Universität Stuttgart

Institut für Straßen- und Verkehrswesen
Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik
Univ.-Prof. Dr.-Ing. M. Friedrich

A Framework for Traffic Assignment with Explicit Route Generation

Yaohua Xiong

Veröffentlichungen aus dem
Institut für Straßen- und Verkehrswesen

Heft 52 (September 2014)

ISSN 0932 - 402X
ISBN 978 - 3 - 9816754 - 2 - 9
D 93 (Dissertation der Universität Stuttgart)

52

Universität Stuttgart

Veröffentlichungen aus dem Institut für Straßen- und Verkehrswesen

Universität Stuttgart

Institut für Straßen- und Verkehrswesen
Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik
Univ.-Prof. Dr.-Ing. M. Friedrich

A Framework for Traffic Assignment with Explicit Route Generation

Yaohua Xiong

Veröffentlichungen aus dem
Institut für Straßen- und Verkehrswesen

Heft 52 (September 2014)

Herausgeber	:	Institut für Straßen- und Verkehrswesen Universität Stuttgart
Copyright	:	Das Copyright liegt beim Verfasser.
Eigenverlag und Vertrieb	:	Institut für Straßen- und Verkehrswesen Universität Stuttgart Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik Pfaffenwaldring 7 70569 Stuttgart

Hinweis / Note:

Diese Veröffentlichung ist auch als "Elektronische Dissertation" online unter **http://elib.uni-stuttgart.de** verfügbar und kann dort im PDF-Format heruntergeladen werden.

This paper is also published online as "Electronic Dissertation" at **http://elib.uni-stuttgart.de** and can be downloaded there as PDF file.

A Framework for Traffic Assignment with Explicit Route Generation

Von der Fakultät Bau- und Umweltingenieurwissenschaften der Universität Stuttgart zur Erlangung der Würde eines Doktors der Ingenieurwissenschaften (Dr.-Ing.) genehmigte Abhandlung

Vorgelegt von

Yaohua Xiong

aus Chongqing, VRC

Hauptberichter: Prof. Dr.-Ing. Markus Friedrich

Mitberichter: Prof. Dr.-Ing. Keping Li

Tag der mündlichen Prüfung: 18.09.2014

Institut für Straßen- und Verkehrswesen

Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik

Universität Stuttgart

2014

Foreword

Traffic assignment models play a vital role in understanding the internal dynamics between the demand and the supply in a transport system. This dissertation is the cumulative result of my four years' study at the Institute of Road and Transportation Science, University Stuttgart. It attempts to improve the predictive ability of traffic assignment models by providing a framework to formalize and integrate travellers' route identifying and choosing behaviours.

During the past four years, I studied and lived in Germany, a foreign country for me. As a result, life brought special challenges with more stress and excitement. In the end, it were four years that I will always remember. Here I would like to express my gratitude to people who made the four years memorable.

First, I would like to thank Professor Markus Friedrich of University of Stuttgart, for offering me the opportunity to study at the institute. He supervised and supported my research with great skill and patience; without them, I could hardly finish this dissertation. I learned a lot from his insight into the modelling of transportation system, his methodical working approach, and his constantly motivating personality.

I also would like to thank Professor Keping Li of Tongji University, who co-supervised this dissertation. He gave insightful advices on the general methodology of conducting research and on many specific problems.

Additionally I would like to thank my colleagues and friends who provided help to this thesis. Most notably, Alice Lorenz helped to proofread the draft; Juliane Pillat provided assistance in data processing. My gratitude also goes to other friends and colleagues, who may not be involved directly in my research, but whose friendship and encouragement made my life in Germany more enjoyable.

Finally, I want to thank my parents, to whom I dedicate this thesis, for their unconditional love and support.

Yaohua Xiong

Table of Contents

List of Tables

List of Figures

List of Notations

Graph

$\boldsymbol{G} = (\boldsymbol{V}, \boldsymbol{E})$	Directed graph with vertices and edges
$\boldsymbol{V}$	Set of vertices
$\boldsymbol{V}_o$	Set of origin vertices
o, d	An origin vertex and an destination vertex
$\boldsymbol{E}$	Set of edges
$e \equiv [e_t, e_h]$	Edge as a ordered pair of vertices
e_t	Tail vertex
e_h	Head vertex
$p \equiv [v_1, \dots, v_n]$	Path or path segment as a ordered sequence of vertices
$p + q$	Concatenation of two path segments
$\boldsymbol{\Delta}$	Edge-path incident matrix
δ_{ep}	Element of edge-path incidence matrix for edge e and path p
$\boldsymbol{P}_{ij}$	Set of all paths between vertex$i$ and vertex j
$\boldsymbol{R}_{od}$	Set of all routes between OD-pair od

Traffic flows

d_{od}	OD flow(or travel demand) between origin o and destination d
$\boldsymbol{D}$	OD-matrix
h_p	Path flow on path p between od

$\boldsymbol{h}$	Path flow vector
$\boldsymbol{h}_o$	Path flow vector for all paths from origin o
$\boldsymbol{h}_{od}$	Path flow vector for all paths between OD-pair od
f_e	Edge flow on edge e
$\boldsymbol{f}$	Edge flow vector

Cost

c_e	Edge cost
$\boldsymbol{c}$	Edge cost vector
$\boldsymbol{g}$	Path cost vector
g_p	Path cost
g_p^*	Path cost of the shortest path between two ends of path p

Constraints

gd_p	Global detour factor of path p
ld^*	Local detour factor of path p
lt^*	Local optimal threshold of path p
hi	Hierarchical index of path p
$\overline{gd}$	Parameter of global detour factor constraint
$\overline{ld}$	Parameter of local detour factor constraint
$\overline{lt}$	Parameter of local direct threshold constraint
$\overline{hl}$	Parameter of hierarchical index constraint

Abstract

Transport assignment models play an important role in the research and practice of transport planning. Traditionally, transport assignment models are link-based, predicting traffic volumes on network links. However, route-based transport assignment models, which work with route sets and determine traffic volumes of routes, became more important in recent years. They support advanced analysis procedures that require route volumes as input, e.g. estimation of OD-matrices and evaluation of intelligent transport systems (ITS).

The aim of this research is to develop a general framework for a route-based transport assignment, which emphasizes the quality of the resulting route sets. To achieve this goal, this framework differentiates the assignment procedure into two spate stages. In the first stage, it generates the route sets for all OD-pairs explicitly, and then in the second stage it distributes the travel demand between these routes. With this two-stage structure, a comprehensive route set generation procedure can be incorporated considering specific constraints.

In order to generate realistic route sets, it is essential to have a clear definition of the quality of route sets. After examining the assumptions on the travellers' behaviour in identifying and choosing routes, a set of indicators is introduced to quantify the quality of route sets. These indicators concern two aspects of the quality of route sets: the quality of individual routes and the composition of route sets. Constraints such as global detour factor, local detour factor, local optimal threshold for route sections and road hierarchical index are introduced to ensure the quality of individual routes while constraints such as choice set size and minimal dissimilarity address the composition of route sets.

With these indicators in place, a two-stage route set generation procedure is developed to generate route sets whose quality satisfies certain constraints. The first stage employs the branch-and-bound technique to enforce the constraints on individual routes. For a given OD-pair, it enumerates all possible routes systematically and discards routes whose indicators do not satisfy the constraints. The second stage enforces the constraints on the route set composition. It produces route sets that contain a given number of routes where the routes are dissimilar with each other.

Using the route set generation method a framework for distributing travel demand on given route sets is presented. It applies a nested logit model that takes advantages of the hierarchical structure of the route choice sets. For the feedback between route choice and route travel time it uses a simple convergence procedure based on successive average. A pilot implementation of this assignment framework is developed and some implementation issues such as network pre-processing are discussed. The pilot implementation is validated with a small test network and a larger test case in the Munich region. Here the generated route sets are compared with observed route sets. The result shows that the framework is capable of reproducing observed route sets with sufficient accuracy and that it is flexible enough to be calibrated for different situations.

Kurzfassung

Für die Forschung und Praxis der Verkehrsplanung spielen Verkehrsumlegungsmodelle eine wichtige Rolle. Üblicherweise sind Verkehrsumlegungsmodelle streckenbasiert, d. h. sie prognostizieren die Verkehrsstärke für einzelne Strecken des Verkehrsnetzes. In den vergangenen Jahren haben routenbasierte Umlegungsmodelle an Bedeutung gewonnen, die Routenmengen generieren und die Verkehrsstärke jeder Route bestimmen. Sie sind Grundlage für weitere Verfahren, wie die Schätzung von Quelle-Ziel-Matrizen oder die Bewertung verkehrstelematischer Maßnahmen, die Routenbelastungen als Eingangsgrößen benötigen.

Das Ziel dieser Arbeit ist es, ein allgemeines routenbasiertes Verkehrsumlegungsmodell zu entwickeln, bei dem die Qualität der verwendeten Routensets im Vordergrund steht. Zu diesem Zweck erfolgt die Umlegung in zwei getrennten Stufen: In der ersten Stufe werden die Routensets für alle Quelle-Ziel-Relationen in einem eigenständigen Verfahren generiert. In der zweiten Stufe wird die Verkehrsnachfrage dann auf diese Routen umgelegt. Die Trennung von Routengenerierung und Routenwahl ermöglicht eine Routengenerierung, die besondere Randbedingungen erfüllen kann.

Die Generierung realistischer Routensets erfordert eine eindeutige Definition der Qualität von Routensets. Aufbauend auf einer Untersuchung der Annahmen zum Routenwahlverhalten von Verkehrsteilnehmern wird eine Reihe von Kenngrößen eingeführt, um die Qualität eines Routensets zu bewerten. Diese Kenngrößen bewerten zwei Aspekte der Qualität von Routensets: die Qualität der einzelnen Routen und die Zusammensetzung des gesamten Routensets. Die Qualität einzelner Routen wird durch die Vorgabe eines globalen und eines lokalen Umwegfaktors, eines lokalen Grenzwerts für Routenabschnitte und eines Straßenkategorieindexes sichergestellt. Die Zusammensetzung der Routensets wird durch die maximale Größe des Routensets und eine geringe Ähnlichkeit der Routen beeinflusst.

Auf der Basis dieser Kenngrößen wird ein zweistufiges Verfahren zur Generierung von Routensets entwickelt, das Qualitätsgrenzwerte einhält. Die erste Stufe nutzt einen Branch-and-Bound Ansatz um die Grenzwerte für einzelne Routen einzuhalten. Für jede Quelle-Ziel-Relation werden alle möglichen Routen systematisch erzeugt, wobei Routen, deren Kenngrößen den Grenzwerten nicht entsprechen, aussortiert werden. In der zweiten Stufe wird die Zusammensetzung des Routensets anhand weiterer Grenzwerte überprüft, so dass Routensets mit einer vorgegebenen Anzahl von Routen generiert werden, deren Routen möglichst eigenständig sind.

Aufbauend auf dem Routengenerierungsverfahren wird ein allgemeines Verfahren vorgestellt, mit dem die Nachfrage auf die Routen aufgeteilt werden kann. Es basiert auf einem Nested Logit Modell, das den hierarchischen Aufbau des Routensets nutzt. Für die Rückkopplung zwischen Routenwahl und belastungsabhängiger Fahrzeit wird ein einfaches Konvergenzverfahren basierend auf der Method of Successive Averages eingesetzt. Das Umlegungsverfahren wird als Prototyp implementiert und Fragen der Netzaufbereitung behandelt. Der Prototyp wird mithilfe eines kleinen Testnetzes und anhand eines größeren Testfalls in der Region München validiert. Die Ergebnisse zeigen, dass das Routengenerierungsverfahren die beobachteten Routensets mit angemessener Genauigkeit darstellen kann und dass es für unterschiedliche Anforderungen kalibriert werden kann.

1 Introduction

1.1 Background and Motivation

Transport, the movement of people and goods from one place to another, is a fundamental aspect of human society. It has great impacts on social and economic development by providing the mobility that allows people to take part in social and economic activities at different locations. It also enables increasing degrees of economic specialization, which shapes the modern world.

The overall transport process within a given area can be regarded as a complex system, or a **transport system**. In a broad sense (CASSETA, 2001), a transport system consists of different interrelated elements such as passenger, goods, vehicles and infrastructures. It also includes non-physical elements such as operation plans, regulations, and policies. Transport researchers, planners and engineers are striving to understand the interactions among these elements and devise measures to improve the performance of the system. However, this is a non-trivial task due to the complexity of transport systems. Various **transport models** are developed to assist this task.

A transport model is a simplified and idealized description of a transport system. It handles the complexity of a transport system by concentrating on aspects that are relevant for a specific purpose while omitting aspects that are irrelevant. This research focuses on one particular type of transport model, the **traffic assignment model**, which predicts the traffic flows in the transport network by modelling the interaction between travellers' travel decisions and the traffic condition in the transport network.

Assignment models can roughly be divided into two categories, **link-based assignment models** and **route-based assignment models**, according to how the traffic flows are represented. The link-based assignment models predict traffic flow on each link while the route-based assignment models can predict traffic flow on each used route. While link-based assignment models, such as the traditional Deterministic User Equilibrium (DUE) model (DAFERMOS & SPARROW, 1969), are sufficient for common planning tasks where only link flows are required, they are insufficient for more advanced analysis, such as estimating the OD-matrix or evaluating an ITS system, where the distribution of traffic flows among routes is critical. It is trivial to convert route flows into link flows, while the reverse is not true. In this sense, the route-based assignment models produce more detailed results than the link-based assignment models.

A route-based assignment model involves two essential components: identifying routes for all OD-pairs and distributing travel demands among these routes. The routes for a given OD-pair define a route set. Depending on the way how route sets are identified route-based assignment models can be further divided into those **with implicit route sets** and those **with explicit route sets.** In an assignment model with explicit route sets, the route sets for all OD-pairs are generated explicitly before distributing the travel demand. On the other hand, in an assignment model with implicit route sets, the route sets are identified within the distribution process according to some implicit rules.

Traditionally, assignment models with implicit route sets received more attention as they are computationally less demanding. A choice set generally comprises a large number of routes even for a moderately sized transport network. Assignment models with implicit route sets reduce the demand for processing power and storage capacity by not generating and storing these routes explicitly. The origin-based assignment (BAR-GERA, 1999) and the stochastic assignment model developed in (DIAL, 1971) are two such examples.

With the rapid development of computational technology in recent years, the processing power and storage capacity of computers is no longer a major issue. On the other hand, the advantages of explicit route set generation, respectively that it provides the flexibility to incorporate comprehensive behaviour rules and enable advanced choice models handling advanced concepts such as overlapping routes, non-linear and route specific costs become more appealing. Some preliminary work to incorporate explicit route set generation into the assignment model by (BLIEMER & TAALE, 2006) shows promising results.

1.2 Research Objectives

The objective of this research is to develop a general framework for a route-based traffic assignment with a focus on the quality of the explicitly generated route sets. In order to achieve this objective, the following problems need to be solved:

- How to define the quality of a route set in an objective way, so that an automatic procedure can be developed to assess the quality.
- Given a clear definition of route set quality, how to generate route sets efficiently.
- How to compare the generated route sets with the route sets observed in real world and calibrate the route generation procedure accordingly.
- With explicitly generated route sets, how to distribute travel demand between them.

1.3 Thesis Outline

This thesis presents a traffic assignment framework that incorporates an explicit route generation method. The route generation method is carefully designed in order to generate realistic routes efficiently and the generated routes are used as the basis for the following route based traffic assignment.

Chapter 2 reviews research relevant to route generation, route choice and traffic assignment. Section 2.1 provides a formalized description of transport system models, establishing essential concepts for the rest of the text. Section 2.2 categorizes the different assumptions about modelling route choice behaviour of travellers. Section 2.3 analyses existing route choice set generation methods, comparing their advantages, shortcomings, and underlying assumptions, in order to lay the ground of developing a new method. Sections 2.4 reviews different forms of route choice function, highlighting their ability to handle overlapping routes. Section 2.5 reviews various network loading models and analyses the trade-off between fidelity and computational performance for each model.

Chapter 3 develops a new route choice set generation method based on the branch-and-bound technique. Section 3.1 discusses assumptions on travellers' behaviour in forming a route choice set and defines quality indicators for both individual routes and the composition of route sets. Section 3.2 gives an overview of the route set generation procedure. Section 3.3 describes procedures of enumerating routes between OD-pairs in a systematic way. It includes the basic branch-and-bound procedure and the generalized procedure based on segments instead of single links. Section 3.4 describes the procedure to assemble reasonable routes into route choice sets with a reasonable composition. Section 3.5 discusses the relationship between generated route sets and observed route sets and how to calibrate constraint parameters based on reference route sets.

Chapter 4 describes a route based assignment procedure that is able to find an equilibrium assignment solution based on pre-generated route choice sets. Section 4.1 describes the overall structure of the assignment procedure. Section 4.2 introduces a nested route choice model that takes advantage of the hierarchical structure of the consideration set. Section 4.3 describes simple network loading models for the reference implementation. Section 4.4 describes the procedure to balance travel demands among routes in order to achieve equilibrium.

Chapter 5 puts the developed framework to test. Section 5.1 addresses several issues in developing and applying a reference implementation for this framework. Section 5.2 applies the route set generation procedure to an artificially constructed network to highlight the effect of different constraints. Section 5.3 applies the assignment framework to a real world network. It first describes the data sources, including characteristics of the road network and the reference route sets. Then it presents the generated route sets and compares them with the reference route sets.

Chapter 6 summarizes this thesis and suggests directions for further research.

2 Literature Review

2.1 Transport System Model

A transport system refers to the overall transport process within a given area. A transport system model is a simplified and idealized description of a real world transport system. In order to reduce the complexity and provide a framework to facilitate understanding, a transport model is conceptually split into two sides: the **demand side** and the **supply side**.

The travel demand results from the cumulative decisions made by all travellers within a transport system. These decisions occur across different time spans, ranging from long term decisions lasting for years, such as choosing residence location or work location, to middle term decisions lasting for days and weeks, such as choosing a transport mode and a departure time and a to short term decisions, such route choice, travel speed and lane changes.

The supply side of a transport model comprises all elements that provide travel opportunities, including vehicles, infrastructures and their operating plans. The supply of a transport model can be represented as a network of transport services connecting different places. Therefore, the supply is also referred to as the **transport network**. There are different **modes** of transport services, using different vehicles, infrastructures and operating plans. A transport system in the real world usually comprises more than one mode. For example, common modes of passenger transport in an urban area include private cars, buses, trams, metros, bicycles and walking. If a transport model considers only one mode, then it is called **unimodal** model. Otherwise, it is called **multimodal** model.

2.1.1 Travel Demand

The demand of a transport system can be represented by **trips** within a given geographical area, or the study area of the transport system. These trips result from various travel decisions, involving complex psychological and cognitive mechanisms. As common practice, the study area is divided into non-overlapping **traffic analysis zones**, and the starting point (origin) and end point (destination) of a trip always falls in certain zones. With further assuming that a trip always starts and ends at the centroid of these zones, a collection of trips can be recorded in a concise form, called an **OD-matrix**:

$$\boldsymbol{D} = \begin{bmatrix} d_{11} & \dots & d_{1n} \\ \vdots & \ddots & \vdots \\ d_{m1} & \dots & d_{mn} \end{bmatrix}.$$

Each element d_{ij} of an OD-matrix denotes the number of trips from zone i to zone j. The zone i and zone j are called the **origin** and the **destination** of these trips, hence the name OD-matrix.

The travel demand in a transport system can be represented by several matrices, each of which represents a different **demand segment**. A demand segment refers to all the trips that comply with certain criteria. These criteria can be based on various properties of trips, such as the distance (e.g. local, regional, long-distance), the purpose of the trips (e.g. home, work, shopping, leisure), the traffic modes used for the trips (e.g. car, bus, metro) and the departure times. In particular, a set of OD-matrices that represent travel demands in consecutive time periods are called **dynamic OD-matrices**.

2.1.2 Transport Supply

A network of transport services represents the supply of a transport system. A real world transport network contains different transport modes and each mode could have different representations. To represent transport network of multiple modes is a research topic by itself and out of the scope of this research, see (BENJAMINGS & LINDVELD, 2002) for more details.

This research concentrates on a road network for private car traffic. A road network can be represented based on three types of **network elements**: **nodes**, **links** and **turns**. A node represents a point where several road segments converge, such as an intersection or a ramp. A link represents a road segment between two adjacent nodes. A turn represents the turning movement from one link to another.

Beside these elements that represent physical entities in a road network, virtual elements called **connectors** are used to represent connections between zones and nodes. Travellers are assumed to enter and leave the road network through connectors. Although connectors do not physically exist in a road network, they are considered as part of the supply model. Figure1 illustrates the relationship between these elements in an idealized road network.

In a road network, a **route** is a sequence of network elements connecting an origin zone to a destination zone. It represents the travel trajectory followed by a traveller when he makes a trip in the network. For a trip between a given OD-pair multiple routes may exist which are referred to as a **route set**. Different types of route sets need to be distinguished to facilitate the discussion of route generation.

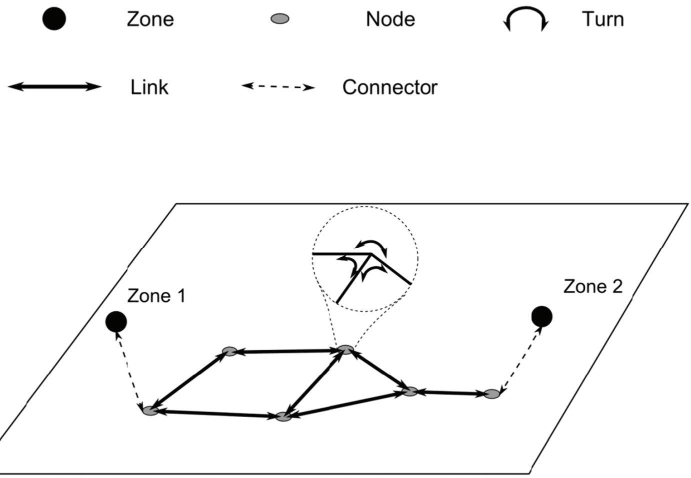

Figure1: A schematic representation of a road network with the following network elements: nodes, links, turns, zones and connectors

A **graph** $\boldsymbol{G}$ is a mathematical structure which comprises a pair of sets $\boldsymbol{G} = \{\boldsymbol{V}, \boldsymbol{E}\}$, where $\boldsymbol{V}$ is the set of **vertices** and $\boldsymbol{E}$ is the set of **edges**. To simplify the analysis, a graph that represents a transport network is further assumed to be:

- directed, which means the edge connecting vertex a with vertex b is different from the edge connecting vertex b to a.
- simple, which means that at most one edge is allowed for one direction between any two vertices, and that no edge is allowed to start and end at the same vertex.

Generally, a vertex corresponds to a node or a zone centroid; an edge corresponds to a link or a connector. $\boldsymbol{V}$ has two subsets $\boldsymbol{V_o} \subset \boldsymbol{V}$ and $\boldsymbol{V_d} \subset \boldsymbol{V}$, where $\boldsymbol{V_o}$ represents the set of vertices where trips start and $\boldsymbol{V_d}$ represents the set of vertices where trips end.

In a simple and directed graph, an edge e from vertex a to vertex b can be denoted as an ordered pair of vertices $e \equiv [a, b]$. Vertex a and vertex b are called the **tail** and the **head** of edge e respectively. Giving an edge e its tail and head are denoted with notations e_t and e_h. According to these definitions, the invariant $e \equiv [e_t, e_h]$ always holds.

In a graph, a **path** p is a consecutive sequence of edges $p \equiv [e_1, \ldots, e_m]$ such that

$$e_{i_h} = e_{(i+1)_t}, \forall 1 \leq i \leq m - 1.$$

This means that in a given path the head of the former edge is the tail of the latter edge. For generality, an empty sequence [] is also regarded as a valid path.

In a simple graph a path can also be represented as a sequence of vertices $p \equiv [v_1, \dots, v_n]$ such that

$$[v_i, v_{i+1}] \in E, \forall 1 \le i \le n-1.$$

This means that an edge exists between v_i and v_{i+1} and it can be presented as $[v_i, v_{i+1}]$. The two representations of a path p are equivalent when $n = m + 1$ and

$$e_{i_t} = v_i, e_{i_h} = v_{i+1}, \forall 1 \le i \le n-1.$$

For a path p the tail of the first edge and the head of the last edge are called the tail and the head of the path, denoted as p_t and p_h. For $p = [e_1, \dots, e_m]$, $p_t = e_{1_t}$, $p_h = e_{m_h}$.

If a path goes through any vertex not more than once, then it is a **simple path**. If it goes through any edge not more than once, then it is **semi-simple path**. A simple path must be a semi-simple path; however, the reverse is not necessarily true.

Given two path p and q if $p_h = q_t$ then a new path, denoted by $p + q$ can be built by concatenating them. However, even if both p and q are simple, $p + q$ is not necessarily simple.

Given a transport network and a corresponding graph representation, a route in the network has a corresponding path in the graph. On the other hand, a path in the graph has a corresponding route if and only if for the path p the tail $p_t \in \boldsymbol{V_o}$ and the head $p_h \in \boldsymbol{V_d}$. A path that has a corresponding route is called a **complete path**.

Each edge has different values to measure the performance of transport on the network element corresponding to this edge. These values are called **performance values.** Common performance values for road networks are length, travel time, monetary cost and discomfort caused by heavy traffic. Each performance value represents one aspect of the disutility perceived by travellers when they travel through this edge. Different performance values on an edge e can be combined into a single value c_e that represents the generalized **edge cost**. Different travellers may have different perceived edge cost depending on their personal preference and on information availability.

Although the network model and the graph are closely related, they describe transport networks at different levels of abstraction. It is important to realize this distinction to avoid confusion. The general reviews and discussions in this chapter will use terms such as link, node, turn, zone, connector and route. The more specific description of methods and algorithms will use terms such as edge, vertex and path.

2.2 Overview on Traffic Assignment Models

Traffic assignment models are studied extensively by many researchers because of their importance for the theory and practice of transport planning. Various models are built while most of them follow a common concept structure.

2.2.1 Concept Structure of Traffic Assignment Model

Traffic assignment models simulate the interactions between travel choices of the demand side and the performance of the supply side. Conceptually these interactions can be summarized with a diagram in Figure 2 (FRIEDRICH, 2010).

- Route choice requires decisions that individuals select from a set of alternatives. They assess each alternative and choose the alternative they believe to be optimal, i.e. they maximise their personal utility (or minimize the cost).
- For each individual the perceived utility of an alternative is influenced by his/her personal preference. Individuals can be categorized into certain person groups with identical or similar preferences. For individuals in the same person group, the perceived utility is further influenced by the quality of the information available. The persons can have perfect or incomplete information.
- The individual choices can influence the state of the transport network. This is the case when high demand affects the traffic flow and the service quality. As shown in Figure 2 this affects the utility of alternatives and the choices of other individuals.
- The travel demand and the network are in a steady state over a long time period. This fact permits a learning process where individuals can collect information on the traffic state and adapt their choices finally leading to an equilibrium state. In this state the travellers stick to their choices so that the traffic state and the resulting utility of alternatives is consistent.

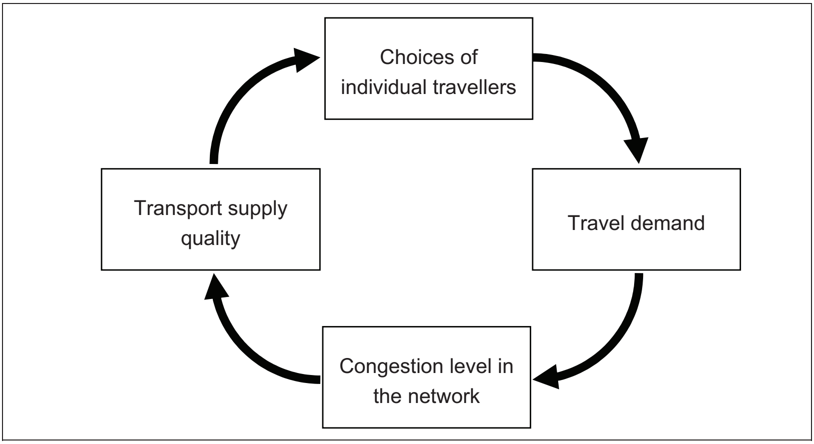

Figure 2: Feedback loop between individual choices, travel demand and supply quality in a transport network.

This structure is highly simplified to provide a high level overview. From this diagram it is clear that the modelling of route choices is a crucial part of the traffic assignment. There are two separate tasks in modelling the route choices:

- route generation, which identifies possible routes between the origin and the destination of a given trip,
- the route choice function, which determines which route among these routes to be chosen.

However in traditional traffic assignment models they are mixed.

2.2.2 Assumptions of Traffic Assignment Models

The basic assumption of traffic assignment is that the demand and the supply will finally achieve an equilibrium state. Assignment models that comply with this assumption are called **User Equilibrium** (UE) assignment models. The equilibrium state of a transport model is first summarized by Wardrop's first principle (WARDROP, 1952):

"The journey times on all the routes actually used are equal, and less than those which would be experienced by a single vehicle on any unused route."

There are also traffic assignment models in which the demand and the supply are assumed to achieve a **System Optimal** (SO) state, however they are not the focus of this research and will not be discussed further.

Wardrop's first principle is based on several implicit assumptions about the transport system. It is helpful to state these assumptions explicitly and discuss how they can be generalized. Table 1 summarizes the assumptions discussed in this section.

The first assumption is that the congestion effect is significant. The congestion effect describes the phenomenon that the performance of the transport supply degenerates as the number of trips within a time unit increases. For example, the travel time on a road increases as the hourly number of vehicles entering this road increases. The congestion effect causes a feedback loop between the demand and the supply. Traffic assignment models that consider this feedback loop are called **congested models**.

When the demand level stays low, the influence of the demand on the supply can be ignored. In this case, traffic assignment models can be simplified by only considering the influence of the supply on the demand while omitting the influence of the demand on the supply. This simplification breaks the feedback loop and makes the assignment problem trivial to solve. These traffic assignment models are called **uncongested models**. An uncongested model can be regarded as a special case of a congested model and the uncongested model can be used as a building block of the congested model.

The second assumption is that the travellers have perfect information on the travel costs of different routes when they make route choices. In this case, the user costs are object values and perceived in the same way by all travellers, therefore represented by constant numbers. Traffic assignment models based on such an assumption are called **deterministic models**.

On the other hand, the **stochastic models** assume that travellers make decisions based on incomplete information and unknown personal preferences. As a result, different travellers perceive route costs differently. To handle these uncertainties user costs are represented by random numbers in a stochastic model. The deterministic assignment can be seen as a special case of the stochastic assignment, where the variance of the uncertainty is zero.

Dimension	Assumption	Summary
Congestion effect	Congested model	Feedback loops exist between demand and supply
	Uncongested model	Supply determines demand
User cost perception	Stochastic model	Uncertainty in the perception of user cost
	Deterministic model	Perfect information of the user cost
Temporal variation	Static model	Demand and supply are time-invariant
	Dynamic model	Demand and supply are time-variant

Table 1: Overview of assumptions in traffic assignment models

The final assumption is that the demand and the supply stay in a steady state during time period of the analysis. All measures of the state of the transport system, such as traffic flow, traffic density, are independent from the instance when the measures are done. Traffic assignment models that follow this assumption are called **static models**.

The static assumption is valid only when the demand is constant for an analysis period that is long enough for the traffic state to propagate through the entire road network. The **dynamic models** simulate the temporal variation of the demand and the supply explicitly. They require an explicit modelling of the propagation of traffic along the routes.

Besides the above assumptions, another way to generalize the traffic assignment model is to generalize the route costs. In Wardrop's original statement, the route costs are determined by the travel times alone. In a generalized traffic assignment model the user cost may incorporate other performance values. For example, the monetary cost and the level of inconvenience caused by heavy vehicles can also influence the cost of a route.

2.2.3 Components of Traffic assignment Models

To address the feedback loop between the demand and the supply in a congested traffic assignment model, the solving procedure generally involves an iterative process which is illustrated in the following diagram.

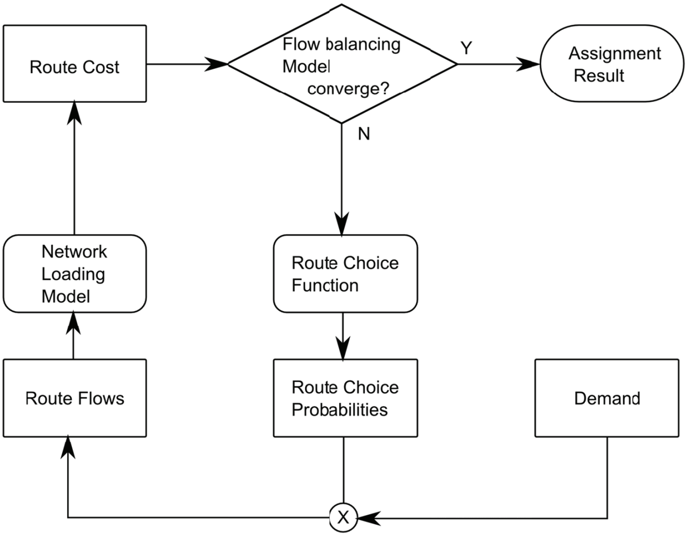

Figure 3: The general structure and components of a traffic assignment model

This process has the following components:

- Route generation model: The route generation model identifies possible routes for given OD-pairs. It provides route choice sets onto which travel demand is distributed.
- Route choice function: A route choice function models the route choice behaviour for travellers. With a given route choice set and the associated route costs, a route choice function predicts the fractions of travel demand distributed on each route.
- Network loading model: The network loading model simulates travellers moving in a network along their pre-determined routes. It produces the experienced route cost for travellers on different routes.
- Flow balancing model: The flow balancing model redistributes traffic flow among routes based on the experienced cost from the network loading model. It repeats the process until the equilibrium state is reached.

The relevant researches carried out on each of these components will be reviewed in the following sections.

2.3 Route Generation Models

This section reviews techniques, which are developed to explicitly generate reasonable routes between given OD-pairs for different purposes. According to(BOVY P. H., 2009) the route sets generated for different purposes have different requirements:

- Supply analysis and route guidance, in which the availability of routes between a given OD-pair is of the main interest.
- Demand model estimation, such as the estimation of route choice function parameters, in which small but representative route sets are sufficient.
- Traffic assignment, in which all routes that are likely to be chosen should be included.

The basic assumption underlying route generation models is that travellers will use routes that are they consider “optimal” in certain senses. For example, for a trip between a given OD-pair the travellers may use the route of the shortest distance, of the shortest travel time, or of the lowest general cost. Following this assumption trips between an OD-pair will use only one route, the “optimal” route. In reality, however, multiple routes can be observed for the OD-pair, which can be explained by the uncertainties in the definition of “optimal”. The source of the uncertainties can be summarized in the following categories:

- The first source of uncertainty is the imperfection of the information. Each traveller knows the performance values of only a certain part of the network. Furthermore, different travellers can perceive an object performance value differently.
- The second source of uncertainty is the variation of personal preference for different travellers. For a group of travellers, even if travellers have the same perceived performance values, the different personal preference will lead to different costs. For example, given the same choices some travellers prefer faster routes while others prefer scenic routes.
- The third source of uncertainty is the congestion effect. When travellers make pre-trip route choices they cannot predict the congestion effect precisely because they cannot know the route choices made by other travellers. If every traveller uses the fastest route, the route will become congested and no longer the fastest.

Because of these uncertainties, travellers consider not only the optimal route but also routes that are close to optimal. The route generation models are designed accordingly to find these close-to-optimal routes.

Based on how the route sets are constructed, the route set generation methods can be divided into two categories: the additive methods and the subtractive methods. An additive route generation method starts with an empty route set and progressively adds new routes to the set. A subtractive route generation method works in the reverse way: it starts with a large route set and gradually removes unrealistic routes from the set.

2.3.1 The Additive Route Generation Methods

The most common route generation method is based on the shortest path problem in the graph theory. Here the shortest path refers to the path with the minimal abstract cost. The shortest path problem itself is a well-defined problem, which can be solved by well-known algorithms such as Dijkstra's algorithm (DIJKSTRA, 1959).

A direct extension to the shortest path search is the K-shortest paths method which finds not only the shortest path but also the 2nd to the K-th shortest path. The K-shortest paths problem is a hot topic in algorithm research and many algorithms are proposed, for example (POLLACK, 1961).

However, the strict K-shortest paths methods are rarely applied directly as they tend to produce highly overlapping routes, which usually are not considered as valid alternatives. (VAN DER ZIJPP & FIORENZO-CATALANO, 2005) propose to incorporate constraints about detour and similarity of route into the K-Shortest path search algorithms to spread out the generated route sets.

In (AZEVEDO, COSTA, MADERIA, & MARTINA, 1993) a link elimination algorithm is proposed to address the spread out problem by searching shortest route iteratively and eliminating links of found routes from the network between iterations. This method can be implemented easily and requires low computational performance. However after removing links for several iterations, the connectivity of the remaining network cannot be guaranteed Further developments of this method try to maintain network connectivity by eliminating only a part of the links on found routes.

The link penalizing method developed by (DE LA BARRA, PEREZ, & ANEZ, 1993) has a similar structure like the link elimination method. But instead of eliminating links on found routes, it penalizes those links with additional costs. The advantage of penalizing over elimination is that it maintains the connectivity of the network. Further developments of this method involve different ways of determining the penalizing factors. For example, (RAMMING, 2002) determines the link penalty by incorporating travellers' network knowledge. (SCOTT, PABON-JIMENEZ, & BERNSTEIN, 1997) choose an appropriate penalizing factor to maintain a certain overlapping level among routes that are found in subsequent iterations.

The labelling method, initially developed by (BEN-AKIVA, BERGMAN, DALY, & RAMASWAMY, 1984), finds multiple routes between a given OD-pair by making repeated shortest path searches .Each iteration uses a different object function to define the link costs. These object functions are called labels, which consider factors such as time, distance, scenic, signal, and capacity. The original labelling method has two problems. First the number of routes between an OD-pair is restricted by the number of possible labels. Second, labels other than time and distance tend to produce indirect routes that are usually regarded as unreasonable by travellers.(DIAL, 2000) proposed a generalized labelling method which calculates link costs as linear combination of several labels. With a given set of labels, it can produce large numbers of routes by using different combinational weights.

The random methods model the traveller's perception of link costs as random numbers.(SHEFFI & POWELL, 1982) suggested a Monte Carlo based traffic assignment algorithm. In this algorithm, the link costs are assumed to be random numbers. The algorithm involves a given number of iterations. In each iteration, a sample that contains all links is drawn from the random link costs and based on this sample a shortest path is found. This assignment model can also be regarded as a route generation model, if the network loading part is omitted. (RAMMING, 2002) proposed to encode traveller's network knowledge into the distribution of link costs. (FIORENZO-CATALANO & VAN DER ZIJ, 2001) manipulated the variance of link impedance distribution to keep the rate of discovering new routes constant.

2.3.2 The Subtractive Route Generation Methods

Branch-and-bound methods have different algorithmic structures from the models reviewed in the previous section. Conceptually they start with a route set that contains all possible routes and then discard unrealistic routes according to a set of specifications. In practice, it is impossible to store the set of all possible routes explicitly, since it is too large. Instead, the Branch-and-bound methods enumerate these routes procedurally using a tree structure (see section 3.3 for more details). (FRIEDRICH, HOFSÄSS, & WEKECK, 2001) apply this method to identify routes in public transport networks. (HOOGENDOORN-LANSER, 2005) extended this idea on multi-modal transport networks with network partition and access and egress point identification.

(PRATO & BEKHOR, 2006) applied the branch-and-bound method to road networks and introduced several bounding constraints such as directional, temporal, loop, similarity and movement. (SCHLAICH, 2009) developed another set of bounding constraints with emphasis on the detour factors including the detour factor for the whole route and the detour factor for the back-end part of the route. (PILLAT, MANDIR, & FRIEDRICH, 2011) further extend this method by considering the back-end part detour factor for different backend part lengths. They also calibrated parameters of detour constrains based on observed routes from GPS trajectories and questionnaires.

2.4 Route Choice Functions

With a given route set, a route choice function calculates the probability of each route been chosen by travellers. Generally, the result is determined by both the properties of the given route and the preference of the given traveller.

2.4.1 Random Utility Theory

Random utility theory is the basis for modelling compensatory choice behaviour. It builds on the economical concept of **utility,** which is a generalized value assigned to each alternative in the **consideration set** by a decision-maker to represent his preference. The utility of an alternative is generally a weighted aggregation of several characteristics of this alternative, and the lack in one aspect can be compensated by another.

A consideration set contains a limited number of discrete, reasonable alternatives. Decision-makers try to maximize individual utility in the choice making process. However, from the modellers' perspective, information about the real choice conditions is incomplete and the utility assigned to each alternative by decision-makers cannot be known with certainty. (MANSKI, 1997) divided sources of uncertainty for the utility into four categories: unobserved alternative attribute, unobserved individual preference, measurement error and instrumental variable. In addition, the incomplete information availability to the decision-maker also increases the uncertainty.

Based on these assumptions the utilities are represented by random variables and usually it is impossible to predict the exact alternative that a decision-maker will select. Instead, the probabilities of choosing each alternative are predicted. Specifically the utility of alternative r considered by decision-maker n is represented as:

$$U_{nr} = V_{nr} + \varepsilon_{nr}$$

where V_{nr} represent the measurable or systematic part of the utility and ε_{nr} represent the random part capturing various uncertainties.

The systematic utility is a function of observable attributes of each alternative and decision-maker. It can be represented as:

$$V_{nr} = V_{nr}(X_{nr})$$

where $\boldsymbol{X}_{nr}$ represent the vector of attributes for alternative r considered by decision-maker n. This functional relationship is often assumed to be linear for the sake of simplicity of the application.

$$V_{nr} = \alpha_{nr} + \sum_{k \in K} \beta_{nk} \cdot X_{nrk}$$

where K represents the set of considered attributes, α_{nr} and β_{nrk} are parameters to be estimated. Specifically β_{nrk} represents the importance of attribute k of choice r for decision-maker n.

With these uncertainties, it is only possible to predict the probability of choosing each alternative. A utility is represented by the sum of a deterministic observable part and a stochastic latent part.

Based on the utility maximization assumption, with the utility of each alternative represented by a random variable, the probability of the decision-maker n choosing a specific alternative r is

$$\begin{aligned} P_{nr} &= P\left(U_{nr} \geq U_{nj}\right) \\ &= P\left(V_{nr} + \varepsilon_{nr} \geq V_{nj} + \varepsilon_{nj}\right) \forall j \neq r \end{aligned}$$

Different choice models are derived from this equation based on different specification of the random term.

2.4.2 Forms of Route Choice Function

In this section different route choice models are reviewed. The scope is restricted to the pre-trip choice in which the whole route is chosen before starting the trip. The different formulations of the route choice function follow the same framework provided by the random utility theory. The main differences between formulations lie in two aspects:

- The distribution form of the random error
- The approach to handle correlation between alternatives
- In the following sections, the major forms of route choice function will be reviewed with respect to these two aspects.

Multinomial Logit (MNL)

The Multinomial Logit is the simplest form in the Logit family. It assumes that the error terms of alternative utilities follow an independent and identical Gumbel distribution. One of its biggest advantages is its simplicity, because it can be written in a simple analytical form

$$P_{nr} = \frac{e^{V_{nr}}}{\sum_{j \in R_n} e^{V_{nj}}}$$

where R represents the set of alternative routes. The details of deduction can be found in (CASSETA, 2001). In a typical network, alternative routes may consist of many common links, thereby establishing complex correlation structures among alternatives. If MNL is applied, it would overestimate the use of routes with common links resulting in serious prediction errors. Nevertheless, this model is applied widely for its simplicity and many models are developed based on it, to address the correlation problem.

Nested Logit (NL) and Cross-Nested Logit (CNL)

The Nested Logit model is used when alternatives can be differentiated into groups in which members are close substitute for each other compared to members of other groups. Alternatives within each nest are correlated and alternatives across nests are not correlated. The probability of choosing an alternative is the product of choosing its nest, and choosing the alternative from within the nest, that is,

$$P_{nr} = P(r|R_n) = P(r|R_{nm})P(R_{nm}|R_n)$$

where R_{nm} is the nest m for traveller n.

The NL is more flexible than the MNL and can account for some similarities between alternatives if routes can be assigned into the nested hierarchy structure. However there is no general way to organize routes into nests. One particular example of nesting routes is provided by (BEN-AKIVA, BERGMAN, DALY, & RAMASWAMY, 1984) in which routes are generated by labelling methods. Because different labels can result in the same physical route, for example the shortest route and the fastest route can easily be identical, each physical route is regarded as a nest which contains all labels attached to it.

The Cross Nested Logit is an extension of the NL. Each route may belong to more than one nest with different degrees of membership. Links are often used as the basis for nesting. The problem of this method lies in its high complexity that cannot be transferred into a real network.

These two models and the Logit models are based on the Generalized Extreme Value (GEV) family of distributions. The random variables may be specified with different kinds of distributions such Gaussian, logistic and lognormal.

C-Logit (CL) and Path Size-Logit (PSL)

Efforts have been made to extend the MNL models to handle the correlation between alternatives especially for route choice modelling. (CASCETTA, NUZZOLO, RUSSO, & VITETTA, 1996) proposed the C-Logit model which extends the MNL model maintaining the simplicity of computation while improving the prediction accuracy for cases where correlations exist between alternative routes.

$$P_{nr} = \frac{e^{V_{nr}-CF_r}}{\sum_{j\in R_n} e^{V_{nj-CF_j}}}$$

When compared to MNL, the C-Logit model adds an adjustment term to the route utilities, which quantifies the amount of overlapping among routes. The correction term is called “commonality factor” (CF) and is computed for every route r. Different definitions of CF are proposed, for example

$$CF_r = \beta \ln \sum_{j \in R_n} \left(\frac{L_{rj}}{\sqrt{L_r \cdot L_j}} \right)^{\gamma}$$

The Path Size Logit (PSL) can be considered as an improvement of the C-Logit that incorporates behaviour theory in the utility adjusting process. As in the C-Logit the PS-Logit also introduces a correction term “Path Size” (PS), and the distribution function has a similar form.

$$P_{nr} = \frac{e^{V_{nr} + \ln PS_r}}{\sum_{j \in R_n} e^{V_{nj} + \ln PS_j}} = \frac{PS_r \cdot e^{V_{nr}}}{\sum_{j \in R_n} PS_j \cdot e^{V_{nj}}}$$

The path size is defined as

$$PS_r = \sum_{a \in \Gamma_r} \frac{l_a}{L_r} LSC_{ar}$$

where $\frac{l_a}{L_r}$ is the weight to the path size, LSC_{ar} is the link size contribution. In the simplest case the LSC_{ar} can be defined by the number of routes which share this link, or it can be extended by incorporating the relative length of different routes in the route choice set.

$$LSC_{ar} = \frac{1}{\sum_{j \in R} \left(\frac{L_j}{L_r} \right)^{\gamma} \cdot \delta_{ar}}$$

Probit and Hybrid Logit

The Probit model is based on error terms having a multivariate normal distribution, it is not constrained by the IIA property and an arbitrary covariance structure may be specified. (DAGONZO & SHEFII, 1977) and (YAI, IWAKURA, & MORICHI, 1997) applied the Probit model assuming the covariance is proportional to overlap length and the variance of the error term itself is also proportional to route length.

The main problem with the Probit model is that because there is no closed form assignment function, numeric techniques such as the Monte-Carlo method need to be used

The hybrid Logit model with combination of Gaussian and GEV error terms have been proposed by (MCFADDEN & TRAIN, 2000).

2.5 Network Loading Models

In the static traffic assignment the travel time of vehicles along a link is simply determined by a volume-delay function. A volume-delay function can take different forms based on different theoretical assumptions and empirical measurements. Of all the volume-delay functions, the one developed by Bureau of Public Roads (Bureau of Public Roads, 1964) is the most widely applied. An overview of researches on volume-delay functions can be found in (DEL CASTILLO & BENITEZ, 1995)

In the dynamic traffic assignment model, the network loading models can also be called traffic flow models, as they generally involve the simulation of traffic flow propagation. Traffic flow models can be roughly divided into three categories: macroscopic, mesoscopic and microscopic. Even for models in each category, there are still great variations in modelling traffic flows. However, the general tendency is that, from macroscopic to microscopic, the detail level of the model increases while the computational performance decreases. Which model is more suitable depends on the specific application context.

One advantage of dynamic network loading models is that they can propagate congestion effects through the network. In the static models, once the flow rate changes at origin, the effect will propagate immediately through the whole network. Another advantage of dynamic models is the ability of modelling queue formation and dissipation explicitly. The main disadvantages of dynamic models are that they are more computationally intensive.

2.5.1 Macroscopic Traffic Flow Models

The common property of macroscopic models is that they model traffic flow as a continuous indivisible flow; individual vehicles cannot be identified. The main advantage of a macroscopic model is its computational efficiency and low requirement on data preparation.

Dynamic macroscopic models such as the LWR model (LIGHTHILL & WHITHAM, 1955) and (RICHARDS, 1956) describe the evolution of traffic over time and space using a set of differential equations. Analogues of physical phenomena are used in defining the differential equations, such as those describing traffic like flows in fluids or gases. The solution to these equations can be obtained analytically or using numerical simulation.

(DAGANZO, 1995) discretized the LWR model, dividing roads into so called cells, small sections with equal length. The solution algorithm keeps track of the number of vehicles in each cell, and calculates the number of vehicles that enter and leave a cell as the simulation clock moves forward. These numbers are determined by the traffic densities in two adjacent cells.

2.5.2 Microscopic Traffic Flow Models

The microscopic models are on the other end of the spectrum of granularities. They model traffic flow as a group of individual vehicles with complex interactions. While in the macroscopic models the speeds, flows and densities of traffic flows are treated as fundamental variables; in the microscopic model these are only aggregated measures, resulting from the driving behaviour of all vehicles.

Examples of microscopic models are VISSIM (FELLENDORF, 1996), AIMSUN/2 (BARCELO & FERRER, 1997), Paramics (SMITH & DRUITT, 1994), MITSIMLab (TOLEDO & KOUTSOPOULOS, 2003)(YANG, 1997) and CORSIM (HALATI, LIEU, & WALKER, 1997).

2.5.3 Mesoscopic Traffic Flow Models

The mesoscopic models cover the spectrum of granularities between the macroscopic and the microscopic models. On the one hand, mesoscopic models model traffic flow broken down to single vehicles but on the other hand they greatly simplify the behaviour of individual vehicles. For example, many mesoscopic models determine the speed of a vehicle by the average traffic density on the road instead of modelling the car following behaviour explicitly.

Mesoscopic models can take various forms. CONTRAM (LEONARD & POWER, 1989) groups vehicles into packets according to their destination. Packets instead of vehicles act as entities to be routed through the network, and their speed on each road segment is derived from a speed-density function.

DYNAMIT (BEN-AKIVA M. E., 1996) took a similar approach to group individual vehicles into cells. The density in each cell determines the speed of the vehicles. The vehicles enter and exits cells but not overtake. The speed of the vehicles is determined by the condition in the cell, not the individual drivers' decisions.

Alternatively, DYNASMART (JAYAKRISHNAN & MAHMASSANI, 1994), FASTLANE (GAWRON, 1998), DTASQ (Mahut, 2001) adopted a queue-server approach, in which the road segment is modelled as a queue and the intersection is modelled as a service counter.

Cellular Automata (CA) are another type of mesoscopic models. In CA models the road is discretized into cells, each of which is either empty or occupied by one vehicle. The vehicles move from one cell to another following a minimalist set of behaviour rules, e.g. the rules developed by (NAGEL & SCHRECKENBERG, 1992).

3 Route Set Generation for Traffic Assignment

In this chapter a route choice set generation method is developed. It is the foundation of the traffic assignment framework using an explicitly generated route choice set before assigning the travel demand. It should produce realistic route choice sets, as they are the prerequisite of producing plausible assignment results. It also should be efficient so that it can handle large networks in real world applications.

3.1 Behaviour Assumptions

This section examines behaviour assumptions of travellers in identifying and choosing routes. These assumptions are drawn from related researches fields of transport modelling, econometrics and psychology. Based on these assumptions, a conceptual framework for measuring the quality of route sets can be established

3.1.1 The Nature of Route Choice Modelling

The route choice problem is a special case of general decision making problems that are studied in a wide variety of disciplines. In an abstract sense, choice models predict decision-makers' choice or probability of choice when they face a given set of alternatives. In the context of route choice models, travellers play the role of decision makers and the choice set consists of possible routes for given trips.

The general techniques of modelling choice set have already been studied in the field of transport planning (BEN-AKIVA & LERMAN, 1985), however modelling the route choice set is different from modelling other types of choice sets in the following aspects.

- The number of possible routes for a trip is very large because of the combinational nature of the route search. Typically, hundreds, even thousands of routes can be identified for a single OD-pair in a medium sized network.
- Different routes in a choice set tend to have overlapping links with each other. These overlapping links create correlation between different routes, which violates the assumption of independence and irrelevant alternatives (IIA) and makes the route choice problem more complicated.

Route choice is a complicated mental process, and the exact mechanism behind this process is not completely clear. There are active researches in disciplines such as psychology and geology (GOLLEDGE, 1999). According to (GARLING & GOLLEDGE, 2000), the route choice process involves two types of memories, the long-term high capacity memory that is permanent, and the short-term low-capacity memory that is volatile. Although past travel experiences and spatial knowledge are stored in the travellers' long-term memory, the decision-making must take place in the short-term memory. When facing a trip-planning task, relevant information is recalled into short-term memory where alternative routes are carefully considered.

Because of the limited capacity of the short-term memory, a mental mechanism should exist to limit the number of alternatives transmitted into the short-term memory. Route choice is carried out in two steps:

- In the first step, a large number of routes, which are clearly undesirable, are filtered out without carefully weighting every single aspect of each route to be left out. This step is called non-compensatory decision making since shortcomings in one aspect cannot be compensated by other aspects. Elimination by Aspects (EBA) developed by (TVERSKY, 1972) is one example for this step.
- In the second step, a limited number of remaining routes are carefully examined; different aspects are weighted to make the final decision. This step is called compensatory decision. The fundamental behaviour assumption in compensatory decision-making is that travellers will try to maximize their utility in choosing routes.

Following the discussion above, a theory framework for modelling route choice behaviour proposed by (BOVY & STERN, 1990) is adopted. In this framework, the route choice modelling for a given traveller is a two-stage procedure. In the first stage the route choice set is identified given the traveller's network knowledge such as network topology and traffic condition. In the second stage, the choices are made based on the utility maximization assumptions.

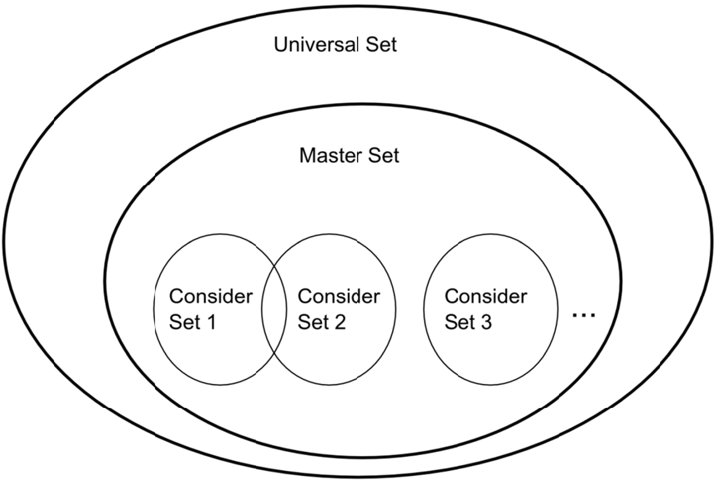

Figure 4: Relationship among choice set notions

To facilitate further discussions, it is important to distinguish several concepts about route sets. The relationships between these concepts are illustrated In Figure 4. For a given OD pair:

- The **Universal set** refers to the set of all physically possible routes. The number of routes in the universal set is very large. However, most of these routes are not likely to be considered by travellers.
- The **Master set** contains all reasonable routes which are likely be considered by some travellers. The definition of reasonable routes is crucial and will be further discussed in the next section. The master set generally is much smaller than the universal set, but the size is still beyond the mental capability of a single traveller. Besides many routes in the master set are very similar with each other, and will not be perceived as distinctive alternatives for one single traveller. The master set can be regarded as the union of the consideration sets of many travellers.
- A **Consideration set** contains routes that will be considered in a compensatory manner by a specific traveller. The number of routes in a consideration set must be within the mental capability of this traveller. Different travellers or traveller groups may have different consideration sets.

A consideration set can be constructed from travellers' perspective and from modeller's perspective. From the perspective of a specific traveller, the considered route set and the final chosen route can be regarded as deterministic. Although different travellers have different knowledge about the network, knowing different sets of routes and perceiving the cost of a route differently, for a specific traveller these factors are definite.

On the other hand, from the perspective of modellers, both the route sets and the final chosen routes are stochastic in nature, because it is impossible to measure the information availability and personal preference of each traveller with certainty. The possible errors are categorized as unobserved attributes, unobserved taste variation, measurement error and instrumental variations (MANSKI, 1997). On top of that, modellers usually deal with not the route set for one individual traveller, but the route set for a person group in traffic assignment. Even for a group of travellers with similar socio-economic background, there is still variation from one person to another within this person group.

3.1.2 Indications for Reasonable Routes

A reasonable route is a route that is likely to be considered by some travellers. The definition of reasonable route is the basis for generating master sets. This section introduces several measures to determine whether a route is reasonable.

The determinant factor for a reasonable route is its perceived cost. The basic premise of travellers' route choice behaviour is that they try to minimize the cost in travelling from the origin to the destination. Ideally, there is only one optimal route between each OD pair if the costs of all routes can be accurately determined.

As analysed in the last section, the route costs should be modelled as stochastic values. However, the link costs are normally represented as deterministic values in a transport model. These deterministic route costs should be regarded as one sample instance of the stochastic counterparts. Therefore, according to the probability theory, for a particular sample instance, routes that have larger costs than the shortest (in the sense of cost) route but still within a certain range are regarded as reasonable routes, as they are likely to be the shortest routes for other sample instances.

Global Detour Factor

The global detour factor of a path p is defined as

$$gd_p = \frac{g_p}{g_p^*}$$

Where gd_p is the total path cost of p, g_p^* is the direct cost of p, which refers to the cost of the shortest path between the tail and the head of path p.

$$g_p^* = \min(\{g_q : q \in \boldsymbol{P}_{p_t,p_h}\})$$

gd_p measures the cost of a path comparing with the shortest path. $gd_p \in [1,\infty)$, and $gd_p = 1$ if and only if p is the shortest path.

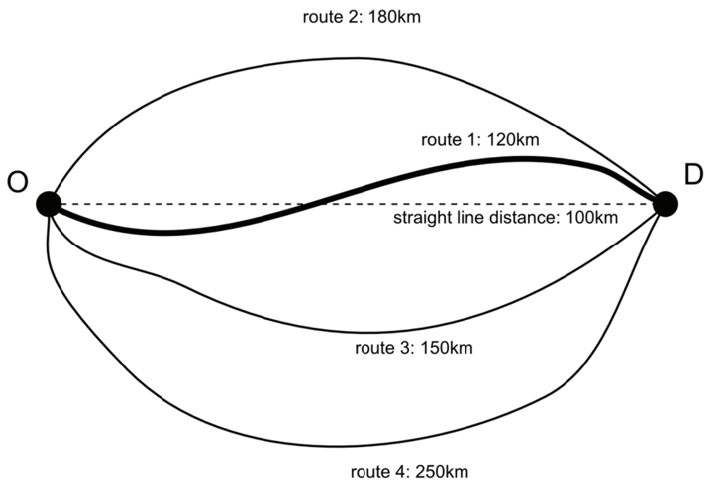

Figure 5: Example network for illustrating global detour factor

For example, In the simple network illustrated in Figure 5, the straight-line distance between the OD-pair is $100km$. Path 1 is the shortest path with a length of $120km$. According to the definition in this research, the global detour factor of path 2 is $180km/120km = 1.5$, while the global detour factor of path 1 is $120km/120km = 1.0$.

Local Detour Factor

A route-searching task for a trip from the origin to the destination can be regarded as many sub tasks of searching routes for consecutive legs. If the whole route is reasonable, then it is logical to assume that all sub-routes are also reasonable. It means each sub-route should not have large detour factor. For example, route 4 in Figure 6 is not reasonable because the segment between node 1 and node 2 have a large detour factor.

To differentiate between these two concepts, the detour factor for the whole route is called global detour factor, the detour factor for a sub-route is called local detour factor. Different segments of a given route may have very different local detour factor. The maximal local detour factor can be defined as:

$$ld^* = max\left(\left\{\frac{g_{p'}}{g^*_{p'}}, \forall p' \subset p\right\}\right)$$

It indicates the largest detour factor of all segments of a route.

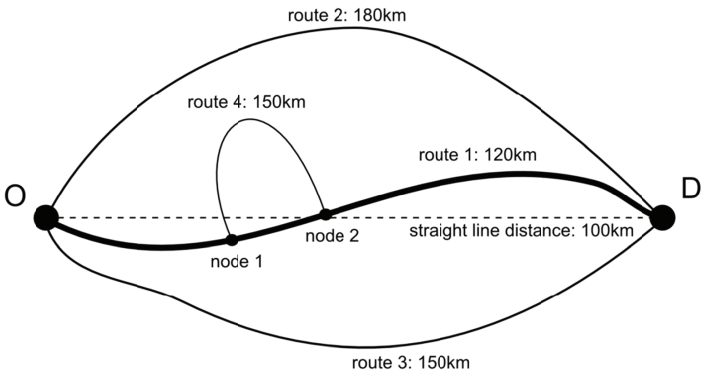

Figure 6: Example network for illustrating local detour factor

Local Optimal Threshold

Following the same thought of the local detour constraint, the route-searching task for a trip can be regarded as many sub-tasks. For a sub-task of a small scope, it can be assumed that travellers have the ability to assess the situation with certainty and make the objective optimal choice. This property of a route segment is called local optimal.

For example, when a traveller drives along a freeway and approaches an intersection, he faces two alternatives

- Keep driving on the freeway,
- Leaving freeway at the off-ramp and entering again at the subsequent on-ramp.

In this scenario the travellers almost always choose alternative 1, the optimal choice, even if the cost of two alternatives are close.

For a route segment p' that satisfies the local optimal property

$$g_{p'} = g^*_{p'}$$

For a given cost value lt, if all segments that have direct costs less than or equal with lt, are local optimal segments, then lt is a local optimal threshold.

$$g_{p'} = g^*_{p'}, \forall\{p' \subset p: g^*_{p'} \leq lt\}$$

There are many valid local optimal threshold for a given p, as if lt is a valid threshold, any $lt' < lt$ is also a valid threshold. However there is a unique maximal local optimal threshold for a given p

$$lt^* = max(lt_p)$$

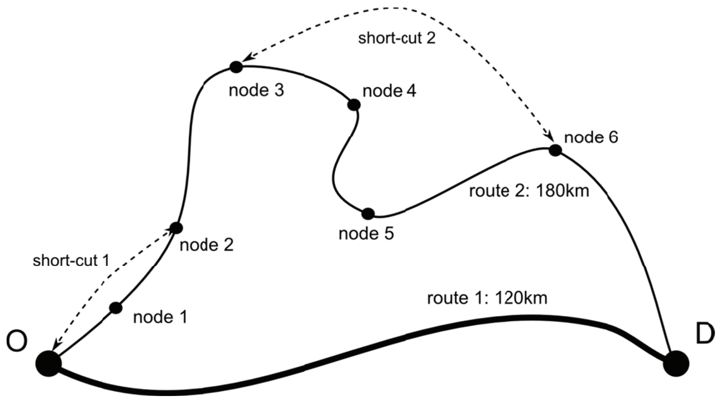

Figure 7: Example network for illustrating local optimal threshold

For example, the network in Figure 7 has two routes. There are two short-cuts on route 2 between origin and node 2, node 3 and node4. According the definition, all route segments that do not contain $[o, 1, 2]$ or $[3, 4, 5, 6]$ are local optimal segments. Any value smaller than the cost of the short-cut 1 is a valid local optimal threshold.

Hierarchical Index

In the last three indicators, the links in a network are treated uniformly. In the real world, however, transport networks are designed with an intrinsic hierarchical structure. Links in a network are assigned with different functional purposes in the network planning phase and are designed accordingly with different physical characters. For example, links in a road network are commonly categorized as follows:

- freeway,
- arterial road,
- collector road,
- local road.

Links of each level are designed with different standards, such as number of lanes, lane width, and design speed.

Links of higher level in this hierarchy are designed with high design speed, high capacity but with a low density in the network, while links of lower with lower speed and capacity but higher density. The ideal of this hierarchical design is that low level links play the role of collecting traffic to high level links, while high level links play the role of connecting places.

A reasonable route should respect this ideal. For example, in Figure 8 a reasonable route connecting OD should follow a certain sequence. After leaving origin, it first uses links of ascending levels, from local road to collector road, to arterial road, to freeway. Then it travels on the freeway until it approaches the vicinity of the destination. Finally, it uses links of descending levels to reach the destination. The hierarchical trait of this route can be drawn with a pyramid shape.

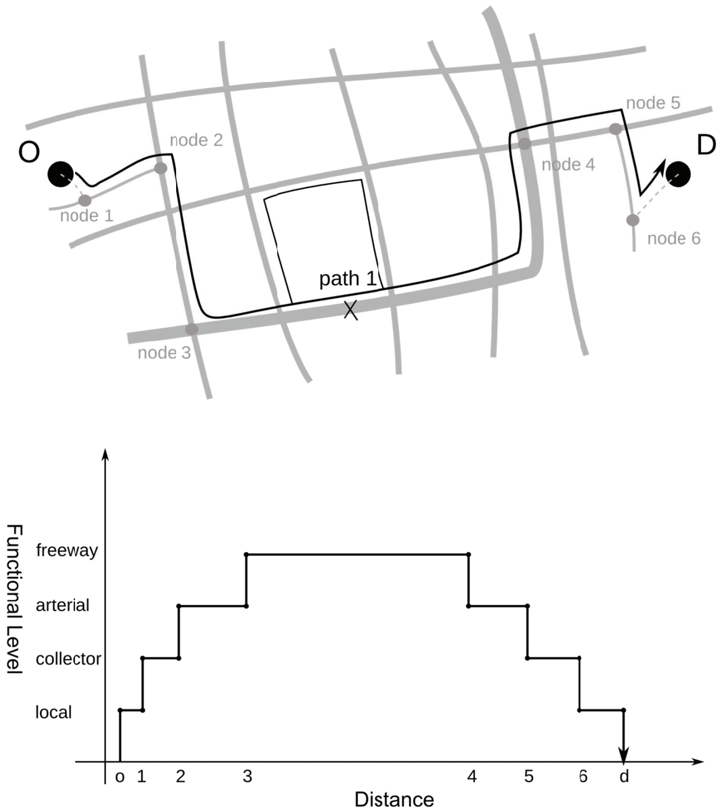

Figure 8: Example network for illustrating hierarchical index

In reality the hierarchy trait of a reasonable route is more complicated. It is not necessary to use links of all functional levels for each route. For instance, routes between OD-pairs within the same district of a city are not likely to use freeway. Also it does not always follow the pyramid shape if the high level links are interrupted in the middle. There is no simple one-formula formulation for the hierarchical constraint, and it also heavily influences the route enumeration process, thus it will be further dealt with in section 3.3.4.

Based on the indicators given in this section, a set of constraints can be defined to identify a master set of reasonable routes. These constraints and an efficient enumeration procedure are describes in section 3.3.

3.1.3 Composition of Route Sets

Usually a master set contains large number of routes, which is far beyond the number that can be considered by an individual traveller. At the same time, many routes are very similar with each other. For these reasons, the master set is treated as the union of consideration sets of many travellers and consideration sets for individual travellers need to be assembled by picking small sub-sets from the master set.

The composition of a consideration set can be assessed by the following measures.

- The first measure is the size of the consideration set, or the number of routes contained. As discussed in section 3.1.1, routes in a consideration set need to be fit into traveller’s short-term memory, which has a limited capacity.
- The second measure is the dissimilarity level of routes in a consideration set. The routes in a consideration set should be mutually dissimilar enough, so that the can be regarded as independent alternatives.

The procedure for assembling consideration sets based on these two measures is described in section 3.4.

3.2 Overall Architecture

In the section above the assumptions of traveller’s behaviour in identifying route set are discussed. This section describes the overall structure of the route set generation procedure.

The overall structure of route set generation contains two stages as illustrated in Figure 9.

- The first stage is the branch-and-bound enumeration, which generate the master set from the universal set. The basic idea behind this generation procedure is called **constrained enumeration**, which enumerates possible routes systematically while omitting certain routes based on given constraints.
- The second stage is the consideration set assembling, which assembles routes from the master set into consideration set for individuals and groups. It maintains the constraints on the composition of consideration sets, namely the size and the similarity index.

Effectively, the generation procedure works as a filter from the universal set to the master set and then to consideration set.

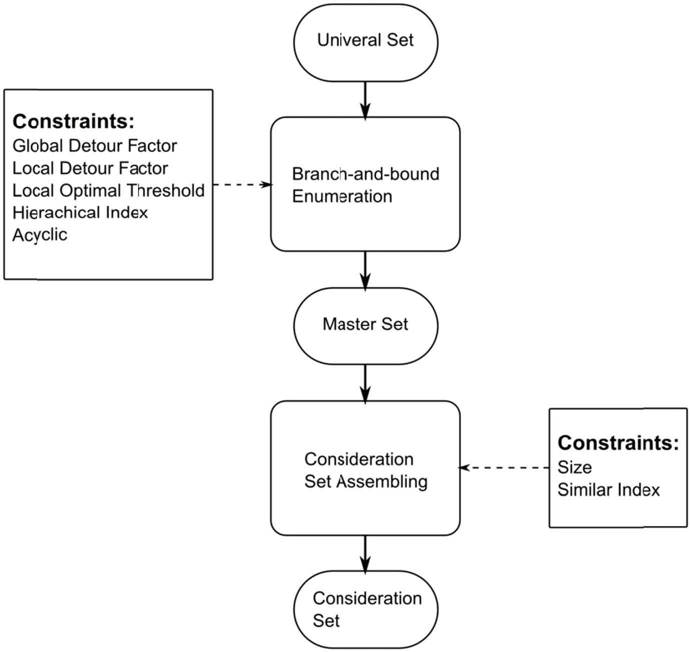

Figure 9: Overview of architecture of the route set generation procedure

3.3 Branch-and-Bound Enumeration

This section describes the actual procedure of identifying reasonable routes based on the measures presented in section 3.1.2. Section 3.3.1 describes the principle and tree-building procedure of the branch-and-bound technique, which enumerates routes in a network. Section 3.3.2 describes the quality constraints for individual routes. Section 3.3.2 extends the basic technique by building the enumeration tree with node sequences instead of single nodes. This extension reduces the size of the route search tree while produce equivalent route sets.

3.3.1 Basic Enumeration

The basic strategy of enumerating possible routes in a road network is called branch-and-bound. It is originally developed in the field operational research to solve discrete and combinatorial optimization problems (LITTLE, MURTY, SWEENEY, & KAREL, 1963). A branch-and-bound procedure involves a **branch method,** which splits a large initial solution domain into several smaller sub-domains, and a **bound method,** which estimates whether a sub-domain contains desired solutions. Based on this estimation, non-promising sub-domains are discarded to reduce the size of the solution domain. Then, both methods are applied to the remaining solution domain to further reduce the size. This procedure is repeated until solution domain is small enough to pinned down the solutions.

As reviewed in section 0, the Branch-and-bound technique has been applied to different types of route generation problems. In the context of a route set generation problem, the universal set is the initial solution domain. The branch method organizes routes in the universal set into a tree structure. The bound method determines whether a branch of this tree structure contains reasonable routes. Branches that do not contain reasonable routes are discarded en masse.

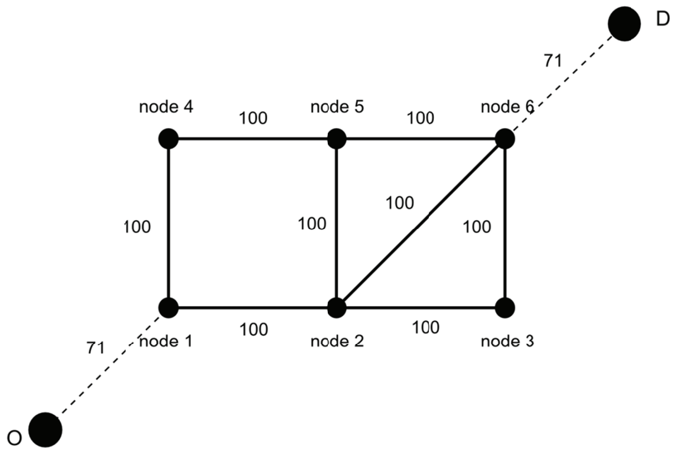

Figure 10: Example network (with link costs) to illustrate the branch-and-bound enumeration procedure

A simple network, as depicted in Figure 10 is introduced to illustrate the tree building procedure. This network contains 1 OD-pair, 6 nodes and 7 links. Travelling costs for all links are given. A route or a partial route can be denoted as a sequence of nodes. For example the shortest route r^*_{od} from o to d is denoted as $o126d$. For two node sequences r_1 and r_2, they have a parent-child relationship if r_2 can be constructed by concatenate r_1 with one more node

$$r_2 \equiv r_1 + [n], \exists n \in \boldsymbol{N}$$

where $\boldsymbol{N}$ is the set of nodes in the network. In this relationship, r_1 is the parent of r_2, while r_2 is the child of r_1. For example $o126$ is the parent of $o126d$. A parent-child relationship between r_1 and $r_2$can be denoted as

$$r_1 \subset r_2,$$

hence $o126 \subset o126d$.

The parent-child relationship is transitive, which means If $r_1 \subset r_2$ and $r_2 \subset r_3$ then r_1 is the **ancestor** of r_3. Based on this definition every route has an ultimate ancestor that contains only one node. All node sequences starting with o can be organized into a tree structure, as they all share an ultimate ancestor o. This tree structure can be called as the **universal route tree** as it contains all routes in the universal set. A sub tree of the universal route tree that contains all routes in the master set can be called as the **master route tree**.

It is clear that the universal route tree is large even for such a simple network. It is difficult to calculate the size of a universal route tree precisely, however the magnitude can be roughly estimated. In a universal route tree each route is represented by a leaf node and the universal route tree is almost full. The number of routes can be estimated by the number of leaf nodes in the route tree.

According to graph theory, the number of leaf nodes in a full n-tree is

$$S \propto d^n,$$

where S is the number of routes in the route tree,d is the depth of the tree, n is the degree of a non-leaf node. For a network of m nodes, the depth of route tree can be estimated as

$$d \propto \sqrt{m}.$$

If the network is further assumed to be grid-like, then most intersection has 4 legs, and the degree of non-leaf nodes can be estimated as

$$n \approx 4 - 1 = 3.$$

The number of routes is estimated as

$$S \propto d^n \propto 3^{\sqrt{m}}.$$

The above equation indicates that the size of a universal route tree is polynomially related with the width and exponentially with the depth. For a network of 2000 nodes, the size of the universal route tree is of an intractable magnitude

$$S \propto 3^{\sqrt{2000}} \propto 10^{21}$$

In a branch-and-bound enumeration procedure, the universal route tree is not generated at one-time but constructed gradually for two reasons. One reason is that, a universal route tree is generally huge and they simply cannot be fit into the computer memory. The more important reason is that, in the constructing process a new added routes can be checked and under certain condition the whole branch under it, which may contains large number of routes, can be discarded without inspecting each of them. This characteristic improves the enumeration performance greatly.

```
1   search(seq):
2       node = seq.last
3       for new_node in adjacent_nodes(node):
4           if has_cycle(seq + [new_node])
5               continue
6           If test_constraints(seq + [new_node]):
7               continue
8           if is_complete(seq + [new_node]):
9               record(seq + [new_node])
10          else:
11              search(seq + [new_node])
12      return
```

Figure 11: Pseudo code for the basic branch-and-bound enumeration

The algorithm for constructing a master route tree procedurally is illustrated with the pseudo code in Figure 11. It is a recursive procedure. The input is a sequence of nodes **`seq`**; as the initial value **`seq=[o]`**, where **`o`** is the origin. Line 2 and line 3 get the head of the current node sequence and iterate over all adjacent node of it. For each adjacent node **`new_node`** the following conditions are tested:

- If the **`new_node`** has already presented in **`seq`**, then it is a cyclic route and should be discarded (Line 4 and 5).
- If the new sequence produced by concatenating **`seq`** and **`new_node`** can not pass all constraint tests, then it should be discarded (Line 6 and 7).
- If the **`new_node`** is a destination node, then a reasonable route is found by concatenating **`seq`** and **`new_node`**, and it will be recorded in the master set (Line 8 and 9).
- If the **`new_node`** is neither discarded nor recorded then the procedure is called recursively with and extended node sequence. (Line 10 and 11)

In this recursive procedure if a node sequence is discarded at one search step, then all its child sequences can also discarded. This property improve the efficiency of the enumerating procedure greatly. In order to maintain this property, each constraint must holds the following **invariant,** that if a node sequence fail to pass the test then all child sequences will also fail. This invariant holds for all constraints proposed in this research and it must be taken into consideration in defining new constraints.

The recursive searching process initiated by applying this algorithm to the network in Figure 10 can be depicted as a tree structure, which is illustrated in Figure 12. Each edge in the tree represents a searching step; each vertex represents a newly incorporated node. The node sequence found at a vertex can be reconstructed by concatenating nodes marked in current vertex and all its ancestors. For example, the vertex marked with 3 in the left bottom corner represents the node sequence of $o1452363$. Edges are marked with different symbols to indicate their status. Edges leading to a discarded node sequence are marked with a dashed line. Edges leading to a complete route are marked with a bold line. As the tree structure indicated, there are 6 valid routes between o and d, which form the master set.

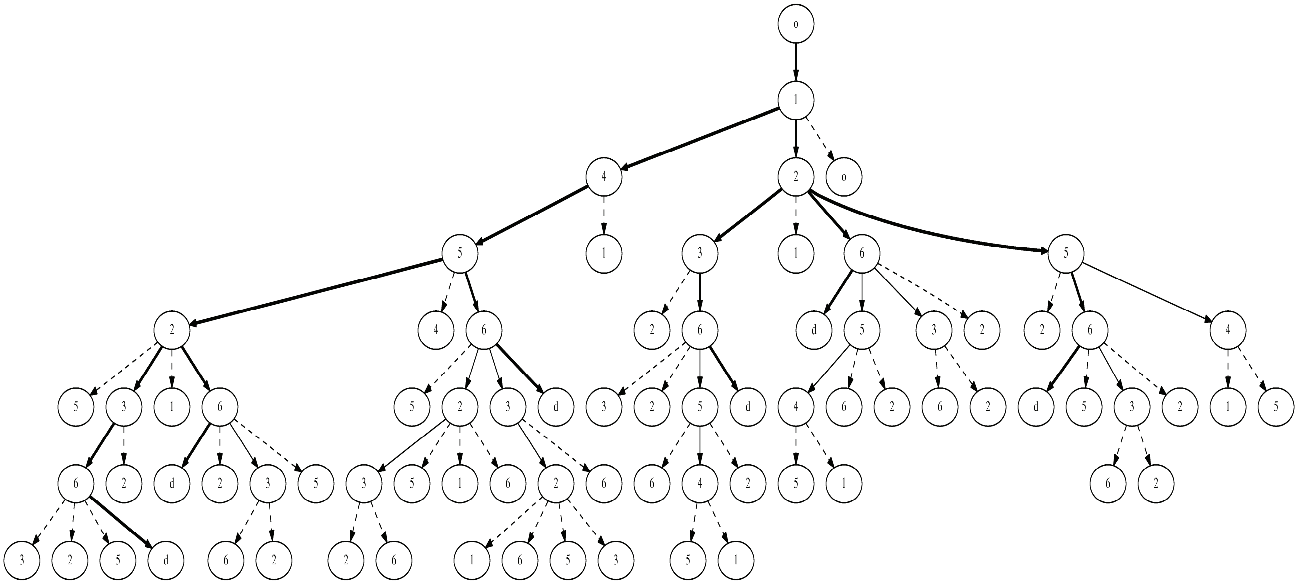

Figure 12: Route search tree for applying branch-and-bound method to the network in Figure 10

3.3.2 Constraints

In this research the following constraints are used in the enumeration procedure.

Acyclic Constraint

Cyclic routes are regarded as unreasonable even though they may have small detour factors. For each cyclic path, travellers can always reduce the travel costs by avoiding the cycles. For example in Figure 13 the cyclic path 3 has a corresponding simple path 2. Although path 3 and path 1 have the same detour factor and both comply with the global detour constraint, path 3 is not likely been chosen by travellers. The difference between path 3 and path 1 is not the actual costs, but how the travellers perceive the costs. Because of the cycle, it is easier for the travellers to realize path 2 is a better choice than path 2.

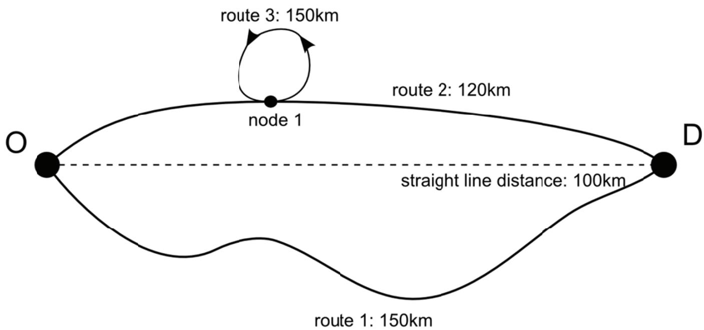

Figure 13: Example network for illustrating acyclic constraint

This idea can be formalized as acyclic constraint. It requires that for a reasonable path $p = \{v_1, \dots, v_n\}$ the following condition holds:

$$v_i \notin p \backslash \{v_i\}, \forall 1 \leq i \leq n \qquad (1)$$

Where $p \backslash \{v_i\} \coloneqq \{v \in p : v \neq v_i\}$ which means the path may contain each vertex only once.

Global Detour Factor Constraint

The global detour factor constraint requires that for a reasonable path p the following condition holds.

$$gd_p \leq \overline{gd}$$

Where $\overline{gd}$ is the parameter of the global detour factor constraint.

One way to estimate the parameter is to use the German "Richtlinien für die integrierte Netzgestaltung RIN" (FGSV, 2008). It defines the detour factor with two slight differences comparing with the definition in this research. First, the path length is measured in distance instead of generalized cost. Second, the path length is compared with the straight-line distance instead of the direct length. Nevertheless, the reference values of the detour factor limit for different level-of-service (LOS) can be adopted. As depicted in Figure 14, for a given LOS the allowed detour factor limit decreases as the direct distance between OD pair increases.

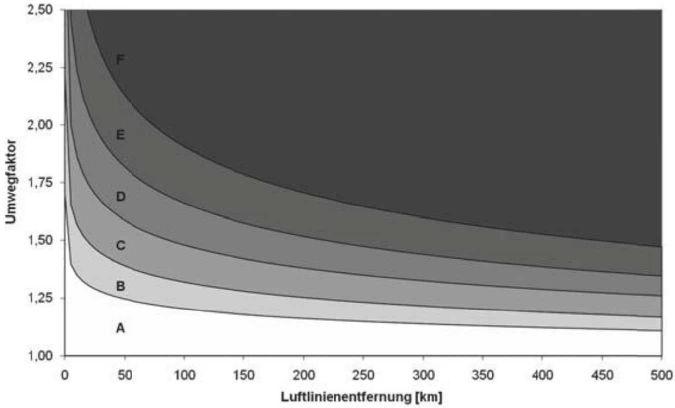

Figure 14: Global detour factor limit for different LOS from "Richtlinien für die integrierte Netzgestaltung" (FGSV, 2008)

For example, if the LOS of reasonable routes are assumed to have at least C level, then for path 4 in Figure 5

$$gd_{p4} = \frac{250km}{120km} = 2.08 > A(120km) = 1.48$$

which means it is not reasonable. Here $A^C()$ represents the function curve for the C level of LOS in Figure 14.

Local Detour Factor Constraint

The local detour factor constraint requires that the following condition holds.

$$ld_{p'} \leq A\left(g_{p'}^{*}\right), \forall p' \subset p$$

Which means for each sub path p' of the path p, the detour factor constraint holds. If one detour factor limit is applied to all sub-path then the local detour constraint can be simplified as

$$ld^* = max\left(\left\{\frac{g_{p'}}{g^*_{p'}}, \forall p' \subset p\right\}\right) \leq \overline{ld}$$

where $\overline{ld}$ is the parameter of the local detour factor constraint.

In the example network depicted in Figure 6, the local detour factor constraint can be used to discard paths that make big detours for small parts of the whole trip. For example, path 4 and path 3 both have the same global cost 150km and comply with the global detour factor. However, for path 4 the sub-path between node 1 and node 2 violates the local detour factor and path 4 should be discarded.

Local Optimal Threshold Constraint

The local optimal threshold constraint requires that the following conditions hold:

$$lt^*_p = ma\,x(lt_p) \leq \overline{lt}$$

where lt^*_p is the maximal local optimal threshold for a given p, $\overline{lt}$ is the parameter for the local optimal threshold. This constraint requires that each path segment p' of a reasonable path p to be an optimal path segment if its optimal alternative has a cost lower than $\overline{lt}$.

The route set that comply with a given parameter $\overline{lt}$ is recorded as $\boldsymbol{P}^{\overline{lt}}$. It can be shown that, for two different parameters $\overline{lt}_1, \overline{lt}_2$if $\overline{lt}_1 > \overline{lt}_2$ then $\boldsymbol{P}^{\overline{lt}_1} \subseteq \boldsymbol{P}^{\overline{lt}_2}$. According to the definition, for any path $p \in \boldsymbol{P}^{\overline{lt}_2}$,

$$g_{p'} = g^*_{p'}, \forall\{p' \in p: g^*_{p'} < \overline{lt}_1\}$$

which means all route segments of a size smaller than β_1 are optimal.Because $\beta_1 > \beta_2$, it is apparent that

$$\{p' \in p: g^*_{p'} < \overline{lt}_1\} \supseteq \{p' \in p: g^*_{p'} < \overline{lt}_2\}$$

Which implies

$$g_{p'} = g^*_{p'}, \forall\{p' \in p: g^*_{p'} < \overline{lt}_2\}$$

also holds. It means $p \in \boldsymbol{P}^{\overline{lt}_1} \Rightarrow p \in \boldsymbol{P}^{\overline{lt}_2}$, which is equivalent to $\boldsymbol{P}^{\overline{lt}_1} \subseteq \boldsymbol{P}^{\overline{lt}_2}$.

Because of this property, the local optimal constraint is not only a constraint for the generation of individual routes. It also plays an important role in the assembling of route sets, which will be further discussed in section 3.4.

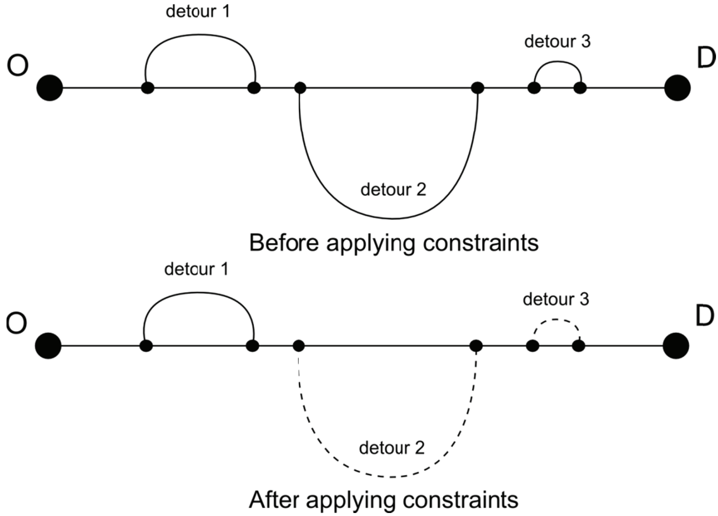

Figure 15: An example to show the different roles of the global detour factor, local detour factor and local optimal threshold

The roles of the constraints on global detour factor, local detour factor and local threshold are summarized in Figure 15. In the network there is a direct connection and several detours between the OD-pair. The three detours have different sizes. After applying all three constraints, routes that uses detour 2 and detour 3 are discarded. The intention of the global detour factor constraints is to discard routes that uses detour 2. These routes are apparently worse than the shortest route. The intention of the local detour factor constraint and the local optimal threshold constraint is to discard routes that uses detour 3. These routes contains detour in a small range, which does not count as a valid alternative.

3.3.3 Generalized Enumeration

The generalized enumeration is inspired by the application of the branch-and-bound technique in public transport network(FRIEDRICH, HOFSÄSS, & WEKECK, 2001). The main difference between the public transport network and the private transport network is that the public transport network consists of transport lines interconnected with transfer stops and limited number of transfers, e.g. 4, are allowed for a trip. On the other hand, in private transport networks travellers make decisions about which turn to take at each intersection.

These characteristics lead to broader but shallower universal route set. Because the size of a universal route tree is polynomially related with the width but exponentially with the depth, as analysed in last section, the universal route tree for public transport in a network is generally smaller than that for private transport in the same network.

```
1   search(seq):
2       node = seq.last
3       for new_node_seq in adjacent_node_seqs(node):
4           if has_cycle(seq + new_node_seq)
5               continue
6           If test_constraints(seq + new_node_seq):
7               continue
8           if is_complete(seq + new_node_seq):
9               record(seq + new_node_seq)
10          else:
11              search(seq + seq + new_node_seq)
12      return
```

Figure 16: Pseudo code for the generalized branch-and-bound enumeration

Based on the above analysis the basic enumeration algorithm presented in the last section is generalized. The algorithm of this generalized enumeration procedure is shown in Figure 16. It has the identical structure, compared with the algorithm of the basic enumeration in Figure 11. The main difference is that, in each search step (line 4) a new node sequence instead of a node is examined.

There are different strategies of building adjacent node sequences. One trivial method is that each adjacent node sequence contains a single adjacent node as depicted in the left half of Figure 17. In this case, the new algorithm is equivalent to the basic algorithm. This strategy is not very useful, but it shows that the generalized enumeration is an extension to the basic enumeration.

A more meaningful strategy, called local optimal strategy, is illustrated in the right half of Figure 17. It is based on the assumption that within a small context the travellers are capable of find the optimal route. Here the context is defined as an adjacent area around the traveller within a given diameter, and in the diagram it is depicted as the shaded circle. In order to build adjacent node sequences the following steps are followed.

- With the current node n and an adjacent area of diameter co, all **boundary nodes** are identified. A boundary node is a node outside the adjacent area while has at least one adjacent node inside the adjacent area.

- From the current node n, a path to each boundary node is found. Each path is an adjacent node sequence.

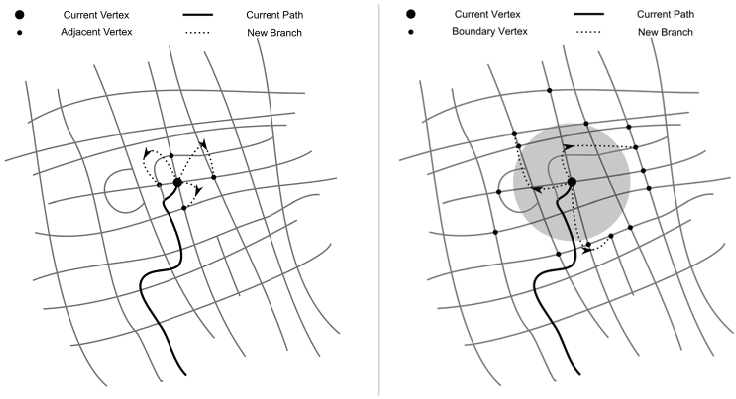

Figure 17: Comparing two strategies of finding adjacent node sequence. (Left for immediate neighbours. Right for direct connection within cutoff co)

When co is smaller than the cost of any single link the boundary nodes are exactly the adjacent nodes. An increasing co leads to more boundary nodes for one search step but less steps before reaching a destination zone, hence a wider and shallower search tree. In this respect, the boundary nodes play a similar role like the transfer stops in the public transport network.

If the cut-off co equals the parameter of the local optimal threshold constraint, and the paths between the current node and boundary nodes are the shortest paths, then it can be shown that the basic enumeration and extended enumeration produces the same result.

3.3.4 Enumeration with Road Hierarchical Structure

To enforce the hierarchical index is different than other constraints. The road hierarchical level has not only physical effects such as different capacity and free flow speeds, it also has psychological effect, which cannot be measured easily. For example, research shows that whether and how a road is marked on the map will have influence on travellers' route choice, despite the measurable travel cost has not changed.

Link costs describe the effect of each individual link while the hierarchical structure gives a high level overview of the network. In real life travellers are not working link by link in planning their trips, instead they will take advantage of the high level knowledge of the network. For example, based on their understanding of the network structure travellers will naturally divide the route searching task for a long distance trip into three steps:

- find a way entering the highway network,
- find a route in the highway network leading to exit near the destination,
- and find a way from exit to the destination.

This behaviour can be addressed implicitly or explicitly in the route set generation procedure. The implicit method maintain the structure of the enumeration structure while adjust the link costs to reflect the effect of the road hierarchy. For example, turns that connect high level roads to low level roads can be penalized to dissuade travellers from leaving the high level roads.

The explicit method extend the enumeration procedure into following steps:

- Separate the network into two sub-networks: the network of primary road and the network of the subordinate road. One possible scheme is that freeway and arterial road make up primary roads; collector road, and local road make up subordinate road.
- For a given OD-pair find the route segments connecting origin with nearby ingress nodes for the primary network. This is done with shortest path search within the subordinate network.
- For a given OD-pair find the route segments connecting destination with nearby egress nodes for the primary network. This is also done with shortest path search within the subordinate network.
- Enumerate route segments in the primary network that connecting ingress nodes and egress nodes found in the last two steps.
- Concatenate ingress segments, segments on the primary network and egress segments to form complete routes.

3.4 Assembling of Consideration Sets

The final step of route set generation is to assemble consideration set for individuals and groups based on the master set. In the assembling process, two constraints are maintained: the size of the consideration set and the dissimilarity between routes. Section 3.4.1 introduce a similarity index that considers both topology and geology discrepancies between two routes. Section 3.4.2 describes a stochastic procedure to assemble consideration set for individuals, while Section 3.4.3 describes the procedure of clustering routes in the master set, based on which consideration sets for groups are assembled.

3.4.1 Similarity Index

A **similarity index** is needed to measure the mutual dissimilarity of routes. For example, the "commonality factor" defined in (CASCETTA, 2001) is a popular form of similarity index:

$$C = \frac{L_{ij}}{\sqrt{L_i \cdot L_j}}$$

where L_{ij} is the length of the common parts of route i and route j, while L_i and L_j are the length of these two routes.

However, this definition only regards the topology of routes and the geographical disparities are ignored. Two parallel routes that runs close with each other should be considered more similar than two routes that runs more separately, even if both pairs have identical commonality factor. The importance of the geographical disparity of routes has long been realized in the field of route choice modelling (RAMMING, 2002) and hazard evacuation (AKGÜN, ERKUT, & BATTA, 2000).

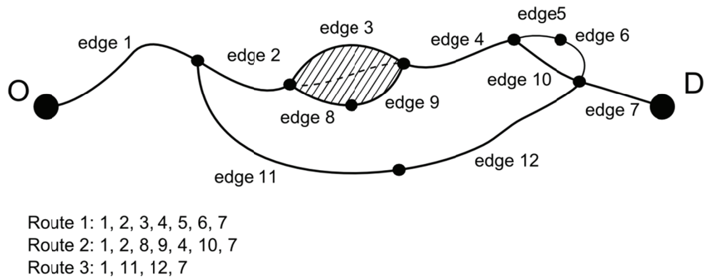

Figure 18: Example to illustrate the similarity index that considers geographical disparities

A similarity index that also considers the geographical disparity is proposed here. For two routes $p_1$and p_2, if a route segment p' belongs to both routes and no extension of it belongs both routes, then p' is an **overlapping segment.** In the network of Figure 18, $[o, 1, 2]$ is the overlapping segemtn of p_1 and p_2, while $[1, 2]$ is not.

Comparing two routes between a given OD-pair, a series of overlapping segments can be found. For any two routes at least two overlapping segments exist, namely the one contains origin and the one contains destination. The part between two adjacent overlapping segments is called a **mesh**. For example, between node 2 and node 4, node 5 and node 7, are two meshes for p_1 and p_2

A weight is defined for each overlapping segment or mesh. For an overlapping segment p', the weight equals the length

$$w(p') = L(p')$$

For an mesh p'', the weight is calculated as

$$w'(p'') = L^*(p'') \cdot s\left(\frac{A(p'')}{L^*(p'')}\right)$$

$$w''(p'') = L^*(p'') \cdot \left(1 - s\left(\frac{A(p'')}{L^*(p'')}\right)\right)$$

Where $L^*(p'')$ is the direct distance between two ends of the mesh, and $A(p'')$ is the area surrounded by the mesh. $\frac{A(p'')}{L^*(p'')}$ estimates the average distance between two branches of the mesh. $s()$ is a continuous monotonic increasing function where $s(0) = 0$, $lim_{x\to\infty} s(x) = 1$. The function separate the size of mesh into the overlapping portion and non-overlapping. $\frac{A(p'')}{L^*(p'')} = 0$ means that the two branches of the mesh are completely overlapping.

The similarity index of p_1 and p_2 can be calculated with

$$\theta_{p_1,p_2} = \frac{\sum_{p' \in P^{os}_{p_1,p_2}} w(p') + \sum_{p'' \in \boldsymbol{P}^{mesh}_{p_1,p_2}} w'(p'')}{\sum_{p'' \in \boldsymbol{P}^{mesh}_{p_1,p_2}} w'(p'') + \sum_{p'' \in \boldsymbol{P}^{mesh}_{p_1,p_2}} w''(p'') + \sum_{p' \in \boldsymbol{P}^{os}_{p_1,p_2}} w(p')}$$

Where $\boldsymbol{P}^{os}_{p_1,p_2}$ is the set of overlapping segments, and $\boldsymbol{P}^{mesh}_{p_1,p_2}$ is the set of meshes. $w()$ is the weight of a mesh or an overlapping segment.

Based on the definition above, the similarity index can be calculated in three steps:

- The first step compares two routes and divides them into overlapping segments and meshes. By treating routes as node sequences, Finding overlapping node segments is conceptually similar with the problem of finding common strings in computer science. Algorithm proposed in(MYERS, 1986) can be adopted.
- The second step is to calculate the weight of each overlapping segment or mesh. One important task in this process is to calculate the area surround by a mesh. The essence of this task is to calculate the areas of arbitrary polygons, which is a well-studied problem in GIS. See (NIEVERGELT & WIDMAYER, 1997) for a detailed discussion.
- Combine the weights of all overlapping segments and meshes and get the similarity index.

With this definition of the similarity between routes, route choice sets for individual travellers and person groups can be build up.

3.4.2 Individual Consideration Sets

Given the constraints on route set composition in section 3.1.3, the consideration set for an individual traveller can be generated in the following procedure.

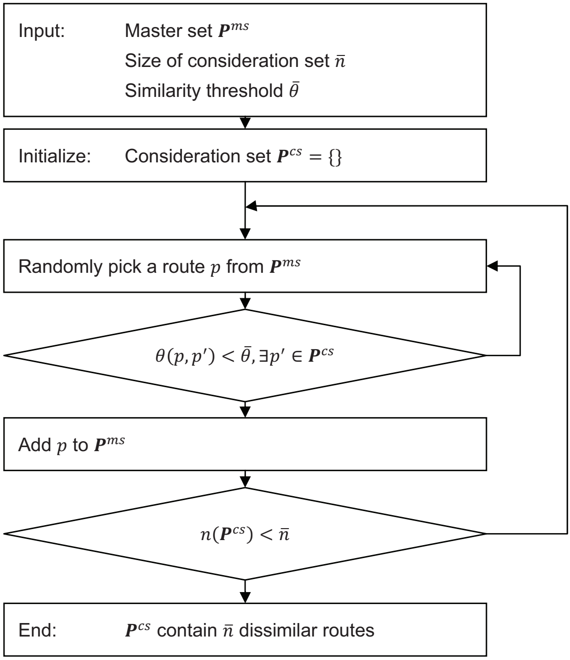

Figure 19: The procedure of assembling consideration set for an individual traveller

As depicted in Figure 19, the consideration set is initialized as an empty set. There are two levels of iteration in the process. In the inner iteration a route is picked randomly from the master set, and it is compared against each route that already exists in the consideration set. If this route is too similar to any existing route in the consideration set, it will be rejected and a new route is picked. The inner iteration repeats itself until a valid new route is found and added to the consideration set. The outer iteration repeats itself until the given number of routes are found for the consideration set.

The output of this generation procedure is a route set of a given size, and routes in this set are dissimilar enough with each other so that they are perceived as valid alternatives by a traveller.

3.4.3 Group Consideration Sets

The procedure of generating individual route choice sets is sufficient for route choice modelling research, in which routes only need to be generated for a small number of travellers. On the other hand, in applications such as traffic assignment the route choice sets of a large number of travellers need to be generated. Generating a consideration set individually for each traveller is computationally infeasible.

One common approach to deal with this problem is to divide the travellers into several person groups based on their socio-economic backgrounds. Then all travellers in one person group are assumed to share an identical consideration set. With this assumption only a small number of consideration sets need to be generated, one for each person group. Although this approach makes handling large number of travellers computationally feasible, it fails to capture the individual variations of travellers, which leads to the unrealistic concentration of traffic on a small number of routes.

This research solves this contradiction between performance and fidelity by modelling the group consideration set as a nested structure. A group consideration set contains limited number of **strategies**. Each strategy is a set of routes that are perceived as similar by a group of travellers. For example, there are five routes in the network depicted in Figure 20. From the network, it is clear that route 1 and route 2 are similar, route 4 and route 5 are similar. These routes can be divided into 3 strategies and two routes from different strategies are perceived as different from each other.

The prediction of the route choice probabilities for a person group becomes a two-stage process: the first stage predicts which strategy is chosen, the second stage predicts which route in that strategy is chosen. The small number of strategies in the first stage makes this computationally efficient, while the second stage reflects the individual variations in a person group. This section focus on building the nested route set, the details of the route choice is given in section 4.2.

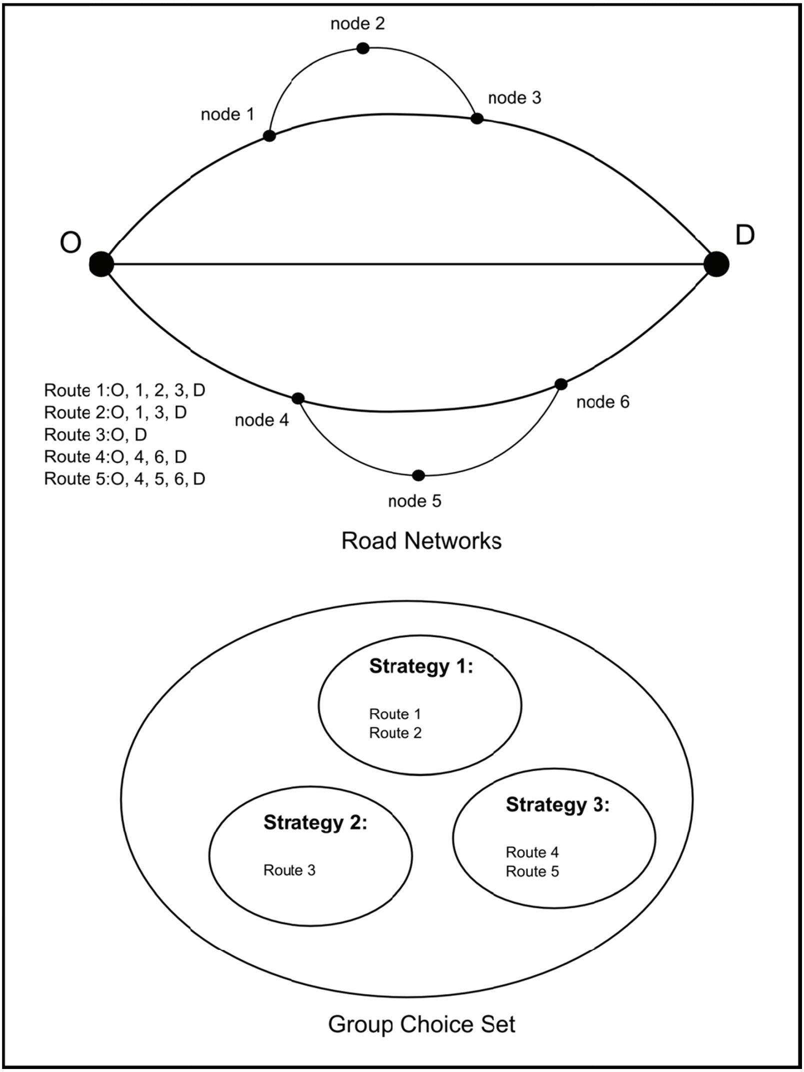

Figure 20: Example of building group choice set of strategies in a network of 5 routes

The local optimal threshold is the key to cluster the routes in a master set into strategies. According to the analysis in section 3.1.2, for a given OD-pair the number of reasonable routes decreases, as the parameter β increases. No matter how large β is, the shortest path is always reasonable, because any segment of a shortest path is an optimal segment(DIJKSTRA, 1959). This means for a given size n a β^* can be found so that $\boldsymbol{P}^{\beta^*}$ contains no more than n routes.

In order to find the β with a given n, a binary search procedure as depicted in Figure 21 can be adopted. The initial interval of β is $[\beta_0, \beta_1]$where β_0 is the threshold used in the enumeration, β_1 is the cost of the shortest path between the OD-pair. Then the middle point $\beta_2 = (\beta_0 + \beta_1)/2$ is inspected. If $\boldsymbol{P}^{\beta_2}$ has more than n routes then the target is in the right half interval $[\beta_2, \beta_1]$, or else the target is in the left half $[\beta_0, \beta_2]$. The search step is repeated on the new interval until β^* is found.

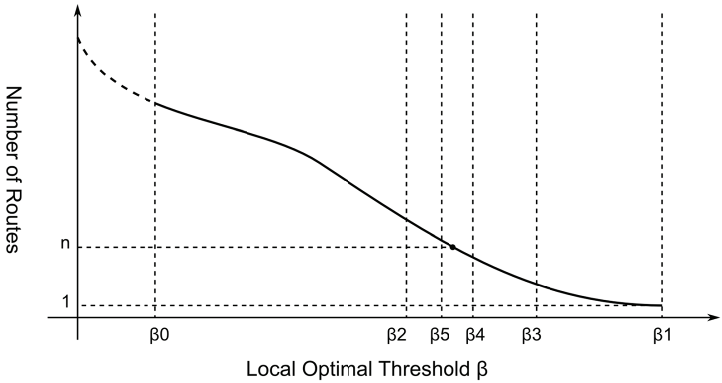

Figure 21: Binary search for the local optimal threshold that confines the number of routes in the master set to a predefined value n.

This $\boldsymbol{P}^{\beta^*}$, which contains no more than n routes, is called the **representative set**. These routes have the largest distinction between each other, measured by mesh size. Each route in $\boldsymbol{P}^{\beta^*}$represents a strategy, called a representative route. The other routes in the master set are compared with each representative route to decide which strategies these routes belong to. This procedure is depicted in Figure 22.

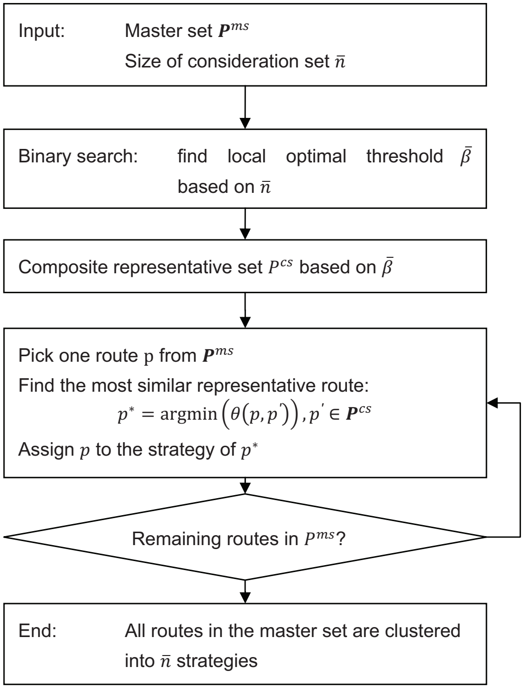

Figure 22: The procedure of clustering routes in the master set into strategies

3.5 Calibration of Parameters

One important factor that directly influences the route generating results is left out in previous sections: How to determine the proper parameters for the constraints in enumeration and assembling procedures. This question will be addressed in this section.

According to the constraints introduced in section 3.3.2, the following parameters need to be determined. There are 3 parameters for generating the master set:

- Limit of global detour factor,
- Limit of local detour factor,
- Limit of local direct threshold,

The basic idea of calibrating parameters is that the actual route sets can be reproduced by choosing proper parameters. The actual route sets considered by travellers is unknown to the researchers and they need to be observed. One method of observing route choice sets is to record the routes used by a group of travellers in a given time period. This can be done with technologies such as GPS, Mobile phone tracking, Bluetooth or ANPR. Another method is to design an experiment, presenting a route choice situation, real or hypothetical, to a group of people and ask them to state their possible choices. Route sets collected in these ways are called **observed route sets**.

Although observed route sets can be collected for individual travellers, it is difficult to compare them with generated individual consideration sets. The reason is that the individual consideration sets are assembled with a stochastic procedure and it is difficult to establish the correspondence between an observed individual and a generated individual. Instead, a collective observed route set, the union of all individual route sets, is compared with the generated master set.

In order to compare a generated master route set with an observed route set, an indicator called **coverage rate** is defined.

$$C(\boldsymbol{P}^{os}, \boldsymbol{P}^{ms}) = \frac{|\{p \in \boldsymbol{P}^{os}: \theta(p, p') > \bar{\theta},\ \exists p' \in \boldsymbol{P}^{ms}\}|}{|\boldsymbol{P}^{os}|}$$

Where C is the converge rate; $\bar{\theta}$ is the lowest similarity index that an observed route p can be regarded as reproduced; $|\boldsymbol{P}|$ is the size of set $\boldsymbol{P}$. The coverage rate indicates what percentage of observed routes are reproduced in the generated master set. C should be 1 for an idealistic route set generation method.

Apparently the universal set has a coverage rate 1 for any observed set. The master set contains all routes from the universal set that is not discarded by the constraints. If the constraint parameters are choosing in a way that no routes in the observed route set will be discarded, then the coverage rate of the master set is also 1. At the same time, the size of the master set should be as small as possible. Following this thought, the constraint parameters can be determined in the following ways.

Limit of Global Detour Factor

The limit of the global detour factor can be determined by

$$\overline{gd} = max(\{gd_p: p \in \boldsymbol{P}^{os}\})$$

The limit is larger than the global detour factors of all routes in the observed set, which means no route in the observed set will be discarded by this constraint.

Limit of Local Detour Factor

The limit of the local detour factor can be determined by

$$\overline{ld} = max(\{ld_p^*: p \in \boldsymbol{P}^{os}\})$$

The limit is larger than the local detour factors of all routes in the observed set, which means no route in the observed set will be discarded by this constraint.

Limit of Local Optimal Threshold

The limit of the local threshold can be determined by

$$\overline{lt} = min(\{lt_p^*: p \in \boldsymbol{P}^{os}\})$$

The limit is smaller than the local direct thresholds of all routes in the observed set, which means no route in the observed set will be discarded by this constraint.

It is worth noting that achieve a coverage rate of 1 should not be the sole consideration of the calibration. The coverage rate is an important measure for shortest path search based route generation methods, which tend to generate too few routes. However, it is not adequate for measuring the master set produced by the branch-and-bound method, which tend to produce too many routes. A more suitable way of using the observed set is to treat it as a reference. By using stringent parameters, a few observed routes will be discarded from the master set, but the size of the master will be greatly reduced. This can be regarded as a method to handle the outliers in the observed route sets.

4 Traffic Assignment with Pre-Generated Route Sets

With realistic routes for all OD-pairs generated, this chapter describes the procedure of distributing the travel demands among these routes and predict the traffic condition in the transport network.

The core issue of the traffic assignment is deciding how to tackle the influence of the congestion effect. In a transport network increased traffic loads will cause congestions, which in turn lead to increased travel costs. The congestion effect may create inconsistencies in a traffic assignment model. In an assignment model the distribution of traffic loads among different routes is the cumulative result of route choices made by each traveller. The travellers make the pre-trip route choices based on a set of **expected route costs**, however with all their choices the traffic load pattern changes and the travel costs of links change accordingly through the congestion effect. This means the route costs based on which travellers make route choices is inconsistent with the route costs which travellers actually experienced.

There are two methods to solve the inconsistency problem. The first method is to assume the congestion effect is insignificant and can be disregarded altogether. The second method is to find an equilibrium travel demand distribution in which the decision route costs are similar to the experienced route costs. These two methods are called uncongested assignment and congested assignment respectively. In this research both methods are addressed.

4.1 Overview

4.1.1 Procedure of Equilibrium Assignment

The traffic assignment procedure in this research is depicted in Figure 23. It receives route choice sets for all OD-pairs as input from the route set generation procedure, and the route choice sets stay unchanged in the subsequent steps. Then the route choice model distributes travel demands among all these routes based on the initial route costs. In the initial state, it can be assumed that the network is empty with no traffic flows. The route costs can be calculated based on free flow link travel times. The link cost can also be calculated based on the historical link travel times, if they are available.

After the initial route choice the assignment procedure repeats the following steps:

- The result of the route choice model is a set of route flows, each of which represents vehicles that move along the given route. The network loading model simulates the propagation of vehicles in the network and establishes the incidence relationship between route flows and link flows at different time steps. Based on this relationship the link flows at different time steps can be calculated, based on which the experienced route costs can be derived.

- If the assignment model does not consider the congestion effect, which means the experienced route costs do not change with the traffic flows, then the procedure stops here and the output of the network loading model is taken as the assignment result. If the assignment model considers the congestion effect, the distribution of traffic flows between routes need to be adjusted based on the experienced route costs.

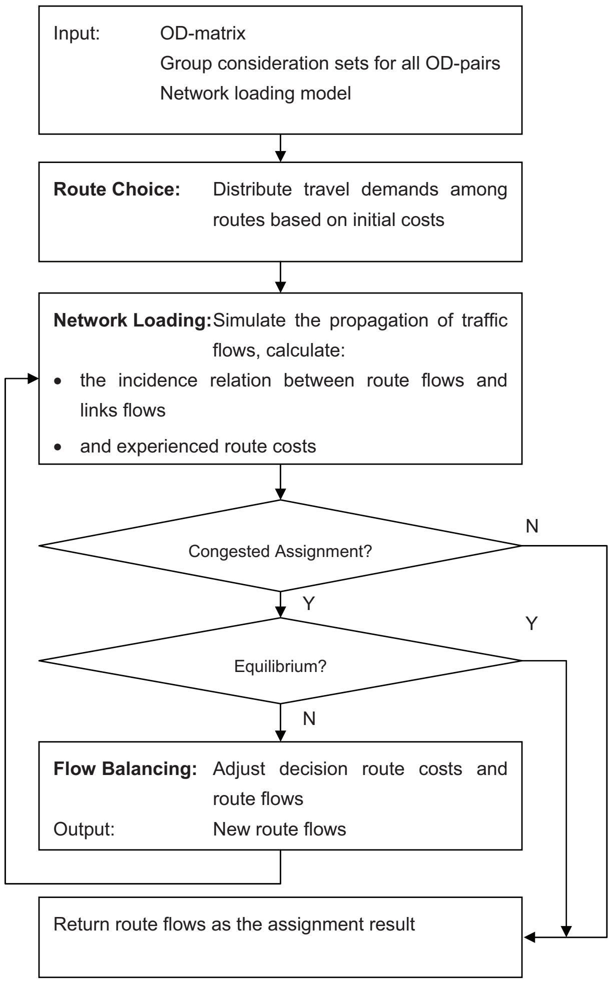

Figure 23: Overall procedure of the traffic assignment with pre-generated route choice sets

With the experienced route costs the subsequent process first checks whether the assignment result is in an equilibrium state. If yes, the process ends and the route flows are returned as the result. If not, the flow balancing model adjusts the distribution of route flows based on experienced route costs. The assignment process goes back to network loading with adjusted route flows. are feed into the network loading model again to calculate the new experienced routes.

4.1.2 Model Assumptions

Until now both the concept of equilibrium and the assignment procedure to achieve it were described in a general way, so that these descriptions could be applied to different modelling contexts.

The modelling of transport is a complex task, it is impossible to build a perfect model to capture every aspect of the transport. Different simplification assumptions are made in different contexts. In this research two aspects are of special interest: the temporal variance of travel demand and the perception of route costs.

In the aspect of temporal variance assumption the assignment models can be classified as static models and dynamic models. The static models assume that the traffic volumes in the network are constant over time while the dynamic models assume that the traffic volumes can change over time. Under the aspect of route cost perception, the assignment models can be classified as deterministic models, which assume that travellers have perfect information of route costs, and stochastic models, which assume that there are random errors in the perception of route costs.

For the stochastic assumption, the origin definition of equilibrium is adapted as “under equilibrium conditions traffic arranges itself in congested network in such a way that all routes between any OD-pair have equal and minimum **perceived costs** while all unused routes have greater or equal costs.”

For the dynamic assumption, the origin definition of equilibrium is adapted as “under equilibrium conditions traffic arranges itself in congested network in such a way that, **for each departure time period**, all **connections** between any OD-pair have equal and minimum costs while all unused **connections** have greater or equal costs.” Here a connection refers to a route with a specific departure time period.

4.2 Route Choice Model

The route choice model assigns travel demand of each OD-pair onto available routes according to the route costs. The basic behaviour assumption of route choice is the minimization of costs. According to the deterministic assumption, travellers have perfect information of route costs. Therefore, for travel demand of one OD-pair, the route choice problem is reduced to a trivial task of finding the route with the lowest cost and all demand is assigned to that route. This section will mainly discuss the assignment model for the stochastic case.

4.2.1 Route Choice in the Static Assignment Model

As discussed in section 2.3, route choice models are special cases of discrete choice models based on the random utility theory. How to handle the similarity between routes is the core problem. One feature of the route choice set generation method in this research is that it identifies the similarity among routes explicitly, and groups similar routes into clusters. As a result, the generated route choice sets for groups have a hierarchical structure. A group route choice set contains several strategies and each strategy contains many similar routes. This cluster structure is depicted in Figure 24.

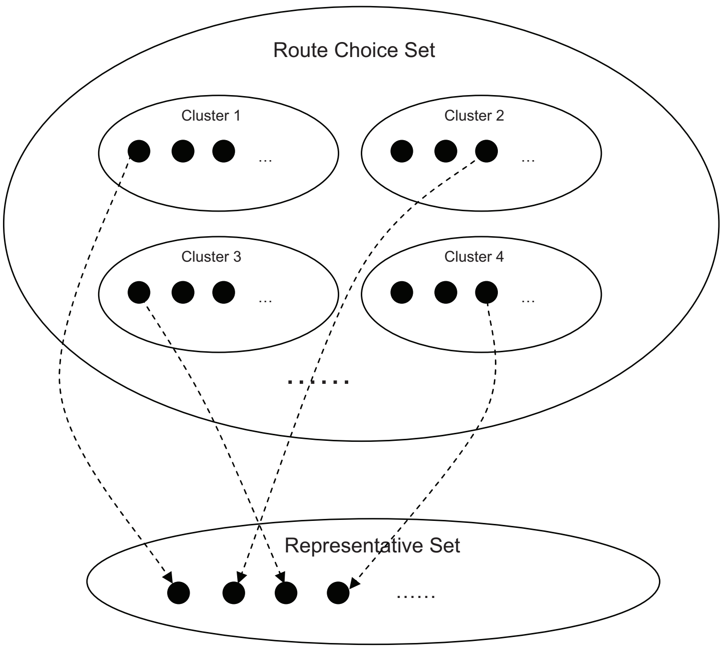

Figure 24: Structure of clustered route choice set

For each cluster there is one special route which represents the average of all routes in this cluster, hence it is called the **representative route.** All representative routes from each cluster in a route choice set make up the **representative route set** for this route choice set.

Given a Clustered route choice set R, each cluster is represented as R_i, where i is the index of this cluster. The set of all clusters for R makes a mutual exclusive partition of R, which means

$$R_i \cap R_j = \emptyset, \forall i \neq j$$

$$R \equiv \bigcup_{i \in [1 \dots N]} R_i$$

For cluster R_i the representative route is denoted by r_i^{rep}, the representative set of R is denoted by $R^{rep} \equiv \{r_1^{rep} \dots r_N^{rep}\}$.

To take advantage of this cluster structure in route choice sets, a two-stage route choice model is proposed. The probability of choosing a route r_{ij}, the jth route in R_i is the product of the probability of choosing ith cluster and probability of choosing jth route within this cluster; that is,

$$P(r_{ij}|R) = P(j|R_i)P(i|R)$$

In this research $P(i|R)$, the probability of choosing cluster i, is assumed to be equivalent to the probability of choosing its representative route r_i^{rep} in the representative set R^{rep}. For this task all existing route choice models can be used. In considering the balance between computational efficiency and modelling fidelity, the C-Logit mode is preferred for its clearer handling of overlapping routes. Refer to section 2.3.3 for a more detailed description.

In the second stage, $P(j|R_i)$, the probability of choosing the jth routes in R_i is calculated. For this stage a simple Logit model without considering the overlapping will be sufficient, since the routes in the same cluster should have similar costs enforced by the route choice set generation procedure.

4.2.2 Route Choice in the Dynamic Assignment Model

In this research, the dynamic network is handled by dividing the time horizon into **time steps** of equal length. To simplify the problem, the traffic state is assumed to be constant within each time step. Based on this assumption, travellers' trip choice involves not only the choice of route, but also the choice of departure time steps. To simplify the discussion, the concept of **connection** is defined as a route plus a departure time step. Within the dynamic models the route choice problem is extended into a connection choice problem.

Travel demand in the dynamic assignment context is represented by a series of dynamic OD matrices. Each matrix in this series corresponds to one time step, representing the travel demand entering the network in this time step. There are two interpretations for these matrices. In the **rigid demand** interpretation a matrix represents the actual travel demand entering the network in the corresponding time step and no shifting among time steps is allowed. In the **elastic demand** interpretation a matrix represents the travel demand to enter the network in the corresponding time step and shifting among time steps is allowed.

In the rigid demand assignment problem the departure time is pre-defined for all travellers, therefore the route choice model is equivalent to that of the static case. For the elastic demand assignment problem, a joint departure and route choice model developed and calibrated in (MANDIR, 2012) will be adapted for this research.

4.3 Network Loading Model

The purpose of a network loading model is to calculate the congested route costs incurred by given route volumes. The input route volumes are represented by a vector $\mathbf{h}$. Each element h_r of $\mathbf{h}$ represents the volume of traffic on route (or connection, in dynamic case) r. The network loading model produces a vector $\mathbf{c}$ as output. Each item c_{r} of this vector represents the travel time of route r. In this sense the network loading model is conceptually equivalent to a multi-dimensional function $L\colon \mathbf{h} \to \mathbf{c}$

4.3.1 Static Model

In a graph the relationship between a set of edges $\boldsymbol{E} = \{e_1, \ldots e_n\}$ and a set of paths $\boldsymbol{P} = \{p_1, \ldots, p_m\}$can be represented by an edge-path **incidence matrix**:

$$\boldsymbol{\Delta} = \begin{bmatrix} \delta_{11} & \cdots & \delta_{1m} \\ \vdots & \ddots & \vdots \\ \delta_{1n} & \cdots & \delta_{nm} \end{bmatrix}.$$

The value of each element is determined as

$$\delta_{ij} = \begin{cases} 0 & e_i \notin p_j \\ 1 & e_i \in p_j \end{cases},$$

which means $\delta_{ij} = 1$ if there is a incidence between edge e_i and path p_j; $\delta_{ij} = 0$ otherwise.

With the edge costs determined, the path costs can be calculated. For a given path$p$ the path cost g_p contains two parts:

$$g_p = g_p^A + g_p^{NA}$$

whereg_p^A is called the **additive path cost**, which is the accumulation of edge costs of all edges in the path. g_p^{NA} is the non-additive path cost, which is specific for the path. For

$$g_{p_j}^A = \sum_{e \in p_j} c_e = \sum_{1 \le i \le n} \delta_{ij} \cdot c_{e_i}$$

or in matrix form

$$\boldsymbol{g} = \boldsymbol{\Delta}^{\boldsymbol{T}} \cdot \boldsymbol{c}$$

where $\boldsymbol{g}$ is the vector of path cost, $\boldsymbol{c}$ is the vector of link cost.

A path flow h_p can be associated to each path p, representing the traffic volumes using this path in a given time unit. Likewise an edge flow f_e is associated with each edge e, representing traffic volumes using this edge. Under the stationary assumption, the path flow is constant on all edges alone the path. Traffic flows follow the conservation law that maintains two constraints. First, the flow of an edge is equal to the accumulation of path flows, which pass through this edge.

$$f_{e_i} = \sum_{p \ni e_i} h_p = \sum_{1 \le j \le m} \delta_{ij} \cdot h_{p_j}$$

Or in matrix form

$$\boldsymbol{f} = \boldsymbol{\Delta} \cdot \boldsymbol{h}$$

where $\boldsymbol{f}$ is the vector of edge flow$\boldsymbol{h}$ is the vector of path flow. Second, the accumulation of path flows is equal to the travel demand. For each OD-pair od

$$d_{od} = \sum_{p \in P_{od}} h_p$$

Where d_{od} is the demand between OD-pair od, $\boldsymbol{P}_{od}$ is the set of paths.

According to the static assumption, the traffic state is independent of the time, which means the travel time for each link is independent of the time when traffic entering this link. As common practice, the link travel time is assumed to be a monotonously increasing function of the traffic volume on this link. This function is called volume-delay function. Several forms of volume-delay functions are developed over the years, however the form developed by Bureau of Public Roads (BPR) is by far the most common one,

$$t = t(f) = t_0 \cdot \left(1 + \alpha \cdot \left(\frac{f}{cap}\right)^{\beta}\right)$$

where c_0 is the free flow travel time, f is the volume, cap is the capacity, α and β are parameters.

The traffic volume on a link is the aggregation of volumes of all routes which pass this link. The travel time of a route is the aggregation of travel times of all links along this route. The relationship between traffic volumes and route travel times can be represented as:

$$\boldsymbol{f} = \Delta \cdot \boldsymbol{h}$$

$$\boldsymbol{c} = \Delta^T \cdot \boldsymbol{t}$$

where $\mathbf{f}$ is the vector of link volume, $\mathbf{t}$ is the vector of route travel time. Δis the link-route incidence matrix. The network loading function can be written as:

$$L(\boldsymbol{h}) = \Delta^T \cdot \boldsymbol{t}(\Delta \cdot \boldsymbol{h}) = \boldsymbol{c}$$

4.3.2 Dynamic Model

In a dynamic assignment model the travel time of a link varies according to not only the time of entering this link, but also the distribution of traffic density on this link. A traffic flow model is required to simulate the propagation of traffic and calculate the travel time.

The main purpose of the traffic assignment model developed in this research is to support strategic planning. To this aim several requirements need to be fulfilled.

First and foremost, the dynamic network loading model should be consistent with the static network loading model. This means that the two models should produce identical results if they are applied to the same scenario where travel demand is constant over time. In theory if a dynamic model and a static model describe the same network, they are just different representations of the same reality. The only difference is that the dynamic model contains more detail than its static counterpart. However, the results from both models are generally not identical. Special requirements must be fulfilled in order to obtain similar results from the two model types.

The second requirement is to maintain the first-in-first-out (FIFO) property, which means for any two vehicles travelling through the same link, the vehicle entering the link earlier should exit the link earlier as overtaking is not allowed. This property is reasonable for most cases and it greatly simplifies the analysis

The third requirement is high computational performance. Since the network loading is invoked repetitively in the process of finding the equilibrium solution, the performance of network loading is decisive to the performance of the whole assignment procedure. One important way to increase the computational performance is to allow large time steps.

The first requirement is especially important for practical reason, since it allows making meaningful comparison between the static assignment result, which is well-understood and the robust and dynamic assignment result, furnishing more details.

Link Travel Time Function

The classic macroscopic traffic flow model is the so called LWR model developed by LIGHTHILL and WHITHAM (1955), and RICHARDS(1956). The LWR model assumes the traffic flows are similar with water flows or air flows, which can be modelled as stream of infinitely divisible particles. The state of traffic flows on a link is governed by the following conservation equation

$$\frac{\partial k(s,t)}{\partial t} + \frac{\partial q(s,t)}{\partial s} = 0$$

$$q(s,t) = k(s,t)v(s,t)$$

The LWR model assumes that the relationship between speed v and density k can be depicted with a simple equitation developed by Greenshied(1954)

$$v(s,t) = v_0(1 - k(s,t)/k_{jam})$$

According to Greenshied’s relation, the traffic flow q is a parabola function of k. $q = 0$ when k is 0 or k_{jam}, and q attains the maximal value at $k = k_{jam}/2$.

Although queuing and shockwave behaviour of traffic flow can be devised from the Greenshied’s relation, however it is not consistent with the volume-delay function when the flow rate exceeds the capacity of links. A new vk relationship needs to be derived from the volume delay function.

A volume delay function can be written in a general form as

$$t = t_0 \cdot \theta(q)$$

where θ is a monotonic increasing function of q and $\theta(0) = 1$. Then for a link of length l, the travel speed is determined by

$$v = \frac{l}{t} = \frac{l}{t_0 \cdot \theta(q)} = \frac{v_0}{\theta(q)}$$

$$q = kv = k\frac{v_0}{\theta(q)}$$

$$k = \frac{q\theta(q)}{v_0}$$

where $v_0 = \frac{l}{t_0}$ is the free flow speed. Because $\theta(q)$ is a positive monotonic increasing function of q, k is also a positive monotonic increasing function of q.

The Greenshied's relationship and this alternative speed-density relationship are compared in Figure 25. The main difference is that the density k can increase indefinitely in the alternative relationship.

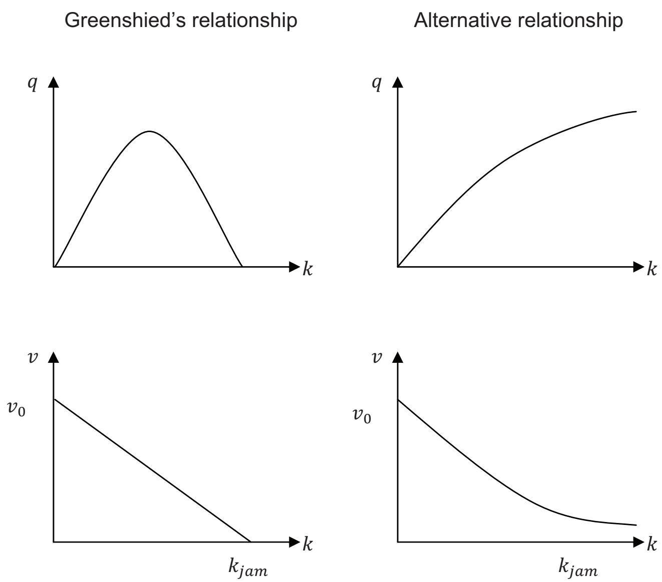

Figure 25: Fundamental diagram derived from Volume-Delay Function

A simple example is given in Figure 26 to illustrate the link travel time function. For a given link, the traffic enters at the tail end of the link and travel towards the head end. The flow rate of entering traffic flow is constant in time interval $[t_0, t_1)$and $[t_1, t_2)$, and they are q_1 and q_2 respectively. According to the steady state assumption, the link travel time will converge to a constant value, if the traffic flow stays constant for a relative long time.

In this example the traffic flow stays constant until t_1, so the link travel time is also a constant value of $tt_1 = tt_{free} \cdot \theta(q_1)$. At time t_1, the traffic flow starts to increase, and the travel time will also increase. After a period of disturbance until t_x,, the traffic will converge to a new steady state $tt_2 = tt_{free} \cdot \theta(q_2)$. Here the travel time can be further assumed to change linearly in the disturbance time interval, the remaining key question is to find the length of this time period $\Delta t = t_x - t_1$.

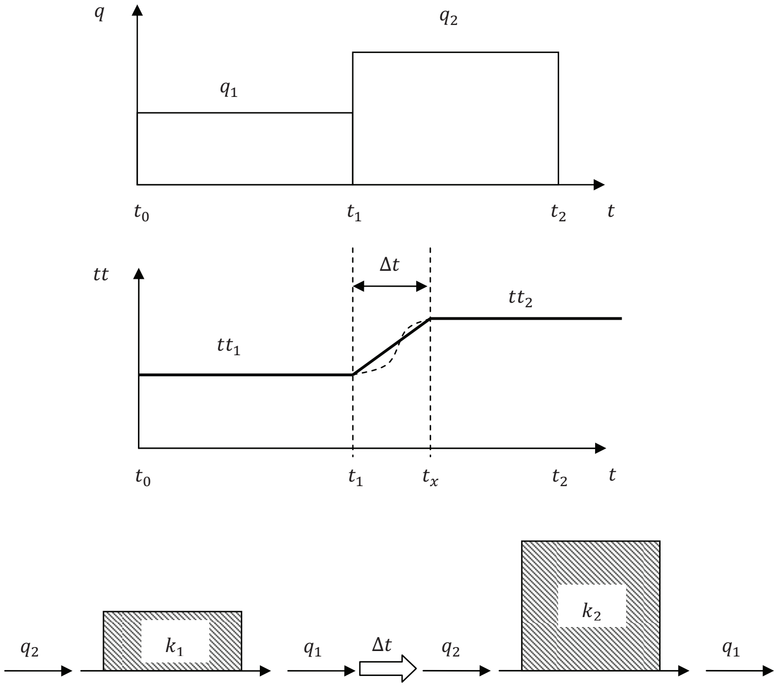

Figure 26: Link travel time function

Furthermore the assumption is made that the flow rate of entering and leaving the link change instantly. From the former state to the latter state, the number of vehicles on the link changes from $k_1 l$ to $k_2 l$, and the traffic flow rates entering and leaving the link are q_1 and q_2 respectively. On the basis of this traffic flow conservation, the time needed for the state shifting is

$$\Delta t = \left| \frac{(k_2 - k_1)l}{q_2 - q_1} \right|$$

The simulation is event-based. Each event describes a change of the flow rate of a stream at the beginning of a link. An event contains the following information:

- The corresponding route of the stream
- The corresponding link
- The new flow rate
- Time stamp

All events are organized in a priority queue using the time stamps as their priorities. At the beginning of a simulation, the route flow changes at different origins are known from the dynamic OD matrices, and these events are pushed into the queue.

During the simulation process, events from the queue head, i.e. events with smallest time stamp, are processed. There can be more than one events with the same time stamp, e.g. change of the route flow at the same time step. When processing a flow change event, the time of flow change on the downstream link can be calculated according to aforementioned method and new flow change events are generated and inserted into the queue.

4.4 Flow Balance Model

The purpose of the flow balance model is to redistribute traffic volumes among routes for the same OD-pair so that the equilibrium condition of route volumes can be achieved. Since the deterministic and stochastic assumptions lead to different interpretations for the equilibrium condition, the flow balance models are also different.

In this research, the vector of route flow shares $\boldsymbol{\alpha}$ is taken as the optimized variable instead of the vector of link flows $\boldsymbol{f}$, as in a conventional assignment model. An element α_i^{od} of the vector $\boldsymbol{\alpha}$ represents the ratio between traffic volumes on ith route of the OD-pair od and the travel demand of the OD-pair od:

$$\alpha_i^{od} = \frac{h_i^{od}}{d_{od}}$$

Based on its definition, a feasible vector $\boldsymbol{\alpha}$ complies with the following constraint. For a given OD-pair, the aggregation of the volume shares of all routes is 1:

$$\sum\nolimits_{i \in R_{od}} \alpha_i^{od} = 1 \ \forall od$$

Based on the definition of $\boldsymbol{\alpha}$, with a given $\boldsymbol{\alpha}$ the element of route flow vector $\boldsymbol{h}$ can be determined as

$$h_i^{od} = d_{od} \cdot \alpha_i^{od}$$

or in matrix form

$$\boldsymbol{h} = \boldsymbol{\Lambda} \cdot \boldsymbol{\alpha}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix with d_{od} as elements. The link flow vector f can be further determined as

$$\boldsymbol{f} = \Delta \cdot \boldsymbol{h}$$

where Δ is the incident matrix. Combining the two formulas we get

$$\boldsymbol{f} = \Delta \cdot \boldsymbol{\Lambda} \cdot \boldsymbol{\alpha}$$

It can be shown that if $\boldsymbol{\alpha}$ is feasible then $\boldsymbol{h}$ is a feasible route flow vector and $\boldsymbol{f}$ is a feasible link flow vector. This means solving for $\boldsymbol{\alpha}$ or solving for $\boldsymbol{f}$ are equivalent.

The stochastic assignment problem can be modelled as a fix point problem. For a given vector of route costs $\boldsymbol{c}$the ration of route volume can be calculated with the route choice model described in section 0, and the result is the ratio vector $\boldsymbol{\alpha}$. This calculation is actually an uncongested assignment and it can be written as

$$\boldsymbol{\alpha}^k = \boldsymbol{A}(\boldsymbol{c}^{k-1})$$

where $\boldsymbol{A}$ is the assignment function .On the other hand with a given ratio vector $\boldsymbol{\alpha}$ a new route costs can be recalculated by a network load model and it can be written as

$$\boldsymbol{c}^k = \boldsymbol{L}(\boldsymbol{\alpha}^k)$$

where $\boldsymbol{L}$ is the network loading function. Combining the two equitation we get

$$\boldsymbol{\alpha}^k = \boldsymbol{A}(\boldsymbol{L}(\boldsymbol{\alpha}^{k-1}))$$

If the limit

$$\boldsymbol{\alpha}^* = \lim_{k\to\infty} \boldsymbol{\alpha}^{\boldsymbol{k}}$$

exists, then $\boldsymbol{\alpha}^*$ is the equilibrium solution.

To solve this fix point problem the Method of Successive Average (MSA) is employed. The MSA can be described by the pseudo code in Figure 27.

1 `initialize` α^0

2 `for` i `in range(1,k):`

3 $\qquad \alpha^{\mathrm{new}} = \mathrm{A}(\mathrm{L}(\alpha^{\mathrm{i}-1}))$

4 $\qquad \alpha^i = \alpha^{i-1} + \frac{1}{i}(\alpha_{new} - \alpha^{i-1})$

Figure 27: Pseudo code for MSA

At the beginning of the assignment process, the ratio vector $\boldsymbol{\alpha}$ is initialized (line 1). One possible initial value is that the ratio distributes evenly on al routes. Then the process enters loop of k times, where k is a predefined parameter. In each iteration, the new ratio vector is calculated based on current ratio vector (line 3), then the old ratio vector is updated by having a weighted average with the new ratio vector. The weight of new ratio vector decreases as the number of iterations increases.

5 Case Study

5.1 Issues in Implementation

In this section several practical issues in implementing the traffic assignment framework is discussed.

5.1.1 Network Pre-processing

The route generation and assignment method developed in this research has specific requirements for the input road network. This section addresses three important issues of converting network data from various sources into the usable form.

Intersection Simplification

Networks for car navigation systems (e.g. NAVTEQ, Open Street Map) provide a common data source for the road network. The problem with these networks is that they have a detailed representation of intersections. Each intersection, especially a complex highway intersection, is represented by multiple nodes and links to depict the structure of interconnected ramps. While this structure is critical for car navigation it is of less interest for planning purposes as generally only the network topology is required. In this case, each intersection can be represented with one node.

Reducing complex intersections into single nodes will reduce the number of nodes in the network, which has two advantages. First, it will reduce the scale of the problem. The performance of route generation and traffic assignment has a positive correlation with the network size measured by number of nodes. With network simplification, larger problems can be handled with the computational resources at hand. Second, it will help the researchers to understand and interpret the results at a higher degree of abstraction. With a simplified network the generated routes concern the strategic choice of travellers at intersections without irrelevant details of how to manoeuvre the ramps.

Two stages are necessary to simplify a network. The first stage identifies nodes and links that belong to complex intersections and their membership relation. The second stage breaks down all nodes and links for a complex intersection into one node, and recalculates its properties.

For the navigation network data the membership relation between intersections and nodes and links is generally manually identified and provided. The network simplification algorithm developed in (MANDIR, 2012) is adopted for this research.

Road Hierarchy Level Identification

As discussed in Section 3.1.2, the road hierarchy levels plays an important role in the process of identifying routes in a network. Navigation network data generally provide the road type of each link. With a given mapping between road types and road hierarchy levels, the road hierarchy level for each link can be determined. If such information is not available in the network data, it needs to be identified manually.

Network Expansion

The route generation and traffic assignment models in this research are designed based on the assumption that a transport network is represented by a simple and directed graph. It is possible to convert a road network in a graph by direct mapping nodes to vertices and links to edges. However, the topological structure of turns at each node is lost through this process, e.g. the turn prohibitions are not honoured, which leads to unrealistic routes and unrealistic traffic flow distribution.

In order to convert a network to a graph while preserving the topological structure of turns, a special procedure as illustrated in Figure 28 is needed. The vertices in the graph have the following sources:

- Each node is converted into n vertices, where n is the total number of the incoming and outgoing links of this node. Each vertex represents an entrance or exit of this node for a specific link.
- Each zone centroid node is converted into 2 vertices. One represents the origin of trips and the other represents the destination of trips.

At the same time, the edges in the graph have the following sources:

- Each link is converted into an edge, connecting the exit of upstream nodes to the entrance of the downstream nodes.
- Each turn is converted into an edge, connecting the entrance of the incoming link and the exit of the outgoing link.
- A pair of entering and leaving connectors between a given zone and node are converted into n edges, where n is the number of the incoming and outgoing links of this node. The entering connector is converted into edges connecting origin vertex to entry vertices, while the exiting connector is converted into edges connecting exit vertices to the destination vertex.

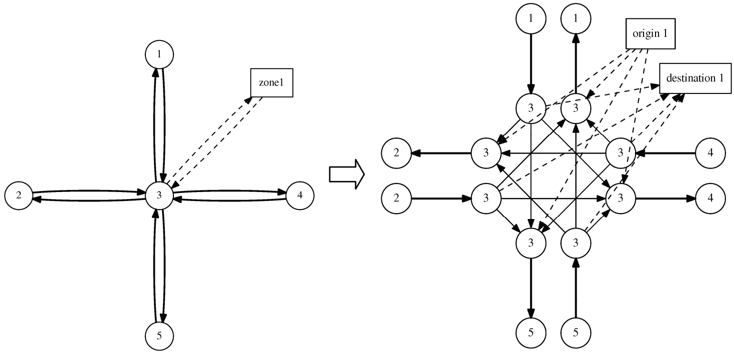

Figure 28: Expanding a node(intersection) into vertices while preserving the topological structure of turns.

It is worth noting that network expansion and intersection simplification undermine the effect of each other. The intersection simplification combines multiple nodes, while the network expansion converts one node into multiple vertices. Despite this conflict, they are complementary as they have different purposes. The purpose of intersection simplification is to reduce the detail of a complex intersection, while the purpose of network expansion is to find a topologically equivalent graph for a network. The former is an optimization for performance, while the latter is essential for the correctness.

5.1.2 Effective Constraint Tests

The constraints are operable rules determining reasonable routes. They are formalized based on assumptions concerning travellers' behaviour when choosing a route. Besides the basic requirement that they reflect certain aspects of the assumptions, the constraints must fulfil additional requirements to make them operable.

- The constraints should be unambiguous. A constraint is formalized so that a route can be tested and the result is clearly either fail or pass. Additionally it requires that if a route fails, the test of its child routes should also fail.
- The constraints should be efficient. The computational cost to test a path should be moderate as the tests will be repeated in large numbers in the route generation procedure thus determining the overall efficiency of the method.

In order to fullfill these two requirements, an invariant need to be maintained for all constraints. A constraint should be defined in a way that if a node sequence fails the constraint test then all its child sequences should also fail the test. All constraints defined in this research comply with this invariant.

5.2 Validation

In this section, the route generation model and the traffic assignment model are validated using synthesized test networks. These networks are built up to highlight certain shortcomings of the conventional methods. The route generation model is validated in 5.2.1 and the traffic assignment model is validated in section 5.2.2.

5.2.1 Route Generation Model

To illustrate the working of the route generation model the test network shown in Figure 29 is employed.

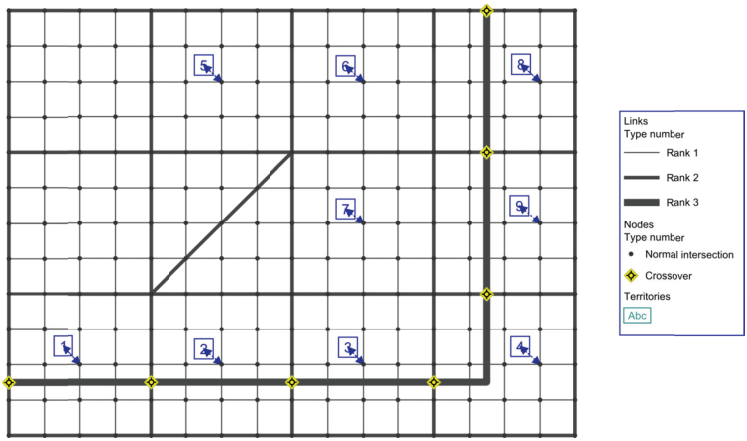

Figure 29: Test network for validating route generation model

This network mimics an urban road system with a grid pattern and the size of each grid is 0.5km by 0.5km. The links are categorized into 3 ranks:

- urban express road,
- trunk road and
- secondary road.

Links of higher rank have higher travel speed and capacity. The essential properties for each rank is summarized in the following table

Rank	Free Flow Speed	Capacity	Description
1	$25km/h$	$1000pcu/h \times 2$	Secondary Road
2	$45km/h$	$1800pcu/h \times 3$	Trunk Road
3	$80km/h$	$1800pcu/h \times 4$	Express Road

Table 2: Road ranks and their essential properties. Capacity is given in hourly passenger car units (PCU) per lane.

Costs Calculation

The costs of network elements such as links and turns have great influence on the route generating process. In this network, the costs are calculated in two different ways.

In the first case, the link costs are the free flow travel times, which can be calculated as

$$c_l = tt_l = \frac{L_l}{v_l}$$

where c_l is the cost of link l, tt_l is the free flow travel time, L_l is the length and v_l is the free flow speed. For example, the cost of a link on the secondary road is

$$\frac{0.5km}{25km/h} \cdot \frac{3600s}{h} = 72s$$

The travel time of turns are assumed to be constant for a given turn type.

$$c_t = tt_t = \begin{cases} 5s & t\ is\ a\ straitforward\ or\ right\ turn \\ 10s & t\ is\ a\ left\ or\ U\ turn \end{cases}$$

Left turns and U-turns are assumed to have longer travel times than that of right-turns and straightforward-turns.

In the second case, the link costs and turn costs are also based on the travel time, but they are adjusted according to the road network hierarchies. The link costs are calculated as

$$c_l = tt_l \cdot f_l^{ha}$$

where f_l^{ha} is the hierarchy-adjust factor for link l. In this test we use

$$f_l^{ha} = \begin{cases} 1.0 & rank\ of\ l\ is\ 3 \\ 1.2 & rank\ of\ l\ is\ 2 \\ 1.5 & rank\ of\ l\ is\ 1 \end{cases}$$

Links of lower ranks are penalized based on the assumption that travellers prefer high rank roads which are more comfortable to drive on.

The turn costs are adjusted in a similar way

$$c_t = tt_t \cdot f_t^{ha},$$

where

$$f_t^{ha} = \begin{cases} 1.0 & turning\ from\ low\ rank\ road\ to\ high\ rank\ road \\ 2.0 & turning\ from\ high\ rank\ road\ to\ low\ rank\ road \end{cases}$$

The **hierarchy-adjusted costs** need to be normalized against the travel times. The normalize-factor is calculated as

$$f^{norm} = \frac{\sum_{t \in T} tt_t + \sum_{l \in L} tt_l}{\sum_{t \in T} c_t + \sum_{l \in L} c_l}$$

where T is the set of turns and L is the set of links. With this factor, the **normalized hierarchy-adjusted** cost is

$$\bar{c} = c \cdot f^{norm}$$

The purpose of the normalization is to maintain

$$\sum_{t \in T} c_t + \sum_{l \in L} c_l = \sum_{t \in T} tt_t + \sum_{l \in L} tt_l$$

that the total cost in the network keeps constant after adjusting. This property is important for the local optimal constraint. Without normalization, the local optimal parameter is not transferable from one scenario to another.

Scenarios

The effects of different constraints are illustrated with different scenarios. In each scenario, the origin zone (zone 1) and the destination zone (zone 8) are situated far away from each other and near the express road. From a traveller's intuition, it is easy to identify the optimal route which mainly uses the express road. The constraint parameters used in each scenario and the number of reasonable routes are summarized in Table 3.

No.	Cost	Global Detour Factor (GD)	Local Detour Factor (LD)	Local optimal Threshold(s) (LT)	Resulting no. of reasonable routes
1	HA	1.2	1.2	200	1
2	HA	1.2	1.5	200	8
3	TT	1.2	1.5	200	4
4	HA	1.5	1.5	200	21
5	HA	1.5	1.5	300	4
6	HA	1.5	1.5	400	2

Table 3: Scenarios for showcasing the effect of different constraints. HA stands for Hierarchy-Adjusted cost; TT stands for travel time.

In scenario 1 (Figure 30) all constraint parameters are rather strict. The global detour factor is 1.2 and the local detour factor is 1.2. The local optimal threshold, or the minimal length of a detour is set to 200 seconds. With these strict constraints only one route, the shortest route, is generated.

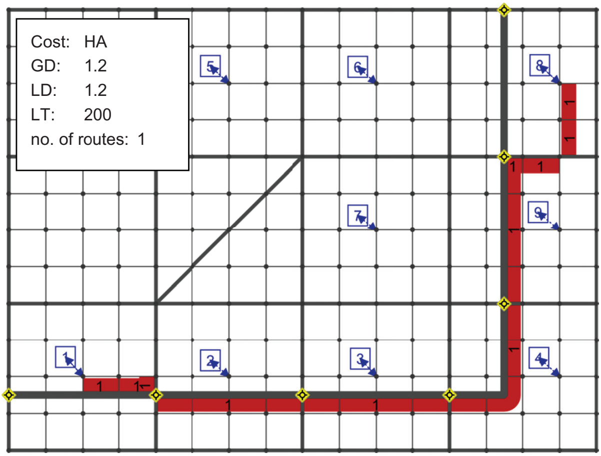

Figure 30: Route set for scenario 1

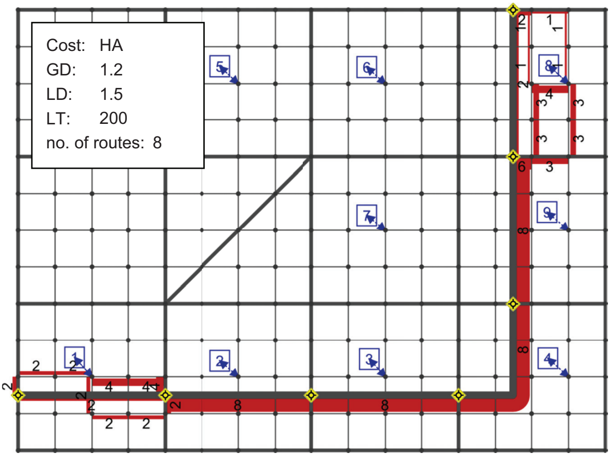

Figure 31: Route set for scenario 2

In scenario 2 (Figure 31) the local detour factor constraint is relaxed to 1.5. In this case, all generated route still converges on the express road, as it is necessary to maintain a small global detour factor. More variations of the ways to enter and leave the express road emerges as the result of the relaxed local detour factor.

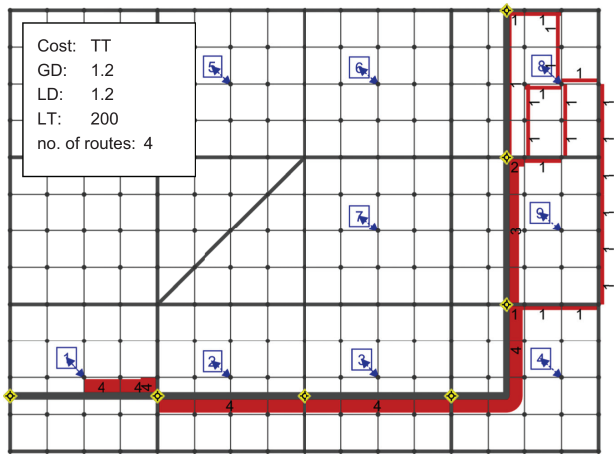

Figure 32: Route set for scenario 3

In scenario 3 (Figure 32) the link costs and turn costs are calculated based on travel time directly without hierarchical adjusting. Compared to scenario 2, although constraints with the same parameters are used, the resulting route set spreads wider from the express roads to trunk roads. This comparison shows that the hierarchical adjusting gives more weight to high-level roads.

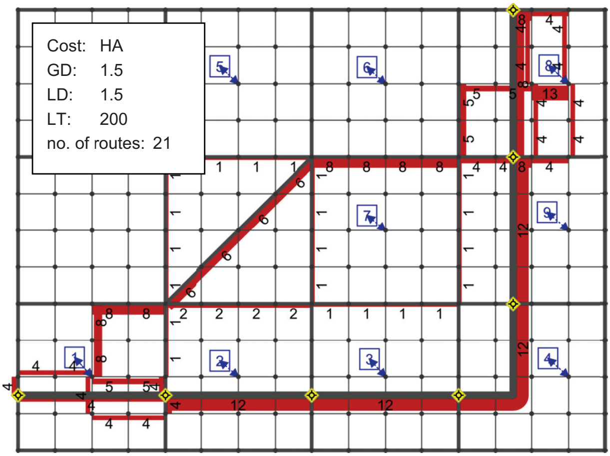

Figure 33: Route set for scenario 4

In scenario 4 (Figure 33) the limit of the global detour factor is relaxed to 1.5. As a result, the number of reasonable routes increased dramatically to 21, and spread out far away from the shortest path. Also the number of routes in the master set exceeds the limit (e.g. 2 to 6 according to (HOOGENDOORN-LANSER, 2005).

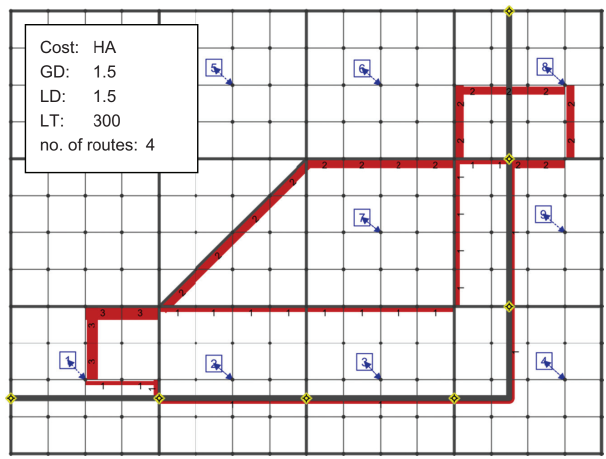

Figure 34: Route set for scenario 5

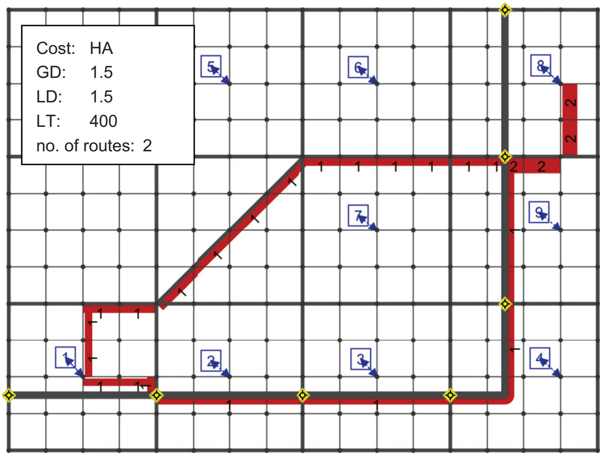

Figure 35: Route set for scenario 6

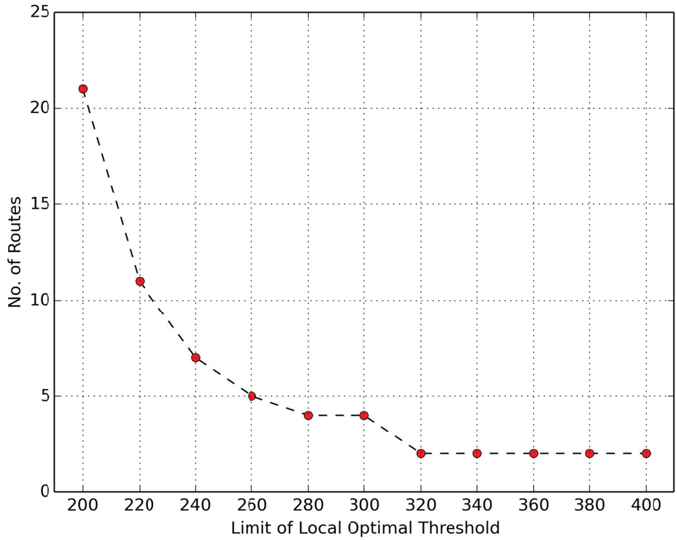

Figure 36: The relationship between size of generated route set and the local optimal threshold

To solve the problem, the local optimal threshold can be increased to reduce the size of the generated route set. In scenario 5 (Figure 34) and 6 (Figure 35), the local optimal thresholds are 300 and 400, and the sizes of generated route set are reduced to 4 and 2. Comparing scenario 4, 5 and 6, it is clear that as the local thresh hold limit increase the number of routes decrease and the remaining routes have a good geographic disparity. The tendency is shown in Figure 36. According to this figure, LD=260 is a suitable for generating a representative set that contains no more than 6 routes.

5.2.2 Traffic Assignment Model

One main problem of a DUE assignment is that it produces unrealistic results in a scenario with low traffic volumes. This problem can be illustrated with a simple symmetric network shown is in Figure 37.

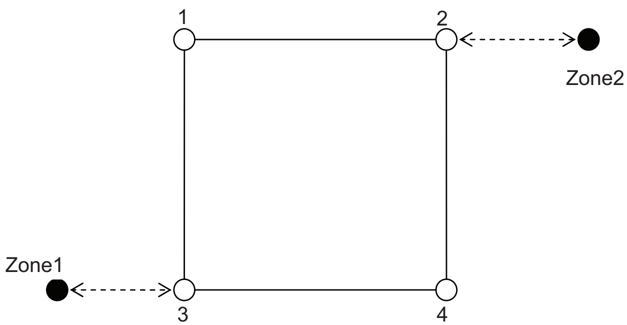

Figure 37: Simple symmetric network

This network consists of 2 zones, 4 nodes, 4 links and 2 connectors. The impedances of network elements are assumed to follow BPR function:

$$t = t_o \cdot (1 + \alpha \left(\frac{q}{c}\right)^{\beta})$$

For simplicity parameter α is assume to be 1, β is assume to be 2, and other parameters are assumed to be

- $t_0 = 3600$s, $c = 2000$veh/h for links;
- $t_0 = 20$s, $c = 2000$veh/h for left turns;
- $t_0 = 10$s, $c = 2000$veh/h for right turns.

Routes from zone 1 to zone2 and their impedance are listed in the following table:

Route No.	Nodes	Route Cost
1	3, 1, 2	$7210 \cdot (1 + \left(\frac{q}{3600}\right)^2)$
2	3, 4, 2	$7220 \cdot (1 + \left(\frac{q}{3600}\right)^2)$

Table 4: Routes of simple symmetric network

Route 1 and route 2 have almost identical impedances. The difference is that route 1 has one left turn instead of a right turn in route 2. In a DUE assignment with high travel demand the traffic volume will distributed almost evenly between these two routes. For example when the demand $d_{12} = 3000$, the equilibrium solution is depicted in the following figure.

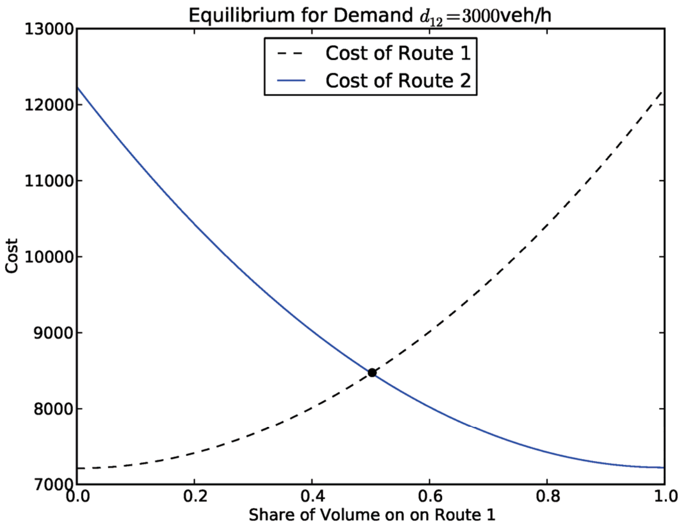

Figure 38: Equilibrium for a demand of 3,000 vehicles

However as the travel demand decreases, the congestion effect become less prominent and larger proportion of traffic stay on route 1.

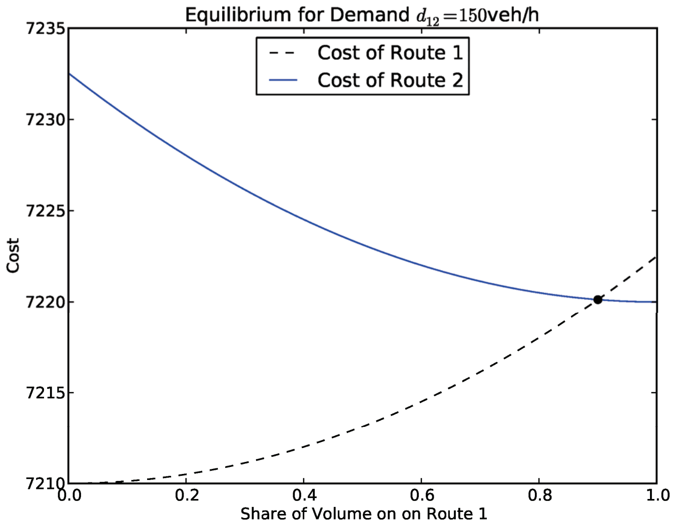

Figure 39: Equilibrium for a demand of 150 vehicles

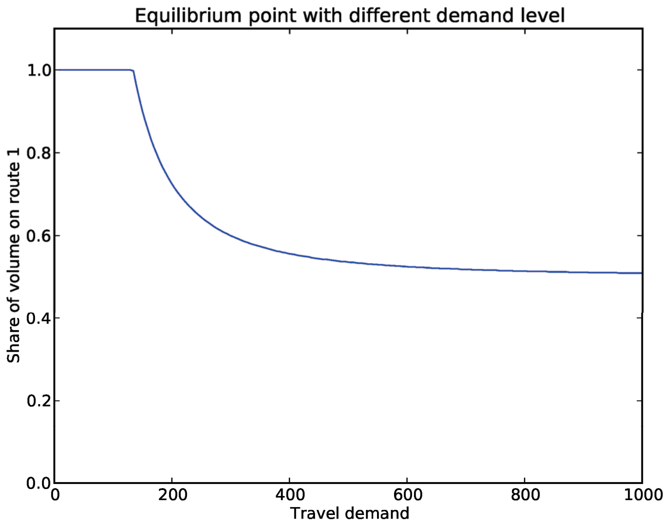

Figure 40: Equilibrium point moves as the travel demand increases

When the demand falls below a certain threshold all travellers choose the shortest route, which is basically equivalent to an all-or-nothing assignment. This is not realistic for practical applications. This problem becomes even more apparent if bidirectional travel is considered. Traffic from zone1 to zone2 will use route $3,1,2$ while traffic from zone2 to zone1 will use route $2,4,3$. This causes imbalance for bidirectional traffic on a specific link.

Although the example in this section is idealized, similar pattern can be easily identified in real world cases. For example, when modelling a city with densely populated downtown and large low density suburbs, traffic in the suburb area will well below capacity so that here the DUE traffic assignment tends to produce an all-or-nothing like results.

5.3 Test Case

In this case study, the route generation procedure is applied to a real world network in the Munich region, Germany. The idea is to show that the route generation procedure is capable of reproducing given a reference route sets when the constraint parameters are calibrated.

5.3.1 Overview of Wiki Project

This test case is based on a data set acquired in the project “Wiki - Wirkungen individueller und kollektiver Verkehrsinformation auf den Verkehr in Ballungsräumen” (FRIEDRICH, MANDIR, & PILLAT, 2012). The main purpose of the Wiki project was to explore the potential of individual and collective traffic information to influence the traffic behaviours of travellers in metropolitan areas in order to reduce time and fuel consumption.

The Wiki project chooses greater Munich area as the test bed and conducted surveys in which the following data are collected:

- A highly detailed NAVTEQ network of the metropolitan area of Munich, including both major roads and most of the minor roads.
- A transport planning model from PTV AG, Karlsruhe, that covers a wider area of southern Bavaria leaving out minor roads to reduce the network size. The supply side of this model includes a network that contains 7,700 nodes and 22,600 links; the demand side includes an hourly demand data of 844 zones for a whole day.
- Commuter routes for 300 commuters working in Munich city and living in the northern suburbs. The survey covers 8 weeks and it collected revealed routes from GPS trajectories, recorded by special devices that were handed out to the commuters at the start of the project. It also collected stated routes reported by commuters in an interview.
- State of traffic flows and traffic information devices during the survey period. These data are used as context to explain the route choice behaviour of commuters.

Based on these data the Wiki project produced the following results:

- A route choice set generation procedure
- Calibrated route and departure time choice model considering the influence of various information devices.
- A traffic assignment model incorporating the route choice model and information availability models.

More details on the data collection and processing in the Wiki project are provided in (MANDIR, 2012).

5.3.2 Networks and Reference Route Sets

The Road Network

In this test case, the road network and the reference route sets are extracted from the data set of the Wiki project. The road network shown in Figure 41 covers the area of the city centre and the strategically important road network in the north of Munich.

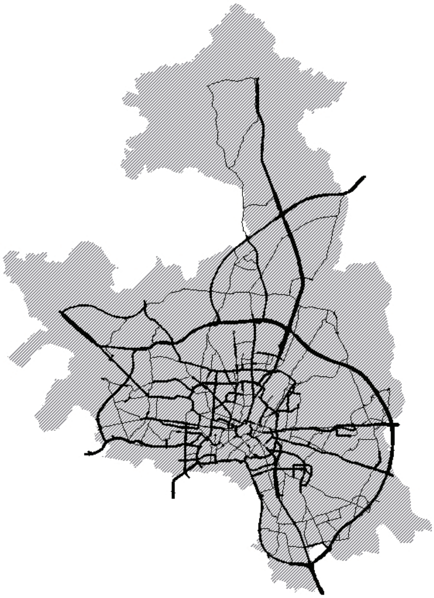

Figure 41: Road network of Munich, Germany

The road network consists of 2,607 nodes, 7,776 links and 24,376 turns. The key attributes of links and turns are listed in Table 5 and Table 6.

Name	Describption
FromNodeNo	The id of the upstream node
ToNodeNo	The id of the downstream node
Length	The length of the link
CapPrt	The link capacity for private transport
V0Prt	Free flow speed for private transport
LinkType/Name	The name of the link type. Several kinds of information are encoded In the this string, including the connection level hierarchy, road category and number of lanes
LinkType/Rank	A numerical rank of the link. Important links have a lower rank number

Table 5: Key attributes for links

Name	Describption
FromNodeNo	The id of the from node
ToNodeNo	The id of the to node
ViaNodeNo	The id of the via node
FromLink/LinkType/Rank	The rank of the from link
ToLink/LinkType/Rank	The rank of the to link

Table 6: Key attributes for turns

To consider turns during the route generation process, the road network is expanded into a graph following the procedure described in section 5.1.1. The expanded graph contains 9,950 vertices and 13,956 edges. For each edge, the hierarchy-adjusted cost is calculated based on the method described in section 5.2.1.

The Reference Route Set

The reference route set in the test case is taken from the synthesized route set of the Wiki project, which combines the observed routes from GPS trajectory, stated routes from the interview and generated routes from a route generator. The composition procedure is described in (PILLAT, MANDIR, & FRIEDRICH, 2011). The synthesized route set contains routes whose origin and/or destination are located outside network used in this research. These routes are truncated to fit to the Munich network. For each route in the composited route set, the truncation procedure delete all links that do not exist in the Munich network and reassign the origin and destination according to the position of the first and last link in the remaining links. If no links of a route exist in the Munich network then this route is discarded.

After the truncation procedure 83,860 routes remain. These routes needs to be further cleaned in the following steps:

- Remove duplicated route segments. The truncation procedure may create duplicated route segments. For example if two routes share identical link sequences in the range of the Munich network but diverge outside this range, the truncation procedure will produce identical result for each of them.

- Remove non-continuous segments. The truncation procedure could produce non-continuous segments for routes entering and leaving the Munich network multiple times.
- Assign origin and destination for truncated route segments based on adjacent zone centroids.
- Add connectors to route segments. Complete the route by connecting origin zone to the segment tail and the destination zone to the segment head.
- Remove cycles in the routes.
- Remove OD-pair that has less than 10 routes.

After the clean-up, the remaining routes constitute the reference route set for this test case. This reference route set contains 25,910 routes between 731 unique OD-pairs. As shown below the routes in the reference route set distribute unevenly among OD-pairs.

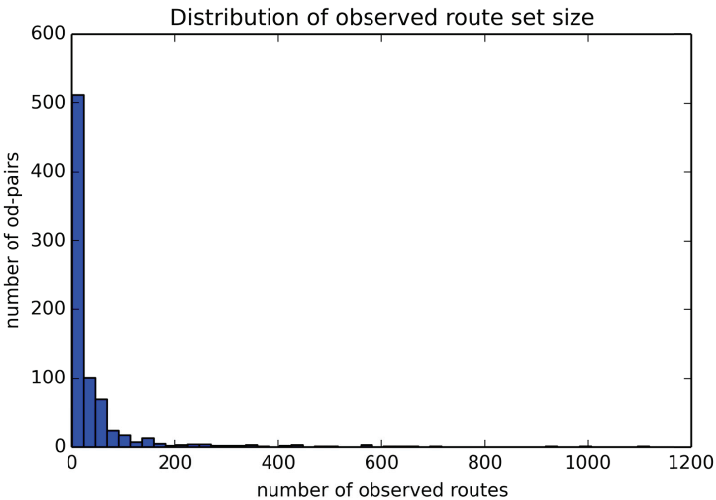

Figure 42: Observed route set size distribution among different OD-pairs

5.3.3 Result of Route Choice Set Generation

With the observed route set as the reference, the parameters of the general constraints are calibrated with the method described in section 3.4. The result for OD-pair 9162018-600023606 and 9162026-600023606 are shown in the following tables.

Constraints	Parameter
Global Detour Factor	1.83
Local Detour Factor	1.95
Local optimal Threshold	416

Table 7: Parameters for OD-pair 9162018-600023606

Constraints	Parameter
Global Detour Factor	1.77
Local Detour Factor	1.95
Local optimal Threshold	226

Table 8: Parameters for OD-pair 9162026-600023606

The route set for all OD-pairs are generated with the respective constraint parameters. The total running time on a laptop with Intel i7-2640 quad core CPU and 4GB RAM is 5:24:13, and 70910 routes are generated.

Figure 43 shows the generated master set for OD-pair 9162018-600023606. The left side is generated with parameters in Table 7. It produces 49 routes, which has a 92% coverage of the reference set. The right side shows a route set generated with local optima threshold increase to 1000. Only 4 routes remain. This example shows that local optimal threshold reduce the size of the generated route set and maintain the geographical disparity, which is consistent with the validation scenarios.

Figure 44 shows the generated master set for OD-pair 9162026-600023606. The left side is generated with the parameters in Table 8. It produces 342 routes, which has a 95% coverage of the reference set. The right side shows a route set generated with local detour threshold increase to 2.5 and local optimal threshold increases to 300. The coverage rate degenerate a bit to 79% however the number of routes reduced drastically to 54. This result indicates the automatic calibration procedure still can be improved to match the quality of the parameters manually adjusted.

Figure 43: OD-pair 9162018-600023606

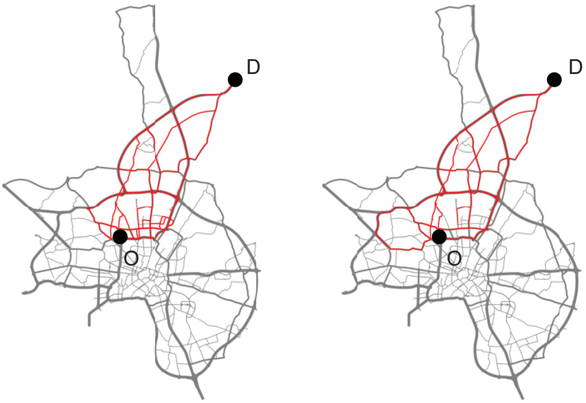

Figure 44: OD-pair 9162026-600023606

The generated routes of all OD-pairs are compared with the reference routes of all OD-pair. The result is shown in Figure 45. It shows that as the Overlapping Threshold increases, the coverage rate drops. Only 31% of observed routes are perfectly reproduced by the route generation procedure. However if 80% of overlapping is regarded as a successful reproduction, then the coverage rate is 71%.

Still this coverage rate result is noticeably lower than similar branch-and-bound route generation methods (PRATO & BEKHOR, 2006). One possible reason is the quality of the reference route sets. The Wiki reference route set is significantly larger than the route set in other research projects where the quality of each route cannot be checked. Another possible reason is that the calibration procedure chooses stringent parameters for the constraint parameters. This avoids excessive routes but results in a lower coverage rate. The coverage rate can always be improved by using loose parameters.

Since for selected OD-pair the coverage rate is much better than the overall coverage rate, the route generation procedure and/or the parameter calibrating procedure works not so well for some OD-pairs. It is interesting to identify these OD-pairs and analyse the reason.

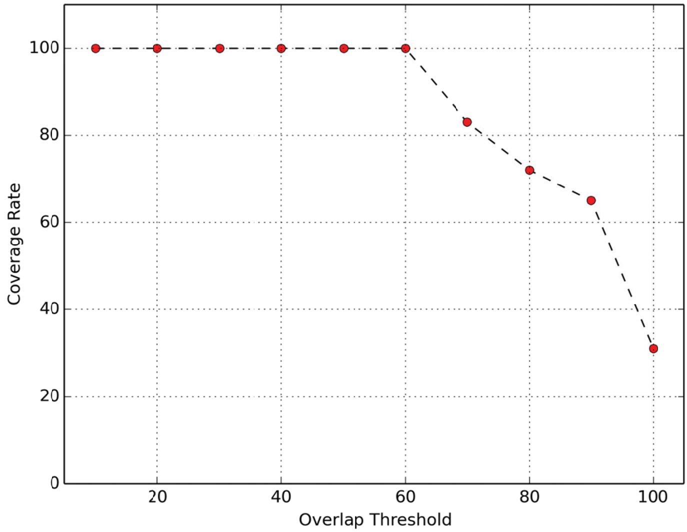

Figure 45: Degenerate of the coverage rate as the overlap threshold increases

6 Conclusion

6.1 Summary

Traffic assignment models play an important role in both research and practice of transport planning. Traditionally traffic assignment models are link-based predicting traffic flows for each link in a transport network. Link-based assignment models are computationally efficient. However, advanced applications require more detailed information on traffic flow distribution between alternative routes. Link-based assignment models cannot fulfil these requirements. Developments in computer technology, however, permit route-based traffic assignment.

The aim of this research is to develop a general framework for a route-based traffic assignment, with a special focus on the quality of the route sets. In order to achieve maximum flexibility, a two-stage scheme is adopted. The assignment procedure is divided into an explicit route set generation stage and a travel demand distributing stage.

Before developing the route set generation procedure, the behavioural assumptions of travellers in identifying and choosing routes are discussed. Based on this discussion, constraints for good route sets are proposed including both general constraints, which apply to all trips and special constraints, which apply to specific individuals or person groups. The general constraints used in this research include global detour constraints, local detour constraints, local optimal constraints and hierarchical constraints.

The route set generation stage is further divided into two stages. The route enumeration stage identifies the master set of reasonable routes using general constraints. The route set composition stage further filters the master set with different special constraints for different individuals. It also composes the remaining routes into a consideration set.

Based on this route set a general framework for distributing travel demand among routes is presented. It involves a route choice model taking advantage of the hierarchic structure of the route choice set. It uses a simple convergence procedure based on the method of successive average. A reference implementation of this framework is given in this research and is applied to a small-scale example and a real world scenario of Munich, Germany for validation. The results show that the proposed framework can reproduce observed reference route sets well. At the same time, the framework is flexible enough to cover other criteria.

6.2 Contributions

The main contribution of this research can be summarized as follows:

- The first contribution of this research is a new method for assessing the quality of route choice sets. Instead of following reference route sets rigidly, this method defines the quality of a route set by its compliance with a set of constraints. Each constraint represents one a priori assumption about an adequate route set. The reference route sets provide information for calibrating these constraints. The constraints are divided into two categories. The general constraints are applicable for all travellers in a network, while the special constraints are applicable for different individuals or person groups.
- The second contribution is the two-stage route set generation procedure. The first stage uses a branch-and-bound enumeration to find a master set for all travellers that operationalizes the general constraints. The second stage filters the master set for each individual or person group, using the corresponding special constraints. It also assembles the remaining routes into a consideration route set while enforcing constraints such as the number of considered routes. The application scope of this route set generation procedure includes conventional traffic assignment models, which requires group route sets and microscopic simulation, which requires individual route sets. It can also be applied to supply side analysis such as evaluating the accessibility of a road network.
- The third contribution is a general route based assignment procedure in which specific network loading models and flow balancing models can be incorporated for various applications. An explicitly generated route set enables more sophisticated route choice models handling non-additive utilities, route specific utilities and improves the correlation of overlapping routes.

6.3 Future Research

Based on the work presented in this thesis, the following directions for further work become apparent.

The assignment framework presented in this thesis is general enough to incorporate different kinds of network loading models and flow balancing procedures. In the reference implementation, the simplest variants, the volume-delay function as network loading model and the successive average as flow balancing procedure, are used as the route set generation is the main focus. For real world application, incorporating more sophisticated network loading models and more efficient flow balancing procedures may be beneficial.

In this research, the two-stage route set generation procedure is applied to the Munich scenario. The result shows that it can reproduce the reference route set relatively well. However, the constraints, especially the individual constraints are set as a prior assumption based on the understanding of this specific scenario. It is important to apply the route set generation procedure to other networks to test the transferability of these constraints among different travellers and networks.

In order to limit the scope of this research, the route set generation procedure is restricted to private road traffic in a static framework. The basic idea of a branch-and-bound enumeration and similarity-based composition can be expanded to generate routes for other modes or even intermodal transport. It would also be interesting to analyse how dynamic network information influences the route set generation.

List of References

AKGÜN, V., ERKUT, E., & BATTA, R. (2000, 3). On finding dissimilar paths. European Journal of Operational Research, Volume 121, Issue 2 , pp. 232-246.

AZEVEDO, J. A., COSTA, M., MADERIA, J., & MARTINA, E. (1993). An algorithm for the ranking of shortest paths. European Journal of Operation Research, 69 , 97-106.

BARCELO, J., & FERRER, J. (1997). Microscopic traffic simulation for att systems analysis. Universitat Politecnica de Catalunya, Barcelona.

BAR-GERA, H. (1999). Origin-based algorithms for transportation network modeling. Chicago, IL: Ph.D. thesis, Civil Engineering, University of Illinois at Chicago.

BEKHOR, S., BEN-AKIVA, M. E., & RAMMING, M. S. (2006). Evaluation of choice set generation algorithms for route choice models. Annals of Operations Research, Volume 144, Issue 1 , 235-247.

BEN-AKIVA, M. E. (1996). Development of a deployable real-time dynamic traffic assignment system, task D interim report: analytical developments for DTA system. ITS Program, Cambridge (MA), MIT ITS Program.

BEN-AKIVA, M. E., & LERMAN, S. R. (1985). Discrete choice analysis: Theory and application to travel demand. MIT press.

BEN-AKIVA, M., BERGMAN, M. J., DALY, A. J., & RAMASWAMY, R. (1984). Modeling inter-urban route choice behaviour. Proceedings of the 9th International Symposium on Transportation and Traffic Theory (pp. 299–330). VNU Science Press, Utrecht, The Netherlands.

BENJAMINGS, M., & LINDVELD, C. (2002). Multimodal travel choice modelling: a supernetwork approach. 81st TRB Annual Meetings. Washington DC.

BLIEMER, C. M., & TAALE, H. (2006). Route generation and dynamic traffic assignment for large networks. In Conference proceedings 1st DTA conference. Leeds, UK.

BOVY, P. H. (2009). On modelling route choice sets in transportation networks: A synthesis. Transport Reviews, 29(1) , 43-68.

BOVY, P., & STERN, E. (1990). Route choice: way finding in transport network. Kluwer Academic Publisher.

Bureau of Public Roads. (1964). Traffic assignment manual. U.S. Dept. of Commerce, Urban Planning Division.

CASCETTA, E. (2001). Transportation systems engineering: theory and methods. Kluver Academic Publishers, Dordrecht, Netherlands.

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 Proceedings of the 13th International Symposium on Transportation and Traffic Theory, (pp. 697–711). Lyon, France.

CASSETA, E. (2001). Transportation Systems Engineering: Theory and Methods. Kluver Academic Publishers, Dordrecht, Netherlands.

DAFERMOS, S. C., & SPARROW, F. (1969). The traffic assignment problem for a general network. J. of Res. of the National Bureau of Standards, 73B, , 91-118.

DAGANZO, C. F. (1995). The cell-transmission model, part II: Network traffic. Transportation Research, 29B (2) , pp. 79-93.

DAGONZO, C. F., & SHEFII, Y. (1977). On stochastic models of traffic assignment. Transportation Science, 11(3) , 253-274.

DE LA BARRA, T., PEREZ, B., & ANEZ, J. (1993). Multidimensional path search and assignment. In Proceedings of the 21th PTRC Summer Meeting.

DEL CASTILLO, J. M., & BENITEZ, F. G. (1995). On the functional form of the speed-density relationship-I: General theory. Transportation Research, 29(B), 5 , pp. 373-389.

DIAL, R. B. (1971). A probabilistic multi-path traffic assignment mode which obviates path. pp. 83-111.

DIAL, R. B. (2000). Bicriterion traffic equilibrium: T2 model, algorithm, and software overview. Transportation Research Record: Journal of the Transportation Research Board, No. 1725 , 54–62.

DIJKSTRA, E. W. (1959). A note on two problems in connection with graphs. Numerische Mathematik , S. 269-271.

FELLENDORF, M. (1996). VISSIM for traffic signal optimisation. Traffic Technology International'96. Annual Review Issue , pp. 190-192.

FGSV. (2008). Richtlinien für die integrierte Netzgestaltung. Köln: Forschungsgesellschaft für Straßen-und Verkehrswesen.

FIORENZO-CATALANO, S., & VAN DER ZIJ, N. (2001). A forecasting model for inland navigation based on route enumeration. Proceedings of the European Transport Conference (pp. 1-11). PTRC Education and Research Services Ltd., London.

FRIEDRICH, M. (2010). Route choice in road networks - observing and modelling route choice. Proceedings of the 5th International Symposium Networks for Mobility. Stuttgart.

FRIEDRICH, M., HOFSÄSS, I., & WEKECK, S. (2001). Timetable-based transit assignment using branch & bound. Transportation Research Record, 1752 , pp. 100–107.

FRIEDRICH, M., MANDIR, E., & PILLAT, J. (2012). wiki – Wirkungen von individueller und kollektiver ontrip Verkehrsbeeinflussung auf den Verkehr in Ballungsräumen, Schlussbericht der Universität Stuttgart. Forschungsbericht Förderkennzeichen 19 P 7048 B (BMWi).

GARLING, T., & GOLLEDGE, R. G. (2000). Cognitive mapping and spatial decision making. In Cognitive mapping: Past, present, and future (pp. 44-65).

GAWRON, C. (1998). Simulation-based Traffic Assignment; Computing User Equilibria in Large Street Networks. University of Cologne, Cologne.

GOLLEDGE, R. G. (1999). Wayfinding behavior: Cognitive mapping and other spatial processes. JHU Press.

HALATI, A., LIEU, H., & WALKER, S. (1997). CORSIM-corridor traffic simulation model. Congestion and Traffic Safety in the 21st Century: Challenges, Innovations, and Opportunities. Chicago.

HOOGENDOORN-LANSER, S. (2005). Modelling travel behaviour in multi-modal networks, Ph.D Thesis. TRAIL Research School, Technical University of Delft, The Netherlands.

JAYAKRISHNAN, R., & MAHMASSANI, H. S. (1994). An evaluation tool of advanced traffic Information and management systems in urban networks. Transportation Research C 2C(3) , pp. 129-147.

LEONARD, D. R., & POWER, P. (1989). CONTRAM: structure of the model. Transportation Research Laboratory, Crowthorn.

LIGHTHILL, M. J., & WHITHAM, G. B. (1955). On Kinematics waves. Part II: A theory of traffic flow on long crowed roads. Proceedings Royal Society, (London), A229, (pp. 281-345).

LITTLE, J. D., MURTY, K. G., SWEENEY, D. W., & KAREL, C. (1963). An algorithm for the traveling salesman problem. Operations research, 11(6) , pp. 972-989.

Mahut, M. (2001). A multi-lane extension of the space-time queue model of traffic dynamics. TRISTAN IV, Azores Islands.

MANDIR, E. (2012). Potential of Traffic Information to Optimize Route and Departure Time Choice, Ph. D Thesis. Institut für Straßen- und Verkehrswesen, University of Stuttgart.

MANSKI, C. (1997). The structure of random utility models. Theory and Decision, 8 (3) , pp. 229–254.

MCFADDEN, D., & TRAIN, K. (2000). Mixed MNL models of discrete response. Journal of Applied Econometrics 15 , pp. 447–470.

MYERS, E. W. (1986). An O(ND) Difference Algorithm and its Variations. Algorithmica, Vol. 1 No. 2 , pp. 251-266.

NAGEL, K., & SCHRECKENBERG, M. (1992). A cellular automaton model for freeway traffic. J. Phys. I France 2(12) , pp. 2221-2229.

NIEVERGELT, J., & WIDMAYER, P. (1997). Algorithmic Foundations of Geographic Information Systems. Springer Berlin Heidelberg.

PILLAT, J., MANDIR, E., & FRIEDRICH, M. (2011). Dynamic Choice Set Generation Based on Global Positioning System Trajectories and Stated Preference Data. Transportation Research Record, 2231 , pp. 18-26.

POLLACK, M. (1961). The kth Best Route Through a Network. Operational Research(9) , 578-580.

PRATO, C., & BEKHOR, S. (2006). Applying branch & bound technique to route choice set generation. Transportation Research Record, 1985 , 19–28.

PTV Planung Transport Verkehr AG. (2012). VISSIM 5.40 user manual.

RAMMING, M. S. (2002). Network Knowlege and Route Choice, Ph.D Thesis. Massachusetts Institute of Technology, Cambridge, USA.

RICHARDS, P. I. (1956). Shockwaves on the highway. Operations Research, 4 , pp. 42-51.

SCHLAICH, J. (2009). Nutzung von Mobilfunckdaten fuer die Analyse der Routewahl, Ph. D Thesis. Institut für Straßen- und Verkehrswesen, University of Stuttgart.

SCHLAICH, J., OTTERSTAETTER, T., & FRIEDRICH, M. (2010). Generating trajectories from mobile phone data. In Proceedings of the 89th Annual Meeting Compendium of Papers. Transportation Research Board of the National Academies.

SCOTT, K., PABON-JIMENEZ, G., & BERNSTEIN. (1997). Finding alternatives to the best path. Proceedings of the 76th Annual Meeting of the Transportation Research Board. Washington, D.C.

SHEFFI, Y., & POWELL, W. B. (1982). An algorithm for the equilibrium assignment problem with random link times. Networks, 12 , pp. 191–207.

SMITH, M., & DRUITT, S. (1994). Paramics Final Report. University of Edinburgh,Edinburgh.

TOLEDO, T., & KOUTSOPOULOS, H. (2003). Calibration and validation of microscopic traffic simulation tools: Stockholm Case Study. Transportation Research Record: Journal of the Transportation Research Board(1831) , pp. 65-75.

TVERSKY, A. (1972). Elimination by aspects: A theory of choice. Psychological review 79(4) , p. 281.

VAN DER ZIJPP, N. J., & FIORENZO-CATALANO, S. (2005). Path enumeration by finding the constrainted k-shortest paths. Transportation Research Part B, 39 , 545-563.

WARDROP, J. (1952). Some theoretical aspects of road traffic research. Proceedings of Institute of Civil Engineering Part II, 1(2), (pp. 325-378).

YAI, T., IWAKURA, S., & MORICHI, S. (1997). Multinomial probit with structured covariance for route choice behavior. Transportation Research Part B, 31(3) , 195-207.

YANG, Q. (1997). A simulation laboratory for evaluation of traffic management systems, Ph.d Thesis. Massachusetts Institute of Technology.

Glossary

ATIS	Advanced Traveller Information System are systems that provide travel related information to travellers.
BPR	Bureau of Public Roads is a division of the United States Department of Transportation that specializes in highway transportation
Choice set	Set of alternatives to choose from in decision making. In this research it refers to a set of routes.
Connector	An element in a transport network model that connects a zone to the network. It represents the access and egress cost from the zone centroid to a node in the network.
Consideration Set	The general idea of a route set considered in a compensatory manner by travellers in route choice process.
Cost	The generalized disabilities of a network element. Common factors contribute to the cost includes travel time, distance, road toll, etc.
DUE	Deterministic User Equilibrium is an assumption of traffic assignment modelling. It assumes that travellers have perfect information of the traffic condition in the network and each traveller seeks to minimize the personal travel cost.
Edge	An element in a graph. It connects two vertices.
GIS	A geographic information system (GIS) is a computer system designed to capture, store, manipulate, analyse, manage, and present all types of geographical data.
GPS	The Global Positioning System (GPS) is a space-based satellite navigation system that provides location and time information.
Group Consideration Set	The group consideration set is the union of many individual consideration sets.
Individual Consideration Set	The individual consideration set is the route set considered by an individual travellers. It is generally of limited size because the traveller's short-term memory

	have limited capacity.
ITS	Intelligent transportation system integrates advanced computer information processing, communications, sensors, electronic technologies, and management strategies to increase the safety and efficiency of the surface transport system.
Link	An element in a transport network model that connects two nodes. It generally represents a road or a lane.
LOS	Level of service (LOS) is a qualitative measure used to relate the quality of traffic service on roads based on performance measure like speed, density, etc.
Master Set	For a given OD-pair, the master set is the set of all reasonable routes.
MSA	Method of Successive Averages is an iterative heuristic for linear optimisation.
NAVTEQ	NAVTEQ is an American Chicago-based provider of Geographic Information Systems (GIS) data and a major provider of base electronic navigable maps.
Node	An element in transport network model that Generally represents an intersection.
OD-matrix	An OD-matrix represents the travel demand in a transport system. Each element represents the number of trips between the respective origin and destination.
OD-pair	An OD-pair includes an origin and a destination for trips in a transport network.
OSM	Open Street Map is a collaborative project to create a free editable map of the world. It is an important source for free digital map.
Path	A series of consecutive vertices in a graph.
PCU	Passenger Car Unit is a metric used in Transportation Engineering, to assess traffic-flow rate on a highway. A Passenger Car Unit is essentially the impact that a mode of transport has on traffic variables (such as headway, speed, density) compared to a single car

RIN	Richtlinien für die integrierte Netzgestaltung. A German guideline for integrated network design.
Route	General terms for a sequence of nodes and links in a transport network.
SO	System Optimal is assumption of traffic assignment modelling. It assumes that travellers cooperate to minimize the total cost of the transport system.
SUE	Stochastic User Equilibrium is an assumption of traffic assignment modelling. It assumes that travellers have imperfect information of the traffic condition in the network and each traveller seeks to minimize the personal travel cost.
Turn	An element in a transport network model that represents a turning movement at an intersection.
Universal Set	For a given OD-pair, the master set is the set of all possible routes.
Vertex	An element in a graph.
VISUM	A transport modelling software package from PTV Group, Germany
Zone	An element in a transport network model. It represents a land area that generates and attracts trips.

Schriftenreihe des Instituts für Straßen- und Verkehrswesen

Heft 1	1. Fachkolloquium Straße und Verkehr - Stuttgart 1986	9/1987
Heft 2	2. Fachkolloquium Straße und Verkehr - Stuttgart 1988	4/1989
Heft 3	Arbeiten aus dem Institut für Straßen-und Verkehrswesen	7/1989
Heft 4	Beiträge zum Ruhenden Verkehr	8/1989
Heft 5	Festkolloquium anlässlich der Vollendung des 60.Lebensjahres von Prof. Dr.-Ing. Gerd Steierwald	5/1990
Heft 6	3. Fachkolloquium Straße und Verkehr - Stuttgart 1990	4/1991
Heft 7	Wolfgang Hertkorn Veränderungen des Kraftstoffverbrauchs und der Abgasbelastungen durch Geschwindigkeitsreduktion in untergeordneten städtischen Straßennetzen	2/1992
Heft 8	City Stuttgart - attraktive Innenstadt im Zentrum einer Agglomeration (Autofreie Innenstadt = Attraktive City?)	3/1992
Heft 9	Arbeiten aus dem Institut für Straßen-und Verkehrswesen	4/1992
Heft 10	Reinhard Unkhoff Der Einfluß des Schwerverkehrs auf Geschwindigkeitsverhalten und Pulkbildung auf Autobahnen	6/1992
Heft 11	Stefanos Bazis Interaktiver Straßenentwurf mit Hilfe der EDV	7/1992
Heft 12	Jiann-Sheng Wu Stochastische Simulation des Überholverhaltens auf zweistreifigen Landstraßen	7/1992

Heft 13	4. Fachkolloquium Straße und Verkehr - Stuttgart 1992	5/1993
Heft 14	Stephan Herkt Abstimmung der Leistungsfähigkeit von Knotenpunkten mit Lichtsignalanlage und knotenpunktfreien Strecken an innerörtlichen Hauptverkehrsstraßen	4/1994
Heft 15	Von den Anfängen bis zur Gegenwart -Straßen- und Verkehrswesen an der Universität Stuttgart	9/1994
Heft 16	Ashraf Hamed Abwicklung des Lieferverkehrs innerörtlichen Hauptstraßen	10/1994
Heft 17	5. Fachkolloquium Straße und Verkehr - Stuttgart 1994	3/1995
Heft 18	Chronik und Retrospektive - Gerd Steierwald und das Institut für Straßen- und Verkehrswesen	10/1995
Heft 19	Heiko Tempel Zusammenhänge zwischen Verkehrsablauf, Kraftstoffverbrauch und Schadstoffemissionen im städtischen Busverkehr	11/1995
Heft 20	Volker Mörgenthaler Strategien zur Verminderung der Emissionen von Ozonvorläufersubstanzen des Nutzfahrzeugverkehrs in Ballungsgebieten	2/1996
Heft 21	Martin Schmid Auswirkungen der Kraftstoffbesteuerung auf die Pkw-Fahrleistungen im Freizeitverkehr	11/1996
Heft 22	Burkhard Flasche Fahrraumbewirtschaftung mit Sonderfahrstreifen für ausgewählte Stadtverkehre	2/1998
Heft 23	Szabolcs Törgyekes Ableitung spezifischer Abgas-Emissionsfaktoren für Kraftfahrzeugkollektive anhand von Konzentrationsmessungen in einem städtischen Straßentunnel	3/1998

Heft 24	Konrad von Kirchbach Zur Entwicklung der Straßen in Württemberg zwischen 1700 bis 1918	3/2000
Heft 25	6. Fachkolloquium Straße und Verkehr - Stuttgart 1999	3/2009
Heft 26	Symposium Stadt und Verkehr – Mobile Konferenz	11/1999
Heft 27	„Institut 21“ Aufbruch in Forschung und Lehre (nicht erschienen)	
Heft 28	Arbeiten aus dem Institut für Straßen-und Verkehrswesen	7/2001
Heft 29	Stefan Denzinger Auswirkungen alternierender Telearbeit auf das Verkehrsverhalten	4/2001
Heft 30	7. Fachkolloquium Straße und Verkehr - Stuttgart 2001	3/2009
Heft 31	Prof. Elisabeth Deakin Changing Importance of Central Cities	8/2002
Heft 32	Walter Vogt, Matthias Lenz, Henrik Schwarz, Wilhelm Glaser, Margrit Glaser, Thomas Kuder Verkehrliche Auswirkungen von Teleshopping und Telecommerce auf die Mobilität privater Haushalte	6/2002
Heft 33	Christoph Hecht Modellierung des Zeitwahlverhaltens im Personenverkehr	2/2005
Heft 34	Henrik Schwarz Vergleich von SchallImmissionsmessungen bei Straßenverkehrslärm mit Standardisierten und Binauralen Geräten	11/2002
Heft 35	Peter Schick Einfluss von Streckenbeeinflussungsanlagen auf die Kapazität von Autobahnabschnitten sowie die Stabilität des Verkehrsflusses	6/2003

Heft 36	Volker Schmid Reaktivierte Nebenbahnen und alternative Buskonzepte Vergleich von ökologischen Wirkungen	7/2003
Heft 37	8. Fachkolloquium Straße und Verkehr - Stuttgart 2003	3/2009
Heft 38	Matthias Lenz Auswirkungen des Ausbaus der verkehrlichen Infrastruktur auf das regionale Fernpendleraufkommen	2/2005
Heft 39	Steffen R. Herrmann Simulationsmodell zum Wasserabfluss- und Aquaplaningverhalten auf Fahrbahnoberflächen	2/2008
Heft 40	Manuel Galster Modellierung von Anbindungen in Verkehrsplanungsmodellen	8/2009
Heft 41	10. Fachkolloquium Straße und Verkehr - Stuttgart 2008	8/2009
Heft 42	Johannes Schlaich Nutzung von Mobilfunkdaten für die Analyse der Routenwahl	6/2010
Heft 43	Florian Bitzer Mikroskopische Modellierung und Simulation des Fußgängerverkehrs	9/2010
Heft 44	Eileen Mandir Potential of Traffic Information to optimize Route and Departure Time Choice	7/2012
Heft 45	Anne Wolff Simulation of Pavement Surface Runoff using the Depth-Averaged Shallow Water Equations	3/2013
Heft 46	Stefan Alber Veränderung des Schallabsorptionsverhaltens von offenporigen Asphalten durch Verschmutzung	3/2013

Heft 47	Thomas Otterstätter Methoden zur Erfassung von Verkehrsströmen und Fahrzeiten mit stationären fahrzeugwiedererkennenden Detektoren	8/2013
Heft 48	Bruno Arbter Numerische Bestimmung der akustischen Eigenschaften offenporiger Fahrbahnbeläge auf Basis ihrer rekonstruierten Geometrie	5/2014
Heft 49	Juliane Pillat Methoden zur Analyse und Prognose des Verkehrsaufkommens unter Berücksichtigung des Wetters auf Autobahnen	7/2014
Heft 50	Jochen Lohmiller Qualität des Verkehrsablaufs auf Netzabschnitten von Autobahnen - Bewertung unter Berücksichtigung der Zuverlässigkeit und Analyse von Einflussfaktoren	7/2014
Heft 51	Katrin Keller Modellierung und Optimierung des Verkehrsablaufs in Evakuierungssituationen	9/2014
Heft 52	Yaohua Xiong A Framework for Traffic Assignment with Explicit Route Generation	9/2014