bus and train services).

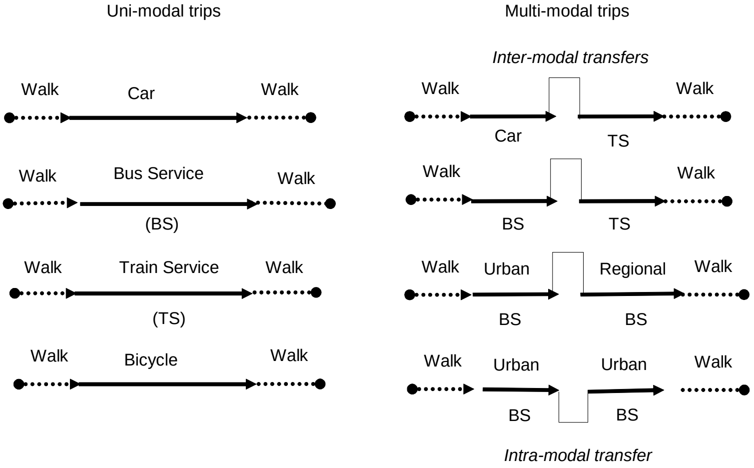


Figure 2.3: Examples of uni-modal and multi-modal trips.

As shown in Figure 2.3, *walking* is considered to be always part of a trip. While this is obvious in the case travellers have to walk to and from the stops of the public transport system, using car or bike also requires some walking to and from the parking place, although these distances might be short. In addition, transferring between modes in most cases also involves some walking between alighting and boarding the vehicles. Walking can thus be considered as a universal component of any trip, and is therefore not considered as a separate mode in the definition of a multi-modal trip. Despite this, walking can constitute a substantial part of a trip, in terms of distance and time, e.g. for accessing urban public transport stops, railway stations or at transfer points. Therefore, in analysing multi-modal travel behaviour the walking legs of trips cannot be neglected (see Hoogendoorn-Lanser (2005)).

The definition of multi-modal transport should ideally be based on *tours* and not on single trips. A tour is a series of trips that starts and ends at the traveller's home base or other activity base (e.g. work place). Tours can be either uni-modal, in which a single mode is used (the same mode for each trip of the entire tour), or multi-modal, in which two or more different transport modes are used in a single trip of the tour. However, a multi-modal tour does not necessarily contain a multi-modal trip. For example, a traveller taking a bus to go to work and being brought home as a car passenger by a colleague, is an example of a simple multi-modal tour (bus and car are the two used modes) consisting of two uni-modal trips. This example also illustrates that tours become essential in describing traveller's behaviour. It is not very logical, in this example that the traveller drives his own car for the return trip. The use of transport modes in earlier trips in the tour determines the availability and use of transport mode in later trips in the tour. For trips starting at home there may be various vehicle modes available than for trips starting at other locations. Furthermore, it must be taken into account that a private mode used for the 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 train, should return at the end of his tour to the railway station where the car was parked.

Ideally, in travel choice modelling, tours should be the basic unit instead of single trips (as is the case in the Dutch National Modelling System). However, data on travel choices in tours are neither always available nor of sufficient detail for modelling multi-modal travel behaviour; therefore, travel research in most cases has to rely on single trips instead of tours. Also in this thesis, given available data we have to confine our analyses on multi-modal trips instead of tours, leaving research on the tour level for the future.

The composition of a multi-modal trip might be quite complex as it may consist of a series of several different travel modes connected by walking legs. Within a trip we may in principle distinguish three different trip parts, namely the main part, the access part, and the egress part. The *main trip part* is that part performed over the largest distance with the highest possible speeds compared to the access and egress trip parts. Typical main trip modes are car and train, sometimes also bus instead of train. The main trip part is in-between two end nodes that connect the access and egress trip

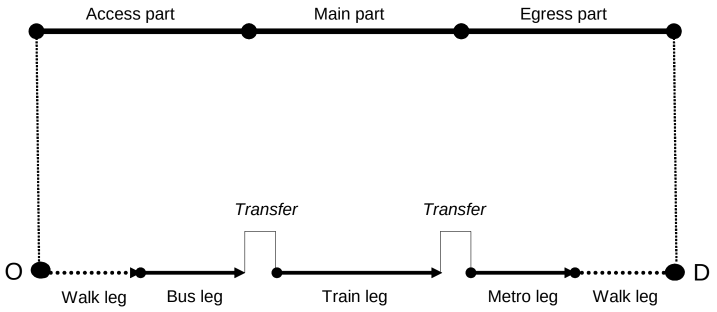


Figure 2.4: Multi-modal trip components and legs (O=origin, D=destination).

parts to the trip origin and trip destination respectively (see Figure 2.4). If the trip origin or the trip destination is the traveller's home address, the corresponding access and egress trip parts are called the home-end part of the trip, else these are called the activity-end part of the trip. With each of the three distinguished trip parts (access, main, egress) we may associate corresponding travel modes. In principle, all modes may fulfil the role of an access, an egress, or a main mode. For example, walking is the main mode in a walking-only uni-modal trip, while the bike maybe the main mode in a walk-bike-walk uni-modal trip. In the practice of interregional trips, typical main modes in multi-modal trips are car, train or regional bus, whereas typical access and egress modes are walking, bike, bus, tram, metro, etc. In principle, car may also play a role as an access-egress mode, such as in car-train-walk trips or walk-train-taxi trips.

Each of the three trip parts (main, access, egress) may consist of one or more legs (see example in Figure 2.4).

Table 2.1 gives a summary of the definitions of the introduced concepts with respect to multi-modal travelling, used in the sequel of this thesis.

For an empirical analysis of the composition of multi-modal trips in general, see Van Nes (2002), and for multi-modal train trips in particular, see Hoogendoorn-Lanser (2005) and Ministrie van Verkeer en Waterstaat (2002).

Table 2.1: Summary of concept definitions introduced on multi-modal travelling (adapted from Hoogendoorn-Lanser, 2005).

Concept	Definition
Transport service network	Defined by a function type (urban, regional or national) and a vehicle type (car, bike, tram, metro, bus or train) used for travelling.
Tour	Trip or series of trips starting and ending at the traveller's base address (usually the traveller's home address).
Single tour	Single trip starting and ending at the traveller's base address without intermediate stops at other locations.
Simple tour	Sequence of two trips, i.e. an outbound trip from the traveller's base address to another location for performing an activity, and a return trip from that location back to the traveller's base address.
Trip	Sequence of travel modes and transfer nodes used in connecting a given OD pair.
Travel alternative	Uni-modal or multi-modal trip with corresponding departure time.
Mode	Means of transport defined by a service type (private or public), a function type (urban, regional or national), and a vehicle type (car, bike, tram, bus or train) used for travelling.
Main trip part	Part of multi-modal trip performed with main modes over the largest distance with the highest possible speeds and consisting of one or more legs.
Access part	Part of multi-modal trip with access modes connecting the trip's origin to the start node of the main trip part (often a transfer point) and consisting of one or more legs.
Egress part	Part of multi-modal trip with egress modes connecting the end node of main trip part (often a transfer point) to the trip's destination and consisting of one or more legs.
Leg	Part of a multi-modal trip for which a single mode is used without intermediate transfers.
Uni-modal trip	Trip with a single mechanised mode leg that might also contain one or more walking legs.
Multi-modal trip	A trip type consisting of at least two mechanized modes (apart from walking) between which at least one transfer is required.
Transfer	Change of transport vehicle within or between transport service networks.
Transfer point	Node in a mechanized transport service network where a change to another vehicular mode is possible.
Inter-modal transfer	Transfer type where travellers change transport service networks of different modes.
Intra-modal transfer	Transfer type within a transport service network (e.g. within an urban bus network or train network) where travellers change vehicle of the same mode.

2.4 Modelling multi-modal travel demand

2.4.1 Modelling requirements

In modelling inter-urban trips, or for example, in evaluating locations for Park & Ride facilities, transport models must be able to deal with multi-modal trips appropriately. The specific theoretical challenge with modelling multi-modal trips is in the multi-dimensional character of these trips. Multi-modal travelling involves complex alternatives consisting of several different legs, the planning of which by the traveller involves complex travel decision making choosing from a variety of transport services, vehicular modes, boarding and alighting stations, routes and the like.

Multi-modal passenger travel modelling needs to deal with a wide range of mode combinations: various vehicular modes might be used in the access and egress parts of a multi-modal trip, such as car, bicycle, bus, tram, and metro. If we combine these possible access/egress travel modes with the possible main modes, such as train, this results in a huge number of multi-modal alternatives. This variety in multi-modal trip composition is further enlarged by the various possibilities of transfer points for boarding, alighting, and switching main modes.

In order to model multi-modal trips it is necessary to have a tool that is able to process such complex combinations of travel modes. Traditional transportation modelling approaches based on the segmentation between car and public transport trips are not able to model correctly such variety of mode combinations within a trip. Actually, in the classical transportation model, the analysis of such a multi-modal trip is simply ignored. The justification of this classical approach might be based on the observation (see Van Nes (2002)) that currently *on average* the percentage of multi-modal trips is only about 3%. However, of the inter-urban trips from and to the major cities in The Netherlands over 20% are multi-modal. In the light of this percentage, there appear to be non-negligible niches of travel demand for which advances in multi-modal transportation modelling are needed. The remainder of this section will address the issue of adapting the classical uni-modal travel demand modelling procedure towards a multi-modal travel demand modelling approach. To that end, we first will summarize from the literature a few attempts in practical modelling to include some forms of multi-modality. This will then be followed in next Section 2.5 by an elaboration of our proposed supernetwork approach for dealing with multi-modal travel demand modelling.

2.4.2 Current practice in transport demand modelling

Current practice in transport demand modelling centres around the four-step procedure, in which each of the stages: production and attraction, distribution, modal split, and traffic assignment, are modelled separately (see Figure 2.5).

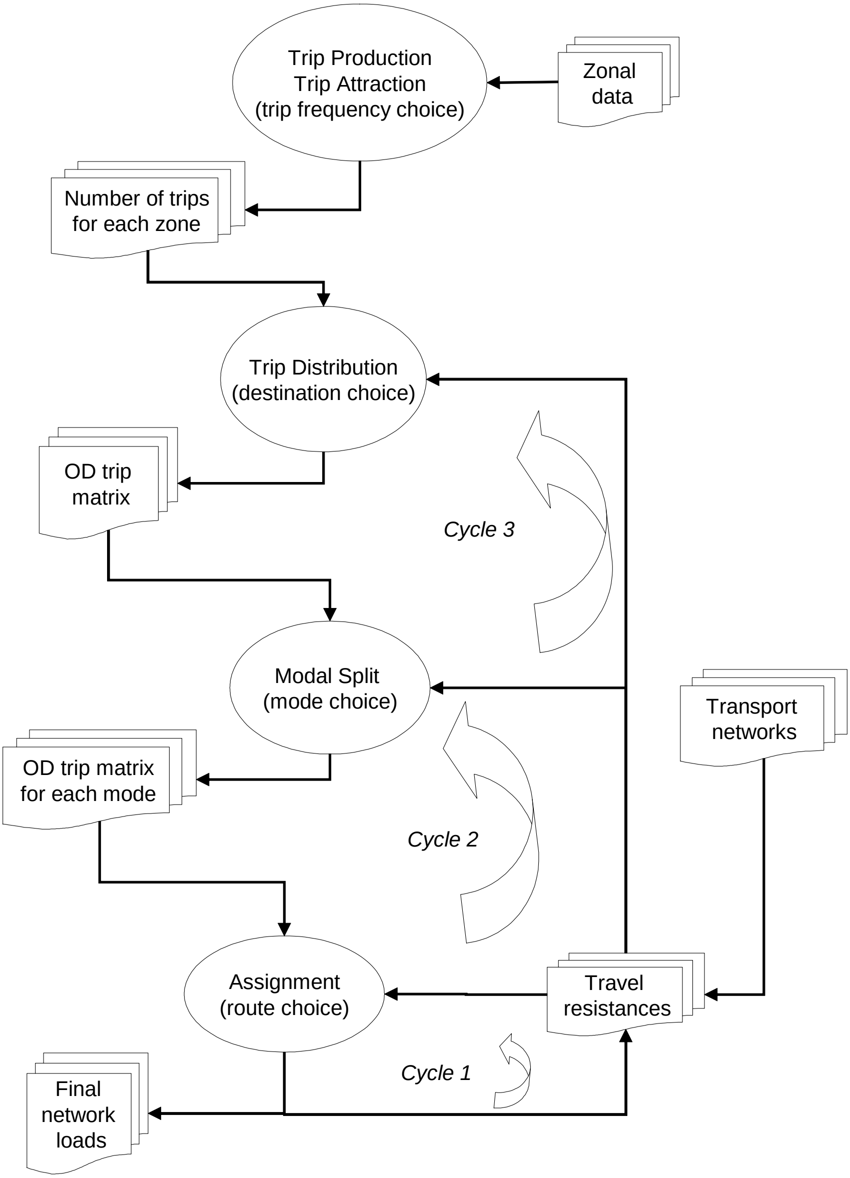


Figure 2.5: Classical four-stage travel demand model system.

The complete model consists of a concatenation of these sub-models. The exogenous input data of each sub-model are shown in the right part of the figure, whereas the endogenous demand prediction data produced by each sub-model are shown in the left part of the figure. Travel resistances or disutilities consisting of travel times, costs, distances play an important role in nearly all travel choices. These resistances depend on the level of usage of the network elements: the higher the network loads, the higher the travel resistances. Since the level of usage is only known at the very end of the calculation cycle (after the network assignment step) an iterative calculation procedure is required in order to achieve consistent final results. Mostly travel resistances resulting from the assignment are assumed to influence the route choice (feedback loop 1), the mode choice (feedback loop 2) and the destination choice (feedback loop 3) steps. Ideally, these iterative processes should be performed several times for each distinct analysis case in order to achieve a satisfactory consistency between the demand prediction results at the various stages. In practice, feedback loops 2 and 3 are seldom performed, whereas feedback loop 1 is often but not always included and processed within the assignment step. (Most recently, departure time is added to this system to allow a more adequate dynamic modelling, which however is not yet state of the practice.)

In most application cases, the assignment cycle is restricted to the car network, implying that travel resistances in the other networks (public transport, bicycle) are fully determined by the input assumptions of these networks and are independent of the flows in these networks. There are however some interdependencies between the predicted demand levels in the various modes: predicted flow levels in the car network determine the resistances in the car network which in turn influence via the mode choice mechanism the flow levels in the other modal networks. In this classical demand modelling approach the various modal networks are completely separated such that only uni-modal trips can be predicted while multi-modal trips are neglected.

An advanced example of the classical approach is described by Cohn et al. (1996) with a model to predict and analyse mode competition between rail and other modes where access and egress level of service directly affect the attractiveness of the main modes; the network assignments however are completely uni-modal. This current practice could be characterized by ’joint uni-modal’ transportation demand modelling.

The next subsections will give a short review of recent proposals for adapting the classical approach to account for multi-modal trips, going from the simplest to the most complicated approaches proposed to multi-modal transport modelling (see Table 2.2).

2.4.3 Multi-modal transport demand modelling approaches

Modelling multi-modal trips centres on the issue of how to model (and represent) traveller’s route and mode choice decisions. According to Fernandez et al. (1994) there are three ways of doing this:

Table 2.2: Characteristics of multi-modal demand prediction methodologies.

Features	Approaches				
	Separate uni-modal (Classic)	Joint uni-modal (Classic)	Multi-modal (Extended Classic)	Intermodal fixed	Intermodal free (Super-network)
Modal networks	Separated	Separated	Separated	Coupled	Integrated
Mode choice modelling	Independent	Joint	Joint	Joint	Absent
Mode/route choice	Sequential before assignment	Sequential before assignment	Sequential before assignment	Sequential	Simultaneous
Combined modes?	No	No	Yes, fixed	Yes, fixed	Yes, free
Mixed modal flows?	No	No	No	Yes	Possible
Equilibrium?	Only in car network	In car network, dependent on PT-shares	In car net., dependent on PT-shares	Simultaneous	Simultaneous
Sources	Ortuzar and Willumsen (1994)	Cascetta (2001)	Fernandez et al. (1994)	Florian et al. (2000) De Cea et al. (2003)	Sheffi (1985) This dissertation

1. All relevant combined-mode alternatives are seen as distinct (and artificial) modes (here called the extended classical approach);
2. All transfer nodes are modelled as a mode alternative (extended classical approach with explicit modelling of transfer nodes and stops);
3. All choice decisions emerge from route choice in an integrated multi-modal network (here called the supernetwork approach).

Along these lines, below we will consider a few significant attempts to develop extensions to the classical approach to handle multi-modal transportation modelling more appropriately. Each proposed approach can be categorized according to the degree of consideration of the interdependencies of travel modes at various levels:

- the modal transport *networks* (e.g. public transport and car networks) are treated separately or are integrated;
- the *modal split* is performed before the assignment (mode and route choice are modelled separately);

- the *combination of modes*, such as car and train, within a trip, are explicitly accounted for;
- the *interactions* between vehicular modes within the same infrastructure, such as e.g. bus and car on the roads, is explicitly accounted for.

These types of approaches together with the uni-modal classical approach are summarized in Table 2.2.

The *separate uni-modal approach* is the classical approach of the last four decades exhibiting complete separation of the modal networks; only pure uni-modal trips are considered which are modelled independently; there is no feedback from the public transport flows to the mode choice (see Figure 2.6 without arrows from the PT network loads to the mode choice). In the more recent so-called *joint uni-modal approach*, there is interdependency between the modes via the joint consideration of flow-dependent modal resistances in the mode choice mechanism; however, only pure uni-modal trips are considered (see Figure 2.6). The *extended classical approach* introduces new artificial separate modes by pre-specifying fixed modal combinations (combined-mode trip alternatives) being combinations of the given main modes. In all other aspects, the modelling follows exactly the same lines as the classical joint uni-modal approach. A further extension of this is the so-called *intermodal-fixed approach* which, apart from specification of fixed artificial combined modes, adopts a high level of interdependency among the modes in trying to achieve a simultaneous equilibrium among all involved modes. Finally, the most advanced category adopts a supernetwork approach defined such that OD-routes in this network may represent arbitrary ('free') uni-modal and multi-modal trips (no fixed modal combinations) where the route choice and network assignment automatically includes modal choices. In this approach the modal sub networks are completely integrated, the combined modes and the interactions among the vehicular modes on the roads might be explicitly taken into account. Below we will discuss in more detail the extended classical and the supernetwork approaches.

2.4.4 The classical approach: joint uni-modal

The classical approach typically uses a mode choice model at origin/destination level, splitting the OD-trip matrix into mode-specific OD-matrices, followed by a network assignment procedure in which often a user-equilibrium is sought. Trips are always uni-modal, meaning that they can only have a single main mode (such as car or train or bike); access or egress modes are not explicitly distinguished.

The advantages of the classical approach are familiarity and simplicity; the disadvantage is a limited ability to investigate the entire spectrum of available modes.

Figure 2.6 shows a scheme of the joint uni-modal approach (omitting some of the non-relevant boxes and feedback cycle 3 from the basic scheme of Figure 2.5) in which

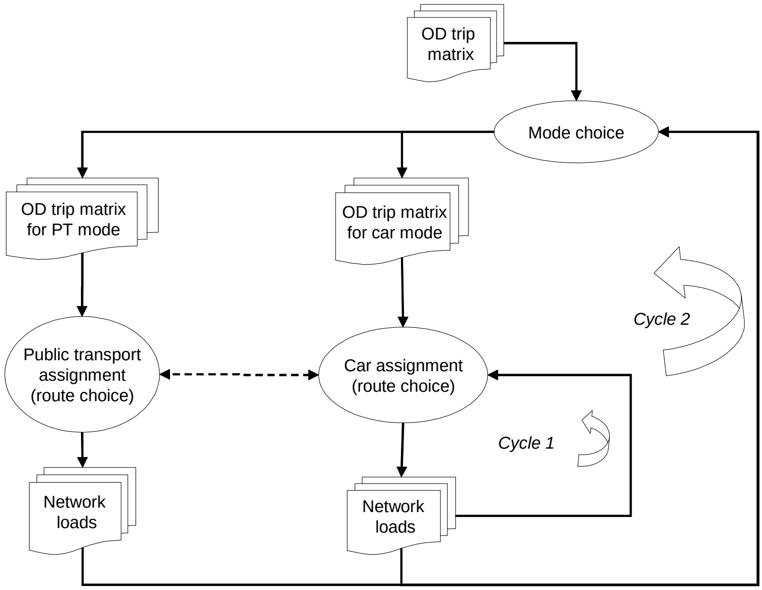


Figure 2.6: Classical travel demand model system: the joint uni-modal approach.

the mode and route choices are modelled separately and sequentially; the assignment and the modal networks are completely uni-modal. This approach is not able to model the choice of combined mode options such as a car and train within a trip, nor the choice of transfer nodes such as Park & Ride facilities. In those classical approaches the feedback in the assignment (cycle 1) is almost always performed, whereas the feedback into the mode choice (cycle 2) and destination choice (cycle 3) can be an option. The dashed line from the car to the public transport assignment box represents the modal flows dependencies.

2.4.5 The extended classical approach

In practice, the classical modelling approach may be extended to handle multi-modal travel by introducing specific pre-specified multi-modal alternatives as *artificial modes* as shown in Figure 2.7, in which as an example the Park & Ride option is specified as a new separate modal alternative additional to the standard main modes. Although the added combined mode trip alternatives consist of two or more main modes, they are considered as an additional main mode, without specifying access or egress submodes.

By introducing an additional artificial mode for each modal combination to be studied,

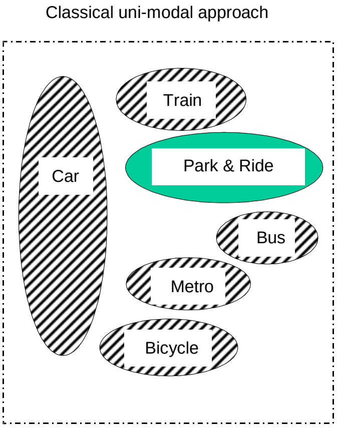


Figure 2.7: Classical uni-modal way to handle multi-modal trips.

the problem can be recast (from a purely formal point of view at least) as a formulation that standard transportation modelling toolkits can handle. This means that the mode choice model now is applied to a wider set of modal options, and the OD trip-matrix now is split up into several mode-specific OD-matrices including matrices for the new artificial modes (see Figure 2.8). The advantages of extending the classical approach to handle multi-modal trips in this way are obvious: practitioners can continue to leverage their standard toolkits, datasets, and skills. The disadvantages emerge upon consideration of the implementation consequences. The mode choice alternatives now include artificial modes that represent multi-modal options. For example, an inter-modal transfer can be modelled quite artificially by introducing a new mode consisting of two travel modes and a transfer at a predefined transfer node in the network. It will be obvious that such an approach might be suitable when only one or two transfer nodes are considered, but that in the case of multiple multi-modal combinations and multiple transfer nodes the number of alternatives will explode. In addition, this has to be done for all relevant OD-pairs in the area.

Florian et al. (2000) report a large-scale application of a multi-class multi-mode traffic equilibrium model in Santiago. The model computes a joint equilibrium between travel demand, mode choice and route choice, where interaction between modes such as e.g. bus and cars is explicitly accounted for. In this advanced model system the model addresses a *fixed set* of possible modes, some of which have two main modes, such as car-subway. Examples of software packages that can handle the modelling approaches previously described (the classical joint uni-modal and the extended classical approach) are Emme/2, VISUM, and the Transims microsimulation package (Los Angeles National Laboratory, 2001).

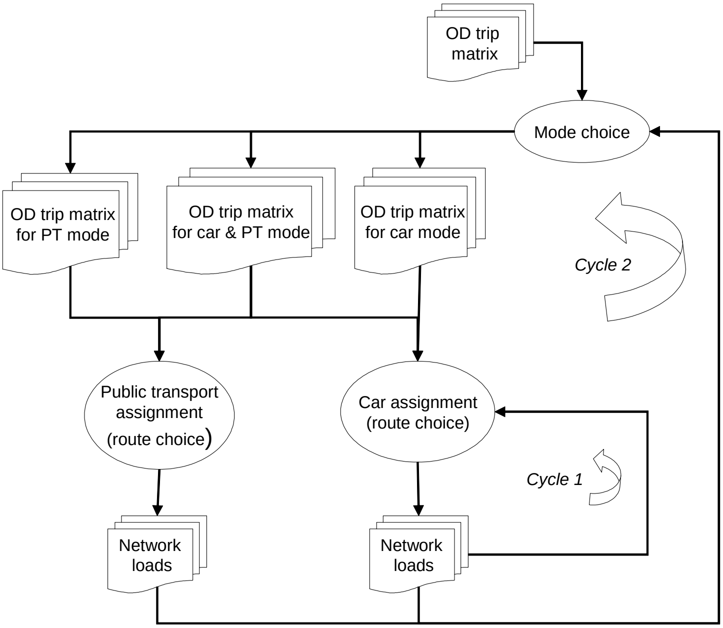


Figure 2.8: Classical travel choice model system: the extended classical approach.

The approaches described above have in common that all available *main* transport modes (uni-modal and multi-modal) considered in the analysis must be predefined for each OD-pair. Mode choices result from a dedicated mode choice model at origin and destination level, based on a network skim for each of the predefined modes. Route choices result from traffic assignment (equilibrium or otherwise) being completely separate from the mode choice model (see Figure 2.8).

2.4.6 The supernetwork approach

A supernetwork as defined in Sheffi (1985) combines several uni-modal networks into a single integrated multi-modal network (see Figure 2.9). At appropriate places, the uni-modal networks are interconnected through transfer links representing the possible transfers between modes. The attributes of these transfer links include times and costs (including those for parking vehicles) on the basis of which transfer disutilities can be derived.

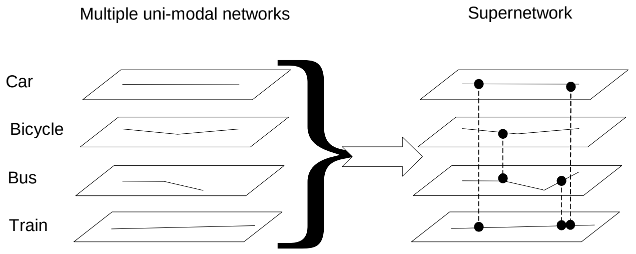


Figure 2.9: Building a supernetwork from single uni-modal subnetworks.

Such a network has the interesting property that *mode choice now is part of route choice*, thus providing a natural unification of the two. In addition, this multi-modal route choice naturally includes the choice of transfer locations, boarding/alighting stops, and possibly the choice of train type (intercity, express, local). So, in a supernetwork, the multidimensional travel choice situation travellers are faced with in reality, is transformed into a one-dimensional choice situation of alternative routes in the supernetwork. In this thesis, we propose and develop this supernetwork approach as a powerful methodology for valid and effective multi-modal travel demand modelling. For illustration, Figure 2.10 presents a simplified example of a choice situation where travellers between an origin and a destination have the main-mode options train-only or car-only and several train-car combinations for their line haul trip part. There are two Park & Ride facilities where a switch from car and train is possible. For their access and egress transportation, several submodes are open. In addition, several boarding and alighting stations may be used, we neglect for the time being the possibility of several train service types calling at different stations. All these options combined result into a choice set of more than ten route alternatives between O and D through the supernetwork including one car-only route, eight different train-only routes, two Park & Ride routes.

In reality, choice situations maybe even more complex due to availability of bus, tram, and metro as access and egress modes, of bus as main mode, of multiple routes for the car mode, etc.

Although concatenation of modal subnetworks seems straightforward, the notion of a supernetwork is not. The main reason is the combination of networks with continuous flow properties (such as car and bicycle) and networks with discontinuous flow properties such as the public transport service networks characterized by regular or irregular service intervals. These crucial differences need to be bridged somehow by an appropriate structuring of the supernetwork and by the specification of appropriate attributes of the constituting network elements among which the transfer links (see Chapter 3).

The biggest advantage of the supernetwork approach is that there is no need to prespecify the available multi-modal options. These emerge in a natural way by searching in the supernetwork for attractive trip opportunities from origin to destination (designing this searching process is central topic of this thesis). Additionally, there is no limit required to the number of trip options (size of choice set). Another advantage of the proposed supernetwork approach is that the rigid separation between mode choice and route choice disappears including the question of their right sequence.

The supernetwork approach thus is an elegant and generic way to model a wide variety of mode-combinations, and allows one to explicitly model transfers as transfer links, and waiting times as waiting links. This is a substantial advantage over the use of separate mode choice models where the choice sets with respect to modes have to be defined rigidly in advance for each OD pair.

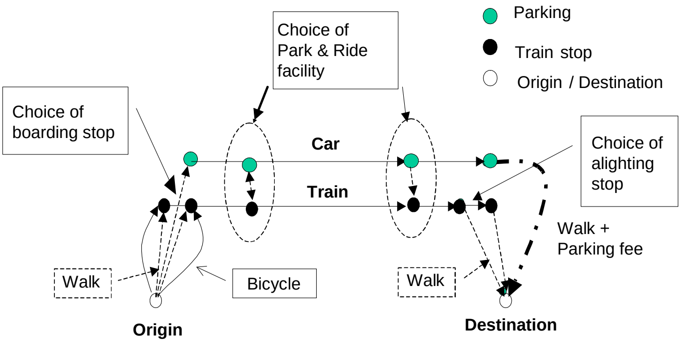


Figure 2.10: Examples of multi-modal routes that consist of route and mode components.

In order to fulfil the requirements of multi-modal transportation planning, that is sufficient consideration of multi-modal trip alternatives in terms of numbers, variety and quality, the route choice modelling applied to the supernetwork should be able to handle this richness and variety appropriately.
This requires first of all a multiple route choice modelling approach that can handle mixed private-public combined-mode route alternatives and the complex composition of choice sets. A pre-condition for that is a separate choice set generation module that generates feasible choice sets from the supernetwork given the attributes of its constituting elements.
A priori and explicit enumeration in a network context not only offers a number of theoretical advantages, such as the application of flexible route choice modelling, non-linear cost functions and route-based attributes; it also offers implementation and computational advantages in iterative network assignment approaches since no repeated optimal route search is necessary.

In the next section, an architecture is proposed for modelling multi-modal travel demand using the supernetwork approach as its basis.

2.5 The proposed multi-modal travel demand prediction procedure

The purpose of the new model system is to analyze and predict travel demand in a multi-modal transport system environment. Although the choice of interchanges between public transport modes has been investigated in the literature, the choice of entry-exit points of the transit network has not been modelled in a general way so far. This can be seen as a lack of classical transport modeling techniques, as suggested in Fernandez et al. (1994). An approach to handle this is essentially to hard-code all possible entry-exit points of public transport as separate modes. As argued in Catalano et al. (2001), this is an ad-hoc approach which is hardly recommendable when it comes to real-world networks. In addition it lacks a systematic and coherent theoretical basis, whereas an elegant unified theoretical basis already exists in the form of supernetwork modeling see Sheffi (1985) and especially Nagurney and Dong (2002).

Therefore, basic elements of the proposed system are the concept of the supernetwork as an integrated description of the multi-modal transport system, a priori choice set generation step, and a route-based travel choice modelling approach, which assumes that a traveller has a set of possible route alternatives available for a specific trip in the multi-modal network (i.e. choice set) from which (s)he chooses the alternative that is perceived as most suited.

Compared to the classical travel demand modelling procedures, the development of this model system poses three specific problems to be solved:

1. How to specify and build a supernetwork including transfer possibilities?

2. How to generate a realistic set of alternative routes in this multi-modal network?
3. How to assign the OD trip demands to these routes?

The assignment of trips in a multi-modal network context has been subject of a parallel thesis project (Hoogendoorn-Lanser, 2005). It requires a route choice model that can handle the prediction of choice probabilities of using routes, and thus also modes, given a multi-modal composition of the choice set. Chapter 5 will discuss the state of the art in route choice modelling as a precursor to the route generation step discussed in the sequel.

The question of how to generate realistic choice sets of routes that account for all realistic combinations of modes among routes and within a route is more complicated in a supernetwork setting than in uni-modal networks because of all kinds of constraints that should be satisfied. This question is the main focus of this thesis (Chapters 6 and 7).

The specification and construction procedure of the supernetwork is the subject of another separate but related research project carried out by the Netherlands Organisation of Applied Scientific Research (TNO) (Carlier, Fiorenzo-Catalano, Schrijver, & Van Nes, 2005). Since the supernetwork specification directly affects the choice set generation step, Chapter 3 will elaborate on this subject in more detail. The development and implementation of the whole multi-modal modelling system (currently called TRANSFER) is a joint development project of TU Delft and TNO (see Carlier et al. (2005)) of which this thesis develops one of the essential building blocks, that is the multi-modal route choice set generation model.

The next subsection describes the architecture of the proposed model system in more detail.

2.5.1 Model system architecture

In this section we will briefly present the system architecture for multi-modal transportation demand modelling in which the choice set generation module to be developed in this thesis constitutes an essential part.

The proposed model architecture circumvents and solves problems of multi-modal travel modelling apparent in the previously presented classical modelling approaches. The proposed architecture is shown in Figure 2.11, which only partly resembles that of classical systems. The main differences are the construction and use of supernetworks, and the a priori multi-modal choice set generation step. A separate modal choice/modal-split step is not needed since this follows from the route choice/assignment step. The new modelling system includes all the elements within the dotted box lines, while the elements outside these lines, such as establishing the trip demand OD matrix, do not differ from the classical approaches. The system is characterized by a number of

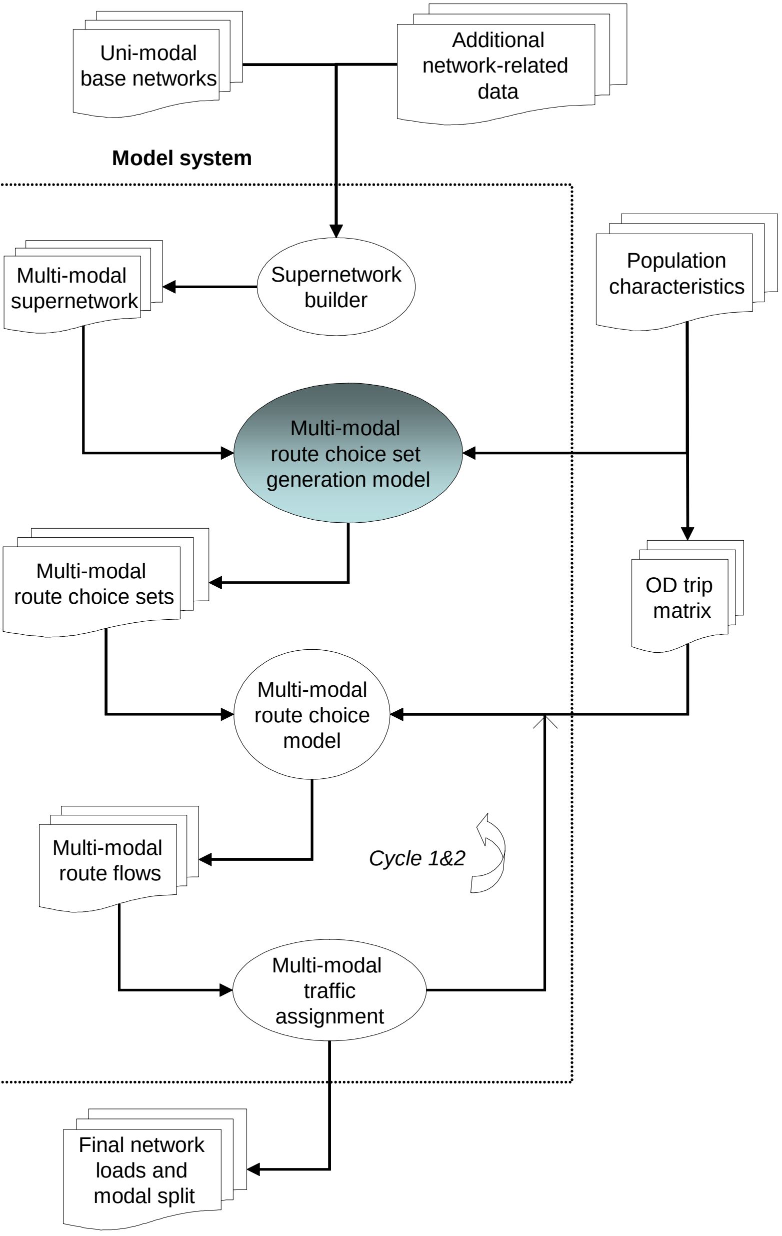


Figure 2.11: Proposed model system architecture for multi-modal transport demand prediction.

possible feedback cycles needed for the correct equilibration of the loaded flow pattern and the resulting trip cost pattern.

Shown on the top are the inputs of the model system. This input consists of network data (top left) and non-network data (top right). Basic inputs are the network descriptions of all available submodes in the multi-modal transport systems including walk, bicycle, car, public transport modes such as bus, tram, metro and train. These network descriptions may come from separate data bases or may already be synthesized in an overall network data base.
The internals of the model system are shown inside the dotted box of which the main components are shown in ovals. The model system consists of four separate models to be applied consecutively, namely the supernetwork builder, the choice set generator, the route choice model and finally the network assignment model. The output (shown bottom left) consists of network loads, travel times and travel costs in the multi-modal supernetwork that easily can be split up by link type and mode, so that the modal split can be derived from the assignment.

The *multi-modal supernetwork* is constructed out of uni-modal networks and additional network-related data by a pre-processor module ”Supernetwork builder”. The additional network-related data consists of items that must be present in the multi-modal supernetwork (e.g. transfer walking links, transfer penalties, waiting time during transfer, parking cost, etc.), but are not part of any of the uni-modal networks. The network builder produces an integrated multi-modal supernetwork with all required transfer links and travel cost attributes. This network is the essential input for each of the three following models, being route generation, route choice and network assignment.

The non-network data consist of the population characteristics and the travel demand (e.g. in the form of an origin-destination trip matrix per population segment). The travel demand matrix shows between which OD pairs there is any demand including for which trip purposes and population groups. The travel demand is assumed exogenous to the model system although consistent prediction of this matrix may require travel resistance input from the assignment outcomes.

Given the supernetwork and the non-zero OD travel demands, the second step is the *generation of route choice sets*. This model produces for each non-zero OD pair a limited number of different routes through the supernetwork that are considered realistic options for travellers. This route set contains uni-modal as well as multi-modal routes. Since the requirements for route options may differ between trip purposes and travellers groups, the generation model should be designed such that it can produce purpose and group specific route choice sets per OD pair. These sets are the input to the route choice model step.

The *route choice step* involves calculation of choice probabilities for each of the routes contained in the given route set for each OD pair. This is done separately for trip purposes and user groups as contained in the trip demand matrix. Each of these categories

may have its own predicted choice sets from the former step as well as its own choice model (utility function). In contrast to most current applications, because of the a priori determined route set, the route choice modelling is fully route based which allows nearly unlimited freedom in the type of choice attributes and the type of choice models to be applied. This is especially advantageous in a multi-modal context with heterogeneously composed multi-modal routes.
Important input to the choice model are various kinds of travel resistances. These are derived from link attributes in the supernetwork, of which the flow-dependent resistances maybe iteratively updated through a feedback cycle from the network assignment step (feedback cycle 1&2).

The calculated choice probabilities combined with the demand levels from the OD trip matrix produce a set of route flows through the multi-modal supernetwork consistent with the route costs. These route flows feed into the multi-modal traffic assignment model. The feedback from the assignment, which might be optional (cycle 1&2), indicates the computation of adapted multi-modal route choice probabilities in the supernetwork based on the modified network flows. The issue is to find a set of multi-modal route flows and resultant multi-modal network loads that are mutually consistent, i.e. resultant route costs satisfy the Wardrop equilibrium conditions.

The final step is calculating the network loads by *assigning* the predicted OD route flows onto the multi-modal network. These loads have a function in their own right in the planning process, but they are also required in calculating flow-dependent resistances such as travel times by car that are input to the choice modelling. In contrast to classical approaches the modal split is one of the results of this assignment step instead of an input by summarizing the loads by link type.
The assignment procedure is simpler than with classical approaches since no iterative optimal route search is needed; this route search has already been performed in the choice set generation step before in a single stroke. The model user is free in adopting a static or a dynamic assignment approach. A dynamic approach leads to better predictions of flows and costs in congested conditions, however at the cost of an additional flow propagation module and more computation time (see for an application Bliemer, M.J.C. and Versteegt, E. and Castenmiller, R. (2004)).

This new way of multi-modal demand modelling offers a number of paramount advantages:

- No explicit specification of multi-modal alternatives is needed; these alternatives emerge automatically from the generation and choice steps;
- The explicit generation of variable sets of feasible alternative routes dependent on OD-pair, population group, and trip purpose enables better route choice modelling and easier trip assignment to the network;
- A wider scope for route choice modelling, enabling the use of route based attributes and great freedom in type of choice models to be applied;

- Simpler network assignment without (repeated) route search, facilitating multi-user class and dynamic flow prediction and analysis;
- Less computation time, because route search need to be performed only once (in the generation step).

Please note that in specific situation the proposed multi-modal demand modelling approach may pose some limitations. For example, in case of heavy congested situation (that are not taken into account in this thesis) adequate checks of the route sets must be applied. Moreover, in case of reliability study, or uncertainty, a different approach for choice sets must be applied.
Finally, the proposed multi-modal demand modelling approach is equally applicable to the static and the dynamic case, which means dynamic in the supply and demand side, i.e. in flow propagation, in departure time choice, and PT time tables. However, since the main focus of this thesis is the generation of choice sets for demand prediction purposes, in which the dynamic aspects are less relevant, in the rest of this thesis we will consider only the static case.

The main topic of this thesis is confined to the choice set generation step, in particular the generation of choice sets in multi-modal networks for prediction purposes in a strategic planning context (Chapters 6 and 7).

Table 2.3: Multi-modal routes in Rotterdam-Dordrecht corridor [Source: Benjamins (2002)] (W is total waiting time in minutes, D total travel distance in km, T total travel time in minutes).

Nr.	Modes					Total		
	Access1	Access2	Main	Egress1	Egress2	W	D	T
1			**Car**				26	35
2			**Car**				25	38
3			**Car**				26	35
4			**Car**	Metro	Walk	2	27	42
5		Bike	**Train**	Bike		6	27	59
6		Bike	**Train**	Bike		6	27	58
7		Bike	**Train**	Bike		6	27	58
8		Bike	**Train**	Bike		6	27	58
9		Bike	**Train**	Bike		6	27	58
10		Bike	**Train**	Bike		6	27	58
11		Bike	**Train**	Bike		6	27	58
12		Bike	**Train**	Bike		8	26	59
13		Bike	**Train**	Bike		8	26	59
14		Bike	**Train**	Tram	Walk	11	26	61
15		Bike	**Train**	Tram	Walk	10	27	61
16		Bike	**Train**	Tram	Walk	10	27	61
17		Bike	**Train**	Metro	Walk	9	25	62
18		Bike	**Train**	Metro	Walk	9	25	66
19		Walk	**Train**	Metro	Walk	17	25	64
20		Walk	**Train**	Tram	Walk	11	26	66
21	Walk	Bus	**Train**	Tram	Walk	12	27	59
22	Walk	Bus	**Train**	Metro	Walk	12	25	61

2.5.2 Implementation of supernetwork approach in practice

A precursor experimental implementation to test the feasibility of the proposed supernetwork approach has been successfully carried out by Benjamins for the Rotterdam-Dordrecht multi-modal corridor in The Netherlands (see Benjamins (2001), and Benjamins et al. (2002)). This precursor implementation involved the first three of the four steps of the model system. The transportation system for interregional trips in this corridor consists of car and train as main modes, Park & Ride facilities for car and bicycle, and walk, bike, bus, tram, metro as access modes to railway stations and Park & Ride facilities. Implementation of the supernetwork builder has been based on the general purpose network management software package NETTER (for details see Chapter 3).
The implemented choice set generation model has been based on a stochastic optimal route search approach based on a link-additive travel cost function of several link attributes such as distance, time, and cost. Different routes are generated by stochastically perturbing the link attribute values in the network. The parameters of the route cost functions (such as value-of-time) differ per user group. The feasibility of route alternatives differs for user groups dependent on their vehicle availability (for details see Chapter 7).

As a route choice model the so-called Paired Combinatorial Logit (PCL) Model has been implemented because of its analytical simplicity in dealing with the complexity of mutually overlapping routes in the choice set (for details see Chapter 5). The experimental implementation by Benjamins used parameters for the route set generation and route choice models taken from literature, without a local calibration. Application of this implementation for the interregional travel in the Rotterdam-Dordrecht corridor proved to be computationally feasible and showed reasonable mode usage outcomes compared to available statistics from the National Travel Survey data.

Table 2.4 gives an example of the route choice prediction outcomes for multi-modal trips (shown in Table 2.3) from Dordrecht-Southeast area to Rotterdam-Northwest area (about 30 km trip distance) achieved with the new model system (source: Benjamins (2001)). The model has been applied to six user groups (see columns 1 to 6). In total, 22 uni-modal and multi-modal route alternatives (including different boarding or alighting railway stations) were generated (see rows 1 to 22) combined over all 6 user groups. Car-only trips appear to be by far the most attractive having the shortest travel times (see column T in Table 2.3). Columns numbered from 1 to 6 show that the generated choice sets differ by user group, showing for example that user groups 5 and 6 do not have car available, while user groups 2 and 6 don't have a bicycle available. The numbers in columns numbered from 1 to 6 indicate the predicted shares of using the respective alternatives available per user group.

Table 2.4: Example results of multi-modal assignment of trips by user groups in % [Source: Benjamins et al(2002)].

Nr.	Modes					User groups					
	Access1	Access2	Main	Egress1	Egress2	1	2	3	4	5	6
1			**Car**			23	50	10	10		
2			**Car**			24	10	14	18		
3			**Car**			19	40	8	8		
4			**Car**	Metro	Walk	8		10	10		
5		Bike	**Train**	Bike		2		2		7	
6		Bike	**Train**	Bike		3		4		13	
7		Bike	**Train**	Bike		4		4		14	
8		Bike	**Train**	Bike		3		3		11	
9		Bike	**Train**	Bike		3		3		12	
10		Bike	**Train**	Bike		3		3		11	
11		Bike	**Train**	Bike		3		3		12	
12		Bike	**Train**	Bike		1		2		4	
13		Bike	**Train**	Bike		1		2		4	
14		Bike	**Train**	Tram	Walk			4	4	1	
15		Bike	**Train**	Tram	Walk			4	7	2	
16		Bike	**Train**	Tram	Walk	1		8	11	3	
17		Bike	**Train**	Metro	Walk	2		14	28	6	
18		Bike	**Train**	Metro	Walk			2	4		
19		Walk	**Train**	Metro	Walk						11
20		Walk	**Train**	Tram	Walk						52
21	Walk	Bus	**Train**	Tram	Walk						14
22	Walk	Bus	**Train**	Metro	Walk						23
Total						100	100	100	100	100	100

Recently, in a joint endeavour of TU Delft and TNO-Inro the preliminary experimental implementation has been superseded by a new multi-modal software package called TRANSFER including all four modeling steps, including network assignment (Carlier et al., 2005). In this version, parameter values of the generation module have been optimized by calibrating them using observed choice sets.

2.6 Conclusion

This chapter focused on main characteristics and definitions of multi-modal transport and on modelling approaches for analysing multi-modal travel choices. The main characteristic of multi-modal transport is that more than one transport mode is used for a single trip and that transfers between travel modes are thus an essential element of

multi-modal transport.
In order to analyse multi-modal planning problems such as the location and design of inter-modal transfer points, the demand modelling need to be capable of analysing and predicting the use of multi-modal trips.

Several modelling approaches have been identified attempting to extend the classical approach to handle multi-modal transportation modelling. One of these approaches is the so-called extended classical approach in which combinations of modes can be specified as distinct artificial modes where the transfers can be modelled quite artificially by introducing a new travel mode consisting of the combination of two modes and a transfer at a predefined transfer node. Such an approach is however suitable only when one or two transfer nodes are considered, but that in the case of multiple multi-modal combinations and multiple transfer nodes the number of new alternatives will explode.

As an alternative concept, the supernetwork concept is proposed as a generic framework for modelling multi-modal networks. A supernetwork is an integration of all relevant uni-modal networks coupled by transfer links between the transport modes. These transfer links represent the possible transfers between modes. In the supernetwork concept, the traditional steps of mode choice and route choice are integrated into a single route choice problem in a multi-modal network. This concept facilitates analysing the impact of new transfer concepts such as transferia, and taking combined modes within a trip explicitly into account.

An architecture for the new multi-modal demand prediction system based on the supernetwork approach has been presented. The supernetwork approach requires accurate modelling of traveller's perceptions of combined modes and transfers. Moreover, in order to better model route choice in the supernetwork, the explicit a priori generation of the choice alternatives is preferred. This not only offers a number of theoretical advantages, such as the application of flexible route choice modelling, non-linear cost functions and route-based attributes; it also offers implementation and computational advantages in iterative network assignment approaches since no repeated optimal route search is necessary. However, in specific situation the proposed multi-modal demand modelling approch may pose some limitations in memory requirements and in case of heavy congested situation.

The proposed multi-modal demand modelling approach is equally applicable to the static and the dynamic case, however, since the main focus of this thesis is the generation of choice sets for demand prediction purposes, in which the dynamic aspects are less relevant, in the rest of this thesis we will consider only the static case.

An experimental implementation and application of the supernetwork concept proved to be computationally feasible and producing reasonable modal usage outcomes.

This thesis will adopt the multi-modal transportation modelling approach based on the supernetwork concept and will focus on one of its modules, namely the a priori choice set generation approach. Specifically this thesis will develop alternative choice set generation approaches and will analyse their empirical validity and performance in practical conditions.

The structure of the remaining of this thesis is as follows. First, attention will be paid in Chapter 3 to the network modelling concepts for uni-modal and multi-modal trip analysis of which the supernetwork concept will receive ample attention. This network concept has to facilitate the proposed approach in route set generation and route choice analysis. In Chapter 4 a start will be made with the route set generation subject by introducing the necessary choice set concepts. This chapter outlines a theoretical framework with conceptual notions about choice sets, and presents empirical figures on route choice sets from the literature. Chapter 5 then gives an overview of route choice modelling, constituting the behavioural context for the route set generation modelling in Chapters 6 and 7. In Chapter 6, we first deal with the route set generation for uni-modal networks. After formulating quality requirements for choice sets, available route set generation procedures will be presented and systematically assessed. Chapter 7 is devoted specifically to the multi-modal case. The specific quality requirements in the multi-modal case are elaborated. Current route set generation approaches for multi-modal networks are reviewed and assessed, after which a new doubly-stochastic generation approach will be presented. The validity and effectiveness of this approach will be tested with observations. The findings and conclusions of the thesis and its prospects for future work will be summarized in the final Chapter 8.

Chapter 3

Uni-modal and multi-modal network representations

3.1 Introduction

After having introduced the new supernetwork approach to multi-modal travel demand modelling in the previous chapter, this chapter focuses on the representation modelling of transport networks, both uni-modal and multi-modal ones. Network models and related path search algorithms are powerful tools for modelling transportation systems. A network model is a simplified mathematical description of the physical network phenomenon in terms of infrastructure elements and transport services on these for the purposes of analysis, design and evaluation of a given transport system. Thus transportation network model representations depend on the purpose for which they are used.
With the *notion of 'network representation'* we mean the specification of the nodes and their connecting links of the network model and their relevant attributes (identification numbers, co-ordinates, distances, capacities, mode types, service types, etc). Representation includes the selection of real-world nodes and links into the network model as well as the addition of artificial links and nodes in order to make the network model suitable for its purposes. The specification includes a classification in link types and node types.

Network modelling needs special attention for many reasons. First of all, a network representation should facilitate the travel demand analysis since the travel costs predominantly are determined by the network. Secondly, more specifically in the context of route choice modelling, the importance of network representation is related to the choice set generation approaches, the appropriate consideration of route overlap, and computational issues. Because of these relationships, network representations are highly determined by the way of demand modelling.

Because of the focus of this thesis, three issues are of special concern in this chapter, namely the specific requirements posed by a multi-modal approach, the requirements stemming from the proposed supernetwork approach, and in particular from the a priori choice set generation step.

Figure 3.1 shows a classification of considered network types among which we first distinguish the uni-modal and the multi-modal networks. In uni-modal networks, two types of transportation services (flow) might be available: the continuous and simultaneous private transport services (car, bicycle, and walk) on the one hand, and the discontinuous and non-simultaneous public transport services (train, tram, metro, bus) on the other. In the case of multi-modal networks, two possible analysis approaches might be applied, the one in which all layers of uni-modal networks are separated from each other (the classical approaches), and the other in which all layers of uni-modal networks are inter-linked (our supernetwork approach). A special challenge of the multi-modal approach now is the representation of the linkages between continuous and discontinuous services.

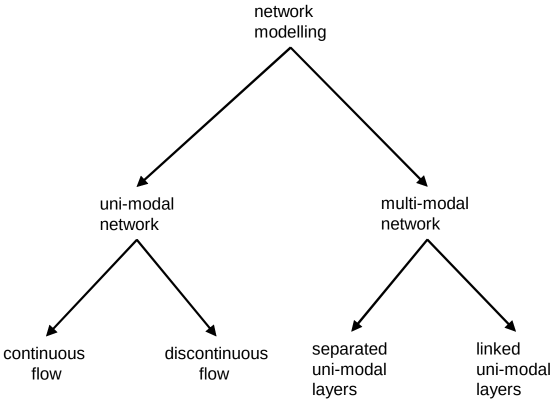


Figure 3.1: Classification of network types.

Questions to be addressed in this chapter therefore are: how may single continuous-flow private networks and single discontinuous-flow public transport networks be represented in a multi-modal travel demand analysis context? Which are the specific requirements needed for a multi-modal network representation in which all these network types are combined? What is the best multi-modal network representation to be suitable for the supernetwork approach (Chapter 2) and the choice set generation approach as proposed in this thesis (Chapters 6 and 7)? Given the several layers of uni-modal networks that constitute the multi-modal network, an important aspect to be analyzed is how the several layers of uni-modal networks maybe linked. What are (if there are) the restrictions per type of mode combinations available in the multi-modal network?

In order to answer these questions, first of all, a general transport network representation and an overview of the main characteristics and definitions of uni-modal and multi-modal network representations from the literature are presented in subsequent Sections 3.2, 3.3, and 3.5. In particular, an overview of uni-modal network representations in cases of private-continuous and public-discontinuous networks (Sections 3.3.1 and 3.3.2 respectively), and multi-modal network representation are presented (Section 3.5).

Second, the requirements needed for an appropriate multi-modal network representation based on the supernetwork approach presented in the previous chapter are derived in Section 3.4, and a precursor approach of our proposed multi-modal representation is presented in Section 3.6.

The proposed approach for multi-modal networks representation is presented in Section 3.7 and the main conclusions are drawn in Section 3.8.

The main contributions of this chapter concern the intelligent comparison of multi-modal network representations presented in Section 3.5.3, and then our proposed supernetwork representation approach with a better and improved multi-modal network structure, presented in Section 3.7.

3.2 Transport network representation

The basis of transportation network models is a *graph model* (for details see Cascetta (2001)). A *graph* is defined as an ordered pair of sets: N, the finite set of elements known as nodes, and L a finite set of pairs of nodes belonging to N, known as links. Symbolically, a graph G can be represented by $G = (N, L)$. The graphs used to represent transportation services are generally oriented; i.e. the links have a direction and the node pairs defining them are ordered pairs. A link connecting the node pair (i, j), can also be denoted by a single index, say l, representing its position in the list of all the links of the graph or by the pair of indices (i, j), relative to the initial and final nodes of the same link. Attributes such as cost, length, speed, or capacity may be associated with each link (Figure 3.2).

The network graph can also be represented by a so-called *node-link incidence matrix* indicating with zero-one variables whether a particular link directly connects to a particular node.

Nodes correspond to significant events delimiting the trip phases (links). Nodes can correspond to points with different space and/or time coordinates in which the event, represented by the nodes, occur. Important nodes in transportation graphs are the so-called *centroid nodes*. They represent the beginning and/or the end of individual trips.

A graph usually includes *links* of different types: *connectors* and *real links*. Connectors are introduced into the network when centroid nodes do not correspond to a physical

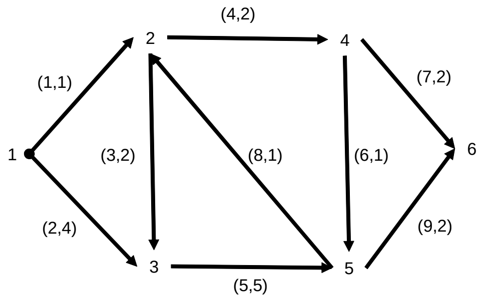


Figure 3.2: Graph model with 6 nodes and 9 links; (l, c) represents the link index and the link cost associated with each link.

element. These links represent the phase between the terminal point (zone centroid) and a physical element of the network. Real links represent phases and/or activities of possible trips between different traffic zones. Thus, a link can represent an activity connected to a physical movement (e.g. covering a road) or an activity not connected to a physical movement (such as waiting for a train at a station). Links are chosen in such a way that physical and functional characteristics can be assumed to be homogeneous for the whole link (e.g. the same average speed). In this sense, links can be seen as the partition of trips in segments of equivalent characteristics; the level of detail of such partition can clearly be very different for the same physical system according to the objectives of the analysis.

In so-called *synchronic networks*, nodes are not identified by a specific time coordinate, and the same node may represent events occurring at different moments (instants) of time. For example, the different entry or exit times in a road segment, an intersection, or a station, may be associated to a single node, representing the entry/exit events.
In *diachronic networks*, also known as *time-space networks*, on the other hand, nodes may have an explicit time coordinate and therefore represent an event occurring at a given instant. The graphs considered in this section are synchronic, diachronic graphs for scheduled services are described in more detailed in the next section. Although our developed supernetwork approach is applicable to both types of chronic networks, the modelling approach as presented in this thesis deals with static networks only.

In a graph representing transportation supply, a *path*, k, is a sequence of consecutive links connection the initial node (path origin) and a final node (path destination). Thus a path is a sequence a trip phases. Usually, only paths connecting centroid nodes are considered in transportation graphs. These paths are sequences of phases allowing travel from a given origin to a given destination and therefore represent possible trips. On this basis, each path is unambiguously associated with one and only one OD pair, while several paths can connect the same *OD* pair. An example of graph with different paths connecting the centroid nodes is shown in Figure 3.3.

Note that we may distinguish with the term *path* a collection of links (representing streets, walkways, railways, etc.) and nodes (representing intersections, stations, etc.) from origin to destination in a *model* network, and with the term *route* a collection of links and nodes from origin to destination in a *real* network. However, in the rest of this thesis the two terms, route and path, are used as synonyms.

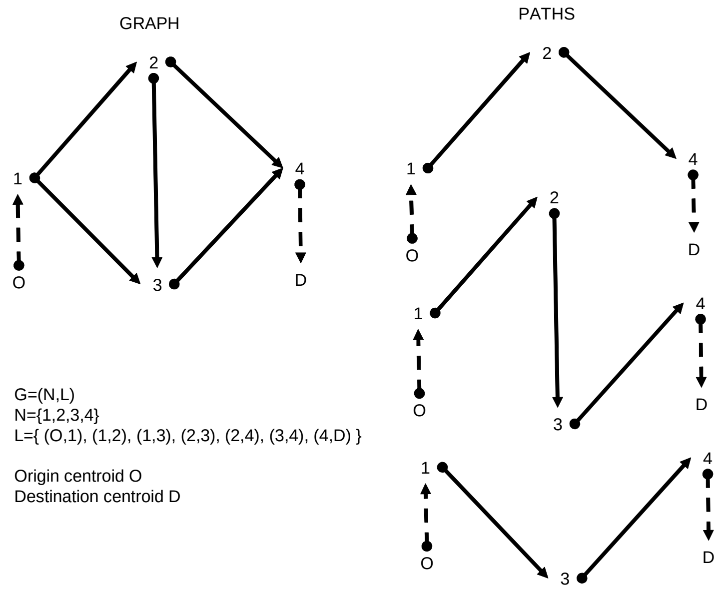


Figure 3.3: Examples of paths connecting centroid nodes from 1 to 4.

3.3 Network representation overview for uni-modal networks

This section provides a general formulation of transportation network models, focussing on continuous systems (such as car and bicycle networks) and scheduled-based or run-based systems (such as transit networks).

A transportation network consists of a finite set of nodes N, a finite set of links L, consisting of a set of pairs of nodes belonging to N, and a finite set of attributes associated with nodes and/or links. An example of a link attribute might be the length of a link; another example might be the link cost or the link travel time. An attribute is not necessarily a directly observed quantity; it might be any function of available data.

3.3.1 Private transport network representation

Continuous and simultaneous services are typical for private transport services. These are available at every instant and can be accessed from a very large number of points. Typical examples are individual travel modes such as car drivers, bicyclists and pedestrians using road systems.

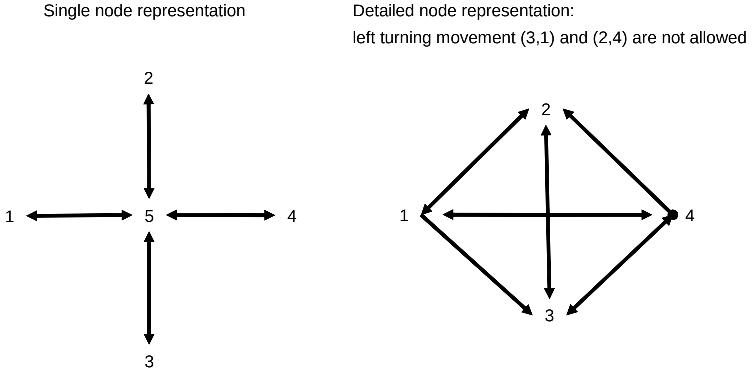


Figure 3.4: Graphs for a road intersection.

In graphs representing road systems, nodes are usually located at the intersections of road segments included in the network model. Nodes can also be located where significant variations of the geometrical and/or functional characteristics of a single segment occur (such as changes in a road cross-section and lateral friction). Links usually correspond to connections between nodes allowed by the circulation scheme. Therefore, a two-way road will be represented by two links going in opposite directions, while a one-way road will have a single link going in the allowed direction.

The level of detail of the road system depends on the purpose of the demand analysis. This is especially true for road intersections. In a coarse representation, a road intersection is usually represented by a single node in which the connected links join. Alternatively, it is possible to adopt a more detailed representation that distinguishes different turning movements and excludes non-permitted turns or turn penalties (if any). Such a representation is obtained by using a larger number of additional nodes and links, with an extended network representation. Figure 3.4 shows two possible representations of a four-arm road intersection. Note that in a single node representation, paths requiring the left turn cannot be excluded if this turning movement is not allowed; furthermore, different waiting times (delays) cannot be assigned to manoeuvres with different green phase durations, such as the right turn. Both of these possibilities are allowed by the detailed representation.

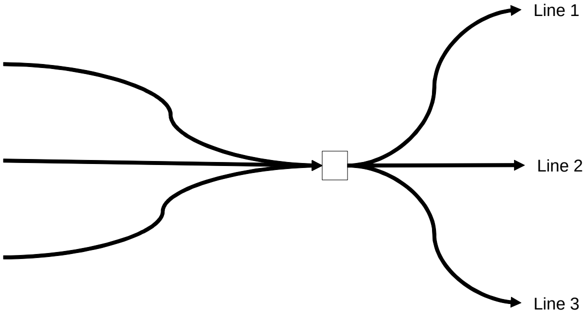


Figure 3.5: Simple line-graph representation of a public transport system.

3.3.2 Public transport transport network representation

Discontinuous and non-simultaneous transportation services are typical for public transportation networks. They can be accessed only at a limited set of given points and they are available mostly only at a limited set of given instants. Typical examples are scheduled services (buses, trains, airplanes, etc.), which can be used only between terminals (bus stops, stations, airports, etc.) and are available only at certain instants (scheduled departure times). Scheduled services can be represented by different supply models according to their characteristics and to the consequence assumptions on users' behaviour.

Public transport or transit services have two characteristics that make modelling public transport networks more difficult than private transport modes such as car, namely the time dimension, i.e. frequencies and schedules, and the concept of lines and thus the need for transfers. The adopted network representation is a balance between network size, network complexity, and algorithmic complexity. Furthermore, the purpose of the planning study itself has a significant influence on the way of network modelling. For short-term design-oriented studies a higher level of detail might be attained and required than for long-term planning studies.

In general, supply models for public transport services consist of a network model (graph plus cost functions) and a set of relationships connecting link costs to path costs and link flows to path flows. In order to represent public transport networks, two different modelling approaches can be used (see e.g. Cascetta (2001)): the first one refers to services represented in terms of lines (line-based supply model), whereas the second one considers services represented as single runs (run-based supply model), as described in the following subsections.

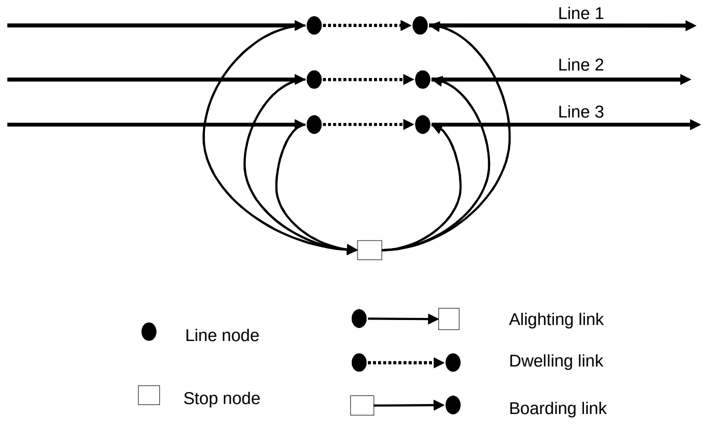


Figure 3.6: Expanded line-based graph for public transport system (taken from Cascetta, 2001).

Line-based supply models

In the line-based, or headway based approach the transit network is represented by a graph in which services between adjacent stops are represented by links between these stops. Every service between stops is represented by a different link, but no distinction is made between the runs of the service. The waiting time at transfers is approximated by an average waiting time (which is a function of the inverse of the service frequency). Stops are represented by a set of nodes and links. The links at the stop represent boarding, alighting or transfers between services.

If the scheduled services have high frequencies (e.g. one run every 5-15 minutes) and low regularity, it is usually assumed that the users do not choose an individual run, but rather a service line or group of lines. A *service line* is a set of runs sharing the same terminals, same intermediate stops, and the same performance characteristics, as in the case of urban bus or underground lines. In this case, a *line graph* is typically used (Figure 3.5). It consists of three public transport lines arriving at and departing from a stop node. In order to correctly model transfers and line choice, such a public transport network might be represented in more detail by an expanded line-based graph (network model). Figure 3.6 shows the expanded stop node of the stop node from Figure 3.5. In this graph, nodes correspond to stops, and more precisely to the relevant events occurring at the stops: *line nodes* represent arrival and departure of vehicles of a given line at a given stop, *stop nodes* represent the boarding to and alighting from a given line at a given stop. The links represent activities or phases of a trip: boarding to the vehicles of a line (*boarding links*), alighting from the vehicle of a line (*alighting links*), and vehicle dwelling at the stop (*dwelling links*). In order to represent the transfer disutilities as accurately as possible, each boarding link is assigned a waiting time at the stop as a function of the frequency of the line, while each alighting link may be assigned some line-dependent constant cost.

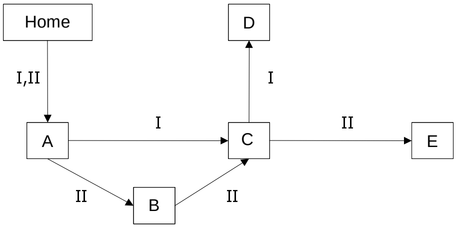


Figure 3.7: Transit network according to Dial (1967) using trunk line links consisting of 2 transit lines (I, II) (taken from Bell, 2003).

The following transit network representations are also discussed in Bell, M.G.H. (2003).

Dial, R.B. (1967) proposed a network structure in which the transit network is represented by "trunk line links" (Figure 3.7). These links have as attributes the travel time and the line numbers using the link. The typical transit modelling problem can be illustrated by the fact that the shortest route to C depends on the final destination. For travelling to C line I might be interesting, while for trips to D line I is the obvious choice. For travelling to E, however, line II is the best choice. Dial adapted the shortest path algorithm of Moore in order to account for transfers. The transfer penalty depends on the expected waiting time of the line that will be boarded. The waiting time is usually defined as half the headway.

An alternative approach is presented by Fearnside and Draper (1971), in which line-specific nodes and links are introduced (Figure 3.8). This approach requires transfer links between stops related to the same physical location. A problem that is not dealt with, however, is the so-called common lines problem that occurs when a traveller waiting at a stop might use different lines for reaching his destination. The traveller then has to decide which line to use. This might be a specific line, yielding the shortest travel time from the stop to the destination, or it might be the first run that arrives. In this example, for instance, a passenger travelling to A or to C might choose between lines I and II. An approach for dealing with the common lines problem is discussed in the paragraph on

Another approach for representing transit networks was suggested by De Cea and Fernandez (1993) (Figure 3.9). In their model they introduced "direct links" connecting two stops that might be reached by a line, with or without transfers. In the case of multiple lines serving the same pair of stops, an artificial aggregate direct link might be created. Consequence of this approach is that the network size increases enormously. In particular, the approach in De Cea and Fernandez (1993), applied also by Benjamins et al. (2002), requires for a PT network consisting of v PT lines each having u_v stops additional $\sum_v \frac{u_v(u_v-1)}{2}$ of travel and transfer links. On the other hand, the approach proposed by Fearnside and Draper (1971) uses for each PT line v having u_v stops additional $(u_v - 1)$ alighting links, $(u_v - 1)$ boarding links, and $(u_v - 1)$ travel links, in total for the complete PT network additional $\sum_v 3(u_v - 1)$ links.

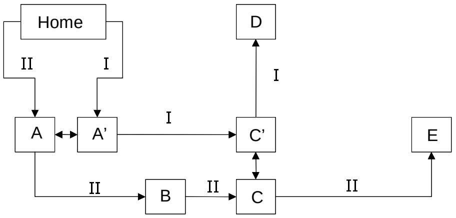


Figure 3.8: Transit network according to Fearnside and Draper (1971) with line-specific nodes and links (taken from Bell, 2003).

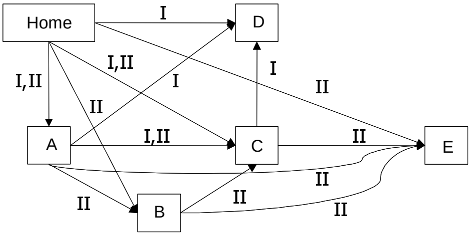


Figure 3.9: Transit network according to De Cea and Fernandez (1993) with route sections (taken from Bell, 2003).

Therefore, the number of transfer and travel links in Fearnside & Draper's approach increase linearly with the number of PT lines and stops on each line, whereas the number of transfer and travel links in De Cea & Fernandez's approach increase linearly with the number of PT lines but are quadratic with respect to the number of stops on each line (see Table 3.1).

Table 3.1: Numbers of additional travel and transfer links for different PT network representations.

	Fearnside and Draper (1971)	De Cea and Fernandez (1993)
Transfer and travel links	$\sum_v 3(u_v - 1)$	$\sum_v \frac{u_v(u_v-1)}{2}$

v = number of PT lines; u = number of stops of line v.

The line-based network approaches are mainly suited for static assignment algorithms which calculate the average system performance and average line occupancy. More advanced network algorithms (using database information) can be applied on a line-based network in order to obtain exact waiting times, enabling this representation to be used for modeling line occupancies or for dynamic assignment. Applying a path-based route choice algorithm allows non-linear costs and transfer feasibility checks. This approach combines the efficient line-based network storage with advanced modeling opportunities at the cost of more complex network algorithms.

Run-based supply models

All transit network approaches discussed above deal with lines having frequencies. A more detailed approach is proposed by Nuzzolo, A. and Russo, F. (1994), in which the individual runs are the basic components. They explicitly account for the time dimension, as can be seen in Figure 3.10. The main advantage is that this approach makes it possible to model transfers properly, which is certainly relevant for low-frequency networks.

The *diachronic network* is a network expanded with a time dimension (every node in a diachronic network has a time coordinate) in which the runs of the transit services are represented both spatially and temporal. Nodes in the diachronic graph represent events such as the departure or arrival of a run, not only spatially but also in time. Exact waiting times (assuming the timetables are reliable) at transfers are obtained from the link connecting the boarding or arrival node at the stop with the departure node. A shortest path in a diachronic network is the path that yields the earliest arrival time given the time instant of departure. An application of this approach for multi-modal choice set generation is given by Hoogendoorn-Lanser (2005), summarized in Section 3.5.3. The diachronic graph is useful for dynamic assignment models, for calculating vehicle loadings and evaluating time tables. Standard algorithms can be used, however at the cost of an expanded network.

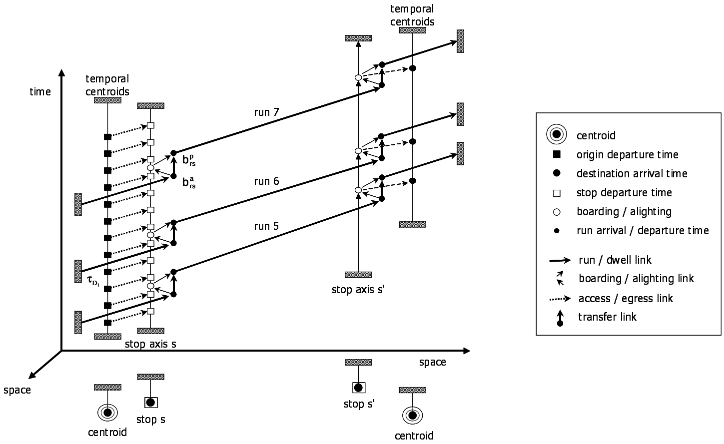


Figure 3.10: Diachronic run-based representation of transit services (taken from Nuzzolo, 2003).

Hypernetworks and strategies

The main difference between the two public transport network representations (line-based and run-based models) presented in the previous paragraphs concerns the representation of the time component: while in the first case the frequencies of the public transport lines are taken into account, in the second case the time table is the basis. From a representational point of view of the graph model, in both cases the theoretical graph model discussed in Section 3.2 is applied, i.e. the graph consists of a finite set of nodes and a finite set of links that connect two nodes.

Another public transport network representation based on a different graph theory is the hypernetwork. A hypernetwork is an extended network in which multiple outgoing links are represented by only a single link, the so-called hyperlink, which is a link that can connect more than two nodes; more details about hypernetwork can be found in Gallo et al. (1993).

The advantage of the hypernetwork representation is that it can capture the so-called strategy concept (Spiess & Florian, 1989). They assume that users travelling in a transit network adopt a particular choice strategy to capture prevailing service conditions when going from a given stop to another one. Users are assumed to consider a set of competitive lines from all lines serving the origin stop, in such a way that they can travel towards their destination by boarding one of the selected lines depending on encountered service conditions. The attractive set, represented by all selected lines, defines the users' strategy since travellers will board the first transit line of the attractive set departing from the origin stop.

Therefore, the hypernetwork concept can be exploited for networks in which paths for user's trip cannot be defined ”a priori” but depend on the state of the network. For example, the hypernetwork framework has been applied to model user behaviour in transit networks (see Nguyen and Pallottino (1988)) and in capacitated road networks (Marcotte & Nguyen, 1998). The behaviour of the user travelling on a transit network can be modelled as a hypernetwork in which hyperlinks represent the boarding links from a given stop to the transit lines servicing that stop and choice alternatives defined by the attractive lines can be represented by a hyperpath (see also Chapter 5). Figure 3.11 shows a hypernetwork in which the hyperlink models the boarding link on the lines belonging to the attractive set.

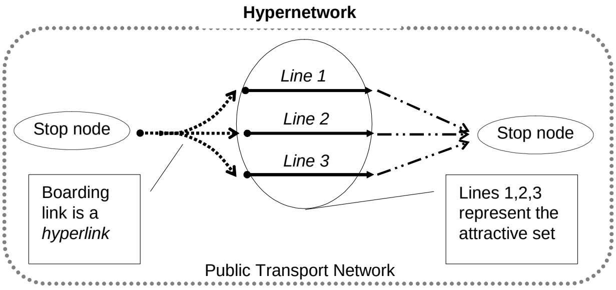


Figure 3.11: Example of a hypernetwork representation of a transit trip.

3.4 Requirements for a multi-modal network representation

The requirements posed on the network representation of a multi-modal supernetwork first-of-all relate to all four main components of the proposed supernetwork approach in which the network plays a role (see Figure 2.11), namely:

1. the supernetwork building process,
2. the choice set generation,
3. the route choice modelling,
4. and finally the presentation of the assignment outcomes.

Of course, the supernetwork should match the original uni-modal networks as closely as possible, allowing a clear representation of the network for sake of network data management and for visualization purposes. In addition, an efficient computation and handling of the whole analysis process poses requirements on an efficient representation.

The specification of the supernetwork should be such that its construction from the basic uni-modal networks can be performed easily and more or less automatically. This is especially true for the addition of the transfer links mutually connecting the uni-modal networks as well as for the addition of other artificial or notional links such as boarding and alighting links.
The *choice set generation* process requires from the network specification that feasible uni-modal and multi-modal individual paths from origin to destination can be determined efficiently with path search algorithms. This implies that link attributes (such as time and distance) should be additive, and that the properties of the added artificial links should be specified in line with those of the real-world links. It implies that each network link may be part of a route, and that the generated routes can be represented in a so-called link-route incidence matrix (sometimes also called assignment map).
The *route choice modelling* requires that the properties of the generated routes can be easily determined, be it by addition of link attribute values such as times and distances, or otherwise such as sometimes needed for cost variables that maybe non-additive such as public transport fares. In particular the network representation should allow maximum flexibility and freedom in the route choice models to be adopted. In addition, the representation should facilitate easy determination of the mutual overlap among routes since this route property significantly influences the quality of the route choice modelling outcomes. For a sufficiently accurate modelling of the route choice behaviour of travellers each distinct travel resistance (or disutility) component in the transportation process between origin and destination should be given a separate link. This implies for example that the transfer from one vehicle to another need to be represented by several

artificial links (representing respectively the transfer components alighting, walking, waiting, boarding, etc) in order to correctly account for the various involved disutilities.
Finally, the specification of the network model should facilitate easy establishment of a rich variety of *assignment outcomes* for evaluation purposes such as modal split figures by user class, or the usage of particular sub-modes or transport services. This may require the specification of specific link types and link attributes in the network model, even if these do not play a role in route generation or route choice.

A transportation network may have different levels of complexity, depending on the system being represented and the details needed for its representation. On the one hand, for short-term projects, such as road circulation plans, or design of transit lines, a very detailed representation of the real system is required. On the other hand, for strategic or long-term planning projects usually a larger-scale network with less detail is sufficient because of the geographical size of the area and the number of elements included in the system. As we have seen in the previous sections, different network models might be associated with the same basic network, depending on the aim of the demand modelling.

Therefore, an aspect that plays an important role in the network representation is network size. The required network size depends on the planning issues at hand; more detailed networks are required for short-term design projects, whereas for long-term planning projects less detailed and less complicated networks may suffice. The network size is a crucial point even for a practical point of view. A very complicated and difficult network structure is not likely to be applied in a practical and feasible application. Moreover, another aspect related to the network size is the computational issue. The more complicated and complex is the network representation the more computational time is required in the analysis. For example, in the case of a multi-modal network representation, a large size of the single uni-modal network layer and complex network structure for generating the multi-modal network requires extra work both from an implementation and computational point of view. Consequently, with regard to our model approach for building the multi-modal network and for generating choice sets in that multi-modal network, the requirements needed for the network representation consist of maximum simplicity and clearness in network representation.

Other requirements consist of maximum flexibility in choice modelling, especially in transfer modelling in the case of multi-modal networks. The network structure should not have any influence on the type of routes that might be generated and available on the network. Moreover, due to the choice set composition (to be analysed in Chapter 4) in which a large variety of mode combinations may appear, no restrictions on mode combinations (e.g. no limitations on number of transfers) or path composition should be put for the network modelling.

For example, a relevant characteristic of multi-modal networks is the transfer from one travel mode to another, e.g. transfer from train mode to car mode. In a good multi-modal network structure no limitations about the type of allowed transfers should be applied. Indeed, given the previous example, the transfer from train mode to car mode might be not allowed in the case of home based trips (trips from home to work in which the car mode might be assumed available only at home side), however in the case of return trips from work to home car mode might be available and therefore such transfer might be assumed also feasible. Due to those reasons the multi-modal network structure should not put any types of constraints a priori; the types of possible paths in the multi-modal network and the type of mode combinations constraints are not be considered in the phase of the definition of the network structure beforehand, but it depends on the aim of the application and the demand model. The same type of network structure should be taken into account for different types of applications.

As stated in Chapter 2, transfers are essential parts of a multi-modal trip, and walking is always part of a trip at the beginning and at the end of the trip, especially for each transfer from one transport mode to another. Therefore in the multi-modal network representation special attention has to be paid to the pedestrian network that should be seen as the generic link among all uni-modal network layers.

3.5 Network representation overview for multi-modal networks

The previous sections focused on uni-modal networks, private and public transport ones. This is because in traditional transportation modelling the choice between modes is dichotomous, that is, travellers can choose only transit or private modes. In reality, however, many travellers combine modes while making their trip, especially when using transit. For instance, 80% of the train users also use another mode than walking to reach their destination, for instance a combination of private car and train (Van Nes, 2002).

This section presents an overview of multi-modal network representations available in the literature for modelling inter-modal trips. Route search in an inter-modal network yields not only the chosen paths, but also the modes used and the locations of boarding, transferring, and alighting. We focus here on the specific inter-modal aspects of these approaches.
As can be seen from Figure 3.1 showing a classification of the possible approaches for uni-modal and multi-modal network representations, in the case of multi-modal network two possible approaches might be applied. First, the one in which all layers of uni-modal networks are separated from each others, which is presented in Section 3.5.1, and the other in which all layers of uni-modal networks are linked to each other, which is discussed in Section 3.5.2. A comparison of the presented approaches with respect of the requirements for a good multi-modal representation as discussed in Section 3.4 is presented in Section 3.5.3.

3.5.1 Multi-modal network representation with separated layers

Based on the classical modelling approach, the multi-modal route choice and assignment models proposed by Montella et al. (1999) adopt the use of two distinct models: one route choice model for each mode, and another one for the mode choice. In this case, the supply model of the multi-modal network is represented by two (or more) non-connected uni-modal networks, to each of which the respective share of demand obtained by modal split is assigned. Figure 3.12 (taken from Montella et al. (1999)) shows a graphical representation of the multi-modal network approach proposed by Montella et al. (1999) in which several planes can be identified, one for the origin, destination and modal diversion nodes, and one for each transportation mode. In this approach, centroids are replicated at each network layer and no multi-modal routes can be generated. All routes are in effect uni-modal. For this reason the approach developed by Montella et al. (1999) might not be defined as multi-modal as we do in this thesis.

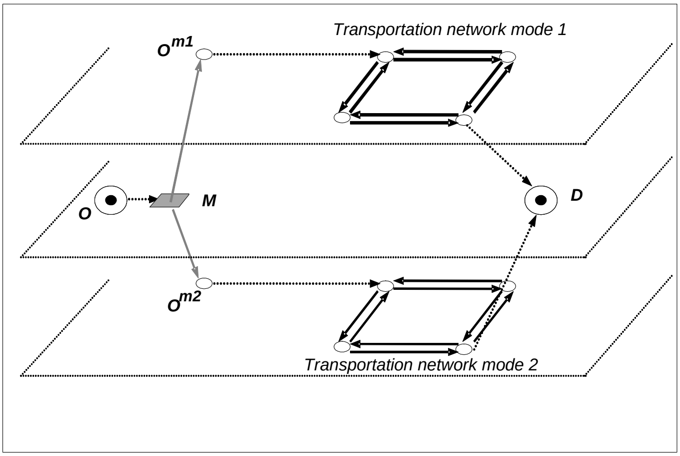


Figure 3.12: Example of a separable multi-modal network representation of the Montella approach.

3.5.2 Multi-modal network representations with linked layers

A supernetwork as defined in Sheffi (1985) combines several uni-modal networks into a single multi-modal network (see Figure 3.13). The supernetwork, which is also defined in Chapter 2, is a concatenation of uni-modal networks interconnected by transfer links in which mode choice and choice of access/egress/transfer location are captured as route-choice in the supernetwork. Transfer links are used to model the cost and/or restrictions of choosing or transferring between the alternatives. Such a network has the interesting property that mode choice can be seen as a special form of route choice, thus providing a natural unification of the two. The advantage of the supernetwork approach is on the one hand that the rigid separation between mode choice and route choice disappears, and on the other hand that route choice models automatically take account of the network structure in defining the choice set and calculating the choice probabilities of the alternatives. The supernetwork approach is an elegant and generic way to model a wide variety of mode-combinations, and allows one to explicitly model transfers as transfer links, and waiting times as waiting links. This is a substantial advantage over the mode choice models, where the choice sets with respect to modes has to be defined in advance.

The field of freight modelling have used supernetwork for multi-modal networks representation, some examples are given by Pedersen, M.B. (2005), Oark, M. and Regan, A. (2005), Loureiro, C.F.G. and Ralston, B. (1996), and Jourquini et al. (1999).

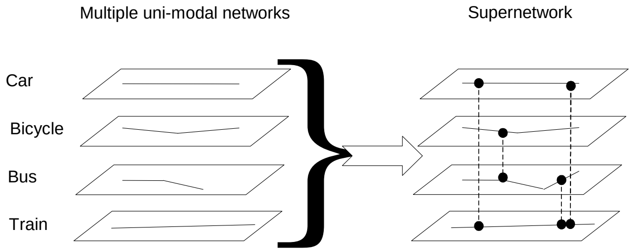


Figure 3.13: Supernetwork principle.

Abdelghany, K.F.S. (2001) represents the multi-modal transportation network as a graph $G(N,L)$ decomposed into a number of layers such that each layer represents a sub-network $G_m(N_m,L_m)$ of a certain mode $m \in M$ as shown in Figure 3.14. Each sub-network G_m consists of a finite set of nodes N_m and a finite set of links L_m. Each link is represented by an upstream node i_m and a downstream node j_m. Each transfer between two modes is represented by an artificial link. The upstream node of each artificial link lies in the sub-network of the mode from which the transfer occurs, while

the downstream node lies in the sub-network of the mode to which the transfer occurs. A list of attributes is associated with each link; attributes representing the different travel attributes are associated with travelling on this link. Artificial transfer links may have different cost elements associated with the transfer movement they represent. In this type of multi-modal representation, centroids for origins and destinations may belong to each of the network layers. Moreover, transfers from one travel mode to any other travel mode may occur for any possible combination, although for illogical and non-feasible mode combinations transfer links are not included in the network.
On the one hand, the number of transfer links might be very large due to the number of possible mode combinations available in a multi-modal network, and the network size could increase enormously. On the other hand, the constraints on mode combinations have to be defined in advance before taken into account the type of application and demand model to be applied. In this case, the network should, theoretically, be changed at each different application purposes (such as type of trips to be modelled, trip purposes, etc.).

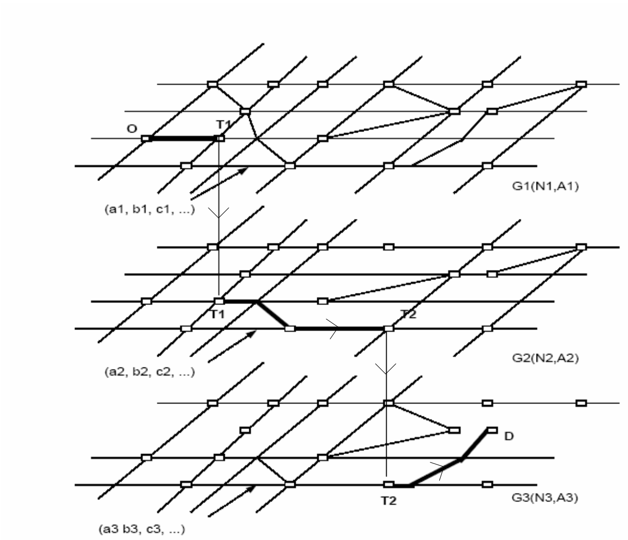


Figure 3.14: Interconnected multi-layer representation of a multi-modal network (taken from Abdelghany, 2001).

In a similar way, Lozano and Storchi (2001) define a multi-modal network as a set of uni-modal networks, one for each travel mode, interconnected by transfer links. The authors also apply a hypernetwork representation (see Section 3.4) for the public transport network. They define a state information system to avoid the generation of illogical multi-modal paths and have limitations on mode combination in the paths generated through the multi-modal network. In this case, as in the case of Abdelghany's multi-modal network, the network structure is defined a priori and it is dependent on the type of the application to be applied. Moreover, the mode combinations that are considered illogical and unfeasible have to be defined in advance, although it is often difficult to know in advance the type of multi-modal paths to be considered, which often depends on the type of application. Therefore, in this case, the network need to be re-defined and re-build each time a different application is applied.

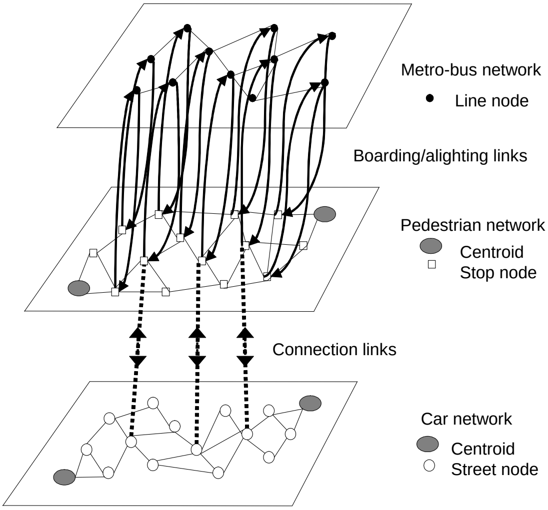


Figure 3.15: Multi-modal network representation (taken from Caramia and Storchi, 1997).

Figure 3.15 shows the representation of the multi-modal network model according to Caramia and Storchi (1997) applied for evaluating the effects of parking prices. The multi-modal network represents four different modes (car, bus, metro, and walk). The boarding and alighting links between pedestrian and PT networks are uni-directional, whereas the connection links between car and pedestrian mode are bidirectional. The authors aim to demonstrate that the use of Park & Ride facilities, as opposed to on-street parking, is more economical, since the price for on-street parking increases reason why car users might find it more attractive to park their cars in a Park & Ride facility and continue their trip taking one or more transit modes.

One characteristic of this multi-modal representation is that the centroid nodes may belong either to the walk or to the car network. The centroids are replicated also in the car network because the authors consider that origins and destinations are very close to parking places due to the evaluation of the parking price effects. The authors pose also restrictions to the composition types of multi-modal paths allowed in the network, assuming that a multi-modal path only can start or finish with the walk or car mode.

Lo, H.K. and Yip, C.W. and Wan, K.H. (2003) propose a third way of multi-modal network representation to overcome two difficulties encountered with traffic assignment in a multi-modal network. The number and kind of transfers are explicitly considered and nonlinear fare structures are accommodated in the network modelling adopted. The authors transform the multi-modal network through a state augmentation technique to a single network, called State-Augmented Multi-modal (SAM) network. The probable transfer rules as well as non-linear route fares or utilities are automatically captured in the SAM network.

Every node in the SAM network is represented by four state variables: location i, transfer state ts, number of prior transfers n and alight or board indicator l (0/1). The transfer state ts indicates the series of consecutive modes that have been used to travel from the origin to the node considered. According to the example in Figure 3.16, a node with transfer state 4 is reached by a subway ride followed by a bus ride. Note that in an inter-modal network, the number of transfer states exceeds the number of modes.

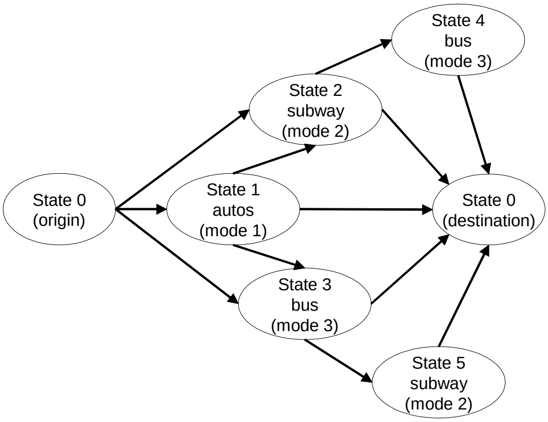


Figure 3.16: Probable transfer states (taken from Lo et al, 2003).

The transfer states list the valid mode-combinations in the model. Every node in the base network is represented $|ts|$ times in the SAM network, where $|ts|$ is the number of transfer states by which this node can be reached. At locations where more than one mode is available, transfer links connect the modal sub-networks if the transition from

mode A to mode B corresponds to a valid transfer state.
After an inter-modal network has been transformed to a SAM network, it can be considered as a simple network with single valued link costs without the need to care for transfer feasibility and non-linear fares. On the other hand, a route in the SAM network automatically combines the mode-transfer choices, which can be decoded for the specific modes used and transfer locations selected.
The transfer states list the valid mode-combinations in the model. This implies a loss of generality (all mode-combinations that are allowed for have to be defined in advance). The chosen set of allowed mode combinations can be incomplete, forcing trips to be assigned to other mode combinations, or could include combinations that are not used, causing an unnecessary computational burden. The disadvantages of this approach are the complexity of the network representation and the enormous increasing of network size since each node location is replicated many times according to the transfer state and the number of prior transfers. Another limitation is due to the mode combination constraints, they have to define in advance depending on the type of the application and applied directly on the network. It might be considered quite difficult either to fix the number of transfers or the type of combinations before knowing the type of application, and it also might imply the adaptation of the network structure as soon the application changes a little bit.

Hoogendoorn-Lanser (2005) developed a multi-modal choice set generation (CSG) model for estimation purposes for establishing individual choice sets for multi-modal trips using the train as main mode. She adopted is a so-called *timetable-based* or *run-based* approach meaning that each run of a public transport service is explicitly taken into account in both space and time. Each transport service is represented by a separate so-called *time-space graph* or *diachronic graph*. These diachronic graphs are combined in a *diachronic super-network* by adding transfer walking legs connecting the various PT services and extended the diachronic super-network with transfer-free in-vehicle legs. The generated choice sets are sets of routes in this *extended diachronic super-network*.
The diachronic graph (see Figure 3.17) consists of three different directed sub-graphs (see Cascetta (2001), and Nuzzolo, A. and Russo, F. (1994)):

- a service sub-graph G_v in which each run of each PT line is defined in both space - through its stops - and in time - according to its arrival and departure times at stops (see Figure 3.17);
- a demand sub-graph G_d in which each node represents a temporal centroid in order to simulate the time-space dimension characteristics of the trip;
- an access / egress sub-graph G_{ae} which allows the connection of the demand sub-graph with the access / egress stops.

The service sub-graph G_v for each PT line v consists of different sub-graphs G_{vr} - one for each run r of that PT line. Sub-graphs G_{vr} consist of nodes representing the arrival

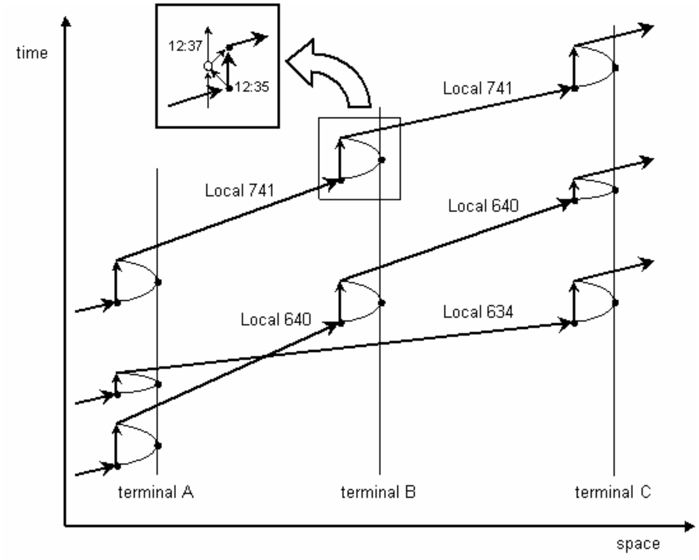


Figure 3.17: A run-based representation of transit service sub-graphs G_{vr} (taken from Nuzzolo, 2003).

and departure times at stops, legs representing *travel* from one node to another (run section), and the *dwelling* of the vehicle at a given stop (see Figure 3.18). Other nodes represent the time in which travelers *board* or *alight* a run at a stop. These nodes are connected to the nodes representing run arrival and run departure through boarding and alighting legs. The complete sub-graph G_v is built by connecting all the sub-graphs G_{vr} through legs representing travelers' transfer from one run to another run at the same stop (*stop axis*).
For each PT service, a diachronic graph $G_{PTservice}$ is specified by adding transfer-walking legs connecting the different line graphs G_v.

The demand sub-graph G_d represents the temporal demand segmentation made up by a number of sub-graphs G_{dcen}, one for each centroid *cen*. For each spatial centroid *cen* this sub-graph G_{dcen} consists of temporal centroids. The nodes are located at the position of spatial centroid *cen* and temporarily according to the user's departure times.

The access / egress sub-graph G_{ae} is made up of:

- legs connecting temporal origin centroids to nodes on the first boarding stop axis to represent the private mode access to transport services;
- legs connecting the alighting nodes of the stop axis to the spatial destination

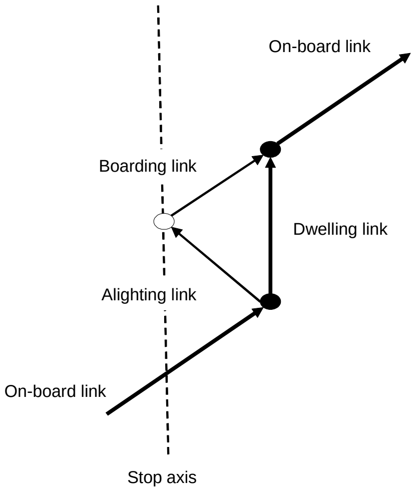


Figure 3.18: Leg classification at stops (taken from Nuzzolo, 2003).

centroids to represent the private mode egress from the PT system;

- legs connecting stops to represent possible interchanges between different stops (walking mode).

The private mode access/egress legs represent the available private modes between centroid and stops such as walking, bicycle, car, etc.

To represent a multi-modal transport network having different PT services, like trams, city busses, regional buses, metro, local trains, express trains and intercity trains, by a graph G, the different diachronic graphs have to be interconnected. This is achieved by adding connecting transfer-walking legs. The resulting graph, accounting for different PT modes, is called a diachronic super-network.
The diachronic super-network is extended by defining additional transfer-free in-vehicle route segments between every stop of a public transport line and all downstream stops of the line (see De Cea and Fernandez (1993)). These route segments are defined for each run of the public transport line using timetables. All these transfer-free in-vehicle route segments are added to the diachronic super-network, resulting in an extended diachronic super-network.

3.5.3 Comparison of the multi-modal network representations: advantages and disadvantages

This section summarises the main characteristics and limitations of the multi-modal network representation approaches described in the previous subsections. Two types of multi-modal network representations are distinguished, namely the separated uni-modal network layers, and the linked uni-modal network layers respectively. Each discussed approach is listed in a row in Table 3.2, while the columns list the main characteristics that from our point of view are related to the requirements for a useful multi-modal representation, as discussed in Section 3.4. These main characteristics, which are also relevant for the choice set generation approach, are network structure, mode combination constraints, generation of multi-modal routes, and location of centroids.

The *network structure* provides an indication about the relative network size of each approach and the network representation's clarity and consequently its complexity for building the multi-modal network from the uni-modal networks. The approach with separated uni-modal layers has a simple network structure which is too simple to enable from our point of view, a realistic multi-modal approach. The examples of the linked uni-modal network layers have complicated network structures except for the supernetwork approach (Sheffi, 1985). Abdelghany, K.F.S. (2001)'s network is quite complicated because of its number of transfer links that might be very large due to the number of possible mode combinations available in a multi-modal network, implying that the network size may increase enormously. Lozano and Storchi (2001) and Caramia and Storchi (1997) have defined a multi-modal network using a hypernetwork for the public transport system together with a state information system that make the network structure more complicated. Finally, Lo, H.K. and Yip, C.W. and Wan, K.H. (2003) combine several uni-modal networks into a single multi-modal network through the SAM network that make the network structure extremely complicated and leading to a large network size.

Another important characteristic of a multi-modal network representation is the possibility of *generating true multi-modal routes* and the degree of *flexibility in generating* such routes. This latter point relates for example to the flexibility in generating modal combinations within a route that, depending on behavioural aspects and the demand model, are defined in advance at the level of network structure. The classical approach proposed by Montella et al. (1999) has big limitations in this respect since the mode and route choice are modelled separately which does not allow a more realistic and behaviourally sound travel choice modelling. In fact, such an approach might not be considered multi-modal in our sense because no multi-modal routes can be generated. All generated routes are in effect uni-modal, reason why no mode combination constraints are defined nor applied at the level of network structure.

The multi-modal representation proposed by Sheffi (1985) allows the generation of multi-modal routes with a high flexibility in mode combination constraints, no restrictions are defined at the level of network structure. The multi-modal network representation proposed by Abdelghany, K.F.S. (2001) may generate multi-modal routes, although with low flexibility because it models the transfers from one transport mode directly to another for avoiding unrealistic transfers and mode choices at the level of network structure in advance depending on the type of application.

Table 3.2: Summary assessment of alternative multi-modal network representations (N.A. not applicable).

MM network type	Reference	Network structure	MM routes generated	Mode combination constraints	Location of centroids
Separated uni-modal layers	Montella et al. (1999) Vrtic (2003)	Simple	No	N.A.	At each uni-modal network
Linked uni-modal layers	Sheffi (1985)	Simple	Yes, high flexibility	No	N.A.
	Abdelghany (2001)	Quite complicated	Yes, low flexibility	Yes, due to transfer links	May belong to each UM network layer
	Lozano and Storchi (2001)	Quite complicated	Yes, low flexibility	Yes, due to state information system	Connected to private and PT network layers
	Caramia and Storchi (1997)	Quite complicated	Yes, low flexibility	Yes due to state information system	At car and walk network layers
	Lo et al. (2003)	Very complicated	Yes, low flexibility	Yes, due to the SAM network	Connected to each UM network layer
	Hoogendoorn-Lanser (2005)	Complex	Yes, high flexibility	Yes	Connected to private modes

However, it might also be too rigid and not enough flexible approach for the choice modelling. Lozano and Storchi (2001) define a multi-modal network together with a state information system that generates multi-modal routes with a low flexibility, because they take into account limitations on path composition. Caramia and Storchi (1997) explicitly define the composition of multi-modal paths, which have to be generated in advance depending on the specific purpose of their application namely the evaluation of the effects of parking price. Also in this case, the flexibility in multimodal route generation is limited. Finally, Lo, H.K. and Yip, C.W. and Wan, K.H. (2003) combine several uni-modal networks into a single multi-modal network to deal with transfers and non-linear fare structures through the addition of state information, with a low flexibility in the generation of multi-modal routes and the choice modelling.

The *locations of centroids* are defined for each uni-modal layers in the Montella et al.

(1999) approach, with high replications of nodes, like the approach by Lo, H.K. and Yip, C.W. and Wan, K.H. (2003). Caramia and Storchi (1997) have replicated the origin and destination nodes in the pedestrian and car layer networks, instead of having them only in one network layer, due to the aim of the application of evaluating the parking price effects; differently from Abdelghany, K.F.S. (2001)'s multi-modal network in which centroid nodes may belong to each uni-modal network layer. Hoogendoorn-Lanser (2005) only applies a single centroid connected by private mode links to the public transport modes.

In summary, the network representation approaches found in literature do not meet all requirements for efficient multi-modal demand modelling described in Section 3.4. This refers to the generation flexibility, to the sufficient consideration of route composition constraints, and also to the choice modelling flexibility. From a computational point of view, these approaches are too cumbersome in building the supernetwork, and too costly in computing time.
Below we will therefore propose and elaborate a representation approach that meets the requirements for efficient multi-modal travel demand modelling.

3.6 The Benjamins supernetwork approach as precursor

A precursor of the proposed supernetwork methodology (see Section 3.7) has been developed by Benjamins (2001). We summarize his approach in order to show better the merits of the proposed slimmer supernetwork version.
Benjamins' supernetwork version tried to meet various requirements. On the one hand, it should match the original uni-modal networks as closely as possible, allowing a clear representation of the network for sake of network data management and for visualisation purposes. On the other hand, it is used to compute multi-modal route sets using least-cost algorithms. Therefore a distinction is made between a base network and a supernetwork. The distinguishing feature of this approach is that it combines the level of spatial detail found in GIS-based base networks with an explicit representation of available transfers at each stop.

The base network.

The base network is for network data management and for visualisation. For the base network a lightweight GIS-environment called NETTER is used in which the network is coded using links and nodes. The NETTER software (*DEMIS*, 2001) serves as a geographical front-end for computational models dealing with spatial information. It provides support for models dealing with traffic networks, and allows a close match with topological maps.
The links represent streets or segments of railways, and have a large number of attributes describing the modes that may use that link. Associated with each node is a transfer table that specifies all possible transfers at this node along with their (default or node specific) characteristics (e.g transfer distances, times and costs).

Public transport is described using lines having a service frequency and defined by route segments connecting successive stops. Route segments have a travel time according to the timetable and correspond to a sequence of nodes and links in the base network. Geographical zones and their corresponding zone centroids are defined in the base network.

The supernetwork.

The supernetwork is meant for route set generation, calculation of route choice probabilities and trip assignment. In this network a distinction is made between transport links, access links, transfer links, and waiting links. All links have as attributes: link length, travel time, and travel cost. Transport links for private modes connect the same nodes as private mode links, but a single link will be created for each mode and for each direction. Access links connect the centroids with the nodes of the private modes, that is walking, cycling, and car. Transport links for public modes are created between all possible pairs of stops of a public transport line. Transfer links are situated at those nodes for which a transfer table has been defined and connect transport links with waiting links and transport links. Waiting links connect transfer links and links representing public transport modes. The nodes in the supernetwork are located at the same location as the nodes in the base network, but more nodes might be located at the same position. For an example see Figure 3.19. Shown on the left side are part of the base network and a transfer node, which is accessible by bus, train, bicycle, and on foot. Although the base network would imply that all transfers are possible, this is not always the case. In the example all available transfers are possible except those to and from car, since the transfer node don't offer parking facilities. This situation is represented explicitly in the supernetwork shown on the right. The car network is not connected to the other modes, even though the transfer node in the base network is physically accessible to cars. Walking links connect to all modes. A waiting link represents the average waiting time for the train. Note that the transfer table is encoded in the structure of the supernetwork.

Conversion.

The supernetwork is generated automatically using a dedicated conversion program. In this conversion all references to the original links and nodes are preserved enabling a return-conversion between supernetwork and base-network for the presentation of the results.
The conversion from a base-network to a supernetwork obviously has a huge impact on the dimensions of the network. Every link in the base-network is split up into unidirectional links per mode. Each public transport line is transformed into a set of stop-to-stop links and their waiting links. Finally, all transfer nodes are split up into all relevant intermodal transfer links. The result is called the *full* supernetwork.

In order to reduce the number of public transport links in the supernetwork, links serving identical pairs of stops are aggregated if the travel times are almost equal. In such cases the frequency of the resulting waiting link is the sum of the frequencies of the

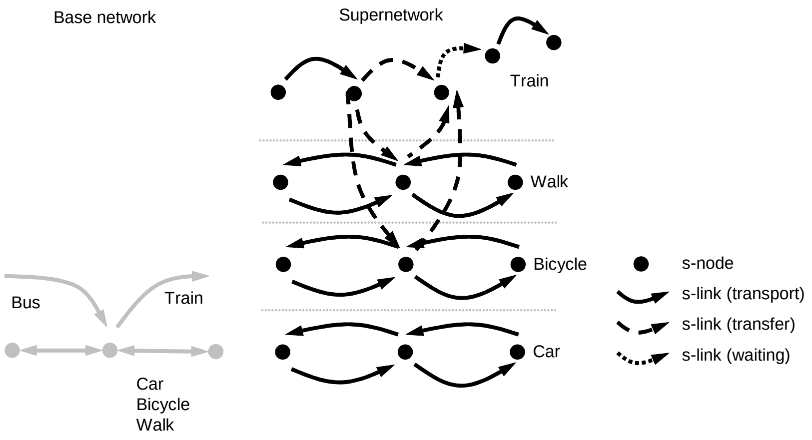


Figure 3.19: Relation between base-network and supernetwork (taken from Benjamins et al, 2002).

original waiting links. In addition, only those stops are retained that are located close to the centroids and stops where a transfer between public transport lines is possible. This is called the *reduced* supernetwork. However, other approaches of supply modeling of the public transport services (see e.g. this chapter) may conveniently be adopted as well. The network dimensions of the full and reduced networks of the case study corridor are shown in Table 3.3.

For the reduced network the number of nodes and links increase by a factor 2.4 and 7.3 respectively. More than 50% of the links in the supernetwork are related to the private modes, 40% to public transport links and waiting links, and 8% to transfer links. For the full network these factors are 13 and 32 respectively.

3.7 Proposed multi-modal supernetwork representation

In this section, our recently developed multi-modal network modelling approach will be presented. It is a simplified version of the precursor representation developed by Benjamins. This proposed network model has been implemented in our multi-modal demand model called TRANSFER (Carlier et al, 2005) and will be adopted for the choice set generation step to be presented in Chapter 7. For details on the network modelling, see Carlier et al (2005). First, arguments concerning the multi-modal network structure are discussed in Section 3.7.1, then definitions about what is the multi-modal network are introduced in Section 3.7.2, and an example of the multi-modal network representation is provided in Section 3.7.3. Finally, a comparison among our

Table 3.3: Dimensions of the base-network and the super-network (case study corridor) [Source: Benjamins et al, 2002].

Base network		Reduced Supernetwork		Full Supernetwork	
Type	Number	Type	Number	Type	Number
Nodes	2,388	Nodes	5,836	Nodes	41,000
Transfer nodes	903				
Centroids	136				
Links	3,099	Links	22,498	Links	77,979
		Transport links	19,631	Transport links	42,094
		Car	3,454	Car	3,454
		Bicycle	3,984	Bicyle	3,984
Lines	134	Pedestrian	3,984	Pedestrian	3,984
Bus	66	Bus	6,047	Bus	23,900
Tram	99	Tram	1,753	Tram	5,404
Train	55	Train	75	Train	447
Line segments	2744	Transfer links	1884	transfer links	5213
		Waiting links	983	Waiting links	30672

multi-modal network approach and the previous approaches presented in Section 3.5 is presented in Section 3.7.4.

3.7.1 Arguments for the multi-modal network approach

The first requirement we have considered very important is the feasible and easy applicability of the network model for practical demand prediction applications. From a practical point of view, a complicated and difficult network structure is not likely to be used by analysts. For this reason we have decided to choose a very simple network structure with a maximum clarity, being the supernetwork approach as defined in Chapter 2 and Section 3.5.2.

The supernetwork consists of the networks of the individual modes, which are connected at appropriate places by transfer links between the vehicular modes (see Figure 3.13). These transfer links represent the actual transfer between vehicular modes, for instance, walk time, as well as activities related to the parking of vehicles when relevant. This approach also makes it possible to analyze the impact of new transfer concepts such as Park & Ride locations. The application of the supernetwork approach allows maximum flexibility in the choice modelling due to the combination of mode and route choice in a single route choice in the multi-modal supernetwork.

We have chosen the supernetwork approach also for another important requirement being the possibility of generating multi-modal routes with a maximum flexibility and without any constraints defined in advance of the demand model. Another characteristic we considered relevant is the freedom and flexibility in generating all kinds of travel mode combinations without posing any restrictions or constraints at the network level. It is very difficult to define in advance which are the travel mode combinations that might be allowed according to the application aim and the demand model. In some cases very illogical transfer modes, such as the case of transferring from train to car, might be allowed if the application considers return trips from work to home. Our network model allows maximum flexibility in generating all types of multi-modal paths. Finally, another advantage of this approach is that the generation of the multi-modal network by building the supernetwork is quite simple and easy to be applied.

Since we aim to develop a practically applicable multi-modal model approach for long term planning projects, uni-modal (such as road and PT) network representations are required as simple as possible; therefore the hypernetwork approach is not considered for our model approach. The adopted a priori route generation approach (see Chapter 7) combined with a stochastic route choice modelling (see Chapter 4) sufficiently facilitates complex forms of choice behaviour of passengers. Moreover, given the long term planning purpose static modelling is preferred to the dynamic one because of simplicity and data requirements. Therefore a static network representation will suffice for our modelling aims.

Finally, centroids (origin/destination nodes) are located only in the pedestrian network layer because we assume that a trip begins and ends always with a walk mode, how short this may be. The walk mode is considered essential because walking is always a part of multi-modal trip, especially in transferring from one mode to another. Therefore, the walk network layer is an important layer to be connected to all other network layers. In the proposed multi-modal supernetwork network it is impossible to transfer directly from one mode to another, it always uses the walk network as an intermediary. For example transferring from car to train always involves going from the car network to the walk node and then to the train boarding link.

3.7.2 General formulation of multi-modal network

A multi-modal transportation network might be defined by a directed graph $G = (N, L, M)$, where N is the finite set of nodes, L the finite set of links, and M the finite set of travel modes. In a multi-modal network a node is a place where one has to select either continuing with the current mode or changing it. A link connecting two nodes by only one travel mode is called travel link. A change of mode or (inter-modal or intra-modal) transfer is represented by transfer links. In our formulation walk is always considered a travel mode, and transfer links mainly consist of parking links in the case of private travel mode (such as, car and bicycle) or boarding or alighting links in the case of PT travel mode.

On the graph G, one and only one transport mode $m \in M$ is associated with each link $(i, j) \in L$. Hence, there exist $|M|$ simple networks, one for each travel mode (including walk mode, which belongs to the pedestrian network), interconnected by transfer links.

The sets of nodes and links forming each one of the $|M|$ networks are respectively denoted by N_m and L_m, where m ($m \in M$) stands for a travel mode.

3.7.3 Multi-modal network representation

Based on the experiences with the Benjamins precursor approach, a network representation has been developed for modelling multi-modal networks (Carlier, K. and Fiorenzo-Catalano, S. and Lindveld, C. and Bovy, P.H.L., 2003) and (Carlier et al., 2005). The travel modes included in the current network structure are the following:

- Private modes as: bicycle, car driver, car passenger, and walk.
- Public modes as: bus, metro, tram, and train.

Since walking is always a part of a multi-modal trip, pedestrian network is also included as a network layer and it plays an important role, especially in the connections to all other network layers.

The private networks consisting of car (driver and passenger), bicycle, and walk networks each are represented as a finite set of nodes and links. The car network consists of motorway and secondary road streets, whereas the bicycle network consists mainly of secondary road streets, and the walk network includes mainly local streets. The public transport network part consists of all public transport services, characterised by the geographic location of its stops, frequency information, and parameters for the service model. Each service is modelled as a sequence of alternating travel links and dwelling links (the latter representing the time spent by the vehicle at a stop) (see Figure 3.6 and Figure 3.20).

For long term planning purposes, simple road and PT network representations are adopted. In particular, private network representation of a road intersection is based on the single node representation (see Figure 3.4) while the transit network representation is based on the Fearnside and Draper (1971) approach (see Figure 3.8), in other words, it is based on the transit network with line-specific nodes and links. Please note that the route section approach of De Cea and Fernandez (1993) (see Figure 3.9) as adopted in the precursor version of Benjamins might also be used to model the transit network. The disadvantage of this method, however, is that the number of links needed to represent the transit network increases enormously (see Table 3.1 in Section 3.3.2). The transit network representation is the main difference between the model approach applied by Benjamins and the one developed by Carlier et al. (2005). Moreover, the hypernetwork model is not considered since the strategy concept can well be handled by our combined a priori choice set generation and choice modelling approach. Finally, since the main focus of this thesis is the generation of choice sets for demand prediction purposes, in which the dynamic aspects are less relevant, a run network model is not taken into account and in the rest of this thesis we focus on a static model.

In our network representation, all public transport stops are connected to the walk network as well as are all relevant nodes of the car and bicycle networks, i.e. access and egress nodes such as parking facilities. The pedestrian network thus becomes the universal transfer network or intermediary between all vehicular modes. Please note that walk is always considered a mode within small zones, whereas walk links are not considered in case of large zones. Therefore, the pedestrian network exists only if the distance between nodes is likely that walking is an option.

The pedestrian network is connected through boarding and alighting or parking links to:

- the bicycle network at each node;
- the car (driver and passenger) network at a limited set of nodes where car parking is allowed (free or paying a fee), in the neighbourhood of each centroid node, railway stations and at P&R facilities, carpool locations;
- the public transport network at each public transport (bus, tram, metro and train) stop.

A change of mode or inter-modal transfer is represented by specific links called boarding and alighting links, which usually are get-on/get-off links from and to the public transport network such as metro, tram, bus and train networks, and parking links which usually are get-in/get-out links from and to the private transport network such as car (driver or passenger) and bike networks. Figure 3.20 shows an example of such a multi-modal supernetwork; in particular, three network layers are connected. On the one hand, the pedestrian network is connected to the car network with so-called parking links, and on the other hand connected with the PT network, such as metro, bus, tram or train network, via boarding and alighting links (it should be noted that the PT nodes in this layer are coded as detailed as given in Figure 3.20). As can be seen from the picture, centroid nodes are only connected to the network layer of walking mode.

The car (driver and passenger) and bicycle networks are accessible only through special parking links. These represent the disutility of finding, opening, boarding and starting the car or bicycle (and vice versa for leaving the car/bicycle network). Also a parking fare may be due, which is then coded as a cost on the parking link. Parking links are only present near origins, destinations, at stations, at carpool parking and at Park & Ride (P&R) facilities. The networks for car passengers and for taxis are identical to the car driver network, except for the link cost parameters and the occasional dedicated link for taxis.

In the network representation, much attention has been paid to the representation of inter-modal transfers and line-to-line transfers in the supernetwork (see Figure 3.20). All transfers are modelled as a sequence of an alighting link (or parking link), pedestrian nodes (and maybe also pedestrian links, but not necessarily) and a boarding link (or access link to the car or bicycle network). The source node of a boarding link (or

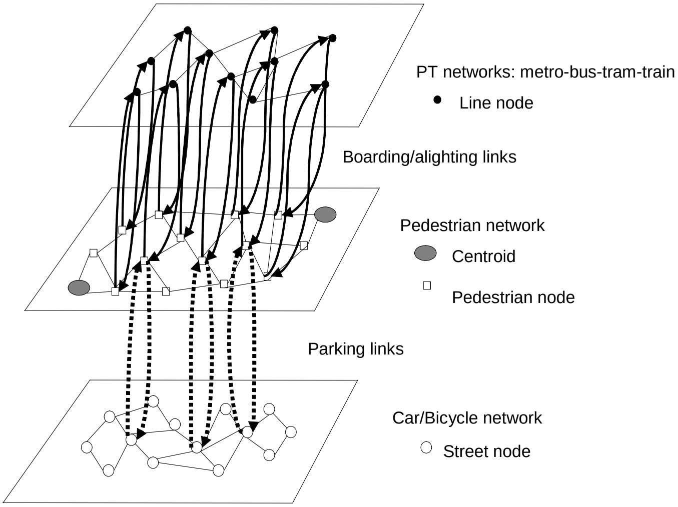


Figure 3.20: Multi-modal supernetwork representation.

parking link) and the tail node of an alighting link (or access link to the car or bicycle network) are in the walk network, thereby isolating each mode-specific network. In big transfer points (e.g. railway stations) walk links may be added to represent the walking structure of the transfer point. The transfer disutilities (involved with walking and waiting) can be attached to the respective walk and boarding links.

Table 3.4 summarizes the introduced concepts of our supernetwork specification.

Figure 3.21 shows as example a multi-modal path through the multi-modal supernetwork (again, the PT layer is a simplified representation, since the PT nodes have to be coded as given in Figure 3.22).

The path starts at the origin centroid in the bottom-left part of the pedestrian network. After two walk links, the path switches via walking through the to-parking link connecting the pedestrian and the car network to get to the car. The path then proceeds with a short sub path in the car network (two links) at which end the car is parked. Using a from-parking walk link to the walk network, the path proceeds via a boarding link to a PT stop, the pedestrian node is used for transferring from the car mode to the train mode. Finally, the sub path in the train network and the alighting link gets to the pedestrian network and then to the final destination centroid in the right-up part of the pedestrian network.

To illustrate in more detail some of the concepts introduced previously, Figure 3.22 gives a representation of a railway station at which three different train services stop.

Table 3.4: Summary overview of link and node types in the supernetwork.

Link name	*Mode*	*Function*
Street link	Car/Bicycle	Connects two adjacent street nodes.
To-parking	Walk transfer	Connects a walk node to a street node.
From-parking	Walk transfer	Connects a street node to a walk node.
Walk link	Walk	Connects two adjacent walk nodes.
PT link	Train, Metro, Bus, Tram	Connects two adjacent PT stops.
Boarding link	Walk transfer	Connects a walk node to a PT stop.
Alighting link	Walk transfer	Connects a PT stop to a walk node.
Node name		
Street node	Car/Bicycle	Joins two or more street links.
Line node	Train, Metro, Bus, Tram	Joins two adjacent segments of the same line where of passengers can alight and board.
Walk node	Walk	Joins two or more walk links
Centroid	Walk	Walk node where trips can begin (origin) or end (destination

Each train service is represented by separate links. Travellers boarding a train at this platform arrive using the pedestrian network and board a train service using a boarding link. This link has a boarding disutility and a waiting time equal to the time until the first departure of the chosen train service after the traveller arrived at the platform. Travellers can transfer from the train network to the pedestrian network via the alighting link. Using the pedestrian network the destination or other transport modes can be accessed. Travellers on a train passing the stop, without boarding or alighting the train, use the dwelling link. All travellers alighting and boarding the train services pass the same node representing the platform. The platform itself is connected using a walk link with the station square, which is connected with transfer links to and from all other modes (e.g. bicycle, car, urban bus, regional bus).

The multi-modal network representation of stations/transfers presented in Nielsen, O.A. and Frederiksen, R.D. (2001) has some similarities to the ones proposed here.

3.7.4 Multi-modal network comparisons

Advantages of our proposed multi-modal network representation compared to the other possible network representations proposed in the literature are the following. First, this concerns the simplicity of our network representation in terms of structure (role of the walk network as interconnecting transfer mode) and network size (numbers of nodes, travel links for the PT networks but also number of transfer links), leading to very little

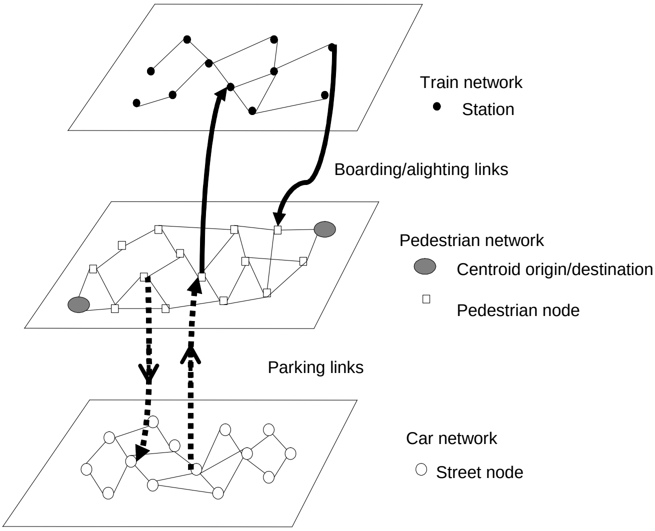


Figure 3.21: Example of a complete multi-modal path form origin to destination in the multi-modal supernetwork.

computational complexity for building in a very easy way the multi-modal network given the uni-modal network layers. In our approach there is no replication of nodes, in effect centroids are only located to the pedestrian network layer, since walk is considered as an essential parts of a multi-modal trip.

Second, it concerns the possibility of generating multi-modal paths and its high flexibility in generating all kinds of multi-modal paths and travel mode combinations, without any restrictions or constraints defined in advance before knowing the type of application to be applied. Our approach might be applied for different practical applications without any type of changes in terms of network structure and data requirements.

3.8 Conclusion

This chapter has dealt with the uni-modal and multi-modal network representations. Several types of representation have been proposed in the literature of which an overview of the uni-modal network representations distinguished in the continuous (e.g. road networks) and discontinuous (e.g. PT networks) networks have been presented. Subsequently, an overview of multi-modal network representation approaches has been given and a comparison of these multi-modal network representations has been discussed on the basis of the requirements established in this chapter.

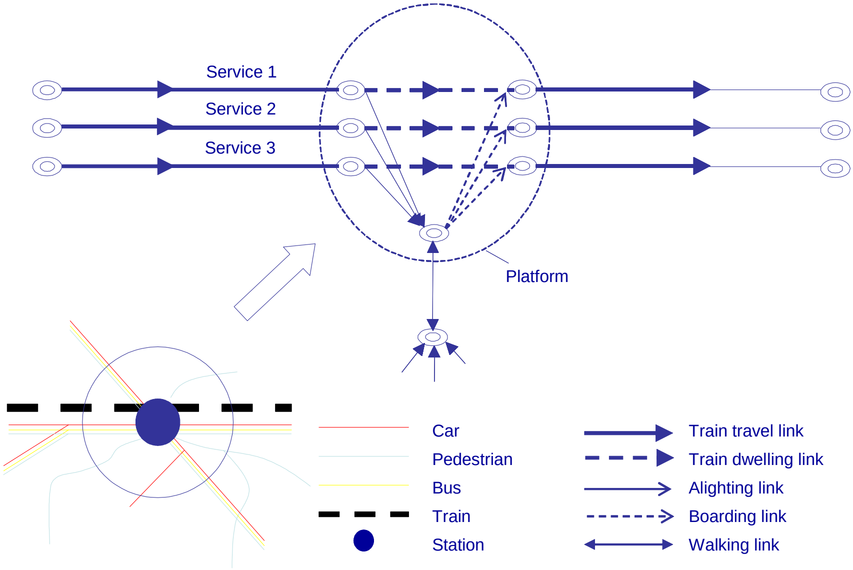


Figure 3.22: Supernetwork representation of a railway station (taken from Carlier et al, 2003).

In order to develop an efficient multi-modal network representation suitable for choice set generation, the following aspects play an important role. First, the chosen network structure influences the network size and therefore its computational complexity. A very complicated and difficult network structure is not likely to be applied in a practical and feasible application. Another requirement is maximum flexibility in choice modelling, especially in transfer modelling and in the generation of all types of multi-modal paths (with all mode combinations). The network structure should not have any influence on the type of routes that might be generated and available in the network. Moreover, due to the choice set composition in which a large variety of mode combinations has been observed, no a priori restrictions on mode combinations (e.g. no limitations on number of transfers) or path composition should be put for the network modelling.

As a result of the previously stated requirements, due to the complexity and data requirements of dynamic model, a run network model is not considered and a static model has been developed, also because the main focus of this thesis is the generation of choice sets for demand prediction purposes, in which the dynamic aspects are less relevant. Moreover a supernetwork approach for multi-modal network representation has been presented with which the traditional steps of mode choice and route choice can be integrated into a single route choice problem. An important characteristic of this approach is the use of the walk network as the intermediary mode interconnecting all uni-modal vehicular network layers. To represent intermodal transfers correctly a number of artificial links are added as connecting links between the walk network and the vehicular networks. Typically, there are no direct links among the vehicular networks.

The main property of the proposed multi-modal network representation is maximum flexibility in choice set generation and choice modelling and in generating all possible mode combinations due to the use of supernetwork and simplicity in the uni-modal network representation. Simple road and static PT network representations maybe exploited in the supernetwork building process. Finally, an advantage of this approach is that the generation of the multi-modal supernetwork is quite simple and easy to be applied. The chosen supernetwork representation keeps the increase in network size relative to the sum of involved uni-modal networks within acceptable limits.

In the coming chapters, the network representations for uni-modal and multi-modal networks presented in this chapter will be exploited in order to apply choice set generation models in the case of private uni-modal networks (Chapter 6) and in the case of multi-modal networks (Chapter 7).

Chapter 4

Conceptual framework for choice set concepts

4.1 Introduction

The elaboration of the proposed multi-modal modelling approach in Chapter 2 pinpointed at the crucial importance of the a priori generation of route choice sets in advance of the choice modelling and network assignment steps. This chapter will elaborate on the notion of choice set as a preparation for the route choice modelling chapter and the establishment of choice set generation procedures (Chapters 6 and 7). It appears that the notion of choice set is not self-explanatory. There is nearly no scientific literature available dealing with this subject in relation to route choice (an exception is Bovy, P.H.L. and Stern, E. (1990)). In doing empirical research with route choice sets, it is important to distinguish between actual choice sets considered by travellers in their decision making, observed choice sets as reported in a survey, or generated choice sets established by an analyst or researcher. In addition, we may distinguish choice sets of individuals (disaggregate level) or choice sets referring to groups of individuals (aggregate level). So, it is important to distinguish the perspective of the traveller and that of the researcher. Finally, it makes sense to look at choice sets from an application point of view (which always is the researcher's perspective): for which purpose are they needed? We may for example consider the purpose of scientific analysis of the choice set formation behaviour of travellers, or the purpose of estimating parameters of choice models, or prediction of choice outcomes given such models.

The purpose of this chapter is to establish a theoretical framework for the various choice set concepts being used in the remainder of this thesis. To that end, the chapter will introduce definitions of the various choice set notions and will deal with the role of choice sets in transportation modelling in general, and in a priori generation in particular.

For several reasons it is important to deal explicitly with choice sets. A basic usage of choice sets is in providing insights into available travel opportunities in particular network conditions which will be paramount in explaining choice behaviour. This will support better knowledge of traveller behaviour such as choice set formation and choice decision making. Moreover, the usage of choice sets is essential in determining best parameter values of travel choice models and with calculation of usage of alternatives (in prediction and forecasting models). For all these reasons, the generation of choice sets, the main subject of this thesis, based on sound empirical knowledge of actual choice sets is a research subject of the most importance.

The main contribution of this chapter is a conceptual framework, presented in Section 4.3, for the various choice set notions from a choice behavioural point of view and from researcher perspective in a nice context system. This will be based on the concise theoretical framework (preceded in Section 4.2) showing the choice set notions as basic elements in a spatial travel choice theory. The notions are clearly defined in relation to the various application purposes. A synthesis of all these notions is given in Section 4.4. Section 4.5 provides a wide variety of empirical figures about choice set characteristics in various transport network types such as choice set size and composition in terms of numbers and types of alternatives. This chapter concludes with implications of the new notions and findings on choice sets for the next chapters of this thesis.

4.2 A behavioural framework for route choice sets

In analyzing the specific aspect of human behaviour dedicated to choice decisions, assumptions must be made about the way individuals make decisions. Assume, for example, that some individuals decide to have dinner in a restaurant, and face the question to which restaurant? Among all restaurants existing in a given area, individuals might consider only a certain subset of restaurants, for example, the ones where they have been once and of which they have appreciated the kind of cuisine or the good price, or the restaurants known by friends or by advertisements. In this case, it might be stated that people make a choice among all restaurants that are known by them. Obviously, if a restaurant is not known it cannot be chosen. In order to make this choice, it might be stated that people make a sort of list of restaurants based on certain criteria. Several classifications of the restaurants are possible referring among others matter to kind of cuisine, location, price, special interests, etc. For example, the list of restaurants may consist of Chinese, Greek, Italian, Mexican, Japanese restaurants. This is a typical example of the so-called choice set of which the alternatives are characterized by the type of food cooked in the restaurant. Several other aspects may characterize the identification of the list of alternatives, for example, the choice set may contain all restaurants located in the city centre, or all cheapest restaurants, or all vegetarian restaurants, etc.

Given this (trivial) example, it might be noted that a basic assumption in a generic choice problem is the concept that the decision-maker chooses from a finite non-empty set of available alternatives known to him. From a logical point of view the assumption that the choice set is finite and non-empty is due to the fact that at least one alternative belongs to the choice set, moreover, from a psychological point of view it is well known that the decision maker is only able to consider a very limited number, say about 7 alternatives Bovy, P.H.L. and Stern, E. (1990). The characterization of the choice set containing these alternatives depends on the context of the application.

Usually the decision-maker is assumed to be an individual. For some applications, this assumption is not restrictive. However, the concept of individual may easily be extended, depending on the specific application. At aggregate level, we may consider that a group of persons (a household, for example) is the decision-maker. In doing so, we decide to ignore all internal differences within the group, and we assume that individuals within the group have similar preferences, and characteristics, and behave at similar way. Depending on the application context the traveller might be an individual or a group of individuals having similar travel demand and supply conditions (same origins and destinations), and behaving in similar way.

As a basis for modelling the choice set formation process, this section will present a concise overview of elements determining individual route choice behaviour as presented in Bovy, P.H.L. and Stern, E. (1990), with emphasis on the way route finding is structured. According to their theory, the route selection process of a traveller appears to be a complex task (see Figure 4.1).

Figure 4.1 synthesizes a number of linked successive transformations assumed to be performed by the traveller in his mind in order to eventually choose a travel option. On top of the figure there are the main external factors of the choice process: the traveller with its characteristics, the transport network with its various route options and their various attributes. Most essential to our thesis is the sequence of sets of route alternatives indicated in the most right column of Figure 4.1, and the way how these sets result from a sequence of mental processes in the mind of the traveller, indicated in the column directly to the left. The most left column describes the most essential personal attributes of a traveller that are determinants in the various transformation processes leading to choice.

The first process called information acquisition refers to the ways travellers acquire information about potential travel options in transport networks (cognition process). The most common ways are learning the environment through a map (known as survey-based knowledge), and experiencing it through a travel process (known as route-based knowledge). Several studies (see in Bovy, P.H.L. and Stern, E. (1990)) have investigated the difference between knowledge acquisition through maps or travel experiences. Although some researchers have pointed out that the map-based (survey) knowledge should be more effective than the route-based learning process (Golledge, R.G. and Dougherty, V. and Bell, S., 1995), others have argued that route knowledge should contribute most to the environment representation. Besides of those manners of acquiring information there are other possible ways such as asking friends, relatives or colleagues, consulting telephone information line, consulting transit schedules, attempting to navigate when needed, referring to signs and strangers' advice as necessary,

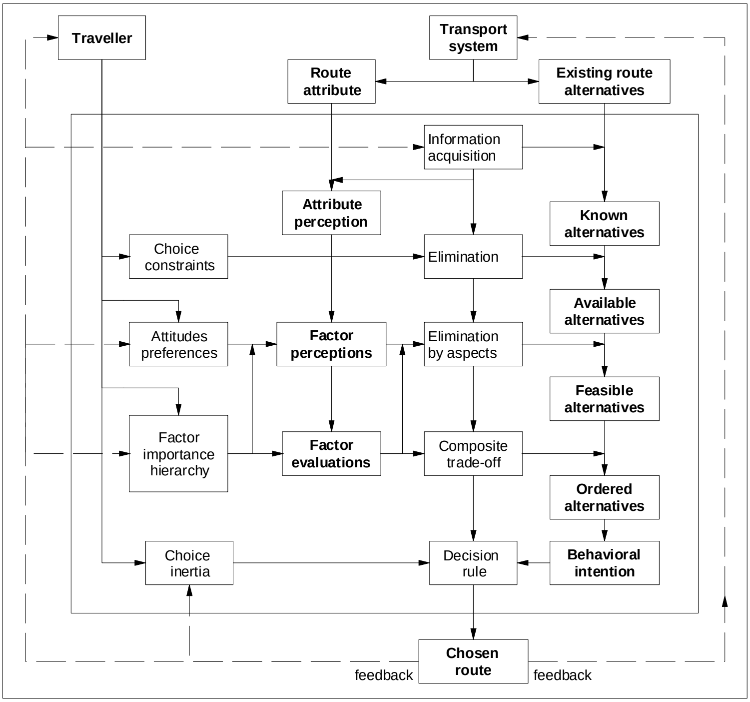


Figure 4.1: Elements of individual choice behaviour focusing on the route choice process (taken from Bovy and Stern, 1990).

exploring the city during spare time and travelling new routes based on combination of known routes. Since dealing with the complicated psychological phenomenon of ”cognition” is not the aim of this thesis, we confine ourselves to the class of travel situations in which no information needs to be acquired on behalf of the journey such as, for example, commuter trips, home-to-work, and home-to-school trips. The traveller in such situations is so well acquainted with the network, or at least one route to his destination, that he does not need to consult maps or to ask questions about it on the way. For such travel, we assume that relevant route options are well known to the traveller and are part of his choice set.

The traveller’s cognition of the network is associated with his travel experiences (feedback from usage of chosen routes), implying that his knowledge of route options open to him is limited: not all existing opportunities are known to him. Only a subset of the existing routes, the so-called known alternatives can be considered in the decision process. Based on the traveller’s personal characteristics such as choice constraints, preferences, and perceptions, only the known alternatives satisfying these constraints form the set of feasible alternatives among which the traveller will make a deliberate choice.

An essential feature of the work of Bovy, P.H.L. and Stern, E. (1990) is the notion that travellers choose their routes on the basis of their perceptions of the transport network. This perception is known as the mental map. As noted above, travellers know only a subset of all existing route alternatives; they also have a certain subjective perception of the characteristics of the known alternatives due to their travel experience and information acquisition behaviour. Consequently, a person's mental map strongly depends on how the objective route attributes are perceived and distorted into the subjective route factors that are successively evaluated in the decision process.

Apart from the cognition, perception, and elimination filters that route options have to pass in order to be feasible for an individual traveller, the evaluation process plays an important role as further filtering process in the choice set formation. This filtering process of the feasible alternatives is partly based on elimination by aspects and partly on comparison of similar alternatives in a trade-off of route attributes. Some travellers preclude the use of motorways or tolled roads, whereas others may disregard routes having a detour of more than a factor two or a number of public transport transfers of three or more.
Only after a fairly limited group of attractive feasible alternatives remains (the consideration set), the traveller will make a more in-depth evaluation of options making a trade-off among their counter-balancing characteristics using some composite combination rule.
The successive choice sets resulting from each of the mentioned processes are assumed to be successively narrowed down to the behavioural intention that through the traveller's decision rule is transformed into the ultimately chosen route.

The route choice decision-making is characterized by dynamic components, the most important being the feedback from usage of chosen routes to nearly all phases in the decision process. In this way, choice set formation is a learning process through dynamic adaptations of cognition and perception of available and feasible options.
Given the type of process involved in route choice decision-making, we may expect strong individual differences in behaviour due to differences in cognition, preconceptions, and preferences of options among individual travellers.
Despite these strong idiosyncracies, different individuals may however take the same decision, that is, choose the same route, though on different grounds.

For our choice set formation subject we may conclude the following (see Bovy, P.H.L. (2007)):

- choice set formation and choice from considered options are *distinct mental processes* for a traveller that follow different rules;
- whereas choice set formation predominantly is governed by *non-compensatory* decision-making on the basis of constraints and elimination by aspects, choice from considered options appears to be largely of a compensatory nature;

- apart from constraints, choice set formation is also *preference driven*, influenced by the traveller's most important choice factors;
- a traveller's choice set formation largely follows an experimental process of *trial-and-error* of route use and information acquisition;
- consideration sets may differ strongly between individuals even under the same conditions.

In the following, we basically adopt the series of choice set notions and their background developed in Bovy, P.H.L. and Stern, E. (1990) for our purpose of developing choice set generation methods. However, since our scope is much wider than purely scientific insight into the individual route choice process, that is including also observation, estimation, and prediction aspects of route sets, we will adapt the notions somewhat and will introduce a different terminology. In particular, we have to extend the Bovy-Stern notions towards a multi-modal network context.

4.3 Conceptual notions about choice sets

4.3.1 Various considered perspectives

The following subsections introduce specific notions about choice sets in the context of route choice where a decision-maker who wants to travel from an origin to a destination at a specific period of time, and for a given purpose, chooses from a set of available trip alternatives. In dealing with traveller's choice behaviour it appears necessary to clearly distinguish between a couple of perspectives. On the one hand, we will distinguish the situation of a single individual traveller from that of a group of travellers that make more or less the same trip between similar origins and destinations and exhibiting similar behaviour. An individual has its own choice set pertaining to his individual situation whereas a group's choice set is the envelop of individual choice sets of the group's members. On the other hand for both the individual and group situation we will distinguish the perspectives of the traveller actually making choices and that of the researcher studying such travel choices for purposes of analysis, estimation, or prediction. In the traveller's perspective it is the actually occurring behaviour that is of interest, much of which is not known to an external observer, whereas the researchers perspective is that of an external observer who tries to observe and model as good as possible the actual behaviour. To that end, the following subsections introduce specific notions about choice sets pertaining to the travellers and to the researchers respectively, and in both types of notions special definitions are given at disaggregate level (individual level) and at aggregate level (group level). In order to help the reader in finding his way in these complex matters, a dedicated terminology and notation is used. Table 4.1 introduces the notational system adopted in this chapter to distinguish the various choice set notions to be explained and defined in the sequel of this section. The distinction between the traveller's and the researcher's perspective is relevant for this thesis in at least the following two respects. First, since in the following Chapter 7 we will work with route set observations in calibrating generation models one should be aware of the difference between actual choice sets (perspective of traveller) and observed reported choice sets (perspective of researcher). Secondly, in testing choice set generation procedures (Chapters 6 and 7) predicted (generated) choice sets (perspective of researcher) are compared with observations of actual routes (perspective of traveller).

Travellers and researchers have different knowledge of the transport system, and hence of alternatives available for a specific trip. Often the number of *available* alternatives is large. Only a subset of them is *known* to the traveller, *feasible* to him, and *considered* in his choice process. Travellers make travel choices based on *incomplete* and *inaccurate information*, which is partly derived from past experiences. In this context, especially discrepancies between anticipated and actual experiences are important. Past experiences determine to a large extend travellers' perception of the transport network and the characteristics of available alternatives. Researchers, however, estimating or predicting choice sets often have multiple data sources and choice set generation algorithms available to them (*full information*), and use them to determine sets of available alternatives for individual travellers and groups of travellers.

Table 4.1: Legenda of terminology and notation to be used in Chapter 4.

	Traveller's perspective		**Researcher's perspective**		
	Actual behaviour		*Observation, estimation, prediction*		
					Font type
Individual q **situation** n $n = OD, t, q$ t = departure time instant	master c.s.	ms	Master c.s.	MS	*Regular*
			Objective set	OS	
	subjective c.s.	ss	Subjective c.s.	SS	
	consideration set	cs	Consideration set	CS	
	chosen alterative	ca	Chosen alternative	CA	
Group situation	master c.s.	$\mathbf{ms}$	Master c.s.	$\mathbf{MS}$	**Bold**
			Objective set	$\mathbf{OS}$	
	subjective c.s.	$\mathbf{ss}$	Subjective c.s.	$\mathbf{SS}$	
	consideration set	$\mathbf{cs}$	Consideration set	$\mathbf{CS}$	
	chosen alterative	$\mathbf{ca}$	Chosen alternative	$\mathbf{CA}$	
Letter type	*Normal*		CAPITAL		

Choice sets *observed* by researchers are not necessarily equal to *actual* choice sets available to, known by, or feasible to travellers. Differences naturally result from the

fact that researchers are not present during the entire trip, and have to reconstruct trips from e.g. interview reports or travel diaries. Observed travel behaviour can be used to study current trip making, establish travellers' knowledge of trip alternatives and determine travellers' preferences.
Researchers may also *generate* choice sets to be used in *analysis* of choice options (e.g. availability of alternatives), *choice modelling* (e.g. estimation of parameters in a utility function) or *prediction* of choice probabilities in a travel demand analysis (e.g. for an assignment problem). The anticipated use of generated choice sets puts special requirements on the choice sets as well as on the choice set generation technique.

This section aims at properly defining the concept of a choice set thereby accounting for different stages in the travel choice process (available, known, feasible, considered and chosen), viewpoint (travellers or researchers), number of travellers (individual traveller or group of travellers), origin (actual, observed, generated) and application (analysis, estimation or prediction). The resulting definitions are not only applicable in a multi-modal context, but also for more general choice situations.

This section introduces the concepts of *universal sets*, *master sets*, *objective choice sets*, *subjective choice sets*, *consideration sets*, and *chosen alternatives*, and shows the relationship between these sets of alternatives. In doing this, a difference is made between:

1. actual, observed and generated choice sets,
2. sets of alternatives for individual travellers and groups of travellers,
3. analysis, estimation and prediction applications and
4. traveller's viewpoint or researcher's viewpoint.

In Section 4.3.2 and Section 4.3.4 the different concepts are considered for individual travellers from the traveller's and the researcher's perspective, respectively. Sections 4.3.3 and 4.3.5 consider the same concepts for groups of travellers - again from the traveller's and the researcher's perspective. Section 5.4.5 discusses the differences between the various application purposes of choice sets. In section 4.4 relationships among actual, observed and generated choice sets are discussed.

4.3.2 Choice set notions from the traveller's perspective at disaggregate level

In line with the theoretical framework given in Section 4.2, in defining individual choice sets, the following key concepts are adopted: the *universal set*, the *master choice set*, the *subjective choice set*, the *consideration set*, and the *chosen route alternative* (see also Table 4.2).

Table 4.2: Choice set notions applying to the route choice context at individual level from traveller's perspective.

Terminology	Symbol	Definition	Attributes	Referring to a
Universal set		Set of all *existing* routes in a network between an OD pair.	OD addresses	Spatial region
Master (choice) set	$ms_{OD,T}$	Set of all *available* routes in a network between an OD pair in a time period T.	Size OD addresses Time period T	Spatial region including a single OD pair (door-to-door)
Subjective choice set	ss_n	Subset of the $ms_{OD,T}$ containing the routes *known* by and *feasible* to an individual q for his trip between an OD leaving at a departure t.	Size OD addresses Time instant t Preferences**	A specific travel demand of individual q
Consideration set	cs_n	Subset of ss_n containing the routes *considered* by q in his choice process i.e. routes among which he makes a trade-off.	Size OD addresses Time instant t Preferences**	Individual q
Chosen route	ca_n ca_n	The route *chosen* by the traveller.	Time instant t	Individual q

$n = \{OD, t\ , q\}$.
T = time period, such as morning or evening peak.
t = departure time instant, for example, 9:15.
**Preferences that may be related to trip-purposes.

The *universal set* is a generic notion applicable to all perspectives and applications, meaning the set of all existing route alternatives between O and D in a network. This is not a fruitful concept since it is only seldom possible to identify this set because it is too large and sometimes even non-countable.

The *master choice set* contains all *available* alternatives in the context of the choice problem of a particular traveller going from origin O to destination D. Available in this context means that an existing option is *known* to the decision maker: if it is not known it is not available to him.

Considering the transportation mode choice, the master choice set may contain all potential transportation modes, such as walk, bicycle, bus, car, etc. In the context of urban trips the alternative plane, which is also a transportation mode, is clearly not an option, and therefore it is not included in the master choice set. Therefore, the master choice set depends on the context of the choice problem, and for instance in the context of route choice, the master choice set is defined as follows.

Definition 1 *Let the master choice set* $ms_{OD,T}$ *be the set of all available route alternatives in a given network from a given origin* O *to a given destination* D *and in a specific time period* T*.*

We will refer to ”master choice set” and ”master set” interchangeably throughout the rest of this thesis. From Table 4.2 it can be seen that the master set refers to a particular OD pair from a network located in a spatial region. The OD pair refers to individual origin and destination addresses.

From the perspective of an individual traveller who wants to satisfy his particular travel need in the subjectively best way given his personal constraints, the choice set pertaining to this individual traveller is defined as the *subjective choice set.*

Definition 2 *Let the subjective choice set* ss_n *be the subset of the master set containing the route alternatives* known by *and* feasible to *a particular individual* q *for a trip between the given origin and destination (OD) pair and leaving at a preferred departure time instant* (t) *that satisfy his travel needs* (n = *OD*, t , q). n *indicates a triple of personal attributes of the trip maker (OD locations, departure time instant, and individual traveller).*

Thus, the subjective choice set consists of the trip alternatives known by the traveller in his decision process that are all feasible to him in satisfying his travel needs. The term *feasible* refers among other matters to the availability of private transport modes and public transport services, time feasibility (time pressure at origin and destination addresses and time budget), monetary feasibility (monetary budget) and physical (dis-)ability in using alternatives.

The *consideration set* is defined as the set of all alternatives considered in the choice process.

Definition 3 *Let the consideration set* cs_n *be the subset of the subjective choice set containing those route alternatives considered by a particular individual* q *in his choice decision, i.e. alternatives among which he is making a trade-off and is forced to make a choice* ($n = \{OD, t , q\}$).

The difference between the alternatives that are *known* by and feasible to travellers (subjective choice set) and the alternatives that are *considered* in the *choice process* (consideration set) might be somewhat ambiguous (see Hoogendoorn-Lanser (2005)). One way to characterise this difference is by identifying the alternatives in the consideration set as the alternatives that travellers have in *top of their mind* and among which they make a deliberate choice. The larger subjective choice set consists also of those alternatives that travellers only come up with *after some time of thought*, which are alternatives known by and feasible to travellers, but among which travellers never make a choice.

The single alternative actually chosen by the traveller, the so-called *chosen route* belongs to the actual consideration set.

Definition 4 *Let the chosen route alternative ca_n be the route from the consideration set chosen by the individual q for travelling between a given origin and destination (OD) pair and leaving at a preferred departure time instant (t) that satisfy his/her travel needs (n = OD, t, q).*

We will refer to ”chosen route alternative” and ”chosen route” and ”chosen alterative” interchangeably throughout the rest of this thesis. Table 4.2 summarises the introduced notions and definitions together with a related notation given in the route choice context at disaggregate level from the perspective of the individual decision-maker.

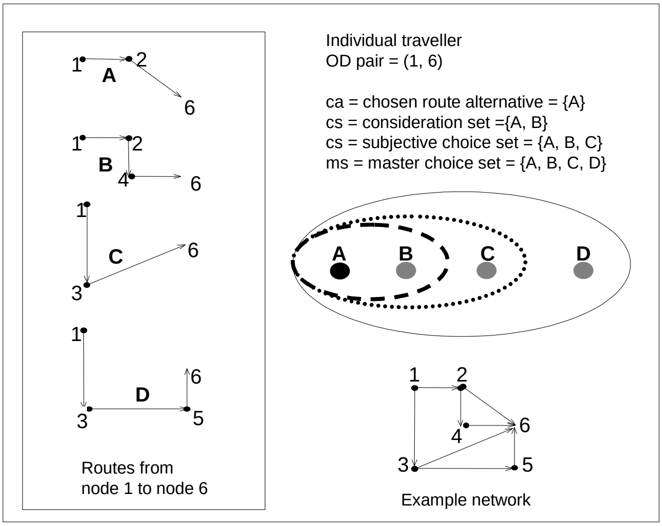


Figure 4.2: Examples of chosen route, consideration set, subjective choice set, and master set.

Figure 4.2 illustrates the choice set concepts introduced. It shows a network of existing connections between an origin 1 and a destination 6 giving rise to 4 alternative paths from A to D. The chosen route of the individual travelling from 1 to 6 leaving at a preferred departure period is *ca* = A, the consideration set is *cs* = A, B, whereas the subjective choice set is *ss* = A, B, C and the master set also includes the non-considered alternative D.

Figure 4.3 illustrates the relationships among choice set concepts for the individual decision-maker from the perspective of an individual traveller.

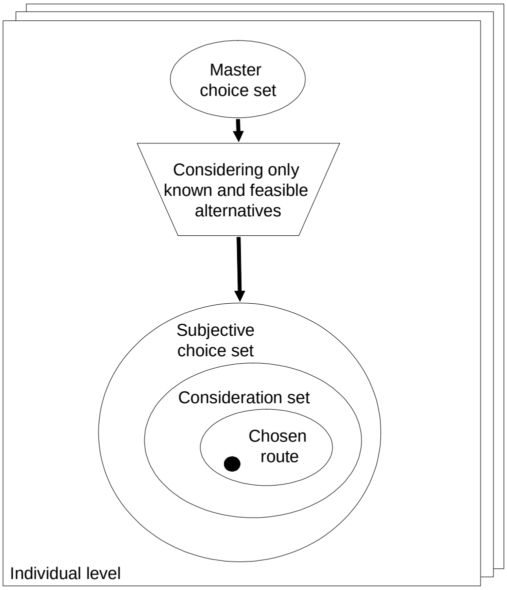


Figure 4.3: Relationships among choice set notions referring to an individual traveller from the traveller's perspective.

4.3.3 Choice set notions from the traveller's perspective at aggregate level

This subsection again focuses on actual choice behaviour, but now the viewpoint of a group of travellers having similar OD-addresses within a spatial area, is taken. Such a group consists for example of people living in the same neighbourhood and working in the city centre, which might be defined using areas' postal codes. Since origins and destinations for different travellers in a group are in the vicinity of one another, alternatives available for one traveller are likely to be also available to the other travellers in the group (given that individual feasibility constraints are satisfied). Individuals differ however with respect to their knowledge of the transport network, availability of transport modes, travel needs and preferences. Therefore, although individuals have similar OD-pairs, their subjective choice sets, their consideration sets, and their chosen alternatives may differ considerably (see e.g. empirical figures in Bovy, P.H.L. and Stern, E. (1990)).

The definitions of choice sets pertaining to a group of travellers making similar trips are a straightforward extension of those at individual level; instead of a pair OD-addresses we now have a pair of OD-areas each covering a set of OD addresses; the choice set notions and definitions at aggregate level are preceded by the term 'joint' (see Table 4.3).

Of course, the joint universal set in the aggregate case is the union of the individual universal sets combining all individual OD-addresses to the joint OD-areas.

The *joint master (choice) set* consists of the union of all master choice sets of each group member. The distinction between the master set and the joint master set is necessary when the alternatives available in the region at the individual level (origin-destination pair) are quite different among the individuals belonging to the group.

The *joint subjective choice set*, i.e. the set of all alternatives known by and feasible to at least one of the travellers in the group, is determined by taking the **union** of subjective choice sets of all individuals in that group. The *joint consideration set* and the *joint set of chosen alternatives* are defined in similar manner.

Table 4.3 summarizes the definitions similarly to Table 4.2. The distinction of the class of joint choice sets is relevant since in many applications, especially in prediction, the choice analysis cannot be done at individual level but rather only at aggregate level.

Figure 4.4 illustrates the developed notions while Figure 4.5 illustrates an example about the choice set concepts introduced.

Figure 4.5 shows a network having four alternative routes connecting origin 1 and destination 6 and two alternative routes E and F connecting origin 2 and destination 6. The consideration set of traveller I1 travelling from origin 1 to destination 6 is equal to the set A, B, whereas traveller I2 travelling from origin 2 to destination 6 has the consideration set equal to E. Their joint consideration set thus is A, B, E. The set of all known and feasible route alternatives (the subjective choice set) for travellers q_1 and q_2 also includes the non-considered alternative C and F, respectively. The joint subjective choice set consists of the union of their subjective choice sets. The master choice set of traveller q_2 is identical to his subjective choice set, while the master choice set of traveller q_1 also includes the non-known and/or non-feasible alternative D. In this example, we assume that travellers do not have exactly the same origin-destination addresses and that the alternatives existing in the area related at the individual level (a single origin-destination pair) are quite different among the two travellers belonging to the group. Thus, in this case a distinction is needed between the master choice set and the joint master set.

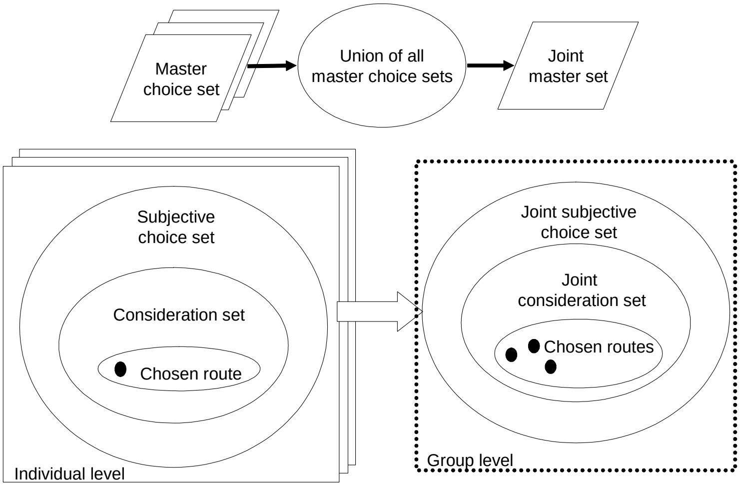
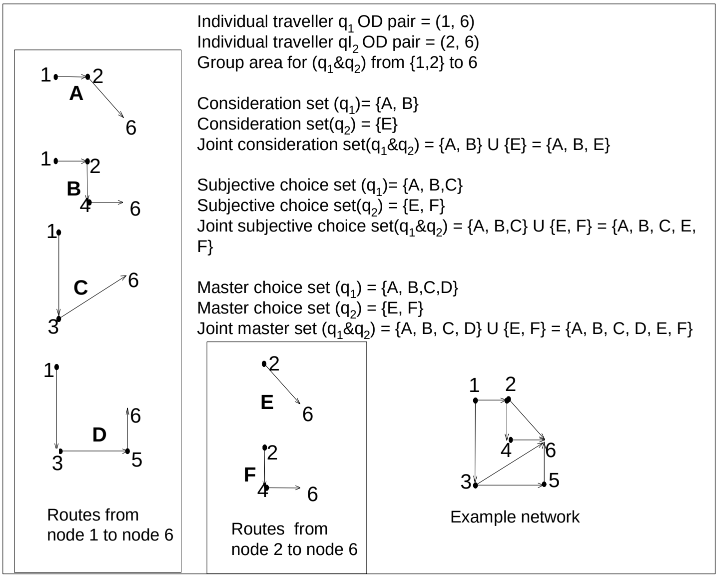


Figure 4.4: Choice set notions referring to a group of travellers from the traveller's perspective.



Individual traveller q_1 OD pair = (1, 6)
Individual traveller ql_2 OD pair = (2, 6)
Group area for (q_1&q_2) from {1,2} to 6

Consideration set (q_1)= {A, B}
Consideration set(q_2) = {E}
Joint consideration set(q_1&q_2) = {A, B} U {E} = {A, B, E}

Subjective choice set (q_1)= {A, B,C}
Subjective choice set(q_2) = {E, F}
Joint subjective choice set(q_1&q_2) = {A, B,C} U {E, F} = {A, B, C, E, F}

Master choice set (q_1) = {A, B,C,D}
Master choice set (q_2) = {E, F}
Joint master set (q_1&q_2) = {A, B, C, D} U {E, F} = {A, B, C, D, E, F}

Figure 4.5: Examples of joint consideration set, joint subjective choice set, and joint master set.

Table 4.3: Choice set notions applying to the route choice context at group level from traveller's perspective.

Terminology	**Symbol**	**Definition**	**Attributes**	**Referring to a**
Joint universal set		Set of all individual universal sets.	OD areas	Spatial region
Joint master (choice) set	**ms**	Union of all ms_n of group's members resulting in the set of *available* routes in a network between OD areas in a specific T.	Size OD areas Time period	Spatial region including a pair of OD areas
Joint subjective choice set	**ss**	Union of all ss_n of group's members resulting in a subset of **ms** containing all routes *known* by and *feasible* to Q travelling between OD areas and leaving at T_w.	Size OD areas Time window Preferences**	Q
Joint consideration set	**cs**	Union of all cs_n of group's members resulting in a subset of **ss** containing all routes *considered* by Q in the choice process i.e. routes among which they make a trade-off.	Size OD areas Time window Preferences**	Q
Chosen route	**ca**	The route *chosen* by the travellers.	Time window	Q

T = time period, such as morning or evening peak.

T_w = departure time window, such as [9:00, 10:00].

**Preferences that may be related to trip-purposes.

Q = group of individuals having similar demand conditions, preferences and characteristics.

Table 4.4: Choice set notions applying to the route choice context at individual level from researcher's perspective.

Terminology	Symbol	Definition	Attributes	Referring to a
Universal set		Set of all existing routes between a specific OD pair.	OD pair	Spatial region
Generated objective master choice set	MS_{OD}	Subset of the universal set containing all generated routes in a network between an OD pair assumed by the researcher to be *relevant* and *logical* for q.	Size OD pair	Spatial region including a single OD pair
Generated objective choice set	OS_q	Subset of the MS_{OD} containing all generated routes assumed by the researcher to be *feasible to* q travelling between an OD pair and leaving at t that satisfy his travel needs.	Size OD pair Time instant Purpose	A specific travel demand of q
Generated subjective choice set	SS_q	Subset of the OS_q containing all generated routes assumed by the researcher to be *known by* and *feasible to* q travelling between an OD pair.	Size OD pair Purpose	A specific travel demand of q
Generated consideration set	CS_q	Subset of the SS_q containing all generated routes assumed by the researcher to be *considered* by q travelling between an OD pair.	Size OD pair Purpose	A specific travel demand of q
Generated chosen route	CA_q	Route belonging to the generated CS_q assumed to be chosen.	OD pair Purpose	A specific travel demand of q

q = specific individual traveller.

t = departure time instant, for example, 9:15.

Table 4.5: Choice set notions applying to the route choice context at group level from researcher's perspective.

Terminology	Symbol	Definition	Attributes	Referring to a
Universal set		Set of all existing routes between OD areas.	OD areas	Spatial region
Generated joint objective master choice set	**MS**	Union of MS of group's members being the generated routes in a network between OD areas assumed by the researcher to be *relevant* and *logical*.	Size OD areas	Spatial region including OD areas
Generated joint objective choice set	**OS**	Union of OS of group's members being a subset of **MS** containing all generated routes assumed by the researcher to be *feasible to* Q travelling between OD areas and leaving at T_w that satisfy their travel needs.	Size OD areas Time window User classes**	Q
Generated joint subjective choice set	**SS**	Union of SS of group's members being a subset of **OS** containing all generated routes assumed by the researcher to be *considered* by Q travelling between OD areas.	Size OD areas Time window User classes**	Q
Generated joint consideration set	**CS**	Union of CS of group's members being a subset of **SS** containing all generated routes assumed by the researcher to be *considered* by Q travelling between OD areas.	Size OD areas Time window User classes**	Q
Generated chosen routes	**CA**	Routes belonging to the **CS** assumed to be chosen.	OD areas User classes**	Q

T_w = departure time window, such as [9:00, 10:00].

**User classes = Class of travellers, or purposes.

Q = group of individuals having similar demand conditions, preferences and characteristics.

4.3.4 Choice set notions from the researcher's perspective at disaggregate level

In contrast to the traveller's perspective, we may take the position of an external observer or researcher, who tries to *define*, *specify* and *analyse* the trip alternatives of an individual traveller or a group of travellers. To this end, the researcher *observes* actual travel behaviour, i.e. observes traveller's actual choice sets and chosen routes, and *generates choice sets* for *analysing* travel behaviour or for *predicting* future travel behaviour.

In this section we focus on the individual level. From a researcher's perspective, individual choice sets can be subdivided into generated and observed choice sets.

The researcher may *generate* choice sets, either to analyse travel behaviour or to make forecasts for a future situation. Generating sets of alternatives from the perspective of the researcher mostly starts with specifying the master set. Since the size of the master choice set (see Definition 1) might be very large, especially in the route choice context, this set might not be workable, nor necessary for the researcher's purpose. Therefore, in generating the choice sets of an individual traveller, for practical reasons, the researcher first excludes alternatives based on *objective criteria*, i.e. not requiring any information about the knowledge and considerations of a specific traveller. To this end, first all *illogical* alternatives are removed. The term illogical refers to alternatives including *loops*, and alternatives that are not *temporally suitable*, e.g. a trip alternative having impossible transfers or consisting of a sequence of trip segments having an illogical time ordering (travellers cannot go back in time). This results in the *objective master set* containing all relevant and logical route alternatives irrespective of traveller characteristics. In dealing with the individual level, the researcher may specify additional selection criteria based on traveller and trip characteristics (among other matters availability of private transport modes, time and monetary resources) resulting in the subset of the objective master set of *feasible routes* for a specific traveller, called the *objective choice set*.

Table 4.4 reports definitions and notation for individual choice sets generated by the researcher. Figure 4.7 (left side) illustrates the developed notions of choice sets referring to an individual traveller from the researcher's perspective.

Besides generating traveller's choice sets, the researcher may *observe* choice sets, by using appropriate survey techniques. The researcher may try to determine not only the chosen trip alternative but to obtain as much information as possible about the traveller's consideration set and subjective choice set as well.

The notions developed from the traveller's perspective can be subdivided according to *actual* and *observed* choice sets to be explained below. The true actual choice sets in most cases may not be known to an external observer. In some cases, the researcher may observe actual travel behaviour (traveller's choice set), for example, through interviews or surveys. Therefore, the notions referring to subjective choice sets and consideration sets might be distinguished between *actual* (mostly non-observed or even non-observable) and *observed.*

Asking travellers about non-used trip alternatives may result in two types of alternatives, the alternatives that are *known to them* (actual subjective choice set) and the alternatives that are *considered in the choice process* (actual consideration set). However, the boundary between actual consideration set and the actual subjective choice set clearly is not fully unambiguous and is difficult to establish from interviewing travellers (see Hoogendoorn-Lanser (2005)).

The collected sets of alternatives are called the *observed consideration set* and the *observed subjective choice set* respectively. These sets of alternatives naturally stem from reports given by the traveller since non-used alternatives cannot be observed independently by an external observer. Since during interviews the traveller may forget to mention feasible or even considered alternatives (highly dependent on the used interviewing technique), may not share some information with the researcher (considers it to be not relevant), may make errors in describing the alternatives or may have incomplete information about the alternatives, observed consideration sets and observed subjective choice sets both are considered to be random samples that only partially cover the actual consideration set and the actual subjective choice set. Besides that, the researcher may make errors in interpreting the traveller's reports. Figure 4.6 illustrates the relationships between what a researcher may observe as choice sets and the actual choice sets of travellers. Observed choice sets may not be considered to be true subsets of the related actual sets since we have to reckon that respondents may incorrectly report about there available, feasible or considered alternatives, or similarly, that the researcher may incorrectly consider existing alternatives as being available, feasible or considered by a traveller (see Figure 4.6).

To summarize, the following definitions are introduced.

Definition 5 *Let the actual subjective choice set* ss^a*, the actual consideration set* cs^a*, and the actual chosen route* ca^a *be respectively the subjective choice set, the consideration set, and the chosen route actually considered and performed by the individual traveller* q *for his trip.*

(see related definitions in subsection 4.3.2 and Table 4.2).

Definition 6 *Let the observed subjective choice set* SS^o*, the observed consideration set* CS^o*, and the observed chosen route* CA^o *be respectively the subjective choice set, the consideration set, and the chosen route of traveller* q *for his trip as observed by the researcher.*

(see related definitions in Table 4.4).

In Figure 4.6 the observed chosen route is assumed to be the same to the actual chosen route, however, it is important to note that in some cases depending on the survey

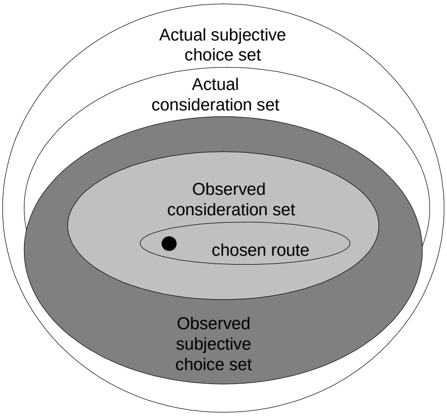


Figure 4.6: Relationships between actual and observed choice sets from the researcher's perspective.

technique it might be that the observed or reported chosen route does not correspond to the actual chosen route.

4.3.5 Choice set notions from the researcher's perspective at aggregate level

The definitions of the actual and observed choice sets pertaining to a group of travellers making similar trips might be easily extended from the previous subsection as unions of sets of alternatives. In the case of a homogeneous group of travellers the *actual joint subjective choice set* and the *actual joint consideration set* vis-a-vis the *observed joint subjective choice set* and the *observed joint consideration set* are taken into account.

Similar to the definitions given in Table 4.4 we now have:

Definition 7 *Let the actual joint subjective choice set* ss^a*, the actual joint consideration set* cs^a*, and the actual chosen routes* ca^a *be respectively the union of the actual subjective choice sets, of the actual consideration sets, and of the actual chosen routes actually considered by the individual members of a group of travellers with similar travel needs.*

Definition 8 *Let the observed joint subjective choice set* SS^o*, the observed joint consideration set* CS^o*, and the observed chosen route* CA^o *be respectively the union of the observed subjective choice sets, of the observed consideration sets, and of the observed chosen routes of the members of a group of travellers.*

The developed distinction between individual and group level stems from the fact that it is not always possible in the analysis of choice behaviour to follow a completely individual-level approach. Especially in predicting choice behaviour the analyst often has to resort to a group-level approach. Table 4.5 reports definitions and notation for the choice sets generated by the researcher while Figure 4.7 (right side) illustrates the developed notions of choice sets referring to a group of travellers from the researcher's perspective.

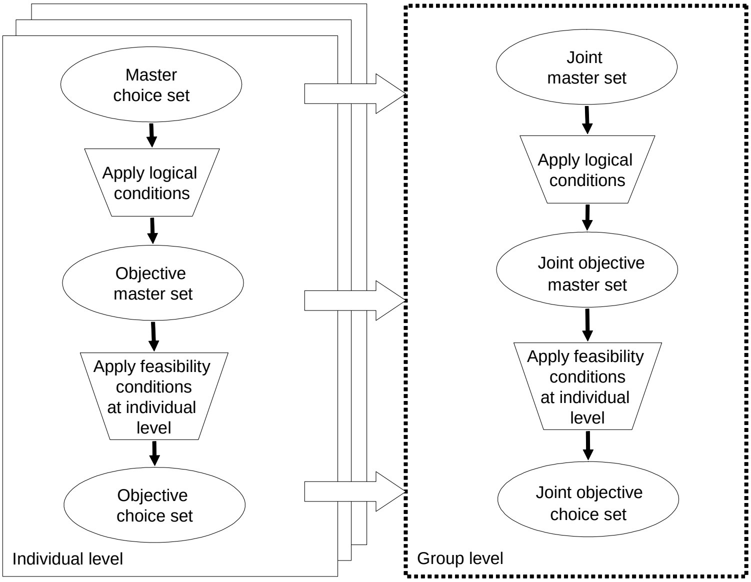


Figure 4.7: Choice set notions referring to an individual traveller (left side) or a group of travellers (right side) relevant from the researcher's perspective.

4.3.6 Notions and terminology of generated estimation and prediction choice sets

For modelling purposes, the researcher needs to generate choice sets, either to analyse observed travel behaviour or to predict choice behaviour for a future situation. Choice set generation methods (see Chapters 6 and 7) can be used to approach as closely as possible individuals' true choice sets. This is a complex task, since individuals and researchers mostly have different information about the transport network, and researchers do not precisely know the traveller's preferences or the additional considerations taken into account in his choice process.

In generating choice sets, we may distinguish four different purposes of the researcher:

- supply side analysis of available travel alternatives;
- estimation of parameters of utility functions or other choice models (demand side analysis);
- prediction of route choice probabilities with route choice models in a travel demand analysis and determining the shares of route usage in networks (assignment);
- data completion of route choice surveys.

These applications of choice set generation pose different requirements to the choice set generation outcomes in terms of choice set sizes and representational quality. Whereas for estimation purposes the choice sets need not be exhaustive nor close to the actually considered routes by the travellers, in the other cases it is to be preferred to have choice sets that match reality as closely as possible in terms of choice set composition and route types. In addition, prediction choice sets need to satisfy the purposes of the demand analysis given by the policy context (such as inclusion of tolled alternatives in a tolling study, or Park & Ride (P&R) alternatives in a P&R policy study).

A characteristic important difference between the estimation and prediction applications in practice is that the estimation process mostly is performed at the individual level (chosen routes and maybe choice sets of individual travellers and individual OD addresses), whereas the prediction process mostly is performed at the aggregate group level (choice sets referring to trips between zones). However, using micro-simulation approaches, choice set prediction at the individual level is feasible.

For an elaboration on the use of generated multi-modal choice sets for supply side analysis, estimation, and data completion, we refer to the companion dissertation of Hoogendoorn-Lanser (2005).

In this thesis, we focus on the prediction application of generated multi-modal choice sets where we analyse the prediction quality of our generation procedures at both individual and aggregate level (Chapter 7).

In line with the different application purposes, we extend our set of notions and terminology and introduce the notions of *estimation choice set* and *prediction choice set* at individual and aggregate levels (see Table 4.6).

In generating the *subjective estimation set* and the *consideration estimation set*, information about the traveller's knowledge and preference of alternatives available from observations may be used. To this end, alternatives are selected based on *behavioural criteria* that refer among other matters to travellers' preferences with respect to trip attributes, such as transport modes, in-vehicle, waiting, and walking time and costs; overlap conditions which refer to the exclusion of alternatives that are quite similar. Partly this may be based on travel information or objective criteria, for example trip

Table 4.6: Types of generated choice sets.

	Estimation purpose	*Prediction purpose*
Individual level	Estimation choice sets OS_q, SS_q, CS_q	Not common (this thesis, Chs. 6 & 7)
Group level	Not common	Prediction choice sets $\mathbf{OS_q}$, $\mathbf{SS_q}$, $\mathbf{CS_q}$ (this thesis, Chs. 6 & 7)

frequency and years of travel experience for a specific origin-destination relation, and partly it maybe based on general travel experiences and specific events which are therefore difficult to incorporate.

Theoretically it is plausible to introduce predicted choice sets at individual level, for example using a micro-simulation approach. However, since in predicting choice behaviour it is impossible to observe individual travellers behaviour, the analyst mostly has to resort to a group-level approach. Therefore, practically, the resulting choice sets that might be considered are the *joint objective prediction choice sets*, *joint subjective prediction choice sets* and *joint consideration prediction sets* at group level (see Figure 4.8 (right side)).
Although in a choice set generation context it makes sense to distinguish these different types of choice sets, current knowledge on actual behaviour of travellers is not mature enough to operationalize this distinction in practical applications of generation models. Therefore, in this thesis we consider the consideration and subjective choice sets to be identical.

4.4 Interrelationships between developed choice set notions

The choice set notions developed in this section have clear interrelationships. Group level choice sets depend on individual choice sets, observed choice sets depend on actual choice sets, and researcher's generated choice sets depend on observed choice sets.
For a better understanding of the outcomes of choice set generation models and choice modelling exercises attention should be given to the possible disagreements between actual, observed/reported and predicted choice sets.
To that end, Figure 4.9 summarizes the developed choice set notions and their mutual relationships. The figure is composed of four columns, the first two referring to the traveller and the last three to the researcher of which the second column to the right contains notions relating to both the traveller and the researcher. While the top three rows refer to the individual level, the last row represents the aggregate group level.

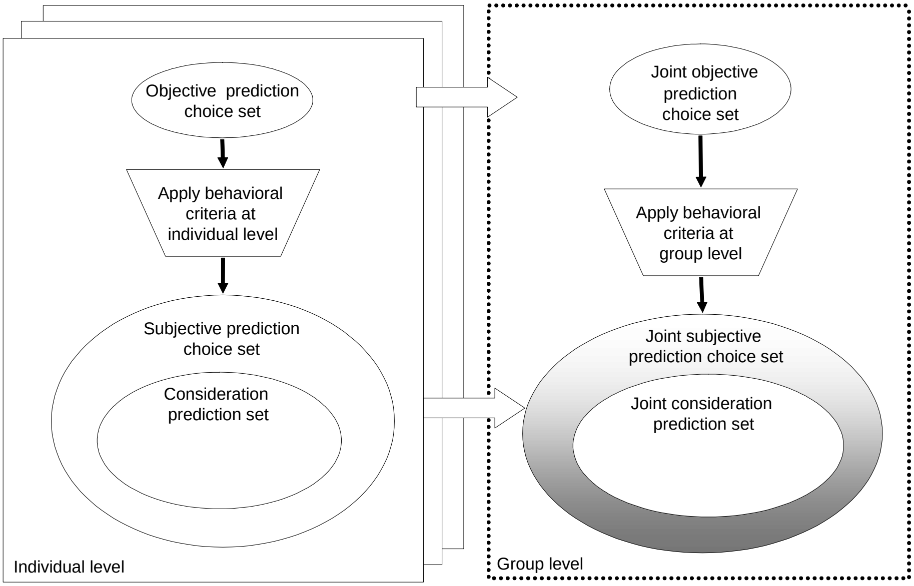


Figure 4.8: Notions for generated choice sets for prediction referring to an individual traveller (left side) or a group of travellers (right side) relevant from the researcher's perspective.

The most left column indicates the notions directly related to the travellers' actual choice process. Apart from the chosen alternative these subjective choice sets cannot be observed independently by an external observer. The second column from the left indicates the reported choice sets observed by an independent researcher. The reported choice sets and chosen routes need not be exactly identical to the true ones. The remaining two columns to the right indicate the choice sets generated by researchers for prediction purposes (predicting the expected use of alternatives) and estimation purposes (in order to derive insights into choice behaviour and for estimation of behavioural parameters). The third column from the left deals with the prediction case where the researcher establishes subjective choice sets to be used for prediction of choices in unobserved conditions, whereas the most right column pertains to the estimation case where the researcher tries to estimate the objective choice sets related to observed choices.

Figure 4.9 synthesizes concisely the theoretical framework developed in this section. An individual has knowledge about feasible trip alternatives (*actual subjective choice set*), makes choices (*actual consideration set*) and travels (*actual chosen trip*). Researchers try to obtain as much information as possible about this travel behaviour, analyse the collected data and, based on the data, draw conclusions about which alternatives are known by and feasible to the traveller (*reported subjective choice set*) and the chosen trip (*reported trip*). Using appropriate generation algorithms, subjective and objective choice sets can be estimated (*generated subjective choice sets* and *generated objective choice sets*). In generating the objective choice set, a researcher tries to generate alternatives considered feasible to the traveller (based on the reported traveller's characteristics). In generating the subjective choice set, a researcher narrows the generated objective choice set down to the alternatives assumed known by the traveller. Using appropriate choice models, researchers estimate route choice *probabilities*. There will be potential disagreements between the actual, reported and generated travel behaviour. Similar relations hold for choice sets of groups of travellers having similar OD relations.

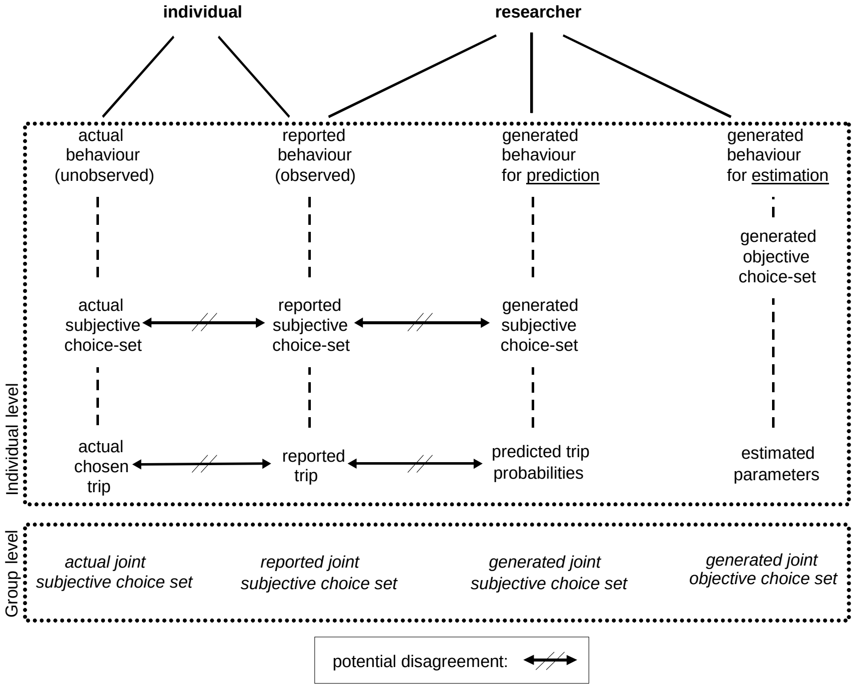


Figure 4.9: Relationships between actual, observed, predicted and estimated behaviour and corresponding choice sets (taken from Hoogendoorn-Lanser, 2005).

4.5 Choice set characteristics: some empirical figures

Relevant characteristics of choice sets from a choice modelling point of view are their size and composition. Whereas the size refers to the number of different alternatives in the set, the composition may refer to a variety of characteristics, such as spatial distribution, spatial structure (mutual overlap), road type composition, and modal composition.
An important characteristic of a route choice set is its size, i.e. the number of routes which is theoretically possible, available, or feasible in a certain choice situation between an origin and destination. This number of routes may be very large, even after the exclusion of so-called topologically senseless routes, such as the ones containing loops. Another important characteristic of a route choice set is its composition in terms of spatial structure and route types that provides insight into service quality of the transport network at hand, for example, by analysing the spatial structure of the set, such as the degree of mutual overlap and crossing of the alternative routes.

Several empirical studies have paid attention to choice set size and composition in various transport network conditions in order to describe the supply situation in those conditions but also to study their effects on individual choice behaviour.
So far, only few studies have investigated the choice set size and composition in a route choice context. Below we will concisely summarize findings from such studies.

A route choice survey carried out by Van Der Waard (1988a) collected data about 1,863 centre oriented urban public transport trips by bus and tram in several cities in The Netherlands for the purpose of estimation choice models. In this case, the individual objective route choice sets contained 2 to 4 alternatives. 60% of travellers had a choice between 2 alternatives only, 34% among 3 alternatives, while only 5% had a choice set of 4 alternatives available from which to make a choice. In total, the number of observed routes is 4,588 among which are the cases with no alternatives available (81 persons).

Bovy, P.H.L. and Stern, E. (1990) report various cases of choice set size and composition in route choice contexts. In one urban case regarding a cycle route network in Delft, The Netherlands, between the central railway station and the university campus, the generated master set consists of more than 1000 existing routes, whereas the joint objective choice set contains approximately 40 routes and the size of the observed subjective set has an average of 4 to 5 route alternatives per individual; in total approximately 15 routes are known by and feasible to all cyclists, therefore 15 is the size of the observed joint subjective choice set. In another case referring to route choice behaviour of 50 inter-urban commuters driving by car between the cities of Gouda and Delft in The Netherlands (distance about 30 km), the joint objective choice set consists of approximately 15 relevant routes, and the observed subjective choice sets had an average size of 4 routes per driver. Finally, similar results are reported about the size of observed subjective choice sets of 2 or 3 routes in the context of motorists' route choice in Newcastle (UK).

A study conducted by AGV (1990) analysed the possibilities of rerouting available in the motorway network of the western part of The Netherlands. The main finding of this study was that the opportunities of rerouting are very limited. The observational study identified 28 situations in which convergences and divergences allow feasible rerouting with 2 or more alternative routes, 25% of them having a traffic intensity of more than 10,000 trips per day. Of those cases having a rerouting option, on average 3 possible routes exist with a maximum of 6 routes in one case, 5 routes in 3 cases and 4 routes in 7 cases (see also the study by Bliemer, M.J.C. and Versteegt, E. and Castenmiller, R. (2004)).
Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001) show that in the waterway network of The Netherlands a minimum of 540 to a maximum of 20,869 routes are generated for 30 OD pairs, depending on the choice et generation approach adopted. In the case of the Monte Carlo Labelling combination approach (see Chapter 6) applied to the Dutch waterway network, a total number of 562 paths has been generated for 30 OD pairs, which is on average 19 routes per OD pair. In this case, the maximum number of generated routes for a single OD pair appears to be 58.
Ramming, M.S. (2002) reports on about 160 observed trips in the urban network of Boston, metropolitan area (Massachusetts, USA). His survey focused on observations of the chosen alternative for a specific origin-destination pair, and data collection about one single trip. The generated objective choice sets have sizes up to 51 alternative routes per trip, with a median size of about 30 routes per trip. The size of the objective choice set is very small (less than 5 routes) for about 5% of the origin-destination pairs, in which cases the origin is very close to the destination, while it is about 40 or more routes for 25% of the origin-destination pairs.

Statistics on observed subjective choice sets collected in 2001 from a large survey conducted among 511 train travellers from/to larger cities in The Netherlands are reported in (Fiorenzo-Catalano, S. and Hoogendoorn-Lanser, S. and Van Nes, R., 2003). In this case the generated objective choice sets contained on average 63 alternatives per trip, with a maximum of 376, whereas the observed subjective choice set has on average 2 alternatives per trip, with a maximum of 6. An in depth analysis of choice set composition for inter-urban train trips (same survey) is reported in Hoogendoorn-Lanser (2005) in which the multiple boarding and alighting stations, train services, access and egress modes in combination yield 24 alternatives, which is less than the full combinatorial set that contained more than 100 combinations of sub-options alternatives. Analysis of mode composition of the generated objective choice sets, observed subjective choice sets, and chosen alternatives shows that there is a substantial difference in the modal composition of those choice sets. The largest variation within route alternatives is found for the home-based part of the trip, because more modes are available in the home-end side than the activity end side. Moreover the objective choice sets appear to be very large with alternatives having a substantial overlap.

Bliemer, M.J.C. and Versteegt, E. and Castenmiller, R. (2004) show that in the motorway network of The Netherlands 181,567 routes are generated for 109,292 OD pairs, which is on average 1.66 routes per OD pair. This may seem little, but taking into account that the network is mostly a freeway system with many triangular shapes, there are many OD pairs for which only one route is actually a realistic option. The maximum number of routes generated for an OD pair is 14. These outcomes confirm the earlier findings of the AGV Adviesgroup voor verkeer en vervoer BV (1990) study.

A data collection carried out by Van De Reijt (2004) on pedestrian behaviour in railway stations in Delft and Breda (The Netherlands) resulted in 534 observed routes chosen by a group of pedestrians going from an origin area (bus and tram stops) to a destination area (station platforms) in Delft, and a total of 180 observed routes in Breda. The observed individual choice sets varied in size between 1 and 4 (see also Daamen, W. and Bovy, P.H.L. and Hoogendoorn, S.P. and Van De Reijt, A. (2005)).

From a variety of studies on regional multi-modal networks we know that the set of available alternatives may be very large, see eg. Benjamins et al. (2002) with 20 or more generated alternatives, or Hoogendoorn-Lanser (2005) with even 50 or more alternatives. The rich travel options especially exist on the access and egress parts of the multi-modal trip. The feasible choice sets differ between user groups depending on vehicle availability (car, PT, bicycle).

Table 4.7 summarizes for each type of network the number of routes being considered by the travellers that have been observed (reported) through travel surveys (Columns 4 and 5). In addition the table shows figures of the numbers of routes generated through simulation tools considered feasible to the travellers (Column 3). Columns 3 and 4 show results at individual level, whereas column 5 at group level. Column 3 show figures pertaining to generated objective choice sets, whereas columns 4 and 5 show results on observed subjective choice sets and observed joint subjective choice sets, respectively. Column 6 shows the number of observations and the number of choice cases such as individuals or OD pairs.

As we can see from the table, the size of the route choice set strongly depends on the type of network and the type of trips. On the one hand, the enormous density of urban networks leads to a multitude of routes in most choice situations; on the other hand, in a very sparse network such as most interurban motorway networks, only few alternatives are commonly available. Anyway, the number of routes that the traveller compares with one another and whose advantages and disadvantages he considers before making a choice appears to be very limited in most cases, between 2 and 6 routes and in few cases even less than 2 routes.

The main implication of these findings for modelling choice sets is that the size of the objective choice set is proportional to the network density, whereas the size of the subjective choice set depends on the limited ability of travellers in considering many alternatives. Because of this, in the case of modelling choice sets in multi-modal transport networks, which are rather complex and dense, large choice sets are expected to be generated (see the figures by Hoogendoorn-Lanser (2005) and Bovy, P.H.L. and Fiorenzo-Catalano, S. (2006)), whereas small choice sets are expected to be considered by the travellers. Then the challenge is to develop generation models capable of generating choice sets that are as close as possible to the subjective choice sets.

Table 4.7: Choice set size of generated and observed routes in several types of networks (sample).

Source	**Network**	Generated objective choice set	Reported subjective choice set	Reported joint subjective choice set	# of obs. # of cases
1	**2**	3	4	5	6
Van Der Waard (1988)	Urban public transport network in NL	1-4	1-4	-	4,588 1,863 individuals
Bovy and Stern (1990)	Bicycle network in Delft (NL)	40	4-5	15	40 1 OD pair
	Road network in NL (car drivers)	15	4	-	50 individuals
	Road network in Newcastle UK (motorists)	-	2-3	-	1,300 individuals
Fiorenzo et al. (2001)	Waterway network in NL	19 (average) 9.5 (median) 58 (max)	-	-	562 30 OD pairs
Ramming (2002)	Road urban network in Boston area	30 (median) 51 (max)	-	-	160 individuals
Benjamins et al. (2002)	Multi-modal corridor in NL	22 (max)	-	-	22 1 OD pair
Bliemer et al. (2004)	Motorway network in NL	1.66 (average) 14 (max)	-	-	181,567 109,292 OD pairs
Van De Reijt (2004)	Pedestrian network in railway stations in NL	-	1-4	-	534 (Delft) 180 (Breda) individuals
Hoogendoorn-Lanser (2005)	Public transport railway network in NL	63 (average) 376 (max)	2 (average) 6 (max)	-	511 individuals

NL=The Netherlands.
obs. = observations.

Outcomes of the choice process depend on choice set composition. Complex choice sets imply the application of complex choice models. For example, the characteristic of having trip alternatives with high degree of mutual overlap needs a complex choice model to be solved. In the case of multi-modal transport network the composition of choice set is expected to be rather complex, therefore in that case complex choice models should be considered.

4.6 Conclusion

This chapter introduced an extended theoretical framework for dealing with the phenomenon of choice sets. Although the framework is dedicated to the route choice context, the developed notions are applicable to other travel contexts as well, such as destination choice or mode choice.
Choice sets of individual travellers play an important role in analyzing travel choice behaviour. Choice sets are defined as the collection of travel options perceived available by individual travellers in satisfying their travel demand. The critical role of choice sets in choice modelling has given rise to profound research into choice set modelling in transportation fields, although largely confined to mode choice (see e.g. Swait (2001)). From a variety of studies (Swait, J. and Ben-Akiva, M.E., 1987) and (Swait, J. and Ben-Akiva, M.E., 1985) it is well known that the size and composition of choice sets do matter in case of model estimation and demand prediction. Incorrect choice sets (e.g. because of captivity) can lead to misspecification of choice models and to biases in predicted demand levels (see Williams, H.C.W.L. and Ortuzar, J.D. (1982) and Ortuzar, J.D. and Willumsen, L.G. (2001)).

Based on a theory of the mental processes involved in choice set formation and choice, the notions of universal set, master choice set, subjective choice set, consideration set, and the chosen route are introduced and defined.
Specific new contributions of this chapter are the introduction of various perspectives deemed necessary for a deeper understanding. These are in particular the perspective of the researcher vis-*à*-vis that of the traveller, and the perspective of the individual traveller vis-*à*-vis that of a spatially aggregated group.

A further contribution of the conceptual framework concerns a clear distinction of purposes for which a researcher needs to specify choice sets, such as analysis, estimation, and prediction purposes. This distinction is important because the different purposes pose different requirements to the comprehensiveness of the generated choice sets in terms of size and composition.
This thesis will focus on the generation of prediction choice sets at aggregate zonal level. In developing such generation algorithms we will evaluate the prediction quality of such algorithms through comparison with observations at both individual (address) and aggregate (zonal) level.

One of the basic hypotheses of this chapter is that choice set formation in the mind of travellers strongly relates to their preferences and experiences with respect to route choice.

The main implications of the findings about choice set size and composition for modelling choice sets is that the size of the objective choice set is proportional to the network density, whereas the size of the subjective choice set depends on the limited ability of travellers in considering many alternatives simultaneously. Because of this, in the case of modelling choice set in multi-modal transport networks, which are rather complex and dense, large choice sets are expected to be generated, whereas only small choice sets are expected to be considered by the individual traveller.

Outcomes of the choice process highly depend on choice set composition. Complex choice sets require the application of complex choice models. Choice sets with high degree of mutual overlap of their alternatives need appropriate advanced choice models to be solved. This is even more true for the case of multi-modal transport network for which the composition of choice sets has been shown to be rather complex.

Before turning to the topic of this thesis, that is generation algorithms for route choice sets, we will first devote some attention in the next chapter to the type of choice models available for handling route choice sets with some complexity because of size, multimodal composition of its routes, and their mutual overlap.

Chapter 5

Route choice modelling

5.1 Introduction

After having introduced the choice set concepts in the previous Chapter 4, the next step of this thesis deals with the behavioural and modelling aspects for route choice in transport networks. According to the modelling framework established in Chapter 2 (see Figure 2.11) the modelling of route choice will be performed on given route choice sets established in a choice set generation (CSG) step performed in advance of the choice modelling and network assignment phases. Although this modelling framework poses no restrictions on the way of route choice modelling, the route choice models we focus on in this thesis represent current state-of-the art, that is, are of the random utility type, predicting choice probabilities of routes available in the choice set.

One of the basic hypotheses followed in this thesis is that choice set formation in the mind of travellers (see Chapter 4) strongly relates to their preferences and experiences with respect to route choice. Our developed choice set generation approach (to be discussed in the next Chapters) therefore will be based on this hypothesis, meaning that route choice factors and behavioural parameters of travellers will determine the generation of route options for the choice sets via a utility or cost function, called route generation function. Since there exist significant differences in choice factors and preferences between user groups and between trip purposes, these route generation functions will be specific for user and trip categories.

It is thus to be expected that an analysis of route choice behaviour and of applicable route choice models may provide requirements and clues for the choice set generation models to be developed.

The purpose of this chapter is to provide insights into route choice factors and derive behavioural rules for route choice on the basis of which requirements for choice set modelling will be derived. In order to achieve these purposes, a concise overview of route choice factors of travellers in various networks and conditions is presented in

Section 5.2. This is followed by an overview of behavioural models for route choice with focus on random utility maximisation (RUM) models presented in Section 5.3. In Section 5.4.1 a classification of route choice models is provided with emphasis on models predicting future use of route alternatives on the basis of given choice sets. Among the choice models presented in this section, special attention is given to the RUM models based on a priori CSG because of many reasons. We recall that a priori generation of choice alternatives has a number of advantages. First, application of a priori CSG allows the adoption of more complex and more realistic utility functions (such as non-additive and non-linear utility functions). Secondly, a priori generated choice set gives more freedom in choice models to be adopted allowing the application of advanced route choice models (see this chapter). Third, explicitly generating choice sets gives the analyst additional insights into the supply of travel opportunities. A final important advantage is the lower computational effort compared to classical demand modelling approaches.

The main contribution of this chapter is presented in Section 5.5 that deals with the main factors that influence route choice behaviour and their implications for generating choice sets. Specific aspects for CSG approaches in uni-modal and multi-modal networks are analysed and the following main consequences for CSG are taken into account. First, different travellers may have different preferences (so-called taste variation) so that the population consists of a mix of several segments each with their own attribute weights which often are correlated with observable traveller characteristics. Second, the size of choice set depends on the type of the network and, for example, in multi-modal networks the size is even bigger than the size of choice set in uni-modal networks due to the many combinations among all route components. Finally, the combination of mode choice and route choice in a multi-modal network requires special attention in order to allow application of standard route choice models. The main conclusions of this chapter are summarised in Section 5.6.

5.2 Route choice factors of travellers

The factors influencing route choice of travellers relate to the traveller, to the trip to be made, and, most importantly, to the attributes of the available routes (Bovy, P.H.L. and Stern, E., 1990).

Personal factors relate to age, sex, occupation, income, vehicle availability, travel experiences, and the like. These factors determine the relative importance a traveller attaches to the various route attributes. For example, high-income groups attach a higher weight to time relative to cost than low-income people. These personal factors are represented by the parameters in the utility or resistance functions of routes. Trip factors relate among other things to trip purpose (work, business, shopping, etc) and trip distance. Also these factors are reflected in the parameter values of the utility or resistance function.

Because of apparent differences in choice factor importance between user groups and trip types it makes sense to adopt different route utility functions, and consequently different choice set generation functions for these groups.

Most important are the choice factors relating to the routes themselves such as travel time, road type, delays, etc. For a particular traveller making a particular trip these are the factors that determine his route choice.
For an overview of the factors identified in literature, see Bovy, P.H.L. and Stern, E. (1990), especially their Table 3.3.

Of special interest in our case is the multi-modal aspect of the network and its routes. The perception and relative importance of choice factors appear to be dependent on the type of transport mode. Several studies have investigated the importance of reasons for route choice regarding different transport modes such as for car, public transport, bicycle and pedestrians.

Table 5.1: Route choice factors related to modes for groups of travellers.

Factors	*Travellers*	*References* (sample)
Travel time Tolls Congestion delay Travel distance Road quality Road safety Number of turns	*Motorists*	Benshoof, J.A. (1970) Bovy, P.H.L. and Stern, E. (1990) Cascetta et al. (1996) Nielsen, O.A. (1996) Ramming, M.S. (2002)
Number of transfers Walking time for transfer Access and Egress time Waiting time at stop Waiting time for transfer Cost / Delay In-vehicle time	*PT users*	Bovy, P.H.L. and Stern, E. (1990) Van Der Waard (1988a) Nielsen, O.A. (2000) Axhausen et al. (2001) Cascetta (2001) Nuzzolo (2003) Hoogendoorn-Lanser (2005)
Travel time Travel distance Safety Pollution Weather protection Number of turns First leg's travel distance Number of curves	*Pedestrians*	Senevirante, P.N. and Morrall, J.F. (1986) Bovy, P.H.L. and Stern, E. (1990) Lam et al. (1995) Golledge, R.G. (1995) Golledge, R.G. (1997) Daamen (2004) Van De Reijt (2004) Van Der Waerden et al. (2004)
Travel time Travel distance Road surface Number of turns	*Cyclists*	Bradley, M.A. and Bovy, P.H.L. (1984) Bovy, P.H.L. and Stern, E. (1990)

Table 5.1 summarizes for several modes those factors mostly found in literature to be important for route choice of travellers. Contrary to the common believe that travel time is the only important reason for route selection, those studies have confirmed that most travellers consider several reasons for their route choices and different travellers have different reasons. Although travel time usually appears to be the most preferred reason, several others are important as well. Furthermore, those findings provide valuable insights into which criteria might be used for choice set formation, because we hypothesize that the criteria applied for route choice are related to the ones governing choice set formation.

One of the most relevant findings concerning the choice factors for motorists in Benshoof, J.A. (1970) is that the choice reasons are related to the trip purposes, the competence of the driver, and the length of the trip to be taken. In addition, 70% of respondents provided two or more reasons for their routes of which the following reasons appear to be the most important: travel time, congestion, and travel distance. However, other motivations connected with ”habit” (drivers prefer to drive along the same route), ”familiarity” (drivers prefer to drive in a familiar environment), road quality, and safety were mentioned as well.

The most important choice factors among urban public transport users are reported among others by Van Der Waard (1988a) and Van Der Waard (1988b). For each choice factor, the weight measuring its relative importance was estimated. Based on Van der Waard’s route choice research, it turns out that urban PT users give the highest weight to the number of transfers (preferring the route that provides the most direct trip) followed by walking time for transfer, access time, waiting time at stop, and waiting time for transfers; the egress and in-vehicle time appear to be less relevant factors. For inter-urban PT trips, Hoogendoorn-Lanser (2005) summarizes relevant choice factors.

According to the analysis performed by Bradley, M.A. and Bovy, P.H.L. (1984) about cyclists’ behaviour, the main choice factors for cyclists appear to be travel time, travel distance, road quality, congestion, and safety.

Senevirante, P.N. and Morrall, J.F. (1986) show that the most important factor for pedestrians is the shortest route even considering several trip purposes such as work, shopping, etc. Despite the relevance of the travel distance, this does not imply that the selected route is indeed the shortest one, because not only this factor but also other reasons related to habit, familiarity, road attractiveness, pollution, road safety, and weather protection were mentioned by the pedestrians. An extensive overview of the pertaining literature is given in Daamen (2004).

To give an example, recent studies carried out by Golledge, R.G. (1995) and Golledge, R.G. (1997) have investigated the criteria used by pedestrians in their route selection. After defining the following route choice criteria:

1. Shortest distance
2. Least time
3. Fewest turns
4. Most scenic/aesthetic
5. First noticed
6. Longest leg first
7. Many curves
8. Most turns
9. Different from previous route taken (variability)
10. Shortest leg first

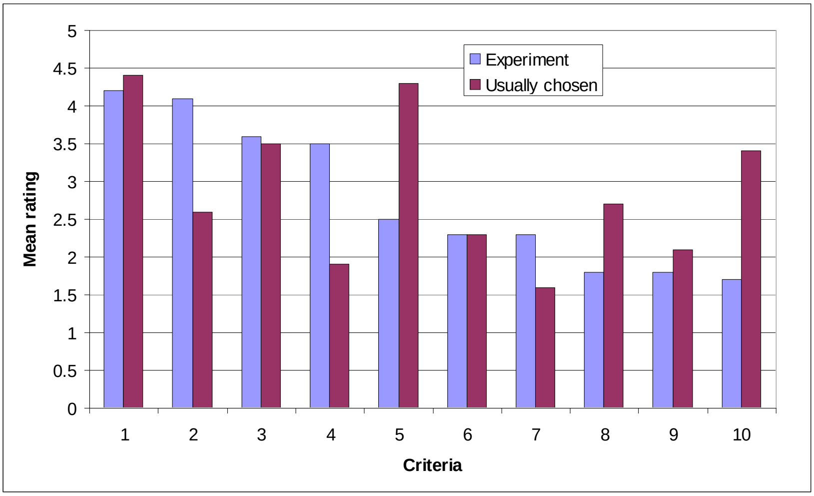


Figure 5.1: Comparison of stated and revealed criteria in route selection, and list of route choice criteria (source: Golledge, 1997).

Golledge compared the criteria used to choose a route in the actual trips with those implemented in the model. It was Golledge's hypothesis that people often are not aware or able of using the criteria used in almost all travel models, such as shortest route or least time. Several experiments in the laboratory and in the field have been carried out by Golledge to identify the criteria used in route selection. In the lab experiments subjects were asked to select a route in a variety of scenarios and give a score from 1 (completely unimportant) to 5 (extremely important) for each criterion considered in their route choice decisions.

Subjects were also asked to rate on the same five-point scale what criteria they usually adopted when choosing routes in their real world travel patterns. In this case the criteria that have the highest score are different from the ones stated in the experiments. Figure 5.1 depicts the choice criteria, the mean rating of each criterion (stated in the experiment) and the mean rating of the revealed criteria.

The consequence of the complex choice process described above is that the choice set involved in this process may be specific for each individual and it depends on many factors such as: network knowledge, travel experience, perception factors, and evaluation factors (see Figure 5.1).

The implications from the findings in the route choice literature for choice set generation are the following:

- the more important choice factors identified may also be assumed to influence the choice set formation process of the traveller, and thus should be considered as influencing variables in the choice set generation algorithms;
- there appear to be significant differences in choice factor importances among travellers and among trip purposes; this asks for specific generation functions (variables and their parameter values) in the CSG for different person and trip types.

5.3 Behavioural models for route choice

5.3.1 Overview

Route choice models described in this chapter attempt to reproduce traveller's choice behaviour. These models indicate the contribution of choice factors to the choice decision and can predict the conditional probability of choice of each alternative in the choice set depending on the choice conditions (choice set size and composition, attribute values, etc).

Most behavioural choice models today are based on the concept of *utility* from classical micro-economic theory assuming that the decision-maker's preference for an alternative is captured by a value, called utility, that is computed taking into account specific evaluation *aspects* of each alternative. Moreover, maximisation utility theory hypothesizes that the decision-maker selects the alternative from the choice set with the highest utility to him (Ben-Akiva, M.E. and Lerman, S., 1985).

The concept of utility associated with the alternatives plays an important role in the context of discrete travel choice models. Figure 5.2 summarizes the essential assumptions made by the various classes of choice models such as with respect to sources of uncertainty.

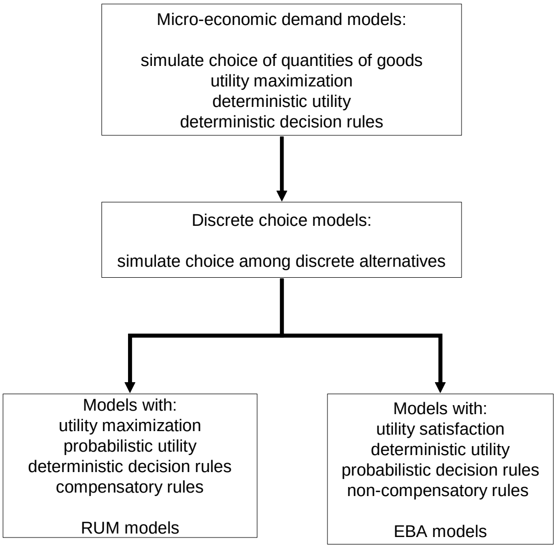


Figure 5.2: Classification of choice models based on behavioural assumptions.

Specific families of models can be derived, depending on the assumptions about the sources of the uncertainty (Ben-Akiva, M.E. and Lerman, S., 1985). Random Utility Maximisation (RUM) models, used intensively for travel behaviour analyses, are designed to capture uncertainty on the part of the modeller. These models are based on utility maximisation theory and assume deterministic decision rules, while the modeller's uncertainty is captured by random variables representing unknown utilities. Moreover, the RUM models have a *compensatory* nature, meaning that a chosen alternative may be worse than all others in terms of specific aspects of the alternative, but it can still be chosen if its overall utility is higher than the other alternatives. In other words, it is possible to compensate a deficiency in any of the aspects with better than average performance on other aspects.

The assumption of utility maximisation, which is fundamental to the RUM, has attracted criticism from some writers. Herbert Simons's contribution to the critical literature has been of particular significance. Simon, H.A. (1990) argued that "*Because of the limits on their computing speeds and power, intelligent systems* [such as physical symbol systems, computers and the human brain] *must use approximate methods to handle most tasks. Their rationality is bounded.*" Simon proposed the principle of

bounded rationality for individual decision making. Here, instead of seeking of maximise utility, the individual employs simplifying heuristics, achieving an approximate solution with modest effort. Models based on Simon's ideas usually employ sequential elimination processes in conjunction with criteria of acceptability, thereby precluding 'trade-off' between aspects of the alternatives.
Models with stochastic decision rules are often known as *non-compensatory* choice models, in contrast to the compensatory structure of most RUM models, since they do not allow the compensation of negative attributes with positive ones. An example of such model is the Elimination-by-Aspects (EBA) model, proposed by Tversky (1972), which assumes a deterministic utility and a probabilistic decision process postulating a hierarchy of aspects to be considered sequentially each of which successively must meet a threshold value in an alternative, in order for that alternative to be considered or chosen.

Although the proposed multi-modal demand modelling architecture poses no restrictions on the way of route choice modelling nor on the type of choice prediction models to be used, the route choice models, we focus on in this thesis, represent current state-of-the art that are the random utility type predicting choice probabilities of routes available in the choice set. It has been shown that non-compensatory models are at present mostly research tools and are not widely used in practice, furthemore, a properly specified RUM model can very often satisfactorily approximate the choice behaviour. Therefore, in the next section, we will focus on basic assumptions for the RUM models, while the rest of this chapter will deal with some examples of RUM models for route choice.

5.3.2 Random utility maximisation models

In this section we summarize the description by Cascetta (2001) of RUM models that assume, as in classical micro-economic theory, that every individual is a rational decision-maker, has a perfect discrimination capability, and maximizes his utility in making his choices. Basic assumptions of these models are as follows. The generic decision-maker, in making a choice, considers a limited number of mutually exclusive alternatives that make up his *consideration set*. The decision-maker assigns a perceived (dis)utility, or *attractiveness* to each alternative belonging to his consideration set and selects the alternative maximizing his utility. The utility assigned to each alternative depends on a limited number of measurable characteristics, or *attributes*, of the alternative corresponding to the evaluation aspects or factors considered by the individual in selecting the alternatives. The analyst, however, is supposed to have incomplete information about the real choice conditions and does not know the utility assigned to each alternative by the decision-maker with certainty, which leads to uncertainty in the model and therefore must be represented by a random variable. Manski (1977) identifies four different sources of uncertainty about the utility on the part of the modeller: unobserved alternative attributes, unobserved individual preferences (called

”unobserved taste variation”), measurement errors, and proxy, or instrumental, variables. In addition to the uncertainty about the decision maker’s utility, the information about the choice set available to the decision maker might be incomplete.

On the basis of the above assumptions, it is usually not possible to predict with certainty the alternative that the generic decision-maker will select. However, it is possible to express the probability of selecting an alternative given the consideration set. In a generic case of travel choices such as destinations, modes, routes etc. the perceived utility can be expressed by the sum of a deterministic observable part and an unknown random part.

Specifically, the utility U_{kq} of route alternative k belonging to the modelled consideration set CS_q, which is subset of a master choice set MS, to an individual q is modelled to reflect this uncertainty, and is therefore given by:

$$U_{kq} = V_{kq} + \varepsilon_{kq} \qquad k \in CS_q \subseteq MS \tag{5.1}$$

with V_{kq} the measurable (or systematic) part of the utility function, and ε_{kq} a random term capturing the various uncertainties. The route alternative with the highest utility is supposed to be chosen. Therefore, the probability that the route alternative k is chosen by the individual q from his consideration set CS_q is:

$$P_{kq} = P[U_{kq} = \max_{h \in CS_q} U_{hq}] \tag{5.2}$$

Consideration sets are in principle individual-specific sets.

The derivation of random utility discrete choice models is based on the specification of the random utility term. Specific assumptions about the random term and the deterministic term will produce specific models. We present here the most usual assumptions adopted in practice about the deterministic part of the utility and the random part.

The utility of each alternative is a function of observable attributes of the alternative itself and maybe also of the decision-maker. We can write the deterministic part of the utility that individual q is associating with alternative k as:

$$V_{kq} = V_{kq}(X_{kq}) \tag{5.3}$$

where X_{kq} is a vector containing the values of the considered *attributes*, both of individual q and alternative k. The function defined in 5.3 is commonly assumed to be linear in the parameters, that is, if a set of Y attributes are considered:

$$V_{kq}(X_{kq}) = \beta_{1q}X_{kq}(1) + \beta_{2q}X_{kq}(2) + \ldots + \beta_{Yq}X_{kq}(Y) = \sum_{y=1}^{Y} \beta_{yq}X_{kq}(y) \tag{5.4}$$

where $\beta_{1q},...,\beta_{Yq}$ are parameters related to individul q to be estimated. However nonlinear effects can still be captured in the attributes definition.

In principle, attributes and their parameters are alternative-specific (such as train time or frequency, bus time or frequency, car time, walking time); however, in many applications for some attributes so-called generic attributes and parameters may be used (such as travel time or cost).

The analyst has to identify the attributes of each considered alternative that are likely to affect the choice of the individual.
In addition to the deterministic part of the utility function, the properties of the model are determined by the unknown random part of the utility. We can rewrite equation 5.2, as:

$$P_{kq} = P[V_{kq} + \varepsilon_{kq} \geq \max_{h \in CS_q} (V_{hq} + \varepsilon_{hq})] \tag{5.5}$$

Only for very specific random distributions this probability can be solved analytically, such as with so-called extreme value distributions. In the special case where the random terms ε_{hq} are assumed to be independently and identically Gumbel-distributed, the well-known MultiNomial Logit (MNL) model results:

$$P_{kq} = \frac{e^{(\mu \cdot V_{kq})}}{\sum\limits_{h \in CS_q} e^{(\mu \cdot V_{hq})}} \tag{5.6}$$

where μ is the scale parameter. If the error terms have e.g. a normal distribution a probit model results for which no closed-form expression is available; calculation of the choice probability then becomes very involved. Among the many potential choice models that can be derived dependent on the random error specification, we elaborate below (in Subsection 5.4.3) on the most popular. The models within the logit family are based on density functions of the maximum of several Gumbel-distributed random variables, whereas probit and probit-like models are based on normality assumptions of the error term.

5.4 Route choice models

The route choice problem of a traveller involves the selection of a route from a given origin to a given destination when faced with a transportation network. From a traveller's point of view, a route from origin to destination in a real network might be defined as a collection of links (representing streets, walkways, railways, etc.) and nodes (representing intersections, stations, etc.). In a model network, such as our supernetwork, the term path might be used instead of route, however from a researcher's perspective both terms might be considered as synonym. An overview of models for analysing route choice is given in Batley et al. (2001).

In contrast to other choice problems such as destination, mode or location choice, a route through a network is a more complex object of choice. This is the more true for a multi-modal route.

First of all, a route consists of a multitude of constituting links connected by nodes. Each of these elements has a set of attributes that contribute to the routes disutility. For some attributes, such as for example, distance and time, the attribute value of the route is the simple sum of the corresponding attribute values of its constituting links and nodes. However, there are attributes for which this additivity does not hold such as with PT fares and road tolls; these attributes are non-linear in the link or sub-route values. Often the unit costs of a trip decrease with its length.

Secondly, there exist attributes at route level that do not have a corresponding attribute at link or node level and thus cannot be derived from the constituting route elements. Examples of such route attributes are winningness, hilliness, and angularity.

Another important feature of routes is that they may partly overlap with other routes in the choice set. The degree of overlap with other routes is another route attribute without a correspondence at link or node level. In addition, physical overlap of routes strongly influences the probability of choice that needs to be taken into account in choice modelling (see Subsection 5.4.2).

Finally, typical for route choice is the large number of routes in the choice sets. With some model types, this poses difficulties in estimating models and in predicting probabilities.

Because of the chosen approach of a priori route set generation in advance of choice modelling and network assignment, these typical characteristics of routes can be captured relatively easily. However, these characteristics also pose special requirements from the choice models (see also Bovy, P.H.L. and Fiorenzo-Catalano, S. (2006) and Bovy, P.H.L. (2007)).

As shown by Expression 5.3 the systematic part of the utility function is defined in terms of the attributes of the alternative and perhaps of the individual, therefore it can be stated that each route alternative in the choice set may be characterised by an own set of attributes.

5.4.1 Classification of route choice models

Basic points of departure in route choice modelling are that travellers because of all kinds of reasons behave differently and will choose different routes in equal circumstances. There is thus a distribution of trip makers over available routes. In addition, because of uncertainty on the part of the modeller only probabilities of choice can be predicted which means that single trips will be split over available alternatives according to these probabilities, summing up to one.

To accomplish this, probabilistic route choice models need to be applied. Models based on random utility maximisation theory automatically have this property; therefore RUM models might be applied for modelling the route choice process.

In modelling the decision making process of route choice in a network, two specific aspects might be clearly distinguished by the researcher:

1. the composition of the choice set;
2. the type of modelling approach for the choice of an alternative from the given choice set.

Whereas in a traveller's mind these processes might be mixed up and not clearly distinguished, these two steps might be clearly separated in a modelling context. For a discussion see Bovy, P.H.L. (2007).
In route choice modelling, we can distinguish two streams of choice modelling approaches: those that require explicit a priori choice set specification and those that do not. In the latter case, the specification of routes and their probabilities result from the route search process. This implies that non-chosen routes are not identified explicitly. In the former case, we can subdivide the choice models according to the type of choice set that is taken as input. This may be for example the master choice set (see definition in Chapter 4) containing all route alternatives available to all travellers, or the subjective choice set containing all relevant route alternatives that are considered feasible and available to particular decision-makers dependent on their personal attributes or trip type. For each type of choice set generated in a first step, a specific route choice model might be applied in a following step (see Figure 5.3).

Examples of route choice models without a priori CSG are Dial's model (Dial, R.B., 1971), most equilibrium assignment models and the stochastic version of the Multi-Nomial Probit model (MNP) model (Bovy, P.H.L., 1990). In modelling approaches without a priori CSG, only chosen route alternatives result from the choice modelling. In Dial's model for example, a route will have a non-zero probability only if it satisfies certain strong criteria such as that every portion in it takes the traveller further from the origin and/or closer to the destination. Dial's model performs a route search (based on link selection) and a choice probability computation simultaneously. In stochastic MNP only those routes will receive a non-zero probability that are the optimal route after a randomization of the utility function of the network links. Both methods rely on optimal path search in the network and require additivity of link utilities. Such approaches are also known as link-based approaches. In principle, both approaches can also be applied to given choice sets, however without benefiting from the advantages that explicit choice sets offer.

Conversely, route choice models based on a priori CSG approach guarantee better control of the ”feasibility” of the generated routes and introduce more behavioural elements into route choice set modelling; they are also known as route-based approaches.

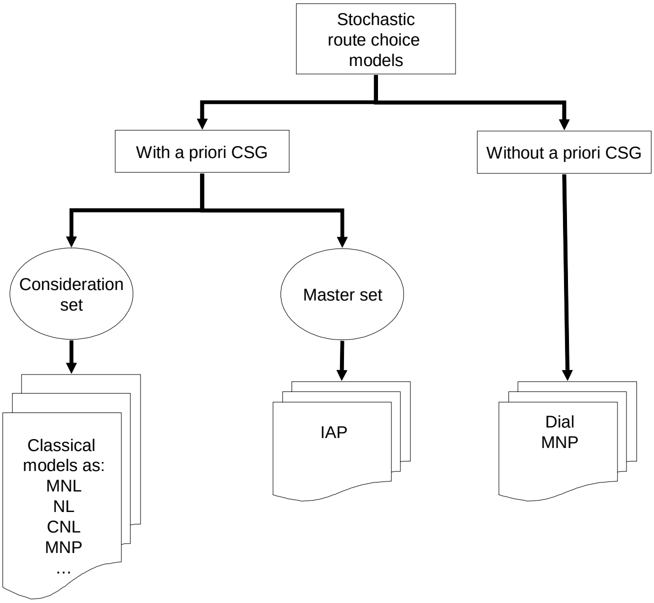


Figure 5.3: Classification of route choice models with respect to adopted choice sets.

These approaches offer nearly any freedom in the utility function specification (non-additivity, non-linearity), the type of attributes to be included (route attributes), and the type of choice models to be adopted. Of course, these advantages are obtained at the expense of greater conceptual and computational complexity. (However, it should be emphasized that no systematic analysis of the computational complexity and memory requirements of the two approaches exists yet and that the literature seems to suggest a tendency towards route choice models based on a priori CSG).

Depending on the route choice model to be applied, the a priori CSG approaches might generate either a master set or a consideration set. For example, adoption of the approach proposed by Manski (1977), with unconditional probability formulation, so far has been limited to very simple choice situations with a master set of only a few alternatives such as mode choice (see e.g. Swait, J. and Ben-Akiva, M.E. (1987) and Louviere, J.J. and Hensher, D.A. and Swait, J.D. (2003)). It seems impossible to adopt this full probabilistic approach to route set generation because of the sheer size of master sets in networks, even after strong selection. If a master set for a particular OD pair consists of about 10 routes (which is a very small size indeed), calculating the route probabilities seems feasible. However, since 10 routes imply $2^{10} - 1 = 1023$ potential consideration sets, determining these set probabilities and using these, for example, in a choice model estimation exercise seems beyond current possibilities. An example of the specification of the Manski approach is given by Morikawa (1995) in which the choice set generation model is a random constraint model that has a non-compensatory nature among multiple constraints and an alternative is included in the choice set if and only if all latent conditioning measures of the alternative satisfy the criteria.

Another approach that requires the master choice set generated through an a priori CSG approach (Figure 5.3) is the Implicit Availability/Perception (IAP) model (Cascetta & Papola, 2001). In this model the probability of choice set membership of an alternative (of a deterministic chocie set) enters the utility function of the choice model where a low membership probability of an alternative naturally leads to a decrease of its choice probability. The IAP model introduces a convenient way to incorporate awareness of routes into route choice modelling. In the IAP model, availability and perception of the alternatives is implicitly modelled as a heuristic function of alternative and decision-maker attributes as part of the overall utility function. The IAP model assumes that all routes belonging to the master set might be available to every traveller, but it uses a correction term to decrease the route's share to reflect the possibility that travellers are unaware of that route, or unable to use it. From a practical point of view, the IAP approach is quite simple and straightforward, but there is a lack of reliability because the "utility" and "availability" attributes are taken into account jointly assuming that the choice set for the decision-maker is equal to the master choice set.

Finally, examples of route choice models requiring the consideration set generated through an a priori CSG approach are the classical RUM models (such as MultiNomial Logit (MNL), Nested Logit (NL), Cross-Nested Logit (CNL), Generalized Nested Logit (GNL), PCL, C-Logit, PS-Logit, Error Components Logit or hybrid logit (ECL), and MultiNomial Probit (MNP), etc.) to be dealt with in the following subsection where we will focus on RUM models that require a priori CSG approaches for their solubility.

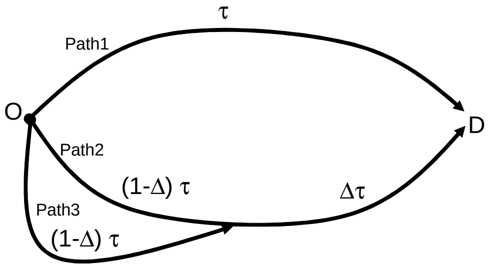
Even within the framework of RUM models, correct route choice modelling presents difficulties. As noted in (Ben-Akiva, M.E. and Bierlaire, M., 1999), the route choice problem is characterised by very large choice sets, of which only a small part will actually be considered and used by decision makers, and by correlation between choice alternatives due to route overlap. These characteristics represent non-trivial modelling challenges and it is in particular this correlation of alternatives requiring methodological innovation in most of the choice models.

5.4.2 Correlation among route alternatives: the overlapping problem

Since route alternatives in general and multi-modal alternatives in particular show significant degrees of physical overlap (common links and nodes), utilities of routes are correlated such that choice modelling should account for that in order to make correct predictions. This represents non-trivial challenges for route choice modelling being only very recently solved satisfactorily. Explicit choice set generation contributes significantly to the correct handling of the overlap correlations since the degree of overlap can be quantified easily if the choice set has been specified.

In this subsection, an example of correlation between route choice alternatives is given using the network shown in Figure 5.4. The choice set of travellers going from origin O to destination D consists of three paths of equal length, two of which (Path2 and Path3) are overlapping to a certain degree (indicated by Δ).

Let us assume that the travel time on Path1 is equal to τ, on the overlapping part $\Delta\tau$, and that the travel time on the small sections is $(1-\Delta)\tau$. Then, the route choice probability predicted with the MNL model (formula 5.6) is an equal share of 33% for each of the three routes for any value of Δ. The MNL result is consistent with our intuition only when the degree of overlap Δ is infinitesimally small ($\Delta \simeq 0$) and the two overlapping routes are quite different (the overlapping part is quite small). However, when Δ is approaching to 1 and the two overlapping routes are quite similar (the overlapping section is quite large), we expect the probabilities of the three paths to be close to 50%, 25% and 25%.



Overlap approaches 100% (Δ ~1), prob. close to 50% 25% 25% of share
Overlap approaches 0% (Δ~0), prob. close to 33% 33% 33% of share

Figure 5.4: The overlapping problem in route choice.

Fortunately, there are currently advanced choice models available that can correctly deal with overlap of routes.

An extensive account of the overlapping problem in car route choice in urban road networks is given by Ramming, M.S. (2002). An empirical analysis of the degree of physical overlap in multi-modal networks is given in Hoogendoorn-Lanser (2005). An overview of solution approaches for correctly handling routing overlap in transport networks is given in Hoogendoorn-Lanser, S. and Van Nes, R. and Bovy, P.H.L. (2005). A quantitative statistical comparison of various modelling approaches to include overlap in the estimation of choice models can be found in Hoogendoorn-Lanser (2005). These empirical and modelling analyses were possible thanks to the explicit generation of (master) choice sets.

5.4.3 Review of RUM models for route choice based on a priori CSG

In this section we will focus on the analysis and comparison of route choice models for prediction purposes. Route choice behaviour might be characterized by pre-trip choice, in which travellers consider, before leaving the origin, available information obtained from previous trips or supplied by a user information system and adaptive choice, in which, besides previous information, travellers consider situations that occur during the trip itself.
Given the wide variety of RUM models that theoretically may be applied, we will focus on those that are specifically suitable for the route choice problem, and among those we will restrict to static models with assumptions on pre-trip choice in which the whole route is chosen before starting the trip.

Most models considered here are from the Logit family such as MultiNomial Logit (MNL), Nested Logit (NL), Cross-Nested Logit (CNL), Generalized Nested Logit (GNL), PCL, C-Logit, PS-Logit, Error Components Logit or hybrid logit (ECL), and are extended with the MultiNomial Probit (MNP). Their model formulations and functional specifications have been elaborated in depth in Ben-Akiva, M.E. and Bierlaire, M. (1999), Batley et al. (2001), Cascetta (2001), and are beyond the scope of our discussion. Since the route overlapping problem is a crucial point for route choice modelling, a classification of the route choice models is made on the basis of route overlap criteria. Which of the models are more suitable for dealing with the route overlapping problem, and how do these route choice modelling approaches cope with the problem of route overlap? Figure 5.5 concisely shows the kind of answers to this question and the resulting classification.

There are four ways to deal with the overlapping problem in route choice models (see also Hoogendoorn-Lanser, S. and Van Nes, R. and Bovy, P.H.L. (2005)):

- overlap is not taken into account (MNL, NL, IAP);
- common links define a nesting structure (CNL, GNL, PCL);
- common links determine a dedicated additional utility component (C-logit and PS-logit);
- common links specify a dedicated variance-covariance structure of the error terms (MNP, Probit with Logit Kernel).

MultiNomial Logit (MNL) is the simplest form of the Logit family. MNL assumes that the error terms of the alternatives are independently and identically Gumbel-distributed. In a typical network (especially in a multi-modal network), alternative routes may consist of many common links, thereby establishing complex patterns of correlation among route alternatives. If MNL is applied to such a network, it would

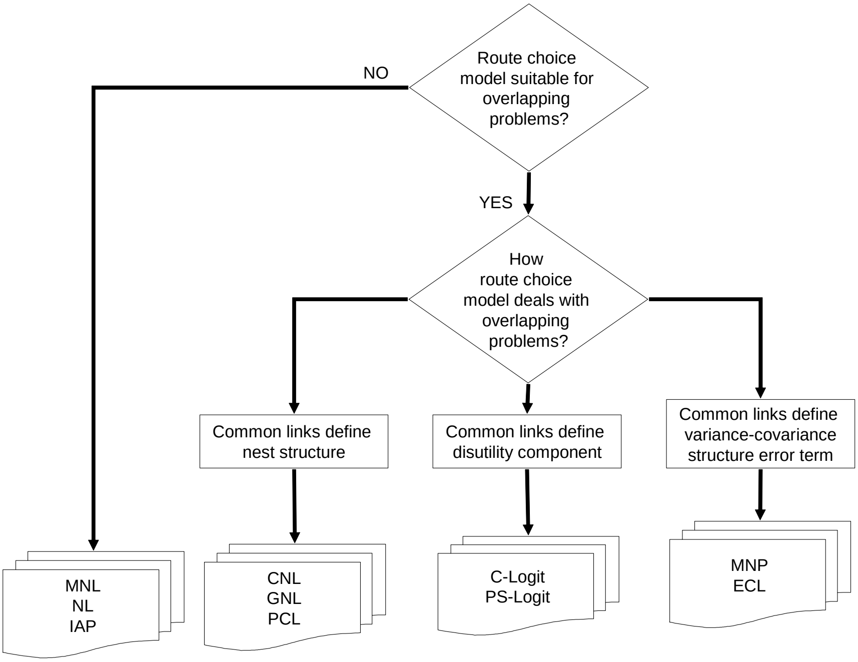


Figure 5.5: Classification of route choice models suitable for the overlapping problem.

overestimate the use of routes with common links and would result in seriously mispredicted link flows. Nevertheless, this erroneous model has been applied widely for route choice modelling (for example, Dial, R.B. (1971)). IAP is a MNL model with an extension for implicit route availability modelling.

Nested Logit (NL) models are designed to capture choice problems where alternatives within each nest are correlated; no correlation across nests can be captured however. NL models have more flexible error structures than MNL and can account for some similarities between alternatives if these can be assigned exclusively to branches in a hierarchical nesting structure. However, when alternatives cannot be partitioned into well-separated nests to reflect their correlation, NL models are not valid. This is the case of most route choice problems since each route alternative can only be assigned to one single nest (Hoogendoorn-Lanser, 2005).

Cross Nested Logit (CNL), which in the case of route choice also is known as link-nested logit, uses the common links of routes as the basis for nesting. Overlapping alternatives are handled by a separate nesting parameter. The CNL is a direct extension of the nested logit model, however where each alternative now may belong to more than one nest with different degrees of membership. On the other hand, CNL is a restricted special case of a GNL model. Common links form upper level nests while routes are lower level alternatives (captures correlation among routes that share a link). However, for a realistic size network, a realistic number of links per route, and especially for a more complex network such as a multi-modal network, the nesting structure of the CNL model will become extra-ordinarily complex (Cascetta, 2001).

Generalized Nested Logit (GNL) extends CNL by allowing the dissimilarity parameters to be different for different nests. For route choice applications it is not clear how the nesting should be applied. If link-based nesting is used, as in the CNL application, variation of the nest parameters is not welcome, because it implies that there is some inherent variability between the links, and that nest parameters values for new links may not be easy to obtain. For those reasons it can be concluded that GNL is not suitable for route choice modelling, except in special cases, see Hoogendoorn-Lanser (2005).

The paired combinatorial logit (PCL) model is a special case of the GNL model allowing correlation of the error terms by the inclusion of a pair-wise similarity parameter. The similarity parameter can represent not only overlapping network links, but as well potentially other route similarities such as facility type, number of turns, levels of congestion and signage. Gliebe, J.P. and Koppelman, F.S. and Ziliaskopoulos, A. (1998) proposes use of the PCL model to overcome the IIA property of the MNL and to provide a more computationally tractable mechanism for route choice problem than MNP. The similarity parameter can be structured to represent the overlap between any pair of route alternatives, based on the distances, travel times or other attributes of links common to pairs of routes. In the precursor version of the supernetwork approach, the PCL-model was implemented for multi-modal route choice predictions. Subsection 5.4.5 will describe this application of PCL.

The MultiNomial Probit model (MNP) takes account of correlations between alternatives by allowing covariance between the error terms for all pairs of alternatives in the choice set. To this end the covariance is expressed in terms of percentage spatial overlap such that simple functional relationships are employed (Bovy, P.H.L., 1990). The MultiNomial Probit model (MNP) is very difficult to solve analytically due to the high complexity of its formulation. However, MNP might also be solved by simulations and in this case it might be employed in most contemporary practical applications on a network basis using a priori CSG or not.

The Error Components Logit (ECL) or Mixed Logit, also known as Hybrid Logit model can be regarded as a mix between probit and logit models. It is intended to bridge the gap between logit and probit models that is to combine the advantages of both of them. ECL does not suffer from the IIA property. Indeed complex patterns of route similarity can be specified through appropriate specification of the density function (Ramming, M.S., 2002). The error components logit model seems free of theoretical disadvantages, but (like the Probit model) requires sampling to calculate the choice probabilities being a clear drawback in an operational model. This model has been applied successfully by Ramming in his Boston route choice study to capture overlap.

Cascetta et al. (1996) specified a route choice model featuring a modification of the MNL model called C-logit, which penalizes the alternative's utility function in the standard MNL by a ”commonality factor”. In fact, the MNL is corrected by subtracting a commonality factor term from the route's utility. The commonality factor represents the overlapping link length of one route with respect to all other ones together.

$$P_{kq} = \frac{\exp(\mu \cdot V_{kq} - \delta \cdot CF_k)}{\sum\limits_{h \in CS} \exp(\mu \cdot V_{hq} - \delta \cdot CF_h)} \tag{5.7}$$

Expression 5.7 shows the choice probability of alternative k for individual q (or user class q), where μ is the scale parameter, δ is the weight coefficient of the route overlap and CF_k the commonality factor of path k, and it is directly proportional to the degree of similarity of path k with other paths h in the CS. Compared to the ECL and MNP approaches, C-logit has to be considered as an approximation. Cascetta et al. (1996) defined the term CK_k as:

$$CK_k = \gamma_0 \ln \sum_{h \in CS} \left(\frac{d_{kh}}{\sqrt{d_k \cdot d_h}} \right)^{\gamma_1} \tag{5.8}$$

where d_{kh} is the length (cost) of links in common to paths k and h, d_k and d_h are the overall path lengths (costs) of paths k and h respectevely, and γ_0 and γ_1 are positive parameters to be estimated or calibrated.
The commonality factor reduces the probability of choosing routes that overlap and increases the probability of choosing an independent route. Cascetta (2001) offers a few variations of the C-logit in which the commonality factor takes on different structural forms. Versions of the C-logit presented to date represent similarities between routes based on overlapping link lengths, but could be extended to account for other similarities on non-overlapping links. The lack of theory or guidance on which form of commonality factor should be used is a drawback of the C-Logit method. Applications of this model have been performed for the choice of inter-city routes by truck drivers in Italy (Cascetta et al., 1996). However, according to applications of Ramming, M.S. (2002), C-logit offers implausible results.

The *path size logit* (PS-logit) model can be viewed is an improved variant on the C-logit model that addresses the issue of the single commonality factor. Path size logit can correct MNL by adding a ”path size” term to route utilities. Similar as C-logit, PS-logit is an approximation since it uses a single path size value for a route irrespective of the varying degrees of overlap of that route with the other routes in the choice set.

$$P_{kq} = \frac{\exp(\mu \cdot V_{kq} + \delta \cdot \ln PS_k)}{\sum\limits_{h \in CS} \exp(\mu \cdot V_{hq} + \delta \cdot \ln PS_h)} \tag{5.9}$$

Equation 5.9 shows the choice probability of alternative k for individual q (or user class q), where μ is the scale parameter, and PS_k is the path size (give the PS-formula). The PS-logit model is an application of discrete choice theory for aggregate alternatives representing an effort to incorporate behavioural theory in the C-logit adjustment process. As in the C-Logit model, PS-logit adds a correction term to the utility of alternative routes. A route with no overlapping links needs no utility adjustment and has a *size* of one. The size of partially overlapping routes can be thought of as composed of the size of links that then are weighed by some appropriate measure, such as link's percentage contribution to total route length. We present the so-called *exponential path-size* formulation introduced by Ramming, M.S. (2002):

$$PS_k = \sum_{a \in \Gamma_k} \frac{l_a}{d_k} \cdot \frac{1}{\sum\limits_{h \in CS} \left(\frac{d_k}{d_h}\right)^{\gamma} \cdot \delta_{ah}} \tag{5.10}$$

where l_a is the length of link a, d_k is the total length of route k, γ is the size assignment parameter, Γ_k is the set of links of route k and δ_{ak} is a binary variable (0/1) indicating whether link a is part of route k or not. For an in depth discussion on PS-logit and its application in route choice analysis, in particular to multi-modal networks, see Hoogendoorn-Lanser, S. and Van Nes, R. and Bovy, P.H.L. (2005). This PS-logit choice model has been implemented in our currently operational model TRANSFER (Carlier et al., 2005).

Of special concern is route choice modelling in public transport service networks. In order to correctly model route choice in cases of parallel lines the concept of route strategy has been developed that can be easily implemented using the hypernetwork concept (see Chapter 3). This concept allows a link-based approach to multiple route choice. However, if a priori generated choice sets are applied with properly considering the parallel lines, the strategy concept will be automatically implemented using the choice models describe above. This shows another advantage of the use of a priori generated choice sets.

5.4.4 Mixed Logit models

Mixed Logit also known as Hybrid Logit or simply Logit Kernel has been introduced by Ben-Akiva, M.E. and Bolduc, D. (1996). The Mixed Logit model is a discrete choice model in which the disturbances (of the utilities) consist of both a Probit-like portion and an additive i.i.d. extreme value portion (i.e. a Multinomial Logit disturbance). The result is an intuitive, practical, and powerful model that combines the flexibility of Probit (and more) with the tractability of Logit. For this reason, Mixed Logit has been deemed the "model of the future" and is becoming extremely popular in the literature.

Mixed Logit (ML), such as the ECL model can be used to capture the route overlap (as shown in Section 5.4.2). However, a much more central use of the Mixed Logitt model is that it allows high flexibility by specifying taste coefficients to be randomly distributed across individuals, and it provides the modeller a tremendous range to specify individual unobserved heterogeneity.

In fact, the Mixed Logit specification, known as random parameter specification, involves specifying each beta parameter associated with an attribute of the alternative as having both a mean and a standard deviation, i.e. it is treated as a random parameter instead of a fixed parameter (a fixed parameter essentially treats the standard deviation as zero such that all the behavioural information is captured by the mean).

The presence of a standard deviation of a beta parameter accommodates the presence of preference heterogeneity in the sampled population. This is often referred to as unobserved heterogeneity. While one might handle this heterogeneity through data segmentation and/or attribute segmentation, the challenge of these segmentation strategies is in picking the right segmentation criteria. A random parameter representation of preference heterogeneity is more general; however, such a specification carries a challenge in choosing the most appropriate parameter's distribution.

Recent research on the mixed choice models for public transport (Mabit, S.L. and Nielsen, O.A., 2006) clearly shows that randomising parameters in the choice models improve the fit to the observed choices and finds a drastic improvement of the choice set.

5.4.5 Application of PCL in multi-modal corridor

In the precursor implementation of the supernetwork approach (Benjamins et al., 2002), the Paired Combinatorial Logit (PCL) model was implemented as the route choice model. This subsection describes its specification and some results.
The PCL-model was selected because of its analytical simplicity compared to for example MultiNomial Probit. The PCL-model solves the overlap problem by looking at all pairs of routes in the choice set and determining the mutual overlap of within each pair.

The PCL-model (described in detail in Koppelman, F.S. and Wen, C. (2000) expresses the probability of choosing alternative k as the sum of choice probabilities of all pairs of alternatives multiplied by the conditional probability of selecting alternative k given that that alternative pair k, h was selected first. The PCL model gives the following choice probability for route k:

$$P(k|CS) = \sum_{h \neq k, h \in CS} P(k|kh) \cdot P(kh) \tag{5.11}$$

where the probability of choosing the pair kh is given by:

$$P(kh) = \frac{\left(\exp\left(\frac{\mu \cdot V_k}{1-\eta_{kh}}\right) + \exp\left(\frac{\mu \cdot V_h}{1-\eta_{kh}}\right)\right)^{1-\eta_{kh}}}{\sum_{r=1}^{|CS|-1} \sum_{p=r+1}^{|CS|} \left(\exp\left(\frac{\mu \cdot V_r}{1-\eta_{rp}}\right) + \exp\left(\frac{\mu \cdot V_r}{1-\eta_{rp}}\right)\right)^{1-\eta_{rp}}} \tag{5.12}$$

and the conditional probability of choosing alternative k from the pair kh, is given by:

$$P(k|kh) = \frac{e^{\left(\frac{\mu \cdot V_k}{1-\eta_{kh}}\right)}}{e^{\left(\frac{\mu \cdot V_k}{1-\eta_{kh}}\right)} + e^{\left(\frac{\mu \cdot V_h}{1-\eta_{kh}}\right)}} \tag{5.13}$$

In these expressions V_k is the generalised cost of route k, $|CS|$ is the cardinality of the consideration set, i.e. the total number of routes in the CS, and η_{kh} is a measure of similarity of routes k and h, while μ is the scale parameter (in the applications et to 1). PCL is consistent with random utility maximisation if $0 \leq \eta_{kh} < 1$.

The PCL model deals with route overlap through its similarity parameters that are defined for each pair of routes. A low value of this parameter implies limited dependency among alternatives. The following measure of similarity η has been used:

$$\eta_{kh} = \frac{d_{kh}}{\sqrt{d_k \cdot d_h}} \qquad 0 \leq \eta_{kh} < 1 \tag{5.14}$$

with
d_{kh} is the length of the common part of routes k and h;
d_k, d_h are the lengths of routes k and h respectively.
The utility function V_k of route k was specified as a sum of generalised costs of the constituting links a of k. The generalised costs variables were mode and link type specific while the corresponding parameters were travel purpose specific. The following variables were used:

- in-vehicle time, by mode;
- waiting time before boarding, by mode;
- length of the link;
- transfer penalty on transfer links;
- parking cost on connector links.

Parameter values were taken from literature. Results of the PCL-application can be found in Section 2.5.2 in Table 2.3. The example table shows calculated choice probabilities (in percent) specified by user group 1 to 6 for the alternatives in a large choice set (N=22) of an OD pair in the Dordrecht-Rotterdam Corridor.

5.4.6 Suitability of models

In evaluating the discussed models for their suitability to route choice modelling in multi-modal networks several criteria are at stake:

- Does the model sufficiently account for the overlap among routes?
- Does the model sufficiently account for preference heterogeneity among individuals?
- Does the model have overlap parameters that need to be estimated?
- How is the analytical tractability of the model?

The best way of accounting for overlap and taste variation is by using a dedicated variance-covariance matrix such as in MNP and Logit Kernel. The specification and estimation of correlation parameters however is very cumbersome. Also application of these models for prediction of choice probabilities requires much effort.
A second-best alternative then is PCL that combines correct overlap modelling with analytical tractability as shown in the application reported in Subsection 5.4.5.
Another feasible approach, although it is an approximation compared to the former two, is PS-logit. This choice model is successfully applied to route choice in multi-modal networks (Hoogendoorn-Lanser, 2005) and has been implemented in our currently operational multi-modal demand model TRANSFER (Carlier et al., 2005).

5.5 Implications of route choice models for CSG

This section summarizes important aspects of route choice modelling considered relevant for the choice set generation approach. The aim of this section is to review the main factors that influence route choice behaviour and derive implications for the set up of CSG algorithms.

As shown in the literature (e.g. Bovy, P.H.L. and Stern, E. (1990), Nielsen, O.A. (1996), and Nielsen, O.A. (2000)) a variety of aspects contribute significantly to the explanation of route choice behaviour. These include aspects describing the route alternatives as well as aspects that characterise the decision maker. A key consideration in this respect is that the perceived attractiveness of a route by a traveler can depend on multiple attributes (as shown in Section 5.2)

The analysis of those criteria is relevant because we hypothesize that the criteria considered important by the traveler for his route choice are related to the ones governing his choice set formation in his mind. For example, if a person is a time sensitive traveler, travel time also influences the formation of his route alternatives in his mind. In particular, only route alternatives with low travel time might be considered in the route choice process.

Aspects characterizing the decision maker are also relevant because within any population of individuals taste variation is likely to occur. Perception and relative weights of the route attributes that make up the generalized cost vary among travelers, known as taste heterogeneity. From a modeling point of view this can be interpreted as the existence of several sets of weights, each of which determines a particular generalized cost function. Taste variation within a population can then be modeled by dividing the population into a number of segments that have approximately the same attribute weights (tastes) in their generalized cost function, and finding least-cost routes for each population segment separately. Taste variation is essential for choice set generation and route choice models. This differentiation among travelers often can be correlated with external characteristics of travelers such as age, income, license holding, and car competition (i.e. less cars in the household than licenses), gender, employment status. Finally, a third source of preference variation is trip purpose: people on their way towards obligatory activities often have different priorities than those on their way to discretionary activities, so that travel purpose is a major determinant of preference weights. Examples of segmentation criteria that often turn out to be of relevance are trip purpose, and socio-economic variables such as income, age, household composition etc.

Finally, public transport fares are usually non-linear, so are the disutility functions of travel time or waiting time. A possible way to allow non-linear costs is adopting a path-based cost function in choice modelling. The difficulty with path-based approaches is the combinatorial explosion of the number of possible paths, especially as these have to be enumerated in advance. A good choice set generation module is needed to reduce the computational burden to acceptable proportions by generating a limited set of feasible paths that encompasses the subjective choice set.

5.6 Conclusion

This chapter provided insights into route choice factors and derived behavioural rules for route choice on the basis of which requirements for choice set generation approaches can be derived.

One of the basic hypotheses followed in this thesis is that choice set formation in the mind of travellers strongly relates to their preferences and experiences with respect to route choice. Route choice criteria of travellers are relevant for CSG approaches because we hypothesize that criteria applied for route choice are related to the ones governing choice set formation in the mind of people. It appeared that travel time, transfers and waiting times are outstanding factors influencing route choice that should be considered as factors playing a role in choice set formation.

Our developed choice set generation approach (to be discussed in the next chapters) therefore will be based on this hypothesis, meaning that route choice factors and behavioural parameters of travellers will determine the generation of route options for the choice sets via a utility or cost function, called route generation function. Since there exist significant differences in choice factors and preferences between user groups and between trip purposes, these route generation functions need to be specific for user groups and trip categories.

Given the complexity of route choice in multi-modal networks, and the special problem of overlap, only advanced choice models qualify for being applied in the supernetwork approach. Since overlap and taste variation, essential for choice set generation and route choice models, especially in a multi-modal context, can be handled by Mixed Logit models, they appear to be the most suitable models, although specification and estimation of ML models are not always so straightforward.
The PCL model is one of the better options in this respect. Another potentially suitable model is PS-logit. The PS-logit model seems to offer a consistent approach to the route choice problem, but remains an approximation because it uses only one single commonality factor for a route despite its varying degrees of overlap with other routes.

The use of choice sets is essential in determining the best parameter values of choice models (estimation purpose) and in using these choice models for calculating the expected usage of alternatives in a network (prediction purpose). It is important to note that estimation and prediction pose different requirements on the completeness and composition of choice sets. For model estimation, the misspecification of the choice set may produce distorted parameter estimates when choice sets vary significantly among decision makers, which is true for route choice (Williams, H.C.W.L. and Ortuzar, J.D., 1982). Equally, a correct specification of choice sets has been shown to be important Louviere, J.J. and Hensher, D.A. and Swait, J.D. (2003). On the other hand, for estimation purposes, choice sets need not be exhaustive, a well-chosen sample may suffice Ben-Akiva, M.E. and Lerman, S. (1985). An example of the impact of choice set composition in the case of estimating pedestrian route choice models is Van Der Waerden, P. and Borgers, A. and Timmermans, H.J.P. (2004). For prediction purposes matters are different. In that case, choice sets need be fairly complete in order to achieve sufficiently good estimates of the use of the routes.

The insights gained in this and previous chapters will be used in the establishment of choice set generation algorithms to be deal with in the next chapters. Chapter 6 will tackle this subject for uni-modal networks, while in Chapter 7 CSG models will be developed specifically designed for application in multi-modal networks.

Chapter 6

Route set generation in uni-modal networks

6.1 Introduction

Having defined in Chapter 2 the role of choice sets in travel demand modelling, this chapter will go into the details of choice set generation methods for route choice prediction. This will be based on the choice set concepts developed in Chapter 4. As a preparation to the establishment of a new choice set generation method, developed in particular for multi-modal networks, this chapter will present a discussion of the choice set generation (CSG) problem itself for uni-modal networks. The extension of choice set generation approaches for multi-modal networks is provided in Chapter 7. In this chapter we will address questions of what the specifications of an adequate route choice set are, in particular in the context of demand prediction purposes, and which requirements should be posed on such route choice sets to be adequate for that purpose. A related question we will deal with is which requirements an appropriate choice set generation approach should satisfy. In this chapter, we tackle these questions with an application to uni-modal networks in mind.

Based on the established adequacy requirements, the main purpose of this chapter is to discuss and evaluate a large number of route generation methods proposed in literature and indicate which of these methods are potentially suitable for generating satisfactory route choice sets. Therefore, apart from the notion of choice set generation (CSG) we introduce in this chapter the notion of route set generation (RSG) which simply is the generation of a set of routes not necessarily for the specific purpose of route choice modelling and prediction.

The main contributions of this chapter are, first of all, the specification of requirements for adequate route choice sets for prediction of route and link flows in uni-modal transport networks. Secondly, we specify a generic CSG procedure for the purpose of characterizing current route set generation procedures. Finally, using this generic procedure we give a comprehensive set of structured comparable descriptions of current route set generation methods and give a comparative evaluation of these for sake of deriving desired properties of an adequate CSG method for use in travel demand prediction.

The type of choice sets to be generated in most cases depends on the application purpose. As stated in Chapter 4, apart from purely scientific analysis reasons, choice sets might be generated for either estimation or prediction and forecasting purposes. The differences between these two application purposes and their consequences for composition of choice sets and generation approaches will be explained in Section 6.2. The rest of this chapter will focus on presenting and discussing choice set generation approaches to be applied for prediction purposes only.

An important topic dealt with in this chapter is the discussion about the main characteristics of adequate choice sets and of appropriate generation processes. In Section 6.3 we will deal with the notion of an adequate choice set, which kind of choice set is most suitable for prediction purposes, and will formulate requirements an adequate choice set should satisfy. Subsequently, we will specify requirements for an appropriate choice set generation process. Thirdly, we will define quality criteria to assess the quality of the resulting choice sets and the quality of the generation process (Section 6.4).

Based on a developed dedicated generic generation scheme in Section 6.5, a classification of the choice set generation methods known from literature will be given, after which each of these methods will be described in a structured comparable way as a basis for our evaluation of their adequacy for our purposes, on the basis of which recommendations are given for the best approach to be adopted.
We then will illustrate in Section 6.6 some of these CSG algorithms by applying them to simple uni-modal networks, namely continuous networks such as road networks and waterway networks. The generated choice sets will be analysed in order to establish if the resulting choice sets match the defined requirements for adequate choice sets. In a concluding section 6.7, the findings relevant for multi-modal applications will be summarized and discussed.

6.2 Purposes and importance of choice set generation

We recall from Chapter 2 that for prediction applications, a priori generation of choice sets for route and link flow calculation is strongly favoured for several reasons. A major advantage of a priori choice set generation for prediction purposes is that a priori given choice sets allow much more flexibility and realism in behavioural assumptions in the route choice models adopted. In that case, no restrictions exist on the type of choice model or utility function specification to be adopted. Applicability of advanced analytical approaches, easy consideration of route overlap, non-linear utility functions, and route-based attributes are only a few advantages of applying a priori generated choice sets. In addition, a priori enumeration in a network context not only offers a number of theoretical advantages, but also implementation and computational advantages in iterative network assignment approaches since repeated optimal route search no longer is necessary, which saves computing time. An example of a priori route set generation is given by Damberg, O. and Lundgren, J.T. and Patriksson, M. (1996).

Route choice sets may be generated mainly for the following three application purposes:

- Scientific *analysis* of travel options in networks where the planner or researcher is interested to know the availability of travel alternatives, their number, their characteristics, their variety, their composition etc.;
- Demand model *estimation* (e.g. estimating behavioural parameters of utility functions of choice models);
- *Prediction* of choice probabilities in a demand analysis for determining route flow and link flow levels in networks using route choice models with known parameters derived from estimation.

These different applications of choice sets appear to pose different requirements on size and composition of the choice sets.

Whereas choice sets need not necessarily be exhaustive for estimation purposes, prediction choice sets must include at least all attractive routes. For estimation purposes, even if not all relevant alternatives are included in the choice set and small well-sampled choice sets are considered in the estimation model, it may nevertheless provide satisfactory results (Ben-Akiva, M.E. and Lerman, S., 1985), (Train, 2003). On the other hand for prediction purposes, the choice sets should include quite all realistic and reasonable alternatives; otherwise computed route choice probabilities may produce wrong predictions.

In this thesis we will focus on the generation of choice sets for prediction of route and link flows, so we will devote our attention to the adequacy of choice sets for that purpose.

In the context of a prediction application, the analyst is interested to achieve proper predictions of route and link flows, especially for those routes and links that have special policy relevance. Such a prediction involves calculating the choice probabilities of all non-zero OD-trips, maybe separately for user groups or trip purposes, and then summing up the number of trips that will use each of the potentially feasible routes, and derived from this, through a network assignment, the use of links. This would require the specification of choice sets in which each route that may attract trips is included. In the case of predicting route flows the requirements on the quality of choice sets are very strict since in order to have correct route flows, predicted choice sets should include all relevant routes. Inclusion of some unattractive routes in the choice set may not necessarily distort the demand predictions, and may not have serious influence on computational efficiency.

If only link flow predictions are of concern to the analyst, less severe requirements on the quality of choice sets are due since link flows mostly are sums of very many different and often small route flows from different OD pairs. In this case, it is sufficient to have the most important routes in the choice sets, while it is not a problem to have some wrong routes in the choice sets, because there will be an error compensation. If a link flow is wrongly predicted having, for example, zero flow, it means that this link does not belong to any routes included in the choice sets. If that link is a crucial link for the network (e.g. a bridge), it implies that some important (relevant) routes including that link have incorrectly not been selected in the choice set generation process.

Therefore, for link flow prediction, generated choice sets need to include the most attractive routes but may miss some routes of less attractiveness. Given this, prediction choice sets should be sufficiently large, maybe even including some irrelevant unattractive routes, in order to benefit from the compensation mechanism of erroneous predictions of small route flows. Consequently, a prediction choice set should likely consist of all relevant routes with high probability of being chosen; inclusion of some unattractive routes may not necessarily introduce problems. An exception of this, however, is the presence of overlap among routes, since unattractive routes via their overlap with attractive routes may influence the choice probabilities of these routes (see for example Bliemer, M.J.C. and Versteegt, E. and Castenmiller, R. (2004) and).

6.3 Requirements for adequate choice sets for prediction purposes

In this section we will discuss the requirements for an adequate choice set and an appropriate choice set generation process. We will first look at properties of single routes by specifying requirements for reasonable routes to be generated. On that basis we then look at the composition of reasonable sets of routes from an individual traveller's perspective, after which we consider the adequacy of route sets for groups of similar travellers. On this basis, we will define criteria for adequate choice sets for prediction purposes, and define quality criteria for the resulting choice set and for the generation process.

6.3.1 Requirements for a *reasonable* route

Hoogendoorn-Lanser (2005) offers a systematic account of conditions that reasonable routes should satisfy from the traveller's perspective to become a member of an individual's choice set. These conditions can be grouped under the following headings:

1. logical conditions;
2. feasibility conditions;
3. behavioural conditions;
4. perceptual conditions.

Logical conditions concern the topological form of routes in space and/or time. The term 'logical' expresses that travellers don't undertake unnecessary or superfluous actions such as travelling in cycles or loops, or taking impossible actions (travelling backwards in time). In this context, a *loop* is defined as a part of a trip visiting the same node or link more than once. Illogical routes will not be selected in choice sets, apart from the fact that most route search algorithms cannot generate illogical routes.

Feasibility conditions regard the suitability of a route alternative in terms of time, space, vehicle availability, and physical (dis-)abilities. Feasibility in *time* can be accounted for by explicitly considering time constraints at origin and / or destination addresses. These types of constraints are especially relevant in a multi-modal context with time-tabled services other than in a uni-modal context (see Chapter 7). For example, in a multi-modal context maximum transfer time constraints can be used to establish which runs of consecutive transport services can be taken. Feasibility in *space* refers to availability of transport modes not only at origin and destination, but also at transfer points that is especially relevant in multi-modal trip making (see Chapter 7). *Physical disabilities* might restrict the use of private transport modes (for example, walking and cycling) and public transport services. To account for physical disabilities during the choice set generation, routes with certain types of transfers can be excluded from the choice set and walking speeds can be adapted accordingly.

Behavioural conditions refer to individual traveller preferences with respect to trip attributes, such as transport modes, waiting times, walking times, costs and number of transfers, etc. A typical example is the distance and time detour relative to the straight OD-connection. A basic premise behind these conditions is that travellers always try to use the subjectively best route: they optimise their behaviour according to a personal utility function. Travellers are not accepting excessive detours or multiple switches between link types during the trip. Moreover, route alternatives that largely *overlap* with others will not be identified as a distinct route by the traveller, reason why such routes might be excluded from the choice set. Another behavioural condition relates to the hierarchical set-up of routes being preferably followed by travellers, see below. Behavioural conditions specifically related to a multi-modal context will be discussed in more detail in Chapter 7.

Perceptual conditions relate to the knowledge on the part of the traveller of available alternatives and their attributes. Of course, these are difficult to specify, especially in forecasting cases.

For a more elaborate account of these conditions see Hoogendoorn-Lanser (2005).

Looking for implementations of these conditions in the literature, apart from Hoogendoorn-Lanser (2005) only few sources are devoted to the specification of reasonable routes for route choice analysis. A summary is given in Bovy, P.H.L. (1988) and in Bovy, P.H.L. and Stern, E. (1990).
The following reasonabless criteria from the literature are relevant for our purpose:

- efficiency (Dial, R.B., 1971);
- route detour (De La Barra, T. and Perez, B. and Anez, J. (1993) and many others);
- optimal in a transportation sense (Ben-Akiva, M.E. and Bergman, M.J. and Daly, A.J. and Ramaswamy, R., 1984);
- maximum size (Borgers, A. and Timmermans, H.J.P., 1984).

Dial, R.B. (1971), for example, defined a route as reasonable if it is "efficient", meaning pursuing to its destination without any backtrack. He operationalized his efficiency principle with the following criteria: a route from origin O to destination D is reasonable if every road segment $(i,\ j)$ in the route takes the traveller:

1. either farther away from the trip origin $F_O(i) < F_O(j)$;
2. or closer to the trip's destination $F_D(j) < F_D(i)$;
3. or, more strict, farther away from the origin and at the same time closer to the destination $F_O(i) < F_O(j)$ & $F_D(j) < F_D(i)$.

where F may represent any disutility measure such as time, distance, number of links etc. and i and j are the begin and end nodes respectively of the link or road segment $(i\ ,\ j)$. $F_O(x)$ may represent the shortest time or distance from origin O to node x while $F_D(x)$ may represent the shortest distance or time from node x to the destination D. These criteria should prevent from getting routes with loops.
The set of routes that satisfy these criteria can be generated by a single simple shortest path tree search from the origin or destination or from both, reason why these criteria have been implemented in many network assignment models. A recent application of Dial's criteria to generate choice sets in a dynamic assignment context is given by Lim, Y. and Heydecker, B.G. (2005).
Although these criteria seem reasonable a prima vista, they are not without problems. First, in the same network, each of these three criteria will give rise to different choice sets. Secondly, it appears that these criteria will overlook other attractive alternatives in the network. A further problem is that the resultant choice set is dependent on the chosen disutility measure for F (for details see Bovy, P.H.L. (1988)). Dial's criteria are easy from a computational point of view (analyst) but not necessarily suitable from a behavioural point of view (traveller).

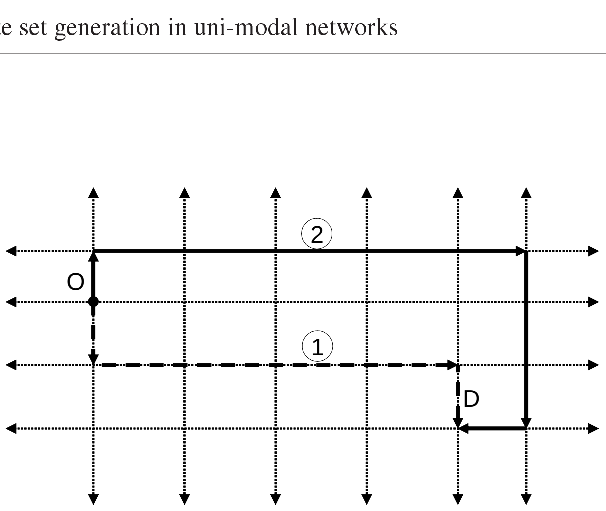


Figure 6.1: Examples of reasonable routes that may (not) satisfy Dial's criteria.

Figure 6.1 shows a grid network with two routes from origin O to destination D: route 1 (dashed line) and route 2 (solid line). Let us assume that all links in the grid network have the same distance therefore the total distance of route 1 is smaller than that of route 2, but some links in the grid network have different speeds in such a way that the total travel time of route 2 is lower than that of route 1.

If we assume that function F represents the distance, route 1 is an efficient route satisfying Dial's criteria, whereas route 2 is not. In this case the resultant choice set contains route 1. However, if we assume that F represents the shortest time, route 2 is reasonable and it does satisfy Dial's criteria, whereas route 1 does not. In this case the resultant choice set includes route 2.

The problems with Dial's criteria indicate that additional criteria are needed. An often used criterion is a maximum detour relative to the straight OD connection in distance or relative to the shortest path in time between O and D. In many car traffic applications a detour of 50% has been is implemented. A related criterion is the maximum route length in terms of number of constituting links, which requires restrictions on minimum and maximum link length (Borgers, A. and Timmermans, H.J.P., 1984). A well-known route generation approach is 'labelling', meaning that every route between O en D that is optimal in some travel disutility sense is eligible for the choice set Ben-Akiva, M.E. and Bergman, M.J. and Daly, A.J. and Ramaswamy, R. (1984). Possible optimality criteria are: shortest route, fastest route, cheapest route, safest route, maximum freeway route, etc. These routes can be easily generated by repeatedly applying a shortest path search.

Apart from the reasonabless properties of routes explicitly discussed in the literature, additional desired properties will be considered that are implicitly used by the route set generation methods developed in the recent past, to be discussed in Section 6.5.

Based on the previous arguments, let us discuss our definition of *reasonable*[1] route.

---

[1] Note that the term *reasonable* used in this thesis differs from the one defined by Dial, R.B. (1971)

Definition 9 *A single route is defined* ***reasonable*** *meaning that this route is eligible as member of the choice set if the following criteria are fulfilled:*

Acyclic criterion *A reasonable route does not contain loops.*

Detour criterion *A reasonable route does not exhibit a detour from the shortest possible connection in terms of one or more measures such as distance or time between origin and destination larger than a maximum threshold* α *(e.g. 100%).*

Hierarchic quality *A reasonable route is constituted by a systematic sequence of functional link levels in the network such as roads (collectors, arterials, and freeways), waterways, avoiding route parts going from higher to lower level links and back, such as for example, repeated entrance to and exit from the same motorway.*

All criteria listed in Definition 9 refer to a *single* route.

The *detour criterion* derives from the fact that travellers try to optimise their routes, which will not show a large detour from the shortest or quickest distance between the origin and destination. We may expect a reasonable route to deviate from the shortest or fastest one only to some maximum level.

In order to formalise the detour criterion let us introduce the following notation. Let $\Gamma_r = \{l_1, l_2, , l_z\}$ denote the sequence of links of route r, which starts from the origin node O and ends to the destination node D. Let $F(l)$ be the function that maps link l to its length, time or generalised cost. Let $d[i, j]$ be the shortest connection (in distance, time, or cost) between any nodes i and j. Let $\alpha \geq 1$.

A route satisfies the detour criterion if the following holds:

$$\sum_{l_z \in \Gamma_r} F(l_z) \leq \alpha \cdot d[O,D] \tag{6.1}$$

This detour criterion refers to the full route and not to parts of the route.
Figure 6.2 shows network examples with reasonable routes and non-reasonable routes from O to D. Again assuming that all links in the network have the same speed, two routes with the same number of links will have the same travel time. Route A is reasonable according to the acyclic and detour criterion, no loops are in the route. Route B cannot be defined reasonable according to the acyclic criterion, because it contains a cycle. Finally, route C is also not reasonable according to the detour criterion because the route contains a large detour relative to the shortest possible distance. In addition, please note that it does not satisfy the Dial-criteria.

The requirement of *hierarchical quality* is derived from the hierarchical structure inherent to transportation networks mostly consisting of different network parts having different functional levels suited for facilitating specific trip speeds and trip lengths.

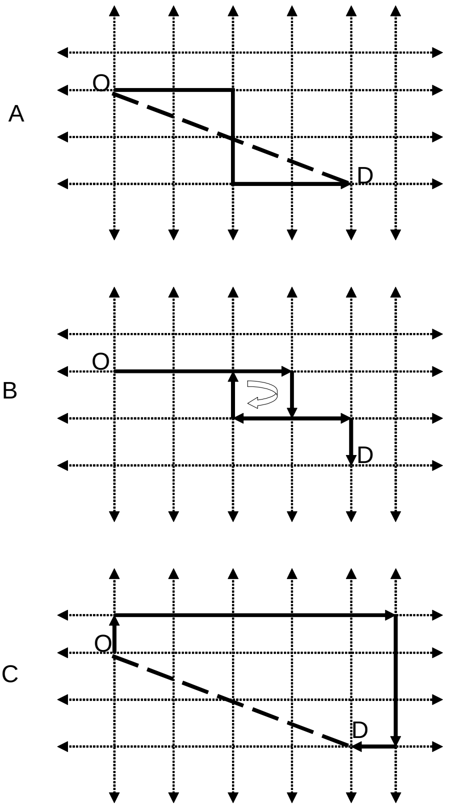


Figure 6.2: Examples of reasonable (dashed) and non-reasonable (solid) routes.

For example, in Figure 6.3 four network levels are distinguished, from local streets at the lowest network level being a spatially very detailed network, having high accessibility and low speeds via collectors and arterials to the motorway level showing decreasing levels of spatial detail but increasing levels of speed and capacity.

Therefore, a reasonable route should have the systematic sequences of functional roads (pyramidal set-up), whereas deviations from a pyramidal set-up (dashed lines) as shown in Figure 6-3 are considered non-reasonable.

The other example in Figure 6.3 (solid line) is a typical example of a reasonable route for specific interurban trip and for a specific location.
The hierarchical quality criterion is a difficult one to realize in a route set generation procedure since it is a typical route property that requires the assessment of the route in its entirety, which is not simply derivable from the properties of the constituent links.

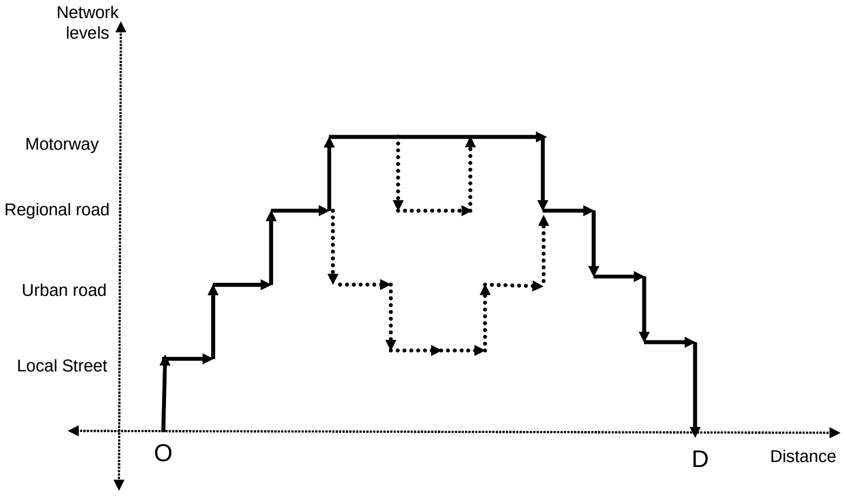


Figure 6.3: Example of a reasonable route and an unreasonable route (dashed line) in a hierarchical network having four network levels.

One should note that the presented criteria are not mutually exclusive, the various criteria to a certain extent overlap and may result in identical routes.

In the next Subsection 6.3.2, the defined requirements for single routes to be reasonable routes will be extended towards the definition of an adequate set of routes for a single individual traveller. Then also criteria related to choice set size and set composition come into play. The formulation of requirements for the individual level serve as an intermediate step towards the definition of the requirements of adequate choice sets for a group of similar travellers (Subsection 6.3.3) In Subsection 6.3.4 the requirements for choice sets for prediction purposes are summarised.

6.3.2 Requirements for adequate choice sets at individual level

Based on the reasonabless criteria for single routes, we now look at adequacy of generated complete choice sets from the perspective of a single traveller. Since we aim at specifying choice sets generated for prediction purposes and thus for a group of similar travellers, this subsection is an intermediate step.

As discussed in Chapter 4 (Section 4.3), from the viewpoint of a single traveller the following characteristics of choice sets are important:

- the spatial structure of the set in terms of the degree of mutual overlap or, oppositely, their spatial difference;

- the composition of the set in terms of comparability of the routes in the set or oppositely their functional variety;
- the size of the set in terms of number of alternatives.

From these arguments the following definition is derived.

Definition 10 *A choice set generated for an individual traveller is* ***adequate*** *if all routes of the choice set are reasonable according to Definition 9, and additionally satisfy the following criteria:*

Overlap criterion *Any two routes of the generated choice set should have a mutual overlap (in terms of number of links, distance, or time) less than Δ percent with respect to the shorter one of the two routes.*

Comparability criterion *Any two routes of the choice set should be comparable in travel disutility (time, distance, or cost) within a given threshold of θ_1 percent with respect to the shorter one of the two routes (irrespective of overlap).*

Detour-max criterion *The non-common parts of two partly overlapping routes in the choice set should have a maximum detour (in disutility terms) not larger than a given maximum percentage ω_{max} of the minimum of the two parts.*

Detour-min criterion *Any two partly overlapping routes in the choice set should have a minimum detour (in disutility terms) between the two routes not smaller than a given minimum percentage ω_{min} of the minimum total route length, (or time or cost).*

Choice set size *The choice set should contain a limited number of alternatives, say less than S (e.g. 10).*

Instead of single routes considered in Definition 9, now in Definition 10 the mutual relations between routes within the choice set are at stake. In Definition 10, we first consider the criteria referring to the comparison of full routes (*overlap* and *comparability criteria*); secondly the criteria regarding the comparison of part of routes (*detour-max* and *detour-min criteria*) and finally criteria relevant for the entire choice set (*choice set size*).

While the overlap criterion refers to the common route part of two routes, the detour-max and detour-min criteria specifically refer to the non-common route parts of two routes.

In order to specify these criteria let us introduce the following notation. Let $\Gamma_r = \{l_1, l_2, \cdots, l_a\}$ denote the sequence (ordered set) of links of route r and $\Gamma_p = \{l_1, l_2, \cdots, l_b\}$ denote the sequence (oredered set) of links of route p. Let $F(l)$ be the function that maps link l to its length, time or generalised cost, and let $F(l) = \{0, 1\}$ representing that link l does not or does belong to route Γ_r or Γ_p.

Overlap criterion

Regarding the *overlap criterion* (see also Subsection 5.4.2), two routes that do not visit precisely the same nodes in the same order are considered different (distinct and non-overlapping). However, from a transportation perspective a less strict overlap definition is required where routes with large overlap might both be considered reasonable alternatives whereas only small variations to a given route are not relevant for the analyst.

Hoogendoorn-Lanser (2005) introduces different measures to quantify overlap of alternatives in a choice set. Different figures, indicating the amount of routes or the use of links, can be derived from the so-called *assignment map* or *link-path incidence matrix*, which can be constructed for each trip. The assignment map shows which links are used in which alternatives and thus contains all information regarding route overlap. The route overlap can be defined in terms of:

- Number of common links;
- Time of common links;
- Distance of common links.

Given the above definitions and let $0 \leq \Delta \leq 1$, a route r satisfies the overlap criterion with respect to route p in terms of length or time of the number of common links, or in terms of the number of common links if the following holds:

$$\sum_{l_a \in \Gamma_r \bigcap \Gamma_p} F(l_a) \leq \Delta \cdot \min \left(\sum_{l_a \in \Gamma_r} F(l_a), \sum_{l_b \in \Gamma_p} F(l_b) \right) \tag{6.2}$$

where Δ is the overlap threshold.

If such criterion is not satisfied the longest (in terms of number of links, or distance, time or cost) route between route r and route p will be eliminated from the choice set.

To illustrate, Figure 6.4 depicts two routes r and p from O to D; the number of links of route p is 6 whereas the number of links of route r is 8. Given the cost associated with each link in Figure 6.4 the total cost of route r is 9 whereas the total cost of route p is 6. If we set Δ to 0.8 (80% of overlap) and we apply Formula 6.2 we observe that route r does not satisfy the overlap criterion. The cost of the links that are in common to the two routes is 5, therefore by applying Formula 6.2:

$5 \leq 0.8 \cdot \min\{6,9\} = 0.8 \cdot 6 = 4.8$

In this case routes r and p do not satisfy the overlap criterion because route r overlaps route p for more than 80% of the shortest route (p), therefore route r has to be eliminated.

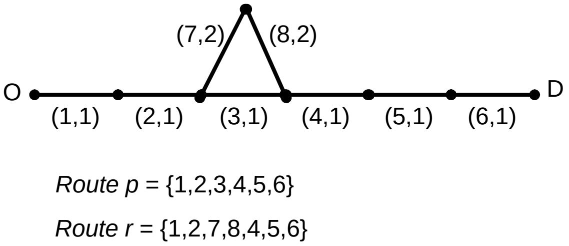


Figure 6.4: Example of overlapping routes r and p. ($(l,\ c)$ represents the link index and the link cost associated with each link).

Schnabel and Lohse (1997) in road network context, for example, propose that paths that overlap for more than 50% are not identified as separate paths. This can be verified by setting Δ to 50% (0.5) in Formula 6.2.

Comparability criterion

The *comparability requirement* is based on the assumption that an individual traveller only considers alternatives below a certain threshold of maximum travel time, travel cost or disutility. Such a maximum might e.g. be double the value of the minimum shortest route.

Given the above definitions and let $0 \leq \theta_1 \leq 1$, a route r satisfies the comparability criterion with respect to route p in terms of distance, time or generalized cost if the following holds:

$$\max\left(\sum_{l_a \in \Gamma_r} F(l_a),\ \sum_{l_b \in \Gamma_p} F(l_b)\right) \leq (1+\theta_1)\cdot \min\left(\sum_{l_a \in \Gamma_r} F(l_a),\ \sum_{l_b \in \Gamma_p} F(l_b)\right) \tag{6.3}$$

where θ is the comparability threshold.

Let assume that two routes r and p in the choice set have a total cost (for example, in time) of 100 and 150 (minutes) respectively. The two routes satisfy the comparability criterion if we set $\theta_1 = 1$, therefore the Formula 6.3 becomes: $150 \leq 2 \cdot 100 = 200$. The longest route is shorter than the double of the shortest, so the criterion is satisfied.

Detour-max and detour-min criteria

The *detour-max requirement* is somewhat similar to the detour criterion from Definition 9. The main difference is that in the detour-max criterion non-overlapping parts of two routes are considered to identify the level of maximum detour relative to the shortest connection, whereas in Definition 9 the entire route is considered and compared to the shortest distance between the origin and destination. Moreover, in Definition 9 a single route is considered whereas in Definition 10 always pairs of partly overlapping routes are taken into account. The detour-max criterion includes spatial variability among routes in the choice set. Routes should preferably differ but not too much.

The *detour-min requirement* is similar to the overlap criterion. The main difference is that in the detour-min criterion non-overlapping parts of two routes are compared to identify the level detour relative to the total length or cost of the other route, whereas in the overlap criterion the full routes are taken into account and compared to each other. The detour-min criterion is applied to eliminate routes that are largely overlapping with only minor deviations however, such as, for example, routes via motorways using off-ramps and on-ramps at the same interchange.

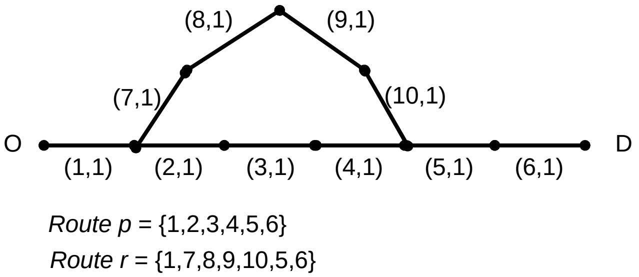


Figure 6.5: Route r satisfies the detour-max constraint relative to route p. ((l,c) **represents the link index and the link cost associated with each link**)

In order to specify the detour-max and detour-min criteria let us introduce the following notation. Given the above definitions, let a and b be the diverging and converging links of route r respectively, let $\hat{a}$ and $\hat{b}$ be the diverging and converging links of route p and let $\omega_{max} \geq 1$ and $\omega_{min} \geq 1$. A route satisfies the detour-max criterion if for any subsequence of links $\{l_a, l_{a+1}, \cdots, l_b\}$ of route r and for any subsequence of links $\{l_{\hat{a}}, l_{\hat{a}+1}, \cdots, l_{\hat{b}}\}$ of route p the following holds:

$$\max\left(\sum_{z=a}^{b} F(l_z), \sum_{\hat{z}=\hat{a}}^{\hat{b}} F(l_{\hat{z}})\right) \leq \omega_{max} \cdot \min\left(\sum_{z=a}^{b} F(l_z), \sum_{\hat{z}=\hat{a}}^{\hat{b}} F(l_{\hat{z}})\right) \tag{6.4}$$

where $l_z \in \Gamma_r$, and $l_{\hat{z}} \in \Gamma_p$ and ω_{max} is the maximum detour threshold.
Given the above definitions, a route satisfies the detour-min criterion if for any subsequence of links $\{l_a, l_{a+1}, \cdots, l_b\}$ of route r and for any subsequence of links $\{l_{\hat{a}}, l_{\hat{a}+1}, \cdots, l_{\hat{b}}\}$ of route p the following holds:

$$\min\left(\sum_{z=a}^{b} F(l_z), \sum_{\hat{z}=\hat{a}}^{\hat{b}} F(l_{\hat{z}})\right) \geq \omega_{min} \cdot \min\left(\sum_{z=a}^{b} F(l_z), \sum_{\hat{z}=\hat{a}}^{\hat{b}} F(l_{\hat{z}})\right) \tag{6.5}$$

where $l_z \in \Gamma_r$, and $l_{\hat{z}} \in \Gamma_p \omega_{min}$ is the minimum detour threshold.

Figure 6.5 illustrates a case where the detour-max constraint is satisfied. Route r satisfies the detour-max constraint by setting ω_{max} equal to 1.5 and applying Formula 6.4:

$\max\{3,4\} \leq 1.5 \cdot \min\{3,4\} = 4.5$

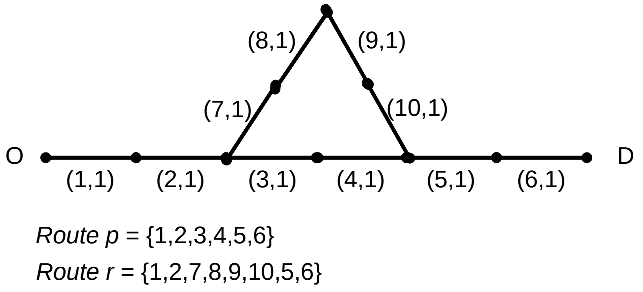


Figure 6.6: Route r does not satisfy the detour-max constraint relative to route p. ((l,c) represents the link index and the link cost associated with each link)

Figure 6.6 illustrates a case where the detour-max constraint is not satisfied. Route r does not satisfy the detour-max constraint by setting ω_{max} equal to 1.5 and applying Formula 6.4:

$\max\{2,4\} = 4 \leq 1.5 \cdot \min\{2,4\} = 1.5 \cdot 2 = 3$

All routes in the previous examples are reasonable, whereas route r depicted in Figure 6-6 is eliminated by pair-wise comparison.

Schnabel and Lohse (1997) discuss how this type of constraints can be used to eliminate behaviorally unrealistic paths. They recommend a value $\omega_{max} = 1.40$ for urban networks and $\omega_{max} = 1.25$ for motorway networks.

Threshold ω_{min} should be interpreted as the minimum required length that a detour should have before the path containing this detour is recognized as a distinct alternative. Figure 6-7 illustrates that the detour-min constraint is not satisfied, whereas

Figure 6.8 shows an example in which the detour-min constraint is satisfied. Route r depicted in Figure 6.7 does not satisfy the detour-min constraint by setting ω_{min} equal to 0.5 and applying Formula 6.5:

$\min\{1,4\} \geq 0.5 \cdot \min\{6,12\}$

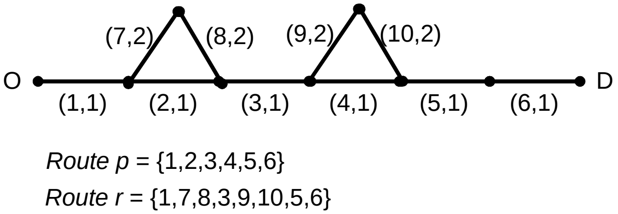


Figure 6.7: Route r does not satisfy the detour-min constraint relative to route p. ((l,c) represents the link index and the link cost associated with each link)

Route r depicted in Figure 6.7 is rejected because it is the longer one. Compared to route p the minimum of its detours is less than the threshold ω_{min} with respect to the minimum of the entire route cost. Note that route r depicted in Figure 6.7 does satisfy the overlap constraint with respect to route p meaning that route r is not relevant for the analyst because of its small variations to route p.

Figure 6.8 illustrates a case where the detour-min constraint is satisfied. Route r satisfies the detour-min constraint by setting ω_{min} equal to 0.5 and applying Formula 6.5:

$\min\{3,4\} = 3 \geq 0.5 \cdot \min\{6,8\} = 0.5 \cdot 6 = 3$

Choice set size

Finally, the requirement about the *choice set size* is derived from Chapter 4 showing that the size of the objective choice set is proportional to the network density, whereas the size of the subjective choice set depends on the limited ability of travellers in considering many alternatives. Thus, in modelling route choice sets for transport networks, large choice sets may be expected to be generated, whereas only small choice sets maybe expected to be considered by the travellers. In fact, since the number of routes that might be known or considered or used by the travellers is limited, a choice set at individual level is expected to have also a limited number of route alternatives, even very few alternatives as is observed in reality.

To a certain extent, the criteria developed and discussed above are overlapping in the sense that different criteria may lead to rejection of the same route.

Table 6.1 summarizes the properties of the criteria introduced in this section. Four types of comparison for route elimination are distinguished:

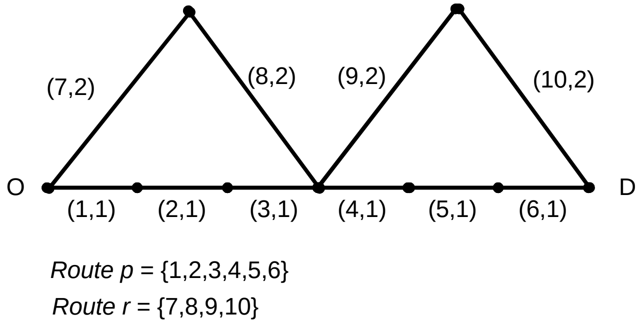


Figure 6.8: Route r satisfies the detour-min constraint relative to route p. ((l,c) **represents the link index and the link cost associated with each link**)

- Comparison at level of a single route on its own;
- Pair-wise comparison of full routes;
- Pair-wise comparison of parts of routes;
- Comparison at level of the full choice set.

Some criteria such as Acyclic, Detour and Hierarchic from Definition 9 are applied at the level of a single route. The overlap and comparability criteria are applicable to full route pairs, whereas the detour-max and detour-min criteria are applicable only to parts of route pairs. The choice set size is applied at the level of the full choice set.

Table 6.1: Criteria classified with respect to four types of route comparison.

Criterion	*Single route*	*Full route pairs*	*Parts of route pairs*	*Full choice set*
Acyclic	X			
Detour	X			
Hierarchic	X			
Overlap		X		
Comparability		X		
Detour-max			X	
Detour-min			X	
Choice set size				X

6.3.3 Requirements for an adequate choice set at group level

The requirements for an adequate choice set for a group of similar trips directly relate to our focus on the prediction application of choice sets. Therefore the arguments related to the prediction of flows (discussed in Section 6.2), will play an important role in this section. Moreover, when a group of individuals is considered, which are likely to be based on different preferences, aggregation of individual choice sets has to be taken into account. Therefore, all criteria introduced in the previous subsection need to be discussed again at aggregate level.

From the viewpoint of prediction of route choice, the following characteristics of choice sets are important:

Overlap and comparability criteria In prediction purpose applications, the analyst is much more interested in knowing the flows in the main arterials of the network than in collector and local links, which are mainly used in the network for modelling purpose only. Therefore, an adequate choice set should contain the routes consisting of these main arterials. Routes that overlap and are different only for local links (for example, on/off ramps) might be rejected because only variations of the main arterials are considered relevant for the prediction application. Routes belonging to the prediction choice set should be comparable, but at group level more variation in route types and route costs is allowed, because individuals belonging to the same group may have difference preferences.

Detour-max and detour-min criteria Detour-max and detour-min requirements introduced in Definition 10 derive also from the prediction purposes for which only variations of the main road arterials are relevant which might be included in the choice set, whereas overlapping or non-overlapping of local links are not relevant for the prediction choice set.

Choice set size As stated in Section 6.2 for link flow prediction, generated choice sets need to include all attractive routes but may miss routes of less attractiveness. Given this, predicted choice sets should be sufficiently large, possibly even including 'wrong' (non-used) routes. Consequently, we can state that a prediction choice set should likely consist of all reasonable routes with significant probability of being chosen.

Spatial variability criterion The generated choice set should contain as different route types (with respect to road type composition) as possible. This criterion, in addition to the detour-max and detour-min criteria, emphasises spatial variability among routes, referring to variability with respect to road type composition and other observable properties of roads. This criterion is especially relevant at the group level (see below).

Preferential variability criterion This criterion takes into account the variability in preferences among the groups of travellers. When several groups of travellers are taken into account, they are not homogenous within each others, but they consist of homogenous travellers within the same group (i.e. same trip purpose, transport mode available, etc.). The choice set should thus contain as many as possible routes representing the taste variation of each single group of travellers.

From these arguments the following definition of an adequate route choice set generated for prediction purposes for a group of travellers is derived.

Definition 11 *A route choice set generated for a group of travellers is defined* ***adequate*** *if all routes of this set are reasonable according to Definition 9, the set is adequate according to Definition 10, with the following remarks:*

Overlap criterion *See Definition 10*

Comparability criterion *Two routes belonging to the choice set should be comparable in travel disutility (time, distance, or cost) within a given threshold of* θ_2 *percent.*

Detour-max and detour-min criteria *See Definition 10*

Choice set size *The choice set should contain a limited number of alternatives but should include all attractive routes having a high probability of being chosen.*

and fulfils the following additional criteria:

Spatial variability *Routes of the choice set should be spatially different (dissimilar) with respect to the links used.*

Preferential variability *Routes of the choice set should represent the taste variation of each group of travellers.*

The *comparability criterion* has been adapted at the group level since routes in the joint choice set are less comparable than routes in the individual choice set. In order to take this into account different parameters s might be used. At the individual level the θ_1 parameter may be used in the formula whereas at group level the θ_2 parameter may be applied, with $\theta_1 < \theta_2$.

This implies that at individual level routes are more comparable in terms of distance, time or generalised cost, whereas at group level routes should be also comparable but in a less restrictive way, more variation of route types and route costs is allowed at group level, because individuals belonging to the same group may have difference preferences.

The *choice set size* requirement follows from Chapter 4. The size limit should allow the consolidation of different choice sets from various travellers at group level. In addition, the analyst is interested in selecting routes that vary as much as possible and in excluding routes that are similar and largely overlapping.

The *spatial variability* requirement follows from the observation that choice sets of different travellers in a group may strongly differ in composition due to the fact that the group members have each their own knowledge, preferences, and perceptions of the network. In addition, they travel between (slightly) different OD pairs with different OD locations, so that different types of routes may be chosen even if knowledge, preferences, and perceptions are equal. The main implication of this line of reasoning is that the prediction choice sets should preferably consist of route alternatives that are spatially different, and different with respect to cost and time or other attributes. Because of this a broad variety of options should be available in the generated choice set as far the transportation network allows. The spatial variability is already included in the overlap, detour-max and detour-min criteria. By setting the Δ, ω_{max}, ω_{min} parameters of the overlap, detour-max, and detour-min criteria respectively in a proper way, routes generated with those parameters values might contain enough spread in a spatial sense.

From surveys it is known that many travellers choose routes so as to optimise a particular subjective criterion. Whereas some travellers try to use the shortest route in distance, others prefer to use the shortest route in time, or having minimum traffic lights, minimum right turns or minimising other route attributes. The *preferential variability* criterion addresses this variability in preferences in the group by trying to achieve differences in label routes in the choice set.

6.3.4 Summary of desired properties of adequate choice sets

Table 6.2 summarizes the required choice set properties defined in the previous subsections classified by single route at individual level, and by choice set at individual and group level respectively. These criteria will be used in the sequel to evaluate the quality of the generated choice sets produced by the various choice set generation approaches. Since the main aim of this thesis is to generate adequate choice sets for prediction purposes, mainly the criteria for a group of travelers will be taken into account, additional to the criteria for reasonable single routes.

In Section 6.5 an extensive overview of choice set generation approaches will be presented. Some of these methods use explicitly some of the quality criteria previously defined for generating adequate choice sets. For each method described in the following subsections it will be indicated which of these quality criteria are fulfilled in generating the choice sets, how they are applied by each approach, and which are the results of the generation process.

In Chapter 7 we will deal specifically with the operationalization (chosen threshold and parameter values) of the developed criteria and introduce additional criteria for multi-modal networks. In that chapter as well, the developed dedicated CSG for multi-modal networks will be presented and discussed, and its performance quality evaluated.

Table 6.2: Requirements for a reasonable route and adequate choice set.

	Individual (OD pair)	Group level (OD zone level)
Single route	Acyclic criterion Detour criterion Hierarchic quality	
Choice set	Overlap criterion Comparability Detour-max criterion Detour-min criterion Choice set size	Overlap criterion Comparability Detour-max criterion Detour-min criterion Choice set size Spatial variability Preferential variability

6.4 Requirements for an appropriate choice set generation process for prediction purposes

The generation process, first of all, should ideally generate what we define as adequate choice sets for prediction purposes, a choice set that matches partly (as much as possible) or totally the quality criteria we defined in the previous sections.

Secondly, the generation approach should satisfy as much as possible the following criteria:

1. the approach should be suitable for both single OD pair and multiple OD pairs;
2. the approach should be generic;
3. the approach should be flexible;
4. the approach should satisfy the parsimony requirement.

The first important requirement for generating choice sets for prediction purposes is to have a generation process suitable both for one-to-one OD pair and one-to-many OD pairs, which is especially relevant at group level. Since we are especially interested in methods suitable for single and multiple OD pairs, in the following section we will use this requirement in order to classify the generation methods. Unfortunately some generation methods are suitable only for single OD pairs, as we will see in the next section.

The generation process should also be *generic* in the sense that it should be applicable to different kinds of uni-modal networks (such as roadway or waterway networks) at different spatial scales.

It should also be *flexible*, meaning that it may easily be adapted to be applied for different policy questions, under different network conditions, like peak/off-peak period or for different groups of travellers, such as commuters, students or elderly people; or the approach could be applied with different cost functions or search criteria. The policy analysis context may set requirements on the composition of the choice set, for example in the sense that particular alternatives that are relevant for the policy impact assessment always should be part of the choice set. In such cases (e.g. tolling or P&R policy studies) the policy analyst wishes to know the impact of such policies on travel choice behaviour and demand. Network elements (links and nodes) that are subject of such policies then might be required to be included in travel options. In such cases the policy analyst is assumed to prescribe which alternatives should anyhow be included in the generated choice set. The generation process should be flexible in the sense that it might allow analysts to set the model in order to generate choice set including alternatives that are required from a policy analysis point of view and alternatives that are relevant for policy makers. A generation approach might be not flexible at all, or might allow the analyst to set some given parameters in order to generate some specific choice sets, or it also might force to use some specific links or nodes to be included in the final solution.

Finally, the approach should not be too complex but at the same time include all relevant alternatives; we refer to this as the *parsimony* requirement.

6.5 Overview of exact and heuristic RSG approaches

6.5.1 Principles of a generic RSG approach

In the following, we make a distinction between choice set generation (CSG) and route set generation (RSG). Choice set generation approaches are aimed at generating route sets that fulfil the requirements of an adequate choice set to be used in prediction, whereas route set approaches generate sets of routes for some purposes, not necessarily fulfilling the adequacy requirements. Since however RSG's may have attractive properties for our purpose, and for gaining insight into the complexities of CSG, we analyse and discuss available RSG approaches.

Apart from a few exact approaches, most approaches for generating route choice sets presented in literature are heuristic since it is very complex to specify clear objective functions for best prediction route choice sets. Nearly all procedures assume that the relevant route properties can sufficiently be described by the properties of the constituent links; no attempt is known of adopting route-based characteristics in generating or selecting routes. Route set generation approaches are based on the application of some basic components that we will discuss in the following.

First of all, almost all route set generation approaches to be described in this section are based on shortest path search, except the one based on enumeration approach with Branch & Bound tecnique. Shortest path algorithms are very efficient to determine routes between two given points in networks of some size. The minimum criterion is mostly related to the traveller's choice criteria, such as minimum time and distance.

Two types of shortest path search algorithms are reported in the literature: the approaches that consider a single criterion and the ones that take into account multiple objectives in the search function. The solution of the single-objective shortest path problem determines the route, which optimises only one travel attribute, such as time or distance. The solution of the multiple-objective shortest path problem, on the other hand, obtains the route (set of routes), which optimises simultaneously more than one travel attribute, for example time and distance. The idea of the multi-objective (or multi-labelling) algorithm is that the search function equation is extended from a scalar function to a vector-valued function such that all objectives under consideration are included.

Abdelghany, K.F.S. (2001) presents an example of application of multi-objectives shortest path search. It is well known that the multi-objectives shortest path is $\mathbb{NP}$-hard problem and it is difficult to solve. Abdelghany, K.F.S. (2001) found that the multi-objectives approach may outperform a single-objective approach in terms of quality of generated routes, but it needs much more computation time than a simple single-objective approach. Moreover, as we will show in the coming sections single-objective approaches may performs well not only in computation time but also in terms of quality of generated routes, therefore, for those reasons in this thesis we will focus only on single-objective shortest path approaches.

6.5.2 Generic scheme of a RSG approach

The basic steps in most RSG algorithms are as follows:

- search a best route according to certain conditions;
- evaluate the route according to a set of route criteria;
- select or reject the generated route;
- evaluate the resulting route set according to a set of criteria.

Based on these steps and ingredients (search criteria, selection constraints, etc.) we designed a generic route set generation approach that encompasses nicely all currently available route set generation approaches.

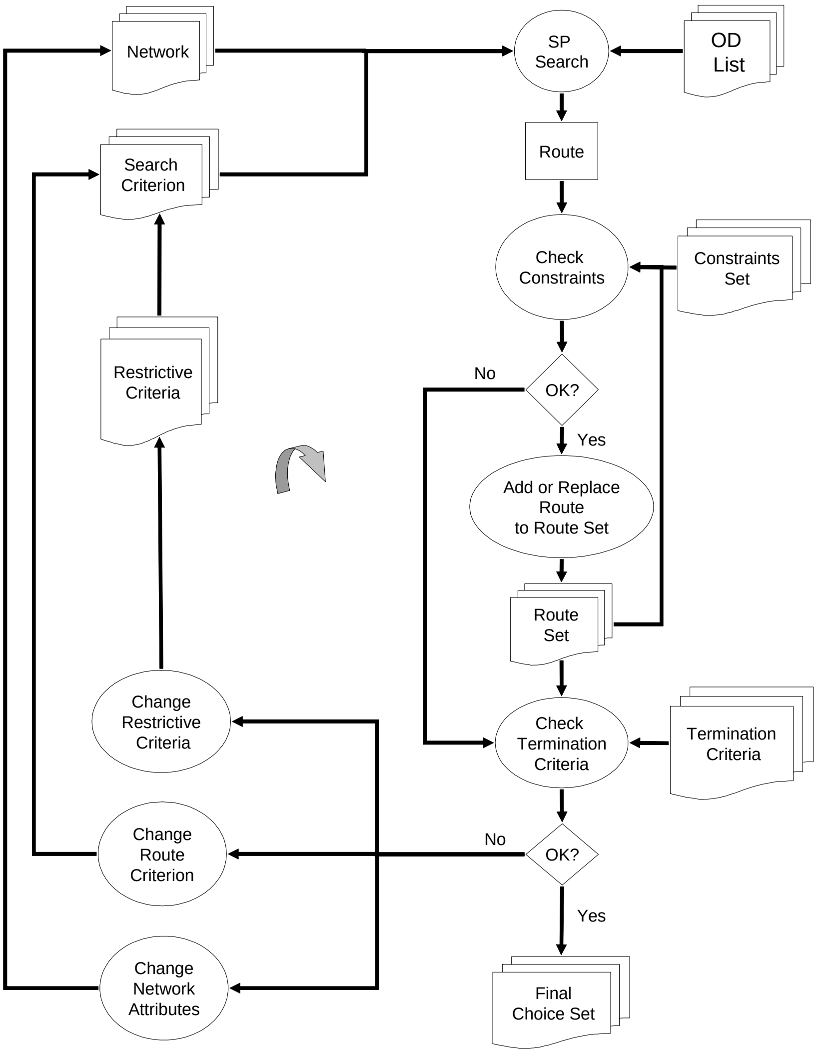


Figure 6.9: Classification of exact and heuristic approaches for RSG

Figure 6.9 shows a generic procedure for generating route sets as a basis for choice sets. As shown, the generic procedure requires six input boxes, namely: the network, in which the paths are searched, a ”restrictive criteria” that may force a route to pass by a specific node or link, the OD list for which the paths have to be generated, the search criteria that are applied in determining the shortest paths, and the constraints set and the termination criteria boxes consist of the constraints used by the approach to eliminate routes and the criteria applied to terminate the process respectively. The OD list might consist of single OD pairs (one-to-one search) or a list of origins from which to establish trees to all destinations (one-to-many search) at once, the multiple OD pairs case. Two output boxes are distinguished in Figure 6.9, namely the final route set and the intermediate route set, that is the set in which at each iteration a new route may be included or replaced. The final route set consists of the characteristics of the choice set (total number of routes generated, routes per OD pair, etc.) and the characteristics of the route set generation process (total number of iterations, etc.)

All methods for generating route sets apply a specific *Search criterion* or multiple search criteria on the basis of which the shortest paths are determined. The search criteria might be related to the notion of the generalized cost of a route or to some other attribute or combination of attributes. Assuming that an individual traveller takes a single or several aspects of a route into account simultaneously and selects the best route with respect to the combination of those aspects, leads us to the notion of generalized cost of a route. The (stochastic) generalized cost $\bar{C}_{qk}$ of a route alternative k perceived by the individual traveller q is usually modelled as a weighted (linear) combination of link attributes X_{ay}:

$$\bar{C}_{qk} = \sum_{a \in k} \sum_{y \in Y} (\beta_{qy} + \hat{\varepsilon}_{qy}) \cdot (X_{ay} + \tilde{\varepsilon}_{ay}) \tag{6.6}$$

where index a refers to links of route k, index y represents the index of the Y attributes considered by individual q, while β_{qy} are the weights of the link attributes y for individual q. X_{ay} is assumed to be the exact and true value of the attributes y for link a of path k considered by the individual q.
For classification purposes, we specify cost function 6.6 with two separate error terms, one ($\hat{\varepsilon}$) attached to parameters β, while the other ($\tilde{\varepsilon}$) is attached to attributes X.

Error terms $\hat{\varepsilon}$ express the variations in parameter values β in the population of travellers (taste variation) and the uncertainty about β on the part of the analyst, while error terms $\tilde{\varepsilon}$ express perception errors of X in the population of travellers and uncertainty about X on the part of the analyst.

Note that a path cost function based exclusively on travel time is a special case of a generalised cost where the weights β of all attributes X except travel time are zero. Therefore, the notion of shortest or least cost routes through a network is preferably cast in terms of generalised cost. The least-cost route with respect to the generalised cost function will be the route that best matches a particular set of attribute weights.

Search criterion function 6.6 shows that by selecting attributes X, parameter values β, error term specifications for $\hat{\varepsilon}$ and $\tilde{\varepsilon}$, and the search function specification, many different search function types may be specified.

Next component that a route set generation algorithm might adopt is a *Constraints Set* (see Figure 6.9). These constraints consist of two subsets to be applied successively. The first group, so-called single route constraints, are to be adopted to check the validity of the optimal route found through the shortest path (SP) search to the single route criteria, and if valid, the second group of so-called relational criteria (or pair-wise route comparison) is adopted to compare the generated route with all previously found routes. If the single route constraints are not satisfied the new generated route is rejected, if satisfied, then the second group of constraints related to the pair-wise comparisons are applied. If these constraints are not satisfied because, for example, one route in the choice set is overlapping with the new generated route, the two route costs are compared and the route with the worst cost is rejected (or replaced) and the route with the best cost is added to (or kept in) the choice set. If the constraints related to the pair-wise comparison are all satisfied the new route can be added to the choice set. These constraints are non-compensatory: if one of the constraints is violated the route under consideration is rejected. Please note that the only pair-wise comparison constraint that could be applied during the generation process is the 100% overlap, due to the random process of route generation it is better to apply the pair-wise comparison only after having generating all routes, in the so-called filtering process. Finally, termination criteria are applied (see Figure 6.9).

Each method might employ simple or complex *Termination Criteria*, for example, required (exact, minimum or maximum) size of the generated route set, or quality of the generated route set. The procedure might also terminate when the generated set is exhaustive (no more new routes can be found). If the termination criteria are satisfied the final route set is ready, if not, three additional steps in the generation procedure might be applied, separately or simultaneously depending on the kind of the approach.

One such additional component (Change Network Attributes action in Figure 6.9) changes the network (structure or attributes) by applying some rules, such as eliminating some links from the network, or changing the value of some link attributes X, for example the travel time of some specific links or randomising the error terms. Another component (Change Routes Criteria action in Figure 6.9) changes the route criterion function C.

Last additional component (Change Restrictive Criteria action in Figure 6.9) changes the criteria that force a route to pass by certain nodes or links. At the beginning, the set of restrictive criteria might be empty, during the generation process some links are forced to be included in the route to be searched; therefore they are included in the ”restrictive criteria”. In the case that some links (close to the origin or close to the destination) are forced to be included the new route to be searched may begin from a different origin or may arrive to a different destination.

6.5.3 Route set generation methods classified

All known route set generation methods will be classified using the generic scheme shown in Figure 6.9. In this scheme, apart from the shortest path search, five basic components are distinguished as dimensions of our classification, namely Change Network Attributes, Change Routes Criteria, Change Restrictive Criteria, Check Constraints, and Check Termination Criteria. The generation methods to be discussed below differ from one another depending on which combination of these basic components is adopted.

In the following we will describe a sample of the most relevant methods to generate route sets in transport networks (each method may be an exact or a heuristic approach and may be applied to single or to both single and multiple OD pairs):

- k-shortest paths (KSP): exact method, only applicable to single OD pairs;
- Constrained k-shortest paths (CKSP): exact method, applicable to single OD pairs;
- Constrained enumeration: exact method, applicable to single OD pairs;
- Link elimination method: heuristic method, only applicable to single OD pairs;
- Link penalty method: heuristic method, only applicable only to single OD pairs;
- k-dissimilar paths: heuristic method, only applicable to single OD pairs;
- Gateway method: heuristic method, only applicable to single OD pairs;
- Least-cost paths: heuristic method, only applicable to single OD pairs;
- Labelling method: heuristic method, applicable also to multiple OD pairs;
- Monte Carlo method: heuristic method, applicable also to multiple OD pairs;
- Accelerated Monte Carlo method: heuristic method, applicable to multiple OD pairs;
- Monte Carlo Labelling combination method: heuristic method, applicable also to multiple OD pairs;
- Accelerated Monte Carlo Labelling combination method: heuristic method, applicable also to multiple OD pairs.

Table 6.3 shows for each indicated route set generation method which procedural components (among Change Network Attributes, Change Routes Criteria, Change Restrictive Criteria, Check Constraints) are applied.

Moreover, the table shows for each method the type of simple or complex Termination Criteria, and the type of single or multiple OD pairs that can be taken into account. A termination criterion is defined simple if the process stops when, for example, a specific number of paths is generated; it is defined complex when more conditions are checked, for example, the number of iterations, the number of new routes found in the last iterations, etc.

In the next section these route set generation methods are described using the generic scheme depicted in Figure 6.9 and the procedural components reported in Table 6.3. Table 6.4 extends on Table 6.3 by showing the type of approach and some sources for each method.

Table 6.3: Basic procedure components applied in each RSG method.

RSG approaches	Change Network Attributes	Change Routes Criteria	Change Restric. Criteria	Check Constr.	Termin. Criteria	OD pair
k-shortest paths	X		X		Simple	Single
Constrained *k*-shortest paths	X		X	X	Simple	Single
Constrained enumeration				X	Simple	Single
Link elimination	X				Simple	Single
Link penalty	X				Simple	Single
k-dissimilar paths	X			X	Simple	Single
Gateway method			X		Complex	Single
Essentially least-cost paths				X	Complex	Single
Labelling		X			Simple	Multiple
Monte Carlo (MC)	X				Simple	Multiple
Accelerated MC	X				Complex	Multiple
MC Labelling combination	X	X			Complex	Multiple
AMC Labelling	X	X			Complex	Multiple

As shown in Table 6.3 (last column) some methods can be applied only to single OD pairs, whereas some might be also applied to multiple OD pairs. We consider applicability to multiple OD pairs an advantage, mainly because of saving computational time, but also because to generate predicted choice set at group level is more relevant to have approach suitable for multiple OD pairs.

The aim of the following subsections is to describe in detail a fairly comprehensive set of route set generation methods, for each of which a definition and the main principle of the method are presented. Moreover, we compare the features of these methods with our requirements of an adequate choice set generation presented in Section 6.3 and Section 6.4. For each generation method a table shows which choice set criterion (for reasonable route and for adequate choice set) the generated choice set satisfies and which choice set generation approach requirement is also satisfied.

In this context, a criterion is considered *satisfied* if and only if the criterion is *explicitly* included and applied in the generation method. On the contrary, a criterion is considered *not satisfied* if it can *never* be included and applied in the generation process, otherwise fulfilment of the criterion remains an pre-question.

Based on this evaluation, for each method a list of advantages and disadvantages and a final comment is provided (referring to the quality criteria of an adequate choice set generation approach for prediction purposes). If possible, we will give some attention to the empirical validity of each approach, in the sense whether a quality check of the proposed approach with observations has been performed, and with what result.

Since we consider that generation methods suitable both for single and multiple OD pairs are more important and especially relevant for generating choice set at group level and for prediction purposes, in the following section we will divide the route set generation descriptions in two subsections, one dedicated to the single OD pair approaches, and the other to the approaches suitable for both single and multiple OD pairs.

Table 6.4: Exact and heuristic choice set generation approaches.

RSG approaches	Approach	References (sample)
k-shortest paths	Exact	Lawler, E.L. (1976); Akgün et al. (2000); Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001); Cascetta et al. (2002); Ramming, M.S. (2002);
Constrained *k*-shortest paths	Exact	Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001); Van Der Zijpp, N.J. and Fiorenzo-Catalano, S. (2002); Van Der Zijpp, N.J. and Fiorenzo-Catalano, S. (2005);
Constrained enumeration	Exact	Prato, C.G. and Bekhor, S. (2006);
Link elimination	Heuristic	Martins, E.Q.V. (1984); Park, D. and Rilett, L.R. (1997); Ramming, M.S. (2002);
Link penalty	Heuristic	Johnson et al. (1992); De La Barra, T. and Perez, B. and Anez, J. (1993); Park, D. and Rilett, L.R. (1997); Scott et al. (1997); Akgün et al. (2000); Ramming, M.S. (2002);
k-dissimilar paths	Heuristic	Akgün et al. (2000);
Gateway method	Heuristic	Lombard, K. and Church, R.L. (1993); Akgün et al. (2000);
Essentially least-cost paths	Heuristic	Hunt, D.T. and Kornhauser, A.L. (1997);
Labelling	Heuristic	BenAkiva et al. (1984); Dial, R.B. (2000);
Monte Carlo (MC)	Heuristic	Sheffi, Y. and Powell, W.B. (1982); Ramming, M.S. (2002);
Accelerated MC	Heuristic	Bliemer et al. (2004);
MC Labelling combination	Heuristic	Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001);
AMC Labelling	Heuristic	Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001);

6.5.4 RSG method descriptions of the single OD-pair group

Method: **k-shortest paths (KSP)**
Original source: Lawler (1976).
Application sources: Cascetta, E. and Russo, F. and Viola, F.A. and Vitetta, A. (2002), Ramming, M.S. (2002), Akgün et al. (2000), Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001).
Approach type: exact.
Definition: Generated set consists of the k-shortest routes according to a given criterion.
Principle: Repetitive applications of shortest path (SP) algorithm to successively changed network. As shown in Table 6.3, in this method the Change Network Attributes procedure is applied, in which the network is adapted by excluding some links. The Change Restrictive Criteria component is also applied by forcing the shortest path to pass by some links, such links are including in the restrictive criteria. The Change Routes Criteria is not changed and no constraints are used, therefore the Check Constraints is not applicable. Finally, the termination criterion is based on the given maximum number k of paths to be generated: the choice set should be exhaustive with a maximum size of k.
Advantage: This method finds in an exact way the k shortest paths. It matches only few requirements for generating *adequate* choice sets with *reasonable* routes. The requirements matched are the acyclic criterion, routes do not contain any loops, and the comparability criterion since the k shortest paths are variation of the shortest paths are all comparable each other in terms of time or distance.
Disadvantage: All other requirements summarised in Table 6.5 are not satisfied. For example, the size of the choice set depends on the number k, which in some cases might be much larger than the desired size. Moreover, the overlap and spatial variability criteria are not satisfied: generated paths tend to be very similar in spatial sense and also largely overlapping (Fiorenzo-Catalano, S. and Van Der Zijpp, N.J., 2001). This method is not flexible in the sense that it is difficult to force the method to generate particular alternatives.
Empirical test: Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001) present empirical tests with simulation networks. Ramming, M.S. (2002) and Cascetta, E. and Russo, F. and Viola, F.A. and Vitetta, A. (2002) carried out empirical tests with real networks and observations. Ramming, M.S. (2002) states that routes generated with KSP result in a good coverage (about 80%) with the observed routes, however it does take a large amount of computation time (about 3 days of computation).
Comment: This method is applicable only to single OD pairs. Because of this and other reasons, finding the k shortest paths tends to be computationally expensive on a network basis, also because this method is complex, especially due to the Change Network Attributes and the Change Restrictive Criteria procedures, which take computational time (parsimony criterion not satisfied). We conclude that the KSP method is not useful for generating choice sets.

Table 6.5: Satisfied quality criteria for the KSP approach.

	Quality Criteria	Satisfied	**Not** Satisfied
Reasonable route	*Acyclic criterion*	X	
	Detour criterion		X
	Hierarchic quality		X
Adequate Choice set	*Overlap criterion*		X
	Comparability	X	
	Detour-max criterion		X
	Detour-min criterion		X
	Choice set size		X
	Spatial variability		X
	Preferential variability		X
Appropriate CS generation approach	*Generic*	X	
	Flexible		X
	Parsimony		X

Method: **Constrained k-shortest paths (CKSP)**
Original sources: Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001), Van Der Zijpp, N.J. and Fiorenzo-Catalano, S. (2002).
Application source: Van Der Zijpp, N.J. and Fiorenzo-Catalano, S. (2005).
Approach type: exact.
Definition: Route set contains the k-shortest routes each of which satisfies predefined constraints in an exact way.
Principle: Repetitive applications of shortest path (SP) algorithm to successively changed network. As shown in Table 6.3, in this method the Change Network Attributes component is applied adapting the network by excluding some links. The Change Restrictive Criteria component is also applied by forcing the shortest path to pass by some links, such links are including in the restrictive criteria. Whereas the Change Routes Criteria box is not applied, the constraints set is defined (the so-called overlap, detour-max, and detour-min constraints) in such a way that the generated paths are more spatially different; therefore the Check Constraints Set is applicable. Finally, the termination criterion is based on the given maximum number k of paths to be generated: the choice set should be exhaustive with a maximum size of k.
Advantage: This method finds in an exact way the k shortest paths that match some of the requirements for generating *adequate* choice sets with *reasonable* routes. The requirements matched are the acyclic criterion routes do not contain any loops, and the comparability criterion since the k shortest paths are variations of the shortest paths being all comparable to each other in terms of time or distance. The overlap and spatial variability criteria are also satisfied since the application of constraints guarantees the generation of more spatially different paths compared to the paths generated by the KSP method.

Table 6.6: Satisfied quality criteria for the CKSP approach.

	Quality Criteria	Satisfied	**Not** Satisfied
Reasonable route	*Acyclic criterion*	X	
	Detour criterion	X	
	Hierarchic quality		X
Adequate Choice set	*Overlap criterion*	X	
	Comparability	X	
	Detour-max criterion	X	
	Detour-min criterion	X	
	Choice set size		X
	Spatial variability	X	
	Preferential variability		X
Appropriate CS generation approach	*Generic*	X	
	Flexible		X
	Parsimony		X

Disadvantage: Several other requirements summarised in Table 6.7 are not satisfied. For example, the size of the choice set depends on the number *k*, which in some cases might be much larger than the desired size. This method is not flexible in the sense that it is difficult to force the method to generate particular alternatives.

Empirical test: Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001), Van Der Zijpp, N.J. and Fiorenzo-Catalano, S. (2002), Van Der Zijpp, N.J. and Fiorenzo-Catalano, S. (2005) carried out empirical tests with simulation networks showing that this method performs better than the simple KSP, however is not completely useful for generating choice sets.

Comment: This method is applicable only to single OD pairs. The check of all constraints during the generation process may slow down the performance of the method (parsimony criterion not satisfied). Application of this method has shown that results are depending on the parameters settings of the method.

Method: **Constrained enumeration**

Original source: Prato, C.G. and Bekhor, S. (2006).

Application source: Prato, C.G. and Bekhor, S. (2006).

Approach type: exact.

Definition: Finds route set based on a Branch & Bound (B&B) technique, which enumerates paths by generating a tree of routes connecting an origin node O to a destination node D.

Table 6.7: Satisfied quality criteria for the Constrained enumeration approach.

	Quality Criteria	Satisfied	**Not** Satisfied
Reasonable route	*Acyclic criterion*	X	
	Detour criterion	X	
	Hierarchic quality		X
Adequate Choice set	*Overlap criterion*	X	
	Comparability	X	
	Detour-max criterion	X	
	Detour-min criterion	X	
	Choice set size		X
	Spatial variability	X	
	Preferential variability		X
Appropriate CS generation approach	*Generic*		X
	Flexible	X	
	Parsimony		X

Principle: In a pre-processing phase of the algorithm all elements of the network, such as links and path segments, are processed. After determining those path segments, given the origin node O, all possible consecutive links of each node are considered. A new node of the next link is inserted into the tree if, and only if some constraints are satisfied. In this phase the Check Constraints Set is applied to eliminate all links of non-adequate alternatives. Finally, the termination criterion is based on processing all nodes into the tree. The methods ends if no more nodes into the tree might be explored.

Advantage: This method finds in an exact way the route set that matches some of the requirements for generating *adequate* choice sets with *reasonable* routes. The requirements matched are the acyclic criterion routes do not contain any loops, and the comparability criterion. The overlap, detour and spatial variability criteria are also satisfied since the application of *temporal*, *loop* and *similarity* constraints proposed by the authors guarantees the generation of more spatially different paths.

Disadvantage: Several other requirements summarised in Table 6.7 are not satisfied. For example, the size of the choice set based on the size of the B&B tree in some cases might be much larger than the desired size. This method is not generic in the sense that the specification of the constrains might be determined by the application under consideration. Moreover, in some cases this method may delete some attractive alternatives; therefore it requires being very careful on defining and applying such constraints, since some relevant alternatives might be cut off.

Empirical test: Prato, C.G. and Bekhor, S. (2006) carried out empirical tests with real urban (car) networks and observations showing that the algorithm generates realistic and heterogenous routes. This method performs better (in reproducing the observed routes) than the simple labelling or link elimination approaches.

Table 6.8: Satisfied quality criteria for the Link elimination approach.

	Quality Criteria	Satisfied	**Not** Satisfied
Reasonable route	*Acyclic criterion*	X	
	Detour criterion		X
	Hierarchic quality		X
Adequate Choice set	*Overlap criterion*		X
	Comparability	X	
	Detour-max criterion		X
	Detour-min criterion		X
	Choice set size		X
	Spatial variability	X	
	Preferential variability		X
Appropriate CS generation approach	*Generic*	X	
	Flexible		X
	Parsimony	X	

Comment: This method is applicable only to single OD pairs. It is computationally inefficient (parsimony criterion not satisfied) because of the possible complex size of the B&B tree. Moreover, the check of all constraints during the generation process may slow down the performance of the method (parsimony criterion not satisfied). For these reasons and because of the type and size of transport networks, exhaustive methods that try to determine and enumerate all available paths within constraints are not considered as realistic candidates.

Method: **Link elimination method**
Original sources: Bellman and Kabala (1968), Martins, E.Q.V. (1984).
Application sources: Azevedo, J.A. and Costa, M.E.O.S. and Maderia, J.J.E.R.S. and Martins, E.Q.V. (1993), Park, D. and Rilett, L.R. (1997), Ramming, M.S. (2002).
Approach type: heuristic.
Definition: Route set contains routes that are variations of the shortest path by deleting some or all links of the shortest path.
Principle: Repetitive applications of shortest path (SP) algorithm to successively changed network. As shown in Table 6.3, in this method the Change Network Attributes procedure is applied, in which the network is adapted by eliminating according to some rules some or even all links of the shortest paths. The search criterion is not changed and no constraints are used, therefore the Check Constraints is not applicable. Finally, the exhaustive termination criterion is applied; the method ends if no more paths are found because of link elimination.

Advantage: This method matches only few requirements for generating *adequate* choice sets with *reasonable* routes. The requirements satisfied are the acyclic criterion; routes do not contain any loops, and the comparability criterion since the k-shortest paths are variations of the shortest paths being all comparable to each other in terms of time or distance. The variability criterion is also satisfied since the elimination of links from the shortest path guarantees the generation of spatially different paths.
Disadvantage: Many other requirements summarised in Table 6.8 are not satisfied. One danger of eliminating all (or some) links on the shortest path at once is removal of 'essential' links resulting in a disconnected network such that no more paths may be found. On the other hand, eliminating one link at a time might result in very similar, highly overlapping paths. This method is not flexible in the sense that it is difficult to force the method to generate particular alternatives that are relevant for the policy impact assessment.
Empirical test: Park, D. and Rilett, L.R. (1997) present empirical tests with simulation networks. Ramming, M.S. (2002) carried out empirical tests with real networks and observations.
Comment: This method is applicable only to single OD pairs. It may be computationally very efficient compared to the k-shortest path methods; however, the quality of the choice set might be very poor. Based on the results presented by all authors we conclude that the link elimination method is not a proper method for generating choice sets.

Method: **Link penalty method** (Iterative penalty method - IPM)
Original source: Johnson, P.E. and Joy, D.S. and Clarke, D.B. and Jacobi, J.M. (1992).
Application sources: De La Barra, T. and Perez, B. and Anez, J. (1993), Park, D. and Rilett, L.R. (1997), Scott et al. (1997), Akgün et al. (2000), Scott and Bernestein (2001).
Approach type: heuristic.
Definition: Generated set consists of k routes that are variations of the shortest path by successively applying a penalty at each link of the shortest path.
Principle: Repetitive applications of shortest path (SP) algorithm to successively changed link attribute values (increased impedances). As shown in Table 6.3, this method applies the Change Network Attributes procedure, in which the network is adapted according to some rules by changing the value of link travel time or by adding a penalty on the impedance of all links of the shortest path (see De La Barra, T. and Perez, B. and Anez, J. (1993)). Park, D. and Rilett, L.R. (1997) propose not to increase impedances of the links in the vicinity of the origin and destination in order to avoid producing minor deviations at the start and end of the route. Moreover, in this method the search criterion is not changed and the constraint set is empty. Finally, the termination criterion is based on the given maximum number k of paths to be generated: the choice set should be exhaustive with a maximum size of k.

Table 6.9: Satisfied quality criteria for the Link penalty approach.

	Quality Criteria	Satisfied	**Not** Satisfied
Reasonable route	*Acyclic criterion*	X	
	Detour criterion		X
	Hierarchic quality	X	
Adequate Choice set	*Overlap criterion*		X
	Comparability	X	
	Detour-max criterion		X
	Detour-min criterion		X
	Choice set size		X
	Spatial variability	X	
	Preferential variability		X
Appropriate CS generation approach	*Generic*		X
	Flexible		X
	Parsimony	X	

Advantage: This method matches several requirements for generating *adequate* choice sets with *reasonable* routes. The requirements satisfied are the acyclic criterion, routes do not contain any loops, and the hierarchic criterion applied by Park, D. and Rilett, L.R. (1997), by searching for a new path on those links that have a higher functional classification before examining links with a lower classification. For example, the major arterial roadways and highways would be examined first, the collectors second, and local streets last. The comparability criterion is also satisfied since the k shortest paths are variations of the shortest path all comparable to each other in terms of time or distance. The spatial variability criterion is also satisfied since the increasing of links impedances of the shortest path guarantees the generation of spatially different paths. Links are not deleted as in the link elimination algorithm. If a crucial link belongs to the shortest path (i.e. a bridge) the link penalty discourages the use of this link but that link may still be part of the new generated routes.

Disadvantage: Many requirements summarised in Table 6.9 are not satisfied. The method does not take into account the evaluation of the quality of the set of the generated paths in terms of satisfaction of the requirements previously defined, for example, overlap, detour-max, and detour-min criteria. This method is not flexible in the sense that it is difficult to force the method to generate particular alternatives that are relevant for the policy impact assessment.

Empirical test: Park, D. and Rilett, L.R. (1997) and Akgün et al. (2000) present empirical tests with simulation networks. Ramming, M.S. (2002) carried out empirical tests with real networks and observations.

Table 6.10: Satisfied quality criteria for the *k* dissimilar paths approach.

	Quality Criteria	Satisfied	**Not** Satisfied
Reasonable route	*Acyclic criterion*	X	
	Detour criterion		X
	Hierarchic quality		X
Adequate Choice set	*Overlap criterion*	X	
	Comparability		X
	Detour-max criterion		X
	Detour-min criterion		X
	Choice set size		X
	Spatial variability	X	
	Preferential variability		X
Appropriate CS generation approach	*Generic*	X	
	Flexible		X
	Parsimony		X

Comment: This method is applicable only to single OD pairs. As stated by Akgün et al. (2000), the Link penalty method satisfies the parsimony criterion because it is computationally much faster than *k*-shortest path because of its simplicity. Especially the change network procedure is very simple to be applied compared to the one applied to the *k*-shortest path methods. However, in this method several penalty strategies might be applied. For example, the level of the penalty might be too small or too large. A small penalty might not achieve sufficient dissimilarity, whereas a large penalty may eliminate many feasible paths from consideration. The elements to which the penalty might be applied might be only links, or also nodes. The penalty might be additive or multiplicative. All those settings make the method strictly problem-dependent and not easily portable (generic criterion not satisfied). Based on the results presented by all authors we conclude that the link penalty method is not a proper method for generating choice sets.

Method: **k dissimilar paths**
Original source: Akgün et al. (2000).
Application sources: Akgün et al. (2000), Carotenuto et al. (2004), Dell'Olmo et al. (2002).
Approach type: heuristic.
Definition: Finds a subset of *k* routes within the set of $K > k$ shortest paths so that the dissimilarity between any two selected paths is maximised above a certain minimum.

Principle: Repetitive applications of shortest path (SP) algorithm to successively changed network. As shown in Table 6-2, in this method the Change Network Attributes component is applied, in which the network is adapted by applying the same rules of either the *k*-shortest path algorithm or the penalty method. Moreover, in this method the search criterion is not changed. In this method the constraints set is defined in such a way that among the generated paths only sufficiently spatially dissimilar paths are selected according to some criteria. Finally, the termination criterion is based on the given maximum number of *k* paths to be generated: the choice set should be exhaustive with a maximum size of *k*.
Advantage: This method matches only few requirements for generating adequate choice sets with reasonable routes. The requirements matched are the acyclic criterion, routes do not contain any loops. The overlap and variability criteria are also satisfied since this method explicitly compares pairs of paths to check the spatially dissimilarity constraints criteria and avoids selecting paths similar to each other.
Disadvantage: Many other requirements summarised in Table 6.10 are not satisfied. For example, the size of the choice set depends on the number *k*, which in some cases might be much larger than the desired size. Some very long paths might be selected in the final set. This method is not flexible in the sense that it is difficult to force the method to generate particular alternatives that are relevant for the policy impact assessment.
Empirical test: Akgün et al. (2000) present empirical tests with simulation networks. Carotenuto et al. (2004) and Dell'Olmo et al. (2002) carried out empirical tests with real networks.
Comment: This method is applicable only to single OD pairs. It is not efficient (parsimony criterion not satisfied) because first the *k*-shortest paths are generated and afterwards a selection of all spatial dissimilar paths is made; from a computational point of view it may be slower compared to the Constraints *k*-shortest paths method. Based on those results we conclude that the *k* dissimilar paths method is not a proper method for generating choice sets.

Method: **Gateway method**
Original source: Lombard, K. and Church, R.L. (1993).
Application source: Akgün et al. (2000).
Approach type: heuristic.
Definition: Finds routes that are spatially dissimilar variations of the shortest path from a given origin to destination (OD pair) through a search of ”gateway shortest path” being the shortest path between the OD that is constrained to go through a specific node (or link), called *gateway*.

Table 6.11: Satisfied quality criteria for the Gateway approach.

	Quality Criteria	Satisfied	**Not** Satisfied
Reasonable route	*Acyclic criterion*		X
	Detour criterion		X
	Hierarchic quality		X
Adequate Choice set	*Overlap criterion*		X
	Comparability		X
	Detour-max criterion		X
	Detour-min criterion		X
	Choice set size		X
	Spatial variability	X	
	Preferential variability		X
Appropriate CS generation approach	*Generic*		X
	Flexible	X	
	Parsimony		X

Principle: Repetitive applications of shortest path (SP) algorithm using successively changed gateways according to some rules. As shown in Table 6.3, in this method the Change Restrictive Criteria procedure is applied by including all *gateway* nodes or links. All gateway links are determined at the beginning of the process and included in the restrictive criteria. In this case two shortest paths are searched, namely one from the origin to the gateway node (link) and the other from the gateway node (link) to the destination, together constituting the new generated OD-path. Moreover, in this method the Check Constraints box is not applicable since the constraint set is empty. Finally, the termination criterion is based on the total number of gateway nodes (links) given as input to the shortest path search; the method ends if all gateway nodes (links) are processed.

Advantage: This method matches only very few requirements for generating *adequate* choice sets with *reasonable* routes. Generation of large number of alternative paths by going through different gateways by using a SP search twice. It is possible to evaluate the similarity between two paths. The method produces an efficient set of paths in terms of their lengths and their dissimilarity but only from the shortest path. It is one of the very few methods that easily can address specific policy questions by specifying particular nodes or links (e.g. tolled links) as gateways (flexible criterion satisfied).

Disadvantage: All other requirements summarised in Table 6.11 are not satisfied. Some of the gateway paths may even contain loops (acyclic criterion). It is impossible to identify some dissimilar paths. The overlap, detour-max, and detour-min criteria are also not satisfied since very similar paths may be generated, because the method focuses only on the similarity check of candidates' paths to the shortest paths and not to the other paths found. It is unclear how to choose the best gateway nodes (links).

Empirical test: Akgün et al. (2000) present empirical tests with simulation networks.

Table 6.12: Satisfied quality criteria for the Essentially least-cost-paths approach.

	Quality Criteria	Satisfied	**Not** Satisfied
Reasonable route	*Acyclic criterion*	X	
	Detour criterion	X	
	Hierarchic quality		X
Adequate Choice set	*Overlap criterion*	X	
	Comparability	X	
	Detour-max criterion	X	
	Detour-min criterion	X	
	Choice set size		X
	Spatial variability	X	
	Preferential variability		X
Appropriate CS generation approach	*Generic*	X	
	Flexible		X
	Parsimony		X

Comment: This method is applicable only to single OD pairs. It is also not generic since the list of gateway nodes or links has to be established for each different network, and this may be slower from a computational point of view (parsimony criterion not satisfied). Based on those results we conclude that the gateway method is not a proper method for generating choice sets.

Method: **Essentially least-cost paths**
Original source: Hunt, D.T. and Kornhauser, A.L. (1997).
Application source: Hunt, D.T. and Kornhauser, A.L. (1997).
Approach type: heuristic.
Definition: Finds route set of all routes that have *locally acceptable* detours from the shortest path within a given cost threshold. This is different from the classical *k* least-cost path problem in that it operates on a cost threshold as opposed to a predetermined number of paths.
Principle: In a preliminary phase a repetitive application of shortest path (SP) search is applied to determine all detours that are acceptable within a given cost threshold, and that might be used in a path going from O to D. These detours are included as constraints in the constraints set and used in the next phase to determine all essentially-least-cost-paths. In this method, the constraints set is the set of all locally acceptable detours in the network for a given OD pair. All those detours form the basic building blocks for generating essentially-least-cost-paths. After determining the detours, the shortest path from O to D is searched and the Check Constraints Set is applied to find, based on those detours, all combinatorial route possibilities associated with the shortest paths. Finally, the termination criterion is based on the total number of detours within a given threshold belonging to the constraints set. The methods ends if no more paths might be constructed based on that detour list and on the shortest path.

Advantage: This method matches many requirements for generating *adequate* choice sets with *reasonable* routes. This method generates paths that are within a certain level of acceptability of detour from the shortest path based on a given cost threshold. Travellers might accept (tolerate) only a certain threshold of maximum travel time, or travel cost. Therefore, routes that exceed such value will not be generated. The overlap and spatial variability criteria are also satisfied since this method explicitly compares pairs of paths to check the overlap. The detour and comparability constraint criteria, which avoid selecting paths similar to each other, are also satisfied.
Disadvantage: Few other requirements summarised Table 6.12 are not satisfied. For example, the size of the choice set depends on the number of the detours, which in some cases might be much larger than the desired size. This method is not flexible in the sense that it is difficult to force the method to generate particular alternatives.
Empirical test: As far as we know there are no empirical tests done and presented in the literature regarding this method.
Comment: This method is applicable only to single OD pairs. It is computationally inefficient (parsimony criterion not satisfied) because of the possible combinatorial explosion of the generation of the essentially-least-cost-paths based on the detour generated in the network. Although this method seems attractive from several points of view is not good enough for complexity and application in real case.

6.5.5 RSG methods description of single and multiple OD-pairs group

Method: **Labelling approach**
Original source: Ben-Akiva, M.E. and Bergman, M.J. and Daly, A.J. and Ramaswamy, R. (1984).
Application sources: Ben-Akiva, M.E. and Bergman, M.J. and Daly, A.J. and Ramaswamy, R. (1984), Ramming, M.S. (2002).
Approach type: heuristic.
Definition: Finds route set that consists of all *labelled* shortest routes that each are optimal for a specific label from a given set of labels.
Principle: Repetitive applications of shortest path (SP) algorithm according to successively changed search criterion. As shown in Table 6.3, in this method the network is not changed. However, different Search Criteria are applied successively through the identification of so called label criteria such as: minimum time, minimum distance, minimum cost, maximum scenery, etc. in order to generate multiple possible paths that representing the fact that different travellers may have different objective functions in seeking best routes. For each label a specific search criterion (cost function) is defined and a related SP search is applied. In this method no constraints are considered; therefore the Check Set Constraints is not applicable. Finally, the termination criterion is based on the total number of labels considered.

Table 6.13: Satisfied quality criteria for the Labelling approach.

	Quality Criteria	Satisfied	**Not** Satisfied
Reasonable route	*Acyclic criterion*	X	
	Detour criterion		X
	Hierarchic quality		X
Adequate Choice set	*Overlap criterion*		X
	Comparability		X
	Detour-max criterion		X
	Detour-min criterion		X
	Choice set size	X	
	Spatial variability		X
	Preferential variability	X	
Appropriate CS generation approach	*Generic*		X
	Flexible	X	
	Parsimony	X	

Advantage: This method matches several requirements for generating an adequate choice set with reasonable routes. The labelling approach (Ben-Akiva, M.E. and Bergman, M.J. and Daly, A.J. and Ramaswamy, R., 1984) takes variation of taste into account by including in the path set all paths that are optimal with respect to a particular label. Heterogeneous choice alternatives are included in the choice set. The preferential variability criterion is then satisfied. This method is also flexible because it can easily address specific policy questions by specifying dedicated labels (e.g. by giving obligatory links or link types specific cost values).
Disadvantage: Several other requirements summarised in Table 6.13 are not satisfied. For example, identical routes may be generated with different labels. The spatial variability criterion is not satisfied.
Empirical tests: Ben-Akiva, M.E. and Bergman, M.J. and Daly, A.J. and Ramaswamy, R. (1984) and Ramming, M.S. (2002) performed empirical tests with real networks and observations by applying the labelling method. Ramming (2002) states that this method results in a good coverage (about 80%) of observed routes only by taking into account many attributes and labels. Results presented in those studies show that this method performs well in generating reasonable routes.

Table 6.14: Satisfied quality criteria for the Monte Carlo approach.

	Quality Criteria	Satisfied	**Not** Satisfied
Reasonable route	*Acyclic criterion*	X	
	Detour criterion		X
	Hierarchic quality		X
Adequate Choice set	*Overlap criterion*		X
	Comparability	X	
	Detour-max criterion		X
	Detour-min criterion		X
	Choice set size	X	
	Spatial variability		X
	Preferential variability		X
Appropriate CS generation approach	*Generic*	X	
	Flexible	X	
	Parsimony		X

Comment: This method is applicable not only to single OD pairs, but also to multiple OD pairs simultaneously at once (tree search from origin). The number of network attributes needed to define the labels may be a problem. It depends on the type of application and on the type of network; in some cases it is easy to define such labels (for example, in the waterway network and multi-modal network) in other cases it is not (generic criterion not satisfied). Form a computational point of view, this method is very easy and efficient (parsimony criterion satisfied). Based on those results we conclude that the labelling method is a proper method for generating choice sets.

Method: **Monte Carlo approach**
Original source: Sheffi, Y. and Powell, W.B. (1982).
Application sources: Sheffi, Y. and Powell, W.B. (1982), Ramming, M.S. (2002), Nielsen, O.A. (2002).
Approach type: heuristic.
Definition: Finds a route set by successive SP search on the network after randomisation of link attributes.
Principle: Repetitive applications of shortest path (SP) algorithm on same randomised network. As shown in Table 6.3, in this method the network is changed by randomly changing the attribute value of all links (e.g. travel times). The search criterion is not changed and the constraint set is empty. Finally, the termination criterion is based on the maximum number of iterations only.

Advantage: This method matches some requirements for generating an adequate choice set with reasonable routes. This method appears to find sufficient paths that are within a certain level of acceptability of the travellers (comparability criterion is satisfied). Ramming, M.S. (2002) has used utility functions in combination with simulation to generate a choice set.

Disadvantage: Several other requirements summarised in Table 6.14 are not satisfied. For example, the detour-max, detour-min and the spatial variability criterion are not satisfied. Randomising the link attributes only may determine more differentiation of the routes in terms of distance, time, or cost, for example, but not necessarily in terms of spatial difference.

Empirical tests: Nielsen, O.A. (2002) and Ramming, M.S. (2002) performed empirical tests with real networks and observations by applying the Monte Carlo method. Ramming, M.S. (2002) states that the method results in a good coverage (about 80%) of observed routes with a reasonable amount of computation time (about 20 hours). Results presented show that this method performs well in generating reasonable routes.

Comment: This method is applicable not only to single OD pairs, but also to multiple OD pairs simultaneously as well (tree search from origin). This method is also generic and flexible in the sense that the number of randomised variables to be included in the search criterion may be varied. Finally, this method may be a bit slow from a computational point of view (parsimony criterion not satisfied) because of the high number of iterations needed for generating an adequate choice set.

Because it is a probabilistic method, there is no guarantee for exhaustiveness of all attractive alternatives. However, most non-probabilistic methods have a similar drawback since most of these are heuristics (for a discussion see Section 7.4.4 ??). Based on those results we conclude that the Monte Carlo method is a proper method for generating choice sets, especially if some additional criteria (such as overlap and detour-max, detour-min criteria) are applied after the generation process.

Method: **Accelerated Monte Carlo approach**

Original sources: Bliemer, M.J.C. and Versteegt, E. and Castenmiller, R. (2004).

Application source: Bliemer, M.J.C. and Versteegt, E. and Castenmiller, R. (2004).

Definition: Finds a route set by successive SP search with randomisation of link attributes and by successively increasing the variance of the random variables. If after a given number of iterations a new path has not been found the variance is increased. The method is called *accelerated* since more paths are generated with less or the same number of iterations than the Monte Carlo approach due to the increase of the variance.

Table 6.15: Satisfied quality criteria for the Accelerated Monte Carlo approach.

	Quality Criteria	Satisfied	**Not** Satisfied
Reasonable route	*Acyclic criterion*	X	
	Detour criterion		X
	Hierarchic quality		X
Adequate Choice set	*Overlap criterion*		X
	Comparability	X	
	Detour-max criterion		X
	Detour-min criterion		X
	Choice set size		X
	Spatial variability		X
	Preferential variability		X
Appropriate CS generation approach	*Generic*	X	
	Flexible	X	
	Parsimony	X	

Principle: Repetitive applications of shortest path (SP) algorithm on same randomised network. As shown in Table 6.3, in this method the network is changed by randomly changing the attribute value of all links (e.g. travel times or costs). The accelerated version of the Monte Carlo approach uses a gradually increasing of the variance of the random components in the search criteria. The variance is increased if no more paths are found after a given number of iterations. In this method the search criterion is not changed and no constraints are applied. Finally, the termination criterion is based on the maximum total number of iterations.

Advantage: This method matches many requirements for generating adequate choice sets with reasonable routes. This method appears to find paths that are within a certain level of acceptability of the travellers (comparability criterion is satisfied).

Disadvantage: Other requirements summarised in Table 6.15 are not satisfied. For example, the detour-max, detour-min and the spatial variability criteria are not satisfied. Randomising the link attributes may determine more differentiation of the routes in terms of distance, time or cost, for example, but not necessarily in terms of spatial difference. Because of the increasing of the variance this method may generate much more routes than the simple Monte Carlo approach (the size of the final choice set might be much larger than the desired size).

Empirical tests: Bliemer, M.J.C. and Versteegt, E. and Castenmiller, R. (2004) present empirical outcomes for real networks but without observations.
Comment: This method is applicable not only to single OD pairs, but also to multiple OD pairs simultaneously (tree search from origin). This method is very flexible because the number of variables to be included in the search criterion as well as the level of variance of each may be varied. From a computational point of view, this method is faster than the simple Monte Carlo approach (parsimony criterion is satisfied); since the same number of routes might be generated with a lower number of iterations, because of the increasing variance. Because it is a probabilistic method, there is no guarantee for exhaustiveness of all attractive alternatives. Based on those results we conclude that this method may be applied for generating choice sets.

Method: **Monte Carlo Labelling combination approach**
Original source: Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001).
Application source: Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001).
Definition: Finds a route set by successive SP search with both randomisation of link attributes and search criteria.
Principle: Repetitive applications of shortest path (SP) algorithm on same randomised network successively using adapted search criteria. As shown in Table 6.3, in this method the network is changed by randomly changing the attribute values of all links (e.g. travel times or costs). Moreover, in this method the Change Routes Criteria is applied through the adoption of different label criteria such as: minimum time, minimum distance, minimum cost, maximum scenery, etc. in order to generate multiple possible paths meaning that different travellers may have different objective functions in seeking routes. For each label, a specific search criterion (cost function) is defined and a related SP search is applied. In this method no constraints are applied; therefore the Check Constraints box is not applicable. Finally, the termination criterion is based on the maximum number of iterations for the randomisation of the network and on the maximum number of iterations for the randomisation of link attributes.
Advantage: This method matches some requirements for generating adequate choice sets with reasonable routes. This method appears to find paths that are within a certain level of acceptability of the travellers (comparability criterion is satisfied). This method appears also to find sufficient paths that are spatially different due to the variation in search criteria that correspond to variation of travellers' preferences (spatial and preferential variability criteria are satisfied).

Table 6.16: Satisfied quality criteria for the Monte Carlo Labelling approach.

	Quality Criteria	Satisfied	**Not** Satisfied
Reasonable route	*Acyclic criterion*	X	
	Detour criterion		X
	Hierarchic quality		X
Adequate Choice set	*Overlap criterion*		X
	Comparability	X	
	Detour-max criterion		X
	Detour-min criterion		X
	Choice set size		X
	Spatial variability	X	
	Preferential variability	X	
Appropriate CS generation approach	*Generic*	X	
	Flexible	X	
	Parsimony		X

Disadvantage: Other requirements summarised in Table 6.16 are not satisfied. For example, because of the variation in the network and in the search criteria this method may generate much more routes than the simple Monte Carlo approach (the size of the final choice set might be much larger than the desired size). The size of the final choice set might be overestimated, but it is well known that generated choice sets for prediction purposes need to include the most attractive routes but may also including unattractive (non-used) routes (see Section 6.3.3). Overlapping routes might be also generated (overlap, detour-max, and detour-min criteria are not satisfied).
Empirical tests: Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001) present empirical results for real networks.
Comment: This method is applicable not only to single OD pairs, but also to multiple OD pairs simultaneously (tree search from origin). It is very flexible because the number of variables to be included in the search criterion as well as the variance of the parameter of each attribute may be varied. From a computational point of view, this method is more time-consuming than the simple Labelling or Monte Carlo approaches (parsimony criterion is satisfied). Because it is a probabilistic method, there is no guarantee for exhaustiveness of all attractive alternatives. Based on those results we conclude that this method is a good method for generating adequate choice sets with reasonable routes, especially if a filter process is applied after the generation process, in order to apply some important criteria such as the overlap, and detour-max, and detour-min criteria.

Method: **Accelerated Monte Carlo Labelling combination approach**

Original source: Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001).

Application source: Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001).

Definition: Finds a route set by successive SP search with randomisation of link attributes, search criteria, and by successively increasing the variance of the random variables. If after a given number of iterations a new path has not been found the variance is increased. The method is called *accelerated* since more paths are generated with less or the same number of iterations than the simple Monte Carlo Labelling approach due to the increase of the variance.

Principle: Repetitive applications of shortest path (SP) algorithm on same randomised network successively using adapted search criteria. As shown in Table 6.3, in this method the network is changed by randomly changing the attribute value of all links (e.g. travel times or costs). Moreover, in this method the Change Routes Criteria is applied through the adoption of different label criteria such as: minimum time, minimum distance, minimum cost, maximum scenery, etc. in order to generate multiple possible paths meaning that different travellers may have different objective functions in seeking routes. For each label, a specific search criterion (cost function) is defined and a related SP search is applied. The accelerated version of the Monte Carlo Labelling approach uses a gradually increasing of the variance of the random components in the search criteria. The variance is increased if no more paths are found after a given number of iterations. In this method no constraints are applied. Finally, the termination criterion is based on the maximum total number of iterations and on the maximum number of iterations for the randomisation of link attributes.

Advantage: This method matches some important requirements for generating adequate choice sets with reasonable routes. This method appears to find paths that are within a certain level of acceptability of the travellers (comparability criterion is satisfied).

Disadvantage: Other requirements summarised in Table 6.17 are not satisfied. For example, because of the increasing of the variance this method may generate much more routes than the simple Monte Carlo Labelling approach (the size of the final choice set might be much larger than the desired size). The size of the final choice set might be overestimated, but it is well known that generated choice sets for prediction purposes need to include the most attractive routes but may also including unattractive (non-used) routes (see Section 6.3.3). Some important criteria are not satisfied such as detour-max, detour-min and overlap criteria. *Empirical tests*: Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001) present empirical tests on simulation networks.

Table 6.17: Satisfied quality criteria for the Accelerated Monte Carlo Labelling approach.

	Quality Criteria	Satisfied	**Not** Satisfied
Reasonable route	*Acyclic criterion*	X	
	Detour criterion		X
	Hierarchic quality		X
Adequate Choice set	*Overlap criterion*		X
	Comparability	X	
	Detour-max criterion		X
	Detour-min criterion		X
	Choice set size		X
	Spatial variability	X	
	Preferential variability	X	
Appropriate CS generation approach	*Generic*	X	
	Flexible	X	
	Parsimony		X

Comment: This method is applicable not only to single OD pairs, but also to multiple OD pairs simultaneously (tree search from origin). This method is very flexible because the number of variables to be included in the search criterion as well as the level of variance of each may be varied. From a computational point of view, this method is faster than the simple Monte Carlo Labelling approach, since the same number of routes might be generated with a lower number of iterations, because of the increasing variance. Anyway, the parsimony criterion is not satisfied because this method is more time-consuming than the simple Monte Carlo or Labelling method. Because it is a probabilistic method, there is no guarantee for exhaustiveness of all attractive alternatives. Based on those results we conclude that this method may be applied for generating choice sets, especially if some not satisfied criteria during the generation process are applied after the generation process.

6.5.6 Comparative evaluation of route set generation methods

In this section, the methods described above are compared based on their matching of requirements described in Sections 6.3 and 6.4. Some requirements appear to be met only very seldom such as hierarchic quality and flexible criteria. The main aim is to establish which methods match better our requirements of generating an adequate choice set and therefore which are the most suitable methods. Our comparison of the methods and the requirements is made on the basis of two tables from which the best methods are selected.

For sake of simplicity we distinguish two groups: one is the group of methods applicable only to generate paths for a single OD pair, and the other for simultaneously generating paths for single and multiple OD pairs. Table 6.18 lists the methods applicable to single OD-pair while Table 6.19 shows the methods applicable to single and multiple OD-pairs. In both tables all examined quality criteria are listed and for each method it is indicated which criteria are satisfied. Please, note that the Constrained enumeration method based on B&B technique is not taken into account in this comparative evaluation since it is not considered as realistic candidate for generating choice sets.

With respect to the quality criteria only very few appear to be considered by nearly all methods. This holds for Acyclic routes and Spatial variability. The important criterion of Hierarchic Quality is only addressed by one method (Link Penalty). On the other hand, there are a few methods that succeed in addressing a majority of the proposed criteria.

Single OD-pair group (Table 6.18): Among all previously described methods applicable to a single OD pair, the CKSP, the constrained enumeration, and the least-cost paths are the three methods that match the majority (7 criteria over 10 are satisfied) of the requirements to generate an adequate choice set. However, the constrained enumeration and the least-cost paths method are computationally very cumbersome while the CKSPT is computationally much more efficient because the check of the constraints is applied during the generation process, without generating non-feasible paths. Therefore, we conclude that the CKSP is the best method among all methods applicable to single OD pair.
The Gateway method is one of the two procedures in this group that easily can fulfil the flexible criterion. This method therefore may be a useful pre-processor for other methods in generating initial solutions.

Single and Multiple OD-pairs group (Table 6.19): Among all previously described methods applicable to single and multiple OD pairs the Monte Carlo Labelling combination (MCL) and the Accelerated Monte Carlo Labelling (AMCL) combination are the only two methods that satisfy 4 criteria over 10. The number of criteria satisfied by those methods seems to appear little; however, we should notice that some criteria, such as the overlap, detour-max and detour-min criteria, even if they are not satisfied during the generation process, might be applied after the generation process, in a filtering process, in order to obtain a much better adequate choice set.
Indeed, not all criteria introduced in Section 6.3 and Section 6.4 have the same relevance, meaning that some criteria, such as the spatial and preference variability and the flexible criteria, appear to be more important since they can be satisfied only if explicitly applied and included in the generation process, such as the MCL and the AMCL approaches do.

Table 6.18: Satisfied quality criteria for Single OD-pair approaches.

Single OD pair methods		KSP	CKSP	Const. enum.	Link elim.	Link penalty	k-diss. paths	Gateway	Least-cost paths
	QC								
Reasonable route	A	X	X	X	X	X	X		X
	D		X	X					X
	H					X			
Adequate Choice set	O		X	X			X		X
	C	X	X	X	X	X			X
	Dx		X	X					X
	Dn		X	X					X
	CS								
	SV		X	X	X	X	X	X	X
	PV								
Appropr. CS gener. approach	G	X	X		X		X		X
	F			X				X	
	P				X	X			

Whereas other criteria that are not satisfied during the generation process can be easily applied after the generation process as it is shown in Figure 6.10, this is the case for the overlap, detour-max and detour-min constraints. The filtering process is a procedure, which compares the routes generated by the generation process by taking into account the constraints that must be satisfied and it produces as result a more adequate choice set that satisfied also the constraints. During the filtering process, pair-wise comparisons of the generated routes are applied and checked, routes that do not satisfy the constraints are eliminated form the choice set.

A special concern of the probabilistic methods in this group is that there is no guarantee that all attractive alternatives are indeed generated (no guaranteed exhaustiveness). Even with very many iterations, there remains always a non-zero probability that one or more important alternatives may not have been generated (see for a discussion Section 7.4.4).

The question of *exhaustiveness* reached with each method is not addressed in the method descriptions. Not only probabilistic methods suffer from this issue, also with all other non-probabilistic methods no guarantee can be given that all attractive alternatives will have been generated after a certain number of steps. It is a subject of future empirical research to derive conclusions on the level of exhaustiveness achievable with the various methods.

Table 6.19: Satisfied quality criteria for Single and Multiple OD pair approaches.

Multiple OD pair methods		Label.	Monte Carlo (MC)	AMC	MCL	AMCL
	Quality Criteria (QC)					
Reasonable route	*Acyclic* (A)	X	X	X	X	X
	Detour (D)					
	Hierarchic (H)					
Adequate Choice set	*Overlap* (O)					
	Comparab. (C)		X	X	X	X
	Detour-max (Dx)					
	Detour-min (Dn)					
	Choice set size (CSS)	X	X			
	Spatial variab. (SV)				X	X
	Preferential var. (PV)	X			X	X
Appropriate CS generation approach	*Generic* (G)		X	X	X	X
	Flexible (F)	X	X	X	X	X
	Parsimony (P)	X		X		

Finally, among all approaches for generating choice sets described in this chapter only few methods have been implemented and tested on real networks and only very few generation methods have tested using empirical route set data (only Ben-Akiva, M.E. and Bergman, M.J. and Daly, A.J. and Ramaswamy, R. (1984), Ramming, M.S. (2002), Cascetta, E. and Russo, F. and Viola, F.A. and Vitetta, A. (2002), Nielsen, O.A. (2002), Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001), Prato, C.G. and Bekhor, S. (2006)).

In the case that a filter process is applied after the generation process for all methods listed in Table 6.19, the following conclusions can be drawn from the comparison of those methods.

1. The Labelling approach has a limited coverage of the generated routes, since the number of labels taken into account is limited.
2. The Monte Carlo approach tends to generate routes that have small variation and do not vary from a spatial point of view.
3. The Accelerated Monte Carlo (AMC) approach generates routes that similar spatial variability of the ones generated by the Monte Carlo approach. However, the computation time of the AMC might be reduced because of the increasing of the variance, but a huge number of routes might be generated with this method, therefore a filtering process is really necessary to obtain a choice set with a workable size.
4. The Monte Carlo Labelling (MCL) combination approach generates a more adequate choice set with respect to the choice set requirements (especially spatial and preferential variability are satisfied during the generation process), since both labels and variation in network link attributes and travellers preferences are applied.
5. The Accelerated Monte Carlo Labelling (AMCL) combination tends to generates also an adequate choice set as the one generated by the MCL method, even with less computation time. However, the choice set size of the AMCL approach might also be greater than the one obtain with the MCL method, therefore in this case a filtering process is highly recommended.

To summarize, route sets generated by the MCL and AMCL methods can be considered adequate if and only if a filtering process is applied after the generation process, in which some criteria, such as the overlap and detour-max and detour-min criteria are satisfied; other most important criteria such as the flexible criteria and the spatial and preference variability are satisfied during the generation process. In this case, those two methods can be considered the most relevant approaches for generating choice sets at group level and for prediction purposes. In Chapter 7 only the MCL method is taken into account since the AMCL is an extended version of the MCL approach and it can be always applied.

Based on this qualitative evaluation, we can conclude that the Monte Carlo Labelling combination approach is the one to be preferred from the other methods in order to generate an adequate choice set for prediction purposes and for groups of travellers. Please note that in this case it is highly recommended to apply a filtering process just after the generation process in order to obtain a much more adequate choice set that satisfies the requirements introduced in Section 6.3.

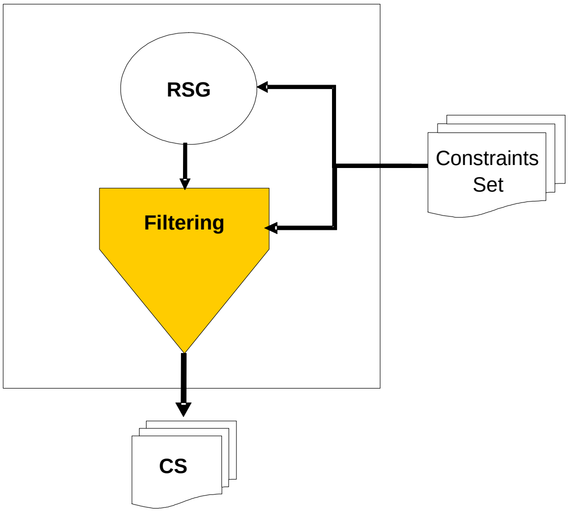


Figure 6.10: Filtering process applied after the RSG.

6.6 Applications of CSG approaches to uni-modal networks

Some of the choice set generation approaches described in the previous sections have been implemented and tested on two real networks: the Dutch waterway network and the Dutch main road network. The selected approaches are the following five:

1. KSP tested on the Dutch waterway network.
2. CKSP tested on the Dutch waterway network.
3. Monte Carlo Labelling (MCL) tested on the Dutch waterway network.
4. Accelerated Monte Carlo Labelling (AMCL) tested on the Dutch waterway network.
5. Accelerated Monte Carlo (AMC) tested on the Dutch main road network.

Only the last approach is tested on the Dutch main road network while all the others on the Dutch waterway network. The last three approaches have been selected for their capabilities in generating adequate choice set with reasonable routes suitable for both single and multiple OD pairs. The other two methods have been implemented mainly for sake of comparison to the results of the previous good methods and among the results of the two KSP and CKSP methods, which are mainly methods suitable only for single OD-pair.
Some characteristics of the real networks used for those applications are presented in Table 6.20 such as number of links, nodes, and OD pairs and so on.

In the following subsections, first, the networks and the application's purposes are described in more details, and then the characteristics of the adopted CSG methods are presented and described, while finally the results of the generated choice sets of all methods are presented.

Table 6.20: Sample of characteristics of the networks analysed with the CSG approaches.

Network characteristics	Main road network	Waterway network
Network length [km]	42,662	11,472
# directed links	25,434	2,500
# nodes	10,801	1,000
# zones	400 (345)	6
# non-zero OD pairs	109,292	30

6.6.1 Networks and applications characteristics

Two applications have been carried out on two different transport networks: one applied on the Dutch waterway network and the other on the Dutch main road network.

The purpose of the application on the waterway network (see Figure 6.11) is related to the forecasting model of freight transport demand (Fiorenzo-Catalano, S. and Van Der Zijpp, N.J., 2001). Its main aim is the prediction of vessel movements over the waterway network. One component of this application deals with the generation of route sets to be used in the assignment framework.

Characteristic for this freight assignment model is the presence of a ship choice model that not only determines the ship type but also the amount of tons for each trip. Depth is an important constraint in the route choice process. The depth of a ship mainly depends on the weight of its cargo. Therefore it is important that the cargo size determined by the ship choice model is responsive to the depth attributes of the network links.

The test network considered in this application is the waterway network of The Netherlands consisting of about 1,000 nodes and 2,500 directed links (see Figure 6.12). The cost associated with each links is simply its distance. A few OD pairs (30 OD pairs) were identified on the basis of the most frequent inland navigation trips and some constraints on the aspects of height and width were associated with a certain numbers of links.

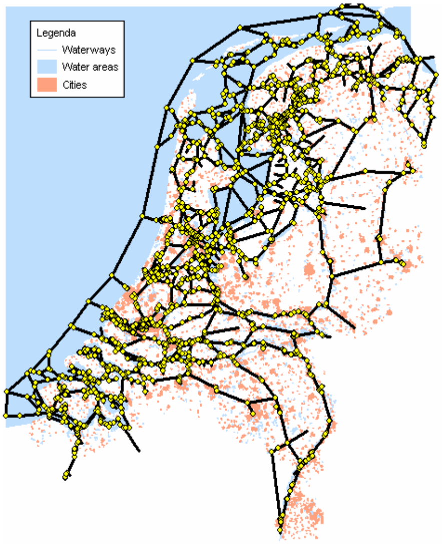


Figure 6.11: Waterway network of The Netherlands.

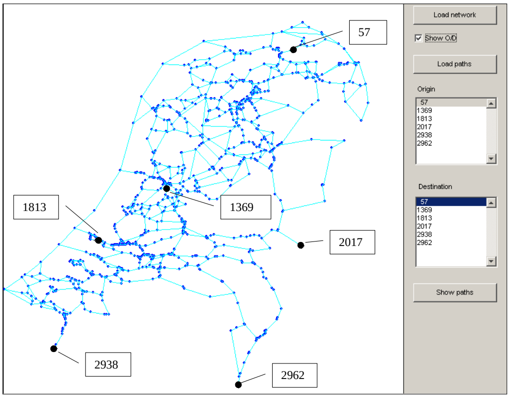


Figure 6.12: Waterway network of The Netherlands with six OD pairs.

The six origins and six destinations shown in Figure 6-12 are the following:

57 Groningen

1369 Amsterdam

1813 Rotterdam

2017 Duisburg (border with Germany)

2938 Antwerp (border with Belgium)

2962 Liege (border with Belgium)

In contrast, the purpose of the application on the main road network is to perform a dynamic traffic assignment (Bliemer, M.J.C. and Versteegt, E. and Castenmiller, R., 2004). Its main aim is to show the validity of a new route-based analytical multiclass DTA model and its capability of dealing with large networks. One component of this application deals with the generation of route sets to be used in the assignment framework.

Figure 6.13 shows the road network considered for the experiment that is the Dutch national road network, which consists of all freeways in The Netherlands and other main arterials and urban roads, with in total 25,434 directed links (with given capacity, speed limit, and road type), 10,801 nodes, and 400 zones, to be more precise 345 are the zone with non-zero traffic flows (see Table 6.20). The costs associated with each link are distance and time.

In this case, choice alternatives are defined as sequences of road segments, which can be represented as routes in a network. The definition of choice set is particularly relevant since the topological complexity of the network might generate an unrealistic high number of routes connecting a single origin-destination pair.

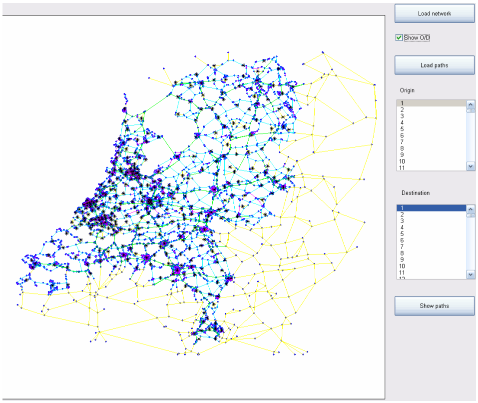


Figure 6.13: National road network of The Netherlands

6.6.2 Characteristics of the selected CSG methods

Table 6.21 illustrates the main characteristics of the five selected methods.

In the KSP method applied to the Dutch waterway network, Formula 6.7 is applied as generation function while distance D is the only link attribute considered as search criterion. In this case the cost function is given by the sum of all link distances belonging to the path.

$$\bar{C}_k = \sum_{a \in k} D_a \tag{6.7}$$

The KSP method does not apply any specific constraint for the detour-max, detour-min, and overlap criteria. The termination criterion applied is the total number of generated paths for each OD pair, set to 100 in this specific application.

In the CKSP method applied to the Dutch waterway network, the same Formula 6.7 of the KSP method is applied as generation function while distance D is the only link attribute considered as search criterion. In this case the cost function is given by the sum of all link distances belonging to the path.

The CKSP method does apply the detour-max and detour-min criteria, but does not the overlap criterion. The value of the parameters used for the detour-max and detour-min criteria are $\omega_{max} = 2$ and $\omega_{min} = 50km$ respectively. We refer to Van Der Zijpp, N.J. and Fiorenzo-Catalano, S. (2005) for more details about the type of constraints adopted in this method. The termination criterion applied is the total number of generated paths for each OD pair, set to 100 in this specific application.

In the Monte Carlo Labelling method (MCL) applied to the Dutch waterway network, Formula 6.8 is applied as generation function. In this case, k represents the path index, a the link index, distance D is the only link attribute considered as search criterion.

$$\bar{C}_k = \sum_{a \in k} (D_a + \varepsilon_a) \tag{6.8}$$

Where ε_a is the standardized random term associated with the link distance attribute being random normal distributed in $(0,\sigma_a^2)$. The variance of the error terms σ_a^2 is equal to:

$$\sigma_a^2 = \varphi \cdot D_a \tag{6.9}$$

Where φ is the variance parameter and the error term ε_a is defined as follows:

$$\varepsilon_a = \rho \cdot \sigma_a = \rho \cdot \sqrt{\varphi \cdot D_a} \tag{6.10}$$

Where ρ is the standardized random term being random normal distributed in (0, 1). Therefore, by substituting Formula 6.10 in Formula 6.8:

$$\bar{C}_k = \sum_{a \in k} \left(D_a + \rho \cdot \sqrt{\varphi \cdot D_a} \right) \tag{6.11}$$

The MCL method does not apply any specific constraint for the detour-max and detour-min criteria, only the overlap constraint has been adopted; only routes that have all links (exactly 100% of the links) of a previously generated route in common are rejected (note that in this specific case the costs of the two routes are not compared, the rejected route is simple the last found, and not the longest). The termination criterion applied is the total number of iterations, set to 2,500,000 in this specific application. At each iteration the all-to-all shortest paths are computed for all OD pairs and several ship sizes (as route labels) are considered to simulate some specific preferences, in order to generate routes that are spatially different.

In the Accelerated Monte Carlo Labelling method (AMCL) applied to the Dutch waterway network the same Formula 6.8 of the MCL approach is applied as generation function while distance D is the only link attribute considered as search criterion. In this accelerated version of the method, if no new path is found after 15 sec. (of CPU time) variance parameter φ at iteration it is increased proportional to the parameter λ as follows:

$$\varphi_{it} = \lambda \cdot \varphi_{it-1} \tag{6.12}$$

where λ is the variance increase parameter, with the initial value equal to 1.001 and the value of the variance parameter φ_0 equal to one.

The AMCL method does not apply any specific constraint for the detour-max and detour-min criteria, only the overlap constraint has been adopted; routes that have all links (exactly 100% of the links) of a previously generated route in common are rejected. Note that in this specific case the random costs of the two routes are not compared, the rejected route is simply the last found and not the longest. However, it is very probable that the latter one is the longer one. The termination criterion applied is the total number of iterations set to 150,000 in this specific application. In this case also, at each iteration the all-to-all shortest paths are computed for all OD pairs and several ship sizes (as route labels) are considered to simulate some specific preferences, in order to generate routes that are spatially different.

In the Accelerated Monte Carlo method (AMC) applied on the Dutch main road network, Formula 6.13 is applied as generation function. Accelerated means that the variance in the randomisation process can successively be increased in order to improve the generation process. In this case, k represents the path index, a the link index, and time T is the only link attribute considered as search criterion.

$$\bar{C}_k = \sum_{a \in k} (T_a + \varepsilon_a) \tag{6.13}$$

Where ε is the standardized random term associated with the time attribute being random normal distributed in (0, σ_a^2). The variance of the error terms σ_a^2 is equal to:

$$\sigma_a^2 = \varphi_{it} \cdot T_a \tag{6.14}$$

Where φ_{it} is the variance parameter at iteration it and the error term ε_a is defined as follows:

$$\varepsilon_a = \rho \cdot \sigma_a = \rho \cdot \sqrt{\varphi_{it} \cdot T_a} \tag{6.15}$$

Where ρ is the standardized random term being random normal distributed in (0, 1). Therefore, by substituting Formula 6.15 in Formula 6.13:

$$\bar{C}_k = \sum_{a \in k} \left(T_a + \rho \cdot \sqrt{\varphi_{it} \cdot T_a} \right) \tag{6.16}$$

Since the approach is ”accelerated”, at each iteration variance parameter φ at iteration it is increased proportional to the variance increase parameter λ as follows:

$$\varphi_{it} = \varphi_{it-1} + \lambda \tag{6.17}$$

with the initial value of ε_0 equals to zero and the value of the parameter λ equals to 0.005.

The AMC method does not apply any specific constraint for the detour-max and detour-min criteria, only the overlap constraint has been adopted; routes that have more than 90% of links of a previously generated route in common are rejected. Note that in this specific case the costs of the two routes are not compared, the rejected route is simply the last found and not the longest. However, it is very probable that the latter one is the longer one. The termination criterion applied is the total number of iterations, set to 50 in this specific application.

Table 6.21: Characteristics of adopted CSG approaches for Dutch waterway and main road networks (N.A. Not Applicable).

CSG characteristics	Waterway network				Main road network
CSG method	KSP	CKSP	MCL	AMCL	AMC
Generation function	6.7	6.7	6.8	6.8 6.12	6.13 6.17
Search criteria	Distance	Distance	Distance	Distance	Time
Detour-max	N.A.	$\omega_{max} = 2$	N.A.	N.A.	N.A.
Detour-min	N.A.	$\omega_{min} = 50km$	N.A.	N.A.	N.A.
Overlap	N.A.	N.A.	$< 100\%$	$< 100\%$	$< 90\%$
Termination criterion	# paths	#paths	#iterations	#iterations	#iterations
Max # iterations	N.A.	N.A.	2,500,000	150,000	50
#paths per OD pair	100	100	N.A.	N.A.	N.A.

6.6.3 Results of the selected CSG approaches

A certain number of routes and a certain type of choice set have been generated by each implemented method for a given number of OD pairs.

Table 6.22 summarizes the results of the applications.

In the case of the KSP approach applied to the Dutch waterway network, a total number of 3,000 paths has been generated for a total number of 30 OD pairs, since 100 was set as input value for the total number of paths to be generated for each OD pair. It may include highly overlapping paths.

Table 6.22: Results from CSG applications to Dutch transport networks.

CSG results	Waterway network				Main road network
#analysed OD pairs	30				109,292
Method	KSP	CKSP	MCL	AMCL	AMC
# Generated routes	3,000	540	562	20,869	181,567
Average # of generated routes per OD pair	100	18	19	696	1.66
Minimum # of generated routes per OD pair	100	6	1	12	1
Maximum # of generated routes per OD pair	100	48	58	5,066	14
Median # of generated routes per OD pair	100	13	9.5	195.5	1

In the case of the CKSP approach applied to the Dutch waterway network, a total number of 540 paths has been generated for a total number of 30 OD pairs, which is on average 18 routes per OD pair. In this case, the minimum and maximum number of generated routes for a single OD pair appears to be 6 and 48 respectively and the median number for all OD pairs is 13. It may include some overlapping paths.

In the case of MCL approach applied to the Dutch waterway network, a total number of 562 paths has been generated for 30 OD pairs, which is on average 19 routes per OD pair. In this case, the maximum number of generated routes for a single OD pair appears to be 58 and the median number for all OD pairs is 9.5.

In the case of AMCL approach applied to the Dutch waterway network, much more routes are generated: in total 20,869 paths for the same 30 OD pairs, which is on average 696 routes per OD pair, which seems a very huge number, but it becomes reasonable if we look at the median number that is 195.5. In this case the maximum number of generated routes for a single OD pair appears to be 5,066.

In the case of the Accelerated Monte Carlo approach applied to the Dutch national road network (Bliemer, M.J.C. and Versteegt, E. and Castenmiller, R., 2004), the route generation model produced 181,567 routes in just 5 minutes for the 109,292 OD pairs, which is on average 1.66 route per OD pair. This may seem little, but taking into account that for long distance trip the only network feasible is mostly a freeway system with many triangular shapes, there are many OD pairs for which only one route is actually a realistic option. Such small average number of routes per OD pair depends also on the assumed error variance, which is in the reported example is very little with very small variation in the accelerated version (only 0.005 for the variance increase parameter λ). The maximum number of routes generated for an OD pair is 14.

As we can see from Table 6.22 (and also Table 6.23 for the waterway network) the size of the route choice set strongly depends on the type of network and the type of CSG approach. On the one hand, the enormous density of the waterway network with many possible detours and overlapping routes leads to a multitude of routes (see, for example, Figure 6.17 the huge number of 5,066 generated routes by the AMCL approach); on the other hand, in a very sparse network such as the road network used for the AMC approach only few alternatives are commonly available.

Table 6.23 shows the choice set sizes for each OD pair generated by each approach in the case of the waterway network. We will analyse the results of the generated choice sets in the waterway network, since only 30 OD pairs are considered in that case. The paths of the choice sets generated for the 10^{th} OD pair will be plotted in the following, and their choice set sizes are highlighted in bold. Excluding the choice set size of the KSP, which are all equal to 100, and the one of the AMCL, for some of them the size is very huge, we now analyse in more detail the choice set sizes generated by the CKSP and MCL approaches.

Table 6.23: Choice set size for each OD pair and each generation method applied to the waterway network.

OD Pair	KSP	CKSP	MCL	AMCL
1	100	8	5	61
2	100	11	8	22
3	100	12	3	26
4	100	11	9	49
5	100	41	44	411
6	100	11	9	107
7	100	15	2	93
8	100	10	3	53
9	100	20	5	306
10	**100**	**30**	**58**	**5066**
11	100	9	8	22
12	100	9	1	57
13	100	12	24	682
14	100	8	16	192
15	100	18	12	265
16	100	32	4	37
17	100	26	2	41
18	100	21	36	839
19	100	6	6	13
20	100	48	40	2403
21	100	10	10	48
22	100	17	3	208
23	100	10	16	199
24	100	6	4	12
25	100	10	33	840
26	100	44	42	432
27	100	30	57	4986
28	100	16	23	367
29	100	25	44	2204
30	100	14	35	828

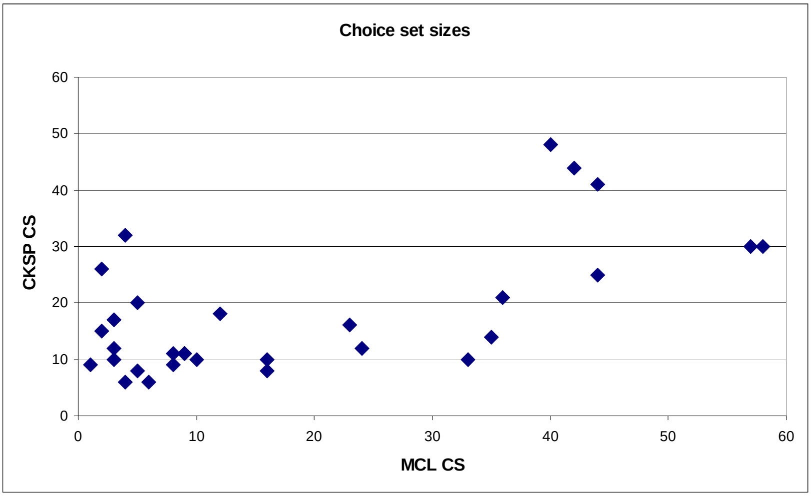


Figure 6.14: Resulting choice set sizes generated by the MCL and the CKSP approach for the waterway network (30 OD-pairs).

If we look more into detail to the choice sets generated by the CKSP approach, we can notice that only few OD pairs have a choice set size more than 40 routes (see Figure 6.14).

If we look more into detail to the choice sets generated by the MCL approach, we can notice that about 50% of the OD pairs (16 over 30) has a choice set size less than 10 (see Figure 6.15). By grouping the choice set size by 10 (from 10 up to 60) only very few OD pairs (2 or 4) have big choice sets (up to 58 alternatives). Moreover, Table 6.24 shows the number of OD pairs for each choice set size, from which it can be seen that only few OD pairs have big choice set sizes, while most of the OD pairs have choice sets with small sizes, in particular only one OD pair has a choice set of one alternative only. These outcomes are in line with the ones about choice set sizes and composition shown in Chapter 4.

If we plot the paths generated by each approach for a specific OD pair we can make some interesting notes. Figure 6.16 and Figure 6.17 show the paths generated by the KSP, CKSP, MCL, and AMCL approaches respectively for the specific 10^{th} OD pair (from Groningen to Antwerp) listed in Table 6.23.

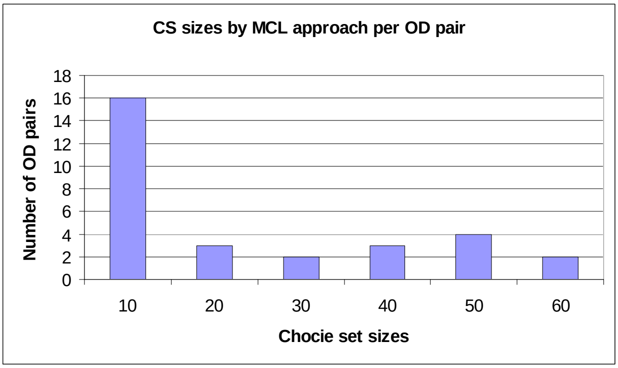
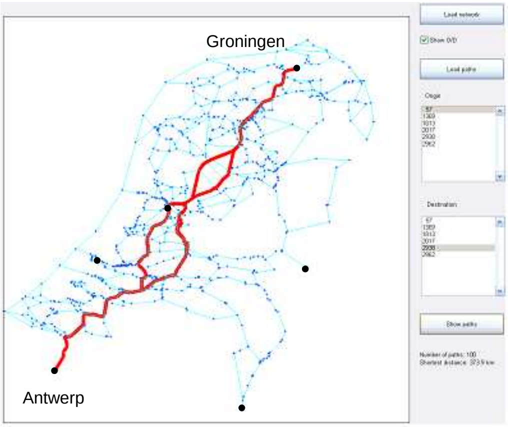


Figure 6.15: Resulting choice set sizes generated by the MCL approach per OD pair.

Looking at those pictures we can observe the following. The KSP approach generates 100 routes that are very overlapping and with little spatial variability (spatial criterion not satisfied). On the contrary for the same OD pair the CKSP approach generates only 30 routes with a much more spatial variability. Based on these results we can state that the CKSP approach outperforms the KSP approach. The MCL approach generates 58 routes with a spatial variability similar to the one generated by the KSP approach.

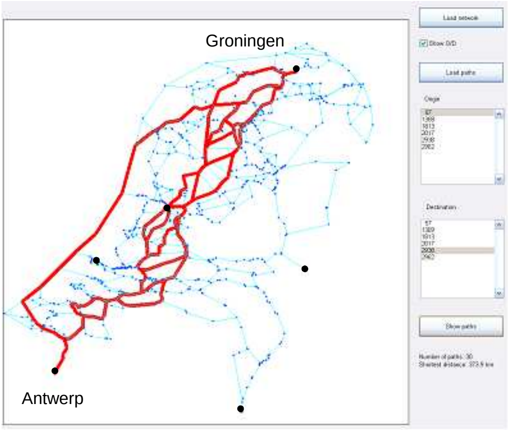
Finally, the AMCL approach generates a huge number of routes 5066 with a better spatial variability of the one generated by the CKSP approach. Ideally the RSG approach should generate an adequate choice set having, for example, a good choice set size (5066 is really too much) with a good spatial variability and not too much overlap, according to the defined criteria in Section 6.3.

Comparing the results of these approaches we can state that among the approaches suitable only for single OD pair the CKSP approach generates an adequate choice set, and among the approach suitable for both single and multiple OD pairs the MCL approach generates an adequate choice set with respect the choice size and the spatial variability and the AMCL approach generates a much better choice set with respect to the spatial variability, but in this case the number of generated routes is very huge, since they are also much overlapping. Both methods generate overlapping routes. To obviate this problem an extra process is needed to filter off all routes that do not satisfy, for example, the overlapping criterion. In the next section an example of filtering process is applied to a simple case to show how the quality of the choice set might be improved.



Number of paths: 100
Shortest distance: 373.9 km

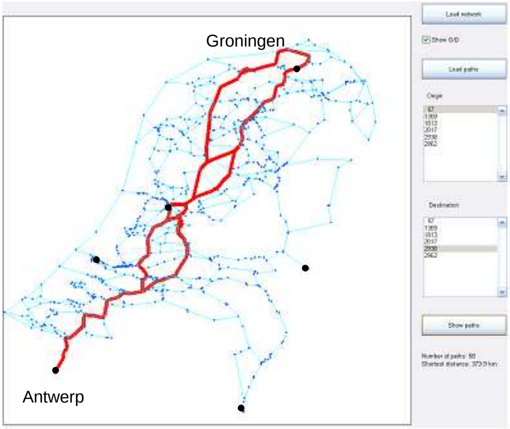
(a) KSP



Number of paths: 30
Shortest distance: 373.9 km

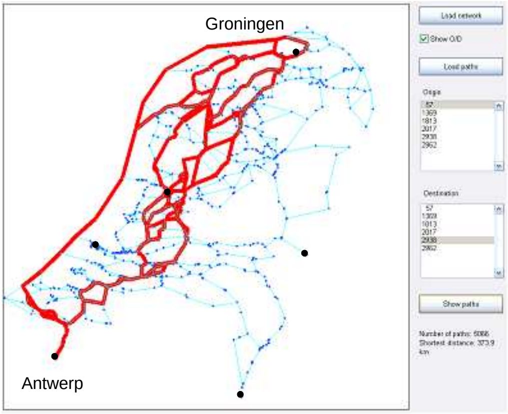
(b) CKSP

Figure 6.16: Choice set paths generated by the KSP approach (a) (100 paths) and the CKSP approach (b) (30 paths) from Groningen to Antwerp (OD n. 10).



Number of paths: 58
Shortest distance: 373.9 km

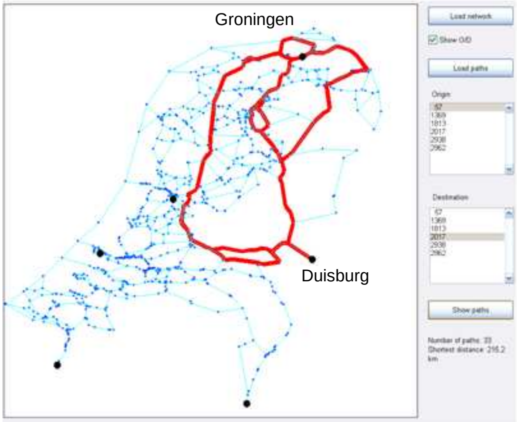
(a) MCL



Number of paths: 5066
Shortest distance: 373.9 km

(b) AMCL

Figure 6.17: Choice set paths generated by the MCL approach (a) (58 paths) and the AMCL approach (b) (5066 paths) from Groningen to Antwerp (OD n. 10).

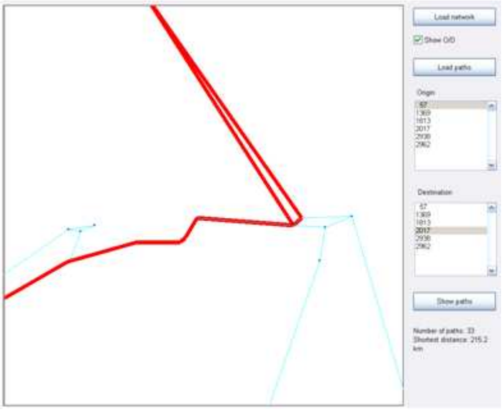


Number of paths: 33
Shortest distance: 216.2 km

Figure 6.18: Choice set paths generated by the MCL approach (33 paths) from Groningen to the German city of Duisburg (OD 25).

6.6.4 Application of a filtering process to the MCL choice set

As shown in the previous section, the sizes of the choice sets generated by the MCL and the AMCL approach are quite big, and in some cases also huge for the choice sets generated by the AMCL approach, with a maximum size of 5066 routes. As we will show in this section, this fact is mainly due to the overlapping problem. Routes generated by the MCL and AMCL approaches are largely overlapping; indeed those methods do not satisfy the overlap criterion defined in Section 6.3. As stated in Section 6.5.6, in order to obviate this problem a filter process needs to be applied after the generation process in order to obtain a much more adequate choice set, reducing the number of overlapping routes and maintaining the same spatial variability. The filter process consists of the application of one or more selection criteria defined in Section 6.3, such as the overlap criterion.



Number of paths: 33
Shortest distance: 215.2 km

Figure 6.19: Detail of two overlapping paths of the choice set generated by the MCL approach from Groningen to the German city of Duisburg (OD 25).

In order to show the effectiveness of the filtering process, we consider the choice sets generated by the MCL approach as example. We apply the filtering process, which consists of the application of the single overlap constraint to a specific OD pair. The idea of the filtering is to compare two routes, check which does not satisfy the overlap criterion, and eliminate those routes from the choice set, in order to obtain a more adequate choice set. For sake of simplicity we decide to consider the 25^{th} OD shown in Table 6.23, in which 33 paths are generated from Groningen to the German city of Duisburg. Figure 6.18 shows the set of 33 paths generated for this relation, from which it seems that it would be much less. Actually, as we zoom in this picture we can see from Figure 6.19 that in the waterway network a lot of detours and small variations are possible, this is the reason of the huge amount of overlap generated by the MCL approach.

In order to analyse the overlapping routes the following data were computed; first, the number of links and the total objective distance for each path (see Table 6.26), then the link-path incidence matrix was built, a path-path matrix was computed in which the number of links in common between two paths were calculated and finally the total number of two overlapping paths of a given overlap percentage was computed (see Table 6.25).

Table 6.24: Frequency of choice set size generated by the MCL approach.

Choice set size	# of OD pairs
1	1
2	2
3	3
4	2
5	2
6	1
8	2
9	2
10	1
12	1
16	2
23	1
33	1
35	1
36	1
40	1
42	1
44	2
57	1
58	1

After computing the mutual overlap as the percentage of the number of links in common between each of the 33 paths generated by the MCL approach for the 25^{th} OD pair, from this path-path matrix the total number of two paths overlapping each other was computed. Given the 33 paths the total number of path couples is given by the formula $n \cdot (n-1)/2$, in this case 528, which are the non-zero elements of the path-path matrix. Given those percentages, they were grouped from 10% to 100% with a 10% step and the total number of overlapping paths were computed based on those groups percentages (see Table 6.25). For example, from the path-path matrix, paths 2 and 3 have 6% of common links with path 1, paths 4 and 5 have 40% of common links with path 1, and path 22 has up to 98% of common links with path 1. From Table 6.25 we can see that most of the paths are largely overlapping (202 paths in total are overlapping more than 90%).

As we can see from Table 6.26 some paths have almost similar number of links and length (in some cases, such as paths 4 and 5, 10 and 11, 20 and 21, exactly the same number of links and the same length).

Table 6.25: Total number of two overlapping paths for each range of overlap percentage for all 33 paths generated by the MCL (OD 25).

# Overlap paths	% Overlap
104	10%
96	20%
5	30%
29	40%
12	50%
2	60%
18	70%
60	80%
102	90%
100	100%
528	

Therefore, we decide to apply the filtering process to obtain a more adequate choice set. In the filtering process the overlap criterion defined in Section 6.2 has been applied with the Formula 6.2 taking into account the distance of the common links and $\Delta = 0.9$; meaning that during the filtering process when two routes overlap for more than 90% the longest route is eliminated.

The result of the filtering process is that about 80% of routes (27) were eliminated. The 6 remaining ones are shown in Table 6.27. Comparing Figure 6.18 and Figure 6.20, please note that the spatial variability of the 6 routes is exactly the same of the 33 routes shown in Figure 6.18.

Table 6.28 shows the number of filtered paths for each OD pair. In bold you can find the 10^{th} OD pair with 58 routes without applying the filter process and with only 6 routes after applying the filter process. With the application of the filtering process there is a reduction of 73% of the number of routes.

As shown in Table 6.28 the statistics data about the choice set generated by the MCL approach are the following: a total number of 562 paths has been generated for 30 OD pairs, which is on average 19 routes per OD pair. In this case, the maximum number of generated routes for a single OD pair appears to be 58 and the median number for all OD pairs is 9.5.

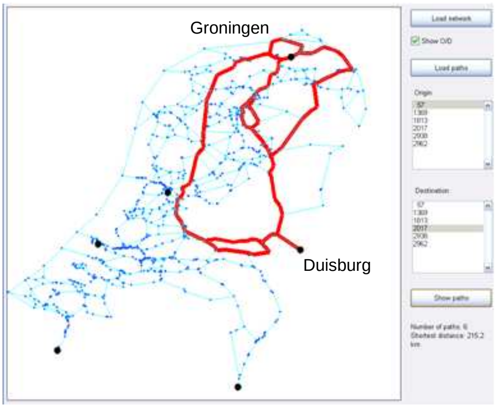
After applying the filter process the following results are achieved. After removing 408 paths (73% of reduction), a total number of 154 paths result for 30 OD pairs, which is on average 5.1 routes per OD pair. In this case, the maximum number of generated routes for a single OD pair appears to be 12 (the minimum is still 1 as the previous case), and the median number for all OD pairs is 5.

Table 6.26: Number of links and total distance for each of the 33 paths generated by the MCL approach for the relation Groningen-Duisburg (OD 25).

Path n.	Num. Links	Length (km)
1	50	267.5
2	61	351.0
3	60	369.8
4	**69**	**267.8**
5	**69**	**267.8**
6	61	351.2
7	59	351.2
8	60	369.9
9	58	369.9
10	**60**	**215.2**
11	**60**	**215.2**
12	71	267.9
13	59	351.3
14	69	267.9
15	71	267.9
16	57	215.7
17	58	370.1
18	66	268.2
19	60	215.4
20	**62**	**215.4**
21	**62**	**215.4**
22	51	267.9
23	66	268.4
24	69	267.9
25	60	215.4
26	71	268.1
27	62	215.6
28	71	268.1
29	62	215.5
30	61	215.6
31	57	215.9
32	70	268.2
33	70	268.1

Table 6.27: Routes resulting from the filtering process for OD 25.

Path	Num. Links	Length (km)
1	50	267.5
2	61	351.0
3	60	369.8
11	60	215.2
15	71	267.9
16	57	215.7



Number of paths: 6
Shortest distance: 215.2 km

Figure 6.20: Selected routes after applying the filtering process to the choice set generated by the MCL approach from Groningen to the German city of Duisburg (OD 25).

As shown in this section, the empirical outcomes show the validity for the implemented approaches. In the case of approaches suitable only for single OD pairs, such as KSP and CKSP, the generated choice sets generally fit the quality criteria introduced in Section 6.3, in the case of KSP a large amount of overlapping paths are generated, whereas in the case of CKSP a more adequate choice sets are generated, especially in terms of spatial variability and choice set size.

In the case of approaches suitable for both single and multiple OD pairs, the generated choice sets also fit the quality criteria for adequate choice sets. As we have shown in the case of the MCL and AMCL approaches a huge number of paths are generated, some of which are very overlapping, which is due also to the network structure. The spatial variability criterion is almost satisfied, whereas the overlap criterion mostly is not satisfied. For these reasons, it is highly recommended to apply a filtering process to reduce the number of overlapping routes and obtain a more adequate choice set.

Comparing the two types of approaches, one suitable only for single OD pairs (such as the CKSP), the other suitable both for single and multiple OD pairs (such as the MCL), we observe that the MCL approach fits much better the generation process criteria (introduced in Section 6.4) than the CKSP approach, since the MCL pproach is more flexible. Therefore, in the case of generating choice sets at group level and for prediction purposes in the case of uni-modal networks, the MCL and the AMCL approaches result to be the most appropriate RSG methods to be applied, in combination with a filtering process, which produces a more adequate choice set.

Finally, since the AMCL approach is the accelerated version of the MCL method in which increasing the variance might even further enhance its performance, but it also generates a huge number of routes compared to the ones generated by the MCL approach. Therefore, we conclude that Monte Carlo Labelling combination is the preferred method among all methods applicable to multiple OD pairs. However, it must be noticed that this method may require large amounts of computation time. This method may be even further improved by including specific constraints that exclude unsatisfactory routes and address the hierarchic property.

Based on the results about the generation process in the uni-modal case, we can conclude that also for the multi-modal case, in which the network is much more complicated, the multi-modal paths are much more complex, and flexibility in the generation process is needed, a stochastic RSG approach (mainly an extension of the MCL method) appears to be most suitable in order to generate adequate choice sets.

Table 6.28: Routes resulting with overlap $< 90\%$ after the filtering process for all 30 OD pairs (N.A. Not Applicable).

OD Pair	MCL without Filter	MCL with Filter	Rejected
1	5	2	3
2	8	6	2
3	3	3	0
4	9	4	5
5	44	11	33
6	9	3	6
7	2	2	0
8	3	3	0
9	5	3	2
10	58	6	52
11	8	6	2
12	1	1	0
13	24	5	19
14	16	9	7
15	12	3	9
16	4	4	0
17	2	2	0
18	36	9	27
19	6	6	0
20	40	5	35
21	10	4	6
22	3	2	1
23	16	8	8
24	4	4	0
25	**33**	**6**	**27**
26	42	12	30
27	57	6	51
28	23	6	17
29	44	6	38
30	35	7	28
Total	562	154	408
Average	19	5	N.A.
Max	58	12	N.A.
Min	1	1	N.A.

6.7 Conclusion

Following a systematic approach, a set of quality criteria have been derived for assessing the adequacy of generated choice sets suitable for prediction application in travel demand modelling. The criteria are above all based on behavioural premises on the part of the individual traveller with respect to choice set size, route composition, and choice set composition. The criteria to be applied to choice set generation at aggregate level reflect the high variability of choice sets between user groups and trip types. Adequate choice sets for prediction therefore should not be too small and should preferably have a rich variety of their composition in terms spatial distribution and route properties. A choice set generation approach should allow responding to specific requirements.

Using these criteria, a large set of route set generation approaches have been analysed for their suitability as a choice set generation method. It turns out that a wide spectrum of methods exists based on a rich variety of route generation principles. Most of the methods do not meet the criteria for achieving adequate choice sets. From a systematic comparative evaluation only very few methods appear to be potentially suitable as a choice set generation method for our purposes. These are the Monte Carlo Labelling combination approach for multiple OD pairs if extended with a filter process applied after the generation process, and the Constrained K-Shortest Paths for the single OD-pair generation if extended with a pre-processor (e,g, Gateway method) to make the method much more flexible.

The feasibility of the Monte Carlo Labelling combination and the CKSP methods has been demonstrated by several applications to two uni-modal networks, namely the Dutch national road and waterway networks respectively. The effectiveness of the filter process applied to the choice sets generated by the MCL approach has also been demonstrated and in the case that a filter process is applied after the generation process a reduction of 73% is achieved.

Based on this quantitative evaluation, we can conclude that the Monte Carlo Labelling combination approach is the one to be preferred from the other methods in order to generate an adequate choice set at group level for prediction purposes. Please note that in this case it is highly recommended to apply a filtering process after the generation process in order to obtain more adequate choice sets that satisfy the quality criteria.

The MCL approach has been also successfully applied in a number of different networks. Recent applications are done on the Dutch national main road network in the course of traffic flow predictions with dynamic assignment model (see Bliemer, M.J.C. and Taale, H. (2006)). The stochastic choice set generation approach is a separate explicit step prior to the route choice modelling and route flow assignment. This application demonstrated that the way of explicit a priori route generation is feasible, even for very large networks.

In the next chapter the MCL method will be extended for use in multi-modal transport networks. In that chapter a preliminary calibration and in-depth testing of the new generation methods using also empirical route set data is presented.

Chapter 7

Route set generation in multi-modal networks

7.1 Introduction

This chapter deals with the extension of choice set generation approaches to multi-modal networks. This will be based on the choice set requirements and generation approaches presented in Chapter 6. In this chapter we will address the specifications of an adequate route choice set, in particular in the multi-modal context, and which requirements should be posed on such route choice sets to be adequate for the prediction purpose. A related question deals with the requirements an appropriate multi-modal choice set generation approach should satisfy. Other questions relate to the definition of the performance quality of the multi-modal choice set generation approach, the calibration of the model, and the analysis of the quality of the choice set. In this chapter, we tackle these questions specifically with an application to a multi-modal network in mind.

Based on the established adequacy requirements introduced in Chapter 6, this chapter will deal with the introduction of the specific aspects associated with multi-modal transportation, the extension of the requirements described in Chapter 6 for an adequate choice set and an appropriate choice set generation method in multi-modal networks, and the presentation of route choice generation methods proposed in literature for multi-modal transit networks and for mixed private and public multi-modal networks. The main purpose of this chapter is then to introduce and describe the new doubly stochastic approach for generating choice sets in a multi-modal network, to present a demonstration case of the application of the generation approach focusing on the analysis of the performance quality of the approach. The rest of this chapter will focus on the calibration of the model, the impact of the stochasticity of the approach, and the quality of the generated choice sets.

The main contributions of this chapter are, first of all, the newly developed route choice set generation algorithm, the so-called doubly stochastic approach, applicable to uni-modal and multi-modal networks, which is actually an extension of the MCL combination approach (presented in Chapter 6) where not only the link attributes but also the preference (or behavioural) parameters of the cost function are randomized; second, the specification of the requirements for adequate route choice sets for prediction of route and link flows in multi-modal transport networks.

The feasibility of the developed multi-modal choice set generation (MM-CSG) approach has been demonstrated by applying the method to the multi-modal network of the Rotterdam-Dordrecht Region in The Netherlands. In order to prove the effectiveness of the MM-CSG method, the generated choice sets were compared with observed chosen routes and reported choice sets in this multi-modal network. In order to achieve adequate choice sets that satisfy most of the defined requirements, a filtering process after the generation is needed as in the uni-modal case. The MM-CSG approach produces an exhaustive base route set (master set) that needs to be freed from superfluous and non-efficient, routes in order to arrive at choice sets that efficiently suit a particular purpose, in our case the prediction purpose. To that end, a variety of selection constraints may be additionally applied.

The structure of the chapter is as follows. The specific aspects of the choice set generation in multi-modal networks and the requirements for adequate multi-modal choice sets are presented in Section 7.2 while Section 7.3 presents an overview of existing route choice set generation algorithms for multi-modal networks. The doubly stochastic approach to MM-CSG is introduced in Section 7.4 after which the demonstration case applied to the Rotterdam-Dordrecht corridor of the generation approach with the results of its performance are presented in Section 7.5. The calibration of the MM-CSG model is discussed in Section 7.6, while the impacts of stochasticity on the quality of the choice set are analysed in Section 7.7 and the results of the application of a filtering process are reported in Section 7.8. The main conclusions of the chapter are drawn in Section 7.9.

7.2 Choice set generation in multi-modal networks

7.2.1 Specific aspects for CSG in multi-modal networks

In Chapter 6, choice set generation has been dealt with directed at uni-modal networks. In this Chapter 7, focusing specifically on CSG for multi-modal networks, we will elaborate on those aspects that are specific for multi-modal networks on top of those that hold for uni-modal networks.

First-of-all, in multi-modal networks we have a combination of several *continuous-type* transport service systems (all private modes such as walk, bicycle, car) and *discontinnous-type* transport service systems (most public transport modes such as bus, tram, metro, train, etc). Typical for the discontinuous-type systems is, for example, the need to wait for departure of the services and the need to transfer between various services and lines. In order to make these various service types compatible within a single route from origin to destination, special nodes and links have to be added to the basic networks in such a way that a consistent supernetwork results that facilitates classical shortest route search algorithms as the chosen base for CSG. The specification of a multi-modal supernetwork has been elaborated in Chapter 3.

A second typical aspect of multi-modal networks is the *transfer* between modes. Since we know that transferring is a crucial factor in travel choice behaviour (see e.g. Hoogendoorn-Lanser (2005)), the treatment of transfers in route generation needs special attention. This relates to the description of the transfer process in the multi-modal network by using special nodes and link types enabling a correct path search in the network (see also Chapter 3). This relates as well to the specification of the transfer process in the generation function with which the attractiveness of generated paths is determined. Transfers and related times and distances need special disutility parameters in the generation function.

A third typical aspect is the *sequence of modes* within a multi-modal trip. From logic and from observations we can deduce that these sequences of modes within a multimodal trip obey certain regularities and conditions. For example, high speed modes are used for the line haul part of the trip whereas low speed modes are used in the networks near to the origin and destination, while home-based modes such as car and bicycle will not be used for legs in-between public transport legs. Also private mode access distances to PT stops as well as private mode egress distances from PT stops to the destination are characterized by maximum acceptable values. For an in-depth account of the behavioural peculiarities in multi-modal trip-making, see Hoogendoorn-Lanser (2005). Another important aspect of choice sets for multi-modal trips is the variation in transport mode compositions. In order to generate routes in a multi-modal context it is very important to consider this aspect and try to achieve this variation.

A number of other aspects appear typical in multi-modal networks: these networks are very large (in terms of numbers of nodes and links) and exhibit very many potential route alternatives. Because of the discontinuous PT services, the same spatial route with the same multi-modal composition may have nearly identical alternatives at various different departure times of these services. Also the route overlap problem (see Section 5.4.2) appears highly different from the overlap problem in uni-modal continuous networks (see Hoogendoorn-Lanser, S. and Van Nes, R. and Bovy, P.H.L. (2005)).

The consequences of these typical multi-modal aspects for a prediction-oriented choice set generation approach for multi-modal trips are the following (compared with Figure 6.9 Classification scheme of generation methods in Chapter 6):

- the need for a specific network description called supernetwork, as presented in Section 3.7;
- the need for specific constraints in checking the reasonabless of routes and adequacy of choice sets (see this chapter);
- the need for a specific generation function (or search criterion) to be used in the choice set generation algorithm (see this chapter);
- the need for specific termination criteria of the generation algorithm (see this chapter).

7.2.2 Specific requirements for an adequate choice set in multi-modal networks

In this section we will discuss the requirements for a reasonable multi-modal route, an adequate choice set, and an appropriate choice set generation process, specifically in the case of multi-modal networks. Following the typical multi-modal aspects discussed in Section 7.2.1, we will focus on those criteria that are typical for multi-modal cases in addition to the criteria for uni-modal networks already elaborated in Section 6.3.

We will first look at such additional properties of single routes by specifying requirements for reasonable multi-modal routes to be generated. On that basis we then look at the composition of reasonable sets of routes from an individual traveller's perspective, after which we consider the adequacy of route sets for groups of similar travellers. On this basis, we will define criteria for adequate multi-modal choice sets for prediction purposes, and define quality criteria for the resulting choice set and for the generation process given a multi-modal network.

Requirements for a reasonable (multi-modal) route

Hoogendoorn-Lanser (2005) offers a systematic account of conditions that reasonable routes should satisfy from the traveller's perspective to become a member of a choice set. These conditions can be grouped under the following headings:

- logical conditions;
- feasibility conditions;
- behavioural conditions;
- perceptual conditions.

Logical conditions concern the topological form of routes in the multi-modal time-space domain. The term *logical* expresses that travellers will not undertake unnecessary or superfluous actions such as travelling in cycles or loops, and cannot take impossible actions (travelling backwards in time, transferring to an earlier train). In public transport networks, loops may happen without being illogical from a behavioural point of view, namely if different services are used to reach the destination. For example, the express train may pass through the envisaged alighting stop of a traveller while at the next stop the traveller takes a local train to his alighting station that he already had passed through before. Illogical routes will not be selected in choice sets, apart from the fact that most route search algorithms cannot generate illogical routes.

Feasibility conditions concern the suitability of a (multi-modal) route alternative in terms of time, space, vehicle availability, and physical (dis-)abilities. Feasibility in *time* refers to time constraints at origin and / or destination addresses and to constraints with respect to transferring between services. These types of constraints are especially relevant in a multi-modal context with time-tabled services. For example, in a multi-modal context, maximum transfer time constraints can be used to establish which runs of consecutive transport services can be taken. Feasibility in *space* refers to availability of transport modes not only at origin and destination, but also at transfer points which is especially relevant in multi-modal trip making. Maximum walking and cycling distance constraints determine the candidate boarding stops within reasonable access distances and candidate alighting stops within reasonable egress distances, thus, which consecutive transport modes are feasible to be used. *Physical disabilities* might restrict the use of private transport modes (for example, walking and cycling) and public transport services. To account for physical disabilities during the choice set generation, routes with certain types of transfers can be excluded from the choice set and walking speeds can be adapted accordingly.

Behavioural conditions refer to the individual traveller's preferences with respect to trip attributes, such as combinations of transport modes, PT waiting times, transfer times walking times, costs and number of transfers, etc. Empirical research (Van Nes (2002) and Hoogendoorn-Lanser (2005)) has shown that travellers have individual preferences with respect to such trip attributes, have bounds on values of trip attributes, do not distinguish highly similar alternatives, and tend to travel in the direction of the final destination. In the context of multi-modal trip making, behavioural conditions may be formulated specifically in relation to complete trips, access and egress parts, train trip part, or transfers. Table 7.1 and Table 7.2 show such conditions and to which trip part they apply.

A basic premise behind these conditions is that travellers always try to use the subjectively best route: they optimise their behaviour according to a personal utility function. Travellers are not accepting excessive detours or multiple switches between link types during the trip. Moreover, route alternatives that largely *overlap* with others will not be identified as a distinct route by the traveller, reason why such routes might be excluded from the choice set. Another behavioural condition relates to the hierarchical set-up of routes being preferably followed by travellers (see Section 6.3.1).

Perceptual conditions relate to the knowledge on the part of the traveller of available alternatives. Because of the complexity of public transport systems, the cognition of multi-modal alternatives maybe assumed limited. Of course, these are difficult to specify, especially in forecasting cases.

In the multi-modal case, because of multiple modes and public transport services, we need to introduce a hierarchy of six movements levels, as follows from high to low level:

1. tour from base (mostly home) to carry out one or more activities and then back to base;
2. trips within a tour;
3. possible (multi-modal) routes to perform the trips;
4. division of trips (and thus also routes) into three basic trip parts, namely, main trip part and access-egress trip parts;
5. legs within trip parts that are therefore separated by between-leg transfers also called inter-modal transfers;
6. links, the smallest parts within a route.

All these concepts have been introduced and defined in Chapter 2. Table 2.1, which summarizes these definitions, will also be used in this chapter.

Table 7.1: Behavioural conditions for choice set composition in multi-modal trip making (adapted from Hoogendoorn-Lanser, 2005).

Trip parts	Conditions
Complete trip and separate trip parts	Travellers accept only a limited detour; alternatives with larger detours are disregarded in the choice process. A detour may be measured in distance, time, or number of transfers.
	Travellers have different preferences with respect to the various transport services within a trip, which are, among other things, based on image and seating comfort.
	Travellers value the distinct time-elements (in-vehicle time, waiting time, walking time and transfer time) and cost-elements (parking costs, PT-fares and fuel costs) differently.
Transfers	Travellers only accept maximum total transfer times, maximum transfer-waiting times, and maximum transfer-walking times.
	Travellers prefer transferring in higher-order stops and railway stations. Higher-order railway stations and stops offer higher frequencies than lower-order railway stations and stops.
	Travellers do not make unnecessary transfers. For example, transfer to a lower-order or equal train service is considered unnecessary if the current service takes the traveller to the preferred railway station.
Access / egress - UPT modes	Travellers choose the closest access point to a specific UPT line (based on the fact that in-vehicle time is considered less negative than walking time).
	Travellers do not use UPT for short distances. The walking time to and the waiting time at the UPT stop take more time than directly walking towards the destination.
	Travellers do not board (at the beginning of the trip) at stops that are further away from the destination than the destination itself.
	Travellers walk to UPT access and from UPT egress points.
	Travellers have different maximum acceptable access and egress distances to stops.

Table 7.2: Behavioural conditions for choice set composition in multi-modal trip making for access/egress modes (adapted from Hoogendoorn-Lanser, 2005).

Trip parts	Conditions
Access / egress - UPT modes	Travellers choose the closest access point to a specific UPT line (based on the fact that in-vehicle time is considered less negative than walking time).
	Travellers do not use UPT for short distances. The walking time to and the waiting time at the UPT stop take more time than directly walking towards the destination.
	Travellers do not board (at the beginning of the trip) at stops that are further away from the destination than the destination itself.
	Travellers walk to UPT access and from UPT egress points.
	Travellers have different maximum acceptable access and egress distances to stops.
Access / egress - private modes	Travellers do not use a bike for short distances. Retrieving and parking a bicycle takes more time than walking directly to the railway station.
	Travellers have different minimum and maximum acceptable walking, cycling and car distances at home-ends and activity-ends of trips.
	Travellers use only a single private transport mode instead of combinations of private modes) to go to railway stations.
Access / egress - general	Travellers use UPT (metro, tram and bus) or private transport modes (walking, cycling and car) to go to railway stations. Combinations of UPT and bicycle / car are not used to go to railway stations (Van Nes, 2002).

A tour consists of a series of trips departing from home to carry out one or more activities and coming back home. A tour may be divided into a series of trips and a trip may be divided into a series of alternative (multi-modal) routes.

The composition of a multi-modal trip might be quite complex as it may consist of a series of several different travel modes connected by walking legs. In general in a multi-modal network different network levels might be distinguished, from higher to lower levels. The higher levels are characterized by higher speed, lower network density and lower access density. Higher-level parts are usually used for the main part of the trip while for access/egress lower network levels are used. Given that, in a (multi-modal) route we may distinguish three parts: namely, the main part and the access and egress parts, since usually a multi-modal route consists of a part performed over the largest distance and other one or two parts at the beginning and/or at the end of the trip that are made over shorter distances.

Each of these trip parts may consist of several uni-modal legs, and since in our multi-modal formulation walking is always considered a travel mode and therefore a leg (see Chapter 2), it can be always distinguished three parts in a multi-modal trip. The *main trip part* is that part performed over the largest distance with the highest possible speeds compared to the access and egress trip parts.

A Leg is the part of a multi-modal trip for which a single mode is used without intermediate transfers. In Chapter 2, intra-modal and inter-modal transfers are introduced; from a behavioural point of view, the two types of transfers (inter-modal and intra-modal) are different, the intra-modal transfers can be less restrictive since the transfer is within the same transport system with similar frequencies and the same ticket; whereas the inter-modal transfers involve more difficulties for transferring between different transport systems, with different frequencies and tickets. However, it is very complex to take into account both types of transfer in modelling travel choices; therefore, in this chapter, the word 'transfer' is used for inter-modal as well as intra-modal transfers for sake of simplicity. Furthermore, we assume that a leg has the same service network, same frequency, and same ticket and therefore there are no transfers within a leg. Because of the definition of legs, being uni-modal by definition, leg and mode mostly are synonymous terms. Each leg has its own mode, but a particular mode may serve multiple legs within a trip or within trip parts.

Finally, we have links being the smallest parts of the network (for uni-modal private networks) and services (for PT networks or parallel lines). Since public modes use fixed routes, it is sufficient to define links as being direct connections between nodes or between PT stops.

Based on the previous arguments (see also Section 6.3), let us discuss our definition of *reasonable*[1] route in multi-modal networks.

Definition 12 *A single multi-modal route is defined* ***reasonable*** *if the following criteria are fulfilled:*

Acyclic criterion *A reasonable multi-modal route does not contain loops.*

Detour criterion *A reasonable multi-modal route does not exhibit a detour from the shortest possible connection in terms of one or more measures such as distance or time between origin and destination larger than a maximum threshold* α *(e.g. 100%).*

Hierarchic quality *A reasonable multi-modal route is constituted by a systematic sequence of functional levels in the multi-modal network such as transport services (including private modes). For example, an urban bus service may be used for access to a train station, an IC train service may be used to perform the main part of the trip and finally a tram service may be used as egress mode.*

Multi-modal feasibility *A reasonable route is uni-modal, or, if it is multi-modal, the number of different modes used in the route is limited and the availability of modes meets given constraints, for example it is unusual to have a car available at the activity-end railway station.*

The first three criteria are similar to the ones defined in Definition 9 of Chapter 6. If the first two criteria are valid for the whole multi-modal route then they are also valid for smaller parts of a multi-modal route such as trip parts and legs. The third (Hierarchic) criterion is applicable to the whole multi-modal route. Finally, these criteria are not mutually exclusive. To a certain extent they overlap and may result in identical generated routes.

An extra criterion is added for the multi-modal case at the single-route level namely the feasibility concerning the multi-modal composition of the route.

---

[1] Note that the term *reasonable* used in this thesis differs from the one defined by Dial, R.B. (1971)

Multi-modal feasibility criterion In order to illustrate the feasibility in the multi-modal route composition the following definitions are needed. As stated above, in a multi-modal route three parts of the route might be distinguished:

- Main part: part of the multi-modal route performed over the largest distance with the highest possible speeds and low interchange density.
- Access part: part of multi-modal trip connecting the trip's origin to the start node of the main trip part (often a transfer point).
- Egress part: part of multi-modal trip connecting the end node of main trip part (often a transfer point) to the trip's destination.

In order to introduce generic definitions we also propose the following classification of functional mode types and vehicular modes. We may distinguish main modes, home-based modes and activity-based modes as functional mode types.

Main modes are characterized by their ability to be suited for long distances and long travel times because of their higher vehicle speeds, higher comfort, etc. Typical examples of main modes are airplane, train, and car; sometimes also bus instead of train. The service type 'train' can be distinguished in Intercity (IC), Express and Local services. IC and Express services are usually used as main modes, whereas Local train can be used as access mode; the combination of Local and IC/Express services can be considered as main mode, since they are part of the same system with same ticket.

Home-based modes are characterized by their ability to serve access to and from a base which mostly is the home. To that end, home-based modes often are owned by or easily accessible by the traveller at his base. Typical home-based modes are walking, bicycle, car, metro, bus and tram.

Activity-based modes are characterized by their suitability for short distances and mostly non-ownership by the traveller. Typical examples are walking, bus, tram, metro and taxi. In some countries often other vehicular modes are available for the distinguished trip parts such as a shared taxi (sjarut) for the main trip part in Israel.

The main trip part is in-between two end nodes that connect the access and egress trip parts to the trip origin and trip destination respectively. The main part of the trip is the most difficult part to be identified. It can be seen from a functional perspective, such as the long distance of a trip belongs to the main part; but if two parts of a trip have the same distance, which is the main part? If we consider spatial aspects, the main modes should touch at least two different regions, but still there are exceptions. In general it is very difficult to define and identify in a generic way the main part of the trip, but in particular, Van Nes (2002) observes that looking at multi-modal trips, train is the most important transport mode, if the main mode is defined as the mode which is used to traverse the largest distance (80% of all train trips are multi-modal trips); the shares of bus and car trips that are multi-modal trips are much smaller (20.4% respectively 2.1%).

Table 7.3: Currently available modes assigned to different network levels.

Transport Modes	International	National	Interregional	Regional	Local
Airplane	X	X			
Intercity train	X	X	X		
Local train				X	X
Interregional Bus			X		
Car	X	X	X	X	X
Taxi			X	X	X
Regional Bus					X
MRT/Metro					X
UPT Bus					X
UPT Tram					X
Bicycle					X
Walking					X

With each of the three distinguished trip parts (access, main, egress) we may associate corresponding vehicular travel modes. In principle, all modes may fulfil the role of an access, an egress, or a main mode. For example, walking is the main mode in a walking-only uni-modal trip, while the bike maybe the main mode in a walk-bike-walk uni-modal trip. In the practice of interregional trips, typical main modes in multi-modal trips are car, train or regional bus, whereas typical access and egress modes are walking, bike, bus, tram, metro, etc. In principle, car may also play a role as an access-egress mode, such as in car-train-walk trips or walk-train-taxi trips.

Because of their properties, some vehicular modes may well serve for more than one trip part type, depending on the spatial conditions. By introducing the network levels present in a transportation network, such as international, national, interregional, regional and local levels, Table 7.3 gives an example of the assignment of currently available modes to those network levels.

As illustrated in Table 7.3, each vehicular travel mode may be used at different network levels. Depending on the type of trip, if it is international, or regional, or national or interregional, different travel modes may be associated with each of the three distinguished trip parts (access, main, egress). In case of an international trip, the same travel mode (train or airplane) might be used as access and as main mode.

Since our analysis focuses on interregional trips, we can state that some travel modes are more likely to be used for the main trip part, and others are less likely to be used as access and egress mode. For example, IC train can be used as main mode, but not as access mode for an interregional trip.

The following *generic constraints* specify the possible combinations of the three trip parts in the specific case of interregional trips:

- a main mode cannot be an access/egress mode in the same trip, and vice versa;
- home-bound modes cannot be used as main modes except for car;
- activity-bound modes cannot be used as main modes except for car;
- home-bound modes cannot be an activity bound mode in the same trip except for walking, UPT, taxi, and bike;
- home-bound modes such as car and bicycle cannot be used for legs in-between public transport legs.

In addition we may have the following *specific constraints* specifying limitations within each of the three trip parts:

- high speed modes are used for the main part of the trip;
- low speed modes are used in the networks near the origin and destination;
- walking and private mode access distances from the origin to PT stops as well as walking and private mode egress distances from PT stops to the destination are characterized by maximum acceptable values;
- a trip part consisting of multi-leg should have a limit number of transfers;

Because of these constraints and the possible functional mode types of the vehicular modes, in a multi-modal case, many multi-modal routes are possible but not **all** multi-modal routes are feasible. For this reason, we need to define when a multi-modal route is feasible. Before introducing the definition we need the following notation:

- tp indicates the trip part of a route, such as the main part, the access and the egress parts;
- NT_{max}^{tp} indicates the maximum number of allowed transfers within trip part tp;
- r_{OD} indicates the complete route r from origin O to destination D;
- NT_{max}^{r} indicates the maximum number of allowed transfers within route r.

Definition 13 *A multi-modal route is defined as follows:*
$\mathbf{HM}(-\mathbf{HM})-\mathbf{MM}(-\mathbf{MM})-\mathbf{AM}(-\mathbf{AM})$ *where:*

- *Three parts are distinguished: the main part (MM=Main Modes), the access (HM= Home-based Modes) and egress parts (AM=Activity-based Modes);*
- *The expressions (-**HM**), (-**MM**) and (-**AM**) indicate the repetition of the same expression n times with* $n \in \{0, NT_{max}^{tp}\}$ *where* NT *is the maximum number of transfers within a trip part* (tp)*, such as the main, access and egress parts;*
- ***MM** are the main modes of the trip consisting of one or multiple legs;*
- ***HM** are the home-based modes of the trip consisting of one or multiple legs;*
- ***AM** are the activity-based modes of the trip consisting of one or multiple legs.*

A multi-modal route is feasible if the following constraints (previously introduced) are satisfied:

- generic constraints specifying the possible combinations of these three trip parts;
- specific constraints within each trip part or leg;
- specific constraints related to specific vehicular modes;

and the number of transfers within the route r is limited by the maximum number of allowed transfers:
$NT(r_{OD}) \leq NT_{max}^{r}$
For instance, depending on the cases under consideration, such as the case of considering interregional trips, the main modes, home-based modes, and activity-based modes might be defined based on Table 7.3 as follows:

- **MM**={IC train, Interregional bus, Car, MTR/Metro}.
- **HM** = {Car, Local train, Taxi, MRT/Metro, UPT Bus, UPT Tram, Bicycle, Walking}.
- **AM** = {Taxi, Local train, MRT/Metro, UPT Bus, UPT Tram, Bicycle, Walking}.

Please note that according to this definition, the following trips are feasible multi-modal trips:

- Trip where car is used to access a metro station;
- Trip where car is used to access a railway station;
- Trip where a bicycle is used for egress from a railway station to the destination;
- Trip where a city bus is used for egress from a railway station to destination;
- Trip where a local train is used to access an IC railway station; in this case local train and IC train are considered main modes since they belong to the same transport and fare system.

According to this definition, the following two trips are also multi-modal trips:

- Trip using two urban bus lines;
- Trip where two IC train services are used and walking is used as a feeder mode.

Depending on the individual's vehicle availability, car or bike might be excluded from the sets **HM**, **MM**, and **AM**. Examples of multi-modal routes satisfying the multi-modal composition criteria are the following, where the main part of the route is denoted within square brackets [ ]:

1. Car-[Metro]-Walk
2. Car-[Train]-Tram
3. Walk-Bus-[Train]-Metro-Tram
4. Bike-[Local Train-IC Train]-Walk

In the first case, car is the only mode used at the home side of the trip and in the main part of the trip; metro and walk are used as egress modes. In the fourth case, the local train could be considered as access mode, but in our definitions the local and IC trains that are considered the main modes, since they belong to the same transport services and fare systems.

In the previous examples the number of transfers among travel modes is limited at maximum 3 transfers (third alternative: walk-bus is not considered a transfer). The travel leg sequence is logical for all given alternatives, whereas routes such as Car-Bike-Tram-Bike-Train-Bus-Bike-Tram are not reasonable, since some constraints of Definition 13 are not satisfied.

Van Nes (2002) analysed National Travel Survey Data[2] and showed that of multi-modal trips consist of two different modes, while of them is comprised of three or more different modes. Train is often the main mode (59%) in multi-modal trips, followed by bus (14%).

In the next subsection, the defined requirements for reasonable single multi-modal routes will be extended towards the definition of an adequate set of multi-modal routes for a single traveller.

Requirements for adequate multi-modal choice sets at individual level

Based on the reasonabless criteria for single multi-modal routes, we now look at adequacy of generated complete choice sets from the perspective of a single traveller. Since we aim at specifying multi-modal choice sets generated for prediction purposes and thus for a group of similar travellers, this subsection is an intermediate step.

---

[2] The Dutch National Travel Survey is conducted by CBS, which collects travel data for more than 70.000 households annually, resulting in 600.000 trip records.

As discussed in Chapter 4 (Section 4.3), from the viewpoint of a single traveller the following characteristics of choice sets are important:

- the spatial structure of the set in terms of the degree of mutual overlap or, oppositely, their spatial difference;
- the composition of the set in terms of comparability of the routes in the set or oppositely, their variety;
- the size of the set in terms of number of alternatives;
- the composition of the set in terms of multi-modal variety.

From these arguments the following definition is derived (for details see Section 4.3).

Definition 14 *A multi-modal choice set generated for an individual traveller is* ***adequate*** *if all multi-modal routes of the choice set are reasonable according to Definition 12 and 13, and additionally satisfy the following criteria:*

Overlap criterion *Any two routes of the generated choice set should have a mutual overlap (in terms of number of links, distance, or time) less than* Δ *percent with respect to the shorter one of the two routes.*

Comparability criterion *Any two routes of the choice set should be comparable in travel disutility (time, distance, or cost) within a given threshold of* θ_1 *percent with respect to the shorter one of the two routes.*

Detour-max criterion *The non-common parts of two partly overlapping routes in the choice set should have a maximum detour (in disutility terms) not larger than a given maximum percentage* ω_{max} *of the minimum of the two parts.*

Detour-min criterion *Any two partly overlapping routes in the choice set should have a minimum detour (in disutility terms) between the two routes not smaller than a given minimum percentage* ω_{min} *of the minimum total route length, (or time or cost).*

Choice set size *The choice set should contain a limited number of alternatives, say less than S (e.g. 10).*

Multi-modal variety *The generated set should consist of different uni-modal routes and different multi-modal routes, depending on the type of multi-modal network under consideration.*

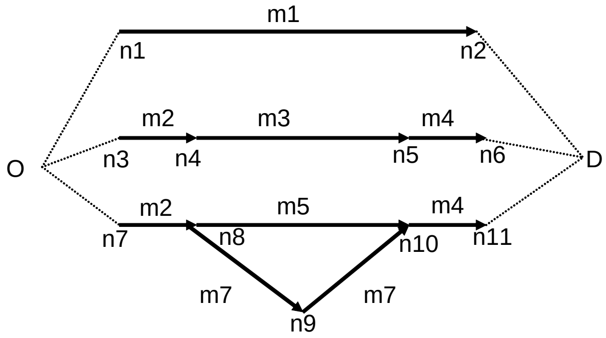


Figure 7.1: Examples of multi-modal routes in a multi-modal network (n nodes, m modes).

Instead of single multi-modal routes considered in Definitions 12 and 13, now in Definition14 the mutual relations between multi-modal routes within the choice set are at stake. In Definition 14 we first consider the criteria referring to the comparison of full routes (*overlap* and *comparability criteria*); secondly the criteria regarding the comparison of parts of routes (*detour-max* and *detour-min criteria*) and finally criteria relevant for the entire choice set concerning the *choice set size* and the *multi-modal variety* within the choice set.

While the overlap criterion refers to the common route part of two multi-modal routes, the detour-max and detour-min criteria specifically refer to the non-common route parts of two multi-modal routes.

The first four criteria are similar to the ones defined in Definition 9 of Chapter 6. If the first two criteria are valid for the whole multi-modal route then they are also valid for smaller parts of a multi-modal route such as trip parts and legs. To a certain extent, the adequacy criteria are overlapping in the sense that different criteria may lead to rejection of the same route.

Overlap criterion at route level In order to satisfy the overlap criterion, any two routes of the generated choice set should have a mutual overlap (in terms of number of links, distance, or time) less than Δ percent with respect to the shorter one of the two routes (see Definition 14). This definition is similar to the one for the uni-modal case; however, in the multi-modal case the application of this criterion is much more complex than in the uni-modal case.

As defined in Chapter 6, the overlap criterion refers to full route-pair comparison. We assume that two routes are considered equal (or similar) only if all their constituting links are equal. Therefore, in order to check the similarity of two routes a detailed check at link level is needed. Let assume the multi-modal network representation proposed and described in Chapter 3, it is simple to show that the overlap criterion may be easily applied by considering the link representations of the two routes and by applying the overlap criterion defined for the uni-modal case (see Definition 10 in Chapter 6).

For example, the multi-modal network shown in Figure 7.1 contains four multi-modal routes with different modes (m) and different nodes (n) that are not overlapping. In our network representation a different link corresponds to each mode; therefore, two routes may overlap only if exactly the same modes are used along the routes with exactly the same boarding and alighting nodes. For example, in the case shown in Figure 7.1 if mode m_3 is the same mode as m_5 (e.g. IC train) and nodes $n_3 = n_7$ and nodes $n_4 = n_8$ and nodes $n_5 = n_{10}$ and nodes $n_6 = n_{11}$ are the same multi-modal transfer nodes then the two routes are also overlapping at link level, otherwise they can be considered different.

Comparability criterion at route level In the multi-modal case, we assume that the comparability criterion can be applied only at route level. Routes are comparable only at route level, since in the multi-modal case, several main travel modes can be used and two routes are comparable only if similar main transport modes are used; for example, routes using train as main mode are comparable with similar routes using train, and vice versa routes using car as main mode are comparable with similar routes using car.

In order to specify the comparability criterion in the multi-modal case, let us introduce the following notation. Let $L_r = \{l_{m1}, l_{m2}, \ldots, l_{mi}\}$ denote the sequence of links of route r and $L_p = \{l_{m1}, l_{m2}, \ldots, l_{mj}\}$ denote the sequence of links of route p, and index m denotes the travel mode m associate with link i or j. Let θ_1 be between 0 and 1. Please note that in this case routes r and p belong to the subset of all routes having the same main mode, such as car, or train.

Given the above definitions and letting $F(l_m)$ be the function that maps link l_m of mode m to its length, time or generalized cost, and let $F(l_m)$ be also equal to 0, 1 representing that link l_m of mode m does not or does belong to route L_r or L_p.

In the case of multi-modal networks, a route r satisfies the comparability criterion with respect to route p in terms of distance, time or generalized cost if the following holds:

$$\max\left(\sum_{m\in M}\sum_{h\in L_r}F(l_{mh}),\ \sum_{m\in M}\sum_{k\in L_p}F(l_{mk})\right) \leq (1+\theta_1)\cdot\min\left(\sum_{m\in M}\sum_{h\in L_r}F(l_{mh}),\ \sum_{m\in M}\sum_{k\in L_p}F(l_{mk})\right) \tag{7.1}$$

where r and p are routes with the same main mode **MM**. Modes belonging to set **MM** refer to modes listed in Table 7.3. Note that access and egress modes might be different for both trips.

Detour-max and detour-min at link level In the multi-modal case, we assume that the detour-max and detour-min criteria can be applied only at link level. If two routes are compared at leg level (see section over the overlap criterion) while being completely overlapping at leg level, a more detailed check is needed at link level.

All criteria defined for the uni-modal case in Chapter 6, such as the overlap, comparability, detour-max, and detour-min criteria, can be applied also to each leg component (the uni-modal part) of the multi-modal route. For this reason it makes sense to apply the detour criteria at link level.

These criteria can be applied to check if large or small variations of some parts of the multi-modal routes are present in the route and if they can be removed from the choice set.

For example, in the case of the detour-min criterion and in the case of having small variations within the same transport mode, the longest route can be excluded.
Figure 7.2 shows an example in which route r does not satisfy the detour-min constraint relative to route p. If we consider the sequence of legs of the two routes, they are as follows:

Route p = Walk-Tram-IC Train-Metro-Walk

Route r = Walk-Tram-IC Train-Metro-Walk

The two routes are apparently identical, however if we check the two routes at link level, they are as follows:

Route p = Walk-Tram-IC Train-Metro-Walk-Walk

Route r = Walk-Walk-Tram-IC Train-Metro-Walk-Walk-Walk

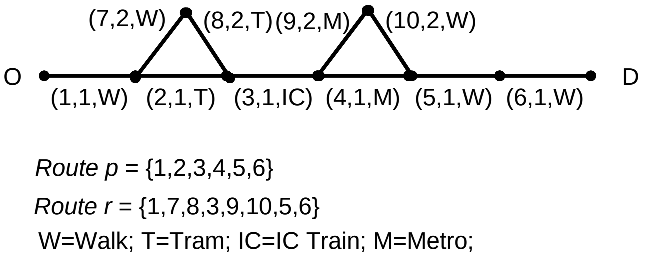


Figure 7.2: Route r does not satisfy the detour-min constraint relative to route p in the multi-modal case at link level. ((l, c, m) represents the link index and the link cost and the transport mode associated with each link).

As we can see from Figure 7.2 there are small variations in the access and egress parts of the trip, therefore the detour-min criterion is clearly not satisfied. In this case the longest route can be excluded from the generated choice set.

Choice set size criterion The requirement about the *choice set size* is derived from Chapter 4 and Chapter 6 showing that the size of the objective choice set is proportional to the network density, whereas the size of the subjective choice set depends on the limited ability of travellers in considering many alternatives. Thus, in modelling route choice sets for transport networks, large choice sets may be expected to be generated, especially in the case of multi-modal networks, whereas only small choice sets maybe expected to be considered by individual travellers. In fact, since the number of routes that might be known or considered or used by the travellers is limited, a choice set at individual level is expected to have also a limited number of route alternatives, even very few alternatives; as it is observed in reality (Bovy, P.H.L. and Stern, E., 1990).

Multi-modal variety Finally, the choice set at individual level consists of several uni-modal and multi-modal routes. An individual might have different preferences for using transport modes; therefore, a variety of routes (uni-modal and multi-modal) is expected in the generated choice set.

However, the multi-modal variety criterion makes sense at individual level only if the multi-modal routes are consistent with the individual preference and with the application of the route generation process in the multi-modal network. At individual level, the travel preferences and the vehicle availability of the individual traveler must be considered for the satisfaction of this criterion. This criterion plays a more important role at group level as we will discuss in the next section.

To a certain extent, the criteria developed and discussed above are overlapping in the sense that different criteria may lead to rejection of the same route.

Table 7.4 summarizes the properties of the criteria introduced in this section. Four types of comparison for route elimination are distinguished:

- Comparison at the level of a single route on its own.
- Pair-wise comparison of full routes, at route level and leg level.
- Pair-wise comparison of parts of routes, at link level.
- Comparison at level of the full choice set.

Some criteria such as Acyclic, Detour, Hierarchic, and multi-modal (feasibility from Definition 12 are applied at the level of a single route. The overlap and comparability criteria are applicable to full route pairs at route and link levels; whereas the detour-max and detour-min criteria are applicable only to parts of route pairs, at link level. Choice set size and multi-modal variety criteria are applicable at the level of the full choice set.

Table 7.4: Criteria classified with respect to four types of comparison in the multi-modal case.

Criterion	*Single route*	*Full route pairs* (route level)	*Parts of route pairs* (link level)	*Full choice set*
Acyclic	X			
Detour	X			
Hierarchic	X			
Multi-modal feasibility	X			
Overlap		X	X	
Comparability		X	X	
Detour-max			X	
Detour-min			X	
Choice set size				X
Multi-modal variety				X

Requirements for an adequate multi-modal choice set at group level

The requirements for an adequate multi-modal choice set for a group of similar trips directly relate to our focus on the prediction application of choice sets. Moreover, when a group of individuals is considered, which are likely to be based on different preferences, aggregation of individual choice sets has to be taken into account. Therefore, all criteria introduced in the previous subsections need to be discussed again at aggregate level.

From the viewpoint of prediction of route choice, the following characteristics of choice sets are important:

Overlap and comparability criteria In a prediction purpose application, the analyst is much more interested in knowing the flows of the main transportation modes than in the access and egress modes. Therefore, an adequate multi-modal choice set should contain the routes consisting of these main transport services. Routes that overlap and are different only for access and egress modes might be rejected because only variations of the main modes are considered relevant for the prediction application. Routes belonging to the prediction choice set should be comparable, but at group level more variation in route types and route costs is allowed, because individuals belonging to the same group may have different preferences.

Detour-max and detour-min criteria Detour-max and detour-min requirements introduced in Definition 14 derive also from the prediction purposes for which only variations of the main services are relevant which might be included in the choice set, whereas overlapping or non-overlapping of access ande egress modes are not relevant for the prediction choice set.

Choice set size As stated already in Section 6.2 for link flow prediction, generated choice sets need to include all attractive routes but may miss routes of less attractiveness. Given this, predicted choice sets should be sufficiently large, possibly even including 'wrong' (non-used) routes. Consequently, we can state that a prediction choice set should likely consist of all reasonable routes with significant probability of being chosen. Moreover, given the complexity of multi-modal routes, the choice set is likely to be larger than in the uni-modal case.

Spatial variability criterion The generated choice set should contain as different route types (with respect to road type composition) as possible. This criterion, in addition to the detour-max and detour-min criteria, emphasizes spatial variability among routes, referring to variability with respect to road type and service lines composition and other observable properties of roads, train lines and UPT lines. This criterion is especially relevant at the group level (see below).

Preferential variability criterion This criterion takes into account the variability in preferences within the group of travellers. The choice set should contain as many as possible routes representing the taste variation of the group of travellers.

Multi-modal variety The generated choice set should contain sufficient different route types with respect to mode type composition. This criterion, in contrast with the spatial variability criterion, emphasizes modal variability among routes, referring to variability with respect to mode type composition within and between routes. This criterion is especially relevant at the group level (see below).

From these arguments the following definition of an adequate multi-modal route choice set generated for prediction purposes for a group of travellers is derived.

Definition 15 *A route choice set generated for a group of travellers in a multi-modal network is defined* ***adequate*** *if all routes of this set are reasonable according to Definition 12 and 13, the set is adequate according to Definition 14, with the following remarks:*

Overlap criterion *See Definition 14*

Comparability criterion *Two routes belonging to the choice set should be comparable in travel disutility (time, distance, or cost) within a given threshold of* θ_2 *percent.*

Detour-max and detour-min criteria *See Definition 14*

Choice set size *The choice set should contain a limited number of alternatives but should include all attractive routes having a high probability of being chosen.*

and fulfils the following additional criteria:

Spatial variability *Routes of the choice set should be spatially different (dissimilar) with respect to the links used.*

Preferential variability *Routes of the choice set should represent the taste variation of the group of travellers.*

Multi-modal variety *The choice set should contain sufficient different route types with respect to mode type composition within and between routes. The generated set should contain several uni-modal and multi-modal routes, depending on the type of multi-modal network under consideration.*

The *comparability* criterion has been adapted at the group level since routes in the joint choice set are less comparable than routes in the individual choice set. In order to take this into account different parameters θ_s might be used. At the individual level a θ_1 parameter may be used in the formula whereas at group level a θ_2 parameter may be applied, with $\theta_1 < \theta_2$.

This implies that at individual level routes are more comparable in terms of distance, time or generalised cost, whereas at group level routes should be also comparable but in a less restrictive way, more variation of route types and route costs is allowed at group level, because individuals belonging to the same group may have difference preferences.

The *choice set size* requirement follows from Chapter 4 and Chapter 6. The size limit should allow the consolidation of different choice sets from various travellers at group level. In addition, the analyst is interested in selecting routes that vary as much as possible and in excluding routes that are similar and largely overlapping.

The *spatial variability* requirement follows from the observation that choice sets of different travellers in a group may strongly differ in composition due to the fact that the group members have each their own knowledge, preferences, and perceptions of the network. In addition, they travel between (slightly) different OD pairs with different OD locations, so that different types of routes may be chosen even if knowledge, preferences, and perceptions are equal. The main implication of this line of reasoning is that the prediction choice sets should preferably consist of route alternatives that are spatially different, and different with respect to cost and time or other attributes. Because of this a broad variety of options should be available in the generated choice set as far the transportation network allows. The spatial variability is already included in the overlap, detour-max and detour-min criteria. By setting the Δ, ω_{max}, ω_{min} parameters of the overlap, detour-max, and detour-min criteria respectively in a proper way, routes generated with those parameters values might contain enough spread in a spatial sense.

From surveys (Bovy, P.H.L. and Stern, E. (1990)) it is known that many travellers choose routes so as to optimise a particular subjective criterion. Whereas some travellers try to use the shortest route in distance, others prefer to use the shortest route in time, or having minimum traffic lights, minimum right turns or minimising other route attributes. The *preferential variability* criterion addresses this variability in preferences in the group by trying to generate as many as possible minimum label routes in the choice set.

Finally, at group level much more variety in mode type composition within and between routes is expected. The generated choice set should also satisfy the *multi-modal variety* criterion.

Table 7.5: Requirements for a reasonable multi-modal route and adequate choice set.

	Individual (OD pair)	Group level (OD zone level)
Single route	Acyclic criterion Detour criterion Hierarchic quality **Multi-modal feasibility**	
Choice set	Overlap criterion Comparability Detour-max criterion Detour-min criterion Choice set size **Multi-modal variety**	Overlap criterion Comparability Detour-max criterion Detour-min criterion Choice set size Spatial variability Preferential variability **Multi-modal variety**

Summary of desired properties of adequate multi-modal choice sets

Table 7.5 summarizes the requirements defined in the previous subsections, classified by single route at individual level, and by choice set at individual and group level respectively. These criteria will be used in the sequel to evaluate the quality of the generated choice sets produced by the various available multi-modal choice set generation approaches. Since the main aim of this thesis is to generate adequate multi-modal choice sets for prediction purposes, mainly the criteria for a group of travellers will be taken into account, additional to the criteria for reasonable single routes.

We will use these desired properties for the establishment of our new choice set generation approach for multi-modal networks (see Section 7.4).

These criteria will be adopted in Section 7.5.5 to check to which extent our new approach indeed fulfils these requirements.

7.2.3 Specific requirements for an appropriate choice set generation approach in multi-modal networks

As already introduced in Section 6.4, the generation process should ideally generate what we define as adequate multi-modal choice sets for prediction purposes, a choice set that matches partly (as much as possible) or totally the quality criteria we defined in the previous sections.

Secondly, the generation approach should satisfy as much as possible the following criteria that have been introduced in Section 6.4:

1. the approach should be suitable for both single OD pair and multiple OD pairs;
2. the approach should be generic;
3. the approach should be flexible;
4. the approach should satisfy the parsimony. requirement

In Section 7.3 an overview of multi-modal choice set generation approaches will be presented. Some of these methods use explicitly some of the quality criteria previously defined for generating adequate multi-modal choice set. For each method described in the following section it will be indicated which of these quality criteria are fulfilled in generating the choice sets, how they are applied by each approach, and which are the results of he generation process.

In Section 7.5 and 7.6 we will deal specifically with the operationalization (chosen thresholds and parameter values) of the developed criteria. In that section as well, the developed dedicated CSG for multi-modal networks will be presented and discussed, and its performance quality evaluated.

7.3 Route choice set generation algorithms for multi-modal networks

In this section, an overview of existing multi-modal choice set generation approaches is presented.

We will describe a sample of the most relevant methods to generate route sets in multi-modal transport networks, in which for sake of simplicity we distinguish between approaches applied to transit networks only and methods applicable to mixed private and public transport networks. Most of these approaches are based on shortest path search as already discussed in Section 6.5.

All these methods are based on one of the approaches described in Section 6.5; therefore we refer to these basic methods to illustrate the main aspects of the multi-modal approaches. Some of these methods use explicitly some of the quality criteria previously defined for generating adequate multi-modal choice sets. For each method described in the following it will be indicated which of these quality criteria are fulfilled in generating the choice sets, how they are applied by each approach, and which are the results of the generation process.

7.3.1 Route generation methods for multi-modal transit networks

A limited number of publications report on the generation of routes in multi-modal transit networks. For example, Lo, H.K. and Yip, C.W. and Wan, K.H. (2003), Montella et al. (1999), Nielsen, O.A. (2000) and Nuzzolo (2003) model travellers' combined modal-route choices in a network of multi-modal transit services, but the approaches are developed to be used in assignment problems and do not explicitly generate a priori choice sets. These approaches explicitly account for aspects related to transferring; Lo, H.K. and Yip, C.W. and Wan, K.H. (2003) also take differences in fare structures between transit modes into account.

Publications dealing explicitly with the generation of routes in multi-modal transit networks are listed in Table 7.6 and grouped according to the type of approach applied for the generation process. Most of these methods are based on shortest path search as indicated in Section 6.5, whereas two are based on the application of path-composition rules while two others are based on the enumeration method based on branch & bound technique.

Table 7.6: Route generation approaches for transit multi-modal networks.

Type of RSG approach	Reference
Shortest path search with path composition rules	Lozano and Storchi (2001) Nuzzolo and Crisalli (2004)
Shortest path search	Ziliaskopoulos and Wardell (2000) Horn, M.E.T. (2002) Florian (2004)
Enumeration technique based on branch & bound	Friedrich et al. (2001) Hoogendoorn-Lanser (2005)

Lozano and Storchi (2001) generate hyperpaths in a multi-modal network by applying a shortest path algorithm. The generation approach explicitly takes into account *path-composition rules* and a maximum number of transfers in order to generate feasible multi-modal routes. Nuzzolo and Crisalli (2004) generate multi-modal routes by first combining uni-modal alternatives generated for separate trip parts (access, egress, etc) and subsequently applied so-called path-composition rules, related to the traveller's perception of alternatives, to evaluate the resulting alternatives.

Ziliaskopoulos and Wardell (2000) present a time-dependent intermodal optimum path algorithm for multi-modal (transit and freight) networks that accounts for delays at mode and link switching points. Horn (2002) uses a generalized cost function and Dijkstra's shortest path algorithm to generate alternatives in a multi-modal network by considering different cost functions.

Florian (2004) introduces the so-called event dominance which he uses to efficiently implement schedule-based transit shortest paths. In this approach, the concept of dominance between two events at the same network element (node or link) is used to reduce the number of alternatives to be explored, because a dominated event does not need be considered further. Florian defines event dominance in terms of time and utility constraints, but implements this in EMME/2 in terms of time and costs constraints.

Friedrich et al. (2001) explicitly generate choice sets using a run-based, selective enumeration method (so-called *branch & bound algorithm*).

Hoogendoorn-Lanser's choice set generation approach extends that of Friedrich and differs in a number of respects, especially since the former approach is typically for multi-modal networks. Generated objective choice sets are 'all' relevant door-to-door trip alternatives given an origin and destination address and a time frame around a preferred departure or arrival time. Thus, not only 'best' but also 'second-best', 'third-best' etc. alternatives are included in the choice set. In a multi-modal context, this further means that alternatives may vary with respect to combinations of used transport modes (Hoogendoorn-Lanser, S. and Van Nes, R., and Bovy, P.H.L., 2006) and (Hoogendoorn-Lanser, S. and Bovy, P.H.L., and Van Nes, R., 2007).

7.3.2 Route generation methods for mixed private and public multi-modal networks

In literature, examples of choice set generation in mixed private and public transport networks are rare. Abdelghany, K.F.S. and Mahmassani, H.S. (2001) generate choice sets in such a multi-modal network using a k-shortest path algorithm. In Abdelghany, K.F.S. and Mahmassani, H.S. (1999) and Abdelghany, K.F.S. (2001) the *labelling approach* or *multi-objective shortest path approach* is used to generate choice sets in this type of network.

Benjamins et al. (2002) use *simulation methods* to generate choice sets in mixed private and public transport networks. To this end, network attributes are randomly drawn from distributions, while accounting for differences in perceptions and preferences of link attributes among travellers by distinguishing user classes. All these applications are *route-based approaches*, where a multi-modal alternative is represented by a route in a single integrated multi-modal network - constructed by interconnecting several uni-modal networks via walking and waiting links.

Publications dealing explicitly with the generation of routes in mixed private and public multi-modal networks are listed in Table 7.7. All these methods are based on one of the approaches described in Section 6.5; therefore, we explicitly refer to those approaches in order to list all relevant publications.

Table 7.7: Route generation approaches for mixed multi-modal networks.

Type of RSG approach	**Reference**
K-Shortest path	Abdelghany, K.F.S. and Mahmassani, H.S. (2001)
Labelling approach	Abdelghany, K.F.S. and Mahmassani, H.S. (1999)
M.O. Shortest path search	Abdelghany, K.F.S. (2001)
Simulation methods	Benjamins et al. (2002)
Branch & Bound	Hoogendoorn-Lanser (2005)

7.3.3 Advantages and disadvantages of the CSG methods for multi-modal networks

This section shortly discusses the main properties of the generation approaches, pointing out their advantages, but also their weaknesses in the case of applying the algorithm in generating choice sets in multi-modal transport networks (including private modes and public transport services) to be used in the choice set prediction.

The following approaches might be distinguished (see Section 6.5):

1. The K shortest path approach;
2. The shortest path search with constraints / with path composition rules;
3. The multi-objective (M.O.) shortest path search;
4. The enumeration technique based on Branch & Bound;
5. The labelling approach;
6. The simulation method.

In order to establish which of the approaches introduced in the previous sections are more appropriate to generate route choice sets in a multi-modal network, we make use of the main conclusions of Chapter 6.

The analysis of Chapter 6 concludes that the k shortest path with constraints is the best method among all methods applicable to a single OD pair. In the case of multi-modal networks we can also exclude the simple k shortest path since none of the quality criteria defined for the multi-modal case (such as multi-modal feasibility and multi-modal variety) are explicitly accounted for in the k shortest path search; in this method also the size of the choice set depends on the number k, which in some cases might be much larger than the desired size.

Moreover, the overlap and spatial variability criteria are not satisfied and generated paths tend to be very similar in spatial sense and also largely overlapping, mainly for the k shortest path approach, while the shortest path search with constraint is more adequate for the spatial variability (as shown in Chapter 6). The main disadvantage of the first two approaches based on shortest path search (with constraints, or path composition rules) and the k shortest path search is that those methods are applicable only to single OD pairs and are not very flexible (see definitions in Section 6.4). Because of this and other reasons, those methods tend to be computationally expensive on a network basis, especially in the case of multi-modal networks. We can conclude that those methods are not useful for generating choice sets in a multi-modal network.

Multi-Objective (M.O.) shortest path search is well known to be a $\mathbb{NP}$-hard problem and difficult to solve. Abdelghany, K.F.S. (2001) found that the multi-objective approach may outperform a single-objective approach in terms of quality of generated routes, but it needs much more computation time than a simple single-objective approach. However, as it is shown in Chapter 6, single-objective approaches may perform well not only in computation time but also in terms of quality of generated routes, therefore, in this thesis we will focus on single-objective shortest path approaches also in the case of multi-modal networks.

The enumeration approach applied, for example, by Friedrich et al. (2001) uses a utility function to select the best routes and a Branch & Bound technique to apply some constraints to eliminate non-adequate alternatives. However, in some cases Friedrich deletes a lot of alternatives; therefore it requires to be very careful in defining and applying such constraints, since some relevant alternatives might be cut off. Hoogendoorn-Lanser (2005) implements various variations of cost functions and various constraint types in order to have more alternatives. Basically Hoogendoorn-Lanser uses similar types of constraints as introduced by Friedrich. However, these constraints are loosened, meaning that they are applied to each combination of transport modes separately. Roughly speaking, for each combination of transport modes, the best alternatives satisfying the constraints are established. By loosening the original constraints, 'irrelevant' alternatives, i.e. alternatives that are counter-intuitive from a behavioural perspective, might end up in the choice set. In order to exclude these alternatives from the choice set, additional constraints were formulated, based on logical as well as behavioural assumptions.

An advantage of the algorithm introduced by Friedrich et al. (2001) and extended by Hoogendoorn-Lanser (2005) is that it does not require the same type of parameters that will be included in the utility functions. Another advantage is that the generated choice sets can easily be reproduced, because no stochasticity is involved in the choice process. The approach is also flexible in a sense that the core of the algorithm consists of a set of constraints which should be satisfied by separate trip parts and complete trips, and new constraints can easily be added to this set in order to account for specific aspects of travellers' behaviour (Hoogendoorn-Lanser, S. and Van Nes, R., and Bovy, P.H.L., 2006).

Based on the conclusions of Chapter 6 we can state that among all the approaches described in this section the labelling and the simulation methods are the best methods for generating adequate choice sets with reasonable multi-modal routes. It is preferable to have a combination of the two approaches, since among all methods described in Chapter 6 and applicable to single and multiple OD pairs the Monte Carlo Labelling combination (MCL) is the best method that satisfies most of the criteria.

Based on this evaluation, we can conclude that also in the multi-modal case the Monte Carlo Labelling combination approach is the one to be preferred from the other methods in order to generate an adequate multi-modal choice set at group level and for prediction purposes. Note that in this case it is highly recommended to apply a filtering process just after the generation process in order to obtain a more adequate choice set that satisfies the quality criteria.

Since no choice set generation approach exists that satisfies all requirements listed in Section 7.2.3, a new choice set generation algorithm based on the MCL approach needs to be developed which is applicable to mixed private and public transport networks.

7.4 Doubly stochastic generation approach to MM-CSG

This section describes the model approach developed in this thesis for generating choice sets in multi-modal networks, the so-called MM-CSG approach.

7.4.1 Principles of proposed MM-CSG method

The MM-CSG approach is based on the MCL combination approach, fully described in Chapter 6, since it results to be the best method satisfying the most quality criteria compared with the other approaches in generating adequate choice sets in multi-modal networks (see Section 7.2.3). The doubly stochastic approach is actually an extension of the MCL approach since not only the link attributes but also the preferences (or behavioural) parameters of the cost function are randomized.
The concept of randomising both link attributes and behavioural parameters in the choice set generation, although done simultaneous with the route choice modelling, is somewhat similar to Nielsen, O.A. and Frederiksen, R.D. (2006), who found that this led to a drastic improvement of the choice set.

As introduced in Section 7.2, especially in a multi-modal context, variations in transport modes and in multi-modal combinations are desirable in the generated routes. It is important to note that an efficient technique that can be successfully applied in order to achieve this variation in the generated routes is provided by the use of user classes and the randomization of the behavioural parameters, which take into account the taste variations within the user classes, and which is provided by the doubly stochastic approach.

Therefore, the proposed MM-CSG approach is based on the following three principles:

- the composition of consideration sets is strongly related to choice preferences of travellers;
- attractive (reasonable) routes are derived from probability assumptions about these choice preferences;
- these attractive routes can be generated through repeated shortest path search by proper randomisation of the network attributes, the behavioural parameters, and the search criterion (disutility function).

The basic premise of our CSG method is that the composition of an individual choice set is related to the traveller's personal choice preferences. This means that the attributes influencing the route choice of a traveller are assumed also to play a role in which routes he considers in his choice set. We assume that, for example, a time sensitive traveller only considers alternatives that fulfil this preference whereas a cost sensitive traveller or a transfer-averse traveller especially considers alternatives that reflect these preferences.

This proposed approach is of a stochastic nature since potentially attractive routes may only be selected in the choice set with a certain probability; it depends on the sampling procedure to make the probability as high as possible. We call this method the doubly stochastic approach since both the preference parameters and the attribute values are made stochastic in the route's disutility calculation.

The basic idea of the MM-CSG approach is that the method focuses on generating routes for prediction purposes at group level, therefore for this reason in the following *groups of travellers* are taken into account. In fact, in the doubly stochastic approach we distinguish *user classes* that vary with respect to expected travel behaviour, for instance based on *trip purpose* (work, business, study, etc), and *traveller groups* that are given by a combination of vehicle availability and preference, and the randomization is based on those user classes.

As claimed in Chapter 6, in order to achieve adequate choice sets that satisfy most of the requirements defined for generating reasonable routes, it is recommended to have a filtering process after the generation for approaches based on the MCL method.

The operationalization of this stochastic approach consists of repeated shortest path search in a randomised network with randomised preference parameters, as will be shown in the next subsection. The cost function used in route search will be explained in subsection 7.4.3.

7.4.2 General scheme of MM-CSG algorithm

This section will present how we operationalize the doubly stochastic approach developed in this thesis for generating choice sets in multi-modal networks. The scheme shown in Figure 7.3 is an elaboration of the classification scheme shown in Chapter 6 (see Figure 6.9).

In the scheme we can distinguish six input boxes, three output boxes and six processes. The inputs of the algorithm are the multi-modal supernetwork, the search function, the random parameters; the OD list of trips for which choice sets are to be determined, the termination criteria, and the set of constraints to be applied in the filtering process. The outputs are the intermediate route set, and the final filtered choice set. The approach is based on repetitive applications of a shortest path (SP) algorithm on the same network successively by applying adapted search functions based on the cost functions computed in the Compute Cost Functions box with the randomised behavioural parameters and network attributes.

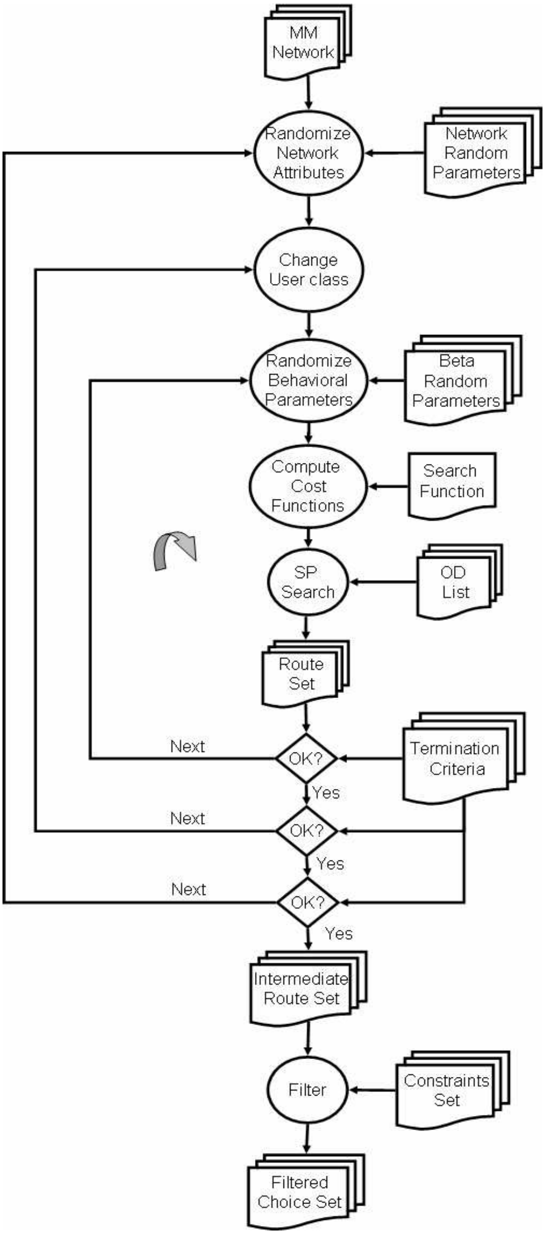


Figure 7.3: Algorithmic set-up for generating choice sets in multi-modal networks.

According to the specification of the algorithm, three main loops can be distinguished in the scheme. First, the outer loop in which the network attributes (link properties X) are randomized. Then, in the middle loop the user class is selected (different labels are applied) according to different traveller groups and trip purposes. Finally, based on the user class selected, in the inner loop the preference parameters β are randomised. After having randomized both link attributes and preference parameters, the stochastic link cost functions are computed based on the search function before applying the path search for each OD pair in the OD list. The new optimal route found is included in the route set if the route appears to be a new one (routes overlapping 100% with a previous route found are excluded during the generation process). At the end of each cycle the termination criteria, which may consist of number of iterations for the randomization process (possibly choice set size), are checked. The three loops depend on three numbers of iterations: the number of iterations for the network randomizations, the number of user classes, and the number of iterations for the parameter randomizations.

In the scheme, the filtering process is applied based on the Constraints Set input after having generated the intermidiate route set. Mainly for efficiency reasons, as stated in Chapter 6, all constraints except the very simple 100% overlapping criterion, which eliminates routes that are 100% identical at link level, are applied after the generation process. A typical property of the proposed approach is that one and the same route maybe found several times as the best one, even with different cost values, during the generation process. Only these superfluous routes are ignored during the generatio which may contribute to inefficiency of the method to some extent, all the other routes are added to the intermediate route set.

The randomisation may refer to several dimensions of the route generation process, separately or combined, namely:

- the various preference parameters β part of the search function C;
- the various link attributes X part of the search function C.

The search function C is the function that will be minimized by the shortest path search in order to generate an attractive route. This criterion is defined by the included link attributes X and by the corresponding β-values. The criterion may be a single attribute or a multi-attribute function. It is a generalized cost function, explained in more detail in subsection 7.4.3. A large variety of criteria may be distinguished and used. The specification of the criterion may be fixed and given in advance or maybe changed during the generation process depending on the outcomes of the process, such as the size or composition of the resulting choice sets. The criterion is a random function since its parameters β and variables X are random variables, being randomized during the generation process. The resulting shortest path and its cost value, therefore, are also random variables. The probabilistic properties of the random outcomes (in terms of type and number of generated routes) will be given attention in Section 7.4.4.

The adopted search criteria are based on different label criteria such as: minimum time, minimum distance, minimum cost, etc. in order to generate multiple possible paths expressing that different travellers may have different objective functions in seeking routes. Depending on trip purposes and vehicle availability, different traveller groups or user classes are taken into account in the generalized cost function in order to reproduce the different label criteria. For each label, a specific search criterion (cost function) is defined (Search Criteria in Figure 7.3) and a related SP search is applied (see Figure 7.3). Based on these randomisations, for each OD pair the minimum-cost path according to search criterion C is computed and will be added to the route set if the route appears to be a new one.

Depending on the intermediate outcomes of the generation process, the *randomisation properties* initially started with might be adapted during the generation process. Important randomisation properties are the variances of the random distributions of the various β-parameters and X-attributes, and maximum numbers of drawings. For example, if the size of the generated choice set does not fulfil predefined thresholds, then these variances maybe increased in an attempt to generate additional alternatives, as is for example the case for the extension of the MCL combination approach described in Chapter 6. However this adaptive method will not be applied in our multi-modal experiments and it is postponed to future research.

Termination criteria are very important since they determine the quality of the final choice set as they relate directly to the requirements of an adequate choice set, presented in Section 7.2. Several termination criteria maybe applied such as based on the maximum number of iterations needed for the randomisation of the link attributes, the maximum number of iterations needed for the randomisation of the β parameters, the attainment of the minimum or maximum size of the choice set, the level of replication of routes already generated in previous steps, the growth speed in the choice set size, etc.

7.4.3 The generalized cost function adopted in MM-CSG

The shortest path route search criterion has the form of a generalized travel cost function. This function transforms various route attributes such as travel time, waiting time, transfer time, etc into a single cost variable. The generalized travel cost function adopted in our route search criterion synthesizes the most important multi-modal trip attributes and their weights as known from earlier studies (Van Der Waard (1988a), Nielsen, O.A. (1996) and Wardman (2001)).

In our combined labelling and randomization procedure we distinguish user classes (represented by suffix s and given by a combination of traveller group and trip purpose). In principle, the generalized cost functions will differ with respect to expected travel behaviour, for instance based on trip purpose (work, business, study, etc), and traveller groups that are given by a combination of vehicle and preference availability at the home-end and the activity-end of the trip, such as car only (car captive), PT only (PT captive), etc. Please note that in our supernetwork every link relates to a single mode and a single trip component.

All methods generating route choice sets described in Chapter 6 apply a specific search criterion on the basis of which the most attractive route is determined. Our model approach also is based on the principle to generate the most attractive routes in the multi-modal network based on specific search criteria. To that end the notion of generalised cost of a route for user class s is introduced as the basis for the search criteria to be considered in our stochastic approach.

First, recall the formulation of a multi-modal transportation network introduced in Chapter 3 (Section 3.7.2), which is defined by a directed graph $G = (N, L, M)$, where N is the finite set of nodes, L the finite set of links, and M the finite set of travel modes. On graph G, one and only one transport mode $m \in M$ is associated with each link $(i, j) \in L$. Hence, there exist $|M|$ simple networks, one for each travel mode, interconnected by transfer links. The sets of nodes and links forming each one of the $|M|$ networks are respectively denoted by N_m and L_m, where m $(m \in M)$ stands for a travel mode. In our formulation walking is always considered a travel mode, and transfer links consist of parking links in the case of private travel mode (such as, car and bicycle) or boarding / alighting links in the case of PT travel mode.

For our generation purpose, the randomized travel cost $\bar{C}_{sk}$ of the multi-modal path k for user class s is a linear summation of the randomized link costs $\bar{c}_{sa}$:

$$\bar{C}_{sk} = \sum_{a \in k} \bar{c}_{sa} \tag{7.2}$$

where $\bar{c}_{sa}$ is the randomized travel cost c on link a for user class s (the randomization is expressed by the upper bar). Link cost $\bar{c}_{sa}$ is considered to be a stochastic quantity reflecting the variation of attribute perceptions X and attribute preferences β among travellers belonging to traveller group s.

Please note that for route choice modelling non-linear cost functions might be applied.

The randomized generalised cost $\bar{c}_{sa}$ of a link a for user class s is modelled as a weighted (linear) combination of link attributes X_{amy} of the constituent links:

$$\bar{c}_{sa} = \sum_{m \in M} \sum_{y \in Y} \delta_{ma} \cdot (\beta_{smy} + \hat{\varepsilon}_{smy}) \cdot (X_{amy} + \tilde{\varepsilon}_{amy}) \tag{7.3}$$

where index a refers to links, m to the travel mode and index y represents the link attributes X, while δ_{ma} is a $\{0,1\}$ variable being 1 if link a belongs to mode m, 0 otherwise.

In expression 7.3, β_{smy} represent the individual preferences of user class s specific for transport mode m and for link attribute y. The stochastic values of the average attribute weights are expressed by introducing a random error term $\hat{\varepsilon}_{smy}$ to consider interpersonal variation in the weights of the perceived attribute y.

X_{amy} represents the y's attribute of link a specific for transport mode m, mostly a time attribute depending on the link type. The stochastic value of the link attribute is expressed by taking into account the random error term $\tilde{\varepsilon}_{amy}$ due to variability in the attribute perceptions X_{amy} among travellers and uncertainty on the part of the analyst.

For our doubly stochastic CSG principle, to be explained in the sequel, the precise form of the disutility function, however, is not at all important. As motivated above, the basic idea of the doubly stochastic approach is to randomize both the link attributes and the β parameters of the cost function 7.2.

The *analyst* having the task to predict a choice set for a group of travellers does not know the exact values of β_{smy}, nor those of the perceived attributes X_{amy}. Apart from being unknown, the values for β and X may vary for the same individual (intra-personal variation) because of all kinds of reasons such as trip type, prevailing traffic conditions, etc., but also at group level interpersonal variations in preferences and in perceived attributes between individuals and between trip types have to be considered as well.

Therefore, the analyst is forced to make assumptions about the probability distributions of both β_{smy} and X_{amy}. With respect to βsmy (s)he may assume that the probability distribution is related to average values found for these parameters in estimations for the corresponding group category, while for the attributes X_{amy} the analyst may assume that the perceived values are related to the objectively measurable values. By establishing such probability distributions and sampling from these distributions the analyst may now calculate by using a predefined disutility function a range of disutility values that potential routes should satisfy in order to be eligible for the consideration set. This sampling annex calculation should be performed such that the most attractive (least disutility) route will be part of the choice set with high probability. If the variation of preferences and perceptions within the group of travellers or for trip types is very large, the analyst may be forced to establish separate sets of probability distributions for these categories in order to come up with adequate choice sets for each group and adopt probability distributions for β and X with a larger variance.

For each traveller group, the parameter values of the link cost function are sampled from some (positive) statistical distributions, followed by computing the generalized link costs with respect to randomized link attributes and parameters. Given the network with its objective link attributes X, randomised attribute values are generated by sampling from some positive statistical distributions (e.g. uniform, truncated normal or Gamma distribution).

Since the considered network is multi-modal, in our model approach we have to take into account link attributes that are usual for private and public transport trip parts supplemented with attributes that are typical for the transfer movements between modes. In particular, distance and mode-specific travel times are taken into account as attributes for private modes (car and bike), whereas mode-specific in-vehicle travel times are taken into account for public modes (train, bus, tram and metro). In addition to the travel link attributes, the following attributes associated with transfer links are considered: waiting time, boarding time, alighting time, parking time and parking cost (the latter only for private modes such as car (driver) and bicycle).

It should be noted that contrary to the cost functions used in the choice modelling, cost functions adopted in the generation modelling need not be representative for real world behaviour. Their primary function is to help in generating reasonable route alternatives in terms of the most important ones in a relative sense. The absolute values of the route properties are not important in this stage, although they are important in the choice modelling stage. That's why questions like correlations among attributes or parameters, and the existence of non-linear cost structures also are not of prime importance in the generation process.

7.4.4 Probabilistic properties of doubly stochastic MM-CSG

Because of its stochastic principle, a typical property of the MM-CSG generation approach is that its results are stochastic (see Bovy, P.H.L. and Fiorenzo-Catalano, S. (2006)). This means that the size and composition of the generated choice sets are stochastic variables. It is a particular case of sampling with replacement. Whether a particular route alternative (e.g. the most attractive route) will have been generated (sampled) after a certain number of randomizations can only be said with a certain probability. Also whether the generated choice set will have a predefined minimum size after a certain number of randomizations can only be said with a certain probability. These probabilities on the one hand depend on the (mostly unknown) properties of the population of route alternatives in the network, while on the other hand they depend on the randomization properties (number of randomizations, seeds, variance levels, etc.).

Thus, in general, the relevant outcomes of the generation approach all are uncertain depending on its stochastic input. Therefore, in applying such generation approaches the question arises, for example, how many draws are needed to generate with a required probability a subset consisting of (at least) a certain number of different routes? Alternatively, the question might be what the probability of generating exactly a certain number of different alternative routes $(r \leq \kappa)$ is given a predefined number of κ draws from a network with R alternative routes? Many more similar questions emerge in performing stochastic choice set generation approaches.

Partly, such questions may be answered by monitoring the generation outcomes during the choice set generation process, such as, for example, the repetition of already selected alternatives. However, in planning the generation applications it is worthwhile to have some general rules at hand that can govern the experimental set up of the generation process such as the minimum number of random draws, or the best termination variable in the monitoring process.

On intuitive grounds, the following trivial probabilistic relationships can be stated:

1. with increasing number of randomizations:
   - choice set size will increase, albeit to a maximum;
   - choice set composition will be more stable;
   - number of selected attractive routes goes towards exhaustive;
   - number of selected unattractive routes is limited.
2. with increasing variance levels (of attributes and/or parameters):
   - choice set size will increase, not necessarily to a maximum size;
   - choice set composition is less stable;
   - number of selected attractive routes goes towards exhaustive;
   - number of selected unattractive routes will increase.

These relationships indicate that variance levels should be kept modest.

An in-depth theoretical analysis of the probabilistic properties of the MM-CSG method has been carried out by Li, H. and Bovy, P.H.L. and Hooghiemstra, G. (2007). From these analyses a number of generally useful statements can be derived, such as the following sample may demonstrate (for details see Li, H. and Bovy, P.H.L. and Hooghiemstra, G. (2007)):

- if the ratio between sample size κ and resulting choice set size k is larger than 5, there is at least 90% confidence that the choice set size includes all alternatives with non-zero selection probability;
- very large samples ($\kappa \gg R$) are needed to guarantee sufficient stability in the composition of stochastically generated route choice sets;
- with large networks (many alternatives R) small choice sets (size k) can be generated already with small samples (size κ draws).

7.5 Demonstration of the doubly stochastic MM-CSG approach

7.5.1 Scope and purpose of demonstration

This section presents a demonstration of the application of the proposed doubly stochastic choice set generation approach (see also Fiorenzo-Catalano, S. and Van Nes, R. and Bovy, P.H.L. (2004a) and Fiorenzo-Catalano, S. and Van Nes, R. and Bovy, P.H.L. (2004b)). The aim of the demonstration is to show the performance of the developed MM-CSG approach. Since in a multi-modal context it is very important to generate multi-modal combinations especially at leg level, we want to show with this demonstration that the randomization of the behavioural parameters together with the randomization of the links attributes plays an important role in achieving this aim. To that end we compare the generated routes with a set of observed routes and compute a set coverage, i.e. how many observed routes match the generated ones and to what extent. The method adopted for demonstrating the MM-CSG approach is described in Figure 7.4.

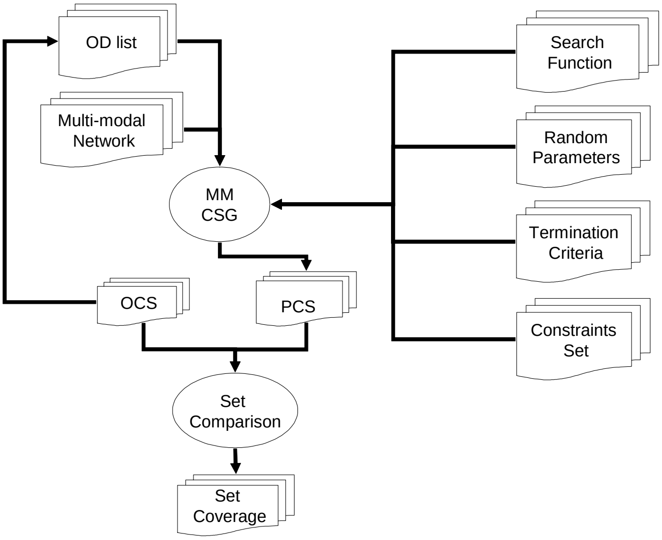


Figure 7.4: General scheme for demonstrating the MM-CSG model (OCS means Observed Choice Set and PCS means Predicted Choice Set).

We can see from Figure 7.4 that the inputs of the MM-CSG model are the same inputs as shown in the algorithm scheme of Figure 7.3: the multi-modal network, the OD list, the search function, the random parameters, the termination criteria and the constraints set input box. Note that the filtering process based on vehicle availability constraints is included in the MM CSG box (as described in Figure 7.3) and it is applied after the generation process based on the constraints set box, therefore the predicted choice set (PCS) generated by the MM-CSG approach has been already filtered when it is compared with the observed routes.

The demonstration of the MM-CSG model focuses on obtaining a certain coverage percentage by comparing (Set Comparison box in Figure 7.4) the predicted choice set (PCS box in Figure 7.4) and the observed choice sets (OCS box in Figure 7.4). In the demonstration we applied default network attributes and behavioral parameter values as input for the MM-CSG model as well as the default user classes based on trip purposes and vehicle availability and preference. Please note that all parameter values used in this analysis are not yet calibrated but rather are based on literature and experience; in line with those found in recent studies (see e.g. Nielsen, O.A. (1996) and Wardman (2001)). The calibration of these parameters will be discussed in Section 7.6.

In the demonstration, the choice set generation algorithm described in Section 7.4.2 has been applied to compute the shortest paths with respect to the randomized generalized cost. The input of the MM-CSG approach as shown in Figure 7.4 are as follows:

- the multi-modal network, to be described in Section 7.5.2;
- the OD list extracted from the observed data, to be presented in Section 7.5.3;
- the search function similar to the one introduced in Section 7.4.3; the formulation of the randomized link cost function adopted in our current implementation and application considers mainly three link attributes: link cost, link travel time, and link length. Depending on the link type travel time may mean walking time, waiting time, transfer time, and the like;
- the user classes (consisting of 4 trip purposes times 4 vehicle availability categories) and the default parameter values that are taking into account in the randomization process, such as the values of the β parameters and the link attributes;
- the constraints set applied in the filtering process, consisting mainly in eliminating all those generated alternatives that cannot be observed because, for example, a specific vehicle (car or bike) was known not to be available to the traveller (no voerlap criterion is applied at this stage);
- a fixed number of total iterations has been adopted as termination criterion.

Given these inputs, three options for generating route sets have been studied, i.e.: variation in the network attributes only, variation in traveller preferences only, and the combination of both. The randomization adopted in the demonstration is the truncated normal distribution both for the link attributes and the preference parameters, the adopted number of randomizations for the link attributes and for the behavioural parameters is 10-by-10, and 16 user classes are taken into account, given by the 4 trip purposes (work, business, study, and shop) times 4 vehicle availability classes (such as, all modes available, car only, PT only, PT and bike).

The predicted choice sets are compared with observed choice sets using coverage measures as performance criterion. Please note that this performance test sets high standards for the comparison. While the generation method is designed to generate route sets for a group of travellers, the comparison is made with respect to individual choice sets.

Section 7.5.4 summarizes the most important findings and outcomes obtained by the generated choice sets; all details about the coverage measure and the description of the comparison applied between the generated routes and the observed ones are dealt with in Section 7.5.5.

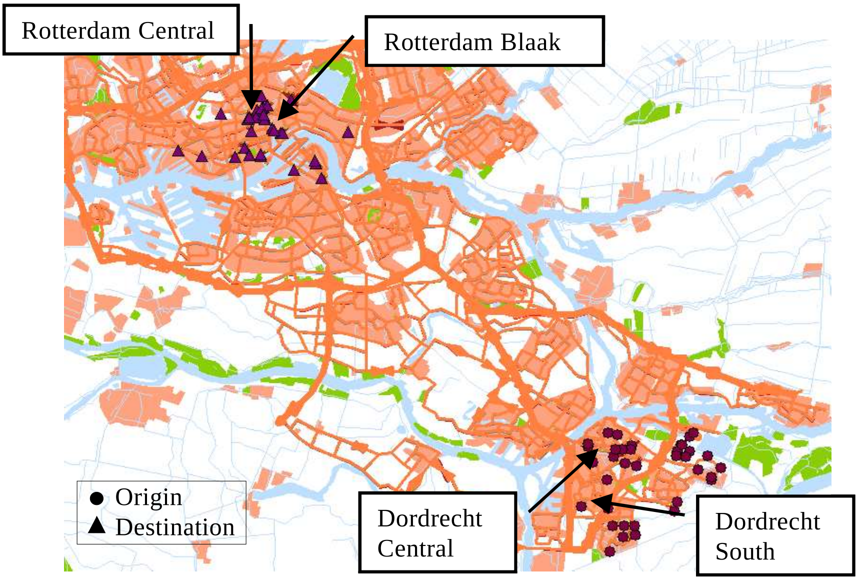


Figure 7.5: Overview of the corridor Dordrecht-Rotterdam and the selected trip origins and destinations.

7.5.2 The network case used for demonstration

The considered case is the corridor between the cities of Dordrecht and Rotterdam in the Netherlands, which are about 30 kilometres apart, with a total population of about one million. As the availability of private vehicles is clearly important in multi-modal route-choice it was decided to focus on home-based trips in which privately owned vehicles are available to travellers. Travellers in this corridor can use car and train as their main mode. In the case study three types of train services are available: local services, express services, and intercity services. Two stations in Dordrecht are considered in the corridor: Dordrecht Central, at which all services call, and one station served by local and express train services only (Dordrecht Zuid). Among all Rotterdam's railway stations four of them are considered in the corridor: Rotterdam Central (all services), Rotterdam Lombardijen (express and local services) and two stations served only by local trains: Rotterdam Zuid, and Rotterdam Blaak. All stations in the area are accessible by foot, bicycle, car, bus, tram and metro (the latter two in Rotterdam only). Both central stations have extensive facilities for bicycle storage and bicycle renting, but car-parking facilities at Rotterdam Central are limited. The resulting supernetwork consists of about 11,000 nodes and 34,000 links.

7.5.3 Observed route data used for demonstration

In order to analyze whether our generated route sets match the subjective choice sets, a comparison is made with observed (reported) route sets (Hoogendoorn-Lanser (2005)). For practical reasons, this analysis is limited to a sample set of 37 OD-pairs in the corridor Dordrecht (home-end) and Rotterdam (activity-end) during the morning peak hour (7.00 to 9.00h). Figure 7.5 shows the locations of the origins and destinations of these trips. To evaluate the performance of our choice set generation approach, use has been made of observed trips collected from a large survey conducted among Dutch train travellers in 2001 (for more details about the survey refer to Hoogendoorn-Lanser (2005)). Travellers reported their chosen alternative, that is, the sequence of transport modes and the transfer nodes, as well as the set of non-chosen alternatives they knew related to access modes, train service types, boarding or alighting stations and egress modes. These reported alternative routes are assumed to be representative for the subjective choice set. In the case of the selected 37 OD pairs the total number of observed routes is 67; the average size of the reported subjective choice sets is between two and three alternatives per OD-pair, with a minimum of one and a maximum of six reported alternatives per individual respondent.

7.5.4 Results of the route generations

Table 7.8 summarizes some of the characteristic outcomes of the generation procedures. Generated routes are always unique, although they may be overlapping with each other to some extent.

During the route generation procedure, all paths generated that are 100% overlapping with a previously found path are rejected. In order to compare the generated subjective choice sets with the sample of observed individual route sets additional constraints are used to account for the traveller's vehicle availability. Furthermore, since the observed route sets only contain train trips, routes by private modes only were skipped from the analysis.

The choice set (CS) generated by the randomized link attributes (RA) approach (CS-RA) contains a total of 286 route alternatives for the 37 OD-pairs with an average size of 8 alternatives per OD-pair, a minimum of 3 and a maximum of 18 route alternatives. The choice set (CS) generated by the randomized preference parameters (RP) approach (CS-RP) contains 283 route alternatives with an average size of 8 alternatives per OD-pair, a minimum of 4 and a maximum of 16 route alternatives. The choice set (CS) generated by the combined (RC=RA and RP) approach (CS-RC) contains 701 route alternatives with an average size of 19 alternatives per OD-pair, a minimum of 10 and a maximum of 36 route alternatives.

Table 7.8 shows the coefficient of variation of choice set size among all OD pairs; the values of this coefficient are reasonable since each OD pair has different possible alternatives, and choice sets differ in size. This is logical because the choice set size depends on the OD pair.

The sizes of these choice sets look plausible and do not conflict with empirical knowledge about objective choice sets (Bovy, P.H.L. and Stern, E. (1990)). Please note, that the average size (see Table 7.8) of the resulting choice sets is reasonable given the fact that the generation method generates choice sets for a group of travellers (at aggregate level) instead of individuals (prediction purpose) and that there is no restriction on the magnitude of overlap (except if this is exactly 100%).

Interestingly the first two generated choice sets (CS-RA and CS-RP) are quite similar in size. However, it might be expected that there are clear differences in route and set compositions. It can be hypothesized that the randomization of network attributes might lead to a smaller variety of routes than the randomization of behavioural parameters. It appears that the combination of randomized attributes and travellers' preferences generates many more different route alternatives with the same number of iterations. This might be explained by the fact that the routes generated by varying network attributes usually vary around the shortest path. If variations of travellers' preferences lead to more different routes, the combined approach would lead to additional variations on these different routes. Finally, the figures in Table 7.8 show that the generated numbers of routes is a lot less than the theoretical maximum.

Table 7.8: Choice Set (CS) generation results for sample of 37 OD-Pairs.

	CS-RA	**CS-RP**	**CS-RC**	**Theoretical extreme**
Total number of distinct routes generated	286	283	701	1600 x 37
Average number of routes per choice set	8	8	19	1600
Maximum Choice Set Size	18	16	36	1600
Minimum Choice Set Size	3	4	10	1
Coefficient Of Variation of OD-Choice Set Size	52%	35%	32%	

RA=randomised link attributes;
RP=randomised preference parameters;
RC= RA and RP combined.

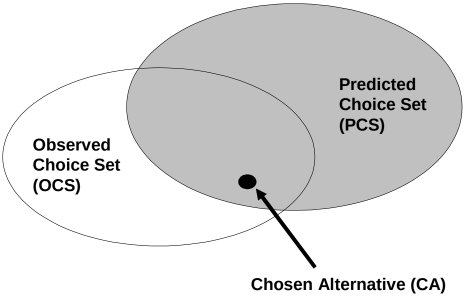


Figure 7.6: Comparison between the Predicted Choice Set (PCS) and the Observed Choice Set (OCS).

7.5.5 Performance comparison

The key question in this section is how well the individual choice sets generated by each approach match the observed individual route sets. In comparing two sets, e.g. A and B, we define the set coverage of A with relation to B as the percentage of alternatives in set A that are also elements of set B. We are looking for the coverage of the observed sets (A) by the generated sets (B). We only look at full (100%) identity of predicted and observed routes. We distinguish three levels of comparison with increasing level of detail:

- Station level: home-end station and activity-end station combination;
- Trip part level: home-end mode, train service types, activity-end mode;
- Trip level: unique combination of home-end mode, home-end station, train service type, activity-end station, and activity-end mode.

The predicted subjective choice sets generated by each of the three approaches (CS-RA, CS-RP, and CS-RC) are compared with the observed alternatives (see Figure 7.6) which are the chosen trip alternative (CA) and the set of trip alternatives (including the chosen alternative) reported to be known (KA) by the traveller (reported subjective choice set). Please note that while the generation method is designed to generate route sets for a group of travellers, the comparison is made with respect to individual choice sets.

Therefore, we are comparing for each OD pair the set of Predicted Choice Set (PCS) and the set of Observed Choice Set (OCS), which may consist of one single alternative (the chosen alternative CA), or the set of reported alternatives (KA). The total coverage of the whole observed set for each level of comparison (station, trip part, and complete trip levels) is computed as follows:

$$TotCov(OCS/PCS) = \frac{1}{|OD|} \sum_{od=1}^{OD} Cov_{od}(OCS/PCS) \tag{7.4}$$

where $Cov_{od}(OCS/PCS)$ is the coverage of the single chosen alternative (CA) or the set of reported alternatives (KA) with respect to the PCS for a single od-pair. This coverage is computed as follows.

$$Cov_{od}(OCS/PCS) = \frac{(OCS \bigcap PCS)}{OCS} \tag{7.5}$$

This coverage measure may vary between 0 and 1. It is zero if the single alternative CA or the KA set are not included in the PCS (the set intersection is null), or it is 1 if the single alternative CA or the KA set are fully covered by the predicted choice set (PCS).

Table 7.9: Set coverage results for each of the three generated choice set types: CS-RA, CS-RP, CS-RC.

N=37 OD-trips	**CS-RA**		**CS-RP**		**CS-RC**	
	$CA \subseteq RA$	$KA \subseteq RA$	$CA \subseteq RP$	$KA \subseteq RP$	$CA \subseteq RC$	$KA \subseteq RC$
Home-end and activity-end railway stations	91.9%	91.0%	86.5%	84.2%	94.6%	93.7%
Home-end trip part modes	54.1%	54.9%	97.3%	96.2%	97.3%	96.9%
Train trip part	89.2%	88.3%	54.1%	53.2%	89.2%	88.3%
Activity-end trip part modes	83.8%	86.5%	83.8%	82.0%	91.9%	91.9%
Complete trip	37.8%	40.5%	51.4%	50.2%	78.4%	77.8%

RA=randomized link attributes; RP=randomized preference parameters;
RC=RA and RP combined.
$CA \subseteq RA$ means: chosen alternative is in predicted choice set (RA, RP, or RC).
$KA \subseteq RA$ means: reported alternatives are in predicted choice set (RA, RP, or RC).

Table 7.9 shows the set coverage results for the choice sets (CS-RA, CS-RP, and CS-RC) generated by the three approaches and their comparison with the chosen alternative (CA) and the known alternatives (KA) respectively. The table shows that the best coverage at trip level is given by the combined approach RC with a percentage of 78%, which might seem obvious given the number of routes found using the RC approach. More interesting is that the RP approach in which the randomization is applied to travellers' preferences only, produces a better coverage than the RA approach in the case of comparing the complete trip. Apparently, the variation in behavioural parameters is more important when considering the complete multi-modal route alternatives than the variation in network attributes.

If we analyze the coverage results of the combined approach we can observe that at the level of home-end and activity-end station, the set coverage is very high for both the chosen alternatives and the known alternatives: about 94%. At the level of individual legs, the set coverage is still high: more than 88% of the reported legs are part of the generated subjective choice set. At the trip level, the set coverage is still high: about 78% for both comparisons, even if it is not as high as the comparison with the trip components. Of course, the classification of high and low might seem arbitrary. The comparison of observed and generated route sets, however, is fairly unique. In a recent study Ramming, M.S. (2002) a comparison has been made between observed routes (chosen route only) and generated routes for a road network. A coverage level was found of 72% for a combined labelling method, 60% for multiple-path algorithms, and 50% for a simulation method with optimized values for the standard deviation. Given these findings, the results of our method can be classified as very promising since we have not yet used optimised parameters.

Therefore, the main conclusions of the demonstration are as follows:

1. the stochastic MM-CSG approach works well since in most cases (78%) the observed route is predicted;
2. the coverage level is very promising, and several and different multi-modal combinations have been generated showing that the randomization of both link attributes and behavioural parameters has been successful;
3. the generated choice sets look plausible and their sizes do not conflict with empirical knowledge about objective choice sets;
4. the coverage may be improved if the default parameters values are better determined by a dedicated calibration.

7.6 Calibration of the MM-CSG model

7.6.1 Scope and purpose of the calibration

This section deals with the calibration of the MM-CSG model to determine the optimal parameter values of the generation function. Given the findings obtained in the demonstration case, the algorithm and its results can be improved in many ways. A sensitivity analysis might be performed with respect to the variances at link level (X attributes) as well as with respect to the weights (β parameters).

In the calibration, the MM-CSG approach is applied to the multi-modal network (as described in Section 7.5.2) to establish the *best* parameter values by comparing predictions with observations through the coverage measure defined in Section 7.5.5. In the calibration phase the adopted number of randomizations for the link attributes and for the behavioural parameters are 10-by-10, with 3 user classes taken into account, given by one trip purpose (work) and 3 vehicle availability classes (such as, all modes available, PT only, PT and bike). Please note that in the calibration phase the number of user classes has been seriously reduced compared with the demonstration case (from 16 user classes to only 3).

The developed MM-CSG method entails a large number of cost function parameters and process parameters, the values of which need to be optimised in order to achieve best generation results. The generation method should establish adequate choice sets (see Section 7.2.2) with respect to their minimum size, variety in composition (especially multi-modal routes), coverage of real-world behaviour, etc. The adequacy most importantly will be judged by comparison with observations of consideration sets and chosen routes. The most simple and easy approach is to look whether the observed chosen route is included in the predicted choice set, or to what extent the observed chosen route is covered by predicted routes. In addition it can be assessed how much effort is needed in the generation to achieve a certain coverage result. A more realistic and more demanding case is when observations of consideration sets are available. In that case, the adequacy of the predicted choice sets may be judged as to whether or to what extent these subjectively considered routes are covered by the predicted routes.

In a first step prior to calibration, default values are set for expectations and variances of which the default expectations were derived from exogenous choice model estimates published in literature.

During the calibration phase, the MM-CSG approach is applied with the same randomization process explained in Section 7.4.2, but adding systematic variations of some values, such as random parameters, and user classes, in order to determine how much the coverage can be improved. Please note that while the generation method is designed to generate route sets for a group of travellers, the coverage measurement is applied to the observed individual choice sets.

The generation procedure introduced in Section 7.4.3 consists of a very large set of parameters of a different kind. First of all, we have the behavioural parameters present in the generalized cost functions of the various user classes based on different traveller groups and trip purposes. Secondly, we have the various process parameters that govern the randomisations of parameters and attributes (expectations, distribution types, variances, minimum and maximum values of random distributions), and termination parameters (number of iterations for each randomization and maximum number of iterations).

In general, the attributes X and parameters β used in the generating function are defined by probability distributions characterized by a distribution type, an expectation (μ), and a variance (σ). Given the formulation of the link cost function as in formula 7.3 in which only costs, travel times and distances are considered as link attributes, the expected values of the attributes (link costs, link times, link distances, etc) in most cases may stem from readily available data and need not be calibrated. Among all these parameters only the travel time attributes and the β parameters are randomized. The values of the link cost and link distance depend on the network values. The variances of the uniform attribute distributions in principle are subject to calibration. In line with our behavioral hypothesis (see Section 7.4.1) that the attribute randomization reflects perception errors and the like, we confined the variance calibration to those time attributes known to be most important in travelers' trip choices (e.g. in-vehicle times of car, PT, and bicycle, PT waiting time, etc.).

For the distributions of the random attribute values of links (X), the adoption of positive uniform distributions at the level of the links seems justified since the summation of random link values to corresponding route values then results in a smooth positive distribution of approximately normal form. Also for preference parameter (β) distributions positive uniform distributions are assumed after some experimenting with other distributional forms such as truncated normal.

Therefore, the randomization adopted in the calibration is based on the uniform distribution both for the link attributes and the preference parameters. The uniform distribution has been chosen for many reasons. In the calibration phase we would like to random generate values in the entire range, not only in the central part of the range (as is more likely with a normal distribution) but also in the extremes. Moreover, from the literature (Bovy, P.H.L. (1990)) it is well known that the uniform distribution is a useful distribution for simulation purposes. From a theoretical point of view, a nice property of the uniform distribution is that if the link costs are uniformly distributed at link level, the route costs are normal/gamma distributed at route level, since the route cost is given by the linear summation of the link costs (Bovy, P.H.L. (1990)). Finally, if the uniform distribution is positive also the normal distribution is positive and from a practical point of view, the uniform distribution is easy to implement. For all these reasons we have applied a uniform distribution in the calibration instead of the truncated normal distribution used in the demonstration case of the MM-CSG approach (see Section 7.5).

7.6.2 Network attributes and behavioural parameters to be calibrated

The calibration involves mainly a set of variance values to be determined by maximizing some measure of resemblance with observations of individual consideration sets or chosen routes. Various calibration performance measures are used for that purpose, such as:

- the number of chosen routes or routes in the consideration sets reproduced exactly by the generation algorithm (strong criterion);
- spatial coverage measures expressing to what extent the observed routes fully or partly are covered by the generated set of routes (weak criterion).

It is tempting to perform the calibration as a constrained optimization problem according to sound statistical rules. However, in-depth analysis shows that the objective function is a complex non-continuous function of the unknown decision variables (parameters and variances). This is the more true for multi-modal networks. Therefore, by a trial-and-error procedure the default values have been adapted to maximize the first criterion, leaving a more sophisticated method for future research.

In this phase, we take into account the parameter values established in the demonstration of the MM-CSG approach in which a result of 78% of coverage of the complete route at leg level is obtained (see Section 7.5.5), and we try to adjust such values in order to improve the coverage percentage. Because of lack of sufficient data we restrict the calibration to a few key parameters while the remaining ones are determined based on literature and engineering judgement.

For the randomization of the link attributes X, such as travel time, waiting time, etc. the values to be determined are the expectation ($\mu(X)$), the variance ($\sigma(X)$), and the ranges within which the attribute value may vary. In this case, the mean value is determined by the free flow travel time, provided with the network as input, the ranges are adapted and calibrated, and the variances are based on these ranges.

For the randomization of the behavioural parameters β associated to travel time, such as travel time (for car, walk, bus, etc.), waiting time, etc. the values to be determined are the expectation ($\mu(\beta)$), the variance ($\sigma(\beta)$), and the ranges within which the attribute value may vary. In this case, both the mean values and the variances are adapted and calibrated, and the ranges are based on the variances.

The expected values of the parameters as well as their distribution types and variances are generally subject to calibration. Because of the multitude of attributes not all parameters are calibrated. Many of them are determined by engineering judgment supported by the travel choice behavior literature. Especially the parameters of the time attributes are considered relevant for calibration.

Table 7.10: Subset of calibrated parameters of time attributes (all uniformly distributed).

β parameters	**Mode**	$E(\beta)$	**Range**	**Min bound**	**Max bound**	**Coeff. of Variation**
PT in-vehicle time	Bus/Tram	1.0	0.4	0.6	1.4	23%
	Metro	0.8	0	-	-	-
	IC Train	1.0	0	-	-	-
	Expr/Stop Train	1.0	0.2	0.8	1.2	12%
PT board/alight time	All modes	1.0	0	-	-	-
PT waiting time	Bus/Tram	1.5	0.7	0.8	2.2	27%
	Metro	1.2	0.2	1.0	1.4	10%
	IC Train	1.2	0.2	1.0	1.4	10%
	Expr/Stop Train	0.9	0.3	0.6	1.2	19%
Walk time		0.6	0.2	0.4	0.8	19%
Car in-vehicle time	Secondary	1.0	0.4	0.6	1.4	23%
	Motorway	1.0	0	-	-	-
Car access/egress time	Secondary	1.0	0	-	-	-
Bicycle time		1.0	0.2	0.8	1.2	12%
Bicycle access/egress time		1.0	0.2	0.8	1.2	12%

The adopted strategy for the calibration is based on a trial-and-error procedure in which the following aspects are taking into account. We try to find a balance of access/egress modes to/from the stations in order to establish the beta parameters values in such a way that all transport modes are comparable within a certain distance interval (e.g. walk and bike should be comparable within 1,5 km). We have noticed that if the bounds of the distribution of the beta parameters are not set correctly, some travel modes, such as walk, or bus were not generated since are dominated by other faster modes, such as bike, or car.

Table 7.10 shows some of the calibration results achieved with observed consideration sets from a survey of multi-modal train trips in the Dordrecht-Rotterdam Corridor in the Netherlands.

The parameter variances (last column) look modest. The decisive variance level in the application however is the product of the variances of the parameters and the attribute values, which is much larger.

The ranges of the travel time distribution for walk and bike modes is set quite small (coefficient of variation (CoV) of 19% and 12% respectively as shown in Table 7-15); also for the urban PT metro a small range of the distributions of the travel and waiting time (CoV of 0% and 10% respectively) is considered since it is perceived more attractive than other travel modes due to the very high reliability; where the ranges of the distributions of travel time and waiting time of some urban PT (such as Bus and Tram) is set quite large (CoV of 23% and 27% respectively) since they have less reliability and actually they are perceived less attractive than other access and egress modes.

Interurban PT modes (Express and Stop Train) are assumed more reliable than urban PT (Bus and Tram), so we assume that they are a bit more attractive than other travel modes. Therefore there is less variation in travel and waiting time (CoV of 12% and 19% respectively).

Interurban PT modes (IC Train) has very high reliability and we assume that there is no variation in travel time and little variation in waiting time (CoV of 0% and 17% respectively) since it is perceived to be more attractive than other travel modes.

Note that variations in secondary road (for behavioral parameters of car mode) are taken into account in order to have this alternative comparable with other access modes (such a bike, walk, bus, etc.) but having variations in secondary roads lead also to the generation of uni-modal routes that are not relevant in this multi-modal context.

Since in peak hours roads are usually congested, we assume that there is a quite large variation in travel time for secondary roads, therefore a large range is considered for the beta parameters in travel time of car mode for secondary roads (23% coefficient of variation in Table 7.10).

It can be also assumed that in peak hours there is a similar congestion also in motorway roads, however since we are not interested in variation in uni-modal car alternatives, but especially in generating multi-modal routes, we have considered a null variation for the beta parameters of the travel time of car mode for motorway networks.

Regarding the range of the link parameters the following assumptions are made.

There is no variation for link attributes referring to private modes (such as car, bicycle), since we focus on the generating routes that are mainly multi-modal, and not variations of uni-modal alternatives. For link attributes associated to public modes (such as train, bus, tram etc.) more variation is taking into account. Large variations in link travel times for Bus mode, Express and Stop Train, medium variations for Tram, and IC Train, and small variations for Metro mode are considered. Variations in in-vehicle travel times are considered in order to make the model able to generate parallel lines, for example, bus lines that serve the same part of the trip.

In order to consider the difference in transferring among different PT travel modes also difference in variations are considered as follows: large variations in boarding link travel times for Bus mode, medium variations for Tram, Express and Stop Train, and small variations for Metro and IC Train.

Table 7.11: Subset of time attribute parameter values before the calibration.

β parameters	$E(\beta)$	Min bound	Max bound
PT in-vehicle time	1.0	0.3	1.7
PT board/alight time	1.5	0.5	2.5
PT waiting time	2.5	1.0	4.1
In-vehicle time car	1.0	0.3	1.7
Board/alight time car	1.0	0.3	1.7
In-vehicle time bicycle	1.0	0.3	1.7
Board/alight time bicycle	1.5	0.5	2.5

7.6.3 Results of the calibration

We have applied the MM-CSG with the calibrated parameters to the same multi-modal network and observed data described in Section 7.5.2 and Section 7.5.3; the observed data consist of 37 trips (OD pairs); however two of those trips (OD pairs) were excluded from the observed set since it appeared to be atypical individual travel behaviour. In one case the car mode is used to access the station at home-side even if the distance is less than 1.5 km. In the other case, the traveller chooses to use the local train even though the intercity train would bring them faster to their activity-end station. Apparently, there are some unaccounted benefits in using the local train service, such as maybe the seat availability. Therefore, 37 OD pairs are considered for the generation process, while a final observed set of 35 trips (OD pairs) is considered for the choice set comparison.

The calibration exercise succeeded in improving the generation performance significantly from a prediction success rate of the chosen routes (leg level) of 78% before the calibration to 86% after the calibration. Please note that in the demonstration phase 16 user classes are used whether in the calibration phase only 3 user classes are used, with much less iterations and computation time.

Table 7.11 shows the parameters of time attributes applied in the demonstration case before calibration. As we compare Table 7.11 with Table 7.10, we can notice that in general less variation of the parameters is considered after the calibration and that the range of the parameters of Table 7.10 is usually smaller than the range used in the analysis before the calibration.

Table 7.12: Multi-modal routes in Rotterdam-Dordrecht corridor generated by the MM-CSG approach.

Nr.	Modes			Time (min)
	Access	Main	Egress	
1		**Car**		38
2		**Car**		45
3		**Car**		46
4	Bike	**IC Train**	Walk	39
5	Walk	**IC Train**	Walk	49
6	Bus	**IC Train**	Tram	46
7	Walk	**Local Train**	Walk	64
8	Bike	**Local Train**	Walk	54
9	Walk	**Express Train**	Walk	58
10	Bike	**Express Train**	Walk	48
11	Walk	**Local Train**	Walk	54
12	Bike	**Local Train**	Walk	64

As an example Figure 7.7 shows the 12 multi-modal routes for an OD-pair in the Dordrecht-Rotterdam area; yellow links represent walk and bike links, car links are in red and train links are in black. Table 7.12 shows multi-modal choice set paths generated by the MM-CSG approach from Dordrecht to Rotterdam containing 12 paths.

Table 7.12 gives an example of the route prediction outcomes for multi-modal trips from Dordrecht-Southeast area to Rotterdam-Northwest area (about 30 km trip distance) achieved with the MM-CSG approach. The model has been applied to three user groups. In total, 12 uni-modal and multi-modal route alternatives were generated (see rows 1 to 12). The most attractive routes with different main mode having the shortest travel times (see column Time, which represents the total travel time in minutes) appear to be the car-only alternative (nr. 1) and the multi-modal alternative with Bike-IC Train-Walk (nr. 4).
The achieved prediction success rate shown above indicates that the performance quality of the doubly stochastic choice set generation approach successfully competes with the best approaches known so far.

7.7 Impact of stochasticity

This section deals with the impact of stochasticity introduced by the number of iterations as well as by the seed number. By systematic variation of the number of iterations we take into account that by increasing the number of iterations more routes might be generated with an increasing complexity.

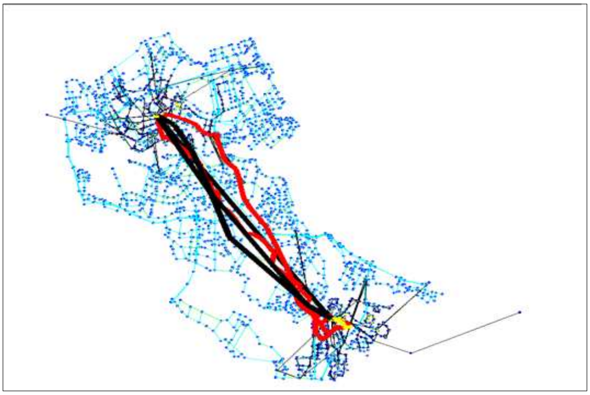


Figure 7.7: Multi-modal choice set paths generated by the MM-CSG approach for an OD-pair from Dordrecht to Rotterdam containing 12 paths; walk and bike links are in yellow, car links are in red and train links are in black.

We are interested in establishing how many draws are needed in order to generate a reasonable route set that maximizes the set coverage, i.e. to generate a predicted choice set that matches as much as possible the observed route set. We refer to Section 7.4.4 for general relationships between choice set size and composition on the one hand and randomisation properties on the other hand, that follow from the chosen way of stochastic route set generation.

This section deals also with the impact of the stochasticity by the random seed number, i.e. how the coverage performance may change as the seed number changes. It is well known that a random process is characterized by the fact that the results depend on the sequence of the drawn random numbers and, of course, on the initial seed number used for the randomization process. Therefore, it is obvious that the results of the generation process (the generated routes sets) will be different as we change the initial seed number. Now the questions are: how much different are the generated choice sets based on different seeds? How sensitive is the generation process for the seed number? By running the MM-CSG with different seeds we check some characteristics of the generated choice set, such as the choice set size, the resulting coverage, and the difference in generated routes.

We have carried out a series of tests with the same multi-modal network and observed data described in Section 7.5.2 and Section 7.5.3; the observed data consist of 37 trips (OD pairs) from which two trips (OD pairs) were excluded from the observed set since some travel modes of those trips were not successfully generated during the calibration phase. Therefore, 37 OD pairs are considered for the generation process, while a final observed set of 35 trips (OD pairs) is considered for the comparisons.

For the comparison of the two sets (generated and observed) we compute the percentage of observed routes included in the set of generated routes as introduced in Section 7.5.5. It is important to note that in the calibration only fixed parameter values are considered (no variations of the variances), whereas in the sensitivity analysis the randomized approach and randomized parameter values are taken into account.

To generate routes in the multi-modal network, the choice set generation algorithm described in Section 7.4 is adopted. Generalized cost functions and parameter values used in the randomization process are as established in the calibration phase (see Section 7.6).

In addition to the theoretical findings shown in Section 7.4.4, we show a number of empirical results of a sensitivity analysis of our generation method to some of the randomization inputs by applying our method to a set of observations on interregional trips in the multi-modal network of the Rotterdam-Dordrecht area in The Netherlands. The generated choice set sizes are generally very large (see Table 7.13) ranging from about 30 to 65. This is because of the following reasons. All routes are unique at the finest level of spatial detail (links) and may overlap considerably because no filtering has been applied yet (except the 100% overlap). In addition, in multi-modal networks, very many public transport alternatives exist because of the multiple modal combinations of the access, egress, and line haul trip parts.

Table 7.13: Average choice set size resulting from increasing numbers (n) of randomizations and 3 different random seeds $s1$ to $s3$ (same variances).

	$n_\beta = 10$	$n_\beta = 20$	$n_\beta = 30$
	$s1$ $s2$ $s3$	$s1$ $s2$ $s3$	$s1$ $s2$ $s3$
n_x=10	35.4 - 29.0 - 28.5	42.3 - 35.2- 33.6	45.9 - 39.2 - 37.3
n_x=15	39.4 - 34.0 - 32.4	46.9 - 41.3 - 38.1	51.2 - 45.7 - 41.9
n_x=20	44.0 - 40.2 - 36.5	52.1 - 49.4 - 43.6	57.2 - 54.5 - 47.9
n_x=25	46.9 - 42.1 - 39.0	54.8 - 51.8 - 46.2	60.1 - 56.9 - 50.9
n_x=30	50.1 - 44.8 - 42.2	58.2 - 55.5 - 50.2	64.1 - 61.0 - 55.3

N=37 observed OD trips, link level, calibrated parameters.

The impact of randomization factors (n_x, n_β and s) on average generated choice set sizes (before filtering) is given in Table 7.13 showing how the number of randomized draws (n_x and n_β from 10 to 30) from the link attribute distributions (rows) and from the preference parameter distributions (columns) influence the average generated size of the choice sets of 37 different trips. In each cell of the table the results achieved with three different seeds are given. The 3 choice set size values in the upper left cell refer to 100 random draws each (shortest path searches) whereas the rates in the lower right cell are based on 900 draws each. These choice set size predictions are stochastic outcomes. Since three user classes are considered in this analysis, the total number of iterations is given by $3 \cdot n_x \cdot n_\beta$.

Expectedly, the generated choice set sizes increase with the number of randomizations though with diminishing growth rates. A doubling of average choice set size emerges when increasing the 10x10 randomization to a 30x30 level (9 fold). At the same time the random seed value appears to have a clear influence which however diminishes with increasing numbers of randomizations. It appears that attribute randomization has a somewhat larger impact on changes in choice set sizes than parameter randomization.

Table 7.14 shows the impact of the number of randomized draws (from 10 to 30) from the link attribute distributions (rows) and from the preference parameter distributions (columns) on the prediction success rate (in percent correctly predicted modal sequences of chosen routes). In each cell of the table the results achieved with 3 different seeds are given. The 3 success rates in the upper left cell refer to 100 random draws each (shortest path searches) whereas the rates in the lower right cell are based on 900 draws each. In contrast to choice set sizes (Table 7.13), the spatial detail now is the so-called trip part level of routes concerning only the correct prediction of the sequence of different modes used in the chosen multi-modal trips. A multi-modal route consists of 3 trip parts at least. These success rates (ranging from 74% to 94%) are stochastic outcomes as well.

Table 7.14: Prediction success rates (in %) of chosen routes in the choice set for increasing number (n) of randomizations and 3 different seeds (same variances).

Parameters Attributes	$n_\beta = 10$ $s1$ $s2$ $s3$	$n_\beta = 15$ $s1$ $s2$ $s3$	$n_\beta = 20$ $s1$ $s2$ $s3$	$n_\beta = 25$ $s1$ $s2$ $s3$	$n_\beta = 30$ $s1$ $s2$ $s3$
n_x=10	86 - 74 - 77	86 - 77 - 80	86 - 80 - 80	86 - 80 - 83	86 - 80 - 86
n_x=15	86 - 80 - 83	86 - 83 - 83	86 - 86 - 83	89 - 86 - 86	89 - 86 - 89
n_x=20	89 - 83 - 86	89 - 86 - 86	89 - 89 - 86	91 - 89 - 89	91 - 89 - 89
n_x=25	89 - 83 - 86	89 - 86 - 86	89 - 89 - 86	91 - 89 - 89	91 - 89 - 89
n_x=30	89 - 83 - 86	91 - 86 - 86	91 - 89 - 86	94 - 89 - 89	94 - 89 - 89

N=35 observed OD trips, leg level, calibrated parameters.

In contrast to choice set size, prediction success rates appear less sensitive to the level of randomization. Already after 100 randomizations a fairly high level of success rate (about 80%) has been achieved. This is to be expected because the chosen route naturally is an attractive one that will be predicted early in the random generation process. Also in this case the impact of attribute randomization appears larger than parameter randomization. There is a systematic difference in results between the 3 seeds that diminishes with increasing randomization.

If instead of chosen routes observed consideration sets are used to measure the prediction success rates these appear expectedly to be somewhat lower while the sensitivity to the randomizations is similar.

Exploring with larger number of iterations confirm the observation of convergence shown in Table 7.14, without an increasing of the converge performance. The outcomes presented in Tables 7.13 and 7.14 suggest that further increasing the randomization might only slightly improve the results. In conformity with theoretical results (see Section 7.4.4), about 1000 randomizations per user group may give optimal results, although the choice set sizes could be in this case nearly doubled.

Finally, let us look at the impact of random seeds at the level of individual OD trips. For 35 of these trips (same data set from Rotterdam-Dordrecht region), we apply a calibrated (see Section 7.6) choice set generation procedure inclusive a few filtering steps such as on vehicle availability constraints. The resulting sets might e.g. be adopted for estimation of route choice models. We show the resulting choice set sizes of three randomizations (different seeds) for the case of 20 randomizations for attributes X and parameters β each.

Table 7.15: Comparative statistics of choice set generation with three random seed values (with calibrated parameters and after filtering) and $n_x = n_\beta = 20$.

N=35 OD pairs	Total unique routes generated (leg level)	Average choice set size	Minimum choice set size	Maximum choice set size	Prediction success rate (route level)
Seed 1	**816**	**23.3**	**7**	**47**	**89%**
Seed 2	**782**	**22.3**	**7**	**50**	**89%**
Seed 3	**723**	**20.7**	**7**	**47**	**86%**

Table 7.15 shows a few statistics on average outcomes demonstrating that averaged over the sample of OD pairs the outcomes are very close for the different randomizations.

However, more interesting are the individual outcomes such as given in Figure 7.8 showing that the generated choice set sizes for individual OD trips are very close, sometimes even identical, for the different seed values.

From these outcomes one may hypothesize that the compositions of the choice sets in terms of unique routes most probably are also very alike; and for a certain number of iterations the choice set size and the generated routes are almost similar to each other even if different seeds are used for the randomization process.

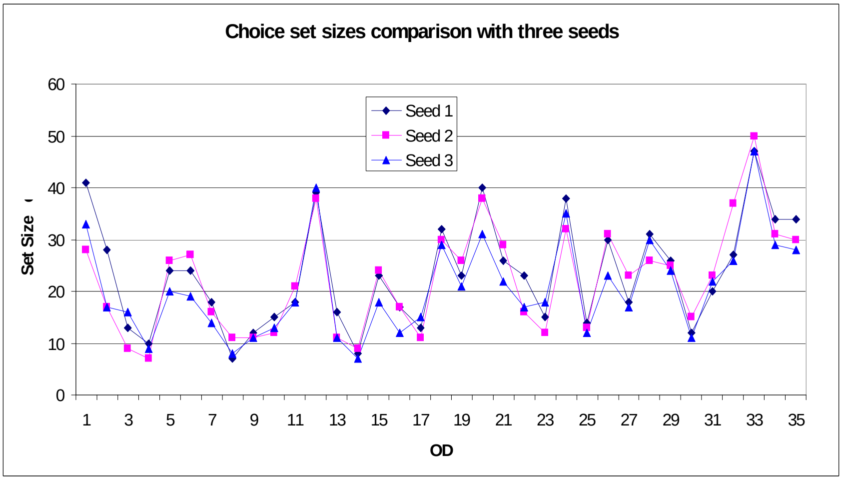


Figure 7.8: Similarity of patterns for the choice set size for each OD pair for different seeds.

The various outcomes shown above indicate that the performance quality of the doubly stochastic choice set generation approach is quite good and the analysis showed that the results are quite robust with respect to the seed number, while an even better performance might be achieved by increasing the number of iterations albeit at the costs of larger choice sets. See for example that the coverage performance has improved from the one obtained in the calibration phase (86% with 10-by-10 iterations) up to 89% for the 20-by-20 iterations with Seed 1 and Seed 2 (Table 7.15).

7.8 Characteristics of the generated choice set and filtering

In Section 6.6.4 we showed the importance of filtering to improve the generated choice sets in the uni-modal case.

In this section we analyse the predicted choice sets in order to establish the reasonability of the single multi-modal routes, and to check the quality of the choice set and how many of the requirements introduced in Section 7.2.2 are satisfied by the generated multi-modal choice sets. This is performed by checking the resulting choice sets vis--vis the quality criteria, mainly with the aim to determine if a filtering process is needed after the generation process in order to obtain a more accurate and adequate multi-modal choice set.

For this analysis we use the predicted choice sets generated with the stochastic MM-CSG approach with the 20-by-20 iterations resulting in a coverage performance of 89% (with Seed1 see Table 7.15).

In the preliminary analysis we focused on generating chosen routes applying only a filtering with regard to vehicle availability. In this section we will extend the filtering according to constraints introduced in Section 7.2.2 in order to improve the quality of these choice sets. A filtering process has been applied to eliminate alternatives that do not satisfy the criteria defined in Section 7.2.2.

The MM-CSG approach produces an exhaustive base route set that needs to be reduced in order to establish choice sets that efficiently suit a particular purpose, in our case the prediction purpose. To that end, a variety of selection constraints may be applied. For prediction purposes, highly overlapping paths may be removed, while at the same time spatial variety is welcomed. In multi-modal network applications, a rich variety in multi-modal composition of the choice set is desirable, while at the same time non-feasible modal sequences (e.g. bus-car-bike-walk) should be removed from the base set.

Since the nature of the generation process is stochastic, it is not recommended to do pair-wise comparisons between the generated routes nor to remove routes already during the generation process. In fact, because of the stochastic nature, the selection of the *best* route depends on the stochastic outcomes and one may risk to reject a good alternative and keep the worst one. Moreover, from a practical point of view it is much more complicated to implement this step during the generation process, whereas some criteria (such as the checks of the overlap, detour-max, and detour-min constraints) can easily be applied after the generation process, as shown in Chapter 6 (Figure 6.10).

Therefore, from a theoretical and practical point of view, we propose to apply the filtering process not during the generation, but rather after the generation process. The filtering process consists of all necessary checks to select the most reasonable multi-modal routes and to satisfy the multi-modal choice set criteria as much as possible. The filtering process should include all constraints needed to have feasible multi-modal routes and enough multi-modal variability in the choice set. The final generated choice set should satisfy the criteria such as the multi-modal feasibility and the multi-modal variability criteria; and include also the overlap, detour-max, detour-min constraints in order to have the spatial and preferential criteria satisfied (see Section 7.2.2).

Therefore, in a preliminary limited exercise the following checks are applied in the filtering process:

- the mode combination constraints;
- the reasonable distance constraints, introduced in Section 7.2.2 as feasibility conditions;
- the comparability criterion.

The mode combination constraints are applied to eliminate all routes with an unlikely combination of walk and bike, of car mode, such as Car-Train-Car, and PT modes such as Bike-Bus, or Train-Bike-Metro. The reasonable distance constraints are applied to eliminate all routes using walk as access mode at home side for more than 2 km, or bike mode at home for more than 5 km. Finally, the comparability criterion defined in Section 7.2.2 is applied within each user class and for the main mode. Please note that, for sake of simplicity, the overlap criterion has not been applied in the final filtering process because of the long trip.

Some of the characteristics of eliminated routes regard the use of some travel modes (e.g. bike) after having walked from the origin for a large distance (up to 2 km) or after having taken the train as main mode and the bus as egress mode in the destination area. Usually some travel modes (e.g. bike) should be available only at certain nodes (e.g. home, railway station) and not at every stop.

The main conclusion of this analysis is that also in the multi-modal case, similar as in the uni-modal case (see Section 6.6.4) the filtering is necessary. The results of the filtering process maintain the same good coverage of 89% while reducing the choice set sizes by about 20%.

7.9 Conclusions

Following a systematic approach, a set of quality criteria have been derived for assessing the adequacy of generated choice sets suitable for multi-modal networks. An overview is given of choice set generation approaches for multi-modal networks, presented in literature. A new choice set generation approach has been defined and a case study has been carried out to demonstrate that this model works already satisfactorily with the default parameter values. Three options for generating route sets have been studied and tested, i.e.: variation in the network attributes only, variation in traveller preferences only, and the combination of both. The latter case proved to yield by far the best match with observed route sets. Furthermore, the analyses show that variation in travellers' preferences is more important than variation in network attributes. The analysis revealed insights into the possibilities of generating realistic multi-modal route sets and it is proved that the randomization approach is able to provide good coverage of the observed routes. By far the best results are obtained by randomizing both network attributes and variation in traveller preferences.

The developed doubly stochastic choice set generation approach (MM-CSG) for multi-modal networks was shown to properly generate choice sets satisfying the most important defined criteria. Applications of such an algorithm to a real size network have been demonstrated to be feasible and adequate, namely specifically in the Rotterdam-Dordrecht region in The Netherlands. The generated choice sets show a large variety of uni-modal and multi-modal route alternatives. Especially the public transport alternatives appear to be frequent and are manifold. This is partly due to the good public transportation provision in the region and the multitude of different modes available to most travellers.

An extensive analysis of the probabilistic properties of this stochastic generation method has been performed (impacts of seeds, number of randomisations, adopted variance levels, etc) showing that satisfactory performance can be achieved with a limited number of randomisations if the adopted variances in preference parameters and attribute values are not too small. Theoretical and experimental results show that despite the stochastic principle of the method, its outcomes in terms of size and composition of generated choice sets are fairly stable already at modest numbers of randomization iterations. The repeated shortest path search principle makes the method computationally very efficient, which has been proven by various applications in very large networks.

The effectiveness of a filtering process applied to the choice sets generated by the MM-CSG approach has been demonstrated showing that the application of a filtering step after the generation process results in a reduction by about 20% of the choice set size, maintaining the same quality performance. We can conclude that in order to achieve adequate choice sets that satisfy most of the defined requirements, it is highly recommended to have a filtering process after the generation. Moreover, the achieved coverage at the level of complete individual trips is very high (about 90% in terms of number of trip parts covered) and is even higher for particular link types such as access legs of multi-modal train trips. Given the complexity of multi-modal routes this is a very satisfying outcome.

Chapter 8

Conclusions

The work presented in this thesis has dealt with the following main topics:

- Presenting main characteristics and definitions of multi-modal transportation, and setting up a new demand modelling approach for predicting multi-modal travel choices, including a new multi-modal transportation network representation (Chapter 2 and 3).
- Extending a conceptual framework for choice set notions from various viewpoints by introducing a clear choice set terminology and defining quality criteria for adequate choice sets in uni-modal and multi-modal networks (Chapter 4).
- Analyzing and comparing choice set generation approaches for uni-modal networks (Chapter 6).
- Developing and testing a new choice set generation approach for multi-modal networks (Chapter 7).

This concluding chapter starts with a short summary of the thesis' contents, after which the main achievements of this thesis are presented. Subsequently, the main conclusions with respect to the new multi-modal modelling approach and the new choice set generation approach are summarized, while the chapter finishes with recommendations and future research directions.

8.1 Summary of research

Multi-modal trips are a common travel phenomenon, which are expected to become more important in the future. The main characteristic of multi-modal transport is that more than one vehicular transport mode is used for a single trip and that transfers between travel modes are thus an essential element of multi-modal transport. In order to analyze multi-modal planning problems, such as the location and design of inter-modal transfer points, the demand modelling procedure needs to be capable of analyzing and predicting the use of multi-modal trips. However, multiple-choice dimensions involved in a multi-modal trip such as with respect to modes, services, transfer locations and routes are difficult to model.

After consideration of literature, it turned out that the classical network model architecture can be adapted to handle multi-modal transport modelling questions, when multimodal transport is cast as travel through a dedicated supernetwork in which mode and route choice are modelled simultaneously. The proposed approach to model multimodal travel choices is based on the assumption that a clear distinction in the route modelling between a choice set generation step and, conditional on that set, the genuine choice modelling step, may significantly improve choice analysis and prediction. Therefore, the route choice modelling is split into *choice set generation* and route choice modelling.

This thesis addresses the generation of choice sets of the travel alternatives for the purposes of route choice analysis and prediction of flows in multi-modal networks. In order to develop such choice set generation approach the following topics are dealt with. First of all, a conceptual framework for choice set notions has been developed by introducing a clear choice set terminology applicable to various analysis viewpoints. Secondly, this thesis has dealt with the notion of an adequate choice set, which kind of choice set is most suitable for prediction purposes, which requirements an adequate choice set should satisfy, and introducing the requirements for an appropriate choice set generation process for uni-modal networks. Quality criteria have been defined to assess the quality of the resulting choice sets and the quality of the generation process. An important step is the extension of these new criteria definitions for the uni-modal case to the multi-modal case, focusing on the differences between uni-modal and multimodal networks.

Finally, a large number of route generation methods proposed in literature for the unimodal networks have been analyzed in order to indicate which of these methods were potentially suitable for generating satisfactory route choice sets. A newly developed route choice set generation algorithm has been presented; the so-called doubly stochastic approach applicable to uni-modal and especially multi-modal networks, in which not only the link attributes but also the preference (or behavioural) parameters of the cost function are randomized, since especially in a multi-modal context, variations in transport modes and in multi-modal combinations are desirable in the generated choice sets. This doubly stochastic method has been tested with respect to its empirical and probabilistic properties, and applied to various networks, in particular to a multi-modal network.

8.2 Main achievements

The main scientific and practical contributions of this thesis concern the following subjects:

- The proposed multi-modal demand prediction approach, based on the supernetwork framework and on the a priori choice set generation, based on a new specification of the multi-modal transportation network representation.
- An original multi-view framework for choice set notions and concepts.
- Original criteria definitions for adequate choice sets to be used for route choice predictions in uni-modal and multi-modal networks.
- An innovative comparative analysis of existing choice set generation approaches.
- The new doubly-stochastic choice set generation approach for multi-modal networks.

8.2.1 Multi-modal modelling approach

In this thesis, a new approach for analyzing multi-modal travelling based on the use of a supernetwork methodology was introduced. In the supernetwork approach, the traditional steps of mode choice and route choice are integrated into a single route choice problem in a multi-modal network. Relevant characteristics of the modelling approach are that a modal split module is absent and a route choice set generation module prior to the route choice model is specified. By explicitly generating the individual choice alternatives prior to the choice modelling, maximum flexibility is available in adopting the most suitable choice modelling approach. Moreover, a new specification of the multi-modal network representation was presented, which takes into account several uni-modal network layers of all available travel modes combined into a single multi-modal network that includes transfer possibilities between modes, via waiting and walking links. A study area in The Netherlands (the Dordrecht-Rotterdam corridor) was represented in this new multi-modal representation and was successfully used in analyzing travel choices in this region.

8.2.2 Choice set notions and concepts

An important topic dealt with in this thesis concerns the systematic choice set definitions and classifications. Starting from a behaviour-based travel choice framework the concepts of *universal sets*, *objective choice sets*, *subjective choice set* and *consideration sets* were introduced, thereby distinguishing between *actual*, *observed*, and *generated* choice sets, and between choice set generation for *estimation* and *prediction* purposes. Furthermore, in transportation modelling it is not common to follow a completely individual-level approach, therefore we have distinguished between individual and group level. Also, the relationships between all different sets of alternatives have been established.

8.2.3 Criteria for adequate choice sets

New scientific insights on choice set concepts have been established with the new systematic definitions of a reasonable route, and adequate choice sets at individual and group level. The specification of requirements and choice set criteria for adequate route choice sets for prediction of route and link flows in uni-modal transport networks have been introduced in this thesis. These are based on sets of logical, feasibility, preference, and other constraints on the part of the individual traveller such as directness, maximum detour, maximum and minimum lengths of modal legs, mode availability, hierarchical logic, etc. At the group level of OD-flows criteria based on preferential and spatial variety are added. Also, the main characteristics of an appropriate generation process for prediction purpose in uni-modal networks have been established.

Based on the defined criteria, adequate choice sets at group level should contain routes that are not highly overlapping (due to the *overlap* criterion), are comparable in time or distance (due to the *comparability* criterion), without excessive detours, with appropriate choice set size and composition, with spatial variability (due to the *spatial variability* criterion) and with variation in route types (due to the *preferential variability* criterion); in the multi-modal case, also variation in multi-modal combinations and in transport modes used (routes should be multi-modal indeed for the *multi-modal variability* criterion).

8.2.4 Comparative analysis of choice set generation approaches

This thesis presents a generic systematic route set generation scheme for the purpose of characterizing and analysing current route set generation procedures. Based on this generic classification scheme, a systematic comparison of the various and sometimes very different route set generation methods known from literature has been established, after which each of these generation methods has been described in a structured comparable way as a basis for the evaluation of their adequacy for the purpose of choice set generation. This evaluation involved testing whether each of the methods might be able to fulfil the requirements for an adequate choice set (see subsection 8.2.3).

On the basis of this comparison, it appears that the simple repeated shortest path generation methods (such as, K shortest paths, etc.) perform unsatisfactory, whereas the best performance was achieved by the Constrained K-Shortest Paths for the single OD-pair generation and the Monte Carlo Labelling combination (MCL) approach for multiple OD pairs generation. It has also been found that some criteria for generating adequate choice sets were not satisfied by the MCL combination approach (for example, many overlapping routes were generated), since these criteria are not applied nor included during the generation process. In order to improve the quality of the generated choice sets it has been demonstrated that the application of a subsequent filtering process on the generated set, which includes set-based criteria such as maximum overlap, is necessary.

8.2.5 Choice set generation approach for multi-modal networks

The newly developed doubly stochastic approach is an extension of the MCL combination approach since not only the link attributes but also the preference (or behavioural) parameters of the cost function are randomized. A behaviour-based rationale has been formulated for this approach founded on two basic principles. Choice set formation on the individual level is highly preference driven meaning that the generation model should consider preferred route attributes. On the other hand the model considers preference differences among travellers. The basic property of the developed multi-modal choice set generation (MM-CSG) approach is that route choice factors and behavioural parameters of travellers determine the generation of route options for the choice sets via a utility or cost functions, called route generation functions. Since there exist significant differences in choice factors and preferences among user groups and trip purposes, but also in order to achieve a rich variety in the multi-modal composition of the choice sets, these route choice generation functions need to be specific for user groups and trip categories.

Therefore, the new method focuses on generating routes for prediction purposes at group level, in which *groups of travellers* are taken into account, and in particular, *user classes* that vary with respect to expected travel behaviour, for instance based on *trip purpose* (work, business, study, etc), and *traveller groups* having different modes available are distinguished. In fact, the randomization of the behavioural preference parameters provides variations in transport modes and in multi-modal combinations by taking into account the variations within the user classes, since in multi-modal network applications, a rich variety in multi-modal composition of the choice set is desirable.

In order to achieve adequate choice sets that satisfy most of the requirements defined, it is highly needed to have a filtering process after the generation. The MM-CSG approach produces an exhaustive base route set (master set) that needs to be freed from superfluous and non-efficient routes in order to arrive at choice sets that efficiently suit a particular purpose, in our case the prediction purpose. To that end, a variety of selection constraints may be additionally applied. For prediction purposes, highly overlapping paths may be removed, while at the same time spatial variety is welcomed.

The feasibility of the MM-CSG approach has been demonstrated by applying the method to the multi-modal network of the Rotterdam-Dordrecht Region in The Netherlands. In order to prove the effectiveness of the MM-CSG method, the generated choice sets were compared with observed chosen routes and reported choice sets in this multi-modal network.

An extensive analysis of the probabilistic properties of this stochastic generation method has been performed (impacts of seeds, number of randomisations, adopted variance levels, etc) showing that satisfactory performance can be achieved with a limited number of randomisations if the adopted variances in preference parameters and attribute values are not too small.

8.3 Main conclusions

This section summarizes the most important conclusions with respect to multi-modal modelling approaches, the comparison of choice set generation methods applicable to uni-modal networks, and the multi-modal choice set generation approach.

The model architecture with the supernetwork approach includes components for the generation of the (route) choice-set, route-choice model, and traffic assignment. Although several of these components have been presented elsewhere, it has been shown how these can be adapted to a multi-modal situation, and integrated into a coherent modelling framework. The main characteristic of this approach is the presence of a priori choice set generation step followed by the route choice modelling step. For demand prediction, explicit generation of route choice sets has a number of advantages. A priori given choice sets allow nearly unlimited flexibility and freedom in route choice models to be adopted. In addition to theoretical advantages, they offer significant computation advantages in iterative demand calculation procedures since repeated optimal path search no longer is necessary.

Based on the systematic comparison of (shortest path based) choice set generation approaches for uni-modal networks, according to which no method meets all requirements, the feasibility of the Monte Carlo Labelling (MCL) combination method, if extended with a filter process applied after the generation process has been demonstrated by several applications to two uni-modal networks, namely the Dutch national road and waterway networks respectively. The effectiveness of a filtering process applied to the choice sets generated by the MCL approach has been demonstrated showing that application of a filtering step after the generation process results in a reduction of 73% of the choice set size. We can conclude that the Monte Carlo Labelling combination approach is the one to be preferred from the other methods in order to generate an adequate choice set at group level and for prediction purposes.

The newly developed doubly stochastic approach is an extension of the MCL combination approach in which not only the link attributes but also the preference (or behavioural) parameters of the cost function are randomized.
In fact, it appears to be an efficient technique that has been successfully applied in order to achieve variations in transport modes and in multi-modal combinations, which are desirable in the generated choice sets, especially in a multi-modal context.
This variation can be achieved by distinguishing different user classes with different parameters and by randomization of these behavioural parameters to take the variations within the user classes into account.
Therefore, we stress the importance of varying the behavioural parameters, which is very essential for generating choice sets in a multi-modal context.

The various analyses performed with the doubly stochastic choice set generation approach have shown that the resulting choice sets show the desired properties in terms of choice set size (minimum and maximum size) and composition (variety in trip properties).

The doubly stochastic choice set generation approach has been applied in real-world multi-modal networks, namely in the Rotterdam-Dordrecht region in The Netherlands. The generated choice sets show a adequate variety of uni-modal and multi-modal route alternatives. Especially the public transport alternatives appear to be frequent and manifold. This is partly due to the good public transportation provision in the region and the multitude of different modes available to most travellers.

The validation study used an observation set of 35 trips (35 different OD-pairs) with chosen routes and reported consideration sets) from the Rotterdam-Dordrecht region in the Randstad, The Netherlands. The face validity (plausibility of generated alternatives) and empirical validity (conformity with observations) are very high; in nearly all cases the reported chosen alternatives are member of the generated choice set. Moreover, in a number of simulations the prediction success rates (predicting the correct sequence of legs in the chosen multi-modal trip, which means among others correct entry, exit and transfer stations, correct train type and correct sequence of public and private mode) was 89% or higher. In the multi-modal case, a filtering process has been applied as well resulting in a reduction by about 20% of the choice set size, while maintaining the same quality performance.

8.4 Recommendations and future research

The findings and conclusions following from the analyses presented in this thesis lead to the following recommendations regarding multi-modal modelling approaches as well as choice set generation methods.

In our analysis it has been assumed that a priori choice set generation is highly recommended in a modelling approach, especially in iterative network assignment. It has also been shown that a priori generation is a feasible approach for realistic multi-modal networks, and through this approach realistic choice sets can be established.

In this thesis the notion of choice set in complex situation, such as in multi-modal networks, has been analysed more from a psychological point of view; however, some behavioural analysis as well might also be required and investigated in future research. In addition, what implications does the new information technology have on choice set generation might be a relevant issue for future work.

The developed doubly stochastic approach by far outperforms the singly stochastic approaches (only randomizing attributes). It is important to note the importance of varying the behavioural parameters especially in generating choice sets in a multi-modal context, because it provides the expected variations in the generated multi-modal routes, especially in a multi-modal context. Note that it is also highly recommended to apply a filtering process on the generated set just after the generation process in order to remove superfluous routes and obtain more adequate choice sets that satisfy the quality criteria.

In the doubly stochastic approach, the adopted initial variances in network attribute values (travel time, waiting time, travel cost, etc.) and the adopted variances in the related parameter values, which reflect the variation in the preferences for these attributes within the population of travellers, are set to specific values and kept constant during the generation process. An alternative approach, which is an extension of the doubly stochastic approach, is the accelerated MCL approach in which the variances are increased in order to speed-up the generation process, might also be developed in future research.

The calibration and validation exercises carried out in this thesis research focused on the multi-modal choice set generation approach. Although some attempts have been done on the calibration and validation of both route generation and multi-modal network assignment (see Carlier et al. (2005)), some research questions have not been answered yet, such as:

- The calibration and validation of the combined multi-modal choice set generation model and multi-modal network assignment.
- The determination of the route and link flows in the multi-modal networks.
- The analysis of which types of single uni-modal and multi-modal routes, and combinations of them, are determined during the iterative process.
- How much the type of routes found during the generation process are useful for the assignment step.
- The application of the supernetwork approach to congested situations.

Finally, due to practical reasons, in this thesis we have as a first step developed a static model meaning that the time dimension has not been considered in the modelling approach. However, transport is a dynamic phenomenon that cannot be completely represented by a static model. Further research on this topic therefore should consider the development of these tools suitable for a dynamic modelling approach.

Bibliography

Abdelghany, K.F.S. (2001). *Stochastic dynamic traffic assignment for intermodal transportation networks with consistent information supply strategies.* Ph.D. Thesis, University of Texas.

Abdelghany, K.F.S. and Mahmassani, H.S. (1999). Multi-objective shortest path algorithm for large scale inter-modal networks. *Presented at the INFORMS Fall annual meeting.*

Abdelghany, K.F.S. and Mahmassani, H.S. (2001). Dynamic trip assignment-simulation model for inter-modal transportation networks. *Transportation Research Record*, *1771*, 52-60.

AGV Adviesgroup voor verkeer en vervoer BV. (1990). *Reroutingmogelijkheden op het hoofdwegennet* (Report). Rotterdam: Rijkswaterstaat Dienst Verkeerskunde.

Akgün, A., Erkut, E., & Batta, R. (2000). On finding dissimilar paths. *European Journal of Operation Research*, *121*, 232-246.

Azevedo, J.A. and Costa, M.E.O.S. and Maderia, J.J.E.R.S. and Martins, E.Q.V. (1993). An algorithm for the ranking of shortest paths. *European Journal of Operation Research*, *69*, 97-106.

Batley, R., Fowkes, T., Watling, D., Whelan, G., Daly, A., & Hato, E. (2001). Models for analysing route choice. In *European Transport Conference (cd-rom).* Cambridge, UK.

Bell, M.G.H. (2003). Capacity constrained transit assignment models and reliability analysis. In W. Lam & M. Bell (Eds.), *Advanced Modeling for Transit Operations and Service Planning* (p. 181-199). Amsterdam: Pergamon.

Bellman, R., & Kabala, R. (1968). On k^{th} best policies. *Journal of the Society for Industrial and Applied Mathematics*, *10*, 582-588.

Ben-Akiva, M.E. and Bergman, M.J. and Daly, A.J. and Ramaswamy, R. (1984). Modeling inter-urban route choice behaviour. In J. Volmuller & R. E. Hamerslag (Eds.), *Proceedings Ninth International Symposium on transportation and traffic theory* (p. 299-330). Utrecht, Nederlands: VNU Science Press.

Ben-Akiva, M.E. and Bierlaire, M. (1999). Discrete choice methods and their applications to short term decisions. In R. Hall (Ed.), *Handbook of Transportation Science* (p. 5-33). Boston/Dordrecht/London: Kluwer Academic Publishers.

Ben-Akiva, M.E. and Bolduc, D. (1996). Multinomial probit with a logit kernel and a general parametric specification of the covariance structure. In *Presented at the* 3^{rd} *Interntional Choice Simposium.* Columbia University.

Ben-Akiva, M.E. and Lerman, S. (1985). *Discrete choice analysis: theory and application to travel demand* (Tech. Rep.). Cambridge, Massachusetts: The MIT Press.

Benjamins, M. (2001). *Assigning trips to multimodal networks by using hypernetworks.* MSc Thesis (in Dutch), Delft University of Technology.

Benjamins, M., Lindveld, C., & Van Nes, R. (2002). Multimodal travel choice modelling: a supernetwork approach. In *Proceedings* 81^{st} *TRB Annual Meeting.* Washington DC. (Paper 03-2948 on CD-ROM)

Benshoof, J.A. (1970, April). Characteristics of drivers' route selection behaviour. *Traffic Engineering and Control*, *11*(12), 604-609.

Bliemer, M.J.C. and Taale, H. (2006). Route generation and dynamic traffic assignment for large networks. In *Paper DTA Conference of Leeds.*

Bliemer, M.J.C. and Versteegt, E. and Castenmiller, R. (2004). Indy, a new analytical multiclass dynamic traffic assignment model. In *Proceedings of the Fifth Triennal Symposium on Transportation Analysis TRISTAN V Conference CD-ROM.* Guadeloupe.

Borgers, A. and Timmermans, H.J.P. (1984). *Multi-purpose trips en routekeuze van voetgangers in binnensteden: theorie, modelvorming en empirische toetsing* (Tech. Rep.). Afs Bouwkunde, Endhoven University.

Bovy, P.H.L. (1988). *Congestion insensitive logit-based assignment methods. (Nietcongestiegevoelige toedelingsmethoden van het logit-modeltype)* (OSPA Werkstuk (in Dutch) No. 11). Delft University of Technology.

Bovy, P.H.L. (1990). *Assignment to congestion-free networks (Toedeling in congestievrije netwerken).* Ph.D. Thesis (in Dutch), Delft University of Technology, Rotterdam, DVK, The Netherlands.

Bovy, P.H.L. (2003). Multimodaal personenvervoer: soms marginaal, maar essentieel. *Special Issue of Tijdschrift Vervoerswetenschap*, *39*(1), 14-19.

Bovy, P.H.L. (2007). Modeling route choice sets in transportation networks: a preliminary synthesis. In *Proceedings of the TRISTAN VI Conference.* Phuket, Thailand.

Bovy, P.H.L. and Fiorenzo-Catalano, S. (2006). Stochastic route choice set generation: behavioural and probabilistic foundations. In *IATRB (Ed) Proceedings* 11^{th} *International Conference on Travel Beaviour Research.* Kyoto.

Bovy, P.H.L. and Stern, E. (1990). *Route choice: wayfinding in transport networks.* Kluwer Academic Publishers.

Bradley, M.A. and Bovy, P.H.L. (1984). A stated preference analysis of bicyclist route choice. In *Proceedings - PTRC Annual Meeting* (p. 39-53). London, UK.

Caramia, M., & Storchi, G. (1997). *Multimodal shortest hyperpaths on transportation networks* (Technical Paper 25). University of Rome, La Sapienza.

Carlier, K., Fiorenzo-Catalano, S., Schrijver, J., & Van Nes, R. (2005). TRANSFER: a new equilibrium model for analysing multimodal passenger trips. In *Proceedings ETC.* Strasbourg.

Carlier, K. and Fiorenzo-Catalano, S. and Lindveld, C. and Bovy, P.H.L. (2003). A supernetwork approach towards multimodal travel modelling. In *TRB (Ed) Proceedings* 82nd *Annual Meeting of the Transportation Research Board.* Washington DC.

Carotenuto, P., Galiano, G., Giordani, S., & Riciardelli, S. (2004). Finding spatial dissimilar paths for hazmat shipments. In *Proceedings of the Fifth Triennal Symposium on Transportation Analysis TRISTAN V Conference CD-ROM.* Guadeloupe.

Cascetta, E. (2001). *Transportation systems engineering: Theory and methods.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Cascetta, E., Nuzzolo, A., Russo, F., & Vitetta, A. (1996). A modified logit route choice model overcoming path overlapping problems: Specification and some calibration results for interurban networks. In J. Lesort (Ed.), *Transportation and Traffic Theory.* Lyon, France: Elsevier.

Cascetta, E., & Papola, A. (2001). Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand. *Transportation Research Part C*, *9C*(4), 249-263.

Cascetta, E. and Russo, F. and Viola, F.A. and Vitetta, A. (2002). A model of route perception in urban road networks. *Transportation Research Part B*, *36*, 577-592.

Catalano, S., Lindveld, C., & Carlier, K. (2001). *A hypernetwork approach towards multimodal transportation analysis. Report 1: literature survey* (Report No. VK 2001-011). Delft University of Technology.

Cirillo, C. and Axhausen, K.W. (2002). Mode choice in complex tours. In *Proceedings of the European Transport Conference (cd-rom).* Cambridge, UK.

Cohn, N., Daly, A., Rohr, C., Dam, A., Oosterwijk, W., & Van Der Star, T. (1996). PROMISE: policy-sensitive passenger forecasting for the Netherlands railways. In *Proceedings of the PTRC Conference.*

Daamen, W. (2004). *Modelling passenger flows in public transport facilities.* Ph.D. Thesis, TRAIL Thesis Series no.T2004/6, TRAIL Research School, Delft University of Technology, Delft University Press, Delft, The Netherlands.

Daamen, W. and Bovy, P.H.L. and Hoogendoorn, S.P. and Van De Reijt, A. (2005). Passenger route choice concerning level changes in railway stations. In *TRB Proceedings* 84th *Annual Meeting of the Transportation Research Board.* Washington DC. (Compendium of Papers (CD-ROM))

Damberg, O. and Lundgren, J.T. and Patriksson, M. (1996). An algorithm for the stochastic user equilibrium problem. *Transportation Research Part B*, *2*, 115-131.

De La Barra, T. and Perez, B. and Anez, J. (1993). Multidimensional path search and assignment. In *Proceedings of the* 21st *PTRC Summer Meeting.*

De Cea, J., & Fernandez, E. (1993). Transit assignment for congested public transport system: an equilibrium model. *Transportation Science*, *27*(2), 133-147.

Dell'Olmo, P., Gentili, M., & Scozzari, A. (2002). Finding dissimilar routes for the transportation of hazardous materials. In *EURO conference.* Bari.

DEMIS. (2001). Delft. (Description of the NETTER software. http://www.demis.nl/DEMIS_UK/Products/Netter32.htm)

Dial, R.B. (1967). Transit pathfinder algorithm. *Highway Research Record*, *205*, 67-85.

Dial, R.B. (1971). A probabilistic traffic assignment model which obviates path enumeration. *Transportation Research*, *5*(2), 83-122.

Dial, R.B. (2000). Bicriterion traffic equilibrium: T2 mode, algorithm, and software overview. *Transportation Research Record*, *1725*, 54-62.

Fearnside, K., & Draper, D. (1971). Public transport assignment - a new approach. *Traffic Engineering and Control*, *12*, 298-299.

Fernandez, E., De Cea, J., Florian, M., & Cabrera, E. (1994). Network equilibrium models with combined modes. *Transportation Science*, *28*(3), 182-192.

Fiorenzo-Catalano, S. and Hoogendoorn-Lanser, S. and Van Nes, R. (2003). Choice set composition modelling in multi-modal travelling. In *IATRB (Ed) Proceedings* 10th *International Conference on Travel Beaviour Research.* Lucerne.

Fiorenzo-Catalano, S. and Van Der Zijpp, N.J. (2001). A forecasting model for inland navigation based on route enumeration. In *Proceedings of the European Transport Conference (cd-rom).* Cambridge, UK.

Fiorenzo-Catalano, S. and Van Nes, R. and Bovy, P.H.L. (2004a). Choice set generation for multi-modal travel analysis. *European Journal of Transport and Infrastructure Research*, *4*(2), 195-209.

Fiorenzo-Catalano, S. and Van Nes, R. and Bovy, P.H.L. (2004b). Choice set generation for multi-modal travel analysis. In *TRB (Ed) Proceedings* 83rd *Annual Meeting of the Transportation Research Board.* Washington DC.

Florian, M. (2004). Finding shortest time-dependent paths in schedule-based transit networks: a label setting algorithm. In Wilson, N.H.M. and Nuzzolo, A. (Ed.), *Schedule-Based Dynamic Ttransit Modelling: Theory and Applications* (p. 43-52). Kluwer Academic Publishers.

Florian, M., Wu, J., & He, S. (2000). A multi-class multi-mode variable demand network equilibrium model with hierarchical logit structures. In *Proceedings of the PTRC Conference* (p. 211-224).

Friedrich, M., Hofsass, I., & Wekeck, S. (2001). Time-table based transit assignment using branch and bound. *Transportation Research Record*, *1752*, 100-107.

Gallo, G., Longo, G., Nguyen, S., & Pallottino, S. (1993). Direct hypergraphs and applications. *Discrete Applied Mathematics*, *42*(2-3), 177-201.

Gliebe, J.P. and Koppelman, F.S. and Ziliaskopoulos, A. (1998). Route choice using a paired combinatorial logit model. In *TRB Proceedings* 78th *Annual Meeting of the Transportation Research Board.* Washington DC. (Compendium of Papers (CD-ROM))

Golledge, R.G. (1995). Path selection and route preference in human navigation: a progress report. In *Proceedings International Conference COSIT '95* (p. 21-23). Semmering, Austria.

Golledge, R.G. (1997). Defining the criteria used in path selection. In Ettema, D.

and Timmermans, H.J.P. (Ed.), *Activity-based approaches to travel analysis* (p. 152-169). Elsevier Science/Pergamon, Oxford.

Golledge, R.G. and Dougherty, V. and Bell, S. (1995). Acquiring spatial knowledge: survey versus route-based knoeledge in unfamiliar environments. *Annals of the Association of American Geographers*, *85*(1), 134-158.

Hoogendoorn-Lanser, S. and Bovy, P.H.L., and Van Nes, R. (2007). Application of rule-based approach to multi-modal choice set generation. In *TRB (Ed) Proceedings* 86th *Annual Meeting of the Transportation Research Board.* Washington DC.

Hoogendoorn-Lanser, S. and Van Nes, R. and Bovy, P.H.L. (2005). Path size and overlap in multimodal transport networks: a new interpretation. In H. Mahmassani (Ed.), *Transportation and Traffic Theory: Flow, Dynamics and Human Interaction* (p. 63-84). Elsevier.

Hoogendoorn-Lanser, S. and Van Nes, R., and Bovy, P.H.L. (2006). A rule-based approach to multi-modal choice set generation. In *IATRB (Ed) Proceedings* 11th *International Conference on Travel Beaviour Research.* Kyoto.

Hoogendoorn-Lanser, S. (2005). *Modelling travel behaviour in multi-modal networks.* Ph.D. Thesis, TRAIL Thesis Series T2005/4, TRAIL Research School, Delft University of Technology, Delft University Press, Delft, The Netherlands.

Horn, M.E.T. (2002). An extended model and procedural framework for planning multi-modal passenger journeys. *Transportation Research Part B*, *37*(7), 641-660.

Hunt, D.T. and Kornhauser, A.L. (1997). Assigning traffic over essentially-least-cost paths. *Transportation Research Record*, *1556*, 1-7.

Johnson, P.E. and Joy, D.S. and Clarke, D.B. and Jacobi, J.M. (1992). *Highway 3.01, an enhanced highway routing model: program, description, methodology, and revised user's manual* (No. ORNL/TM-12124). Oak Ridge, TN.

Jourquini, B., Beuthei, M., Geertsii, J., & KoulaNdjang'Hai, C. (1999). Intermodality and substitution of modes for freight transportation: computation of price-elasticities through a geographic multimodal transportation network analysis. In *Proceedings of the NECTAR Conference.* Delft, The Netherlands.

Koppelman, F.S. and Wen, C. (2000). The paired combinatorial logit model: properties estimation and application. *Transportation Research Part B*, *34*, 75-89.

Lam, S. (2002). Multimodal transit modeling in Singapore. In *Workshop proceedings of the Croucher ASI on Advanced Modeling for Transit Operations and Service Planning.* Hong Kong.

Lawler, E.L. (1976). *Combinatorial optimization: Networks and matroids* (Holt, Ed.). New York: Rinehart & Winston.

Li, H. and Bovy, P.H.L. and Hooghiemstra, G. (2007). *Probabilistic properties of stochastic choice set generation* (TRAIL Studies in Transportation Science). Delft University of Technology.

Lim, Y. and Heydecker, B.G. (2005). Dynamic departure time and stochastic user equilibrium problem. *Transportation Research Part B*, *39*, 97-118.

Lo, H.K. and Yip, C.W. and Wan, K.H. (2003). Modeling competitive multi-modal services. In Lam, W.H.K. and Bell, M.G.H. (Ed.), *Advanced Modeling for Transit Operations and Service Planning* (p. 231-255). Amsterdam: Pergamon.

Lombard, K. and Church, R.L. (1993). The gateway shortest path problem: generating alternative routes for a corridor location problem. *Geographical Systems*, *1*, 25-45.

Los Angeles National Laboratory. (2001). *Transims 2.1 documentation.* Los Angeles. (http://transims.tsasa.lanl.gov/)

Loureiro, C.F.G. and Ralston, B. (1996). Investment selection model for multicommodity multimodal transportation networks. *Transportation Research Record*, *1522*, 38-46.

Louviere, J.J. and Hensher, D.A. and Swait, J.D. (2003). *Stated choice methods: analysis and application*. Cambridge University Press.

Lozano, A., & Storchi, G. (2001). Shortest viable path algorithm in multimodal networks. *Transportation Research Part A*, *35*, 225-241.

Mabit, S.L. and Nielsen, O.A. (2006). The effect of correlated value of travel time savings in public transport assignment. In *Proceedings ETC.* Strasbourg.

Manski, C. (1977). The structure of random utility models. *Theory and Decision*, *8*, 229-254.

Marcotte, P., & Nguyen, S. (1998). Hyperpath formulations for traffic assignment problems. In P. Marcotte & S. Nguyen (Eds.), *Equilibrium and advanced transportation modelling.* Kluwer Academic Publishers.

Martins, E.Q.V. (1984). An algorithm for ranking paths that may contain cycles. *European Journal of Operation Research*, *18*, 123-130.

Ministrie van Verkeer en Waterstaat. (2002). *De markt voor multi-modaal personenvervoer. onderzoek naar de markt- en beleidspotentie van multimodaal personenvervoer* (Tech. Rep.). Rijkswaterstaat.

Montella, B., Gallo, M., & D'Acierno, L. (1999). Multimodal network design problems. In L. Sucharov (Ed.), *Urban transport V - Urban Transport and the Environment for the* 21^{st} *Century* (p. 405-414). Southampton, Boston: WIT PRESS.

Morikawa, T. (1995). A hybrid probabilistic choice set model with compensatory and noncompensatory choice rules. In *WCTR Proceedings of the* 7^{th} *World Conference on Transport Research.* Sydney.

Nagurney, A., & Dong, J. (2002). *Supernetworks: descision making for the information age*. Cheltanham UK, Northhampton, MA, USA: Edward Elgar.

Nguyen, S., & Pallottino, S. (1988). Equilibrium trffic assgnment for large scale transit network. *European Journal Operational Research*, *37*(2), 176-186.

Nielsen, O.A. (1996). Do stochastic traffic assignment models consider differences in road users' utility functions? In *Proceedings of the* 24^{th} *PTRC Annual Meeting, seminar D & E - Part 2 Transportation Planning Methods.* Cambridge, Great Britain.

Nielsen, O.A. (2000). A stochastic assignment model considering differences in passengers utility functions. *Transportation Research Part B*, *34*(5), 377-402.

Nielsen, O.A. (2002). Timetable-based transit assignment with error components in the utility functions. In *Workshop proceedings of the Croucher ASI on Advanced Modeling for Transit Operations and Service Planning.* Hong Kong.

Nielsen, O.A. and Frederiksen, R.D. (2001). Rule-based, object-oriented modelling of public transport systems.

Nielsen, O.A. and Frederiksen, R.D. (2006). Optimisation of timetable-based, stochastic transit assignment models based on MSA. *Annals of Operations Research*, *144*, 263285.

Nuzzolo, A. (2003). Transit path choice and assignment model approaches. In Lam, W.H.K. and Bell, M.G.H. (Ed.), *Advanced Modeling for Transit Operations and Service Planning* (p. 93-124). Amsterdam: Pergamon.

Nuzzolo, A., & Crisalli, U. (2004). The scheduled-based approach and dynamic transit modelling. In Wilson, N.H.M. and Nuzzolo, A. (Ed.), *Schedule-Based Dynamic Ttransit Modelling: Theory and Applications* (p. 1-24). Kluwer Academic Publishers.

Nuzzolo, A. and Russo, F. (1994). An equilibrium assignment model for intercity transit networks. In *Proceedings of the TRISTAN III Conference.* San Juan, Puerto Rico.

Oark, M. and Regan, A. (2005). Capacity modelling in transportation. A multimodal perspective. *Transportation Research Record*, *1906*, 97-104.

Ortuzar, J.D. and Willumsen, L.G. (2001). *Modelling transport* (3^{rd} ed.). Chichester, Great Britain: John Wiley & Sons.

Park, D. and Rilett, L.R. (1997). Identifying multiple and reasonable paths in transportation networks: A heuristic approach. *Transportation Research Record*, *1607*, 31-37.

Pedersen, M.B. (2005). *Optimization models and solution methods for intermodal transportation.* Ph.D. Thesis, Technical University of Denmark, Denmark.

Prato, C.G. and Bekhor, S. (2006). Applying branch & bound technique to route choice set generation. In *TRB Proceedings* 85^{th} *Annual Meeting of the Transportation Research Board.* Washington DC. (Compendium of Papers (CD-ROM))

Ramming, M.S. (2002). *Network knowledge and route choice.* Ph.D. Thesis, MIT, Cambridge, Massachusetts.

Schnabel, W., & Lohse, D. (1997). Verkehrsplanung [Band 2]. In *Grundlagen der Strasssen Verkehrstechnik und Verkehrsplanung* (p. 287-293). Verlag für Bauwesen, Berlin.

Scott, K., & Bernestein, D. (2001). *Finding alternatives to the best path* (Report). New Jersey TIDE Center.

Scott, K., Pab*ó*n-Jim*é*nez, G., & Bernstein, D. (1997). Finding alternatives to the best path. In *TRB Proceedings* 76^{th} *Annual Meeting of the Transportation Research Board.* Washington DC.

Senevirante, P.N. and Morrall, J.F. (1986). Analysis of factors affecting the choice of route of pedestrians. *Transportation Planning and Technology*, *10*, 147-159.

Sheffi, Y. (1985). *Urban transportation networks: equilibrium analysis with mathe-*

matical programming methods (P. Hall, Ed.). Prentice Hall.

Sheffi, Y. and Powell, W.B. (1982). An Algorithm for the Equilibrium Assignment Problem with Random Link Times. *Networks*, *12*(2), 191-207.

Simon, H.A. (1990). Invariants of human behaviour. *Annual Review of Psychology*, *41*, 1-19.

Spiess, H., & Florian, M. (1989). Optimal strategies: A new assignment model for transit networks. *Transportation Research PPart B*, *23*, 83-102.

Swait, J. (2001). Choice set generation within the generalized extreme value family of discrete choice models. *Transportation Research Part B*, *35*, 643-666.

Swait, J. and Ben-Akiva, M.E. (1985). Analysis of the effects of captivity on travel time and cost elasticities. In *Behavioural Research for Transport Policy, Proceedings* (p. 119-134).

Swait, J. and Ben-Akiva, M.E. (1987). Incorporating random constraints in discrete models of choice set generation. *Transportation Research Part B*, *21B*, 91-102.

Train, K. (2003). *Discrete choice methods with simulation.* Cambridge University Press.

Tversky, A. (1972). Elimination by aspects: a theory of choice. *Psychological Review*, *79*, 281-299.

Van Der Waerden, P. and Borgers, A. and Timmermans, H.J.P. (2004). Choice set composition in the context of pedestrians' route choice modelling. In *TRB Proceedings* 83^{th} *Annual Meeting of the Transportation Research Board.* Washington DC. (Compendium of Papers (CD-ROM))

Van Der Zijpp, N.J. and Fiorenzo-Catalano, S. (2002). An algorithm for finding the k-acyclic shortest paths with logical constraints and its application to the route enumeration problem. In *TRB Proceedings* 81^{st} *Annual Meeting of the Transportation Research Board* (p. 1-13). Washington DC. (Compendium of Papers (CD-ROM) Paper 02-3981)

Van Der Zijpp, N.J. and Fiorenzo-Catalano, S. (2005). Path enumeration by finding the constrained k-shortest paths. *Transportation Research Part B*, *39*, 545-563.

Van De Reijt, A. (2004). *Route choice modelling of pedestrians in railway stations (routekeuzemodel voor voetgangers in stations).* MSc Thesis (in Dutch), Delft University of Technology.

Van Der Waard, J. (1988a). The relative importance of public transport trips - time attributes in route choice. In *Proceedings of the PTRC Annual Meeting.* London, UK.

Van Der Waard, J. (1988b). *The relative importance of public transport trip-time attribute in route choice* (Tech. Rep.). VK Report TU Delft.

Van Nes, R. (2002). *Design of multi-modal transport networks: A hierarchical approach.* Ph.D. Thesis, TRAIL Thesis Series T2002/5, TRAIL Research School, Delft University of Technology, Delft University Press, Delft, The Netherlands.

Vrtic, M. (2003). *Simultanes Routen- und Verkehrsmittelwahlmodell.* Ph.D. Thesis, Technical University of Dresden, Dresden, Germany.

Wardman, M. (2001). *Public transport values of time* (Working paper 564). Institute of Transport Studies, University of Leeds.

Williams, H.C.W.L. and Ortuzar, J.D. (1982). Behavioural theories of dispersion and misspecification of travel demand models. *Transportation Research Part B*, *16*, 167-219.

Ziliaskopoulos, A., & Wardell, W. (2000). An intermodal optimum path algorithm for multimodal networks with dynamic arc travel times and switching delays. *European Journal of Operation Research*, *125*, 486-502.

About the author

Maria Stella Fiorenzo-Catalano was born in Venezuela in 1969. At the age of two she moved with her family to Italy. She graduated in Computer Science at the University of Pisa in 1995. In 1996 she worked at the Electronic Engineering and Computer Science department of Perugia University, under the supervision of Prof. Federico Malucelli. From 1997 to 1998 she was employed at the software company Archimede Informatica s.c.r.l. in Pisa, working among others in the development and testing of algorithms to solve the shortest path problems inherent to static and dynamic graphs for transport models in collaboration with Prof. Stefano Pallottino. From 1998 to 2000 she worked at the Centre for Environmental Economics and Management of the MIP University Consortium of Milan Polytechnic and at the Urban Mobility and Innovative Transport Systems Laboratory of Como, under the supervision of Prof. Alberto Colorni.

In 2000 she moved to The Netherlands and she joined the Transportation and Planning section at the Delft University of Technology. Here she started her Ph.D. research in choice set generation in multi-modal transportation networks for a joint research project in collaboration with the Foundation of Applied Scientific Research (TNO) and under the supervision of Prof. Piet Bovy. She also participated in several researches concerning, among others, the design of route generation model to be used for inland navigation on waterways and the route generation model for dynamic traffic assignment model. During these researches she has presented various papers in Dutch and international conferences, and published articles in Dutch and international journals. Since January 2007 she is working at the Mobility and Logistic Business Unit of the Netherlands Organization for Applied Scientific Research (TNO) in Delft, in the area of dynamic traffic assignment and ITS models.

Summary

Introduction

Multi-modal transport is generally seen as a promising approach to help solving today's mobility problems, such as recurrent congestion, deteriorating accessibility of especially city centres, and negative impacts on the environment. Combining private and public transport in a multi-modal transport system offers opportunities to benefit from the strengths of the various systems while avoiding their weaknesses, and might therefore be an interesting alternative to the traditional dichotomous choice between private modes and public transport.

Although promising, multi-modal transport is complex. Multi-modal travelling involves the use of private transport modes (walking, bike, and car), scheduled public transport services (train, bus, tram and metro) and/or non-scheduled public transport services (taxi and demand responsive services). The main characteristic of multi-modal transport is that more than one transport mode is used for a single trip; therefore transfers between travel modes are an essential element of multi-modal transport. In order to analyze multi-modal planning problems, the demand modelling procedure needs to be capable of analyzing and predicting the use of multi-modal trips. However, multiple-choice dimensions involved in a multi-modal trip such as with respect to modes, services, transfer locations and routes are difficult to model.

The approach proposed in this thesis to model multi-modal travel choices is based on the assumption that a clear distinction in the modelling between a choice set generation step and, conditional on that set, the genuine choice modelling step, may significantly improve choice analysis and prediction. Choice set generation consists in finding all feasible routes that a traveller might consider for travelling from his origin to his destination. In a route choice context the choice set composition is a critical aspect because very many routes may be available whereas only a limited subset of those are actually perceived while even less are actually considered by trip makers.

This approach is also based on the use of a supernetwork methodology in which the networks of all available travel modes are combined into a single network that includes transfer possibilities between modes. Multiple-choice dimensions in multi-modal trips are represented jointly as route choice in the supernetwork and by explicitly enumerating the individual choice alternatives prior to the choice modelling; maximum flexibility is available in adopting the most suitable choice modelling approach.

This thesis addresses the generation of choice sets of the travel alternatives for the purposes of route choice analysis and prediction of flows in multi-modal networks. In order to develop such choice set generation approach the following topics are dealt with.

- Presentation of the proposed multi-modal modelling approach and multi-modal transportation network representation.
- Introduction of a conceptual framework for choice set notions and presentation of a clear choice set terminology applicable to various analysis viewpoints.
- Introduction of the notion of an adequate choice set in uni-modal and multi-modal networks; definition of which kind of choice set is most suitable for prediction purposes, and which requirements an adequate choice set should satisfy; presentation of the requirements for an appropriate choice set generation process for uni-modal and multi-modal networks.
- Analysis and comparison of a large number of choice set generation approaches for uni-modal and multi-modal networks.
- Development of a newly route choice set generation algorithm, the so-called doubly stochastic approach applicable to uni-modal and especially multi-modal networks. This doubly stochastic method has been tested with respect to its empirical and probabilistic properties, and applied to various networks, in particular to a multi-modal network.

Multi-modal modelling approach

In this thesis, a new approach for analyzing multi-modal travelling based on the use of a supernetwork methodology was introduced. In the supernetwork approach, the traditional steps of mode choice and route choice are integrated into a single route choice problem in a multi-modal network. Relevant characteristics of the modelling approach are that a modal split module is absent and a route choice set generation module prior to the route choice model is specified. By explicitly generating the individual choice alternatives prior to the choice modelling, maximum flexibility is available in adopting the most suitable choice modelling approach. Moreover, a new specification of the multi-modal network representation was presented, which takes into account several uni-modal network layers of all available travel modes combined into a single multi-modal network that includes transfer possibilities between modes, via waiting and walking links. A study area in The Netherlands (the Dordrecht-Rotterdam corridor) was represented in this new multi-modal representation and was successfully used in analyzing travel choices in this region.

Choice set notions and concepts

An important topic dealt with in this thesis concerns systematic choice set definitions and classifications. Starting from a behaviour-based travel choice framework the concepts of *universal sets*, *objective choice sets*, *subjective choice set* and *consideration sets* were introduced, thereby distinguishing between *actual*, *observed*, and *generated* choice sets, and between choice set generation for *estimation* and *prediction* purposes. Furthermore, in transportation modelling it is not common to follow a completely individual-level approach, therefore we have distinguished between individual and group level. Also, the relationships between all different sets of alternatives have been established.

Criteria for adequate choice sets

New scientific insights on choice set concepts have been established with the new systematic definitions of a reasonable route, and adequate choice sets at individual and group level. The specification of requirements and choice set criteria for adequate route choice sets for prediction of route and link flows in uni-modal and multi-modal transport networks have been introduced in this thesis. These are based on sets of logical, feasibility, preference, and other constraints on the part of the individual traveller such as directness, maximum detour, maximum and minimum lengths of modal legs, mode availability, hierarchical logic, etc. At the group level of OD-flows criteria based on preferential and spatial variety are added. Also, the main characteristics of an appropriate generation process for prediction purpose in uni-modal and multi-modal networks have been established.

Based on the defined criteria, adequate choice sets at group level should contain routes that are not highly overlapping (due to the *overlap* criterion), are comparable in time or distance (due to the *comparability* criterion), without excessive detours, with appropriate choice set size and composition, with spatial variability (due to the *spatial variability* criterion) and with variation in route types (due to the *preferential variability* criterion); in the multi-modal case, also variation in multi-modal combinations and in transport modes used (routes should be multi-modal indeed).

Comparative analysis of choice set generation approaches

A generic systematic route set generation scheme for the purpose of characterizing and analysing current route set generation procedures has been presented. Based on this generic classification scheme, a systematic comparison of the various and sometimes very different route set generation methods known from literature has been established, by which each of these generation methods has been described in a structured comparable way as a basis for the evaluation of their adequacy for the purpose of choice set generation. This evaluation involved testing whether each of the methods might be able to fulfil the requirements for an adequate choice set.

On the basis of this comparison, it appears that the simple repeated shortest path generation methods (such as, K shortest paths, etc.) perform unsatisfactory, whereas the best performance was achieved by the Constrained K-Shortest Paths for the single OD-pair generation and the Monte Carlo Labelling combination (MCL) approach for multiple OD pairs generation. It has also been found that some criteria for generating adequate choice sets were not satisfied by the MCL combination approach (for example, many overlapping routes were generated), since these criteria are not applied nor included during the generation process. In order to improve the quality of the generated choice sets it has been demonstrated that the application of a subsequent filtering process on the generated set, which includes set-based criteria such as maximum overlap, is necessary.

Choice set generation approach for multi-modal networks

The newly developed doubly stochastic approach is an extension of the MCL combination approach since not only the link attributes but also the preference (or behavioural) parameters of the cost function are randomized. A behaviour-based rationale has been formulated for this approach founded on two basic principles. Choice set formation on the individual level is highly preference driven meaning that the generation model should consider preferred route attributes. On the other hand the model considers preference differences among travellers. The basic property of the developed multi-modal choice set generation (MM-CSG) approach is that route choice factors and behavioural parameters of travellers determine the generation of route options for the choice sets via a utility or cost functions, called route generation functions. Since there exist significant differences in choice factors and preferences among user groups and trip purposes, but also in order to achieve a rich variety in the multi-modal composition of the choice sets, these route choice generation functions need to be specific for user groups and trip categories.

Therefore, the new method focuses on generating routes for prediction purposes at group level, in which *groups of travellers* are taken into account, and in particular, *user classes* that vary with respect to expected travel behaviour, for instance based on trip purpose (work, business, study, etc), and *traveller groups* having different modes available are distinguished. In fact, the randomization of the behavioural preference parameters provides variations in transport modes and in multi-modal combinations by taking into account the variations within the user classes, since in multi-modal network applications, a rich variety in multi-modal composition of the choice set is desirable.

In order to achieve adequate choice sets that satisfy most of the requirements defined, a filtering process has been applied after the generation. The MM-CSG approach produces an exhaustive base route set (master set) that needs to be freed from superfluous and non-efficient routes in order to arrive at choice sets that efficiently suit a particular purpose, in our case the prediction purpose. To that end, a variety of selection constraints may be additionally applied. For prediction purposes, highly overlapping paths may be removed, while at the same time spatial variety is welcomed.

The various analyses performed with the doubly stochastic choice set generation approach have shown that the resulting choice sets show the desired properties in terms of choice set size (minimum and maximum size) and composition (variety in trip properties). The doubly stochastic choice set generation approach has been applied in a real-world multi-modal network, namely in the Rotterdam-Dordrecht region in The Netherlands. In order to prove the effectiveness of the MM-CSG method, the generated choice sets were compared with observed chosen routes and reported choice sets in this multi-modal network. Moreover, the generated choice sets have shown an adequate variety of uni-modal and multi-modal route alternatives. An extensive analysis of the probabilistic properties of this stochastic generation method has been performed (impacts of seeds, number of randomisations, adopted variance levels, etc) showing that satisfactory performance can be achieved with a limited number of randomisations.

Conclusion

This dissertation research has brought new insights and results with respect to several research areas and contributions in choice set generation approach in multi-modal transportation networks.

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: Evaluation 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

Nes, R. van, *Design of multimodal transport networks, a hierarchical approach*, T2002/5, September 2002, TRAIL Thesis Series, Delft University Press, The Netherlands

Pol, P.M.J., *A Renaissance of Stations, Railways and Cities, Economic Effects, Development Strategies and Organisational Issues of European High-Speed-Train Stations*, T2002/6, October 2002, TRAIL Thesis Series, Delft University Press, The Netherlands

Runhaar, H., *Freight transport: at any price? Effects of transport costs on book and newspaper supply chains in the Netherlands*, T2002/7, December 2002, TRAIL Thesis Series, Delft University Press, The Netherlands

Spek, S.C., van der, *Connectors. The Way beyond Transferring*, T2003/1, February 2003, TRAIL Thesis Series, Delft University Press, The Netherlands

Lindeijer, D.G., *Controlling Automated Traffic Agents*, T2003/2, February 2003, TRAIL Thesis Series, Eburon, The Netherlands

Riet, O.A.W.T., van de, *Policy Analysis in Multi-Actor Policy Settings. Navigating Between Negotiated Nonsense and Useless Knowledge*, T2003/3, March 2003, TRAIL Thesis Series, Eburon, The Netherlands

Reeven, P.A., van, *Competition in Scheduled Transport*, T2003/4, April 2003, TRAIL Thesis Series, Eburon, The Netherlands

Peeters, L.W.P., *Cyclic Railway Timetable Optimization*, T2003/5, June 2003, TRAIL Thesis Series, The Netherlands

Soto Y Koelemeijer, G., *On the behaviour of classes of min-max-plus systems*, T2003/6, September 2003, TRAIL Thesis Series, The Netherlands

Lindveld, Ch..D.R., *Dynamic O-D matrix estimation: a behavioural approach*, T2003/7, September 2003, TRAIL Thesis Series, Eburon, The Netherlands

Weerdt, de M.M., *Plan Merging in Multi-Agent Systems*, T2003/8, December 2003, TRAIL Thesis Series, The Netherlands

Langen, de P.W, *The Performance of Seaport Clusters*, T2004/1, January 2004, TRAIL Thesis Series, The Netherlands

Hegyi, A., *Model Predictive Control for Integrating Traffic Control Measures*, T2004/2, February 2004, TRAIL Thesis Series, The Netherlands

Lint, van, J.W.C., *Reliable Travel Time Prediction for Freeways*, T2004/3, June 2004, TRAIL Thesis Series, The Netherlands

Tabibi, M., *Design and Control of Automated Truck Traffic at Motorway Ramps*, T2004/4, July 2004, TRAIL Thesis Series, The Netherlands

Verduijn, T. M., *Dynamism in Supply Networks: Actor switching in a turbulent business environment*, T2004/5, September 2004, TRAIL Thesis Series, The Netherlands

Daamen, W., *Modelling Passenger Flows in Public Transport Facilities*, T2004/6, September 2004, TRAIL Thesis Series, The Netherlands

Zoeteman, A., *Railway Design and Maintenance from a Life-Cycle Cost Perspective: A Decision-Support Approach*, T2004/7, November 2004, TRAIL Thesis Series, The Netherlands

Bos, D.M., *Changing Seats: A Behavioural Analysis of P&R Use*, T2004/8, November 2004, TRAIL Thesis Series, The Netherlands

Versteegt, C., *Holonic Control For Large Scale Automated Logistic Systems*, T2004/9, December 2004, TRAIL Thesis Series, The Netherlands

Wees, K.A.P.C. van, *Intelligente voertuigen, veiligheidsregulering en aansprakelijkheid. Een onderzoek naar juridische aspecten van Advanced Driver Assistance Systems in het wegverkeer*, T2004/10, December 2004, TRAIL Thesis Series, The Netherlands

Tampre, C.M.J., *Human-Kinetic Multiclass Traffic Flow Theory and Modelling: With Application to Advanced Driver Assistance Systems in Congestion*, T2004/11, December 2004, TRAIL Thesis Series, The Netherlands

Rooij, R.M., *The Mobile City. The planning and design of the Network City from a mobility point of view*, T2005/1, February 2005, TRAIL Thesis Series, The Netherlands

Le-Anh, T., *Intelligent Control of Vehicle-Based Internal Transport Systems*, T2005/2, April 2005, TRAIL Thesis Series, The Netherlands

Zuidgeest, M.H.P., *Sustainable Urban Transport Development: a Dynamic Optimization Approach*, T2005/3, April 2005, TRAIL Thesis Series, The Netherlands

Hoogendoorn-Lanser, S., *Modelling Travel Behaviour in Multimodal Networks*, T2005/4, May 2005, TRAIL Thesis Series, The Netherlands

Dekker, S., *Port Investment - Towards an integrated planning of port capacity*, T2005/5, June 2005, TRAIL Thesis Series, The Netherlands

Koolstra, K., *Transport Infrastructure Slot Allocation*, T2005/6, June 2005, TRAIL Thesis Series, The Netherlands

Vromans, M., *Reliability of Railway Systems*, T2005/7, July 2005, TRAIL Thesis Series, The Netherlands

Oosten, W., *Ruimte voor een democratische rechtsstaat. Geschakelde sturing bij ruimtelijke investeringen*, T2005/8, September 2005, TRAIL Thesis Series, Sociotext, The Netherlands

Le-Duc, T., *Design and control of efficient order picking*, T2005/9, September 2005, TRAIL Thesis Series, The Netherlands

Goverde, R., *Punctuality of Railway Operations and Timetable Stability Analysis*, T2005/10, October 2005, TRAIL Thesis Series, The Netherlands

Kager, R.M., *Design and implementation of a method for the synthesis of travel diary data*, T2005/11, October 2005, TRAIL Thesis Series, The Netherlands

Boer, C., *Distributed Simulation in Industry*, T2005/12, October 2005, TRAIL Thesis Series, The Netherlands

Pielage, B.A., *Conceptual Design of Automated Freight Transport Systems*, T2005/14, November 2005, TRAIL Thesis Series, The Netherlands

Groothedde, B., *Collaborative Logistics and Transportation Networks, a modeling approach to network design*, T2005/15, November 2005, TRAIL Thesis Series, The Netherlands

Valk, J.M., *Coordination among Autonomous Planners*, T2005/16, December 2005, TRAIL Thesis Series, The Netherlands

Krogt, R.P.J. van der, *Plan Repair in Single-Agent and Multi-Agent Systems*, T2005/17, December 2005, TRAIL Thesis Series, The Netherlands

Bontekoning, Y.M., *Hub exchange operations in intermodal hub-and-spoke networks. A performance comparison of four types of rail-rail exchange facilities*,T2006/1, February 2006, TRAIL Thesis Series, The Netherlands

Lentink, R., *Algorithmic Decision Support for Shunt Planning*, T2006/2, February 2006, TRAIL Thesis Series, The Netherlands

Ngoduy, D., *Macroscopic Discontinuity Modeling for Multiclass Multilane Traffic Flow Operations*, T2006/3, April 2006, TRAIL Thesis Series, The Netherlands

Vanderschuren, M.J.W.A., *Intelligent Transport Systems for South Africa. Impact assessment through microscopic simulation in the South African context*, T2006/4, August 2006, TRAIL Thesis Series, The Netherlands

Ongkittikul, S., *Innovation and Regulatory Reform in Public Transport*, T2006/5, September 2006, TRAIL Thesis Series, The Netherlands

Yuan, J., *Stochastic Modelling of Train Delays and Delay Propagation in Stations*, T2006/6, October 2006, TRAIL Thesis Series, The Netherlands

Viti, F., *The Dynamics and the Uncertainty of Delays at Signals*, T2006/7, November 2006, TRAIL Thesis Series, The Netherlands

Huisken, G., *Inter-Urban Short-Term Traffic Congestion Prediction*, T2006/8, December 2006, TRAIL Thesis Series, The Netherlands

Feijter, R. de, *Controlling High Speed Automated Transport Network Operations*, T2006/9, December 2006, TRAIL Thesis Series, The Netherlands

Makoriwa, C., *Performance of Traffic Networks. A mosaic of measures*, T2006/10, December 2006, TRAIL Thesis Series, The Netherlands

Miska, M., *Microscopic Online Simulation for Real time Traffic Management*, T2007/1, January 2007, TRAIL Thesis Series, The Netherlands

Chorus, C., *Traveler Response to Information*, T2007/2, February 2007, TRAIL Thesis Series, The Netherlands

Weijermars, W.A.M., *Analysis of Urban Traffic Patterns Using Clustering*, T2007/3, April 2007, TRAIL Thesis Series, The Netherlands

Zondag, B., *Joined Modeling of Land-use, Transport and Economy*, T2007/4, April 2007, TRAIL Thesis Series, The Netherlands

Bok, M.A. de, *Infrastructure and Firm Dynamics: A micro-simulation approach*, T2007/5, May 2007, TRAIL Thesis Series, The Netherlands

Fiorenzo-Catalano, M.S., *Choice Set Generation in Multi-Modal Transportation Networks*, T2007/6, June 2007, TRAIL Thesis Series, The Netherlands