MASTER DISSERTATION

ROUTE CHOICE MODELLING BASED ON EMPIRICAL EVIDENCE

CASE STUDY ALKMAAR

T.F. van Duivenbooden, BSc

CENTRE FOR TRANSPORT STUDIES
FACULTY OF ENGINEERING TECHNOLOGY

EXAMINATION COMMITTEE
prof. dr. ir. E.C. van Berkum
dr. T. Thomas

UNIVERSITY OF TWENTE.

March 2012

“A route differs from a road not only because it is solely intended for vehicles, but also because it is merely a line that connects one point with another. A route has no meaning in itself; its meaning derives entirely from the two points that it connects. A road is a tribute to space. Every stretch of road has meaning in itself and invites us to stop. A route is the triumphant devaluation of space, which thanks to it has been reduced to a mere obstacle to human movement and a waste of time.”

Milan Kundera
Immortality

Preface

The statement made by Kundera (see previous page) is quite bold, but touches upon the essence of most travellers: the time spend between an origin and a destination should be as short as possible. The process from starting at University to getting a degree is, in a way, a travel too; and I can certainly say that I have not used the fastest route. Now that I am on the final straight of this travel, with the destination in sight, I can look back to a travel in which every detour I made has been an experience, such that I do not regret taking these detours.

This Master's dissertation is the final hurdle of my travel through University. The subject I have chosen, route choice, has definitely been interesting. It has shown that many routes do exist between an origin and a destination, and that not all travellers choose to travel the, what appears to be, fastest route.

Every destination of one trip is the origin of the next. I am looking forward to the new voyage I am about to embark on, and I can only hope that every stretch along that trip will have a meaning as well.

TvD

Summary

Route choice modelling is one of the major processes in modelling road traffic. The rational thought to use the fastest route available, i.e. a minimisation of travel time, is a logical one, though would require perfect knowledge of travel times with travellers. In reality this will not be the case with all travellers; they may assume that another route is the fastest, and therefore choose that route. Therefore, multiple routes between an origin and a destination would be used, each with a certain ratio.

Many different route choice models exist, from quite simple to very elaborate. The quite simple logit-based models are often used to determine a route's choice probability within a set of routes, though they are still up for improvement on accuracy.

Research into this field requires empirical data. A recent addition in road side systems in the Netherlands is the Bluetooth detection system by the traffic information provider VID. This system is currently available in several regions in the Netherlands, of which the Alkmaar region is one. The data captured by seven detectors in Alkmaar in February 2011, courtesy of VID, has provided the opportunity to look at both the use of VBM systems in route research and to route choice behaviour itself. The aim of this research has therefore been to ascertain the ability of the VBM systems to be used in route choice research and, using the data, to determine relevant route attributes to be included in route choice models to improve accuracy.

Road side systems, like Bluetooth detection, is not able to directly determine the exact route of a detected device, but only can provide a sequence of locations using the passage times and IDs acquired by the detectors. To determine what routes are likely to be used with a certain sequence, a choice set has been generated using the Constrained K Shortest Path method by van der Zijpp and Fiorenzo Catalano (2005), based on 42 OD-pairs – for each of the seven detector locations as an origin, six other locations can be a destination – and a network comprising of collector roads, arterial roads and carriageways. 126 routes have been generated, and compared to the sequences from the data.

Using the unique identifier of every captured Bluetooth device, sequences have been made for each device. The time between subsequent detections has been used to determine if a stop has been made between those locations, which would mean that a trip would have ended and a new one started. The cut-off point has been set at 1.5 times of

the typical travel time between the locations, or in case of a double detection at one location at 30 seconds. 320 different trip sequences have been found, each with 1 or more observations.

Comparison of the generated routes and the observed trip sequences has learned that many sequences are not logical, either overly long or (partly) circular. Another problem has revealed to be detection-based: it has appeared possible that devices can pass a detector, but may not be detected (false-negative), as well as that a device that nears a VBM detector, though its route does not directly pass the detector, may be detected (false-positive). This has seemed to be the case in at least 32 per cent of all observed node sequences. By altering the sequences, i.e. adding or removing a node, nearly half of the sequences involved could be allocated to a likely route. The sequences that could not be allocated have been discarded. The attached number of observations with an illogical sequence however has appeared to be very low, such that over 99 per cent of all observations has been attributable to a generated route. Although requiring correction for errors, data from Bluetooth detectors does provide substantial evidence to be used for route choice research.

Using the observed route ratios and observed travel time differences, two relationships have been estimated, based on logit. The evidence has suggested that route utility is not linearly related to travel time, and it may even include an offset below which routes are evaluated as being equal. However, there still are some large residuals between the estimated model and observations. To improve the accuracy, the effects of easy-to-collect route attributes have been analysed using a regression analysis. Five attributes appeared to be of interest, all based in the type of road, directional changes and signage.

The devised model including route attributes, i.e. using a generalised travel time, outperforms a model using only the observed travel time. However, a comparison with the basic logit model and the model devised by Thomas and Tutert (2009, 2010) has revealed that neither of the estimated models of this research are performing best. The most accurate results have been found by using a model as proposed by Thomas and Tutert (2009, 2010), though with alteration to the scale parameter. The differences however are quite small.

Contents

Chapter 1

Introduction

1.1 Background

1.1.1 Route choice in traffic modelling

Attempts to predict traffic flows in a network date back to the 1950s, but the complexity and non-linearity of traffic behaviour has not yet led to a general theory, applicable to all situations. To date, the basic structure of most transport models is based on the classic four-stage model (Ortuzar and Willumsen, 2001). This approach starts with socio-economic data to estimate the number of trips generated and attracted by each zone in the study area (Step 1: trip generation), the allocation of trips to particular origins and destinations (Step 2: distribution), modelling choice of mode (Step 3: modal split) and finally allocation of trips to the network (Step 4: assignment). Several alternatives in the sequence of the steps are thought up, though assignment is always the final step.

Assignment comprises the determination of a set of routes within each OD-pair, determination of each route's choice factor, calculation of volumes per route, and, as one link in the network may be host to multiple routes, summation of respective route volumes per link. Route choice determination is key in this process. The current practices to determine route choice factors however seem to be inaccurate and therefore not always produce realistic results (Telgen, 2010).

The basic route choice model is based on the rational thought to choose the fastest route available, i.e. a minimisation of travel time, as described by Wardrop (1952). This does require perfect knowledge of travellers on the travel time. In reality, travellers may not have this knowledge and may assume another path to be the fastest, especially when time-similar alternative routes are available. Even though, in practice, methods are often still based on the fastest route principle, due to its simplicity.

Methods incorporating all routes (in a finite set) produce a probability of each of the routes (e.g. multinomial logit), and appear to be more accurate (e.g. Telgen, 2010). A local study in Enschede, as described by Thomas and Tutert (2009, 2010), has however showed that there still might be some difference between an 'ordinary' logit model and reality. Thomas and Tutert; Thomas and Tutert propose a revised logit model, though to what extent this model is valid in other situation remains to be seen.

1.1.2 Applications of travel time data

Travel time is a significant cost to making a trip (Noland and Small, 1995; Noland and Polak, 2002); most travellers will therefore aim to reduce the travel time of a trip. As said earlier, travellers may not have perfect knowledge on the actual travel time, and therefore choose an alternative route over the actual fastest route. However, the assumption can be made that at some point an alternative route is unlikely to be chosen, because the travel time difference to the fastest route is excessively long, i.e. that the route is too much of a detour. Travel time is therefore useful to determine the feasibility of routes in a choice set.

Travel time can also be used to determine if an observed trip from travel data is genuine; e.g. when an observed travel time of an individual trip is significantly higher than the average found travel time, the individual observation is unlikely to be representative for that situation. Such an exercise is particularly useful when sampling rate is low, i.e. when the distance between successive observations of a vehicle (or equivalent) is large, as there would be no detailed information available between observations. If the individual travel time between observations can be considered as an outlier, the observation should not be taken into account, and therefore not be included in determination of volumes.

Average travel time may differ over time, due to e.g. traffic conditions. With low volumes the average travel time would be calculated on a small number of observations, and is therefore not very accurate; in fast changing conditions, taking an average over a longer time leads to inaccuracy, whereas taking a short period would involve less observations, affecting accuracy as well. Choosing suitable, steady data therefore can be done using travel time.

Travel time therefore is a significant factor in route choice research, beyond only being an attribute of route choice. To be able to measure both travel time and route volumes therefore is critical.

1.1.3 Measurement of travel time and route volume ratio

Various technologies are available to measure the travel time of a vehicle, using either devices or features that are unique to respective vehicles, e.g. a dedicated GPS system or licence plate recognition, or using signals from generic devices like mobile phones.

By recording the time at which a vehicle passes certain points in a network, that vehicle's travel time between the recorded points can be calculated. When recording passage times at multiple intermediate points, it may be possible to construct the route taken. Accuracy naturally improves with a higher number of intermediate points, thereby limiting the amount of possible routes between two observations.

Another factor of importance is sample representativeness. For both route and travel time analysis a large sample would be preferred, such that the observed route fractions and average travel times would be similar to the actual values. Furthermore, individuals within the sample should behave naturally. I.e., using only data derived from GPS navigation (e.g. TomTom) to determine route choice, individuals that do not have a navigation device are not included in the research, which therefore may lead to a difference between sample and reality.

Both accuracy and sample representativeness are therefore important aspects in route research. However, accomplishing both aspects in one system would require extensive costs, which is not preferred. Finding a balance between route and travel time accuracy, sample size and research costs is difficult; concessions will have to be made.

The newest addition to data collection methods is Bluetooth measuring; such a system is currently rolled out throughout the Netherlands by VID, the VID Bluetooth Measuring-system (VBM), and is able to provide data on a large scale, with lower costs compared to other large scale systems. Determining route choice behaviour has not been the goal of this system, though can be done. The accuracy in determining travel time has been tested and is adequate; the accuracy to determine route choice is however not known. As this system may be able to provide a suitable alternative to other methods, leading to less concessions to be made, it is subject of this research.

1.2 Objective and research questions

To determine real-life route choice behaviour large-scale data collection is necessary. The newly developed network of VID Bluetooth Measuring (VBM) systems is primarily used for the determination of travel times, but can also be used to determine the order in which a device passes different VBM locations. Reconstruction of routes would therefore be possible, though the accuracy of VBM data in doing so is not yet known.

The case study in Alkmaar provides an opportunity to take a look at both the VBM system and route choice behaviour. Given this case study, this research therefore aims to ascertain the accuracy of VBM system data to determine route choice behaviour, and, based on the data, to determine the influence of route attributes on route choice behaviour.

In order to accomplish this objective two research questions are formulated, each with a set of sub-questions.

1. To what extent are the VBM systems in Alkmaar able to detect routes correctly, and what can be done to correct errors?
   (a) What data manipulations are necessary to determine vehicle routes and route volumes from Bluetooth device detection?
   (b) Can observed routes be explained by a generated set of paths?
   (c) Can observational errors be corrected?
2. What route attributes influence route choice, given the empirical evidence, and does this coincide with previously found relationships?
   (a) What are route attributes of interest?
   (b) What influence do these attributes have on route choice?
   (c) How do the findings compare to previous research?

1.3 Research methodology

This research is divided into two stages. The first stage is a review of the data collection method. To determine to what extend the observed routes, i.e. node sequences constructed from the empirical data, are correct a comparison is made with likely paths.

A generated set of paths, the *choice set*, is build using a set of rules or constraints, and describes the likely, possible paths within the respective OD-pairs. The choice set should ideally include all valid observed routes, though it is expected that this will not be the case. By looking into both observed routes that are not included in the choice set, and not-observed routes that are in the choice set, i.e. the deviations to the likely paths, problems in the accuracy of the VBM system to be used for route determination will show. However, within the set of likely routes problems may show as well, to be determined by looking into route volumes. This can only be done after the expected volumes are known, which is part of the second stage of this research.

The second stage is the actual research into route choice behaviour, and will describe the effect of route attributes on route choice. By performing an empirical cycle to determine a route choice model based on travel time, the route attributes of importance can be found, and be used to make up a route choice model. By testing this model, the accuracy can be determined; possible errors in observed data might show as well. The results of the proposed model can be compared to previous research, to determine if the proposed model will benefit route choice model accuracy.

Based on the results from both stages, conclusions and recommendations will be drawn.

Figure 1.1 on page 12 shows the process described above.

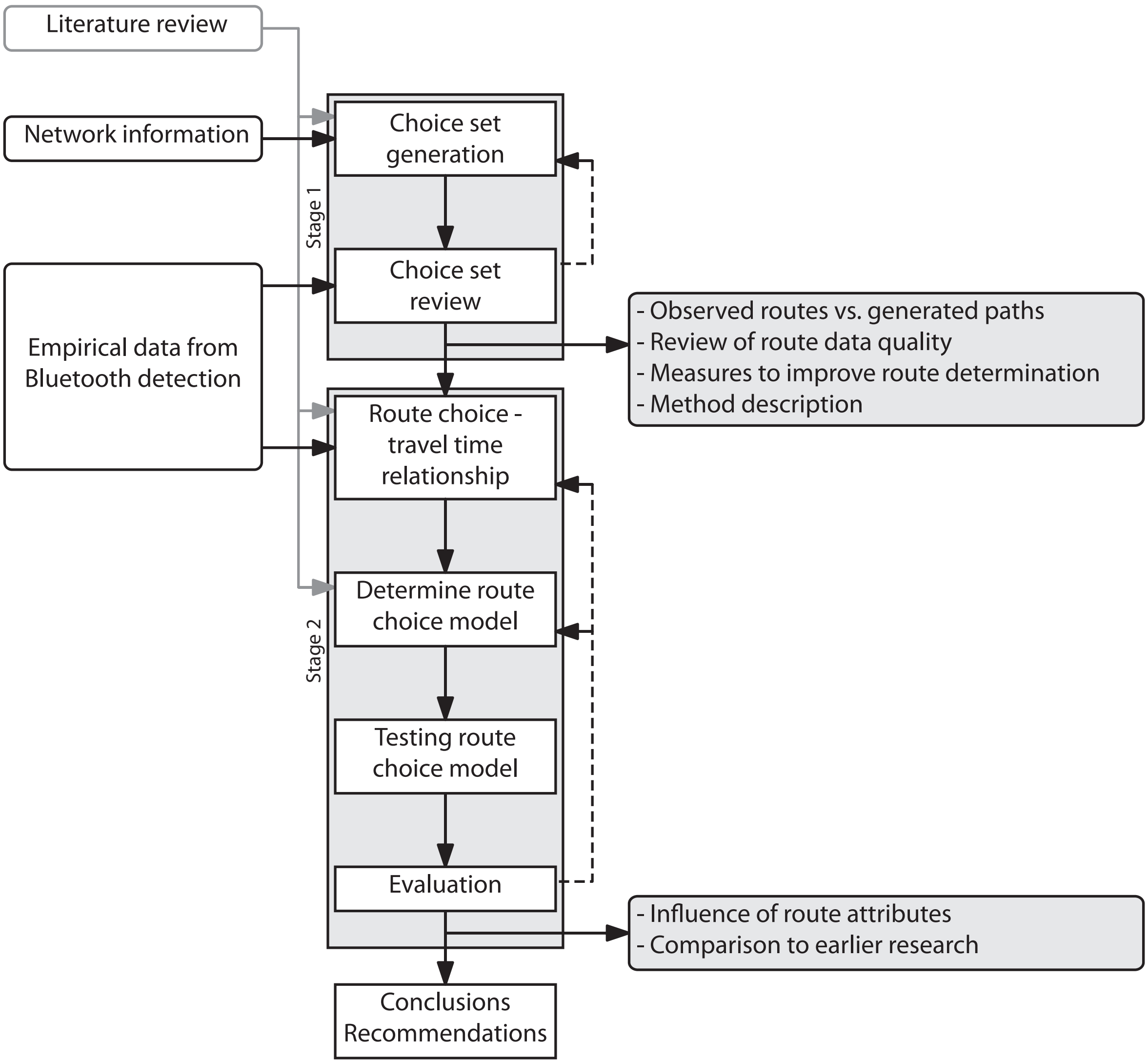

Figure 1.1: Research model

1.4 Dissertation outline

To provide a ‘road-map’ of this dissertation a brief summary of the main objectives and results of each chapter is given below.

Chapter 2 provides a review of existing choice set generation methods, route choice models and data collection methods.

Chapter 3 describes the Alkmaar network, the locations of VBM detectors and subsequent origins and destinations. The used route generation algorithm is explained, as well as the parameters used to determine generated route feasibility. The chapter results in the identification of feasible routes.

Chapter 4 identifies reasons for data variability and decides on the data to be used to reduce variability, after which the method used for route inference and trip identification has been described. Finally the observed sequences are compared to the feasible routes, and possible reasons for differences are given.

Chapter 5 aims to find a relationship between observed travel time and route choice, and to describe the relationship in a model. Furthermore it identifies attributes to be taken into account in a generalised travel time to improve model accuracy.

Chapter 6 presents the conclusions and recommendations of the research undertaken in this dissertation.

Chapter 2

Literature review

Research into route choice behaviour is not a new field of study. In the early days of traffic modelling, Wardrop (1952) developed a set of principles, assuming that travellers try to minimise their travel time. More recent research identified a wider variety of route choice factors (e.g. Papinski et al., 2009). Application of such detailed findings in route choice models will lead to models that, most likely, will deliver a result that approaches reality into detail. Not withstanding the importance of such research and its explanatory power, the implementation of such detail in simple route assignment models is not practical.

Data collection is a major step in route choice research: determining both route ratios and travel times will have to rely on sufficient and accurate data collection. Section 2.1 describes several methods to collect such empirical data.

Not all possible routes between an origin and a destination are of importance. Identification of a suitable set of alternative routes, i.e. the generation of the *choice set*, can be done using several methods, with different quality and calculation time. Section 2.2 discusses this topic.

Section 2.3 discusses several existing models that determine the route choice probability, based on parameters found through research.

2.1 Data collection

2.1.1 Detection methods

The starting point of all methods is the detection of individual vehicles in a network: either by floating positioning, i.e. a vehicle can be followed throughout the network, or by stationary positioning, i.e. a vehicle's position can only be determined when it passes a stationary road-side detector.

Floating probe detection systems can detect the position of a probe, i.e. a vehicle or device, irrespective of its position. After polling the probe, its location can be map-matched to the road network. With regular polling a detailed route can be identified. Depending on the density of the road network and the accuracy of the polled location, differences between the calculated position and real-life position might exist.

Stationary detection systems capture vehicle or device characteristics on several locations in a network. Combined with the time-stamps of capture at the individual capture locations, an algorithm can determine the passage order and travel time of individual probes. However, compared to floating car data, several disadvantages exist. Friedrich et al. (2008) describe several, among those are: (a) vehicles that are detected at two (or more) locations might not be through-traffic, but could have a trip end in between that is both a destination and an origin (e.g. a courier service), and (b) minor detours between two locations can not be detected. The former problem is hard to solve; only when the stationary time between the trips is long enough it would be possible to filter such vehicles out. The latter problem is directly influenced by the number of capture locations. In a dense detection network, i.e. with detectors quite close to each other, there would be less possibility for a detour to exist between two detectors. However, if travel times of such routes differ significantly, it would be possible to detect this (as described by Friedrich et al., 2008).

Detectors can either detect vehicles (or devices stationary to vehicles, like dedicated GPS terminals) or devices that are in vehicles 'accidentally', e.g. mobile phones. With the latter option, vehicle details are mainly unknown, as devices are not linked to vehicles; it is even possible that multiple devices are based in one vehicle.

Table 2.1 shows examples of detection systems within the four quadrants.

	Positioning	
Detection of	**Floating**	**Stationary**
Vehicle	- Dedicated GNSS systems (e.g. GPS)	- Vehicle detection (e.g. ANPR)
Device	- Portable GNSS device (e.g. TomTom) - Mobile phone signal detection	- Electronic device detection (e.g. VBM)

Table 2.1: Detection system examples

Floating vehicle detection

Floating car data is the basis of many studies into vehicle tracking. Most surveys are small-scale, based on global navigation satellite systems (GNSS) data (e.g. Jan et al., 2000; Li et al., 2005). Most studies collect data by placing dedicated GPS receivers in vehicles; Chen et al. (2009) suggest that using GPS capable mobile phones, though less accurate, can be used as well. Due to the small scale of such studies, additional information can be easily collected using, e.g., travel diaries.

An important condition of data collection using GPS devices is that the devices do not give route guidance at the same time. GPS navigation systems, though far more widespread and therefore creating a much larger sample group, calculate the fastest route (either with or without using accurate data) and therefore (nearly) eliminate driver's choice. Route choice behaviour can therefore only be determined for drivers using navigation devices, and is therefore not representative for all traffic. Travel times along a route, however, can be used.

GPS is not perfect: Ochieng et al. (2003) conclude that the difference between the measured GPS position and the actual position can be up to 40 metres, with an average of 10 to 16 meters. In high-density networks this might lead to a mismatch when map-matching a GPS position (e.g. Chen et al., 2009). However, with a good map-matching algorithm (e.g. incorporating driving direction as well) correct matching can be achieved (Ochieng et al., 2003). With a sampling rate of a few seconds a travellers' route can be captured with great detail.

Though very able to describe route choice, the current small scale of GPS surveys might not deliver accurate results when scaling to a full network; large-scale surveys with several thousands GPS units are not deemed to be feasible in costs (Jan et al., 2000), the time and effort necessary to set up such large-scale surveys only add to impracticality.

Floating device detection

Floating car data can also be based on mobile phone signals, e.g. GSM, GPRS or UMTS, called the Cell Global Identity (CGI) method (e.g. van der Zijpp and Van Haastrecht, 2003; Yoo et al., 2005). Most vehicles have one or more mobile phones on board. A mobile phone constantly measures the signal strength of the nearby Base Transceiver Stations (BTS), and connects to the BTS with the strongest signal. Depending on location, a BTS has a range of several hundreds of metres up to several tens of kilometres; the area covered by the system is called a cell, which is typically split into three sectors of each 120 degrees (van der Zijpp and Van Haastrecht, 2003). Furthermore technologies are available to determine distance to a BTS; using triangulation, location precision is typically around 300 metres (Leduc, 2008). Map matching will also be necessary to determine the route followed by the mobile device. In a dense network this would be hard to do, considering the low precision.

Stationary detection of vehicles

A commonly known stationary system is Automatic Number Plate Recognition (ANPR), based on camera surveillance. Though every vehicle has a number plate, not all vehicles can be recognised by ANPR, due to environmental factors (e.g. precipitation, sun glare, etc.) or dirty or deformed number plates. Friedrich et al. (2008) have revealed that a properly set up ANPR-system would recognise at least 80% of all passing vehicles, but in most cases more than 90%. However, number plates are not necessarily recognised correctly. Two types of error might occur: (a) a polluted plate might lead to false identification of resembling characters (e.g. B vs. 8, D vs. 0) and (b) vehicles not in full view of the camera might lead to only a portion of its plate being recorded. When strictly comparing the license plate databases from different locations this can lead to mismatches. A properly set up system, however, does require many ANPR cameras, as one camera can only track one traffic-lane. Considering investment costs of approx. 20,000 Euros per single lane system, rollout over a reasonable network is a costly matter (Friedrich et al., 2008).

Stationary detection of devices

The Dutch traffic information provider VID has recently developed a different infrastructure-based system, the VID Bluetooth Measuring-system (VBM), which senses Bluetooth signals from e.g. mobile phones, and is used to determine travel times along a route. Bluetooth devices broadcast a fixed unique identifier, a Media Access Control (MAC) address. These broadcasts can be received by a roadside detector. Received signals might not necessarily come from vehicles, thus ‘polluting’ the databases. However, when looking into routes, most non-relevant modes can be filtered on travel time, e.g. a bicycle will have a significant higher travel time than a car and thus will be an outlier. In slower traffic however, e.g. inner city, slower modes like mopeds might not be filtered. Wijbenga and Boerma (2010) conclude that the number of captured MAC-addresses is about 50% of the number of vehicles. However, more than one MAC-address could be attached to a single vehicle. VID estimates that around 30% of passing vehicles are taken into account[1]. When intensities are low, the penetration rate can differ considerably, both down- and upward (Wijbenga and Boerma, 2010).

2.1.2 Data requirements

Robinson (2005) mentions several problems concerning measuring link travel time and describes requirements to diminish those problems. Measuring vehicles to determine routes, instead of travel time, does involve similar problems. With some minor alterations to Robinson, the set of requirements for route determination can be set as follows.

- Ability to measure accurately and precisely the links used by each vehicle, i.e. ability to infer the exact route of a vehicle.

[1] As said by Patrick Potgraven, VID, in an interview with author

- Ability to determine the origin and destination of trips.
- Ability to measure the routes of a sufficient and representative number (ideally all) of valid vehicles.

Route inference

With all detection systems a vehicle's route will have to be inferred from the data. Location information from the data is to be matched to a location on a known network: a process known as *map matching*. Using the passage times from the data, a sequence of mapped locations can be constructed. The exact route between such locations is not registered, and therefore unknown. However, if only one possibility exists between two observations, it can be assumed to be the route between those observations. With more than one possibility assumptions will have to be made: a route can then not be fully accurately inferred.

Two factors influence the ability to accurately infer the route: (a) accuracy in location, and (b) sampling rate. With floating car data the detected location is likely to be somewhat inaccurate: GNSS systems may be inaccurate up to several tens of metres, GSM survey are even less accurate. Algorithms as described by e.g. Ochieng et al. (2003) and Chen et al. (2009) are however able to map detected locations of a GNSS device to a link, resp. node, even with the inaccuracy in mind. With stationary detectors the location of the detector is known, with a detected vehicle within a certain range. Location determination will therefore be more accurate with such detectors, i.e. within several meters with the VBM system, up to nearly dead accurate with ANPR.

Sampling rate is an important factor as well. A high number of observations of a vehicle does lead to a short distance between observations, and therefore lower risk of existence of multiple routes between observation locations. GNSS surveys usually involve observations every couple of seconds, which leads to one observation every several tens of metres, depending on vehicle speed. With stationary detection, the number and locations of detectors determine the distance between detections, and therefore accuracy in determining routes. Mainly due to costs the number of detectors is usually limited; it is therefore likely that no connected route can be formed with stationary detectors, nor can the route between two subsequent detectors be determined in an accurate way, leading to ambiguity over the route taken between two detector locations. This is more likely to be a problem in an urban context (Robinson, 2005).

Determination of origin and destination of trips

An observed sequence may consist of more than one trip. In between trips a vehicle would be stationary for a certain time at a fixed position. Not all stationary positions would be classified as a stop between trips: e.g. waiting time at a traffic signal will lead to a stationary position.

GNSS A high sampling rate and quite accurate location makes it easy to spot a stationary position, the location on the network and the time of the stop. Trip ends are therefore quite easy to call.

GSM The poor location accuracy makes it difficult to spot a stationary position, and therefore trip ends.

ANPR & VBM A stop will have to be inferred from travel time between observations as it cannot be observed directly.

Vehicle population

For accurate route choice research a large vehicle population is preferred; each origin-destination pair should have a suitable amount of observations. Each OD-pair would have a different usage: some ODs will be more observed than others. The technology used for route choice research should be able to capture less used ODs too, and therefore would require a large sample size.

The current practise of GNSS surveys usually involves at most a couple of hundred probes (e.g. Jan et al., 2000). This would be only a small sample in the total vehicle population over a network, and therefore can be considered to be less accurate in determination of route ratios of less used ODs. Using common devices like GSM or Bluetooth devices as a probe leads to a far larger sample, up to several tens per cent.

Table 2.2 summarises the performance of the current practise of the earlier mentioned methods.

2.1.3 Choice of detection method

To determine route choice behaviour, empirical data would ideally need to be accurate in measuring a connected set of links, be able to describe only 'valid' vehicles and be based on a sufficient part of the vehicle population. Table 2.2 on page 20 shows that neither of the mentioned methods is fully co-operating.

However, with the route inference possibilities as described in section 6.1, both ANPR and VBM systems, i.e. stationary detection, are usable. As the VBM system is quite new and will be rolled out over an extensive area in the Netherlands, it would be a first to determine route choice behaviour on data from (a part of) the VBM detector network. Using this data might eventually lead to possible problems, assumptions and/or requirements, but would also 'pave the way' for future research.

Technology	Accuracy in measuring used route	Verification of	Capture of vehicle population
GPS survey	Good, measures (nearly) all links with good accuracy	Good	Very poor, usually based on a small sample.
GSM	Poor, due to low precision.	Poor	Moderate to good, though depends on the number of vehicles containing a mobile phone that is switched on.
ANPR	Depends on the number of detectors; detectors itself are very accurate	Poor – needs to be inferred.	Good, approx. 80% to 90% of all vehicles are detected. The number of vehicles matched together is likely to be somewhat less.
VBM	Depends on the number of detectors; detectors itself are quite accurate.	Poor – needs to be inferred	Moderate to good, approx. 30% of all vehicles are detected, though depends on the number of vehicles that contains a detectable device.

Table 2.2: Comparison of capture technology performance

This research project will therefore be based on data from the VBM detectors.

2.1.4 VBM specific detection errors

A VBM detector can receive signals from Bluetooth devices within a range up to several tens of metres, though the range is determined by the lowest emitting device, likely this will not be the detector, but the travellers' devices[2]. A difference of several metres in points of detection of different devices might thus be possible, which would give an uncertainty of several seconds on travel time. The relative effect however is quite minimal under normal circumstances, with detectors at least a few hundred metres apart.

Figure 2.1 shows a situation in which a difference may be significant. Two vehicles (1 and 2) travel from VBM1 to VBM2, over different routes. At VBM1, vehicle 1 is measured almost at the intersection (point A), whereas vehicle 2 is measured quite ahead of the intersection (point B) and thus might be in a queue. At VBM2 the situation is opposite; vehicle 1 is measured sometime before the intersection (point C), vehicle 2 on the intersection (point D). Thus the travel time of vehicle 2 incorporates node delay of both end nodes, whereas the travel time of vehicle 1 does not. Note that in this case VBM3 only captures vehicle 2, thus validating the routes of both vehicles.

Situations do exist, where one vehicle can be detected multiple times, i.e. when a vehicle comes into the range of the VBM, goes out of range and subsequently comes into range again. Such a situation is shown in Figure 2.2.

A perfect data collection method combining accurate route identification, proper origin and destination determination and capture of a large vehicle population, does not yet exist within restricted budgets. A assessment will have to be made between accuracy in route and trip-end information on one hand, and sample size on the other. For route choice research a large sample is preferred, which both ANPR and VBM surveys can provide. The main disadvantage of both systems, i.e. not being able to accurately determine trip-ends, can be largely overcome. With a continuing roll-out of VBM detectors in the Netherlands for travel time determination purposes, the opportunity exists for it to provide large scale data for route research as well. Due to data availability, a VBM survey will be used in this research.

2.2 Choice set generation

Numerous alternative routes may exist in a road network. For a particular origin-destination pair many possible routes are not suitable (e.g. because they are excessively indirect), and would therefore not be considered by travellers (Bekhor et al., 2006). A traveller's

[2]Considered to be the normal range and operation of Bluetooth devices

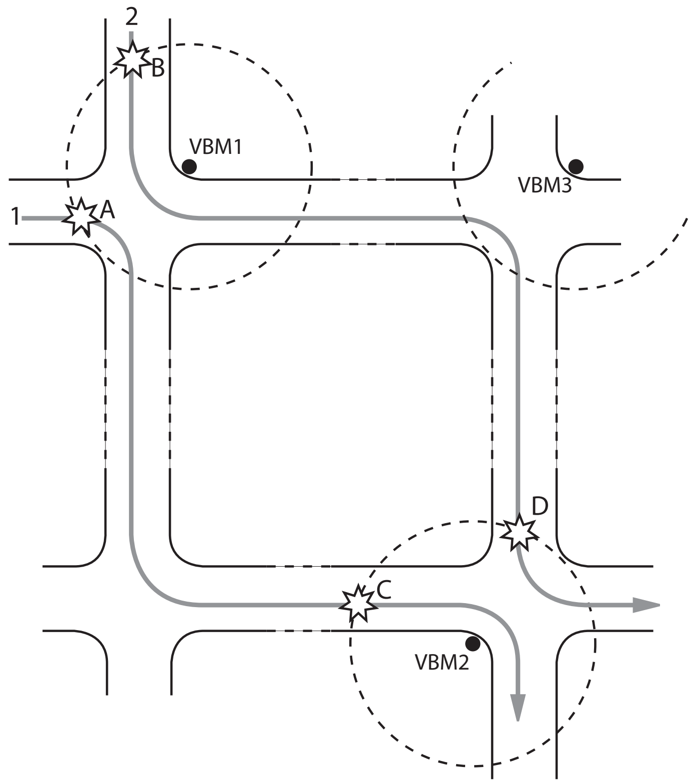

Figure 2.1: Location detection situation

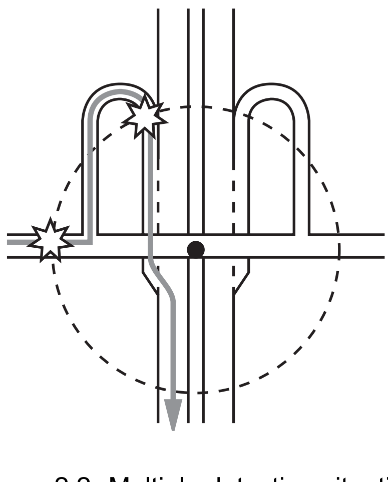

Figure 2.2: Multiple detection situation

knowledge of the road network and travel times differs from one to another; identification of routes that any traveller might consider is therefore necessary, avoiding identification of unconsidered routes to reduce computational effort.

To provide a set of rational alternative routes, in general *K-shortest path* algorithms have been used, which generate the first “k” shortest loop-less paths for a given OD-pair. Several approaches have been identified in previous research.

The most simple algorithm is to enumerate all possible paths, then sort from these the *K* paths that have the shortest length (Bock et al., 1957 ibid. Yen, 1971). Large numbers of computations are necessary, irrespective of the value of *K*. This algorithm will inevitable also identify many irrational routes, which then later on will be deleted; computational time is thus spoiled.

Pollack (1961) introduced a procedure based on the shortest-path problem. Starting with the shortest path between an OD-pair, the distance of each link in that path is , in turn, set to infinity. For each such case the shortest-path algorithm will be repeated, revealing additional routes. The number of computations required exponentially increases with *K*; this method would therefore work best when *K* is not large (Yen, 1971). Several variations of this method are identified, e.g. the *link elimination* approach (Azevedo et al., 1993), which removes a link in turn from the original shortest path, and the *link penalty* approach (De la Barra et al., 1993), which increases the costs of the links in the original shortest path.

The *branching* method (e.g. Hoffman and Pavley, 1959) considers deviations from the shortest path. For each node in the shortest path alternative route are determined by choosing the following node not on the shortest path. This method would determine the *K* shortest paths quickly when the Nth shortest path would branch immediately from the *N-1*th shortest path. Though depending on the network, this is assumed to be unlikely (Yen, 1971).

Not all *K*-shortest paths found might be relevant, e.g. they might be excessively overlapping. Introducing a set of constraints deals with this problem, resulting in the constrained *k* shortest paths. The *Constrained K Shortest Path (CKSP) method* determines the feasibility of the *m*th shortest path after its identification, either accepting or rejecting the route. This method does require identification of all possible routes, feasible or not, until the desired number of feasible paths is found.

If the amount of generated non-feasible routes can be reduced, it would be possible to reduce calculation time. To this effect, Van der Zijpp and Fiorenzo Catalano (2005) describe a method using subsets that are derived from the *m*th shortest path, based on the partitioning rule of Lawler (1976). Their method is based on the presumption that if the sequence of initial links that each path in a subset shares is unfeasible due to constraints, an extension of that set would also be unfeasible. Such a subset will therefore not have to be considered.

The *labelling* algorithm of Ben-Akiva et al. (1984) determines alternative paths based on (a combination of) path attributes. By setting criteria based on attributes, e.g. 'minimise travel time', a single route can be labelled to fulfil the set criteria. By considering multiple criteria, alternative paths may be found, forming the choice set; i.e. the choice set is a set of first shortest paths with different criteria. Note that one path may fulfil several criteria. Bekhor et al. (2006) have found that even when considering a large number of criteria, their research considers 16, a number of observed routes will not be found, as they are not ranked first considering any of the criteria, though travellers may consider other than top-ranked alternatives.

A *simulation* approach can also be used to determine alternative paths. Based on the presumption that travellers perceive path costs with error, link travel times will be drawn randomly from a distribution (Prato, 2009). Each draw will generate a shortest path, though this path may be similar to an earlier found path. A higher number of draws will lead to a larger set of possible paths. Simulation does not outperform the *k* shortest path method on quality, though requires less computational time (Bekhor et al., 2006), as only the shortest route will have to be found with a draw. Especially in large networks this will lead to a significant difference.

As this research does not incorporate a large network, computational time is not a large issue. It is therefore possible to use a method that generates a higher quality choice set, i.e. a *k* shortest path method. Incorporating route feasibility would further improve quality; it is however preferred to generate as less as non-feasible routes as possible. The CKSP-method based on the partitioning rule is therefore chosen as the preferred method in this research.

2.3 Route choice factors

After identifying feasible paths, traffic volumes along those paths will have to be determined: the *route choice factors* will have to be determined.

Most studies of travel have assumed that travellers try to minimise their individual time, cost or distance, based on Wardrop's principles of equilibrium (Wardrop, 1952), where (a) journey times in all routes actually used are equal and less than those which would be experienced by a single vehicle on any unused route and (b) at equilibrium the journey time is minimum.

The premise of the above is the assumption of a rational traveller, i.e. one choosing a route which offers the least perceived (and anticipated) individual costs (Ortuzar and Willumsen, 2001). This behaviour might however not be widespread among individual travellers. Many more types of choice factors exist, e.g. turn minimisation, fewest obstacles, congestion avoidance, road type, route directional signs, but also travel purpose

(Golledge and Gärling, 2002; Hamerslag, 1981; Bonsall, 1992). Variable message signs also influence route choice behaviour, though just a small minority appears to change their pre-determined route (Erke et al., 2007; Wardman et al., 1996).

As it is not practical to incorporate all types of choice in a route choice model, approximations are necessary. Therefore most choice models use a combination of time and monetary cost. Time appears to be the dominant factor in urban situations. However, just 60 to 80 per cent of observed routes can be explained by such a generalised cost function (Ortuzar and Willumsen, 2001).

Even when incorporating all possible choice factors, a difference between modelled and real-life route choice may exist. An important reason is difference in individual perceptions of what constitutes the 'best' route, due to e.g. wrong perception of route features or incorporation of different features in an individual cost function (Ortuzar and Willumsen, 2001). Route choice can furthermore be influenced by prior travel experience and a travellers' ability to (let) forecast travel links (Adler et al., 1993).

Wardrop's principles of equilibrium are therefore not likely to be met in real life, and user equilibrium models based on these principles are therefore argued to be unrealistic (Jan et al., 2000; Batley and Clegg, 2001, as described by e.g.). Furthermore, when incorporating these principles on route choice in uncongested environments, route choice models will only choose the most optimal, e.g. fastest, route.

Incorporating a stochastic element on the dominant route choice factor, i.e. journey time, can lead to more realistic assignment. Inspired by the discrete choice theory, several route choice models are available, of which most commonly known are the proportional method, multinomial logit, C-logit and Path Size logit. Many more logit type models exist though these appear not to be suitable for a range of network sizes or require a long calculation time, and are therefore not suitable for a simple, generic route choice model (Telgen, 2010).

In the above mentioned more suitable methods, travel time, or travel time combination, of route i, TT_i, where i is in a set of routes R in an OD-pair, is used in different ways, to determine the choice probability of route i.

Proportional model

The choice probability in the proportional method can be explained by the relative difference in travel times: routes with an equal relative difference will lead to an equal choice fraction difference. Equation 2.1 shows the expression of this model, with scale factor α a user-defined parameter that captures sensitivity to time and scale.

$$P_i = \frac{TT_i^{-\alpha}}{\sum\limits_{r \in R} TT_r^{-\alpha}} \quad (2.1)$$

When considering two routes with a relative difference factor β, i.e. $TT_j = \beta TT_i$, the choice probability function for route i evaluates as shown in Equation 2.2; irrelevant of the travel time, the relative difference shows to be of importance.

$$P_i = \frac{TT_i^{-\alpha}}{TT_i^{-\alpha} + TT_j^{-\alpha}} = \frac{1}{1 + (\frac{TT_j}{TT_i})^{-\alpha}} = \frac{1}{1 + (\frac{\beta TT_i}{TT_i})^{-\alpha}} = \frac{1}{1 + \beta^{-\alpha}} \quad (2.2)$$

Logit models

The logit based models are based on the exponential values of travel time difference; the choice probability by these models can be explained by the absolute travel time difference.

The multinomial logit method can be expressed as Equation 2.3, with scale factor θ a user-defined parameter.

$$P_i = \frac{e^{-\theta TT_i}}{\sum\limits_{r \in R} e^{-\theta TT_r}} \quad (2.3)$$

When considering two routes with a absolute travel time difference of β, i.e. $TT_j = TT_i + \beta$, the choice probability function evaluates as shown in Equation 2.4; irrelevant of travel time, the absolute travel time difference shows to be of importance.

$$P_i = \frac{e^{-\theta TT_i}}{e^{-\theta TT_i} + e^{-\theta TT_j}} = \frac{1}{1 + e^{-\theta(TT_j - TT_i)}} = \frac{1}{1 + e^{-\theta \beta}} \quad (2.4)$$

Both the proportional method and the multinomial logit method have the drawback that they are not able to distinguish routes with high overlap and thus overestimate the probability of those routes (the IIA property). Based on the multinomial logit model, the C-logit model was developed to overcome this problem (Cascetta et al., 1996 ibid. Casas i Vilaro, 2004). The C-logit model incorporates a route commonality factor *CF* that in fact is a travel time penalty to a certain route. The model can be expressed as:

$$P_i = \frac{e^{-\theta TT_i + CF_i}}{\sum\limits_{r \in R} e^{-\theta TT_r + CF_r}} \tag{2.5}$$

Similarly to C-logit is Path Size logit, which, as C-logit, adds a correction term, *PS*, to the utility function:

$$P_i = \frac{PS_i e^{-\theta TT_i}}{\sum\limits_{r \in R} PS_r e^{-\theta TT_r}} \tag{2.6}$$

Jan et al. (2000) suggest that these kind of path choice methods are developed without a base in empirical evidence. They state that the underlying assumptions have not received an adequate level of validation.

Empirical evidence

Thomas and Tutert (2009, 2010) have found from empirical evidence that a rational power of absolute travel time difference is an explanatory factor, combined with a travel time threshold. Their findings can be translated into a conditional logit model, expressed as:

$$P_i = \frac{e^{-\tau_i \theta((TT_i - TT_1)^\alpha - \beta^\alpha)}}{\sum\limits_{r \in R} e^{-\tau_r \theta((TT_r - TT_1)^\alpha - \beta^\alpha)}} \tag{2.7}$$

With:

$$TT_1 = \min_{r \in R} TT_r \tag{2.8}$$

Factors α and θ are user-defined variables expressing sensitivity to time and scale, β a user-defined time threshold below which routes are considered to be equal and τ_i is a two value variable:

$$\tau_i = \begin{cases} 1 & \text{if } TT_i - TT_1 > \beta \\ 0 & \text{otherwise} \end{cases} \tag{2.9}$$

Note that with $\alpha = 1$ and $\beta = 0$ this model is equal to the multinomial logit model.

A route's travel time is of course dependent to both distance and speed. Assuming a route's distance is known to travellers, the link speeds within this route determine travel time. The actual link speeds are however not commonly known (assuming travellers do not have access to real-time traffic information) and therefore assume a certain speed on a route, given the details of that route. Similar types of routes are therefore assumed to have a similar average speed. Road hierarchy is an important factor as well, whereas routes that are, e.g., more comfortable will be preferred over other routes. Evidence suggests these factors will have to be incorporated in route choice (Thomas and Tutert, 2010). In a similar fashion as Hamerslag (1981) a route's travel time can be calculated as follows:

$$TT_r = \sum_{i \in r} x_{i,w} \upsilon_w^{-1} \theta_w \tag{2.10}$$

Where $x_{i,w}$ is the distance of link i, with road type w; υ_w is the assumed speed; and θ_w is a road hierarchy factor. In case links within a route have similar values for υ and θ, can be simplified to:

$$TT_r = \upsilon_w^{-1} \theta_w \sum_{i \in r} x_{i,w} \tag{2.11}$$

The model of Thomas and Tutert cannot distinguish routes with high overlap, similar to the multinomial logit model, due to the IIA property. Appliance of a commonality factor or path size correction term might be a possibility.

There are many more utility based, and also non-utility based approaches. These approaches mostly incorporate more detail and are therefore less useable in a simple route choice model. Given the prosperous results of an uncomplicated utility-based route choice model based on empirical evidence, this type is further investigated in this research.

Chapter 3

Network and Choice Set

3.1 Study area

3.1.1 Alkmaar study area

The city of Alkmaar is a regional main city, with some 95,000 inhabitants, about 40 kilometres north-west of Amsterdam. It serves mainly as a regional centre, though tourism can be expected as well. The Alkmaar road network consists of a city circular road which is part of the regional road network, i.e. *arterial roads* and *carriageways*, main inner-city roads, i.e. *collector roads*, and *residential roads*. This local network is shown in Figure 3.1. A railway and some waterways dissect Alkmaar, with a limited number of crossings.

On regional level, Alkmaar is strategically located on the main corridors between Amsterdam and Zaanstad in the south, and Schagen and Den Helder in the north (see Figure 3.2). Furthermore, it is on an alternative route (avoiding Amsterdam) between the west of the Netherlands and Frisia. The Alkmaar network is therefore used by local, regional (e.g. Heiloo – Heerhugowaard) and interregional (e.g. Amsterdam – Den Helder) traffic. It can be expected that each of these types have different knowledge on the network; e.g., local drivers might know the inner-city network well, while interregional traffic only might know the city circular.

3.1.2 Road categorisation

All roads in the network can be associated with a certain road type, depending on road characteristics. The four types identified in this research are described on page 31; identification has been done on street level using observations. Figure 3.3 shows the network with the first three road types marked, all unmarked roads within the city circular are categorised as residential roads. Unmarked roads outside the city circular are not categorised.

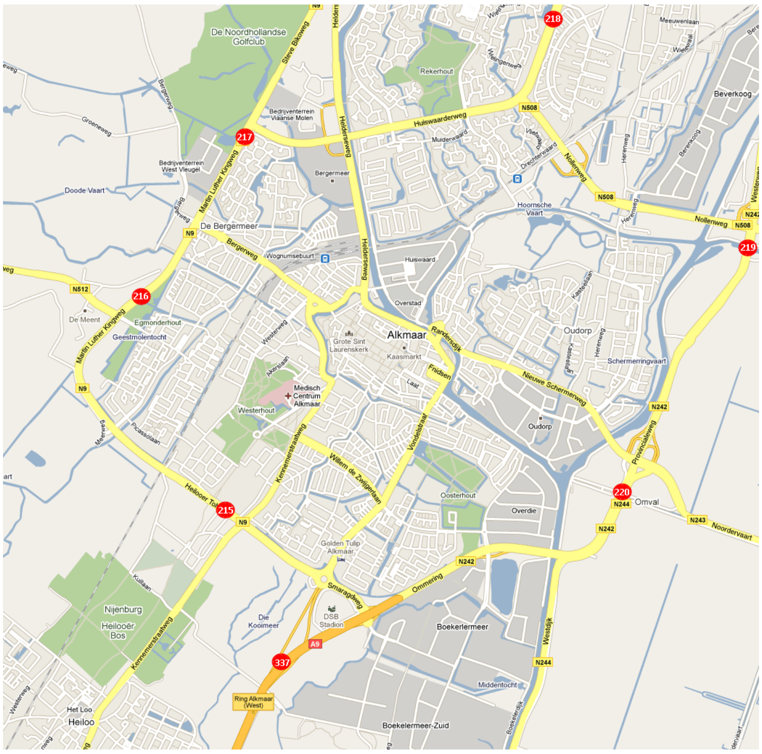

Figure 3.1: Alkmaar local road network

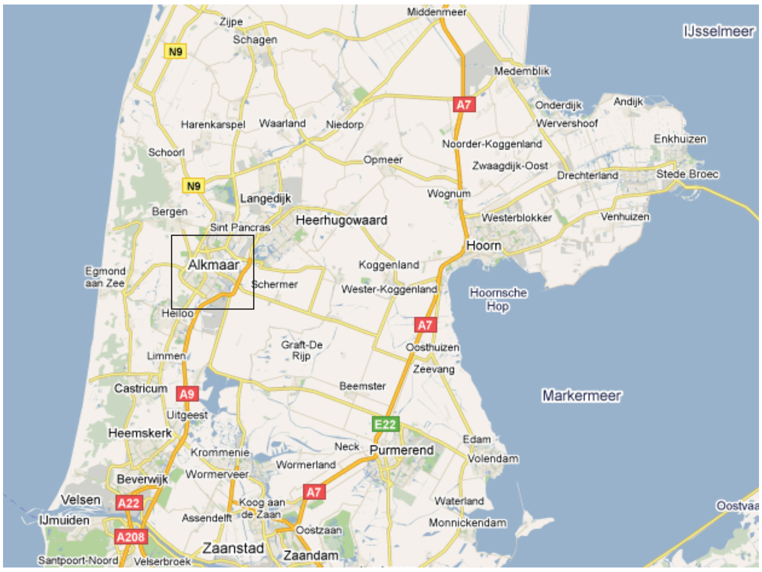

Figure 3.2: Position of Alkmaar in regional network

Dual carriageway Road with two or more traffic lanes per direction, separated by a central reservation; junctions are grade separated; maximum speed is at least 80 km/h.

Arterial road Road with one or more traffic lanes per direction, separated by a central reservation; maximum speed is at least 70 km/h.

Collector road Road with distinguished lanes per direction, i.e. with a separation line, with an asphalt concrete surface; one-way sections with similar specifications are included as well; connects with either a similar road or with an arterial road; maximum speed of 50 km/h.

Residential road All other urban roads, typically with a maximum speed of less than 50 km/h.

Figure 3.3: Road categorisation in the Alkmaar network. *Red*: dual carriageways; *blue*: arterial roads; *green*: collector roads; *unmarked within city circular*: residential roads.

3.1.3 Detectors

At seven locations in the network, Bluetooth detection systems are placed. These detectors scan the Bluetooth frequency band for devices such as mobile phones, car-kits, headsets, etc, and capture the devices' generally unique Media Access Control (MAC) address. For privacy protection reasons the captured MAC-addresses are encoded in a MD5-hash; these hashes are unique to the Bluetooth devices they are derived from, and cannot be easily decoded. It is therefore possible to construct routes and travel times by assembling data from multiple locations.

The locations are shown in Figure 3.4, each with a unique location code.

VBM215 At intersection of city circular and the road to Heiloo. May detect Bluetooth devices travelling on nearby residential roads, especially north-west of the intersection residential roads are close to the intersection.

VBM216 At intersection of city circular and road to Egmond (N512). There are no nearby residential roads, though there may be some pick-up of Bluetooth devices from nearby sports venues.

VBM217 At intersection of city circular and road to Den Helder (N9, N245). A close-by residential street is located south-east of the intersection.

VBM218 On the N245 (road to Schagen), north of the city circular. The road is flanked by residential dwellings, blocking signals from residential roads.

VBM219 At intersection of city circular (north-western corner) and carriageway to Heerhugowaard (N242, N508). No residential roads nearby.

VBM220 At intersection of city circular and road to Hoorn (N242, N243); also picks up devices travelling to/from the provincial road (N244) to the south.

VBM337 Located south of Exit 12 on the A9 motorway. Detects traffic from A9 to N9 and N242 vice versa. Probably also detects traffic on the parallel road and nearby fly-over.

3.1.4 Network generation and reduction

A generated network should represent a real-life transportation network as good as possible. As this research focusses on route choice between seven locations along the city circular and is therefore interested in through-traffic only, not all roads in the study area are of interest; for instance, residential roads are not likely to be used by through-traffic. Using the road categorisation (Section 3.1.2), a node is placed where higher level-of-service roads (i.e collector roads, arterial roads and dual carriageways) intersect each other, and where these roads end. Nodes are connected to each other considering the road type; they are linked if a collector road, arterial road or dual carriageway connects the nodes. One exception is made: if nodes can be connected by a residential road with a maximum distance of 250 meters there are considered to be linked; such a situation is likely where a road is partly degraded for road safety reasons, though may still be considered as a through-route. The generated network is reduced by deleting nodes and links that are certainly not part of a feasible path, e.g. dead-ends. Finally, restrictions in nodes (e.g. turn restrictions) and links (e.g. one-way links) are identified and processed in the network. The resulting network is shown in Figure 3.4.

Figure 3.4: Network used for route generation

3.2 Route identification

3.2.1 Origins and destinations

Each trip made through the Alkmaar area has an origin and a destination; the exact location of trip origins and destinations is not known, as there are no vehicle or device observations on street level. Therefore, irrespective of the actual origin or destination of a device, the *origin* is said to be the first observed node within a devices' trip and the *destination* is the last observed node of that trip. With seven observed nodes, and hence six possible destinations per node, 42 OD-pairs are observed.

For interregional and regional traffic, i.e. with trip starts and ends outside the city of Alkmaar, the first and last observed nodes can be said to be the entry and exit nodes to and from the Alkmaar network. Local traffic is likely to have either or both trip start and trip end in between observed nodes. Neither the exact locations, nor the routing to the first observed node can be determined; it is however still possible to determine route choice behaviour between the first and last observed nodes.

Example

Figure 3.5 shows three possible routes from a trip start location, *O*, to a trip end location, *D*. Devices travelling along either the red or green route are first observed at node 215 and last observed at node 219; these devices are then said to belong to OD-pair 215-219. However, devices travelling on the blue route are first observed at node 220, and last observed at 219, and therefore are included in OD-pair 220-219. Different routes from the trip start location to the trip end location may therefore lead to inclusion of trips in different OD-pairs. The full route choice of these trips can therefore not be determined. However, it still would be possible to use the information from trips along the red and green route to determine route choice between node 215 and node 219.

3.2.2 Route identification method

Identification of a choice set can be done in several ways, as described in the literature review (chapter 2). Considering the choice set algorithm evaluation of Bekhor et al. (2006) and the fact that the Alkmaar network involves only 42 OD-pairs with a relatively simple network, a *k* shortest path method has been considered feasible, using a link elimination approach. The reservations made by Bekhor et al. (2006) concerning the likely close resemblance of alternative paths when eliminating one link at a time has been overcome by using constraints, rejecting both strongly overlapping paths and paths that have a large detour. The Constrained *K* Shortest Path (CKSP) method by van der Zijpp and

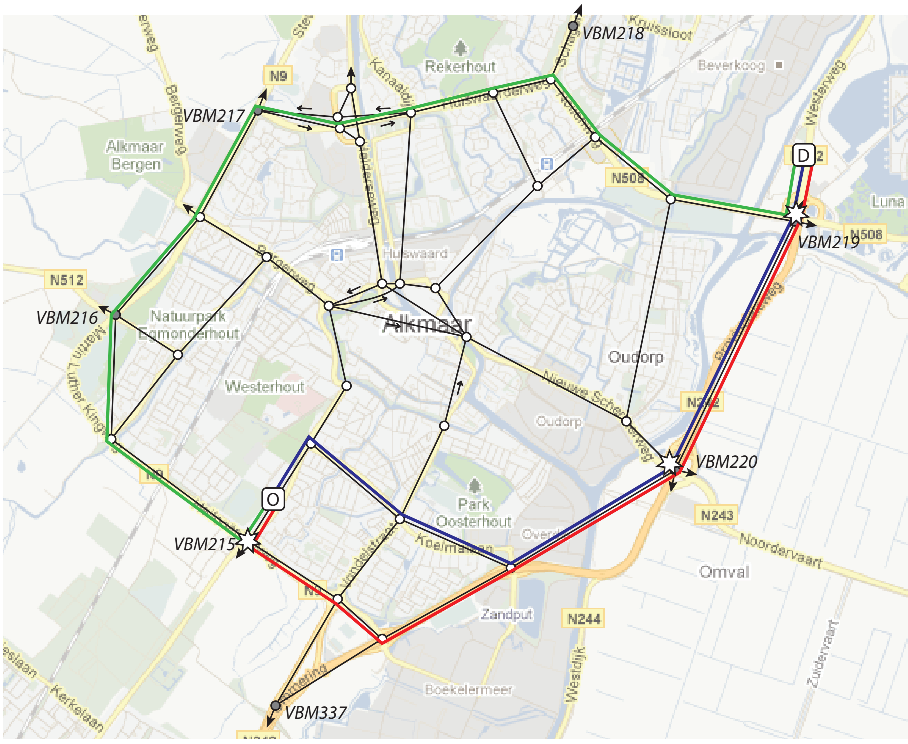

Figure 3.5: Influence of routes on observed OD-pairs

Fiorenzo Catalano (2005) has been used to that effect. The constraint parameters are chosen in such a way that as many observed routes as possible can be described by generated paths, though with few paths that are either unused, or where usage can not be determined.

Identification of a shortest path requires a definition of *path length*, i.e. what attributes of a path are used to determine its length. This length is defined by (a combination of) attributes. With no prior knowledge on the drivers' perception of the network, only data that can be derived from the links itself are certainties. This includes at least basic information linked to road attributes, e.g. distance and maximum speed, but can also include knowledge from prior research, or from captured data, e.g. travel time and delay. Only basic information on links in the Alkmaar network was available in this research.

Bekhor et al. (2006) have found that in a labelling approach the *least free-flow time*, i.e. distance divided by maximum speed, produces the best results, whereas *least distance* does not perform that well. Therefore, combining the least free-flow time and the CKSP method, i.e. determining the constrained *k* least free-flow time paths, has been used to identify the feasible paths in this research. In the remainder of this chapter, the path with the least free-flow time will be named as the *shortest path*.

3.2.3 Shortest path algorithm

Being able to determine a path's length, the *k*-shortest feasible paths can be identified. The CKSP-method has been used; van der Zijpp and Fiorenzo Catalano (2005) have formalised this method in the following algorithm. The shortest path is found using Dijkstra's algorithm (1959).

1. Define $S_{0,1}$ as the set of *all* paths from the origin to the destination.

2. Find the shortest path within $S_{0,1}$ and denote this with $P(S_{0,1})$

3. Divide $S_{0,1} - P(S_{0,1})$ into $q(1)$ mutually exclusive subsets $S_{1,1}, S_{1,2}, \ldots, S_{1,q(1)}$ with feasible initial links and a set of rejected paths R_1.

4. Set $m = 1$.

5. Compute the shortest paths for subsets $S_{m,1}, S_{m,2}, \ldots, S_{m,q(m)}$ and denote these paths with $P(S_{m,1}), P(S_{m,2}), \ldots, P(S_{m,q(m)})$.

6. Find the $(m+1)$th shortest path within all subsets that have been identified until now.

7. Let $S_{a,j}$ be the set that contains the $(m+1)$th shortest path (feasible or not). Divide $S_{a,j} - P(S_{a,j})$ into $q(m+1)$ mutually exclusive subsets $S_{m+1,1}, S_{m+1,2}, \ldots, S_{m+1,q(m+1)}$ with feasible initial links and a set of rejected paths R_m.

8. Set $m = m + 1$.

The partitioning into subsets has been done using the 'alternative partitioning rule' as described by van der Zijpp and Fiorenzo Catalano (2005). In short, a set of paths (not containing the shortest path in that set) is divided into feasible subsets according to the sequence of initial links and defined excluded links. Paths that do not have feasible initial links are rejected in a separate set.

3.3 Feasibility constraints

The feasibility of a route has been determined using a set of constraints, i.e. a *detour constraint*, which rejects paths that are excessively long, and an *overlap constraint* which rejects strongly overlapping routes.

Detour constraint The detour constraint eliminates all (sub)paths that contain a detour larger than factor Φ, relative to the shortest (sub)path, i.e. the length of any sub-sequence of links within a path should not be Φ times more than the shortest length between the begin-node and end-node of that sub-sequence.

Let $A_k = \{a_{1k}, a_{2k}, \ldots, a_{mk}\}$ denote the links of the *k*th shortest path. Let $L(a)$ be the function that maps link a to its length and $D[a_1, a_2]$ the length of the shortest path between the start-node of link a_1 and the end-node of link a_2.
Path k is *unfeasible* if a sub-sequence of links $\{a_{ik}, a_{(i+1)k}, \ldots, a_{jk}\}$ within path k can be found, for which the following is true (van der Zijpp and Fiorenzo Catalano, 2005).

$$\sum_{h=i}^{j} L(a_{hk}) > \Phi D[a_{ik}, a_{jk}] \tag{3.1}$$

Overlap constraint The overlap constraint eliminates all paths that have a strong overlap with shorter paths found, i.e., the total length of the shared sequences of links must not exceed a certain value (as described by Schnabel and Lohse, 1997; Fafieanie, 2009), or that the non-overlapping sequence of links has a minimum length (van der Zijpp and Fiorenzo Catalano, 2005), depending on the viewpoint of the respective author; this research uses the former.

The factor $\Delta_{\max}$ depicts the maximum fractional value of the total length of the overlapping sections in a path, compared to the length of the shorter path.

Use the variables defined earlier. Path k is *unfeasible* if a sub-sequence of links $\{a_{ik}, a_{(i+1)k}, \ldots, a_{jk}\}$ within path k can be found, for which the following two conditions are true.

$$\sum_{h=i}^{j} L(a_{hk}) > D[a_{ik}, a_{jk}] \tag{3.2}$$

$$\frac{\sum_{h=1}^{i} L(a_{hk}) + \sum_{h=j}^{m} L(a_{hk})}{\sum_{h=1}^{i} L(a_{hk}) + D[a_{ik}, a_{jk}] + \sum_{h=j}^{m} L(a_{hk})} > \Delta_{\max} \tag{3.3}$$

3.3.1 Setting constraint parameters

The path constraint parameters are to be chosen such that preferably all logical node sequences will be included in the generated choice set, though without generating an excessive amount of irrelevant routes. Several authors have proposed values; e.g. Schnabel and Lohse (1997); Fafieanie (2009), Table 3.1 shows their proposed values. There does not seem to be consensus, which does suggest the values to be dependent on the local situation.

In the Alkmaar network, values are chosen such that both directions along the city circular for distant OD-pairs (e.g. 216–220, 217–337) are included, though without generation of a large amount of similar routes through the city centre. The value used are a maximum detour of 1.75, a maximum overlap of 0.65 and a maximum number of routes of 7. These values are in between the values used by other authors.

Parameter	**Schnabel and Lohse (1997)**	**Fafieanie (2009)**	**Alkmaar values**
Detour Φ	1.25 – 1.4	1.6 – 2.1	1.75
Overlap $\Delta_{\max}$	0.5	0.75	0.65

Table 3.1: Constraint parameters

Example

Figure 3.6 shows four possible routes from node 215 to node 217. Using the individual link costs (described in Appendix A) the cost of (sub-)sequences of links can be calculated. Each route will be individually examined for constraints and the resulting feasibility of that route.

Blue route The blue route is the shortest path found, with a cost of 6.01.

Green route The green route is quite similar to the blue route. The total cost of the green route is 7.38, compared to 6.01 of the shortest route: the detour considering the entire route is 1.23. However, within the green route there is a sub-sequence of links that does not overlap with the shortest path between the start and end of that sub-sequence. The cost of this sub-sequence is 2.80, the shortest path between the start and end of this sub-sequence has a cost of 1.43, leading to a detour value of 1.96.

The link sequences that do overlap with the shorter blue route have a cost of 4.58. The shortest path using the same link sequences (i.e. the blue route) has a cost of 6.01, which leads to an overlap value of 0.76 (=4.58/6.01).

With a maximum detour factor of 1.96 and a maximum overlap of 0.76, both the detour and overlap values exceed the set limits, this route is therefore not deemed feasible and will not be included in the route set.

Red route The red route is the third shortest route, preceded by the blue and green route. Within the red route, there is one sub-sequence of links that does not overlap with the shortest path between the start and end of that sub-sequence; the cost of the sub-sequence is 6.40, the shortest path has between start and end of that sub-sequence has a cost of 4.72, the detour value is therefore 1.36.

The link sequence that does overlap with a shorter route has a cost of 1.29, the shortest route using the same link sequence is the blue route with a cost of 6.01; the overlap value is 0.21.

Both detour and overlap are within the set limits, therefore this route is deemed feasible. As it is the second shortest feasible route, it will be included in the route set.

Orange route The orange route is completely different to the other described routes; there is no overlap. The cost of the sequence is 14.04, with the shortest route at 6.01: the detour is 2.34.

With no overlap to other routes, only the detour value determines feasibility. As it exceeds the set limit, the route is not feasible and will not be included in the route set.

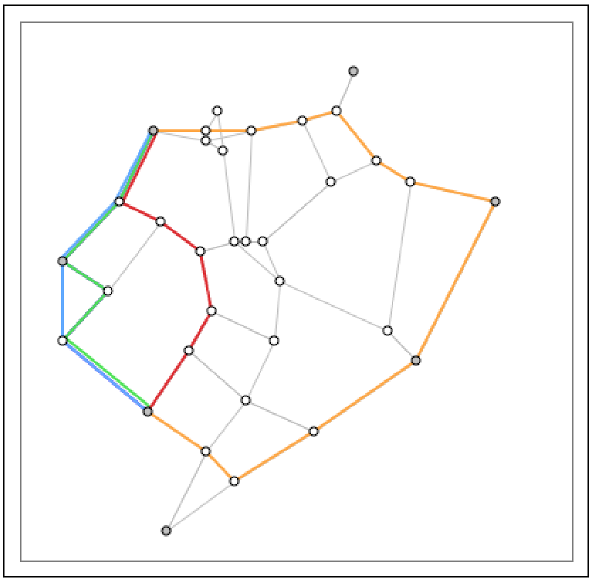

Figure 3.6: Example routes from node 215 to 217

3.4 Choice set generation

Within the data set it is not possible to determine exact origins and destinations of devices and the respective vehicles they are in, irrespective if these locations are within or outside the study area. Therefore only the detected locations can be used, with such a location either identified as an origin, intermediate point or destination of a trip. Considering the limited amount of detectors, a less detailed network omitting residential roads has been able to provide sufficient detail to generate reasonable routes between the OD-pairs. The use of Dijkstra's algorithm for shortest paths and the CKSP-method to identify a number of feasible paths using constraints possible has been possible with manageable calculation times due to the lesser detail in the network, by not having to incorporate irrelevant road sections.

Using a maximum detour value of 1.75, a maximum overlap value of 0.65 and a maximum amount of routes per OD-pair of 7, 126 routes are generated; Table 3.2 shows the number of paths found per OD-pair. Appendix B shows the generated routes in charts per OD-pair.

From–To	**215**	**216**	**217**	**218**	**219**	**220**	**337**
215	-	1	5	7	5	1	2
216	1	-	1	1	4	6	1
217	4	1	-	1	1	7	5
218	7	1	1	-	1	4	5
219	4	4	1	1	-	1	1
220	1	6	7	5	1	-	1
337	2	1	7	7	2	1	-

Table 3.2: Number of identified feasible paths per OD-pair

Chapter 4

Data analysis

4.1 Data selection

Route choice may vary from day to day and even within a day, due to e.g. road circumstances, morning and afternoon peaks and related delay. Route ratios therefore may change when travel times along alternative routes are changing. Variability in travel time therefore is a factor to take into account in route choice research.

4.1.1 Variability in travel time

Wardrop (1952) already noted that trip travel times may have very skew distributions; there tends to be a minimum travel time, though it is possible to have very long travel times as well. The concept of travel time variability therefore is not new. There are several reasons why travel time varies. Robinson (2005) has identified four major components:

Demand for travel is a direct result of peoples' activities and their choice of transport type. Several factors affect demand for road travel, e.g. type of activity, weather, activity time schedule.

Capacity is the maximum rate of traffic a network can manage. The basis for capacity is road layout, though several other factors can reduce capacity, e.g. incidents, parked vehicles, adverse weather.

Vehicle performance describes both the technical capacities of vehicles and vehicle drivers' behaviour in traffic. Factors of influence are e.g. vehicle performance, risk acceptance and human factors

Control by direct traffic management restrict traffic, e.g. by speed limits, traffic lights and crossings with railways or waterways.

Changes in any of these components will cause a change in travel time, and thus causes variability. Such changes are related to time. Several time scales are possible, as described by Robinson (2005), shown in Table 4.1.

Time scale	Description
Vehicle-to-vehicle	Variability in travel time of vehicles that travel over the same link in the same short period of time. It is a short-term travel time variability, and can be described as a random variability.
Period-to-period (within day)	Variability in travel time that occurs over the course of a single day.
Day-to-day	Variability in travel time that occurs between the same period of different days.
Season-to-season	Variability in travel time over a long period, e.g. between summer and winter.

Table 4.1: Disaggregation of travel time variability by time scale

4.1.2 Selection

The data made available for this research is captured in February 2011, i.e. from 1 February 0:00 to 28 February 23:59. As the data only describes one month, season-to-season variability is not present. Day-to-day and period-to-period travel time variability may be present, and should be taken into account.

To determine the variability over time, ideally all data will have to be considered. However, because of the shear amount of data, the necessary calculation time would be high. Not all routes would however be useful, e.g. routes through the city centre are less frequently used and therefore have low hourly volumes; the more often used routes using the city circular are therefore considered to be more accurate. Using a longer route reveals differences more easily.

The hourly average travel times per day of the week of route 337-220-219-218 are shown in Figure 4.1. On weekdays, the morning and afternoon peak do show as an increase in travel time. The day-time off-peak travel times (between 10:00 and 16:00) seems quite steady, whereas after the evening peak travel time gradually goes down to a low in the early morning, after which it picks up towards the morning peak. Other routes along the city circular show a similar behaviour, with slightly increased travel times in the morning and afternoon peaks. On Saturdays and Sundays, no morning and afternoon peaks are observed, with travel time increasing to around midday, and decreasing afterwards.

Day-to-day variability is clearly present, with large differences between weekdays, Saturdays and Sundays. Within weekdays, the period between 06:00 and 16:00 is less affected by day-to-day variability. Period-to-period variability within days is clearly visible as well, with fast changing conditions in peak times. In between peak times the travel time seems to be quite stable, i.e. with low variability. To determine accurate route ratios that depict the actual situation, data collection would have to be split up in short periods when conditions are changing rapidly. However, such short periods will contain less observations, and therefore would lead to inaccuracy as well. Determining route ratios in changing conditions therefore is not to be preferred.

To reduce the effect of day-to-day and period-to-period variability, though still retaining a substantial period, the period from 10:00 to 16:00 on Mondays to Fridays has been chosen for use. Splitting this data in equal parts, to allow for both analysis and validation, leads to one set comprising data from 10:00 to 13:00, and a second set comprising data from 13:00 to 16:00.

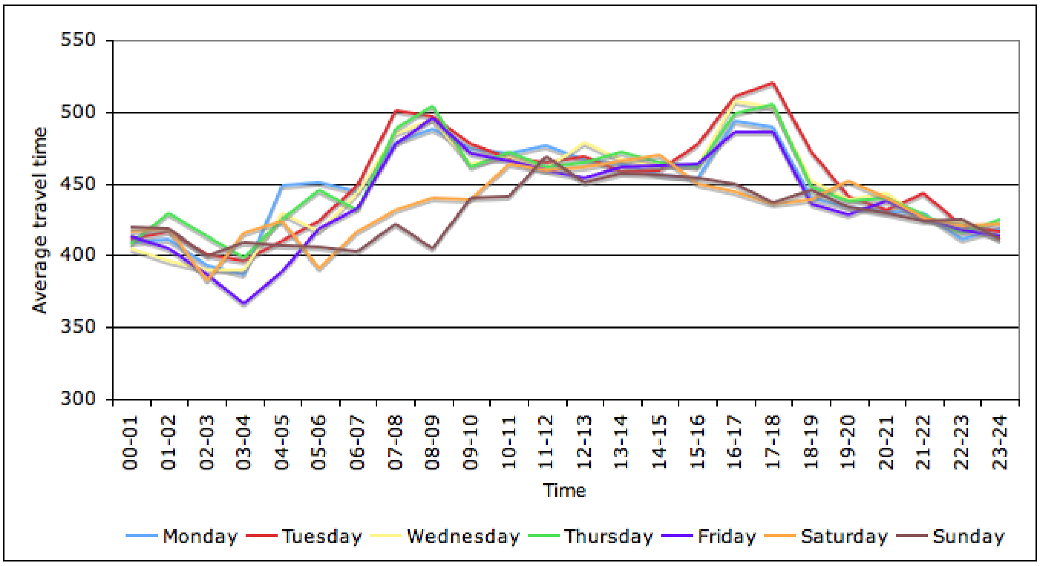

Figure 4.1: Average hourly travel time of route 337-220-219-218 by day of the week

4.2 Route inference

To infer routes from the available data, several steps will have to be taken.

Sequencing Ordering the data on passage times per observed device.

Travel time determination Determination of the average travel times between observed locations.

Trip identification Identification of different trips within the sequences.

Sequence splitting Split of sequences if multiple trips are identified.

4.2.1 Observations, node sequences and links

Each observation in the data sets comprises three variables: a location-ID, a time-stamp and a device-ID. Sorting the data on device-ID will lead to an ordered list of detection locations and times per device. Considering that the data-set comprises three hours per day for 20 days, the location orders are split per day, leading to *node sequences* per device per day. Each of the node sequences can be further chopped into sub-sequences of two successive nodes.

Travel time between observed locations

Using the time-stamp of each observation, the travel time between each of the subsequent observed locations, i.e. a sub-sequence, can be determined. These travel times have a minimum depending on sub-sequence length and the speeds travelled along the sub-sequence, and a maximum of three hours, i.e. 10800 seconds, within the data-set. Due to the possible spread towards high travel times, the travel time distribution is not necessarily normally distributed, which would prevent the use of only the mean, as well as using standard deviations to determine outliers. Having considered the travel time distributions, all node pairs show a large spread to high travel times. However, each node pair has a range of travel times that occurs most frequently. By creating travel time bins, a range can be identified, though this would not lead to a specific travel time value. Therefore, a method is devised identifying individual observations as outliers, and accepting the average of the non-excluded observations as the typical link travel time.

To identify outliers an algorithm is used, which calculates the average travel time using the all observations and excludes observations with a travel time exceeding a 1.5 times the calculated average travel time. This process is repeated with remaining observations until no more outliers are identified, i.e. no more observations are excluded; the average travel time of the remaining observations is then said to be the typical sub-sequence travel time. The result of using this algorithm is a travel time distribution that resembles a normal distribution, with a maximum value distance of two to three standard deviations.

Two examples illustrate the process; the first shows the high volume sub-sequence 215-216 with nearly 18000 observations, the second shows low volume sub-sequence 215-218 with 278 observations. Even though these routes are very different, the outlier identification algorithm does provide good results in both cases.

Example 1 The travel times between nodes 215 and 216 range from 49 to 10140 seconds, though 90 per cent of the observations is below 600 seconds. The upper panel of Figure 4.2 shows the cumulative distribution up to 600 seconds. The average value of all observations is 472 seconds with a standard deviation of 1117 seconds. The travel time distribution, depicted in the lower panel of Figure 4.2, shows that the typical travel time along this link would likely be in the region of 140 to 160 seconds. Using the algorithm that excludes all observations with travel times exceeding 1.5 times the average, the typical travel time is said to be 160 seconds; observations over 240 seconds are identified as outlier, 14 per cent of the observations has been identified as outlier, the standard deviation of the remaining observations is 27.6 seconds.

Example 2 The travel times between nodes 215 and 218 range from 300 to 9220 seconds, 25 per cent of the observations has a travel time less than 900 seconds. The upper panel of Figure 4.3 shows the cumulative distribution, up to 1500 seconds. The average value of all observations is 2603 seconds with a standard deviation of 2002 seconds. If an outlier identification algorithm using the average plus-minus three standard deviations would be used, only three observations would be identified as outliers; the remaining average would then still be very high and therefore unlikely to be the typical travel time. The travel time distribution, the lower panel of Figure 4.3, shows the typical travel time to be between 350 and 450 seconds; and using the algorithm a value of 418 seconds is found. Observations over 626 seconds are identified as outlier, excluding 81 per cent of the observations. The standard deviation of the remaining observations is 56.5 seconds.

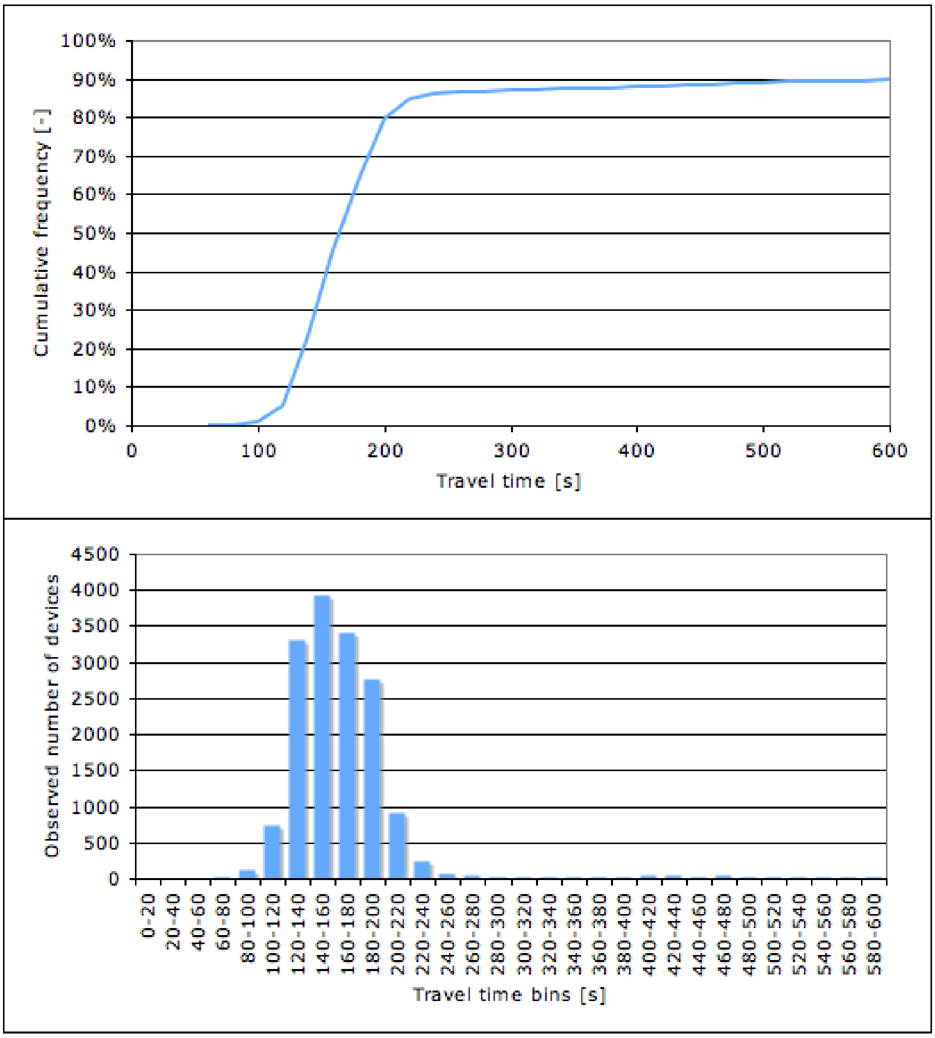

Figure 4.2: Travel time observations between nodes 215 and 216, cumulative (upper panel) and in 20 second bins (lower panel)

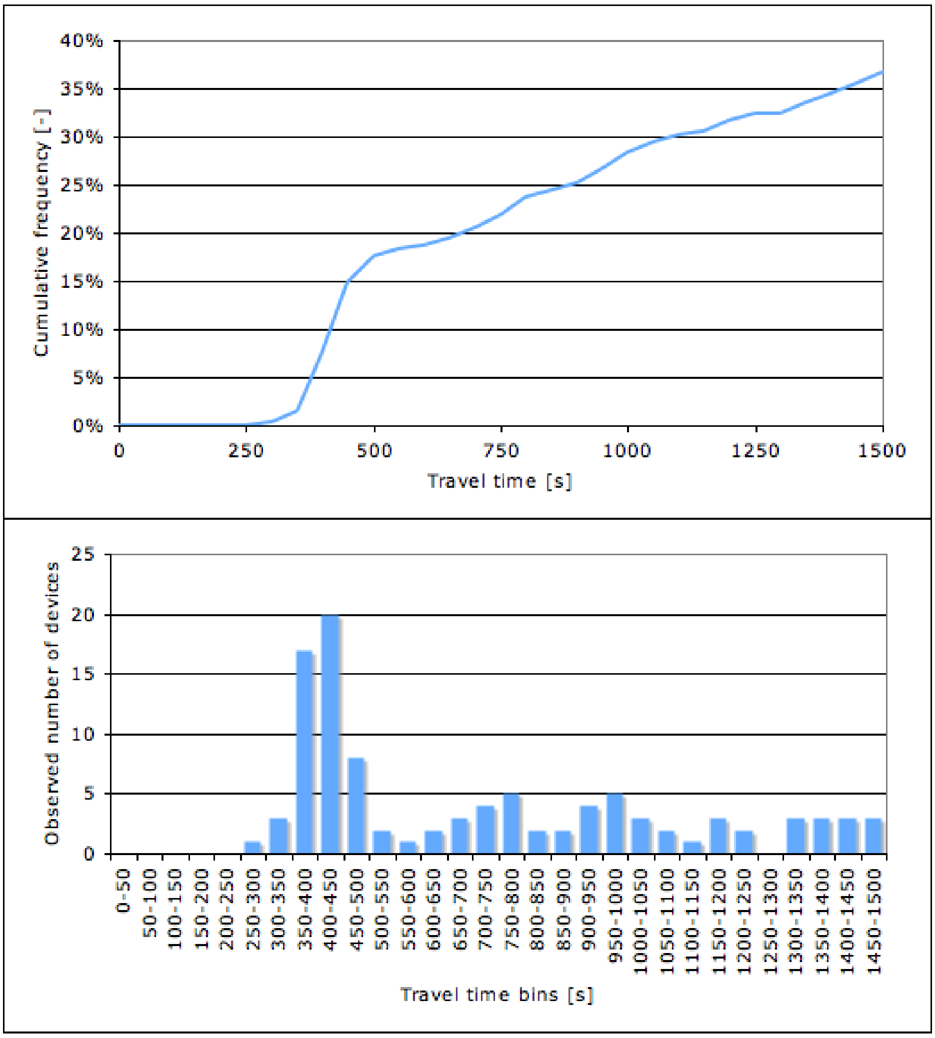

Figure 4.3: Travel time observations between nodes 215 and 218, cumulative (upper panel) and in 50 second bins (lower panel)

From–To	**215**	**216**	**217**	**218**	**219**	**220**	**337**
215	-	160	226	418	1134*	333	161
216	145	-	66	250	378	518	309
217	259	114	-	174	300	1001*	424
218	447	277	167	-	151	375	496
219	818	415	298	136	-	176	296
220	309	493	673	337	196	-	137
337	201	360	434	501	331	137	-

Table 4.2: Average link travel times in seconds

Table 4.2 shows the typical travel times of all sub-sequences, as determined by the devised algorithm. These times have been checked for coincidence with the most frequent occurring travel time range, by means of comparison with the frequency distributions. The distributions of two links, i.e. 215-219 and 217-220, do not show a clear most frequent occurring travel time range; the values provided by the algorithm may therefore not necessarily indicate the typical travel time. This conclusion is strengthened by comparing the concerned links to those in opposite direction: the travel time value for link 215-219 is 1134 seconds, whereas link 219-215's is 818 seconds; link 217-220 is calculated to be 1001 seconds, link 220-217 is 673 seconds. It is expected that the values of opposite links should be more-or-less similar, as is the case with other links in the network. The values of links 215-219 and 217-220 are therefore considered not fit for use; the value of the opposite direction will be used instead. Table 4.2 shows the original values, demarcated with an asterisk.

Influence of route on travel time

Travel times of sub-sequences may be dependent on the route used to the specific sub-sequence, e.g. if there is a significant difference in type of road or experienced delay. Within the Alkmaar network this is specifically the case with routes along the city circular.

Figure 4.4 illustrates this issue; the example network shows a major route and feeder roads. Traffic entering the major route is likely to have to yield to traffic along the major route, e.g. due to traffic signals. The travel time of traffic along the major route, i.e. between detection point A and C, would therefore be less than the travel time of traffic entering the major road, i.e. from point B to C.

The travel times of all sub-sequences along the city circular, incorporating their source, is shown in Table 4.3; the difference to the average sub-sequence travel time has been found not to be statistically significant for any sub-sequence: the average value is therefore used in the research.

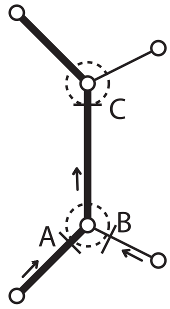

Figure 4.4: Example network with a major route (thick links), feeder roads (thin links), two detectors with a certain range (dashed circles) and some detection points (A, B and C)

Link	From	Link travel time	Difference to average
215–216	337	160	0
	other	160	0
215–337	216	161	0
	other	161	0
216–215	217	141	-4
	other	158	+13
216–217	215	64	-2
	other	78	+12
217–216	218,219	104	-10
	other	120	+6
217–218	216	172	-2
	other	179	+5
217–219	216	295	-5
	other	302	+2
219–217	220	297	-1
	other	298	0
219–218	220	136	0
	other	136	0
219–220	217,218	176	0
	other	176	0
220–219	337	191	-5
	other	202	+6
220–337	219	136	-1
	other	138	+1

Table 4.3: Link travel times in seconds incorporating source

4.2.2 Trips within node sequences

A sequence of nodes of a device may consist of more than one trip, i.e. a trip may be ended and a new one started between two successive observations within an observed sequence. Successive observations can be either a double observation at the same location, or observations at different locations; each requires a different method to decide if successive observations are either within a trip, or mark the end of a trip and the start of a new one. If the latter is the case, the observed node sequence should be split into separate sequences.

Successive observations at same location

Multiple successive observations of a device at the same node are possible and are caused by a device exiting the range of a detector and entering it again. This can be accidental within a trip due to e.g. road layout and can include e.g. signal waiting time, or on purpose, e.g. because of an intermediate stop, hence denoting an end of a trip and start of a new trip. The time difference between the observations is used to determine whether the former or the latter is the case.

About thirteen per cent of all 257,371 successive observations is found to be a double detection at the same location. Figure 4.5 shows the relative amount of successive observations at the same location for time difference categories up to 300 seconds, though the maximum value found nears 10800 seconds – the maximum time in the data set. If successive observations are at the same location, they are most likely to have a time difference of 1 to 30 seconds. Any double entry with a time difference lower than 30 seconds will therefore be considered as belonging to one trip; for route inference only one entry at each location is necessary, the first entry is used. A double entry with a time difference over 30 seconds leads to the first entry being the end of one trip and the second entry to be the start of a new trip.

Successive observations at different locations

There is a distinct possibility that a stop is made in between two successive observations at different locations, especially with sequences transversing the city centre. A stop will lengthen the travel time between the observed nodes; irregularly long travel times between nodes can therefore said to be caused by a stop, indicating different trips.

The method to determine the link mean travel time, as explained in Section 4.2.1, includes observations with travel times within 1.5 times the typical travel time value; these observations are therefore considered to belong to an uninterrupted travel. Hence, observations with longer link travel times are considered interrupted and therefore assumed to indicate trip-ends.

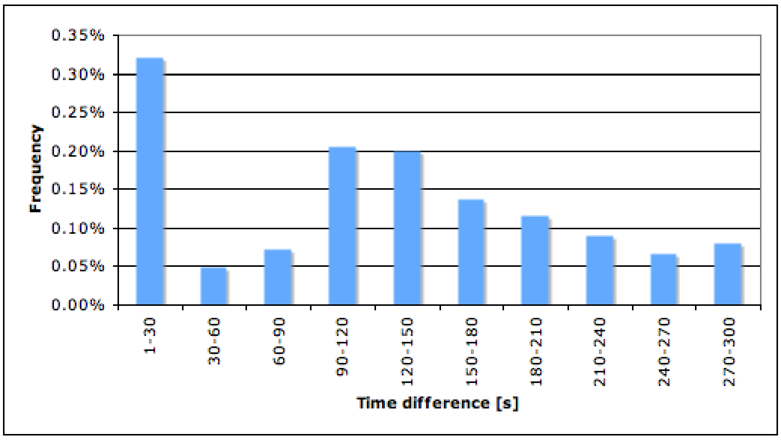

Figure 4.5: Distribution of time difference between successive observations at the same location

It may well be that two trips with a very short stop in between (e.g. a package delivery) is identified as a single trip, because the stop has not lengthened the travel time between observed locations beyond 150 per cent of the calculated typical travel time. Such a short stop is likely to take place within the city, but not on the city circular. Therefore, the number of observations of routes through the city centre, and thereby the route ratio, may be affected; the route ratio of routes through the city centre may therefore be overestimated. This issue cannot be solved with only the available data; a more detailed study would be necessary, which was out of scope for this project. This kind of observation error therefore has been accepted within this project.

4.3 Comparison of observed sequences and generated routes

The previous section has explained the method to determine trips within each device's observed node sequence. Considering all trips in the data set, 320 node sequences are identified. Not all sequences are however fit for use. Comparison with the generated route set identifies routes that can be used.

Observed sequences may not be in the choice set because of several reasons. This research has identified four categories.

Excessive detour and/or overlap The observed sequence can be explained by a path, but has a detour and/or overlap in excess of the set variables. This is expected, as the choice set parameters are chosen such that not all routes are included (see Section 3.3.1); the non-inclusion of some paths, though corresponding with observed routes, is therefore possible.

Circular route Circular sequences are mainly caused by multiple trips that are observed as one, i.e. the time gap in between the trips is smaller than the set value (see Section 4.2.2). This is an observational issue related to the detail of information available; more detailed information, e.g. directional information, may reduce this problem.

Unexpected intermediate nodes The VBM detectors can detect devices from some distance; devices do not necessarily have to pass the detector to be detected, nearing a detector to within detection range will already lead to registration. A device travelling along a route that does not pass a certain detector, though does near it, may therefore be detected. Note that the detection range depends on the device that has been detected. One such situation exists in the Alkmaar network: routes including the link 217-219 or vice versa go near VBM218. Some devices using 217-219 are detected by VBM218, whilst others are not, though they are travelling the same route. This is a clear observation error.

Without expected nodes The VBM detectors either register a complete MAC-address, or not at all: if the Bluetooth signal is blocked in some way, a detector does not register a device. A resulting sequence of nodes may then become illogical. E.g., sequence 337-219-218 is illogical, because it can only be explained by a relevant route that also incorporates node 220, leading to route 337-220-219-218. If an illogical sequence can be amended by adding one node, and this results to be equal to a relevant route, it is said to belong to this category. Again, this is an observation error.

Table 4.4 shows the amount of sequences found in each category as well as the amount of correctly mapped sequences. Sequences that are not included because of the third or fourth reason can still be mapped to a generated route, by assuming they have passed an expected node, or have not passed an unexpected node. The number of useable observed sequences therefore is 127, representing 86 unique sequences.

The generated choice set contains 126 routes; 124 generated routes can be mapped to the 86 unique sequences. 64 observed sequences are explained by a single generated route, the 22 remaining sequences therefore each have multiple generated routes. Table 4.5 shows the amounts concerned. The number of observations per unique sequence can be found in Appendix C. The sequences not mapped to any of the generated routes have been discarded, that is over 60 per cent; the number of observations within those discarded sequences is however very low, being 0.7 per cent of all observations. Therefore, 99.3 per cent of all observations has been useable.

Sequences …	Number of sequences
with excessive detour and/or overlap	146
with a circular part	47
with unexpected intermediate nodes	14
without expected nodes	27
correctly mapped	86

Table 4.4: Number of observed sequences per route category

Although correction for detection errors is required, Bluetooth detectors are very well useable for route choice research.

Observed sequence with …	Number of sequences
1 generated route	64
2 generated routes	12
3 generated routes	6
4 generated routes	2
5 generated routes	2

Table 4.5: Generated routes mapped to observed sequences

Chapter 5

Route choice modelling

5.1 Route choice and travel time

To describe the probability of use of a route among a set of alternative routes, a *logit model* is used. Such a model determines probability based on the route utility, U. The utility of a route can be described in terms of travel time, TT; considering that utility decreases when travel times increase, travel time can be seen as a negative utility. Including a scale parameter, θ, and a random error term, ϵ_i, Equation 5.1 shows this process. Getting from individual route utility to a route probability, P, involves assessment of all routes in the set; a logit model is used to do so. Equation 5.2 shows the derived logit model.

$$U_i = -\theta TT_i + \epsilon_i \tag{5.1}$$

$$P_i = \frac{e^{U_i}}{\sum\limits_{j=1}^{m} e^{U_j}} = \frac{e^{-\theta TT_i}}{\sum\limits_{j=1}^{m} e^{-\theta TT_j}} \tag{5.2}$$

For all routes within an OD-pair, the travel time of the fastest route, TT_1, can be deducted without affecting the route probability, i.e. using the travel time difference of the respective routes; Equation 5.3 shows this derivation.

$$P_i = \frac{e^{-\theta TT_i}}{\sum\limits_{j=1}^{m} e^{-\theta TT_j}} = \frac{e^{-\theta TT_i} e^{\theta TT_1}}{\sum\limits_{j=1}^{m} e^{-\theta TT_j} e^{\theta TT_1}} = \frac{e^{-\theta (TT_i - TT_1)}}{\sum\limits_{j=1}^{m} e^{-\theta (TT_j - TT_1)}} \tag{5.3}$$

Scaling route probabilities to the fastest route in a set requires translation to route ratios, shown in Equation 5.4 and 5.5, with n_i the volume on route i and n_1 the volume on the fastest route. Note that the fastest route will always evaluate to $n_i/n_1 = 1$, as $n_i = n_1$ and $TT_i = TT_1$.

$$P_i = \frac{n_i/n_1}{\sum_{j=1}^{m} n_j/n_1} \tag{5.4}$$

$$\frac{n_i}{n_1} = e^{-\theta(TT_i - TT_1)} \tag{5.5}$$

However, evidence from e.g. Thomas and Tutert (2009, 2010), suggests that route utility might not be linearly related to route travel times, or travel time difference. Equation 5.6 describes different forms that the relationship might have.

$$U_i = f(TT_i - TT_1) \Rightarrow \begin{cases} \text{(a) } (TT_i - TT_1)^\alpha \\ \text{(b) } (TT_i - TT_1 - \beta)^\alpha \\ \text{(c) } (TT_i - TT_1)^\alpha - \beta^\alpha \end{cases} \tag{5.6}$$

5.2 Route choice evidence

For all observed sequences with a generated route equivalent, i.e. 86 routes, the number of observations has been determined. The travel time of each sequence is calculated through a summation of average subsequence travel times, as shown in Table 4.2 on page 49.

Some observed sequences can be described by multiple generated routes; all generated routes are assumed to be feasible and independent. With only the total number of the sequence known, no route ratio can be determined for any of the generated routes, nor can an accurate travel time for any of the generated routes. The found travel time for the observed sequence can be assumed to be the average of all feasible routes within the sequence. Dividing the observed sequence volume by the number of equivalent generated routes would lead to an average number of observations per generated route. In combination with the observed average travel time such routes can be included in research; as they are constructed from one observed sequence, the average values should be used once, irrespective of the number of equivalent generated routes.

The resulting data has been used to determine a relationship between travel time difference and route ratio, assuming that route choice behaviour is equal within the dataset. To this effect, the fastest route within an OD-pair is considered to be the datum of that specific OD-pair, with its volume to be rendered as n_1 and its travel time as TT_1.

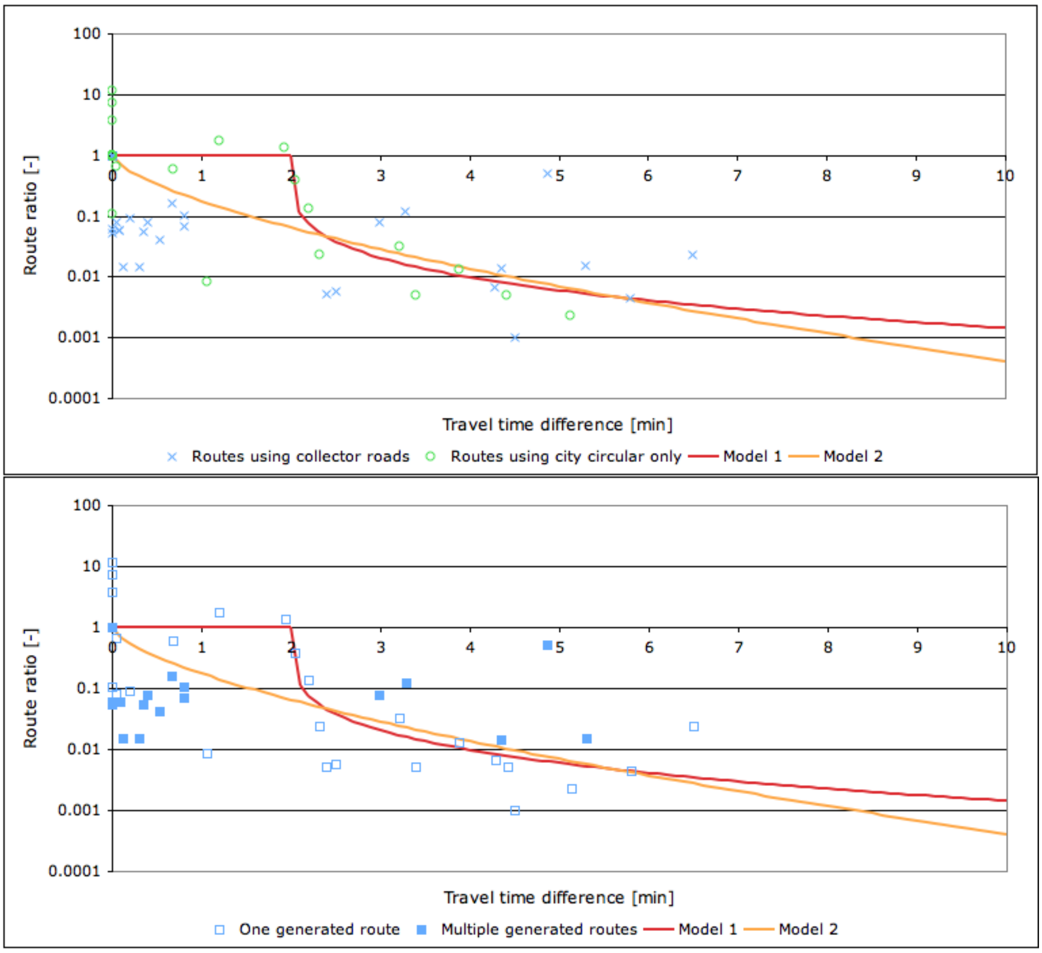

Figure 5.1: Relationship between route fraction and travel time difference

Figure 5.1 shows the observations of the 86 routes; the top panel shows a differentiation between routes using only the city circular and other routes, the lower panel shows a differentiation between the number of generated routes per observed route. With low travel time differences, approximately up to a minute, the data does not show a clear relationship; some large values of n_i/n_1 are found, up to 12, as well as some low values, from 0.008. It does show that some routes over the city circular are most used even though not having the shortest travel time, i.e. having a route ratio over 1, as well as that several routes incorporating collector roads, mostly with multiple associated generated routes, are far less frequently used even though the travel time difference is small. This does suggest that other factors are of influence. With travel time differences from around 1 minute some correlation seems to exist.

5.2.1 Model fits

Two distinct models, including error terms as described in Section 5.1, have been fitted to the data using a least squares method, on observations from 1 minute. To reduce the effect of non-linearity, the logarithmic values of observed and predicted values are used. Equation 5.7 defines the residual sum of squares, with y_i the observed value and $f(x_i)$ the model prediction of y_i. The resulting value of S is minimised by alternating the model's variables.

$$S = \sum_i r_i^2$$
$$r_i = \log y_i - \log f(x_i) \tag{5.7}$$

Model 1

The first model is based on the utility function as shown in Equation 5.6 (b). It assumes a route ratio of 1 up to 2 minutes, i.e. $\beta = 2.0$, corresponding to the observations in that area, after which a steep decline is seen. The least sum of squared residuals is found for $\theta = 3.9$ and $\alpha = 0.25$. The model is shown in Equation 5.8, with a visual representation in Figure 5.1 as the red line. The R-squared value considering all observations, using the logarithmic values, is 0.60.

$$\frac{n_i}{n_1} = \begin{cases} e^{-\theta(TT_i - TT_1 - \beta)^\alpha} & \text{if } TT_i - TT_1 > \beta \\ 1 & \text{otherwise} \end{cases} \tag{5.8}$$

Model 2

The second model does not incorporate an offset, i.e. $\beta = 0$. Equation 5.9 describes this model. The least sum of squared residuals is found for $\theta = 1.75$ and $\alpha = 0.65$. This model is visually represented in Figure 5.1 as the orange line. The R-squared value considering all observations, using the logarithmic values, is 0.68.

$$\frac{n_i}{n_1} = e^{-\theta(TT_i - TT_1)^\alpha} \tag{5.9}$$

5.3 Route attribute regression

Considering only experienced travel times does lead to a model with an amount of error. The probable cause for this error is that travellers do not know the exact travel time, and therefore incorporate other route attributes in their route choice. Identification and inclusion of these attributes and their effects should increase model accuracy.

5.3.1 Attribute regression

Using the observed route ratios, an expected travel time difference of a route to the most used route, ΔTT_i^E, can be calculated using the inverse function of the found relationships from Section 5.2. Equation 5.10 shows this process, the parameter values are equal to those mentioned in Section 5.2.

$$f^{-1}(TT_i - TT_1) = \left(\frac{\ln(N_i/N_1)}{-\theta}\right)^{1/\alpha} + \beta = \Delta TT_i^E \tag{5.10}$$

Assuming that travel times, and therefore travel time differences, are based on route attributes other than experienced travel time, travel time difference can be described by Equation 5.11, with $X_{k,i}$ the value of route variable k within route i, and γ_k a scale parameter of the associated variable.

$$\Delta TT_i^E \approx \sum_k \gamma_k (X_{k,i} - X_{k,1}) \tag{5.11}$$

Regression analysis of the set of independent variables using the expected travel time difference as a dependent value, calculated the extend of each predictor variable's influence, i.e. the parameters γ_k, by a least squares method. The significance of each variable is determined using a t-test; a variable is considered not to be significant if the one-tail p-value exceeds 0.05. The least significant variable, i.e. the variable with the highest p-value, is removed from the set, and the regression of the remaining parameters is recalculated. This process is repeated until all variables are considered to be significant.

5.3.2 Travel time attributes

Based on the available information of the Alkmaar network, i.e. distances and road observations, several attributes are identified for regression analysis.

Distance in road type Each road type in the network has specific attributes like maximum speed and hierarchy. Furthermore, travellers may evaluate each road type differently, i.e. they may prefer certain road types. All have an influence on the expected route travel times. The road types are assumed to be independently reviewed by travellers, and therefore are included in the regression analysis as separate attributes. The variable used is the minimum travel time, consisting of maximum allowed speed, v, and length in road type, L.

Direction changes When not travelling straight on at any intersection, some extend of delay will occur, e.g due to slowing down for turning or waiting for oncoming traffic. Each direction change will increase the total delay of a route. The number of direction changes, N^D, in a route may therefore be considered by travellers.

Traffic signals Traffic signals are another cause of delay. Both the waiting time and the shockwaves associated with it increase travel time. Assuming that all traffic signals operate independently, and therefore each signal results in an additional average delay, the number of signals, N^S, on a route may be an attribute in route choice.

Route signage As it is more easy to follow a signed route, it is likely that such routes are preferred. In terms of travel time this would lead to an assumed shorter route. Each destination node has been mapped to one or more city names, shown in Table 5.1. When an route is signed out at all direction changes using the same city name, the entire route is said to be signed out. It is described by a dummy variable, X^{sign}, which evaluates to 1 when a route is signed out and 0 otherwise.

Bridge In the Alkmaar network a movable bridge is positioned in the city circular, between node 217 and node 218. The uncertainty of a possible bridge opening, and the resulting uncertainty in travel time, may lead to a deterrence to use that stretch of road. It is described by a dummy variable, X^{bridge}, which evaluates to 1 when a bridge is present and 0 otherwise.

Table 5.2 shows the variables involved in each attribute, as used in the regression equation shown in Equation 5.12.

Final node	Associated destination
337	Haarlem, Amsterdam
215	Heiloo
216	Egmond
217	Den Helder
218	Schagen
219	Heerhugowaard
220	Hoorn, Purmerend

Table 5.1: Associated cities per destination node

k	Attribute	Variable
1	Minimum travel time over dual carriageways	$\upsilon_A^{-1}(L_{A,i} - L_{A,1})$
2	Minimum travel time over arterial roads	$\upsilon_B^{-1}(L_{B,i} - L_{B,1})$
3	Minimum travel time over collector roads	$\upsilon_C^{-1}(L_{C,i} - L_{C,1})$
4	Number of direction changes	$N_i^D - N_1^D$
5	Number of traffic signals	$N_i^S - N_1^S$
6	Presence of route signage	$X_i^{\text{sign}} - X_1^{\text{sign}}$
7	Presence of bridge	$X_i^{\text{bridge}} - X_1^{\text{bridge}}$

Table 5.2: Regression variables

Regression variable	Model 1 parameter	Model 2 parameter
Travel time factor of dual carriageways	1.78 [-]	1.73 [-]
Travel time factor of arterial roads	1.05 [-]	1.15 [-]
Travel time factor of collector roads	1.33 [-]	1.32 [-]
Number of direction changes	0.71 [min/change]	0.58 [min/change]
Presence of route signage	1.76 [min]	1.42 [min]

Table 5.3: Significant variables and respective regression parameter values

$$\begin{aligned}\Delta TT_i^E \approx \gamma_1 \upsilon_A^{-1}(L_{A,i} - L_{A,1}) + \gamma_2 \upsilon_B^{-1}(L_{B,i} - L_{B,1}) \\ + \gamma_3 \upsilon_C^{-1}(L_{C,i} - L_{C,1}) + \gamma_4(N_i^D - N_1^D) + \gamma_5(N_i^S - N_1^S) \\ + \gamma_6(X_i^{\text{sign}} - X_1^{\text{sign}}) + \gamma_7(X_i^{\text{bridge}} - X_1^{\text{bridge}})\end{aligned} \tag{5.12}$$

5.3.3 Regression results

The regression analysis has revealed five significant values for either model: the distance over each of the road types, the number of direction changes and route signage. Table 5.3 shows the respective regression parameters. The results of the individual analyses are summed up in Appendix D.

Using the found parameter values, each route's predicted value has been calculated and compared to the value expected by the model. Furthermore, a comparison is made between the observed travel time values and the model expectations. Figure 5.2 shows the values of the first model; the coefficient of determination – or R^2 – of the observed values is 0.64, the R^2 of the predicted values is 0.76. Figure 5.3 shows the values of the second model, with an R^2 for observed values of 0.67 and an R^2 for expected values of 0.77. The predictions based on significant attributes therefore do show to lead to better results than using only observed travel time, with the second model performing slightly better.

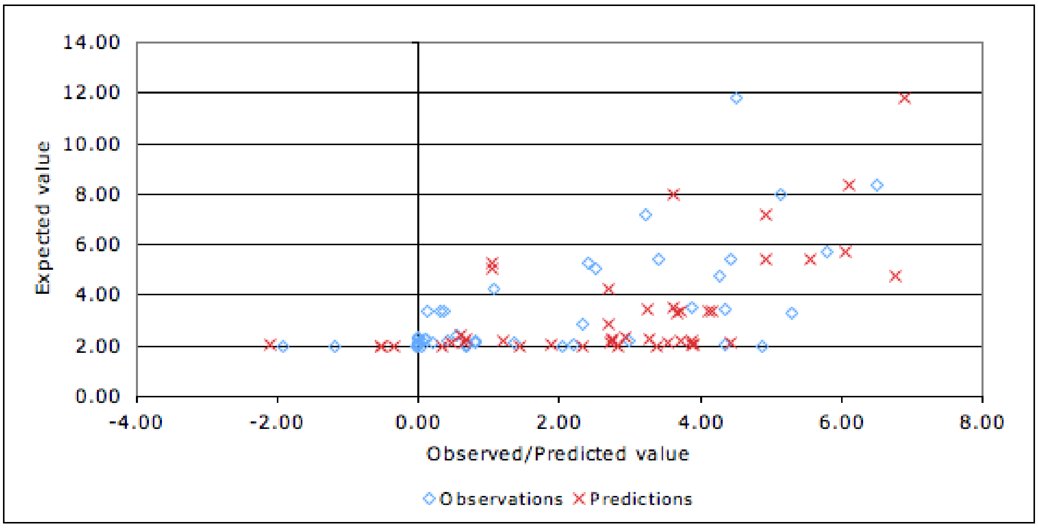

Figure 5.2: Model 1 comparison

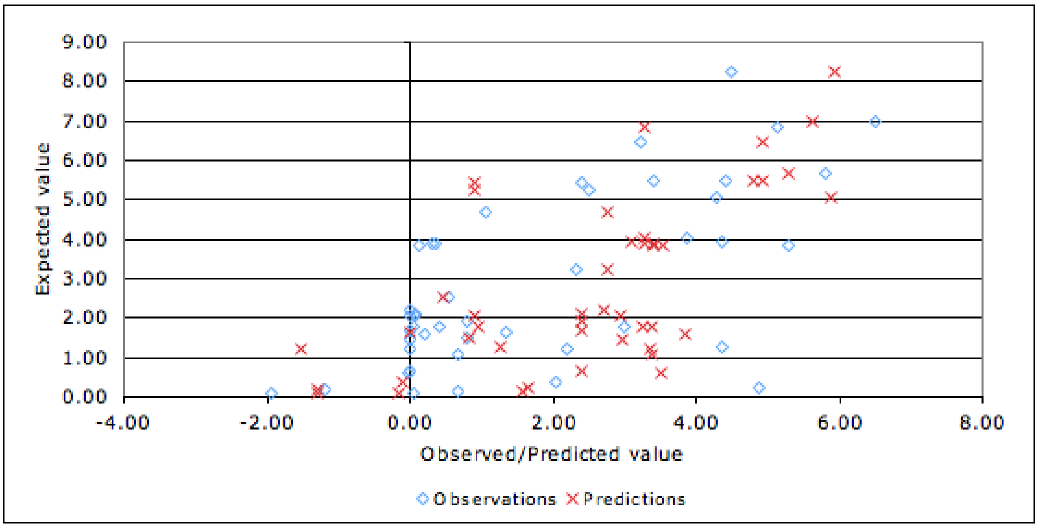

Figure 5.3: Model 2 comparison

5.4 Modelling route choice

5.4.1 Route choice model

The results in the previous section have shown that using a combination of easy-to-determine route attributes, i.e. generalised travel time, outperforms a model using only observed travel time. Therefore, the resulting route choice model should include generalised travel time difference, as described in Equation 5.13 with ΔTT_i^G the generalised travel time difference between route i and the optimal route (i.e. the route with the least generalised travel time).

$$\begin{aligned}\Delta TT_i^G = &\frac{1.73 \cdot 60}{80}(L_{A,i} - L_{A,1}) + \frac{1.15 \cdot 60}{70}(L_{B,i} - L_{B,1}) \\ &+ \frac{1.32 \cdot 60}{50}(L_{C,i} - L_{C,1}) + 0.58(N_i^D - N_1^D) + 1.42(X_i^{\text{sign}} - X_1^{\text{sign}})\end{aligned} \tag{5.13}$$

Route ratios within an OD-pair are to be determined by Equation 5.14, including the generalised travel time difference value.

$$\frac{n_i}{n_1} = e^{-1.75(\Delta TT_i^G)^{0.65}} \tag{5.14}$$

Identification of the optimal route, and its attribute values, is necessary to determine route fractions.

Optimal route identification

The optimal route can be identified using the found significant route attributes and corresponding parameters, though using the actual value of a route instead of the difference between routes. The route with the least calculated travel time within a set of routes in an OD-pair is said to be the optimal route, i.e. the route with the lowest presumed travel time — it may not necessarily be the fastest route considering actual travel time. Equation 5.15 defines the optimal route within a set of routes S. With a route i identified as the fastest route, the fastest route's variables are identified as well, i.e. $L_{A,1}$, $L_{B,1}$, $L_{C,1}$, N_1^D and X_1^{sign}.

$$\exists\, i \in S : TT_1 = \min(1.73L_{A,i} + 1.15L_{B,i} + 1.32L_{C,i} + 0.58N_i^D + 1.42X_i^{\text{sign}}) \tag{5.15}$$

5.5 Model testing and evaluation

Testing a model involves a comparison of model results to an independent set of data. For each of the 86 identified feasible routes a comparison is made between modelled and observed route ratios. The difference between the observed and the modelled route ratio is a measure to determine what model would be best to use.

5.5.1 Comparison of observed and generalised travel time differences

As explained earlier, each route can be described by an observed travel time difference and by a generalised travel time difference. These values are then used to calculate the expected route ratio; this value would be on the red line in Figure 5.4. The modelled route ratios do not fully match the observed route ratios; in Figure 5.4 each route's travel time difference is shown against the observed route ratio of the respective route – generalised travel time difference in blue crosses and actual travel time difference in green circles. Using generalised travel time difference, the residuals between observations and model are less than when using observed travel time differences; the sum of squared residuals are respectively 2.03 and 10.79.

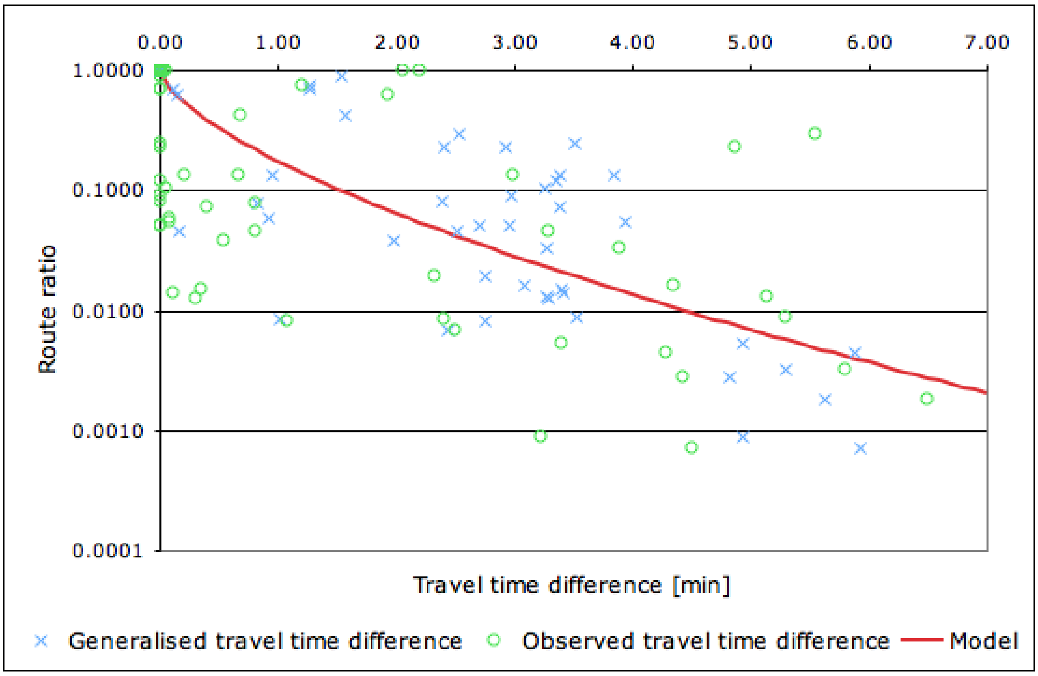

Figure 5.4: Route ratio model as researched versus observations

5.5.2 Comparison to other models

To determine whether route choice modelling would be improved by using the proposed model a comparison is made to other models as well. Two other models have been used: the multinomial logit model, and the model proposed by Thomas and Tutert (2009, 2010). Both models have been scaled to the data, the travel time attributes and corresponding parameters have been redetermined as well.

Multinomial logit

The regression of a multinomial logit model has lead to inclusion of five route attributes, i.e. the three road type attributes, route signage and bridge presence. The resulting route ratio model is shown in Equation 5.16, Figure 5.5 shows the model versus observations. The sum of squared residuals is 1.97.

$$\frac{n_i}{n_1} = e^{-(TT_i - TT_1)} \tag{5.16}$$

with:

$$\begin{aligned} TT_i - TT_1 = {} & \frac{1.89 \cdot 60}{80}(L_{A,i} - L_{A,1}) + \frac{1.17 \cdot 60}{70}(L_{B,i} - L_{B,1}) \\ & + \frac{1.82 \cdot 60}{50}(L_{C,i} - L_{C,1}) + 1.87(X_i^{\text{sign}} - X_1^{\text{sign}}) + 1.09(X_i^{\text{bridge}} - X_1^{\text{bridge}}) \end{aligned}$$

Logit model by Thomas and Tutert

The regression of the model proposed by Thomas and Tutert (2009, 2010), scaled to the Alkmaar data using $\theta = 2.0$, has lead to inclusion of five route attributes, i.e. the three road type attributes, the number of direction changes and route signage. The resulting route ratio model is shown in Equation 5.17, Figure 5.6 shows the model versus observations. The sum of squared residuals is found to be 1.86.

$$\frac{n_i}{n_1} = \begin{cases} e^{-2[(TT_i - TT_1)^{0.7} - (0.5)^{0.7}]} & \text{if } TT_i - TT_1 > 0.5 \\ 1 & \text{otherwise} \end{cases} \tag{5.17}$$

with:

$$\begin{aligned} TT_i - TT_1 = {} & \frac{1.65 \cdot 60}{80}(L_{A,i} - L_{A,1}) + \frac{1.1 \cdot 60}{70}(L_{B,i} - L_{B,1}) \\ & + \frac{1.42 \cdot 60}{50}(L_{C,i} - L_{C,1}) + 0.52(N_i^D - N_1^D) + 1.43(X_i^{\text{sign}} - X_1^{\text{sign}}) \end{aligned}$$

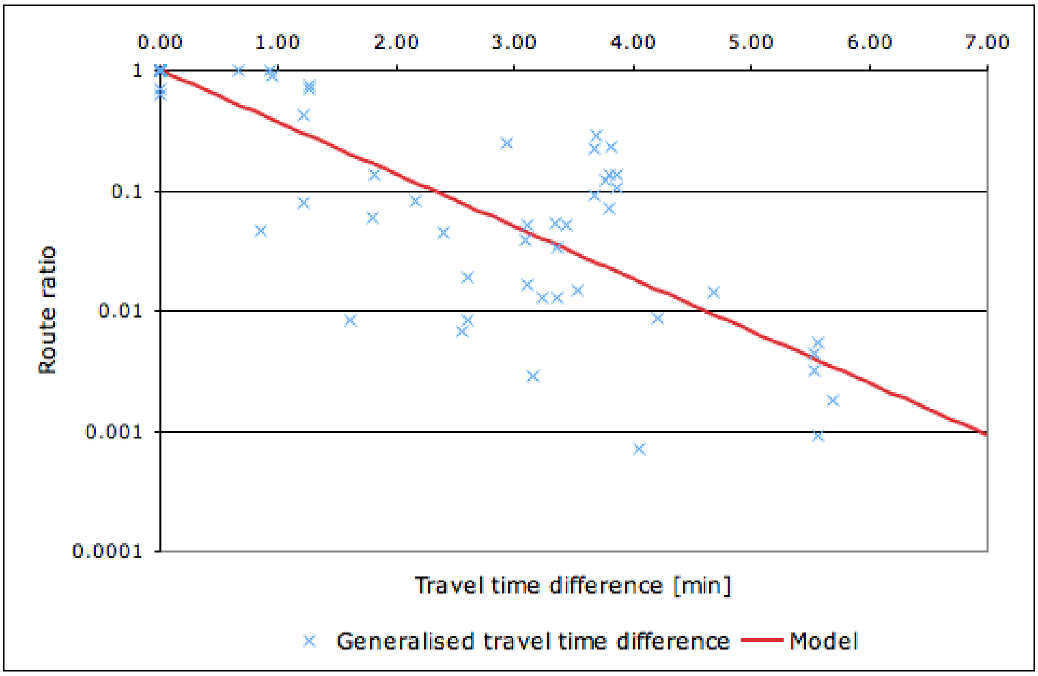

Figure 5.5: Multinomial route ratio model versus observations

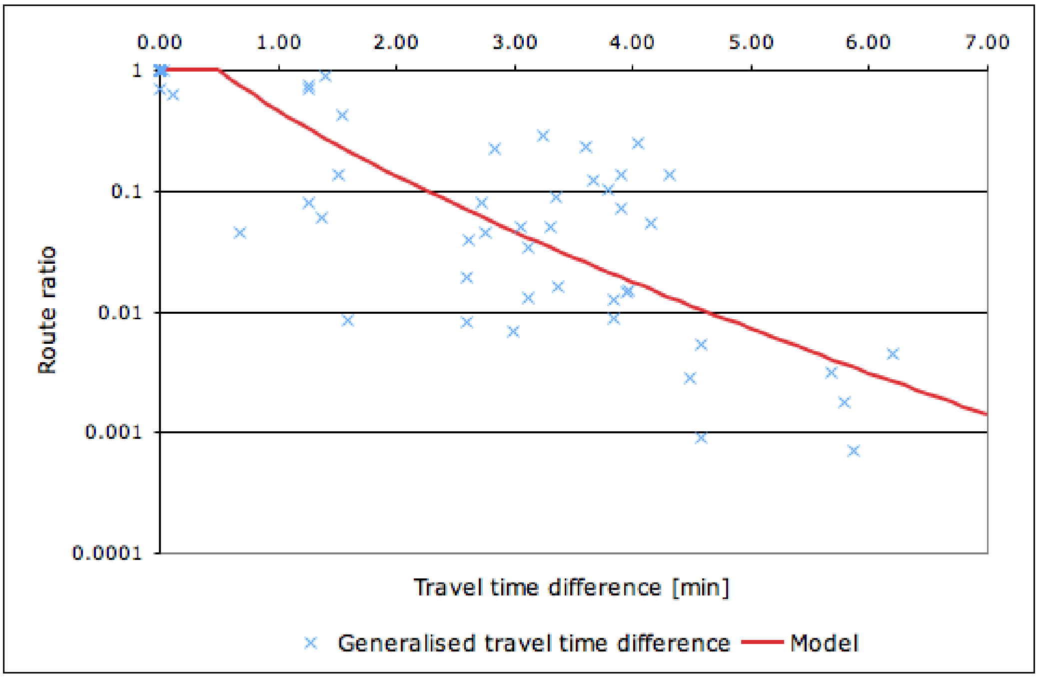

Figure 5.6: Route ratio model based on Thomas and Tutert (2009, 2010) versus observations

The results from the comparison to actual travel time clearly show that using a travel time model does improve model accuracy. The proposed route ratio model however still shows quite some differences to observations. Other models do seem to perform slightly better, of which a model based on Thomas and Tutert (2009, 2010) seems to perform best.

Chapter 6

Conclusions and recommendations

The research objective of this dissertation was to ascertain the ability and accuracy of a network of Bluetooth detectors to be used for route choice research, and, using the data, to determine the influence of route attributes. Two research questions have been formulated:

1. To what extent are the Bluetooth detectors in Alkmaar able to detect routes correctly, and, if not, what can be done to correct for detection errors?
2. What route attributes influence route choice, given the empirical evidence, and does this coincide with previously found relationships?

6.1 Route inference

The Bluetooth detectors in Alkmaar are, similarly to other road side detection systems, not able to detect full routes, but only reveal a sequence of locations where a device have been detected. A found sequence may consist of multiple trips. Only the observed time difference between two successive observations can reveal if (a) those observations belong to one trip – when the time difference is reasonable for the observed pair, or (b) indicate a start and end of a trip – when time difference is overly large. A specific problem has been observed for successive observations at the same location, which amounts to nearly 13 per cent of all successive observations. This can indicate trip ends, but can also be caused by detection issues due to e.g. road layout. Choosing a cut-off time in both cases is an arbitrary decision, which may lead to some error in determination of trip-ends; this is however unavoidable and similar with other other road side detection systems.

The resulting trip sequences are to be translated into routes, to be able to assess route attributes. Trip identification revealed 320 different node sequences, albeit not all are relevant due to e.g. circularity. A comparison with a generated route set, using the Constraint

K Shortest Path-method by van der Zijpp and Fiorenzo Catalano (2005) with a maximum of seven routes per OD-pair, a maximum detour factor of 1.75 and a maximum overlap factor of 0.65, has identified 126 relevant routes through the Alkmaar road network, though some routes would share an observed node sequence. Two generated routes are not observed, the remaining generated routes have been linked to 86 observed sequences: these 86 observations are said to be the relevant observed sequences.

The 234 remaining observed sequences have appeared to be either excessively long, illogical, (partly) circular or incorrectly observed. At least 101 sequences, 32 percent of all observed sequences, seem to be wrongly observed – either missing a likely node or including a node that is close to, but not on the likely route. By including or excluding a node in such a sequence, 41 sequences have been found to belong to relevant routes. The other observed sequences, 60 per cent, have been discarded, albeit this only involves 0.7 per cent of all observed trips.

It is therefore possible to determine likely routes based on found sequences from Bluetooth detectors, though it does require correction for detection errors. However, in a city environment, like Alkmaar, the amount of relevant alternative routes is quite large; with only 7 available Bluetooth detectors some error in route attribution may have occurred as a wrongful (non-)detection easily leads to attribution to another route. The use of more detectors may provide more detail and assurance.

With over 99 per cent of the observed trips by the Bluetooth detectors in Alkmaar attributed to a relevant route, the usability of such systems to determine route ratios is good.

6.2 Route attributes

An uncomplicated utility-based route choice model would use the utilities of all routes within an OD-pair to determine the route ratio of each of the routes. Travel time, being a negative utility, is easy to determine from travel data. The Alkmaar data has revealed a relationship between route ratio and observed travel time, or to be more exact, travel time difference to the fastest route within the observed OD-pair. Equation 6.1 shows this relationship. It is however not a perfect fit, with several large residuals.

$$\frac{n_i}{n_1} = e^{-1.75(TT_i - TT_1)^{0.65}} \tag{6.1}$$

Other factors therefore do seem of importance. Based on the idea that travellers have no knowledge of the actual travel times, they make a decision on route attributes. Analysis of the 86 identified relevant observed routes in Alkmaar has revealed that a traveller's perception can be explained better by a combination of several route attributes instead of by travel time alone, i.e. by a generalised travel time difference. The significant route attributes found for the Alkmaar network have appeared to be:

- Distance over dual carriageways
- Distance over arterial roads
- Distance over collector roads
- Number of direction changes
- Presence of route signage

Each of the three identified road types appeared to be considered differently, even when correcting for maximum allowed speed. In the Alkmaar network, there seems to be a preference for arterial roads, over dual carriageways and collector roads. Each direction change in a route is likely to cause some extent of delay, this is included by assuming a penalty time for each direction change. The presence of route signage has an influence as well in Alkmaar, though quite unexpectedly penalises a route, where it is expected to increase route use and therefore would benefit a route. No explanation has been found for this difference, though it may well be caused by a different but unconsidered route attribute that has similar variable values.

With differences between road types affecting travel time perception, the findings of Hamerslag (1981) seem to fit this research, though additional factors seem to be present. It is recommended that the number of direction changes is to be included in travel time models as well.

The earlier found route ratio relationship has been compared to both the ordinary multinomial logit model, as well as the model based on the findings of Thomas and Tutert (2009, 2010), with all models using generalised travel time difference, with the parameters optimised for each individual model. The model devised by Thomas and Tutert comes out best of the three, with the least squared residuals, although the differences between them are small.

More importantly than the exact model used, is the use of route attributes instead of only travel time. A road type preference seems to have influence too, as well as the number of direction changes. Inclusion of such attributes has majorly benefited the accuracy of the route choice models devised in this research for the Alkmaar network.

6.3 Further research

Considering Bluetooth detectors, it would be advisable to have enough detectors to allow for a more detailed sequence in highly detailed networks. The range of the detectors does require a minimum distance between them, to reduce the number of detections whilst the detected device has not passed the detector. Especially within cities, being highly detailed, this may limit the usability of Bluetooth detectors for route research. To what extend this really is a problem should be up for further research, e.g. by a combined GNSS and Bluetooth research.

An aspect that has not been touched in this research is the amount of knowledge of the network of individual drivers. It is imaginable that drivers that are not from and with a destination outside the region of Alkmaar, i.e. interregional drivers, have less knowledge of the Alkmaar road network, and therefore have a smaller choice set to choose from; local drivers have a more extensive network knowledge and therefore may have a larger choice set. Although this does not influence the results of this research, it may have an effect on the applicability at other locations. Research into this subject would be advisable as well.

Lastly, this research has shown that the expected speed along different types of road are not assessed by drivers to be similar to (a fixed ratio of) the maximum allowed speed. The reasons why a driver expects a certain speed has not been part of this research; such research would be interesting though, and may give even more insight in how drivers assess routes.

Appendix A

Network info

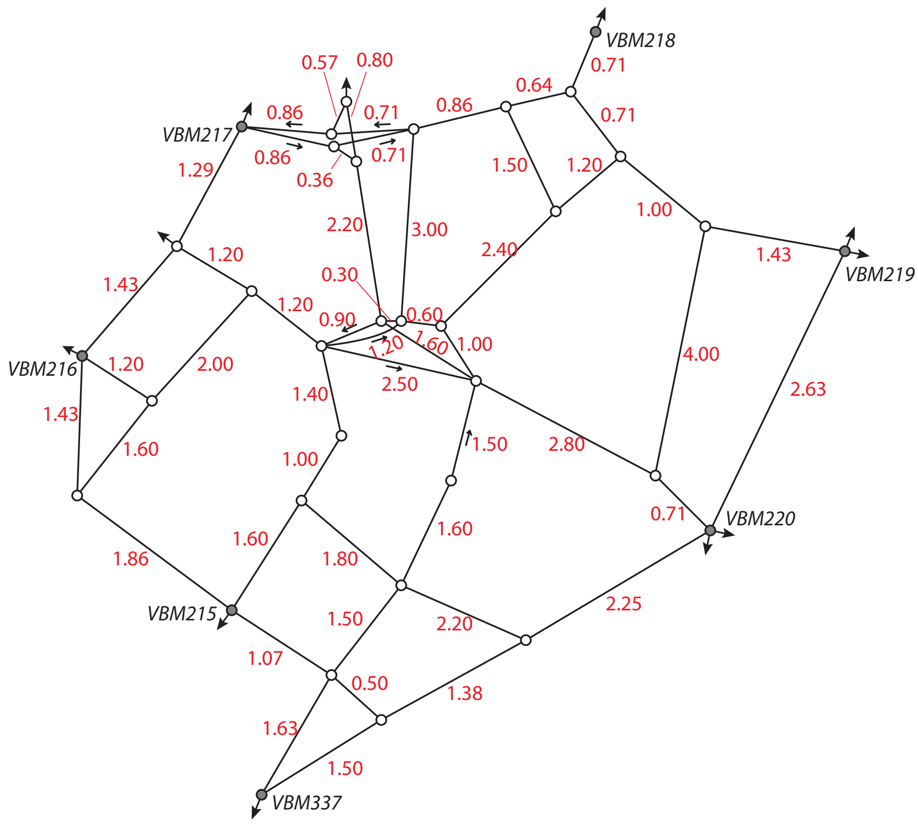

Figure A.1: Link minimum travel time (in 10^{-2} hours)

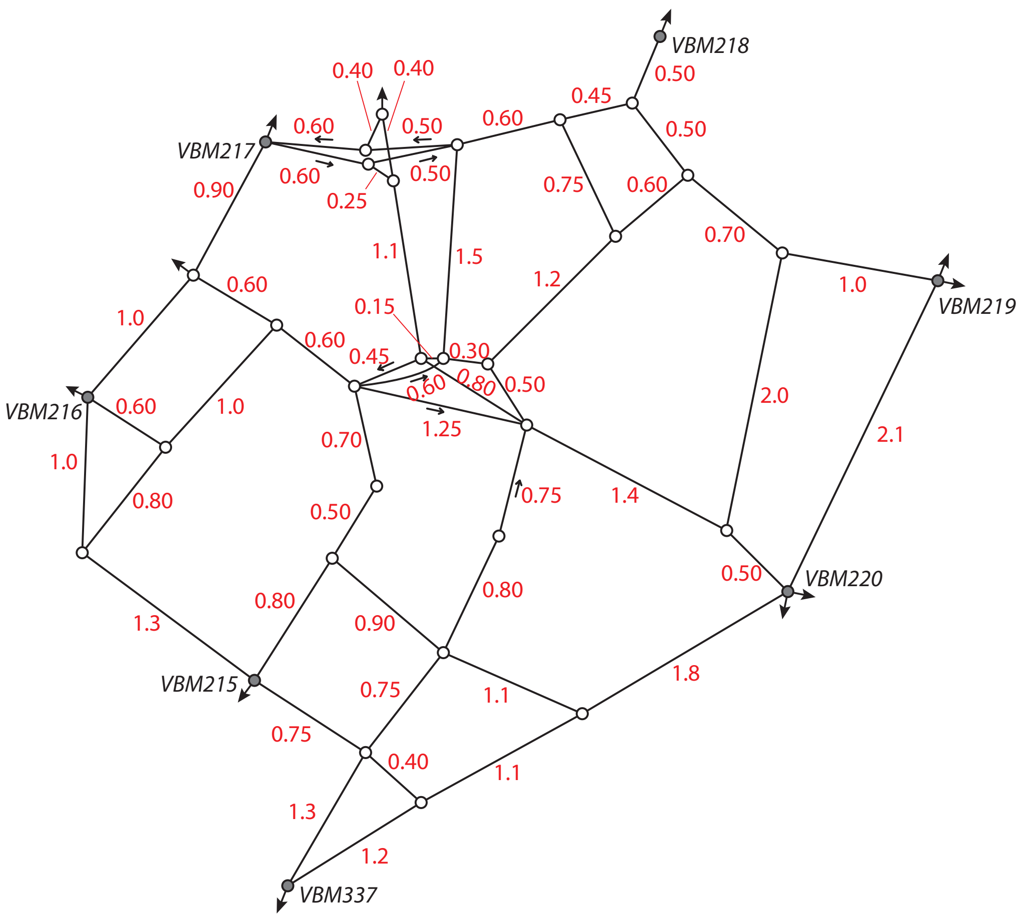

Figure A.2: Link distance (in kilometres)

Appendix B

Generated route set

The figures on the next pages show the result of the route generation process, based on the constrained *k* shortest paths. Table B.1 shows the constraint parameter values used.

Constraint	**Value**
Maximum number of paths	7
Maximum detour	1.75
Maximum overlap	0.65

Table B.1: Constraint parameter values

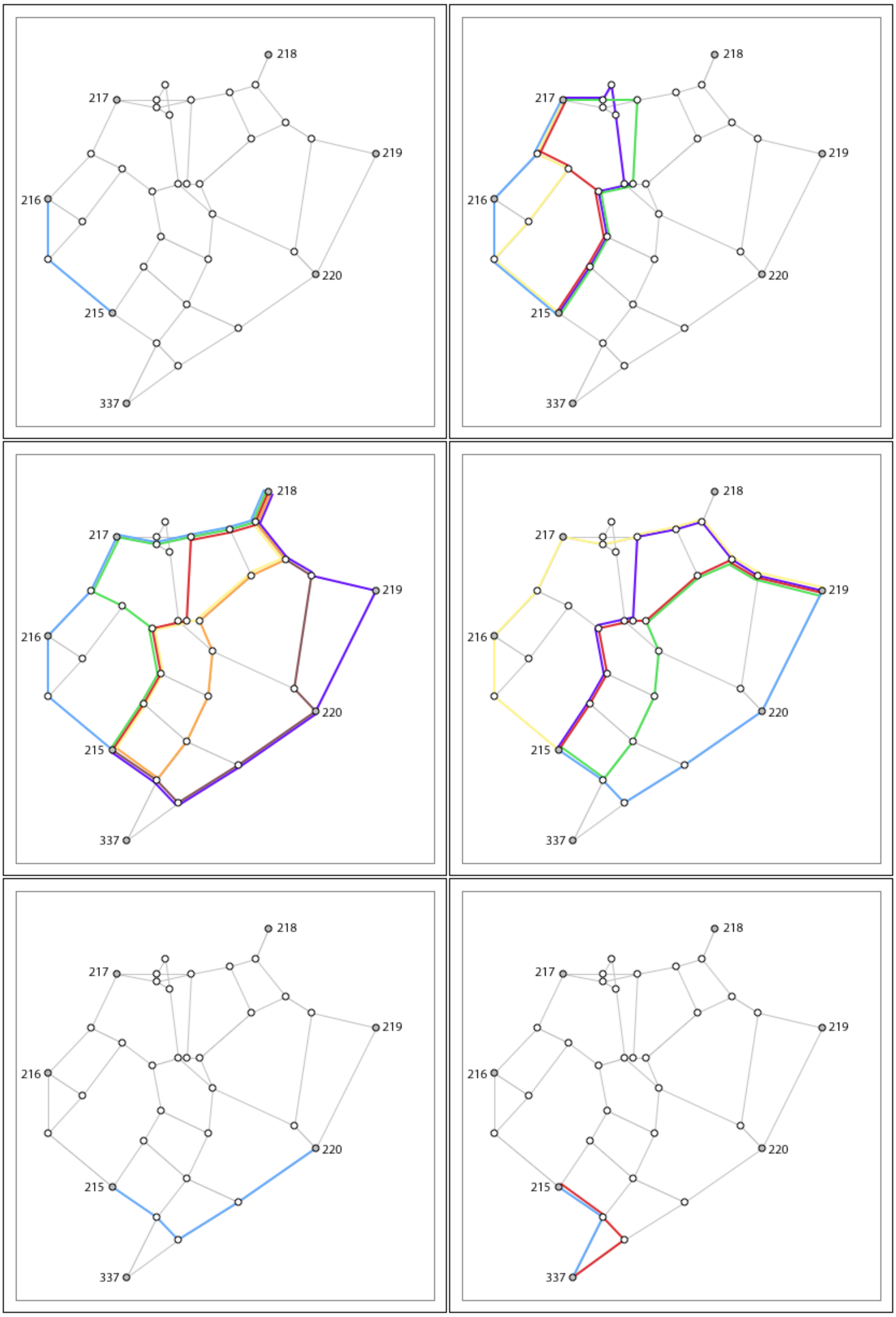

Figure B.1: Generated routes from node 215

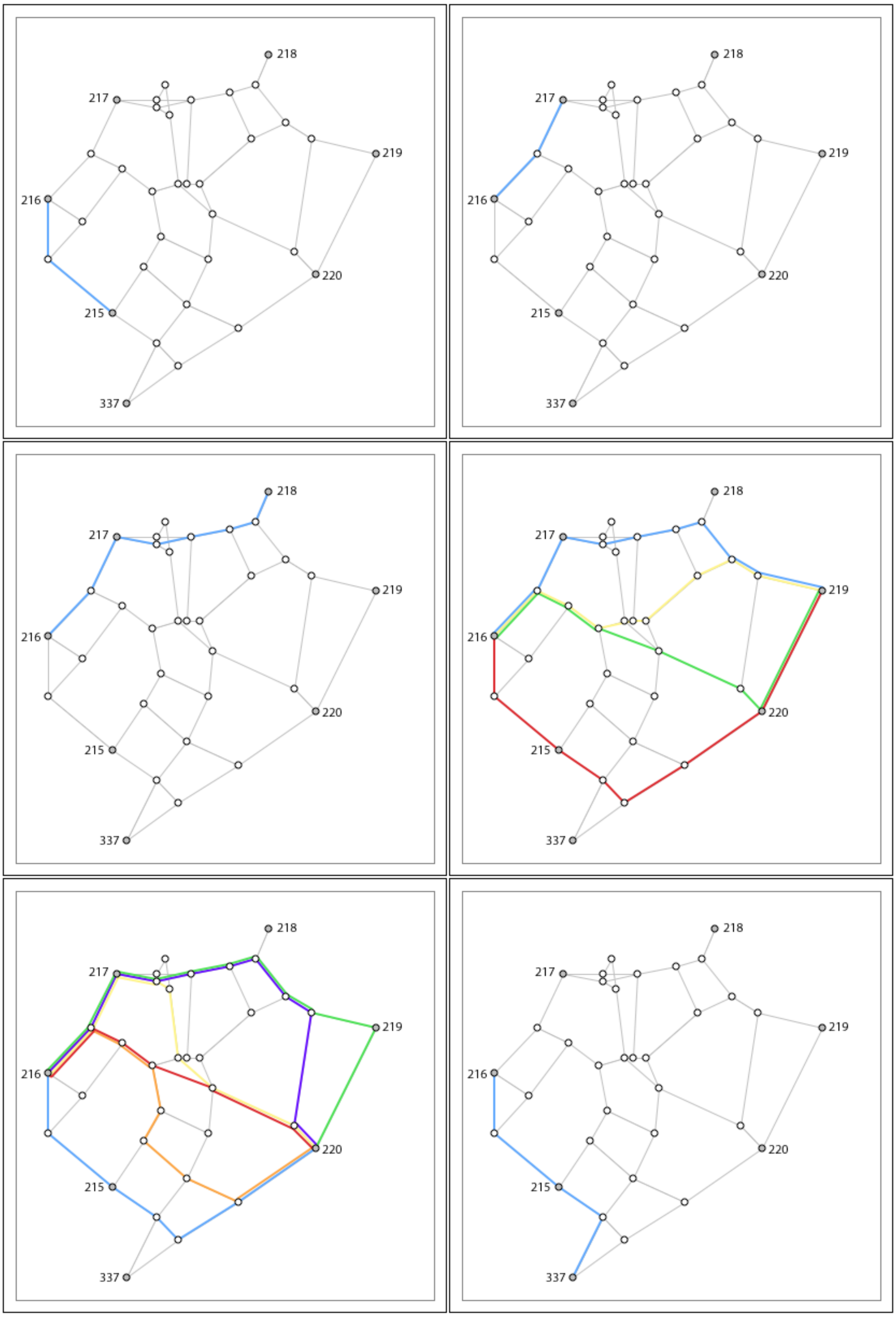

Figure B.2: Generated routes from node 216

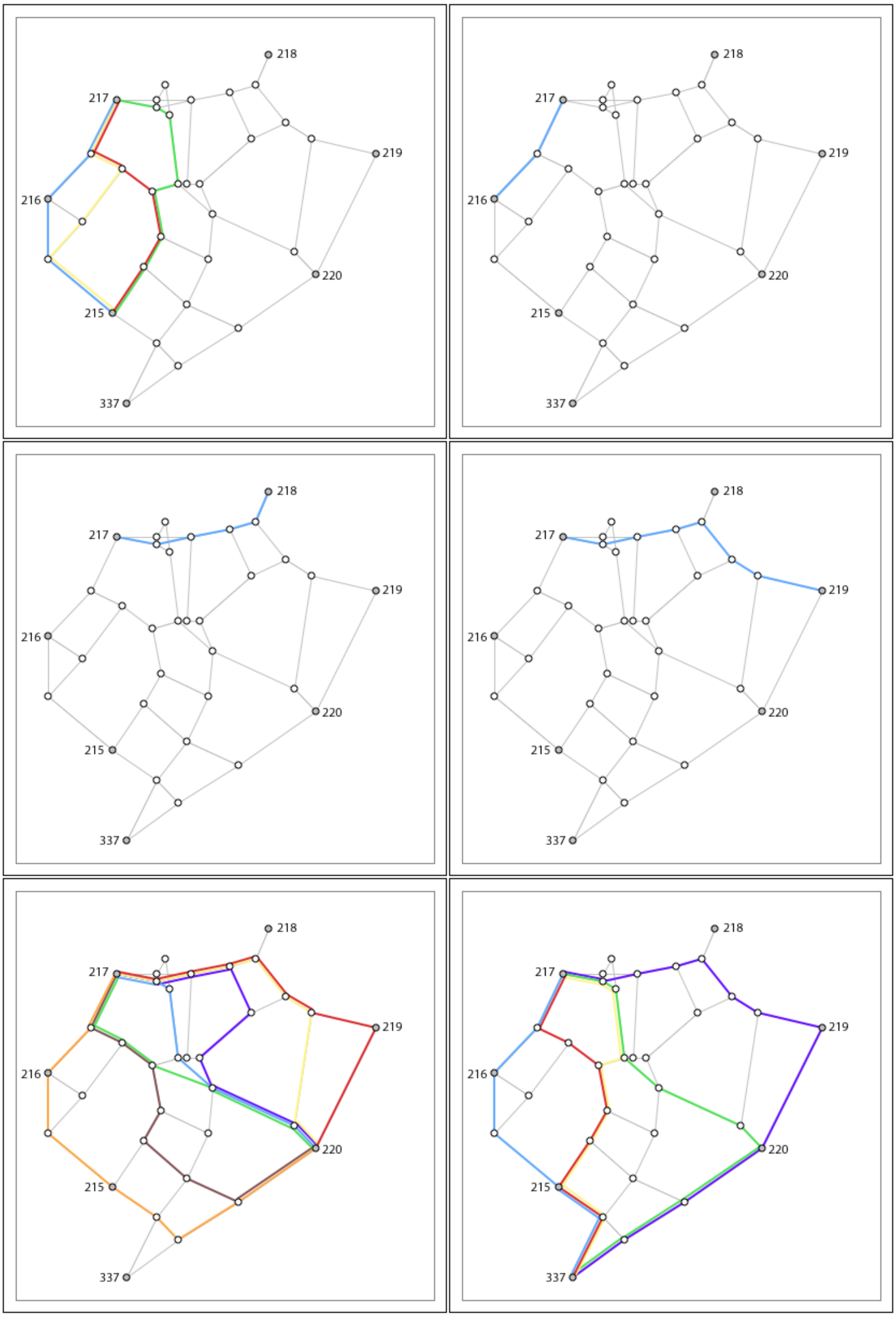

Figure B.3: Generated routes from node 217

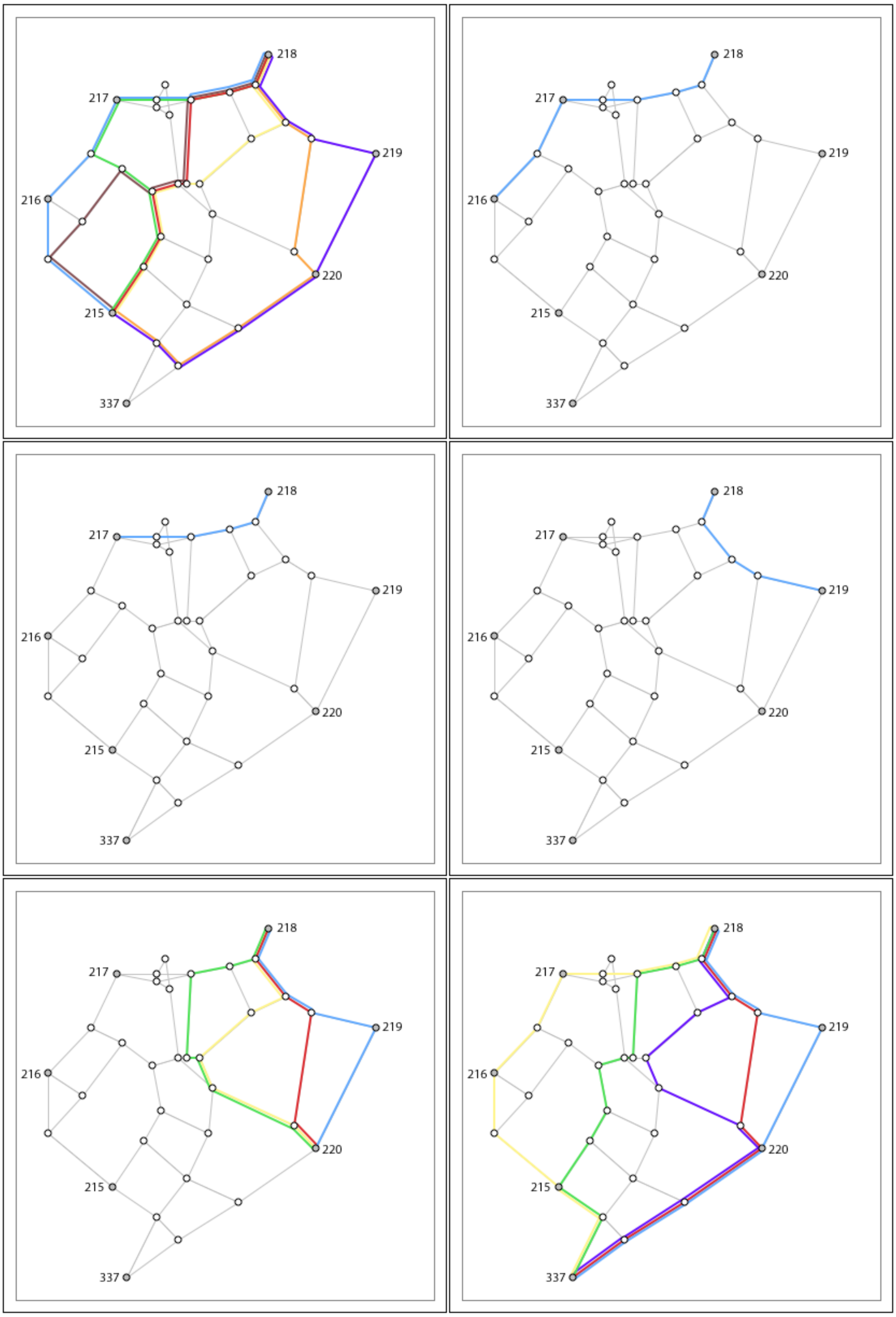

Figure B.4: Generated routes from node 218

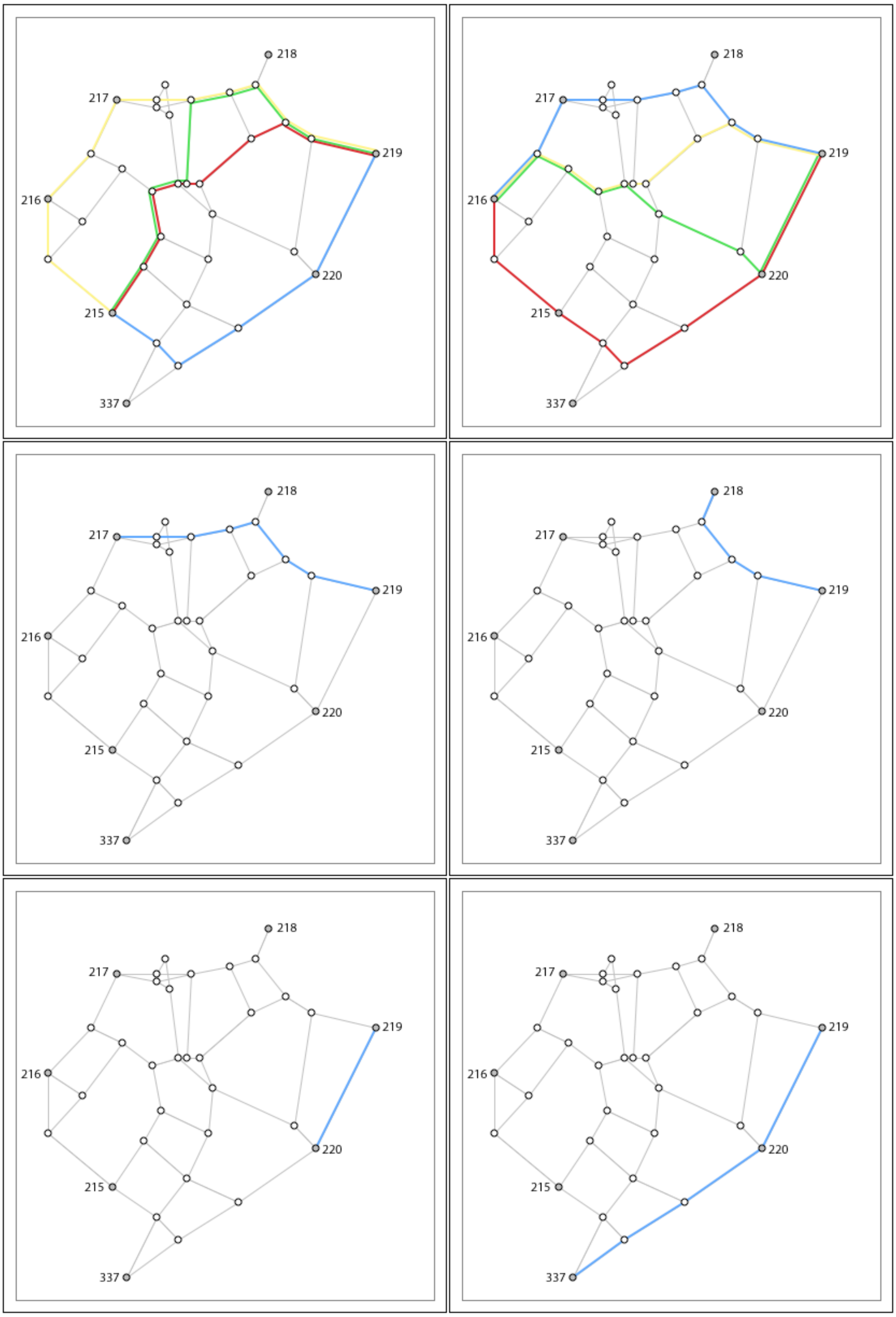

Figure B.5: Generated routes from node 219

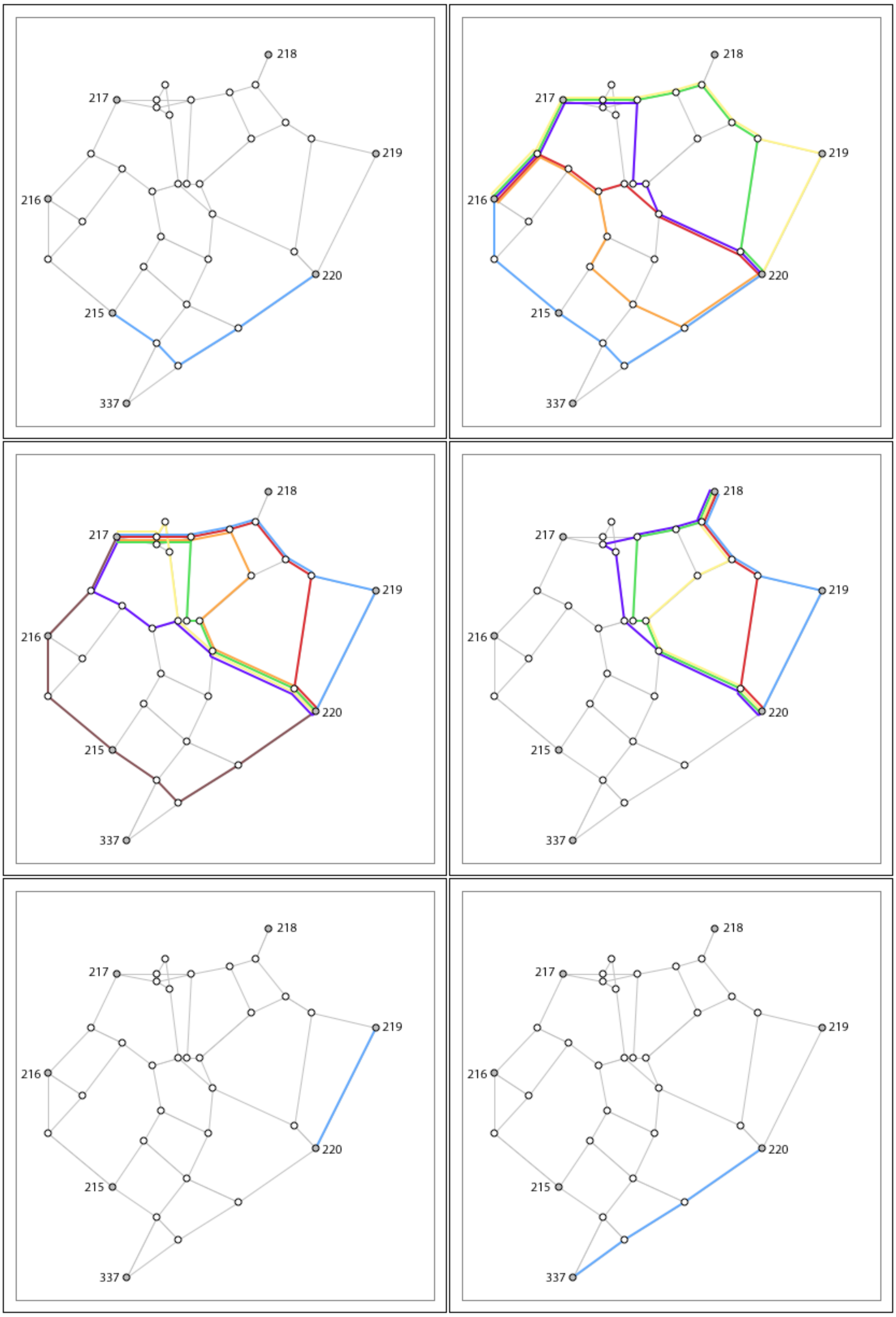

Figure B.6: Generated routes from node 220

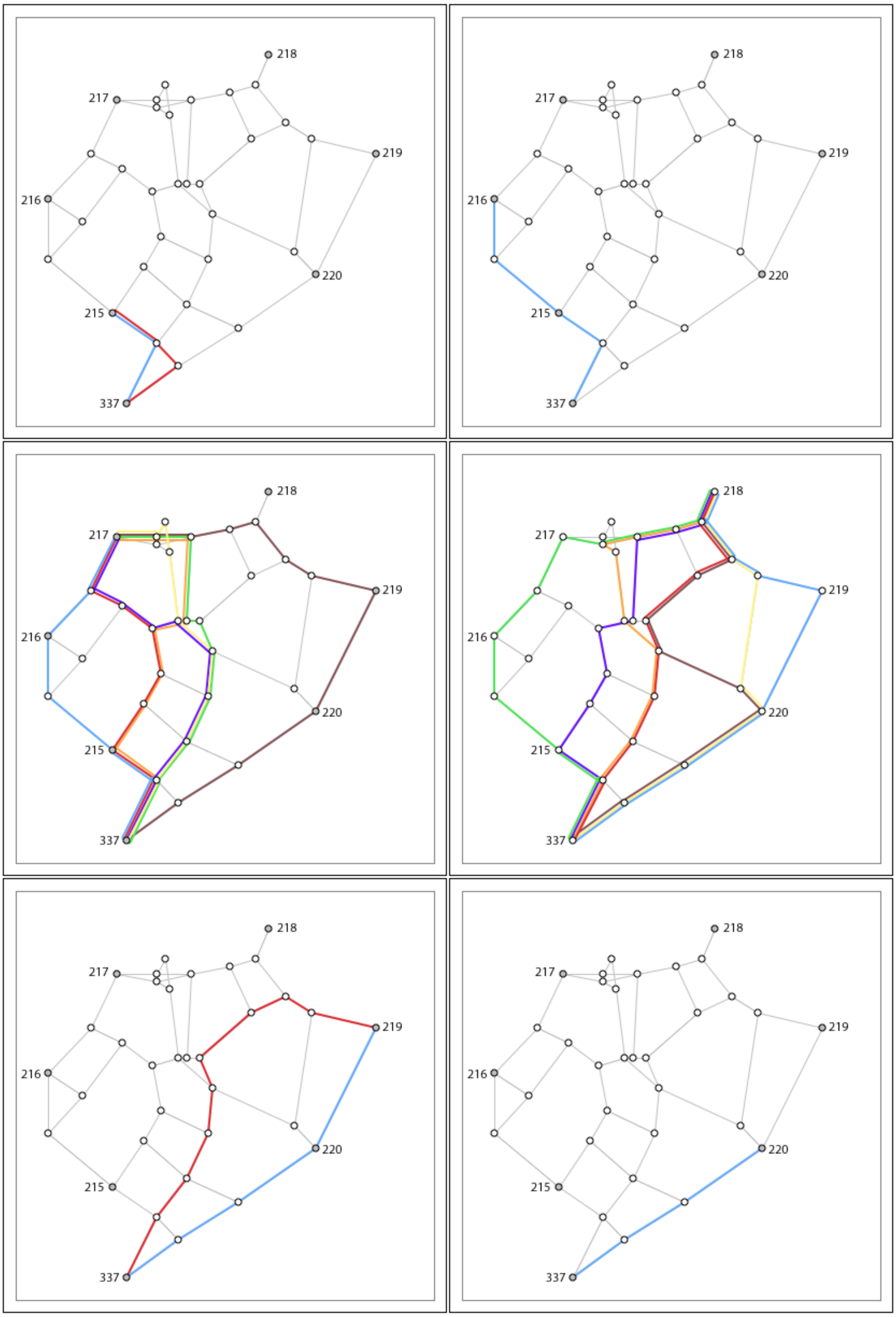

Figure B.7: Generated routes from node 337

Appendix C

Amount of devices detected

OD	Node sequence	Amount of detections
215-216	215-216	3979
215-217	215-216-217	3841
215-217	215-217	821
215-218	215-216-217-218	995
215-218	215-217-218	107
215-218	215-218	44
215-218	215-220-218	1
215-218	215-220-219-218	5
215-219	215-216-217-219	20
215-219	215-219	31
215-219	215-220-219	13
215-220	215-220	872
215-337	215-337	2683
216-215	216-215	3684
216-217	216-217	3822
216-218	216-217-218	1745
216-219	216-215-220-219	1
216-219	216-217-219	447
216-219	216-219	41
216-219	216-220-219	2
216-220	216-215-220	354
216-220	216-217-219-220	3
216-220	216-217-220	10
216-220	216-220	115
216-337	216-215-337	3782

Continued on next page

Continued from previous page

OD	Node sequence	Amount of detections
217-215	217-215	1549
217-215	217-216-215	3738
217-216	217-216	3470
217-218	217-218	3186
217-219	217-219	1755
217-220	217-216-215-220	108
217-220	217-219-220	79
217-220	217-220	48
217-337	217-215-337	938
217-337	217-216-215-337	5366
217-337	217-219-220-337	15
217-337	217-220-337	11
218-215	218-215	47
218-215	218-217-215	286
218-215	218-217-216-215	1063
218-216	218-217-216	1692
218-217	218-217	3710
218-219	218-219	3132
218-220	218-219-220	1817
218-220	218-220	569
218-337	218-215-337	11
218-337	218-217-216-215-337	816
218-337	218-219-220-337	2115
218-337	218-220-337	293
219-215	219-215	8
219-215	219-217-216-215	31
219-215	219-220-215	18
219-216	219-216	24
219-216	219-217-216	302
219-216	219-220-215-216	4
219-216	219-220-216	2
219-217	219-217	1580
219-218	219-218	3498
219-220	219-220	4387
219-337	219-220-337	3415
220-215	220-215	1102
220-216	220-215-216	428
220-216	220-216	67

Continued on next page

Continued from previous page

OD	**Node sequence**	**Amount of detections**
220-216	220-217-216	13
220-216	220-219-217-216	10
220-217	220-215-216-217	75
220-217	220-217	50
220-217	220-219-217	127
220-218	220-218	542
220-218	220-219-218	2278
220-219	220-219	3260
220-337	220-337	15812
337-215	337-215	2566
337-216	337-215-216	3429
337-217	337-215-216-217	4784
337-217	337-215-217	579
337-217	337-217	212
337-217	337-220-219-217	24
337-218	337-215-216-217-218	256
337-218	337-215-218	11
337-218	337-218	155
337-218	337-220-218	223
337-218	337-220-219-218	1893
337-219	337-219	483
337-219	337-220-219	1686
337-220	337-220	11298

Table C.1: Amount of devices detected per unique route

Appendix D

Regression analysis results

Parameter	Coefficients	St Error	P-value
Run 1			
Type1	1.93	0.36	< 0.01
Type2	1.49	0.46	< 0.01
Type3	1.88	0.48	< 0.01
DirCh	0.63	0.36	0.09
Signage	2.35	0.85	< 0.01
Bridge	1.20	0.76	0.12
Signals	-0.40	0.29	0.14
Run 2			
Type1	1.88	0.37	< 0.01
Type2	1.00	0.33	< 0.01
Type3	1.43	0.39	< 0.01
DirCh	0.58	0.36	0.12
Signage	2.07	0.84	0.02
Bridge	0.71	0.70	0.32
Run 3			
Type1	1.78	0.35	< 0.01
Type2	1.05	0.33	< 0.01
Type3	1.33	0.37	< 0.01
DirCh	0.71	0.34	0.04
Signage	1.76	0.78	0.03

Table D.1: Regression analysis model 1

Parameter	Coefficients	St Error	P-value
Run 1			
Type1	1.85	0.32	< 0.01
Type2	1.41	0.41	< 0.01
Type3	1.69	0.42	< 0.01
DirCh	0.49	0.32	0.13
Signage	1.87	0.75	0.02
Bridge	0.94	0.67	0.17
Signals	-0.27	0.26	0.29
Run 2			
Type1	1.82	0.32	< 0.01
Type2	1.11	0.29	< 0.01
Type3	1.42	0.34	< 0.01
DirCh	0.46	0.32	0.15
Signage	1.70	0.73	0.03
Bridge	0.63	0.61	0.31
Run 3			
Type1	1.73	0.31	< 0.01
Type2	1.15	0.29	< 0.01
Type3	1.32	0.33	< 0.01
DirCh	0.58	0.29	0.05
Signage	1.42	0.69	0.04

Table D.2: Regression analysis model 2

Bibliography

Adler, J., Recker, W. and McNally, M. (1993), 'A conflict model and ineractive simulator (FASTCARS) for predicting enroute driver behavior in response to real-time traffic condition information', *Transportation* **20**, 83–106.

Azevedo, J., Santos Costa, M., Silvestre Madeiro, J. and Vieira Martins, E. (1993), 'An algorithm for the ranking of shortest paths', *European Journal of Operational Research* **69**(1), 97–106.

Batley, R. and Clegg, R. (2001), Driver route and departure time choices: The evidence and the models, *in* 'Proceedings of the 33rd Annual UTSG Conference', St. Anne's College, Cambridge, UK.

Bekhor, S., Ben-Akiva, M. and Scott Ramming, M. (2006), 'Evaluation of choice set generation algorithms for route choice models', *Annals of Operations Research* **144**, 235–247.

Ben-Akiva, M., Bergman, M., Daly, A. and Ramaswamy, R. (1984), Modelling inter urban route choice behaviour, *in* J. Volmuller and R. Hamerslag, eds, 'Proceedings of 9th International Symposium on Transportation and Traffic Theory', VNU Press, Utrecht, pp. 299–330.

Bock, F., Kantner, H. and Haynes, J. (1957), An algorithm (the r-th best path algorithm) for finding and ranking paths through a network, Technical report, Armour Research Foundation.

Bonsall, P. (1992), 'The influence of route guidance advice on route choice in urban networks', *Transportation* **19**, 51–73.

Casas i Vilaro, J. (2004), Estimation of the Transport Demand for Real-Time Applications, PhD thesis, Universitat Polytechnica de Catalunya, Barcelona, Spain.

Cascetta, E., Nuzzolo, A., Russo, F. and Vitetta, A. (1996), A modified logit route choice model overcoming path overlapping problems: Specification and some calibration results for interurban networks, *in* 'Proceedings of 13th International Symposium on Transportation and Traffic Theory', Lyon, France.

Chen, J., Newman, J. and Bierlaire, M. (2009), Modeling route choice behavior from smartphone GPS data. Working paper, Transport and Mobility Laboratory, EPFL, Lausanne, Switzerland.

De la Barra, T., Perez, B. and Anez, J. (1993), 'Multidimensional path search and assignment', *Proceedings of the 21st PTRC Summer Meeting* **363**, 307–319.

Dijkstra, E. (1959), 'A note on two problems in connexion with graphs', *Numerische Mathematik* **1**, 269–271.

Erke, A., Sagberg, F. and Hagman, R. (2007), 'Effects of route guidance variable message signs (vms) on driver behaviour', *Transportation Research F* **10**(6), 447–457.

Fafieanie, M. (2009), Calibrating route set generation by map matching GPS data, Master's thesis, Centre for Transport Studies, University of Twente, Enschede, The Netherlands.

Friedrich, M., Jehlicka, P. and Schlaich, J. (2008), Automatic number plate recognition for the observance of travel behavior, *in* '8th International Conference on Survey Methods in Transport: Harmonisation and Data Comparability', Annecy, France.

Golledge, R. and Gärling, T. (2002), Spatial behavior in transportation modeling and planning, *in* K. Goulias, ed., 'Transport Engineering Handbook: Methods and Applications', CRC Press, chapter 3.

Hamerslag, R. (1981), 'Investigation into factors affecting the route choice in "Rijnstreek-West" with the aid of a disaggregate logit model', *Transportation* **10**, 373–391.

Hoffman, W. and Pavley, R. (1959), 'A method for the solution of the nth best path problem', *Journal ACM* **6**, 506–514.

Jan, O., Horowitz, A. and Peng, Z. (2000), 'Using GPS data to understand variations in path choice', *Transportation Research Record* **1725**, 37–44.

Lawler, E. (1976), *Combinatorial Optimization: Networks and Matroids*, Holt, Rinehart & Winston, New York, USA.

Leduc, G. (2008), Road traffic data: Collection methods and applications. Working paper, Joint Research Centre, European Commission, Brussels, Belgium.

Li, H., Guensler, R. and Ogle, J. (2005), 'An analysis of morning commute route choice patterns using GPS based vehicle activity data', *Journal of the Transportation Research Board* **1926**, 162–170.

Noland, R. and Polak, J. (2002), 'Travel time variability: a review of theoretical and empirical issues', *Transport Reviews* **22**(1), 39–54.

Noland, R. and Small, K. (1995), 'Travel time uncertainty, departure time choice, and the cost of morning commutes', *Transportation Research Record* **1493**, 150–158.

Ochieng, W., Quddus, M. and Noland, R. (2003), 'Map-matching in complex urban road networks', *Brazilian Journal of Cartography* **55**, 1–18.

Ortuzar, J. d. D. and Willumsen, L. (2001), *Modelling Transport*, John Wiley and Sons, Chichester, UK.

Papinski, D., Scott, D. and Doherty, S. (2009), 'Exploring the route choice decision-making process: A comparison of planned and observed routes obtained using person-based GPS', *Transportation Research F* pp. 347–358.

Pollack, M. (1961), 'Solutions of the k-th best route through a network—a review', *Journal of Mathematical Analysis and Applications* **3**, 547–559.

Prato, C. (2009), 'Route choice modeling: past, present and future research directions', *Journal of Choice Modelling* **2**(1), 65–100.

Robinson, S. (2005), The development and application of an urban link travel time model using data derived from inductive loop detectors, PhD thesis, Centre for Transport Studies, Imperial College London, London, UK.

Schnabel, W. and Lohse, D. (1997), *Grundlagen der Straßen-Verkehrstechnik und der Verkehrsplanung*, Vol. 2, Verlag für Bauwesen, Berlin, Germany.

Telgen, M. (2010), Realistic route choice modelling, Master's thesis, Centre for Transport Studies, University of Twente, Enschede, The Netherlands.

Thomas, T. and Tutert, S. (2009), Een nieuwe kijk op routekeuze? Bevindingen uit het kentekenonderzoek Enschede, *in* 'Colloquium Vervoersplanologisch Speurwerk 2009', Antwerp, Belgium.

Thomas, T. and Tutert, S. (2010), A new look at route choice: findings from a license plate registration study in Enschede. Working paper, Centre for Transport Studies, University of Twente, Enschede, The Netherlands.

van der Zijpp, N. and Fiorenzo Catalano, S. (2005), 'Path enumeration by finding the constrained k-shortest paths', *Transportation Research B* **39**, 545–563.

van der Zijpp, N. and Van Haastrecht, T. (2003), Gsm als informatiebron voor verkeersbeleid, Technical report, ModelIT, Rotterdam, The Netherlands.

Wardman, M., Bonsall, P. and Shires, J. (1996), 'Driver response to variable message signs: a stated preference investigation', *Transportation Research C* **5**(6), 389–405.

Wardrop, J. (1952), Some theoretical aspects of road traffic research, *in* 'Proceedings of the Institute of Civil Engineers', Vol. 2, pp. 325–378.

Wijbenga, A. and Boerma, S. (2010), Bluetooth pilot: Verkeersinformatie via bluetooth, Technical report, MAP Traffic Management/VerkeersInformatieDienst, Lijnden, The Netherlands.

Yen, J. (1971), 'Finding the k shortest loopless paths in a network', *Management Science* **17**(11), 712–716.

Yoo, B., Chon, K., Kang, S. and Kim, S. (2005), 'Origin-destination estimation using cellular phone base station information', *Journal of the Eastern Asia Society for Transportation Studies* **6**, 2574–2588.