Incorporating Heterogeneity in Route Choice Modeling: Methodology and Case Studies Using GPS Data

（ドライバーの異質性を考慮した経路選択モデルに関する方法論及び **GPS** データを用いた実証的研究）

Dawei LI

Doctoral Dissertation

Incorporating Heterogeneity in Route Choice Modeling: Methodology and Case Studies Using GPS Data

（ドライバーの異質性を考慮した経路選択モデルに関する方法論及び GPS データを用いた実証的研究）

by

Dawei LI

Submitted in Partial Fulfillment of the
Requirement for the Degree of
Doctor of Engineering

December 2013

Graduate School of Environmental Studies
Nagoya University
JAPAN

Abstract

This thesis focuses on the route choice behavior of car drivers. Modeling route choice is essential when modeling travel demand. In the macroscopic travel demand models, such as the classical four steps model, we only need apply the simple route choice models that assume the drivers are homogeneous. However, in the recent years, the activity-based travel demand models attract increasing attentions. Since these models are agent based, it becomes necessary and possible to incorporate heterogeneity in route choice modeling. Therefore, the main objective of this study is to explore the heterogeneity in route choice behavior. Particularly, in this thesis, we investigate the incorporations of three kinds of heterogeneity in route choice modeling: the heterogeneity in perceptions, the heterogeneity in processes and the heterogeneity in tastes.

Heterogeneity in perceptions allows travelers' different perceptions of the same observed route attributes. To consider the heterogeneity in perceptions, especially drivers' en-route experiences, a Bayesian network (BN) based model is proposed and applied in a small network of Beijing, China, with the GPS data collected by taxis. In this model, drivers' dynamic travel time perception process is described by the inference problem of BN.

Heterogeneity in processes means that although the manifested path observations are the same, travelers' really choice processes will be different. To consider the heterogeneity in processes is in fact to consider the en-route choices. We propose a process-based method for analyzing dynamic route choice behavior. The dynamic choice process is defined as the sequence of choices during a trip, including the route choices (both pre-trip and en-route choices) and the choices of making a route choice again at every decision node. The model is estimated and compared with conventional models using probe vehicle data. The results confirm that drivers do not tend to make route choice decisions at all decision nodes. The probability of making an en-route choice is related to a driver's sensitivity to benefit and the cost of making the decision, which is positively correlated with distance to the origin and negative correlated with the spatial scale of the intersection at the decision node.

The differences of tastes on the same observed attributes are referred as heterogeneity in tastes. We at first explore the taste heterogeneity which is related to some observed attributes. Particularly, we explore the effect of familiarity on route choice behavior. Familiarity considered here is both individual and OD pair specific, different from previous researches which only consider the individual specific familiarity. Three methods are applied: class specific parameters, structured scale parameter and structured parameters of explanation variables. The effect of familiarity to O-D pairs on route choice behavior is proved to be statistically significant using the data collected in Toyota by private cars. The model with structured parameters of explanation variables has a better performance than the structured scale parameter model. It is found that drivers will be more easy affected by some unobserved factors and less sensitive to the count of intersections than free travel time when travel between more familiar OD pairs.

At last, we propose a multi-level mixed logit model to incorporate both observed and unobserved characteristics. In the proposed model, the taste coefficients are treated as random and structured as observed characteristics. Further, to deal with the panel data problem, random taste is divided to three components: traveler specific, O-D pair specific and choice situation specific. Models with various assumptions about heterogeneity are estimated and compared using the GPS data collected in Toyota. Beside some behavioral findings, empirical analysis suggests that, to enhance the performance of route choice models, it is more efficient to add more observed characteristics relating to travelers and O-D pairs than to increase the complexity of the specification. It is also proved that the incorporation of O-D pair specific unobserved taste heterogeneity can enhance the performance of a route choice model significantly.

Acknowledgements

This dissertation presents the results of two and a half years of study in the School of Environmental studies, Nagoya University. Here I owe much gratitude to many people who generously supported me.

First and foremost, I owe my most heartfelt thanks to Professor MORIKAWA Takayuki, my principal supervisor and dissertation committee chairman, for his consistent encouragement, valuable suggestions, and insightful guidance in my research, and assistance in my daily life of studying abroad. His solid knowledge of traffic engineering helped me to broaden my view and approach the problems from both computer science and traffic engineering perspectives. I am also particularly grateful to Professor YAMAMOTO Toshiyuki for his kindness and guidance for my research in Nagoya University.

I would like to thank Associate Professor MIWA, Tomio, for his kindness and academic advice for this dissertation. I also greatly appreciate the assistance from Assistant Professor SATO, Hitomi during my study in Nagoya University.

The members of my research group deserve thanks. I would like to express my sincere appreciation to Ms. MU Rui, Mr. CAO, Peng, Mr. YANG Jia, Ms. SUN Xiaohui, and Mr. XU Gang for their suggestion and assistance throughout my studies, and also thanks to all my friends, who always support and assist me.

Finally, it is my greatest pleasure to dedicate this small achievement to my parents. I have totally relied on the continuous support and encouragement of them throughout my study period with great love and patience.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Clothing, food, housing, and travel are the four basic elements for life, and we cannot live without one of them. Compared with the past, in this modern society, travel becomes much easier because of the advanced transportation system. Numerous trips are made for going to work, going shopping, meeting friends, attending activities and so forth. Many of these trips are made by car which has led to a lot of problem, such as congestion and pollution. These problems have a negative impact on peoples' everyday lives. In order to decrease the negative impact of travel, it is essential to first understand travel behavior.

Travel behavior is about all the choices during the process that people make trips, which will raise the questions as following:

- Whether do they have a trip for some purposes?
- Where do they go?
- When do they departure?
- Which transportation mode is used?
- **Which route is used?**

Actually, the core part of travel demand modeling is to answer these five questions. This research focused on the last question, i.e. the route choice behavior given a trip. More precisely, we are interested in identifying which route a given traveler would take to drive from one location to another in an urban road network.

Route choice models play a crucial role in many transport applications. The modeling of route choice behavior is essential if we are to appraise travelers' perception of route characteristics, to forecast travelers' behavior under hypothetical scenarios, to predict future traffic conditions on transportation networks and to understand travelers' reaction and adaptation to sources of information (Prato, 2009).

As the core of traffic assignment and simulation procedures, route choice modeling is essential in both macroscopic transport planning software (e.g. TransCAD) and microscopic traffic simulation tools (e.g. the dynamic traffic assignment module of VISSIM). Route choice models are also required in the dynamic traffic management systems, which aim at improving traffic conditions by controlling the supply of network and by providing real time information to travelers to help them make better route choice decisions. As a core part of these systems is to adjust drivers' route choice decisions using different ways (e.g. restricted passage and real time information), so that the traffic network will be at an optimal state. Therefore the understanding of travelers' route choice behavior will be essential in these systems.

As a representation of individual behavior, route choice models allow for the understanding of travelers' choices under different scenarios. Therefore, route choice modeling is also essential when analyzing the effect of policies (e.g. congestion pricing). For example, the city planners will be interested in this question: which kind of people will pay 500 yen more for driving on a congestion charging route, instead of using another free charge but 5 kilometers longer route?

After this general introduction to route choice analysis, we give a more detailed overview of the modeling process in the following section.

1.2 Route Choice Modeling Overview

People often confuse the route choice modeling problem and the route finding problem. The route finding problem is to find an optimum path for the traveler, according to a pre-set objective. For example, if you want to drive from home to a shopping mall, the route finding algorithms can tell you which route is the shortest, or which route is the most reliable.

The route choice modeling is to identify which route a given traveler would take. For example, there are two routes:

Route 1: 3 kilometers, 30% on the arterial road;

Route 2: 5 kilometers, 70% on the arterial road.

If you ask different people to choose a route, you will get different answers. The route choice modeling is to answer the question whether a given traveler will choose route 1, or route 2.

A good review of the route choice modeling problem can be found in Prato (2009). As summarized by Frejinger (2008), **Figure 1.1** gives an overview of the modeling process.

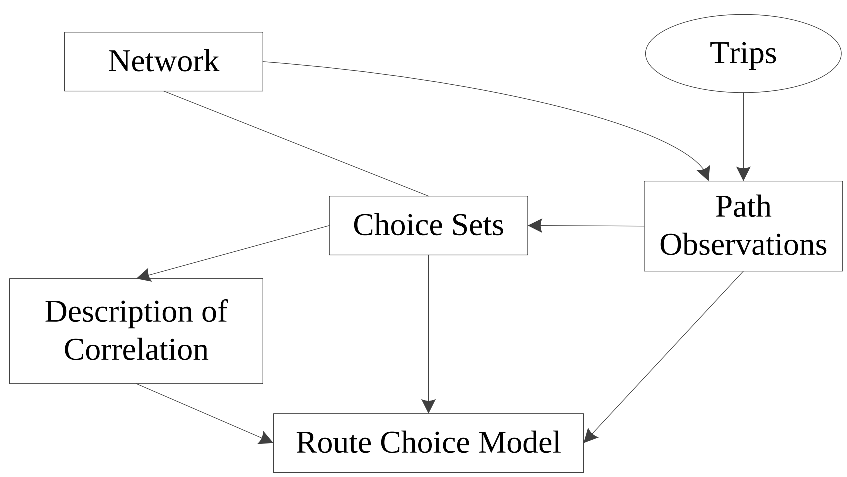


Figure 1.1 Route Choice Modeling Process

As a standard choice modeling problem, before specifying and estimating a choice model, we need obtain the choice sets (i.e. the sets of alternatives that are considered by the decision makers) and the choice observations (i.e. which alternative is chosen by each decision maker).

If the study is based on the state preference (SP) survey data, which is like to ask people to choose route 1 or 2 in the example above, it will be not difficult to get the choice sets and observed path observations. However, for practical applications, at most of the time, the path observations are based on the trips really taken by travelers. Trip observations can be obtained by either asking travelers to describe chosen routes, or by passive monitoring using the Global Positioning System (GPS). In those cases, since the description from the travelers are often ambiguous, and the GPS data have limited accuracy, at first, we need identify which route is really used by the driver, based on both the trip observations and network data.

In a route choice context, the alternatives considered by each traveler are in general unknown to the analyst, except the analysis based on SP survey data. It is therefore necessary to generate a choice set for each path observation. For each origin-destination pair, the number of physically feasible routes is huge, actually unbounded if paths with loops are considered. Therefore choice set generation algorithms are used to define subsets of alternatives as the choice sets for route choice modeling. A review and comparisons on choice set generation methods can be find in Bekhor et al. (2006).

The alternative routes in a choice set often share some same links with each other. This is the so-called "overlapping problem" in route choice analysis. Because of the overlapping problem, the alternatives are highly correlated. Therefore, when modeling route choice, a necessary step is to describe the correlation among alternatives.

The last step of route choice modeling is to specify and estimate the route choice model. The most widely used approach is the random utility model framework. Within this framework, it is assumed that a traveler n associates a utility U_{in} with each route i in the generated choice set C_n. The utility is defined as $U_{in} = V_{in} + \varepsilon_{in}$. V_{in} is the deterministic (systematic) term which is a function of the observed attributes of the alternatives, while ε_{in} is the random term to capture the unobserved utility. In general, a linear formulation is used to specify the deterministic term: $V_{in} = \beta x_{in}$. x_{in} is a vector of observed route attributes. β is a vector of unknown coefficients to be estimated, which can be interpreted as travelers' tastes on the observed route attributes. Travelers are assumed to maximize utility. Therefore the probability that route i is chosen by traveler n from C_n is $P(i \mid C_n) = P(U_{in} > U_{jn}, \forall j \in C_n)$. The detail specifications of $P(i \mid C_n)$ depend on the assumptions on the random term. In next chapter, we will give a further review about this issue.

1.3 Heterogeneity in Route Choice Modeling

In this section, we will give an introduction about the heterogeneity in route choice modeling.

Heterogeneity is opposite to homogeneity. In the context of choice modeling, an assumption on homogeneity means that some differences in different choice scenarios are ignored. To consider heterogeneity is to relax the assumption on homogeneity and consider the ignored differences.

We use the following example to give explicit explanations on the heterogeneity in route choice modeling.

As shown in Figure 1.2, it is assumed that there are only two available routes from node O to node D. The road segments 1 and 2 are composed of main roads, while the road segment 3 is composed of side streets. There are two travelers, **A**-san and **B**-san, who are assumed to have a trip from O to D. We consider the following possible scenarios:

Figure 1.2 An Illustrative Example

- **A:** I think Route 1 will take 5 minutes;
  B: I think Route 1 will take 10 minutes, because I just traveled on road segment 1, and find it is congested.

In the conventional route choice models, the deterministic part of utility $V_{in} = \beta x_{in}$ is a function of the route attributes x_{in} observed by the analyst. In this scenario, we assume that, route travel time is the only route attribute. If **A**-san and **B**-san make route choices at the same time, the travel time of Route 1 observed by the analyst will be the same for both **A**-san and **B**-san. The conventional route choice models will give the same predicted route shares for **A**-san and **B**-san. However, the travelers will make route choices according to their perceived route attributes, not the attributes observed by the analyst. In this case, **A**-san and **B**-san have different perceptions on the travel time of Route 1 because of the different experiences. Therefore, the probability that Route 1 is chosen by **A**-san should be different from that of **B**-san. The conventional route choice models give wrong predictions because of not considering travelers' different perceptions of the same observed route attributes, which is referred as *heterogeneity in perceptions* in this thesis.

- **A:** Yesterday, I chose Route 1 at node O, but I found Route 1 is congested, then I switched to Route 2 at node T;
  B: Yesterday, I chose Route 2 at node O, and didn't change my mind during the trip.

In this scenario, the path observations for **A**-san and **B**-san are both Route 2. We can find that, in route choice modeling, although the manifested path observations are the same, the really choice processes will be different. In the conventional route choice models, the differences in choice processes with the same path observation are ignored. These differences are referred as *heterogeneity in processes* in this thesis. Usually, the conventional route choice models will be assumed that travelers will only make decisions at the origin, and then the choice process of **A**-san is wrong described as the same as **B**-san.

- **A:** I prefer to drive on the side streets, because there are not so many cars;
  B: I hate to drive on the side streets, because there are too many intersections.

The deterministic part of utility $V_{in} = \beta x_{in}$ is related to the parameters β. β can be interpreted as travelers' tastes on observed attributes. In this scenario, the rate of side streets can be seemed as an observed attribute. We can find that, for the same observed attribute, travelers have different tastes. For **A**-san, the tastes on the rate of side streets will be positive, while for **B**-san, the tastes on the rate of side streets will be negative. The differences of tastes on the same observed attributes are referred as *heterogeneity in tastes*. The taste heterogeneity problem is well studied in the field of choice modeling, but not be considered in most of the previous studies on route choice modeling.

In conclusion, there are at least three types of heterogeneity in route choice modeling: heterogeneity in perceptions, heterogeneity in processes, and heterogeneity in tastes. The words "at least" means that there are also other types of heterogeneity in route choice modeling, such as the heterogeneity in choice sets and the heterogeneity in behavioral assumptions. But in this thesis, we only consider the three types of heterogeneity mentioned above. It also should be noted that, there have already some previous studies on the heterogeneity in route choice modeling. A literature review will be given in the next chapter.

1.4 The Objective of This Research

This research has are two main objectives. The first objective of this research is to develop methodologies which explicitly consider the three types of heterogeneity described in last section: heterogeneity in perceptions, heterogeneity in processes, and heterogeneity in tastes. The second objective is to explore travelers' route choice behavior using trip observations extract from GPS data, based on the route choice models incorporating heterogeneity.

Four heterogeneity related problems will be focused in this research:

- How to consider the dynamic perceptions on route attributes?
- How to consider the heterogeneity in processes with path observations?

- How to consider the observed taste heterogeneity?
- How to consider both observed and unobserved taste heterogeneity which is multi-level?

1.5 Outline of This Dissertation

This dissertation is composed of 7 chapters. The outline of this dissertation is presented in the following and for each chapter we give the reference to the publication on which it is based.

- Chapter 2 reviews the literature. We focus on both route choice modeling techniques and methodologies to incorporate heterogeneity.
- Chapter 3 deals with the heterogeneity in perceptions. A Bayesian Networks based method is proposed to consider travelers' en-route experiences and dynamic information explicitly. The GPS data collected in Beijing by taxis are used for a case study.

  Li, D., Miwa, T. and Morikawa, T. (2012) Modeling travelers' perception of travel time for dynamic route choice behavior analysis (scientific paper), Proceedings of the 19th ITS World Congress.
- Chapter 4 deals with the heterogeneity in processes. A choice process based route choice model is proposed to consider travelers' en-route choices. The route choice problem is divided to two choice problems: route choice and decision making choice. Combined with the method proposed in Chapter 3, we do a case study using the data collected in Beijing.

  Li, D., Miwa, T., Morikawa, T. (2013) Dynamic Route Choice Behavior Analysis Considering En Route Learning and Choices. Transportation Research Record: Journal of the Transportation Research Board XXXX, XX-XX.
- Chapter 5 deals with the heterogeneity in tastes. This chapter uses the GPS data collected in Toyota. We first explore the effect of familiarity to origin-destination pairs on the tastes on observed route attributes. Then two models which can consider the observed heterogeneity are proposed and compared.

  Li, D., Miwa, T., Morikawa, T. (2013) Use of private probe data in route choice analysis to explore heterogeneity in drivers' familiarity with origin-destination *pairs. Transportation Research Record: Journal of the Transportation Research Board 2338, 20-28.*

- Chapter 6 also deals with heterogeneity in tastes. This chapter proposes a multi-level mixed logit method to consider both observed and unobserved heterogeneity. The unobserved heterogeneity is further divided to 3 parts: individual specific term, origin-destination pair specific term and observation specific term. This chapter also used the GPS data collected in Toyota for case study.

  (Submitted to Journal) Li, D., Miwa, T., Morikawa, T. incorporating observed and unobserved heterogeneity in route choice analysis. Transportation Research Part B.

- Chapter 7 provides conclusions and future research perspectives.

References

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

Frejinger, E. (2008) Route choice analysis: Data, models, algorithms and applications. École Polytechnique Federale de Lausanne.

Prato, C.G. (2009) Route choice modeling: Past, present and future research directions. *Journal of Choice Modelling* 2, 65-100.

Chapter 2

Literature Review

Modeling route choice is a particular discrete choice modeling problem. Although there are some other behavioral assumptions about choice making, most of the route choice models are based on the random utility theory (RUT). The basic knowledge about RUT can be found in Ben-Akiva and Lerman (1985), an excellent textbook about RUT based discrete choice modeling.

Formally, modeling route choice is, given a scenario (observation) n, to predict the probability that a traveler will choose a given route $i \in C_n$. C_n is a set of routes that will be considered by the traveler under this scenario. This probability is denoted by $P(i \mid C_n)$. The RUT based choice models assume that the traveler will choose the route with maximized utility. The utility of each route $i \in C_n$, which is denoted by U_i, usually has the following formulation:

$$U_i = V_i + \varepsilon_i = \beta^{'} X_i + \varepsilon_i \tag{2.1}$$

where V_i is the systematic part the is a linear function of X_i, a vector of some observed route attributes (e.g. length and travel time). ε_i is a route specific error term. β is a vector of parameters should be estimated.

Then, $P(i \mid C_n)$ is

$$P(i \mid C_n) = P(U_i > \max_{j \in C_n \backslash i}(U_j) \mid C_n) \tag{2.2}$$

With different assumptions about ε_i, many RUT based route choice models have been proposed in the literature. In the following section, we will give a summary of route choice models.

2.1 Route choice models

2.1.1 Logit structures

In route choice models with logit structures, error terms ε_i are assumed to follow a Gumbel distribution. The simplest one is the multinomial logit (MNL) model, in which error terms are assumed to be identically and independently distributed (i.i.d.):

$$P(i \mid C_n) = \frac{\exp(V_i)}{\sum_{j \in C_n} \exp(V_j)} \tag{2.3}$$

However, this i.i.d. assumption cannot hold in the context of route choice due to the well-known overlapping problem: since the alternative routes often share some road links with each other, their utilities cannot be independent.

To address the overlapping problem while keep a simple logit structure, researchers have proposed several route choice models with deterministic correction of the utility for overlapping paths. The most popular two models are the C-logit model (Cascetta et al., 1996) and the Path-size logit model (Ben-Akiva and Bierlaire, 1999).

These two models have similar formulations. In C-logit model, a commonality factor (CF_i) is added in the utility function to measure the degree of similarity of each route with the other routes in the choice set:

$$P(i \mid C_n) = \frac{\exp(V_i + \beta_{CF} CF_i)}{\sum_{j \in C_n} \exp(V_j + \beta_{CF} CF_j)} \tag{2.4}$$

In path-size model, the commonality factor is replaced by a term called path-size (PS):

$$P(i \mid C_n) = \frac{\exp(V_i + \beta_{PS} \ln PS_i)}{\sum_{j \in C_n} \exp(V_j + \beta_{PS} \ln PS_j)} \tag{2.5}$$

Although these two models have similar formulations, they are based on different interpretations about the correction terms: The commonality factor reduces the utility of a path because of its similarity with respect to other routes, while the path size indicates the fraction of the path that constitutes a "full" alternative. Different formulations for the path size and commonality factors can be found in the literature. A summary can be found in (Prato, 2009).

2.1.2 GEV structures

Different from the models with logit structures, the GEV models account for the overlapping problem within the stochastic part of the utility function. Two route choice models in this category are combinatorial logit (PCL) model and cross nested logit (CNL) model.

Prashker and Bekhor (1998) first apply the PCL model in the context of route choice. Then they present a mathematical formulation for the SUE problem with PCL model(Prashker and Bekhor, 2000). In the PCL, the routes are chosen among a pair of alternatives within the choice set, and the choice probability is defined accordingly:

$$P(i \mid C_n) = \sum_{i \neq j \in C_n} P(ij) P(i \mid ij) \tag{2.6}$$

$$P(i \mid ij) = \frac{\exp\left(\dfrac{V_i}{1-\sigma_{ij}}\right)}{\exp\left(\dfrac{V_i}{1-\sigma_{ij}}\right) + \exp\left(\dfrac{V_j}{1-\sigma_{ij}}\right)} \tag{2.7}$$

$$P(ij) = \frac{(1-\sigma_{ij})\left(\exp\left(\dfrac{V_i}{1-\sigma_{ij}}\right) + \exp\left(\dfrac{V_j}{1-\sigma_{ij}}\right)\right)^{1-\sigma_{ij}}}{\sum_{k=1}^{n-1} \sum_{l=k+1}^{n} (1-\sigma_{kl})\left(\exp\left(\dfrac{V_k}{1-\sigma_{kl}}\right) + \exp\left(\dfrac{V_l}{1-\sigma_{kl}}\right)\right)^{1-\sigma_{kl}}} \tag{2.8}$$

where $P(ij)$ is the marginal probability of choosing the pair among the *n*(*n-1*)*/2* possible pairs, and $P(i \mid ij)$ is the conditional probability of choosing route given the chosen binary pair. σ_{ij} is the similarity coefficient between route *i* and *j*, which is similar to *PS* and *CF* in path-size logit and C-logit model.

The CNL model is also first proposed by Prashker and Bekhor (1998), and then apply it in the SUE problem (Prashker and Bekhor, 2000). Bekhor and Prashker (2001) further generalize this model to be the generalized nested logit (GNL) model. The assumption of CNL model is that routes are chosen within nests, which physically correspond to the links in the network, and the choice probability is:

$$P(i \mid C_n) = \sum_m P(m) P(i \mid m) \tag{2.9}$$

$$P(i \mid m) = \frac{(\alpha_{im} \exp(V_i))^{1/\mu_m}}{\sum_j (\alpha_{jm} \exp(V_j))^{1/\mu_m}} \tag{2.10}$$

$$P(m) = \frac{\left(\sum_i (\alpha_{im} \exp(V_i))^{1/\mu_m} \right)^{\mu_m}}{\sum_h \left(\sum_i (\alpha_{hi} \exp(V_h))^{1/\mu_m} \right)^{\mu_m}} \tag{2.11}$$

where α_{im} are inclusion coefficients and μ_m are nesting coefficients. Inclusion coefficients represent the percentage of the generic link *m* used by the generic alternative route *i*.

The PCL model and CNL model can be applied in the SUE model, however because of some computational and behavioral issues, they are difficult to be applied in the route choice analysis on a large real road network.

2.1.3 Non-GEV structures

In all of route choice models described above assume that, the error terms follow Gumbel distributions. However, there are some route choice models in the literature, the error terms are assumed to follow other distributions.

In the multinomial probit (MNP) model, the errors terms are assumed to be normal distributed. Daganzo and Sheffi (1977) first use this model in the context of route choice. The MNP model can directly address the overlapping problem by covariance matrix of the error terms, with the assumption that utilities are link additive. However, it does not have a closed form. Although some efforts are done to decrease the computational cost (Yai et al., 1997), MNP models are still take much more computational time than the logit models. Therefore, they are rarely to be applied in practice.

Another Non-GEV model is the logit kernel (LK) model with a factor analytic specification where some structure is explicitly specified in the model to decrease its complexity. Bekhor et al. (2002) apply this structure in route choice modeling. In this model, the utilities are assumed to have both normal and Gumbel distributed error terms. Then a flexible correlation structure can therefore be defined while keeping the form of a MNL model. Different from Eq. (2.1), in this specification, the utility function is defined by

$$U_i = \beta^{'} X_i + F_i T \zeta_i + \varepsilon_i \tag{2.12}$$

$F_i T \zeta_i$ is the added normal distributed term. F_i is a link-path incidence matrix, ζ_i is the normal distributed i.i.d. vector with zero mean and unit variance. T is the link factors variance matrix. With the assumption that link-specific factors are i.i.d. normal and that variance is proportional to link length, $T = \sigma \,\mathrm{diag}(\sqrt{l_1}, \sqrt{l_2}, \ldots, \sqrt{l_{M_n}})$. l is the length of links in the choice set, M_n is the total number of unique links in the choice set. σ is a parameter to be estimated.

LK model is in factor a case of mixed logit model, and have been used in several studies on real size networks. The estimation of this kind of models need some simulation methods(Train, 2003).

In the Chapter 6, we will propose some other mixed logit based specifications for route choice modeling.

2.2 Choice Set Generation

According to the review in last section, we can find that, for each observation, before modeling route choice $P(i\,|\,C_n)$, we must generate a choice set C_n at first.

Figure 2.1 gives an overview of the choice set generation process. Given a network and a origin-destination (O-D) pair (s_o ,s_d), there is a set of all available paths which is referred to as the universal set U.

If we only study drivers' route choice behavior on a small test network (e.g. the example shown in last chapter), and not try to apply the estimated models on a real large network, we can enumerate U, and assume the choice set C_n is U.

However, from the aspect of application, the developed route choice models must be feasible to be applied in a large real network. On a real network, e.g. the road network of Nagoya City, there will be a very large number of available paths between an O-D pair. Actually, the size of U will be infinitely if the paths contains loops are also considered. Then it is not computationally feasible to replace C_n by U. On the other hand, from the behavioral perspective, it is not reasonable to assume that the travelers can consider too many available paths when they are making choices.

A directly solution for this problem is to generate a subset of U by some algorithms, which is assumed to be the set of paths that may be considered by the travelers. This subset is called a master set and denoted by M. The choice set generation algorithms can be categorized to two classes: deterministic algorithms and stochastic algorithms. For a given O-D pair, deterministic algorithms will give a unique M for different observations, while the stochastic algorithms will generate an observation (or individual) specific master set M_n.

The choice set for observation n, C_n, can be defined based on the master set in either a deterministic way to assume C_n=M (or $C_n = M_n$), or by using a probabilistic model $P(C_n)$ where all non-empty subsets of M (or M_n) are considered. More detail about the probabilistic model can be found in Manski (1977), Swait and Ben-Akiva (1987), Morikawa (1996), and Kaplan and Prato (2012). The probabilistic choice models are complex due to the large number of non-empty subsets, and have never been used in route choice modeling on a real size network. Therefore, in this study, we only use the deterministic choice set generation models. In the following part of this section, we will give a review of existing deterministic and stochastic path generation algorithms.

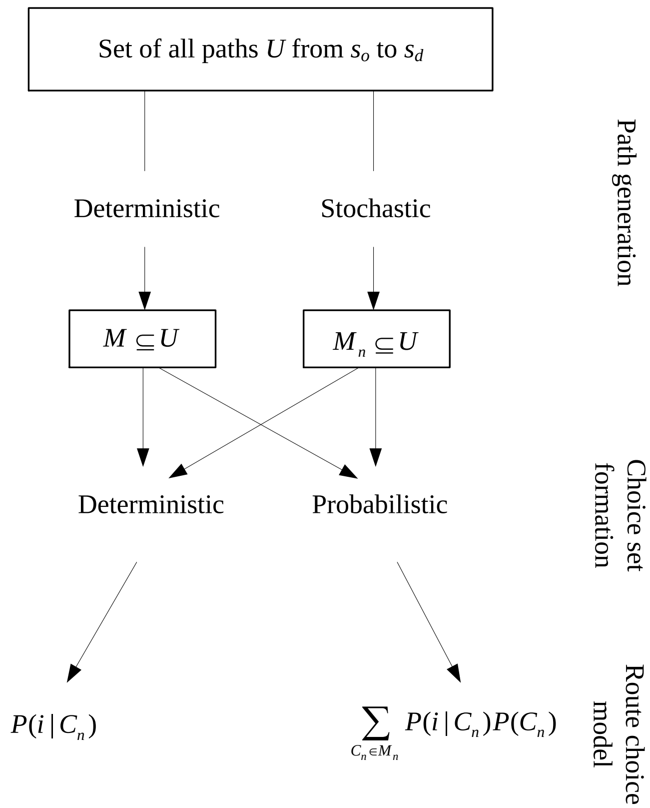


Figure 2.1 Choice Set Generation Overview

2.2.1 Deterministic methods

Deterministic methods can be divided to two categories: shortest path-based methods and constrained enumeration methods.

The shortest path-based methods are based on the repeated shortest path search. The most used algorithms in this category are link elimination methods, link penalty methods, and the labeling method.

The first link elimination algorithm is proposed by Azevedo et al. (1993). The link elimination algorithms at first find the shortest path with respect to the definition of cost and include it in the choice set. Then, each link or some links belonging to the shortest path are removed and a new shortest path is found and included in the choice set.

Instead of eliminating links, link penalty algorithms increase the generalized cost on links in the previously found shortest path for finding a new shortest path (De La Barra et al., 1993). With different plans of increasing link cost, many variants can be found from the literature (Park and Rilett, 1997; Prato and Bekhor, 2007).

Both the link elimination link penalty algorithms will only consider one generalized cost index. Different from these two methods, the labeling approach, first proposed by Ben-Akiva et al. (1984) assume that the drivers may have different objectives. Some of them may wish to minimize travel time, while others may prefer to drive through familiar landmarks. The labeling approach will at first define some labels that a driver may consider e.g. travel time, length and number of links. Then for each label the shortest path algorithm is applied to find the path that minimizes the cost defined by the label. All the found unique paths will be included in choice set.

The shortest path-based methods try to find the "good" paths that the drivers will most likely to consider. On the contrary, the constrained enumeration methods try to exclude the "bad" paths that the drivers will not likely to consider (e.g. the path with more than ten times length of the shortest path).

This method, which is also referred to as branch-and-bound, build a tree where each branch correspond to a path and generate all paths satisfying some constraints. Friedrich et al. (2001) first use this method for public transport network. Hoogendoorn-Lanser (2005) propose this method for multi-modal networks and Prato and Bekhor (2006) apply it for road network.

Compared with the shortest path-based methods, the branch-and-bound methods will take much longer time and generate much more unique alternatives. Since they can generate much more unique alternatives, the branch-and-bound methods will perform better if the coverage of observed choice is the evaluation criterion. However, since the speed of these methods depends exponentially on the depth of the tree, the running time will become unaffordable when apply them on a large road network.

2.2.2 Stochastic methods

Generally, there are two kinds of stochastic methods. The first kind is the shortest path-based algorithms with random cost, which is also called the simulation method (Ramming, 2001). In this method, the generalized cost of each link is assumed to follow some distributions. For each draw, the link costs are sampled, and the shortest path is found and added in the choice set.

Another kind of stochastic methods is the sampling method. It is first proposed by Frejinger et al. (2009), and referred to as the random walk method. In this method, it is assumed that, the choice set is the universal set. Because the universal set is too large, they use random walk method to get some sampled alternatives. They also proposed a method to correct the error caused by using the sampled alternatives to estimate the route choice model. In Chapter 5, we will give a more detail description on this method.

Although there are a lot of methods are proposed to address the choice set generation for route choice modeling, it is always difficult to say which method is the best. The main reason is that the actual choice sets are always unknown to the modelers. We can only propose some measures subjectively to evaluate the choice set generation methods. The following measures, which is proposed by Ramming (2001), are often used: computational time, coverage of the observed routes, number of routes in the choice set and number of links in the choice set. Some deep comparisons of choice set generation methods based on these measures can be found in the previous research (Bekhor et al., 2006; Bovy, 2009).

2.3 Route Choice Data

To estimate a route choice model on a real road network, we usually need three kinds of data:

- the road network, including the topological structure and attributes associated with links and nodes;
- the attributes of travelers who make the route choices;
- and the link-by-link descriptions of the routes used by the travelers.

The road networks now are easy to be obtained, since there are so many digital map providers. The attributes of travelers can be obtained via some questionnaires. The most difficult task for route choice modelers is to observe the routes that really chosen by the travelers. There are two approaches to collect such data: asking travelers to describe chosen routes and extracting the route observations from the location information collected by the Global Positioning System (GPS).

Before the popular of GPS devices, most of the research about route choice is based on the trip data collected via mail, telephone and web-based surveys (Mahmassani et al., 1993; Ramming, 2001). In recent years, most of the studies about route choice are based on GPS data (Bierlaire and Frejinger, 2008; Morikawa and Miwa, 2006).

All of the studies in this thesis are based on the GPS observations. Compared with the conventional methods, using GPS data has some advantages: for instance, multiple days of trip data can be collected automatically and are directly available in electronic format.

However, GPS data also have some additional issues. At first, the raw GPS data are only some location points. Some map-matching algorithms should be applied to match the points to the road links. Second, the GPS data cannot explicitly tell where origins and destinations are. They can only provide the trajectories. Some algorithms should be applied to separate the trajectories to several trips and identify the origin-destination pairs. Third, travel purposes cannot be obtained from the GPS data. At last, in practice, the GPS data often can only describe part of the trips, since some trajectories are not recorded.

2.4 Heterogeneity in Route Choice

Most of the previous studies about heterogeneity in the context of route choice focus on the taste heterogeneity. There are mainly four categories of methods to consider heterogeneity: divide the users to be several classes with different taste coefficients (Yang and Huang, 2004); structure the scale parameter (Chen et al., 2012; Morikawa and Miwa, 2006); add latent variable in the utility function (Prato et al., 2012; Ramming, 2001) and assume the taste coefficients to be random (Srinivasan and Mahmassani, 2003). To avoid too many equations, we will not give the technical details about these methods in this review. The first three methods have some difficulties to deal with panel data. Therefore in this thesis, we will mainly extend the random coefficient method for taste heterogeneity.

About heterogeneity in process and heterogeneity in perception, there are few previous studies about these in the context of route choice. To incorporate heterogeneity in perception, in some feedback-based models, travelers' perceptions of travel time on alternative routes are treated implicitly as an function of the feedback of past experiences on the same route (Ben-Elia and Shiftan, 2010; Joel L, 1984). About the heterogeneity in process, Morikawa and Miwa (2006) first proposed a model to consider this problem. More literature review about the heterogeneity will be given in the beginnings of the following chapters.

References

Azevedo, J., Santos Costa, M.E.O., Silvestre Madeira, J.J.E., Vieira Martins, E.Q. (1993) An algorithm for the ranking of shortest paths. *European Journal of Operational Research* 69, 97-106.

Bekhor, S., Ben-Akiva, M., Scott Ramming, M. (2002) Adaptation of logit kernel to route choice situation. *Transportation Research Record: Journal of the Transportation Research Board* 1805, 78-85.

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

Bekhor, S., Prashker, J. (2001) Stochastic user equilibrium formulation for generalized nested logit model. *Transportation Research Record: Journal of the Transportation Research Board* 1752, 84-90.

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

Ben-Akiva, M., Bierlaire, M. (1999) Discrete choice methods and their applications to short term travel decisions. *Handbook of transportation science*. Springer, pp. 5-33.

Ben-Akiva, M.E., Lerman, S.R. (1985) *Discrete choice analysis: Theory and application to travel demand*. the MIT Press.

Ben-Elia, E., Shiftan, Y. (2010) Which road do i take? A learning-based model of route-choice behavior with real-time information. *Transportation Research Part A: Policy and Practice* 44, 249-264.

Bierlaire, M., Frejinger, E. (2008) Route choice modeling with network-free data. *Transportation Research Part C: Emerging Technologies* 16, 187-198.

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

Cascetta, E., Nuzzolo, A., Russo, F., Vitetta, A. (1996) A modified logit route choice model overcoming path overlapping problems: Specification and some calibration results for interurban networks. *Proceedings of Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, pp. 697-711.

Chen, A., Pravinvongvuth, S., Xu, X., Ryu, S., Chootinan, P. (2012) Examining the scaling effect and overlapping problem in logit-based stochastic user equilibrium models. *Transportation Research Part A: Policy and Practice*.

Daganzo, C.F., Sheffi, Y. (1977) On stochastic models of traffic assignment. *Transportation Science* 11, 253-274.

De La Barra, T., Perez, B., Anez, J. (1993) Multidimensional path search and assignment., *Proceedings of the 21st PTRC Summer Annual Meeting*, Manchester.

Frejinger, E., Bierlaire, M., Ben-Akiva, M. (2009) Sampling of alternatives for route choice modeling. *Transportation Research Part B: Methodological* 43, 984-994.

Friedrich, M., Hofsäß, I., Wekeck, S. (2001) Timetable-based transit assignment using branch and bound techniques. *Transportation Research Record: Journal of the Transportation Research Board* 1752, 100-107.

Hoogendoorn-Lanser, S. (2005) *Modelling travel behaviour in multi-modal networks*.

Joel L, H. (1984) The stability of stochastic equilibrium in a two-link transportation network. *Transportation Research Part B: Methodological* 18, 13-28.

Kaplan, S., Prato, C.G. (2012) Closing the gap between behavior and models in route choice: The role of spatiotemporal constraints and latent traits in choice set formation. *Transportation research part F: traffic psychology and behaviour* 15, 9-24.

Mahmassani, H., Joseph, T., Jou, R.-C. (1993) Survey approach for study of urban commuter choice dynamics. *Transp. Res. Record*.

Manski, C.F. (1977) The structure of random utility models. *Theory and decision* 8, 229-254.

Morikawa, T. (1996) A hybrid probabilistic choice set model with compensatory and noncompensatory choice rules. Volume 1: Travel behavior. *Proceedings of World Transport Research. Proceedings of the 7th World Conference on Transport Research*.

Morikawa, T., Miwa, T. (2006) Preliminary analysis on dynamic route choice behavior: Using probe‐vehicle data. *Journal of advanced transportation* 40, 140-163.

Park, D., Rilett, L.R. (1997) Identifying multiple and reasonable paths in transportation networks: A heuristic approach. *Transportation Research Record: Journal of the Transportation Research Board* 1607, 31-37.

Prashker, J.N., Bekhor, S. (1998) Investigation of stochastic network loading procedures. *Transportation Research Record: Journal of the Transportation Research Board* 1645, 94-102.

Prashker, J.N., Bekhor, S. (2000) Congestion, stochastic, and similarity effects in stochastic: User-equilibrium models. *Transportation Research Record: Journal of the Transportation Research Board* 1733, 80-87.

Prato, C.G. (2009) Route choice modeling: Past, present and future research directions. *Journal of Choice Modelling* 2, 65-100.

Prato, C.G., Bekhor, S. (2006) Applying branch-and-bound technique to route choice set generation. *Transportation Research Record: Journal of the Transportation Research Board* 1985, 19-28.

Prato, C.G., Bekhor, S. (2007) Modeling route choice behavior: How relevant is the composition of choice set? *Transportation Research Record: Journal of the Transportation Research Board* 2003, 64-73.

Prato, C.G., Bekhor, S., Pronello, C. (2012) Latent variables and route choice behavior. *Transportation* 39, 299-319.

Ramming, M.S. (2001) Network knowledge and route choice. Massachusetts Institute of Technology.

Srinivasan, K.K., Mahmassani, H.S. (2003) Analyzing heterogeneity and unobserved structural effects in route-switching behavior under atis: A dynamic kernel logit formulation. *Transportation Research Part B: Methodological* 37, 793-814.

Swait, J., Ben-Akiva, M. (1987) Incorporating random constraints in discrete models of choice set generation. *Transportation Research Part B: Methodological* 21, 91-102.
Train, K.E. (2003) *Discrete choice methods with simulation*. Cambridge university press.
Yai, T., Iwakura, S., Morichi, S. (1997) Multinomial probit with structured covariance for route choice behavior. *Transportation Research Part B: Methodological* 31, 195-207.
Yang, H., Huang, H.-J. (2004) The multi-class, multi-criteria traffic network equilibrium and systems optimum problem. *Transportation Research Part B: Methodological* 38, 1-15.

Chapter 3

Modeling Dynamic Perceptions

Within the random utility model framework, the deterministic part of route utilities depends on two vectors: travelers' perceptions of route characteristics and their tastes on the perceptions.

Travel time is an important route attribute, which must be considered if data is available. In most of the previous studies, travelers' perceptions on route travel time are replaced by the travel time estimated by the analysts. There are mainly two kinds of estimated route travel time for route choice modeling: the fixed travel time (Bierlaire and Frejinger, 2008; Ramming, 2001) and the time-dependent travel time (Li et al., 2013; Morikawa and Miwa, 2006). The fixed travel time will be the free flow travel time or the estimated travel time considering congestion. The time-dependent route travel time will be based on the estimated link cost table, in which the link travel time is specified for each small time interval.

Recall the first example in Section 1.3:

- **A:** I think Route 1 will take 5 minutes;
  B: I think Route 1 will take 10 minutes, because I just traveled on road segment 1, and find it is congested.

We can find that, both the fixed travel time and the time-dependent travel time cannot consider the scenario shown above. In this chapter, a new method for modeling travel time perceptions is proposed to address the problem as shown in this example.

Route choice models can be divided into two categories: static models and dynamic models. The static models only consider the pre-trip decisions, while the dynamic models also consider the en-route decisions (Morikawa and Miwa, 2006). A natural way to build dynamic route choice models is to have a sequence of static choice models at each decision node, where the characteristics of alternative routes reflect updated information (Gao et al., 2010). Some dynamic traffic assignment models, e.g. DynaMIT(Ben-Akiva et al., 2002) and DYNSMART (Mahmassani, 2001), apply dynamic route choice models built in this way.

If the en-route choices are considered, travelers' heterogeneity in perceptions as shown in the above example cannot be ignored. In general, travelers' perception of travel time are affected by three sources of information: historical experiences, external information and current perception during the trip (Adler and Blue, 1998).

To consider the effect of historical experiences, in some feedback-based models, travelers' perceptions of travel time on alternative routes are treated implicitly as an function of the feedback of past experiences on the same route (Ben-Elia and Shiftan, 2010; Joel L, 1984). However, these studies only focus on the static route choice models. There also have been abundant studies of route choice models with external information. However, the previous studies about travel time perception only consider the effect of information on the links with information providing. In fact, travelers can also infer the traffic states of other links based on the provided information. Travelers' current perceptions during the trip are based on the feedback about links that they have passed. Travelers can infer the traffic states of other links according to the feedback. According to authors' limited knowledge, there is no previous research about this.

From the introduction above, it can be found that, there is not a method to model travelers' dynamic perceptions of travel time explicitly for dynamic route choice analysis, which can consider the effect of current perception during the trip. The purpose of this chapter is just to propose such a method. In this chapter, the travelers' historical experiences are represented as a Bayesian network (BN) about network traffic states. Compared with the method in feedback-based models, this method is seemed less "behavioral", but it is more implementable, because it only needs the historical data about network traffic states, such as probe vehicle data to estimate the BN. External information and current perception during the trip are the evidences used to update the BN. The dynamic travel time perception process can be described as the inference problem of this

BN. Unlike the previous studies, during the inferring process, the information about some links can also affect travelers' perceptions about other links without information.

The rest of the chapter is organized as follows. The problem proposed in this chapter is described in the next section. The basic concepts about BN are given in the following section. The BN about network traffic states is then introduced. The inference problem which can describe the dynamic perception of travel time is presented in the following section. Then a case study is presented to illustrate how to estimate and the Bayesian network and how to infer dynamic perception of travel time using probe vehicle data collected in Beijing. The final section presents some concluding comments and discusses future directions.

3.1 The Problem

In most of the previous research, the researchers will assume or estimate the link travel time of each link first, and give an expect value or a probabilistic distribution. Then route choice behavior is analyzed using the estimated or assumed travel time with defined error terms or not. However, in practice, it is the travelers, who estimate the link travel time, and make route choice decision. In this research, we assume that, the drivers have knowledge about the state relationship of links, not the knowledge of the link travel time itself, according to the historical experiences. The link state relationship is presented by a Bayesian network (BN). This Bayesian network can be estimated by history data about link traffic state. The state of the links they have passed and the information they received can be treated as the evidences of the BN. At each decision node, travelers' perceptions of traffic states would be updated. According to the perception of link traffic state, the perception of link travel time can be estimated.

Figure 3.1 is a simple network to illustrate the problem proposed in this chapter. A traveler is driving from O to D. For each link n, the traffic state S_n can be divided into several categories. In this example, there are two categories, i.e. $\{1,2\}$, according to it is congested or not. The travel time of link n is denoted as T_n. It is assumed that T_n is independent with the traffic states of other links. At each traffic state, for link n, the probability distribution of travel time can be estimated using historical data, and denoted as $P(T_n|S_n)$. The probability distribution of S_n can also be estimated, and denoted as $P(S_n)$.

When the driver is at node O, if he has no information about the current link states of network, he will estimate the travel time distribution of each link, $P(T_n)$ according to the law of total probability:

$$P\left(T_n\right) = \sum_{S_n} P\left(T_n \mid S_n\right) \cdot P\left(S_n\right) \tag{3.1}$$

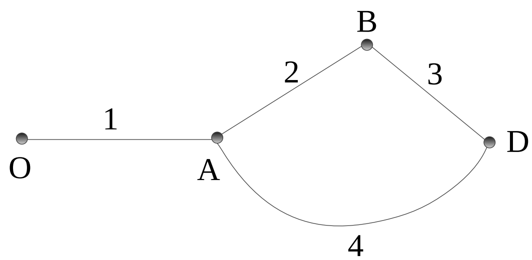


Figure 3.1 An Simple Network for Illustration

When the driver arrives node A, which is a decision node because he can switch his route here. Since he has passed link 1, it is assumed that the driver knows the traffic state of link 1. If the driver also has received the information about link 3, it is assumed that the driver knows the traffic state of link 3. Because T_n is assumed to be independent with the traffic states of other links, the travel time distribution of each link when S_1 and S_3 are known can be estimated:

$$\begin{aligned}
&P\left(T_1 \mid S_1, S_3\right) = P(T_1 \mid S_1) \\
&P\left(T_3 \mid S_1, S_3\right) = P(T_3 \mid S_3) \\
&P\left(T_2 \mid S_1, S_3\right) = \sum_{S_2} P(T_2 \mid S_2) \cdot P(S_2 \mid S_1, S_3) \\
&P\left(T_4 \mid S_1, S_3\right) = \sum_{S_4} P(T_4 \mid S_4) \cdot P(S_4 \mid S_1, S_3)
\end{aligned} \tag{3.2}$$

Therefore, the next problem is how to estimate $P(S_2 | S_1, S_3)$ and $P(S_4 | S_1, S_3)$. The traffic states relationship between links can be described by a BN about network traffic states, which will be introduced in the next section.

3.2 Some Background on Bayesian Network

Bayesian networks have been widely used to solve practical problems in different fields (Li et al., 2011; Ozbay and Noyan, 2006; Zhang and Taylor, 2006). A Bayesian network (BN) is an annotated directed acyclic graph that encodes a joint probability distribution over a set of random variables, and can be presented by a pair $B = (G, \Theta)$. G is a directed acyclic graph whose nodes correspond to the random variables $\mathbf{X}$, and whose links represent direct dependencies between variables (Jensen and Nielsen, 2007). The graph G, which is known as the structure of the BN, encodes independence assumptions: each variable x_i is independent of its non-descendants given its parents $\pi(x_i)$ in G. $\Theta = \{p(x_1 | \pi(x_1)), ..., p(x_n | \pi(x_n))\}$ is a set of parameters that quantifies the network, where $p(x_i | \pi(x_i))$ is the conational probability table attached to node x_i.

Figure 3.2 gives an example of BN. A scenario is assumed as following: Morikawa-sensei lives in Nagoya, where burglaries and earthquakes will also occur. Since Morikawa-sensei often travel for business, there is an alarm installed in his house. If there is a burglary or earthquake, the alarm has a possibility to be activated, and then both of his neighbors, Yamamoto-sensei and Miwa-sensei have a possibility to call him by phone.

In this example, there are 5 random events: burglary (*B*), earthquake (*E*), alarm (*A*), Yamamoto-sensei makes a call (*Y*) and Miwa-sensei makes a call (*M*). All of the five random events are denoted by five binary variables with possible values "*y*" and "*n*". The dependent relationships of the five variables can be described by the acyclic graph shown in Figure 3.2. According to the basic assumption of BN, *B* and *E* are independent on each other, while *M* and *Y* are dependent on each other. However, if the status of *A* is known, *M* and *Y* are independent. This graph which encodes the dependent relationships among variables can be determined by experts' knowledge or learned from data.

For each node, there is a conditional probability table (CPT), which called the parameters of BN. For example, the first row of CPT of *Y* means that, when the alarm is activated, the probability that Yamamoto-sensei calls Morikawa-sensei is 0.7. Most of the time, the parameters should be learned from data.

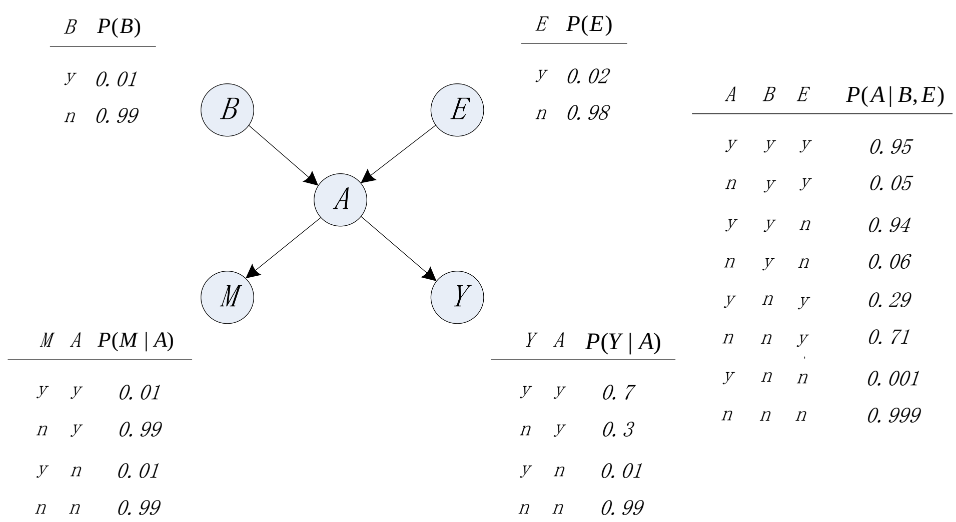


Figure 3.2 An Example of BN

3.3 BN About Network Traffic States

The structure of the BN proposed for modeling travel time perception is presented in Figure 3.3. It is assumed that there are *m* OD pairs and *n* road links in the traffic network. S_i is the traffic state of link *i* and Q_j is the level of traffic demand in OD pair *j*. *U* is a variable that represents the level of total traffic demand in the traffic network. If link *i* is on the route that belongs to the route choice set of OD pair *j*, Q_j is the parent of S_i. *U* is the parent of all the OD pair demands.

This structure is similar to the BN used in previous research (Castillo et al., 2008), because of the similar conditional independent assumptions. The link state nodes $\mathbf{S}=\{S_1,\ldots,S_n\}$ only have the OD demand level nodes $\mathbf{Q}=\{Q_1,\ldots,Q_m\}$ as parents, which implies conditional independence of all link traffic states given the OD demand. This assumption might appear to be unrealistic because the link traffic states, which can be treated as a function of link traffic flow, are corrected. However, accepting that $\mathbf{S}$ are related is not the same as saying $\mathbf{S} | \mathbf{Q}$ are correlated. In fact, as a special case, in the deterministic user equilibrium (DUE) models, under the condition that the OD demands $\mathbf{Q}$ are known, the link traffic states $\mathbf{S}$ are also deterministic, so $\mathbf{S} | \mathbf{Q}$ is not correlated.

The correlations between $\mathbf{Q}$ are obvious. In order to represent these correlations, a parent U is given to all the OD pair demands. U represents the level of total traffic demand in the traffic network. In particular, the OD demands have different characteristics during different periods of the day. In the traffic peak hours, the OD demand distributions are very different from that in the off-peak hours. So in this paper, U represents different time periods in the day.

Regarding the nodes, in this paper U is determined by definition and $\mathbf{S}$ can be estimated using probe data, so U and $\boldsymbol{S}$ are observed nodes. However, the OD demand cannot be observed directly, so $\mathbf{Q}$ are latent variables.

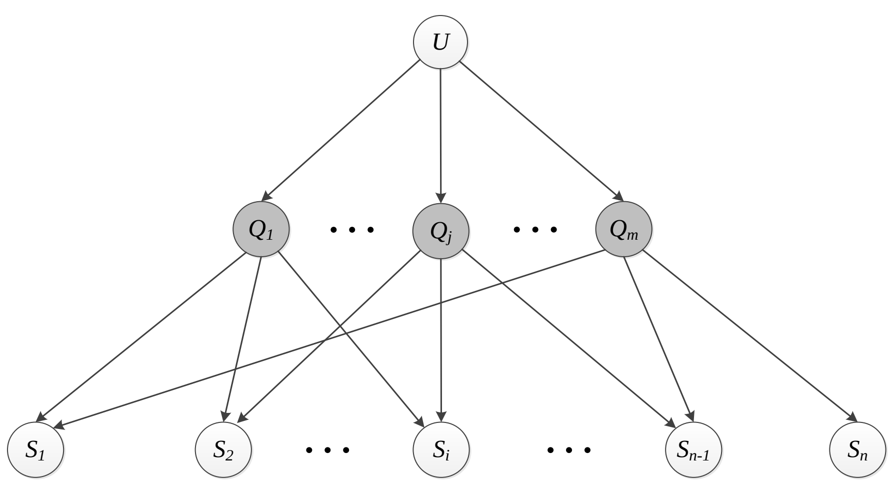


Figure 3.3 Structure of Proposed BN for Network Traffic States

The parameters of the BN are learnt from historical data in this study. Maximum likelihood estimation (MLE) is the most-used method of parameter learning. MLE can be used if the dataset is complete, but in the proposed BN, the OD demand nodes are latent, and there are also some missing traffic state data in practice. Therefore the Expectation Maximization (EM) algorithm (Lauritzen, 1995) is used in this paper. This algorithm basically alternates between two steps:

Expectation Step: complete the dataset by using the current parameter estimates $\widehat{\Theta}$ to calculate expectations for the missing data.

Maximization Step: use the completed dataset to find a new MLE, $\widehat{\Theta}'$, for the parameters. This estimate is then used in the next iteration of the expectation step.

This algorithm has been described abundantly in the literature, so details will not be given in this paper.

3.4 Inferences in BN as Travel Time Perceptions

As explained at the beginning of this section, the problem of modeling travel time perception is to estimate the traffic states of other links when the states of some links are known. For the BN in Figure 3, this means computing the posterior marginal of traffic states of unobserved links $\mathbf{S} \setminus \mathbf{S}'$, where $\mathbf{S}'$ are the set of traffic states that are observed by travelers. For each $\mathbf{S}_i \in \mathbf{S} \setminus \mathbf{S}'$, if X $\mathbf{Q}$ denotes $\{\mathbf{Q}, U, \mathbf{S}\}$ and $\mathbf{e} = \{U, \mathbf{S}'\}$ is the evidence available to the driver, then the posterior marginal distribution of S_i is:

$$P(S_i \mid \mathbf{e}) = \frac{P(S_i, \mathbf{e})}{P(\mathbf{e})} = \frac{\sum_{\mathbf{X} \setminus \{S_i, \mathbf{e}\}} P(\mathbf{X})}{\sum_{\mathbf{X} \setminus \mathbf{e}} P(\mathbf{X})} \tag{3.3}$$

where $P(\mathbf{X})$ can be calculated using the parameters of the BN:

$$P(\mathbf{X}) = \prod_{x_i \in \mathbf{X}} P(x_i \mid \pi(x_i)) \tag{3.4}$$

In order to update the posterior marginal distribution efficiently, certain algorithms are needed. The inference algorithm used in this paper is a junction tree algorithm (Shafer and Shenoy, 1990). A junction tree algorithm can be seen as the mother of all exact inference algorithms for BNs, and has been widely explained in the literature (Kjærulff and Madsen, 2008). With this algorithm, the independence properties of the BN are analyzed to establish a set of clusters and to construct a corresponding junction tree over the clusters. When updating the BN, it is necessary to calculate the posterior distributions of more than one variable. Using a junction tree to solve the inference problem, some processes can be shared, so efficiency is improved.

When the marginal distribution of traffic states of links S are updated, then the distribution of link travel time can be estimated. In this study, we only estimate the mean of travel time, which can then be directly input into most route choice models:

$$M(T_i \mid \mathbf{e}) = \sum_{s=1}^{|sp(S_i)|} P(S_i = s \mid \mathbf{e}) M(T_i \mid S_i = s) \tag{3.5}$$

where $M(T_i \mid \mathbf{e})$ is the mean travel time on link i. $M(T_i \mid S_i = s)$ can be estimated directly using probe data. If there is no information provided, then $\mathbf{e} = \varnothing$.

3.5 Case Study

The data used in this study is provided by CENNAVI company (http://www.cennavi.com.cn/en/index.php). The raw data is collected in real time using more than 10,000 taxies in Beijing.

The raw probe vehicle GPS data is a series of location points with coordinates. CENNAVI did the map-matching and travel time estimation processing and provided us with a one-week (2011.7.25 to 2011.7.31) historical data set giving link traffic states and three months (2011.6 to 2011.8) of trajectory data for part of the urban area of Beijing. The historical traffic state data consists of the link travel time for each link estimated at intervals of five minutes. The trajectory data gives the location and service state of the taxis collected about every minute.

For the purpose of illustration, only a small network is used in this study. This network is shown in Figure 3.4 (with only the main roads considered and shown). There are eight nodes and nine links in this network and all links are bidirectional.

With the probe vehicles being taxis, only data from occupied taxis are used in the study. We only consider trips from node A to B, of which there are 274 observations during the data collection period. All of these observations share the three routes shown in Figure 3.4.

The link traffic states are determined according to travel speed estimated using the probe data. According to CENNAVI's standards, which are similar to those of most traffic information providers, traffic states are divided into three categories:

Congested: $0 < Speed \le 20km/h$

Slow: $20km/h < Speed \le 40km/h$

Uncongested: $Speed > 40km/h$

The means of link travel time under different traffic states are estimated using the historical data and are shown in Table 3.1.

It is assumed that there are four OD pairs in the network {(A, B), (B, A), (C, D), (D, C)} with the following routes:

(A, B): {1+,7+,4+,5+,6+}, {1+,2+,8+,5+,6+}, {1+,2+,3+,9+,6+}

(B, A): {6-,5-,4-,7-,1-}, {6-,5-,8-,2-,1-}, {6-,9-,3-,2-,1-}

(C, D): {4+,5+, 9-, }, {4+, 8-,3+}, {7-,2+,3+}

(D, C): {3-,2-,7+}, {3-,8+, 4-}, {9+, 5-,4-}

The sign given for each link denotes the direction (with left to right and down to up being positive). In this study, as already noted, we only consider trips from A to B, so only the positive direction is considered for simplification.

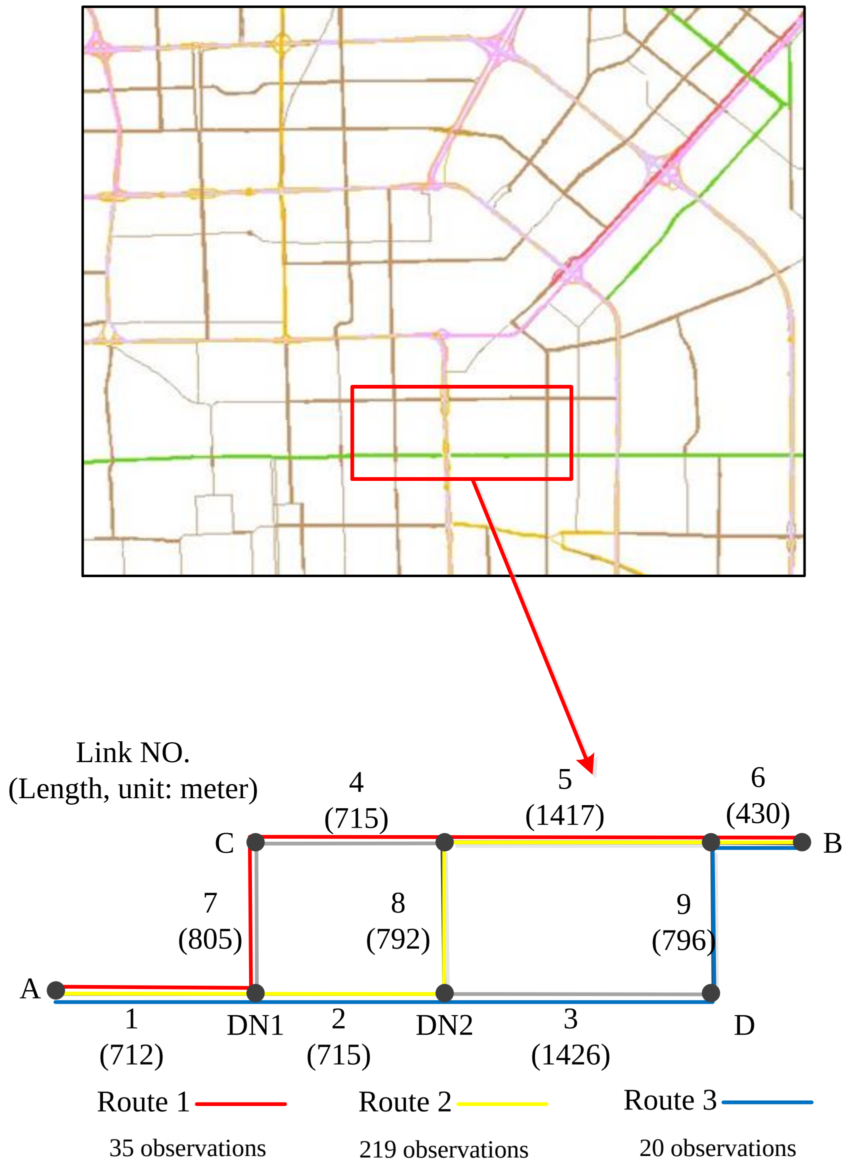


Figure 3.4 Network Used for Study

Table 3.1 Estimated Mean of Travel Time under Different Traffic States (Unit: s)

Link NO.	State 1 (Congested)	State 2 (Slow)	State 3 (Uncongested)	All states
1	280.747	99.578	71.505	126.143
2	178.424	75.471	49.622	103.416
3	317.803	157.764	110.803	193.174
4	170.526	88.521	61.914	106.595
5	377.009	179.737	134.946	198.142
6	174.454	58.164	40.626	76.535
7	215.334	94.7673	63.184	123.511
8	205.168	110.306	70.788	122.614
9	220.940	97.029	64.920	124.514

The BN structure developed for this traffic network is shown in Figure 3.5. The method of determining this structure is explained in Section 3.3. Variable *U*, which indicates the level of total traffic demand in the traffic network, depends on the time of day. In this example, variable *U* has four states: morning peak hours (state 1; 7:00–9:00), flat hours (state 2; 9:00–18:00), evening peak hours (state 3; 18:00–21:00), and valley hours (state 4; 21:00–7:00). $\mathbf{Q} = \{Q_1, ..., Q_3\}$ are the variables representing the traffic demand level of the OD pairs {(A,B), (C,D), (D,C)} respectively. **Q** are latent variables and the count of states is not determined, since the complexity of a BN increases rapidly as the count of states increases. Therefore, it is assumed that **Q** has three states in this example. The variables $\mathbf{S} = \{S_1, ..., S_9\}$ represent the traffic states of links. If link *i* is on the route between OD pair *j*, there is a directed arc from Q_i to S_j in the BN.

The parameters of the BN in Figure Figure 3.5 are obtained from the data using the EM algorithm implemented in BNT, an open source Matlab package (Murphy, 2001).

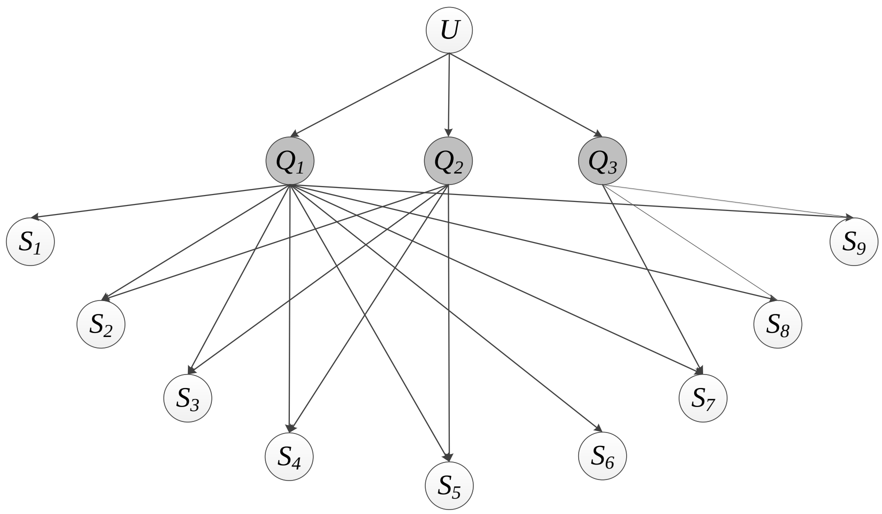


Figure 3.5 BN for the Illustrative Example

Then, according to Section 3.4, the updating of the driver's travel time perception for each link is treated as the inference problem to be solved using the developed BN. At each decision node, the driver's updated knowledge is described by the updated evidence entered into the BN. As a simplification, information about links not included in the studied network is ignored.

Therefore, when a driver arrives at node A, the only information that can be obtained is for the current time period. For instance, assuming arrival is at 13:30 (during the flat hours), then the evidence is $\mathbf{e}=\{U=2\}$. The driver then continues the trip and arrives at DN1. Because the driver has passed link 1, the traffic state of link 1 is known. Assuming link 1 is uncongested, then the evidence is updated to be $\{U=2, S_1=1\}$. Similarly, when the driver arrives at DN2, if link 2 is also uncongested, the evidence is updated to be $\{U=2, S_1=1, S_2=1\}$. From GPS data, the traffic state of passed links during each observation can be estimated, then using Equations (3.3–3.5), we can obtain the dynamic travel time perception for the estimation of choice models (shown in Table 3.2).

If we assume a simple rule that, a driver will choose the path with shortest perceiving travel time and make route choice at every node then:

At the origin, the perceiving travel time of the three routes is 624, 620 and 616, respectively, so the driver will choose route 3 here;

At node DN1, the perceiving travel time of the two available routes (2 and 3) is 513 and 521, respectively, so the driver will switch to route 2 here;

At node DN2, the perceiving travel time of the two available routes (2 and 3) is 418 and 427, respectively, so the driver will still use route 2.

Table 3.2 Updated Mean Travel Time on Links

Link NO.	Mean travel time when **e**=			
	Æ	$\{U=2\}$	$\{U=2, S_1=1\}$	$\{U=2, S_1=1, S_2=1\}$
1	121.737	121.924	-	-
2	96.985	97.038	102.010	-
3	190.455	190.474	197.436	200.226
4	104.165	104.212	112.534	116.023
5	194.807	194.718	198.148	198.870
6	79.341	79.374	89.936	94.118
7	124.741	124.683	126.685	128.107
8	127.601	127.734	123.537	125.015
9	128.175	128.181	131.748	133.589

3.6 Conclusions and Future Directions

Travelers' travel time perceptions are important inputs of route choice models. However there are limited previous researches about this. This chapter proposes a Bayesian networks (BN) based approach to model travelers' travel time perceptions for route choice analysis. The dynamic travel time perception process is described by the inference problem of BN. About the first example in Section 1.3:

- **A:** I think Route 1 will take 5 minutes;
  B: I think Route 1 will take 10 minutes, because I just traveled on road segment 1, and find it is congested.

Their heterogeneity in perceptions can be considered by inputting different evidences to the BN for inferences.

As an illustrative example, a BN about a small part of the road network of Beijing is estimated using the probe data. Using the estimated BN and the mean travel time estimated under different traffic states on each road link, a dynamic route choice process is described with a simple route choice model. Through this example, it can be found that travelers' dynamic perceptions of travel time during the trips can be considered explicitly using this approach.

This study is only a first try to model travelers' heterogeneity in perceptions using BN. There are still some problems should be solved in the future for widely applied in real cases. At first, in this chapter, the case study is on a small network. How to give a structure of BN on a large road network should be considered in the future. Second, in this study, we assume the homogeneity in the knowledge about the link state relationships. In the future, this assumption should be relaxed by adding some characteristics of the travelers to the BN structures. At last, it is also possible to model travelers' day-to-day learning process by the on-line parameter learning algorithms of BN in the future.

References

Adler, J.L., Blue, V.J. (1998) Toward the design of intelligent traveler information systems. *Transportation Research Part C: Emerging Technologies* 6, 157-172.

Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H.N., Mishalani, R. (2002) Real time simulation of traffic demand-supply interactions within dynamit. *Applied optimization* 63, 19-34.

Ben-Elia, E., Shiftan, Y. (2010) Which road do i take? A learning-based model of route-choice behavior with real-time information. *Transportation Research Part A: Policy and Practice* 44, 249-264.

Bierlaire, M., Frejinger, E. (2008) Route choice modeling with network-free data. *Transportation Research Part C: Emerging Technologies* 16, 187-198.

Castillo, E., Menéndez, J.M., Sánchez-Cambronero, S. (2008) Predicting traffic flow using bayesian networks. *Transportation Research Part B: Methodological* 42, 482-509.

Gao, S., Frejinger, E., Ben-Akiva, M. (2010) Adaptive route choices in risky traffic networks: A prospect theory approach. *Transportation Research Part C: Emerging Technologies* 18, 727-740.

Jensen, F.V., Nielsen, T.D. (2007) *Bayesian networks and decision graphs*. Springer Verlag.

Joel L, H. (1984) The stability of stochastic equilibrium in a two-link transportation network. *Transportation Research Part B: Methodological* 18, 13-28.

Kjærulff, U.B., Madsen, A.L. (2008) *Bayesian networks and influence diagrams*. Citeseer.

Lauritzen, S.L. (1995) The em algorithm for graphical association models with missing data. *Computational Statistics & Data Analysis* 19, 191-201.

Li, D., Cheng, L., Ma, J. (2011) Incident duration prediction based on latent gaussian naive bayesian classifier. *International Journal of Computational Intelligence Systems* 4, 345-352.

Li, D., Miwa, T., Morikawa, T. (2013) Dynamic route choice behavior analysis considering en route learning and choices (forthcoming). *Transportation Research Record: Journal of the Transportation Research Board*.

Mahmassani, H.S. (2001) Dynamic network traffic assignment and simulation methodology for advanced system management applications. *Networks and Spatial Economics* 1, 267-292.

Morikawa, T., Miwa, T. (2006) Preliminary analysis on dynamic route choice behavior: Using probe‐vehicle data. *Journal of advanced transportation* 40, 140-163.

Murphy, K. (2001) The bayes net toolbox for matlab. *Computing science and statistics* 33, 1024-1034.

Ozbay, K., Noyan, N. (2006) Estimation of incident clearance times using bayesian networks approach. *Accident Analysis & Prevention* 38, 542-555.

Ramming, M.S. (2001) Network knowledge and route choice. Massachusetts Institute of Technology.

Shafer, G.R., Shenoy, P.P. (1990) Probability propagation. *Annals of Mathematics and Artificial Intelligence* 2, 327-351.

Zhang, K., Taylor, M.A.P. (2006) Effective arterial road incident detection: A bayesian network based algorithm. *Transportation Research Part C: Emerging Technologies* 14, 403-417.

Chapter 4

Modeling Heterogeneity in Processes

Recall the first example in Section 1.3:

- **A:** Yesterday, I chose Route 1 at node O, but I found Route 1 is congested, then I switched to Route 2 at node T;

  B: Yesterday, I chose Route 2 at node O, and didn't change my mind during the trip.

To consider the scenario shown in the above example, it is in fact to consider the en-route choices. Route choice models can be categorized as static models (SM) and dynamic models (DM), where SMs only consider pre-trip decisions, while DMs also take into account en-route decisions (Morikawa and Miwa, 2006).

The development of a dynamic route choice model presents two additional problems as compared with static models, as follows:

- The possibility of making en-route route choices;
- The en-route updating of the driver's perception of route characteristics.

Looking at the first of these problems, a natural way to build dynamic route choice models is to have a sequence of static choice models, one at each decision node, where the characteristics of alternative routes reflect updated information (Gao et al., 2010). Some dynamic traffic assignment models, e.g. DynaMIT (Ben-Akiva et al., 2002) and DYNSMART (Mahmassani, 2001), employ dynamic route choice models constructed this way. This kind of model is referred to as a deterministic dynamic route choice model (DDM) in this paper. Although easily applied in practice, it is obviously not necessary for a driver to make a route choice decision at every decision node. According to the preliminary analysis by Morikawa and Miwa (2006) using probe vehicle data (Morikawa and Miwa, 2006), a SM will sometimes give better numerical results than a DDM in a small network. They proposed a threshold-based model to account for the probability of making a route choice decision at each decision node. This kind of dynamic model is referred to here as a stochastic dynamic model (SDM). However, in their research, a driver's choices are related only to the characteristics of the decision nodes and the links driven. Route choice decisions and expected utility for links not driven are not considered. As a consequence, this model is unable to consider which route will be chosen at the previous node and the utilities of alternative routes at the current decision node.

The second problem is, in fact, the problem of modeling a driver's en-route learning process. The characteristics of a route are various and can be divided into two classes: deterministic characteristics and uncertain characteristics. The driver can have perfect knowledge of the deterministic characteristics before making a route choice decision; they include the number of intersections, the proportion of the route on expressway and toll charges. On the contrary, it is impossible for the driver to have perfect knowledge of uncertain characteristics prior to making a route choice decision. Travel time, which is an important input into route choice models, is one uncertain route characteristic. In most previous studies, the travel time for each route at the time of making a route choice is estimated by the researchers according to the current state of traffic on the network and it is assumed that drivers are able to obtain perfect knowledge of it (Bierlaire and Frejinger, 2008; Xu et al., 2011). However, in reality, a driver cannot know the true traffic state of the whole network and can only infer it from past experience, external information and the state of traffic on links already passed during the current trip. Recently, there has been much research that takes external information and past experience into account using learning-based models (Ben-Elia et al., 2008; Ben-Elia and Shiftan, 2010). However, these works did not explicitly consider a driver's experience on links already passed during the current trip.

In this chapter, corresponding to the first problem, a new random utility-based model is proposed for considering the route choice and decision making choice problem together. Corresponding to the second problem, the Bayesian networks based model which proposed in last chapter is used for modeling a driver's perception of travel time that explicitly considers the driver's en-route inference process. These two models are combined and a simulation carried out in a case study using GPS data from probe vehicles.

The remainder of this chapter is organized as follows. First, a simple example is used to illustrate the dynamic route choice problem in the next section. Then the process-based route choice model is specified in the following section. A case study based on probe vehicle data is carried out in Section 4.3. The final section presents some concluding comments and discusses possible future directions.

4.1 The Problem

In this section, before giving the formal specifications, we illustrate the dynamic route choice model using a simple example. Figure 4.1 is a simple network with five nodes and six links. There are three available routes from O to D: route O-A-B-D (route 1), route O-A-D (route 2) and route O-D (route 3). We now consider the route choice problem for travel from O to D.

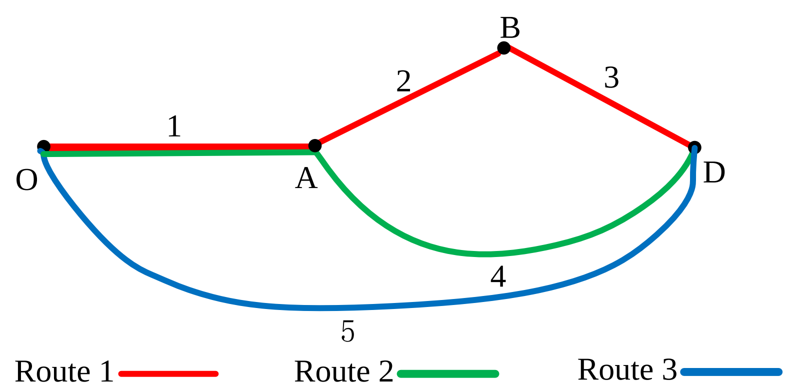


Figure 4.1 A Simple Illustrative Network

The SMs, which only consider a driver's decision-making at the origin, assume that the probability of actually using route 1 (P_{1S}) is equal to the probability of choosing route 1 at the origin (P_1^o):

$$P_{1S} = P_1^o \tag{4.1}$$

However, upon arrival at node A, a driver has a second chance of making a route choice decision. If a route choice decision is made here, the probability of choosing route 1 at node A is P_1^A.

The DDMs assume that a driver will make a decision at each decision node. Decision nodes are defined as the set of intersections at which the driver can change the route choice plan. In this example, because a driver can only change the route choice plans at node A, there is one decision nodes (i.e. node A). Therefore, the probability of actually using route 1 (P_{1D}) is

$$P_{1D} = (P_1^o + P_2^o) \times (P_1^A) \tag{4.2}$$

However, it is not necessary for the driver to make a decision at every decision node. The SDMs assume that there is a probability at each node that the driver will change the route choice again. In this simple example, at node A, for each driver, there will be a binary choice as to whether to make a route choice decision or not. Let q^A denote the probability that the driver will make a route choice decision at node A, then the probability that the driver actually uses route 1 is:

$$\begin{aligned} P_1 &= (1 - q^A) P_1^o + q^A P_1^A (P_1^o + P_2^o) \\ &= (1 - q^A) P_{1S} + q^A P_{1D} \end{aligned} \tag{4.3}$$

It can be found that the SM and the DDM are the two limiting cases of SDM when q^A approaches 0 and 1, respectively. Therefore, the first issue facing dynamic route choice modeling is how to obtain q^A. To solve this binary choice problem, in next section, a random utility-based model will be proposed.

If we use the Multinomial Logit (MNL) model to analyze a driver's route choice behavior over this simple network, and the systematic utility of each route is equal to the sum of the systematic utility of links on this route, then

$$P_{1D} = (P_1^o + P_2^o) \times (P_1^A)$$
$$= \frac{\exp(v_1^o + v_2^o + v_3^o) + \exp(v_1^o + v_4^o)}{\exp(v_5^o) + \exp(v_1^o + v_2^o + v_3^o) + \exp(v_1^o + v_4^o)} \times \frac{\exp(v_2^A + v_3^A)}{\exp(v_2^A + v_3^A) + \exp(v_4^A)} \tag{4.4}$$

where v_i^k is the systematic utility of link i when the driver arrives at node k. Considering the special case in which the characteristics of links do not change during the trip (i.e. $v_i^A = v_i^O$), then

$$P_{1D} = \frac{\exp(v_1^o + v_2^o + v_3^o) + \exp(v_1^o + v_4^o)}{\exp(v_5^o) + \exp(v_1^o + v_2^o + v_3^o) + \exp(v_1^o + v_4^o)} \times \frac{\exp(v_2^o + v_3^o)}{\exp(v_2^o + v_3^o) + \exp(v_4^o)}$$
$$= \frac{\exp(v_1^o + v_2^o + v_3^o))}{\exp(v_5^o) + \exp(v_1^o + v_2^o + v_3^o) + \exp(v_1^o + v_4^o)} = P_{1S} \tag{4.5}$$

This implies that the different results given by SM and DM result from a driver's en-route learning process under the MNL framework. Travel time is an important input for route choice models, so the BN-based model proposed in last chapter is applied in this chapter to model explicitly the process of updating travel time perception.

As shown in this simple example, a process of dynamic route choice can be discribed by Figure 4.2. At the origin, a traveler must make a route choice decision. During the trip, the driver's knowledge is updated. Upon arrival at decision nodes en-route, the driver has the chance to make another route choice decision. To calculate q^n, the probability of making a decision at node n, a binary Logit model is used with two alternatives: making a decision or not. In the next section, we will give a formal specification of a process-based route choice model.

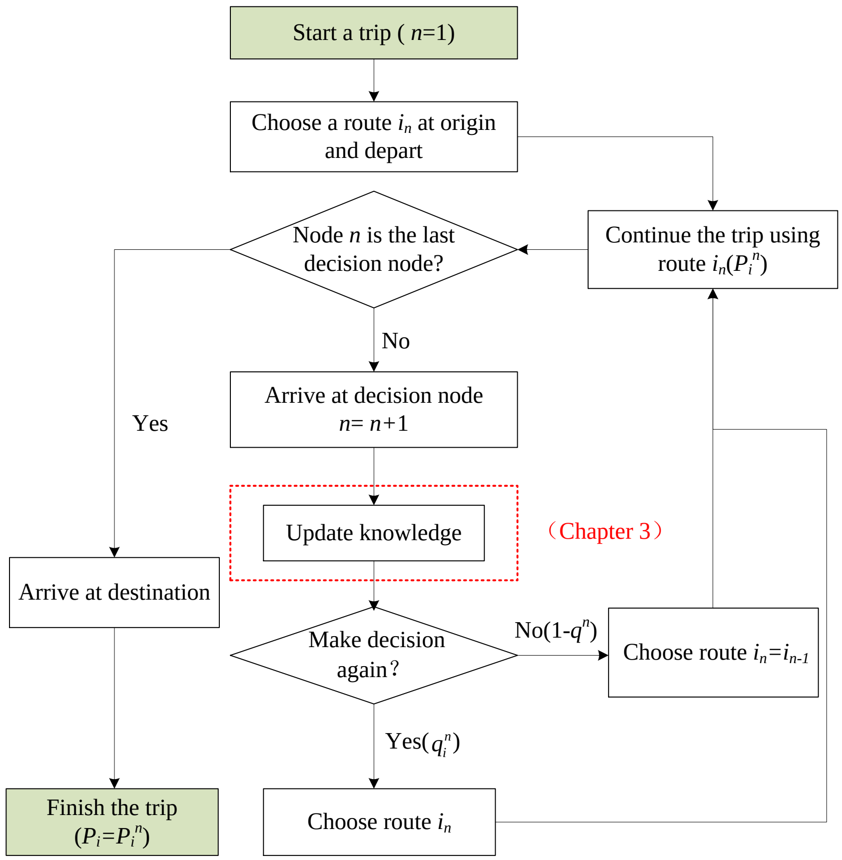


Figure 4.2 The Process of Dynamic Route Choice

4.2 A Process-Based Route Choice Model

In this section, a process-based route choice model is proposed to consider the heterogeneity in processes. The dynamic choice process is defined as the sequence of choices during a trip, including the route choices (both pre-trip and en-route choices) and the choices of making a route choice again at every decision node. Considering the example shown in last section, there are 7 possible dynamic choice processes (coded as *D1* to *D7*) as shown in Figure 4.3.

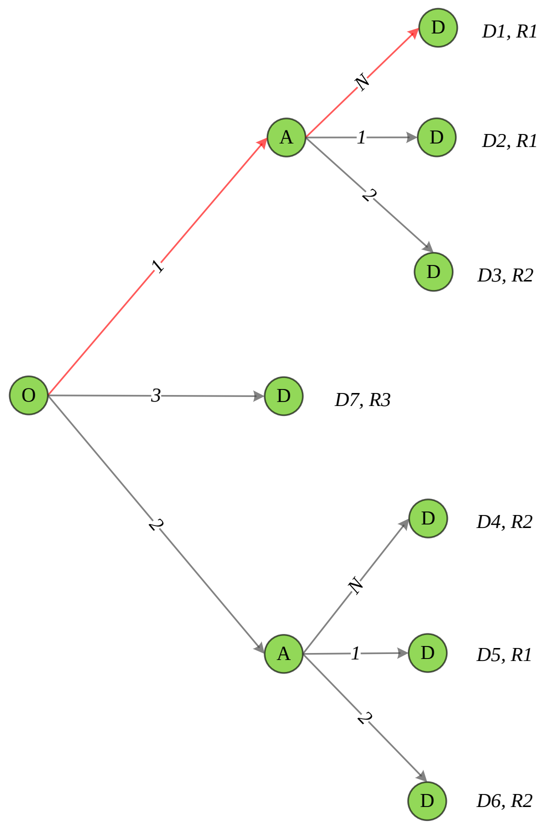


Figure 4.3 Possible Dynamic Choice Processes for the Illustrative Example

As an example, process *D1* (denoted by the red arrows) means that: the driver chooses route 1 at the origin; then at node A, the driver does not make a new route choice. There are 3, 3 and 1 possible dynamic choice processes that result in observation of route 1, route 2 and route 3, respectively. These are labelled by *R1*, *R2* and *R3* in Figure 4.3. Then the probability that the driver actually uses route 1 is:

$$P_1 = \sum_{D \in C_1} P(D) \tag{4.6}$$

where C_1 is the set of dynamic choice processes that result in observed route 1. $P(D)$ is the probability of dynamic process *D*. In the following part of this section, we will discuss how to model *P*(*D*).

Let S_n denote the traffic state for observation *n*, which is defined as the driver's dynamic perception of route characteristics during the trip. Then the probability of the *h*-th dynamic route choice process D_h is

$$P(D_h \mid S_n) = \prod_{t \in T_h} H(j_t^{D_h} \mid S_n, D_h) \cdot \begin{cases} P(k_t^{D_h} \mid S_n), \ if \ \ j_t^{D_h} = Y \\ \quad 1 \qquad , \ if \ \ j_t^{D_h} = N \end{cases} \tag{4.7}$$

where T_h is the set of decision nodes that the driver will pass if the choice process is D_h, $k_t^{D_h}$ is the route chosen by the driver at decision node *t* following process D_h, and $j_t^{D_h}$ is the driver's decision making choice at decision node *t* in following dynamic process D_h. $P(k_t^{D_h} \mid S_n)$ and $H(j_t^{D_h} \mid S_n, D_h)$ are the probability of route choice and decision making choice, respectively.

Therefore, the dynamic route choice problem can be divided into two subsidiary choice problems: route choice and decision making choice. In the rest of this section, we will model these two subsidiary choices.

Because alternative routes overlap, the Multinomial Logit (MNL) model is not appropriate for route choice analysis. The Path-Size model has been proposed to deal with the overlap problem while maintaining the computational simplicity of the logit form(Ben-Akiva and Bierlaire, 1999; Ramming, 2001). Let $U_{kt}^{D_h}$ be the utility of route *k* at decision node *t* when following dynamic process D_h, where this consists of an observed utility $V_{kt}^{D_h}$ and an unobserved component $\varepsilon_{kt}^{D_h}$, such that

$$U_{kt}^{D_h} = V_{kt}^{D_h} + \varepsilon_{kt}^{D_h} \tag{4.8}$$

The unobserved components of the alternatives are assumed to be independent and identically distributed (i.i.d.) Gumbel. The observed utility is assumed to be a linear relationship between attributes and tastes. To consider the correlation of alternative routes, a correction term is added to the utility of alternative routes, such that

$$V_{k_t}^{D_h} = \beta x_{kt} + \beta_{PS} \ln(PS_{kt}) \tag{4.9}$$

where β is a set of parameters of route characteristics, x_{kt}; β_{PS} is the parameter of path-size; and the correction term PS_{kt} is formulated as

$$PS_{kt} = \sum_{a \in \Gamma_k} \left(\frac{l_a}{L_k} \right) \frac{1}{\sum_{l \in C_t^{D_h}} \delta_{al}} \tag{4.10}$$

where Γ_k is the set of links in route *k*; is the length of link a; and δ_{al} is the link-path incidence dummy (that is, 1 if path l uses link a and 0 otherwise).

Then the probability that route $k_t^{D_h}$ is chosen in choice situation *t* is given by the Path-Size model as follows:

$$P(k_t^{D_h} \mid S_n) = \frac{e^{\beta x_{kt} + \beta_{PS} \ln(PS_{kt})}}{\sum_{l \in C_t^{D_h}} e^{\beta x_{lt} + \beta_{PS} \ln(PS_{lt})}} \tag{4.11}$$

where $C_t^{D_h}$ is the choice set at decision node t, following choice process Dh.

A random utility based model is developed for the decision making choice. This is given as

$$H(j_t^{D_h} \mid S_n, D_h) = \begin{cases} \dfrac{1}{1+e^{V_1 - V_0}}, \; if \;\; j_t^{D_h} = N \\ \dfrac{1}{1+e^{V_0 - V_1}}, \; if \;\; j_t^{D_h} = Y \end{cases} \tag{4.12}$$

where V_1 and V_0 are the observed utilities of making a decision and not making a decision, respectively, which depend on S_n, D_h and *t*.

If the traveler has already chosen route *i* before arriving at node *t* and, at this node, the traveler's knowledge is updated, then the benefit of making a decision here (B_{1t}) is the expected maximum utility of the route choice model using the updated knowledge, which is not related to *i*. Since a MNL model is used for route choice, B_{1t} is the logsum of the systematic utility of the routes available at node *t* ($V_{kt}^{D_h}$):

$$B_{1t} = \ln \sum_{l \in C_t^{D_h}} \exp(V_{lt}^{D_h}) \tag{4.13}$$

In order to consider the preference of the driver for the previous route choice decision, in a similar manner to the concept of cognitive cost in reference (Gao et al., 2011), the cost of making a decision, *DC*, is defined:

$$DC = \beta_{dc}{}' x_{dc} \tag{4.14}$$

where x_{dc} is a set of explanation variables that affect the driver's cost of making the decision.

If a decision is not made at node *t*, the driver will continue to use route *i*. The benefit of not making a decision in condition that route $k_{t-1}^{D_h}$ has been chosen beforehand, is the updated expected utility of route $k_{t-1}^{D_h}$:

$$B_{0t} = \text{Mean}(V_{kt}^{D_h} + \varepsilon_{kt}^{D_h}) = V_{kt}^{D_h} \tag{4.15}$$

Then, the systematic utilities of the decision making choice under the condition S_n, D_h and *t* can be specified as follows:

$$\text{Making decision: } V_1 = DC + \beta_{\text{Benefit}} B_{1t} \tag{4.16}$$

$$\text{Not making decision: } V_0 = \beta_{\text{Benefit}} B_{0t} \tag{4.17}$$

where β_{Benefit} is the parameter of benefit.

4.3 Case Study

In this section, we do a case study using the same data set and test network shown in last chapter. Since the data description and the modeling of travel time perceptions have already introduced in last chapter, we only describe the model estimation and analysis work in this section.

We consider two variables in the systematic utility of the route choice model: travel time and intersection density. For the specification of *DC*, we consider two plans:

$$\begin{aligned} &\text{Constant: } DC = b_0 \\ &\text{Related to other variables: } DC = b_0 + b_1 \times d + b_2 \times w \end{aligned} \tag{4.18}$$

Where d is the distance from the origin/mean distance between A and B and w is an index related to the width of the link crossing the current link at the decision node.

Route travel times are estimated according to the departure time and the real travel time on the passed links using the BN developed in last chapter.

In order to evaluate the proposed model, we also estimate a SM and a DDM using the same data. The estimation results are shown in Table 4.1.

Table 4.1 The Estimation Results of Route Choice Models

Parameter	*DDM*	*SM*	*SDM 1*	*SDM 2*
Travel time (min)	-7.46 (-7.19)	-9.08 (-6.22)	-14.56 (-7.95)	-12.70 (-5.21)
Intersection Density (count/km)	-1.63 (-5.23)	-0.71 (-2.71)	-2.27 (-6.27)	-1.78 (-3.55)
LnPS	0.42 (2.15)	0.36 (1.03)	0.61 (0.77)	0.53 (0.92)
Benefit	-	-	5.65 (0.64)	6.62 (1.02)
Constant part of DC	-	-	-4.36 (-0.64)	-3.31 (-0.80)
D (%)	-	-	-	-2.51 (-0.77)
w	-	-	-	7.03 (0.91)
LL at 0	-301.02	-301.02	-301.02	-301.02
LL at convergence	-272.13	-270.57	-258.36	-246.18
Adjusted r^2	0.09	0.09	0.13	0.16

We can obtain several findings from these estimation results. From the aspect of goodness of fit, SDM 2 gives the best result. We find that, in this case, the DDM has a slightly lower LL at convergence than the SM. This is consistent with the results obtained in previous research (Morikawa and Miwa, 2006) and indicates the need to consider a driver's en-route decision-making choice.

In SDM 1, the parameter of the constant part of *DC* is negative. This confirms the assumption that drivers prefer to continue on the initially planned route without making change.

The negative sign of parameter of *d* implies that drivers tend to make a new route choice when they are far from the destination. This is consistent with the results of previous research (Morikawa and Miwa, 2006).

The positive sign of parameter of *w* indicates that drivers tend to change their route choice plans when the spatial scale of the decision node is large.

In order to illustrate potential biases introduced by over-simplifying assumptions about decision-making choices, SDM 1 is used to predict route choices with the same choice scenario resulted the perceived traffic states shown in Table 3.2: A driver arrives from node A at 13:30 (during the flat hours), assuming link 1 and 2 are both uncongested. SDM 2 is not used here because we are interested in the impact of decision costs as a whole on route choice.

We let *DC* change by different multiples of β_{Benefit} (5.65), keeping the other parameters the same as in Table 4.1. The outputs are shown in Figure 4.4. In this figure, P^n is the probability that route *n* is used at last. q^n is the probability of decision making at node *n*, which can be calculated using Equation (4.19):

$$q^n = \sum_{i \in C_n} P_i^{n-} q_i^n \tag{4.19}$$

where P_i^{n-} is the probability that route i is chosen just before the driver arrives at node n. When the driver leaves node n-1, the set of available routes is C_{n-1}. When the driver arrives at node n, the set of available routes at decision node n is $C_n \subseteq C_{n-1}$. It can be known that routes belonging to $C_{n-1} \setminus C_n$ are not chosen when the driver leaves from node *n*-1.Therefore, P_i^{n-} can be calculated using Equation (4.20):

$$P_i^{n-} = \frac{P_i^{n\text{-}1}}{\sum_{j \in C_n} P_j^{n\text{-}1}} \tag{4.20}$$

We find that the decision to make an en-route choice is in fact a tradeoff between the benefit of the perceived change in route characteristics and the cost of making the decision. When DC / b_{Benefit} is large enough, SDM will approach SM, while if DC / b_{Benefit} is negative and small enough, SDM will approach DDM.

It should be noted that the probability of making a route choice again at DN2 (*q2*) does not decrease as a monotonic function of DC / b_{Benefit}. This is because, in the proposed model, the decision-making choice is dependent on route choice. The decision-making choice at DN1 will affect the route choice, and then this will affect the decision-making choice at DN2.

The benefit parameter can indicate a driver's sensitivity to the benefit. According to this study, even taxi drivers (who are more sensitive to benefit than private drivers) cannot be assumed to make a decision at approximately every node. Further, from Figure 4.4, it is found that the share of trips taken by each route will change significantly as the probability of making a decision changes. These results indicate the necessity for considering a driver's en-route choice to make a new decision.

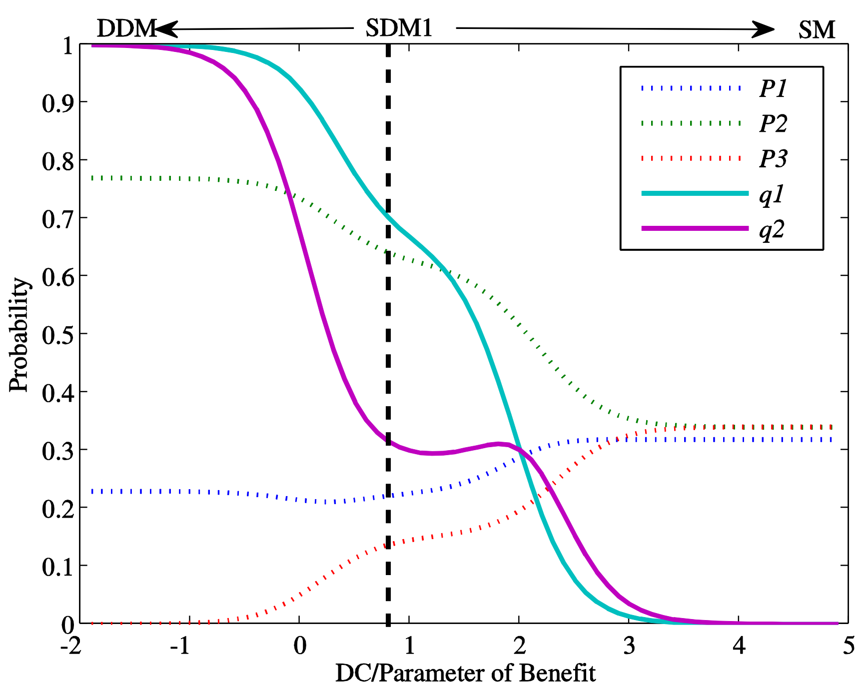


Figure 4.4 Decision Cost Impact on Route Choice

4.4 Conclusions and Future Directions

In this chapter, we have proposed a process-based method for analyzing dynamic route choice behavior. As compared with static models, two additional problems are considered in the dynamic route-choice model: the en-route updating of driver knowledge and the en-route choice to make a new route decision at a decision node. Upon starting a trip, a driver has knowledge about the traffic state from past experience on the links to be passed and external information. This availability of knowledge is described explicitly using a model based on a Bayesian network. Because of knowledge updating, a driver's route choice decisions will differ according to whether an en-route choice is made to change the route at each decision node. A random utility-based method is proposed for modeling the choice to make a decision and the choice of route at each decision node simultaneously.

Using the proposed dynamic route model, a case study over a small network is carried out. The model is estimated and compared with conventional models using probe vehicle data. The results show that drivers do not tend to make route choice decisions at all decision nodes. The probability of making an en-route choice is related to a driver's sensitivity to benefit and the cost of making the decision. The absolute value of the decision cost is positively correlated with distance to the origin and negative correlated with the spatial scale of the intersection at the decision node.

The case study in this chapter is on a small network. When applied on a large network, the number of the possible processes will be huge. Therefore, in the future, some algorithms should be developed to generate a set of choice processes when the proposed model is applied on a large network, which are similar to choice set generation methods for path-base route choice model. About the decision making choice, in this study, travelers' attributes cannot be considered because of the lacking of data. In the future, the traveler specific attributes should be considered in the utility function when data is available.

References

Ben-Akiva, M., Bierlaire, M. (1999) Discrete choice methods and their applications to short term travel decisions. *Handbook of transportation science*. Springer, pp. 5-33.

Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H.N., Mishalani, R. (2002) Real time simulation of traffic demand-supply interactions within dynamit. *Applied optimization* 63, 19-34.

Ben-Elia, E., Erev, I., Shiftan, Y. (2008) The combined effect of information and experience on drivers' route-choice behavior. *Transportation* 35, 165-177.

Ben-Elia, E., Shiftan, Y. (2010) Which road do i take? A learning-based model of route-choice behavior with real-time information. *Transportation Research Part A: Policy and Practice* 44, 249-264.

Bierlaire, M., Frejinger, E. (2008) Route choice modeling with network-free data. *Transportation Research Part C: Emerging Technologies* 16, 187-198.

Gao, S., Frejinger, E., Ben-Akiva, M. (2010) Adaptive route choices in risky traffic networks: A prospect theory approach. *Transportation Research Part C: Emerging Technologies* 18, 727-740.

Gao, S., Frejinger, E., Ben-Akiva, M. (2011) Cognitive cost in route choice with real-time information: An exploratory analysis. *Transportation Research Part A: Policy and Practice* 45, 916-926.

Mahmassani, H.S. (2001) Dynamic network traffic assignment and simulation methodology for advanced system management applications. *Networks and Spatial Economics* 1, 267-292.

Morikawa, T., Miwa, T. (2006) Preliminary analysis on dynamic route choice behavior: Using probe‐vehicle data. *Journal of advanced transportation* 40, 140-163.

Ramming, M.S. (2001) Network knowledge and route choice. Massachusetts Institute of Technology.

Xu, H., Zhou, J., Xu, W. (2011) A decision-making rule for modeling travelers' route choice behavior based on cumulative prospect theory. *Transportation Research Part C: Emerging Technologies* 19, 218-228.

Chapter 5

Considering Observed Taste Heterogeneity:

Exploration on the Effect of Familiarity

Route choice has been widely investigated from both methodology and practice because it is a complex process which depends on many factors. A comprehensive review of this problem can be found in Prato (2009). The random utility theory (RUT) framework(Ben-Akiva and Lerman, 1985), which is also applied in this thesis, is the most widely approach for route choice modeling.

The advantage of the RUT models is the simple specification which makes them easy to develop, estimate and apply. However, in recent years, the behavioral realism of these utility-maximizing models has been questioned by behavioral scientists (Avineri, 2004). In the case of route choice, several researches also show violations of RUT(Avineri and Prashker, 2004; Bogers et al., 2005).

Recalling the last example in Section 1.3:

- **A:** I prefer to drive on the side streets, because there are not so many cars;
  B: I hate to drive on the side streets, because there are too many intersections.

In the conventional RUT based route choice models, travelers' tastes on the observed route attribute (i.e. the parameters in the systematic parts of utility functions) are assumed to be constant and homogeneous. This assumption means that, for all the travelers and all the O-D pairs, the vectors of estimated parameters of route attributes (e.g. the rate of side streets) are the same. Therefore, the conventional RUT based route choice models cannot deal with a scenario shown in the above example.

In order to bridge the gap between economic modelers and behavioral scientists, there are two directions: suggesting alternative approaches to RUT and incorporating psychological and other factors in RUT based models. In the first direction, several alternative conceptual frameworks have been proposed, among which prospect theory (Gao et al., 2010; Xu et al., 2011) and regret theory (Chorus, 2012) have attracted increasing following in the field of route choice. In the second direction, behavioral determinants other than travel time and cost have been considered in the RUT based models to capture the taste heterogeneity across individuals and choice situations. Madanat et al. (1995) explored the effect of attitudes toward route diversion and perceptions of information reliability on route switch behavior following traffic accidents(Madanat et al., 1995). Parkany et al. (2006) explained that attitudinal indicators influence consistency and diversion for both stated and revealed preferences of drivers (Parkany et al., 2006). Bogers (2009) constructed a simulation experiment to explore the influence of day-to-day learning, habit and information. Papinski et al. (2009) examined spatial or temporal deviations between observed and pre-planned routes(Papinski et al., 2009). Prato et al. (2012) proposed a hybrid model and incorporated spatial abilities and behavioral patterns in route choice analysis

In this study, following the second direction, we will explore how to consider the observed heterogeneity in tastes in the route choice modeling. Observed taste heterogeneity means the part of taste heterogeneity which is related to some observed attributes, e.g. age of the traveler and distance between the O-D pairs. Particularly, we will explore the effect of *familiarity* on taste heterogeneity in route choice behavior.

Familiarity is a behavioral determinant associates with drivers' network knowledge. In geographic theory, the concept of *cognitive map* is used to describe drivers' network knowledge. Cognitive maps are the conceptual manifestations of place-based experience and reasoning that allow one to determine where one is at any moment and what place-related objects occur in that vicinity or in the surrounding space(Hensher, 2004). Cognitive map provides knowledge about how to get from one place to another, and naturally interacts with route choice behavior(Golledge and Garling, 2001). Individual differences exist in the degrees of knowledge about places, locations, or landmarks and other components of a route or network(Allen, 1999). To capture these differences in route choice modeling, *familiarity*, which is easy to be measured, is an appropriate term that can be incorporated in the RUT based models.

Familiarity can be divided roughly into two groups (Lotan, 1997): familiarity with the network and familiarity with the information system. In this study, we focus on the familiarity with the networks. This problem has been investigated by several researchers. Ramming (2001) developed a Multiple Indicator-Multiple Cause (MIMic) specification to examine how network knowledge may be related to a person's tenure in the metropolitan area and other socioeconomic factors. He used the questionnaire survey data to estimate the MIMic model and combined it with the route choice models. Some researchers investigated effects of familiarity on route choice behavior using the driving simulators: Adler and McNally (1994) used a hypothetical network and distinguished the levels of network knowledge according to the degree of the network map that was presented to subjects. Lotan (1997) did a similar research using a simplified real network. According to these studies, familiarity has a significant effect on route choice behavior.

In these studies, the level of familiarity is distinguished according to the different characteristics of subjects, therefore it is individual specific. From the perspective of geographic theory, these previous researches have shown the effects of heterogeneity on route choice behavior among individuals with different cognitive maps. However, with the same cognitive map, the same individual's response tastes will also be different when travel different OD pairs.

Therefore familiarity also relates to OD pairs. A driver can be very familiar with the network as a whole and travel between several OD pairs every day, while never travel between many other OD pairs. To analyze the effect of OD pair specific familiarity, we need large size individual specific route choice data between multi-OD pairs. Benefited from the GPS technology, the abundant GPS data collected by private vehicles are available, which can meet the requirements, are used in this research.

The main contribution of this study is to answer the two questions below:

- Whether the effect of heterogeneity in familiarity to OD pairs on route choice is statistically significant?

- How to incorporate the familiarity to OD pairs in route choice models?

About the first question, although the effect of familiarity to OD pairs on route choice seems obvious, we need explore whether the effect is statistically significant and worth to be considered in route choice models. If it is significant, we should answer the second question and get some findings from the estimation results of route choice models.

The rest of this chapter is organized as follows. The next section gives a description about the data used in this study. In the following section, the change of route choice behavior relating to familiarity is explored and the first question is answered. Then, for the second question, two route choice models are specified to consider familiarity to OD pairs explicitly. The estimation results of the proposed models are discussed. The final section provides some concluding comments and future research directions.

5.1 Data

The probe data used in this study are collected by private vehicles. In recent years, benefited from the popularity of vehicle navigation system, GPS data become important resources and have been used in many research about route choice (Bierlaire and Frejinger, 2008; Morikawa and Miwa, 2006; Yamamoto et al., 2012). However, because of the privacy issues, most of these research used the probe data collected by the commercial vehicles, not the private vehicles, which is more useful for exploring the heterogeneity in route choice behavior.

The private probe data was collected in Toyota, Japan in 2011, as a part of the green mobility project. More than 200 drivers participate in this survey. An on-board equipment which installed on their private vehicles will record their operations (e.g. acceleration) as well as the GPS trajectory data real time. These data will be uploaded to internet by the participants every week.

The road network of the urban area in Toyota is used in this study. This is a dense network. Part of the network is shown in Figure 5.1. It covers an area about $20 \times 16\text{km}^2$, and includes 12068 nodes and 35138 links.

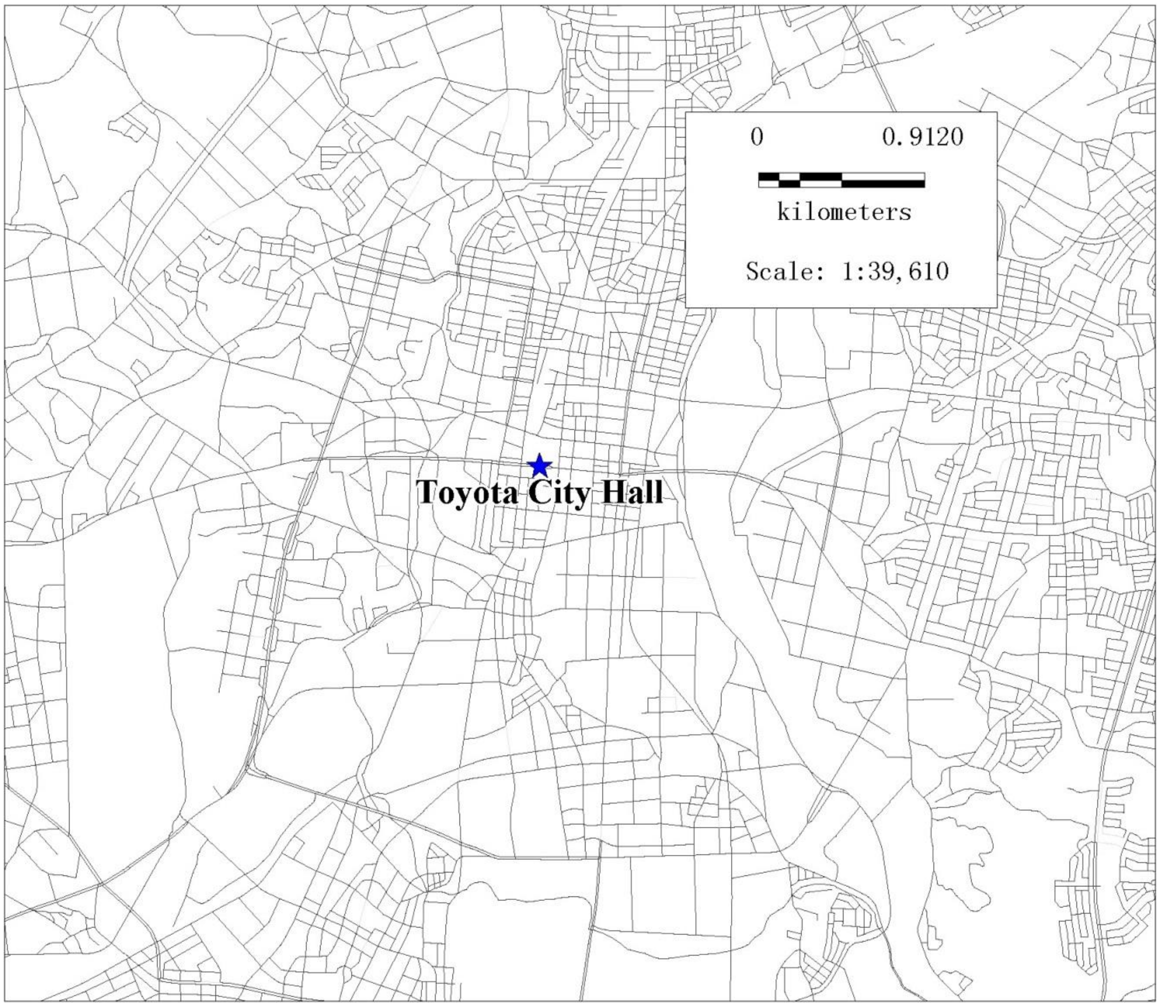


Figure 5.1 Part of the Network Studied in This Study

52 drivers who have more than 400 trips in 10 months (2011.3~2011.12) are selected as the subjects in this study. All of these subjects have abundant experiences on the network, therefore subjects' heterogeneity in familiarity to the whole network would not be considered. After a basic data cleaning process, a data set with 37254 trips is constructed for this study.

The raw GPS coordinates are map-matched to a sequence of links. Bierlaire and Frejinger (2008) have proposed an advanced processing method of GPS data that avoids the ambiguity in map matching. However, its application remains as a task for future studies. For each trip, the outflow node of the first link and the inflow node of the last link are treated as the origin and destination of this trip respectively. Many of these trips start or/and finish outside the target area. In that case, only the part in the target area is considered and treated as a complete trip. The trips with the same or adjacent first and last links are considered to share the same OD pairs.

For each trip, before analyzing the route choice behavior, a choice set should be generated. For route choice analysis, several choice set generation algorithms have been proposed and evaluated (Bekhor et al., 2006). Bovy (2009) stated that random walk method proposed by Frejinger et al. (2009) is promising. Therefore, this method is used in this study.

Given an origin-destination pair (s_o, s_d), a path, which is a ordered set of links denoted by Γ, is generated using the following algorithm:

Initialization: $v=s_o$, $\Gamma = \varnothing$.

Loop: While $v \neq s_d$ perform the following.

Weights For each link $l=(v,w) \in E_v$, where E_v is the set of outgoing links from v, the weights are calculated :

$$w(l \mid b_1, b_2) = 1 - (1 - x_l^{b_1})^{b_2} \tag{5.1}$$

$$x_l = \frac{SP(v, s_d)}{C(l) + SP(w, s_d)} \tag{5.2}$$

In this study, b1=5, b2=1.

Probability for each link $l=(v,w) \in E_v$, the probability of choosing a link is

$$q(l \mid E_v, b_1, b_2) = \frac{w(l \mid b_1, b_2)}{\sum_{l' \in E_v} w(l' \mid b_1, b_2)} \tag{5.3}$$

Draw Random select a link $(v, w)^*$ in E_v based on the above probability

Update path $\Gamma = \Gamma \cup (v, w)^*$

For each trip, 20 paths are random drawn, and construct a choice set that can replace the universal set.

5.2 Route Choice Behavior Relating to Familiarity

In this section, we try to answer the first question stated at the beginning of this chapter. Although the effect of familiarity on route choice behavior seems to be obvious, in order to explore whether it is worth to consider familiarity in choice models, we need test whether this effect is statistical significant.

The most direct measure of a driver's familiarity on an OD pair is the times that the driver has traveled between this OD pair. We use F_n^{rs} to denote the familiarity of driver *n* on OD pair *rs*. Figure 5.2 is the frequency histogram and cumulative curve about F_n^{rs} of the extracted trips.

From Figure 5.2, we can find that, about 1/3 trips with $F_n^{rs}=1$. These trips are categorized to the *very unfamiliar* group. Another 1/3 trips with $F_n^{rs} \in [2, 20]$ are categorized to the *unfamilia*r group. The remaining trips with $F_n^{rs} > 20$ are categorized to the *familiar* group. The description statistics of each group is shown in Table 5.1.

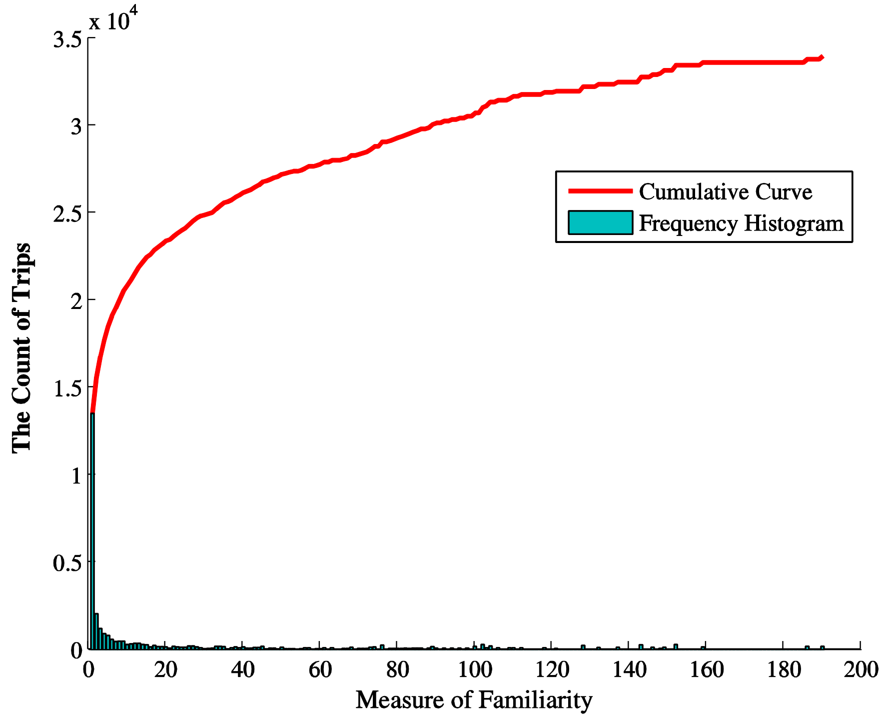


Figure 5.2 The Distribution of Familiarity in the Studied Data Set

This study uses the sampled alternatives to analyze the route choice behavior. Generally, conventional exogenous sample maximum likelihood (ESML) estimation with sampled alternatives provides consistent estimates for a multinomial logit model (McFadden, 1977).

For route choice behavior analysis, Frejinger et al. (2009) proposed an expanding path-size logit model with sampling of alternatives, which is also used in this study. Suppose R_n paths are drawn with replacements from the universal set of paths U, using the algorithm described in last section. Adding the chosen path by the driver, there will be a set C_n for observation n ($|C_n| = R_n + 1$). Suppose that C_n is the choice set contains all the unique paths in C_n, The multinomial logit model can be consistently estimated with the conditional probability given as

$$P(i \mid C_n) = \frac{\exp(\mu V_{in} + \ln(k_{in} / q(i)))}{\sum_{j \in C_n} \exp(\mu V_{jn} + \ln(k_{jn} / q(j)))} \tag{5.4}$$

where μ is the scale parameter; k_{in} is the number of times path *n* appears in C_n; V_{in} is the systematic utility that specified as

$$V_{in} = \beta' x_{in} + \beta_{ps} \cdot \ln(EPS_{in}) \tag{5.5}$$

where β' is the explanation variables with parameters; EPS_{in} is the expanding path size of route *i* in observation *n*, which is given as

$$EPS_{in} = \sum_{a \in \Gamma_i} \frac{L_a}{L_i} \frac{1}{\sum_{j \in C_n} \delta_{aj} \Phi_{jn}} \tag{5.6}$$

where L_a is the length of link a; L_i is the length of path; δ_{aj} equals 1if path *j* contains link *a* and is zero otherwise; Φ_{jn} is the expansion factor given as

$$\Phi_{jn} = \begin{cases} 1 & \text{if route } j \text{ is chosen or } q(j)R_n \geq 1; \\ \dfrac{1}{q(j)R_n} & \text{otherwise.} \end{cases} \tag{5.7}$$

Table 5.2The estimated parameters using different group of data are shown in Table 5.2. Looking at the estimated parameters, all of the estimates in each group are statistically significant and have the expected sign. We can also find the estimated parameters are different between groups. When drivers travel between very unfamiliar OD pairs, they are more sensitive on the count of intersections than free travel time.

Table 5.2Considering again about the estimation results shown in Table 5.2 it is of interest to know whether the differences of estimated parameters in different groups are statistically significant, or those differences occur only because of sampling error.

Because the estimated parameters to be compared are confounded with their respective scale factors in the proposed choice model, the standard Chow test may lead to a wrong conclusion. Consequently, in order to isolate the scale factor differences, a two-stage variant of Chow test is used in this study to show whether the observations in different groups share the same parameters of route choice model.

Swait and Louviere (1993) proposed this test method. In general, it is to consider a situation in which two data sets contain observations of an identical specified multinomial logit model. The estimated parameters of the choice model using different data sets are $\hat{\beta}_1$ and $\hat{\beta}_2$ respectively.

Table 5.1 Description Statistics of the Trips in Each Group

Groups	Familiar	Unfamiliar	Very Unfamiliar	All
Count of trips	10761	9724	13520	34005
Mean free travel time	7.78	6.91	7.02	7.23
Std. deviation of free travel time	4.96	4.76	2.78	4.85
Mean count of intersections	42.11	36.87	37.72	38.87
Std. deviation of count of intersections	24.66	22.63	22.12	23.20
Mean familiarity	78.40	6.99	1	27.21
Std. deviation of familiarity	44.89	4.89	0	43.17

We are interested to know whether the differences between $\hat{\beta}_1$ and $\hat{\beta}_2$ occur because (1) they are simply the result of sampling error, since the true underlying parameters and scale factors are the same in both populations(i.e. $\beta_1 = \beta_2$, $\mu_1 = \mu_2$); (2) the true underlying parameters -are the same but the scale factors are different ($\beta_1 = \beta_2$, $\mu_1 \neq \mu_2$); (3) there are real differences in true parameters and scale factors($\mu_1\beta_1 \neq \mu_2\beta_2$).This problem is identical to test the hypothesis

H1: $\beta_1 = \beta_2$ and $\mu_1 = \mu_2$

The procedure of the two-stage pairwise hypothesis testing is

- Test whether β_1 and β_2 are equal (H_{1A}: $\beta_1 = \beta_2 = \beta$) while permitting the scale factors to differ between datasets. The test statistic is

$$\lambda_A = -2[L_\mu - (L_1 + L_2)]$$

Where L_1 and L_2 are log likelihood values of route choice models estimated using the two datasets respectively. L_μ is the log likelihood value of choice model estimated using the concatenate data of these two datasets, while set $\mu_1 = 1$, μ_2 is a parameter needed to be estimated.

- If H_{1A} cannot be rejected, then test hypothesis: H_{1B}: $\mu_1 = \mu_2 = \mu$, using the test statistic

$$\lambda_B = -2[L_p - L_\mu]$$

Where L_p is the log likelihood value for the model under H_{1B}. Both λ_A and λ_B are asymptotically chi-squared distributed with (K+1) degrees of freedom. K is the number of parameters.

Table 5.2 Estimation Results with Data of Different Groups

Parameter on	Parm (t-stat.)		
	Familiar group	Unfamiliar group	Very unfamiliar group
Free travel time(min)	-0.7389 (-45.13)	-1.26 (-64.13)	-0.549 (-37.10)
Count of intersections	-0.6144 (-231.03)	-0.74 (-220.52)	-0.70 (-270.67)
Ln(*EPS*)	0.59 (14.42)	0.27 (7.91)	0.86 (24.89)
LL0	-217240.21	-174090.73	-236950.91
LL	-44031.78	-19046.70	-40041.81
Rho square	0.7973	0.8906	0.8310
Count of trips	10761	9724	13520

Following this procedure, we test the heterogeneity between groups with different levels of familiarity to OD pairs. The results are shown in Table 5.3. We can find that, in all of the three cases, the hypothesis is rejected at the first stage. This indicates that there are real differences in true parameters and scale factors crossing groups. The estimated value of μ_2 is simply an average multiplier that optimally scales the data of the second group to offset the imposition of the β parameter equality assumption.

In order to test whether the heterogeneity between groups is really because of the familiarity to OD pairs, we do another heterogeneity test within groups. Each group is randomly divided to two sub-groups. The description statistic is shown in Table 5.4. For each group, we do a test on heterogeneity between sub-groups. The results are shown in Table 5.5. We can find that only the test between unfamiliar sub-groups (U1 and U2) should be rejected for 95% confidence level. This can confirm again the heterogeneity found in Table 3 relates to the familiarity to OD pairs. The rejection of test between U1 and U2 implies that the marginal effect of familiarity is larger when the degree of familiarity is lower.

Table 5.3 Heterogeneity Test Crossing Groups

Case	μ_2	L_1	L_2	L_μ	λ_A	Reject H_{1A}?
F, VU	1.07	-44031.78	-40041.81	-84153.19	159.2	yes
F, U	1.30	-44031.78	-19046.70	-41048	225.64	yes
VU, U	1.21	-40041.81	-19046.70	-59563.16	949.30	yes

1. Chi-squared statistic for 4 d.f. and 95% confidence level =9.49.

2. F, VU and U denote the familiar group, very unfamiliar group and unfamiliar group respectively.

Table 5.4 Description Statistics of the Trips in Each Sub-Group

Sub-Groups	VF		U		VU	
	VF1	VF2	U1	U2	VU1	VU2
Count of trips	5438	5323	4910	4814	6784	6736
Mean free travel time (min)	7.71	7.86	6.96	6.85	6.98	7.06
Std. deviation of free travel time	4.92	5.00	4.78	4.74	4.80	4.77
Mean count of intersections	42.03	42.20	37.15	36.58	37.42	38.03
Std. deviation of count of intersections	24.52	24.81	22.76	22.48	21.97	22.27
Mean familiarity	78.54	78.25	7.06	6.91	1	1
Std. deviation of familiarity	44.53	45.26	4.93	4.85	0	0

Table 5.5 Heterogeneity Test Crossing Sub-Groups

Case	F1, F2	U1, U2	VU1, VU2
μ_2	1.00	1.01	1.00
L_1	-20165.59	-9706.45	-21572.84
L_2	-19872.16	-9330.96	-22458.53
L_μ	-40041.58	-19043.98	-44031.78
λ_A	7.66	13.14	0.82
Reject H_{1A}?	*no*	*yes*	*no*
L_p	-40041.81	-	-44031.78
λ_B	0.46	-	0.00
Reject H_{1B}?	*no*	-	*no*

5.3 Route Choice Models Considering Familiarity

According to the findings in last section, route choice behavior significantly relates to the heterogeneity in familiarity to OD pairs. Therefore, it is worth to consider familiarity in route choice models. In this section, we try to answer the second question stated in the first section. Two route choice models are proposed to incorporate familiarity to OD pairs explicitly.

- *Model 1: structured scale parameter*

The heterogeneity among observations resulting from the differences in familiarity to OD pairs can be considered by using a structured scale parameter. The scale parameter μ in Equation (4) can be structured by familiarity. Let ε_{in}^{rs} denote the error term of choice model, which has a variance expressed as

$$Var(\varepsilon_{in}^{rs}) = \frac{\pi^2}{6(\mu_n^{rs})^2} \tag{5.8}$$

The structured scale parameter is given as

$$\mu_n^{rs} = (F_n^{rs})^{\gamma} \tag{5.9}$$

F_n^{rs} is the count of trips that driver n has traveled between OD pair *rs* during the survey period. γ is a parameter needed to be estimated. If γ is negative (positive), the scale parameter is smaller (larger), thus the variance of error is larger (smaller), for more familiar OD pairs. This structuralization of the scale parameter includes heterogeneity, which is assumed in conventional logit models to be a special case when $\gamma = 0$.

- *Model 2: structured parameters of explanation variables*

As analyzed in Section 5.2, the familiarity to OD pairs also affects the parameters of explanation variables in the systematic term of utility. In order to consider the systematic response heterogeneity, Bhat (1998 and 2000) proposed a specification (Bhat, 1998, 2000) which can allow for the observed heterogeneity:

$$\mu_n^{rs} = (F_n^{rs})^{\gamma} \tag{5.10}$$

β' is the parameters in systematic term of utility. The exponential specification is to guarantee the correct sign of parameters for all the possible values of F_n^{rs}. In this study, both the signs of intersection count and free travel time are negative, while the sign of $\ln(EPS_{in})$ is positive. c and α are parameters to be estimated. If α is negative(positive), the more familiar to an OD pair, the less(more) sensitive to the explanation variables.

Because estimated parameters are confounded with their respective scale factors in the MNL models, model 2 can also incorporate the heterogeneity in scale parameter. Therefore, it is not necessary to combine *Model 1* and *Model 2*.

Using the combined data of all of the three groups, *Model 1* and *Model 2* are estimated, as well as the conventional choice model used in Section 5.2 (*Model 0*). The estimation results are shown in Table 5.6.

Table 5.6 Estimation Results for Models Considering Familiarity

Parameter on		Parm (t-stat.)		
		Model 1	*Model 2*	*Model 0*
Free travel time (min)	Constant	-0.61(-88.16)	-0.21(-14.65)	-0.83(-88.15)
	Familiarity (/100)	-	-0.14(-2.31)	-
Count of intersections	Constant	-0.50(-345.24)	-0.34(-226.14)	-0.68(-420.59)
	Familiarity (/100)	-	-0.70(-25.64)	-
Ln(*EPS*)	Constant	0.40(26.16)	-0.62(-116.38)	0.54(25.59)
	Familiarity (/100)	-	-0.01(-25.32)	-
γ		-0.03(-34.57)	-	-
LL0		-628281.86	-628281.86	-628281.86
LL		-106437.15	-105746.72	-107025.90
Rho square		0.8306	0.8317	0.8296
Count of trips			34005	

The three estimated models are applied to a specific setting of choice situation and predictions of route shares are compared to illustrate the potential biases introduced by not considering heterogeneity in familiarity to OD pairs.

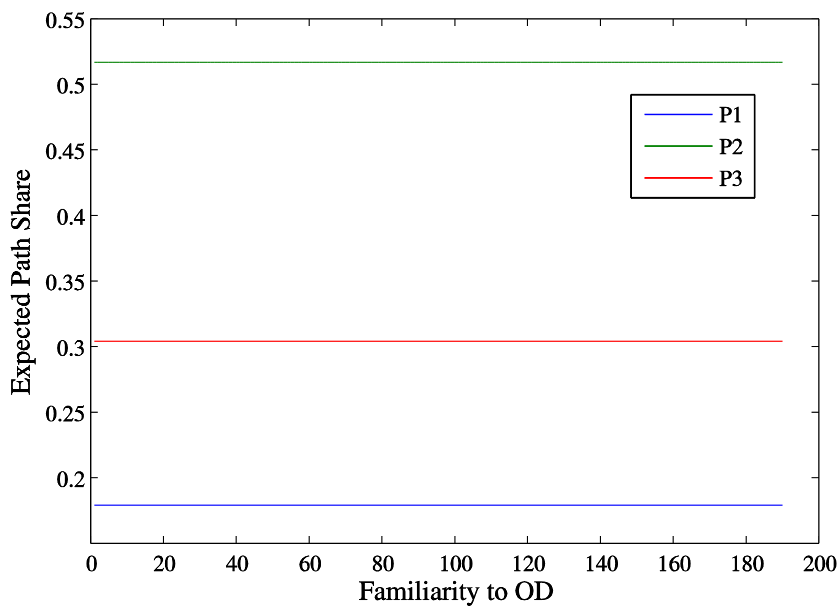
Consider a situation that the universal set of alternatives contains 3 routes without overlap:

Route 1: free travel time is 7 minutes; count of intersections is 40;

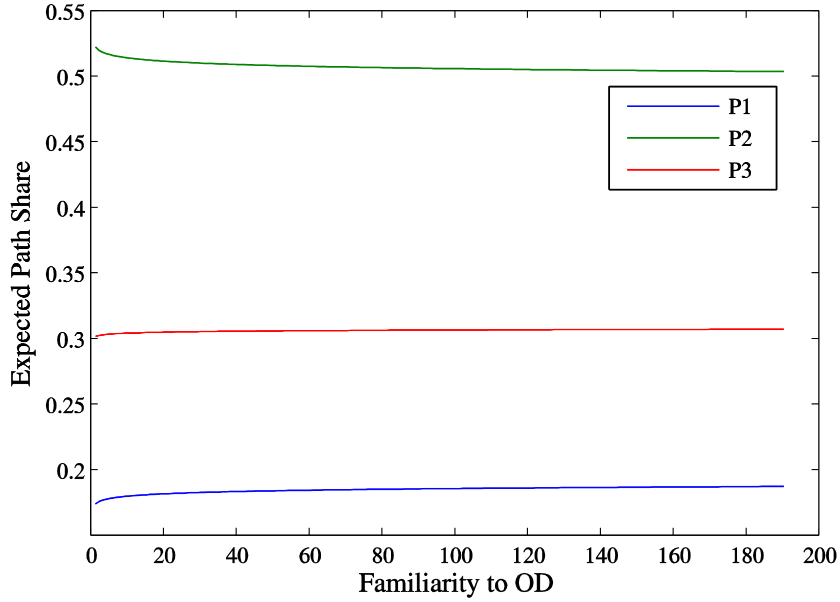
Route 2: free travel time is 9 minutes; count of intersections is 36;

Route 3: free travel time is 8 minutes; count of intersections is 38.

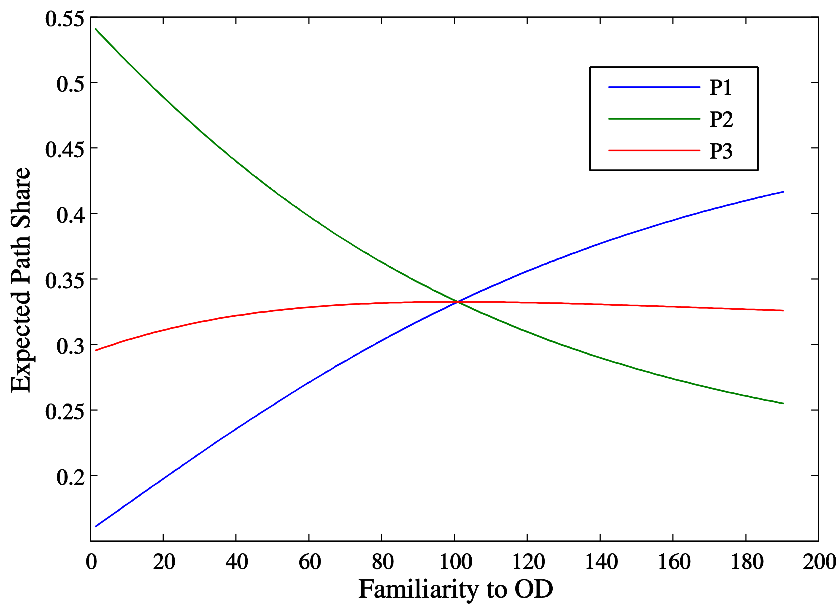
We let F_n^{rs} vary from 1 to 190 (the maximum in the studied dataset), the predicted route shares by 3 estimated models are shown in Figure 5.3.



(a) Model 0



(b) Model 1



(c) Model 2

Figure 5.3 Familiarity to OD Impacts on Expected Path Shares

In Section 3, we try to answer the second question stated in the first section. Two models are proposed to incorporate familiarity for route choice analysis. From the estimation and prediction results shown in Table 5.6 and Figure 5.3, several findings are summarized as follows:

Concerning the goodness-of-fit statistics, the final log-likelihood of *Model 1* and *Model 2* is significant higher than *Model 0*. The adjusted Rho square values of *Model 1* and *Model 2* are both more than 0.001 higher than *Model 0*. Considering the large count of observations, this indicates that the proposed models can capture the effect of familiarity, and confirms that models with explicitly consideration of familiarity on OD pairs merit further consideration(Ben-Akiva and Lerman, 1985). *Model 2* has a 0.001 higher adjusted Rho square value than *Model 1*, this indicates that familiarity not only affects the unobserved term of utility, but also affects the drivers' responses to the explanation variables in the systematic term of utility.

In *Model 1*, all of the coefficient estimates have the expected signs. The estimated sale parameter γ is negative, and statistically significantly different from zero. This suggests that trips between more familiar OD pairs have larger variances of error terms. The error term of utility represents the utility that observed by the drivers but not observed by the analysts. The only observed explanation variables in this study are free travel time and count of intersections. However route choice behavior will be affected by many factors. Some of the dynamic factors could not be observed by the analysts, such as the real time perceived travel time and received traffic information. These factors will be incorporated in the error term. Therefore, the negative sign of γ means that the drivers are more easily to be affected by some dynamic factors when travel between more familiar OD pairs. In fact, when a driver has a trip between an unfamiliar OD pair, he/she will choose a route based on some objective and static knowledge about the network, or even follow the navigation system and choose the shortest path. In opposite, when a driver travels between a very familiar OD pair, he/she will be more easy to change the route choice en-route, because of the dynamic perceptions about the traffic state.

In *Model 2*, the coefficient of familiarity for each parameter of explanation variable is negative, and statistically significantly different from zero. This implies that when drivers drive between more familiar OD pairs, they will less sensitive to the free travel time and count of intersections, and consider more factors that are not observed in *Model 2*. This is consistent to the estimation results of *Model* 1. However, unlike *Model* 1, *Model 2* also captures the effect of familiarity on drivers' response to observed explanation variables. The parameter of familiarity for the response to free travel time has a lower absolute value than that for the response to the count of intersections. This implies that when travel between more familiar OD pairs, drivers will be less sensitive to the count of intersections than free travel time.

The prediction results can show the findings more intuitively. The We can find that, applying *Model 0*, the shares of three routes keep constant as the degree of familiarity increasing. If *Model 1* is applied, as familiarity increasing, because of the increasing variances of error terms, the shares of routes approach to each other. However, because γ is low, the effect of familiarity is not so significant. From figure 3(a), we can also find that the varying of scale parameter will not change the order of route shares. That is because the error terms are assumed to be independent identically distributed (i.i.d.), and the order of route shares only associate with the systematic part of utility. Applying *Model 2*, the shares of routes change significantly as familiarity increasing. When F_n^{rs} equates to 1, *route 1* takes the smallest share of the three routes. As F_n^{rs} increasing, the differences of shares between routes approach to the average (1/3). Unlike Model 1, because Model 2 also consider the effect of familiarity on systematic utility, when F_n^{rs} is larger than about 100, the order of route shares will change, and *route 1* takes the largest share. *route 1* is the alternative with shortest free travel time and most intersections. This result confirms the finding that familiarity has different effects on parameters of the two explanation variables, and drivers will be less sensitive to the count of intersections than free travel time when travel between more familiar OD pairs. The obvious changing of route shares associate with familiarity also indicates the potential biases introduced by not considering heterogeneity in familiarity to OD pairs.

5.4 Conclusions and Future Directions

This chapter provides an exploratory analysis about the effect of familiarity on route choice behavior. Familiarity considered here is both individual and OD pair specific, different from previous researches which only consider the individual specific familiarity. The analysis is based on the probe data collected by private vehicles in Toyota, Japan.

The first contribution of this study is the empirical research about the effect of familiarity to OD pairs on route choice behavior. Although the effect of familiarity on route choice seems obvious, we need explore whether this effect is statistically significant and worth to be considered in route choice models. The route choice observations are categorized into 3 groups according to the levels of familiarity to OD pairs. These three groups of observations are used to estimate the route choice model with the same specification respectively. The differences on estimated parameters across groups show the potential effect of familiarity. To test whether these differences are statistically significant, A two-stage variant of Chow test on the different groups of observations is carried out. The test results prove that the differences on parameters are statistically significant and occur not only because of sampling error. In order to further test whether the differences on parameters across groups mainly relate to familiarity, or other unobserved factors, we divide the observations in each groups into 2 sub-groups randomly, and do the two-stage test on each pair of sub-groups. The test results exclude the effect of unobserved factors on the parameters of route choice model.

The second contribution of this study is the proposed models which can incorporate familiarity explicitly. Since the effect of familiarity on route choice behavior has been proved to be statistically significant, it is worth to consider familiarity in route choice models. 2 specifications of choice models are proposed: a model with structured scale parameter (*Model 1*) and a model with structured parameters of explanation variables (*Model 2*). These two models are estimated using the revealed observations extracted from private probe data, and applied to apply the route shares under a specific setting of choice situation. According to the estimation and prediction results, the proposed models can capture the effect of familiarity. The better performance of *Model 2* than *Model 1* implies the necessary to consider familiarity in both unobserved and observed parts of utility. Finding from the results, drivers will be more easy affected by some unobserved factors (to the analyst), such as en-route information and perception, when travel between more familiar OD pairs. The estimated parameters also imply that drivers will be less sensitive to the count of intersections than free travel time when travel between more familiar OD pairs.

From an application perspective, this study can be used in travel demand prediction and navigation system. Benefited from the popularity of GPS, abundant location data of travelers have made it possible to analyze drivers' behavior and predict travel demand at the individual level. The route choice models considering more behavioral terms (e.g. memory, habit and familiarity) besides cost (e.g. travel time and distance) are required by the activity based travel demand prediction(Ben-Akiva and Bowman, 1998). The findings and proposed models in this study have shown the potential of incorporating familiarity for the improvement of travel demand prediction. This study can also benefit the development of navigation system. One core part of navigation system is to find routes that can give users most satisfaction. This study indicates that drivers' evaluation rules of satisfaction are related to familiarity (more sensitive to the count of intersections when travel between less familiar OD pairs). Therefore, this study has a potential to improve the service of navigation system.

In this research, because it is lack of data about drivers' characteristics and dynamic traffic information, only two explanation variables about static network knowledge are incorporated in the choice models. In future research, more explanation variables can be considered. This study focuses on the familiarity to OD pairs, in the future, the familiarity on information systems, links and routes will also be considered and incorporated in choice models. In this study, familiarity is static. However, it is a dynamic process to get familiar to an OD pair. The dynamic feature of familiarity is another direction for future work.

References

Adler, J.L., McNally, M.G. (1994) In-laboratory experiments to investigate driver behavior under advanced traveler information systems. *Transportation Research Part C: Emerging Technologies* 2, 149-164.

Allen, G.L. (1999) Spatial abilities, cognitive maps, and wayfinding. *Wayfinding behavior: Cognitive mapping and other spatial processes*, 46-80.

Avineri, E. (2004) A cumulative prospect theory approach to passengers behavior modeling: Waiting time paradox revisited. *Proceedings of Intelligent Transportation Systems*, pp. 195-204.

Avineri, E., Prashker, J.N. (2004) Violations of expected utility theory in route-choice stated preferences: Certainty effect and inflation of small probabilities. *Transportation Research Record: Journal of the Transportation Research Board* 1894, 222-229.

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

Ben-Akiva, M., Bowman, J.L. (1998) Activity based travel demand model systems. *Equilibrium and advanced transportation modeling*, 27-46.

Ben-Akiva, M.E., Lerman, S.R. (1985) *Discrete choice analysis: Theory and application to travel demand*. the MIT Press.

Bhat, C.R. (1998) Accommodating variations in responsiveness to level-of-service measures in travel mode choice modeling. *Transportation Research Part A: Policy and Practice* 32, 495-507.

Bhat, C.R. (2000) Incorporating observed and unobserved heterogeneity in urban work travel mode choice modeling. *Transportation Science* 34, 228-238.

Bierlaire, M., Frejinger, E. (2008) Route choice modeling with network-free data. *Transportation Research Part C: Emerging Technologies* 16, 187-198.

Bogers, E.A.I., Viti, F., Hoogendoorn, S.P. (2005) Joint modeling of advanced travel information service, habit, and learning impacts on route choice by laboratory simulator experiments. *Transportation Research Record: Journal of the Transportation Research Board* 1926, 189-197.

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

Chorus, C.G. (2012) Regret theory-based route choices and traffic equilibria. *Transportmetrica* 8, 291-305.

Frejinger, E., Bierlaire, M., Ben-Akiva, M. (2009) Sampling of alternatives for route choice modeling. *Transportation Research Part B: Methodological* 43, 984-994.

Gao, S., Frejinger, E., Ben-Akiva, M. (2010) Adaptive route choices in risky traffic networks: A prospect theory approach. *Transportation Research Part C: Emerging Technologies* 18, 727-740.

Golledge, R.G., Garling, T. (2001) *Spatial behavior in transportation modeling and planning*. University of California Transportation Center, University of California.

Hensher, D.A. (2004) *Handbook of transport geography and spatial systems*. Emerald Group Publishing.

Lotan, T. (1997) Effects of familiarity on route choice behavior in the presence of information. *Transportation Research Part C: Emerging Technologies* 5, 225-243.

Madanat, S., Yang, C., YING-MING, Y. (1995) Analysis of stated route diversion intentions under advanced traveler information systems using latent variable modeling. *Transp. Res. Record*, 10-17.

McFadden, D. (1977) Modelling the choice of residential location. *Cowles Foundation Discussion Papers* 477.

Morikawa, T., Miwa, T. (2006) Preliminary analysis on dynamic route choice behavior: Using probe‐vehicle data. *Journal of advanced transportation* 40, 140-163.

Papinski, D., Scott, D.M., Doherty, S.T. (2009) Exploring the route choice decision-making process: A comparison of planned and observed routes obtained using person-based gps. *Transportation Research Part F: Traffic Psychology and Behaviour* 12, 347-358.

Parkany, E., Du, J., Aultman-Hall, L., Gallagher, R. (2006) Modeling stated and revealed route choice: Consideration of consistency, diversion, and attitudinal variables. *Transportation Research Record: Journal of the Transportation Research Board* 1985, 29-39.

Prato, C.G. (2009) Route choice modeling: Past, present and future research directions. *Journal of Choice Modelling* 2, 65-100.

Prato, C.G., Bekhor, S., Pronello, C. (2012) Latent variables and route choice behavior. *Transportation* 39, 299-319.

Ramming, M.S. (2001) Network knowledge and route choice. Massachusetts Institute of Technology.

Swait, J., Louviere, J. (1993) The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research*, 305-314.

Xu, H., Zhou, J., Xu, W. (2011) A decision-making rule for modeling travelers' route choice behavior based on cumulative prospect theory. *Transportation Research Part C: Emerging Technologies* 19, 218-228.

Yamamoto, T., Takamura, S., Morikawa, T. (2012) Analysis of mode and walk-route choice in a downtown area considering heterogeneity in trip distance. *Proceedings of The 91st Annual meeting of Transportation Research Board*, Washington DC, USA.

Chapter 6

Considering both Observed and Unobserved Taste Heterogeneity: A Multi-Level Mixed Logit Model

In this research, following last chapter, we will explore drivers' taste heterogeneity in route choice analysis. Taste heterogeneity may be incorporated into route choice analysis by introducing observed individual socio-economic characteristics. However, due to the difficulty of data collection for route choice analysis, there is limited research considering observed individual characteristics with revealed preference data. What is more, in the context of route choice, the characteristics of Origin-Destination (O-D) pairs also have an important effect on drivers' route choice behavior. In most previous research into heterogeneity, only a single O-D pair is considered so the O-D pair specific heterogeneity cannot be explored.

On the other hand, it is very likely that taste heterogeneity will remain even when observed characteristics are accounted for. The mixed logit model is a popular mathematical structure for the analysis of unobserved heterogeneity. A method based on the mixed logit model in which both observed and unobserved heterogeneity is considered can also be found in the literature (Bhat, 1998, 2000).

There are two different versions of the mixed logit model: the random coefficient logit model and the flexible error logit model. Since these two versions are proved to be formally equivalent, researchers can choose one version according to their focuses of studies (Train, 2003): The random coefficient version is more straightly when account the correlations among the coefficients. Therefore, it is more appropriate for incorporating heterogeneity and dealing with panel data. The flexible error logit model, which is also often referred as Logit Kernel model, is more appropriate for considering the correlations over alternatives. These two versions both have applied in route choice modeling: Bekhor et al. (2002) apply the Logit Kernel model to consider the overlap problem while Bogers (2009) use the random coefficient logit model to deal with panel data.

Since this study focuses on taste heterogeneity, we apply the random coefficient specifications. Random taste heterogeneity in a mixed logit model is accommodated by random parameters associated with attributes within the utility function specification. Regarding these random parameters, it is often assumed that they are independent between choices. This assumption is only appropriate when the observations are cross-sectional data. However, for route choice, there are often repeated choices for one individual between one O-D pair. To deal with the repeated choice, as applied in almost all previous research in the field of transportation, it is often assumed that tastes vary across individuals, but stay constant across observations for the same individual (Revelt and Train, 1998). However, adopting this assumption is in fact to ignore intra-traveler heterogeneity.

In the context of route choice, because of the complicated nature of the choices, tastes also will vary across choice situations for the same traveler. For example, if the driver is going to be late for work, he/she will be more sensitive to the travel time. This means it is not appropriate to assume intra-traveler homogeneity because travel purpose obviously will affect taste. Understanding this, Hess and Rose (2009) proposed a generalized method to allow for both inter-individual and intra-individual heterogeneity. They have also tested the relative accuracy of this method (Hess and Train, 2011).

The main contribution of this study is to incorporate both the observed and unobserved taste heterogeneity into route choice analysis. The method proposed by Hess and Rose (2009) is extended in two respects for particular application in route choice. Firstly, combined with the method proposed by Bhat (2000), the observed heterogeneity is also considered. Secondly, the intra-individual heterogeneity is divided into two parts: O-D pair specific and choice situation specific heterogeneity. For an empirical analysis, GPS data collected by private vehicles in Toyota city, Japan is used in this study for the estimation of the proposed models.

The remainder of this chapter is organized as follows. Section 6.1 presents the modeling approach developed in this study. Section 6.2 describes the data for route choice analysis. Section 6.3 gives the specifications of models estimated in this study. Section 6.4 shows and analyzes the estimation results. Finally, Section 6.5 presents the conclusions of this study.

6.1Methodology

In this section, we develop a modeling framework of route choice that allows for the joint representation of observed and unobserved heterogeneity. For the unobserved part, both inter-traveler and intra-traveler heterogeneity will be considered.

6.1.1 Path-size Logit model

Because of the overlap of alternative routes in a route choice situation, the MNL model is not appropriate for route choice analysis. The Path-size model was proposed to deal with the overlap problem, while maintaining the computational simplicity of the logit form. As described in the second chapter, Path-size Logit model and C-logit model are both MNL-modifications and only account the overlaps of routes within the deterministic part. Because likelihood values show that the Path-Size Logit model generally outperforms the C-Logit model (Prato and Bekhor, 2006, 2007; Ramming, 2001), although these two models have a similar specification, C-logit model is not chosen in this study. GEV models, Probit model, and Logit Kernel model can consider the overlaps within the stochastic term, however, these models have much more computational requirements than Path-size model (Ramming, 2001). In this study, as shown in the following parts of this section, the random parameter specification has already made the proposed models need a considerable long estimation time. Considering the large network and abundant observations in the case study, to keep an affordable computational expenditure, we apply the path-size model to consider the overlap problem, rather than the more complicate models.

Let $U_{i,n,m,t}$ be the utility of route i for respondent n in choice situation t when traveling between O-D pair m. This consists of an observed utility $V_{i,n,m,t}$, and an unobserved component $\varepsilon_{i,n,m,t}$, such that

$$U_{i,n,m,t} = V_{i,n,m,t} + \varepsilon_{i,n,m,t} \tag{6.1}$$

The unobserved components of the alternatives are assumed to be independent and identically distributed (i.i.d.) as a Gumbel distribution. The observed utility is assumed to have a linear relationship between attributes and tastes, such that

$$V_{i,n,m,t} = \beta_{n,m,t}^{'} x_{i,n,m,t} \tag{6.2}$$

where $x_{i,n,m,t}$ is a vector of observed route attributes; $\beta_{n,m,t}$ is a vector of coefficients that represent drivers' tastes on route attributes.

To consider the correlation of alternative routes, a correction term is added to the utility of alternative routes. The probability that respondent *n* chooses route *i* in choice situation *t* between O-D pair *m* is then given by the path-size model:

$$P(i \mid C_{n,m,t}) = \frac{e^{V_{i,n,m,t} + \beta_{ps} \ln(PS_{i,n,m,t})}}{\sum_{j \in C_{n,m,t}} e^{V_{j,n,m,t} + \beta_{ps} \ln(PS_{j,n,m,t})}} \tag{6.3}$$

The correction term $PS_{i,n,m,t}$ is formulated as

$$PS_{i,n,m,t} = \sum_{a \in \Gamma_i} \left(\frac{l_a}{L_i} \right) \frac{1}{\sum_{j \in C_{n,m,t}} \delta_{a,j}} \tag{6.4}$$

where Γ_i is the set of links in route *i*; l_a is the length of link *a*; L_i is the length of route *i*; $\delta_{a,j}$ is the link-path incidence dummy (that is, 1 if path *j* uses link *a* and 0 otherwise).

6.1.2 Incorporating observed heterogeneity

In the traditional path-size model, tastes β are fixed coefficients. To incorporate observed heterogeneity, as applied by Bhat (1998), β are assumed to have a linear relationship between characteristics and coefficients:

$$\beta_{n,m,t} = \alpha' y_{n,m,t} \tag{6.5}$$

where $y_{n,m,t}$ is a vector of observed variables that related to drivers' tastes on route attributes; α is a matrix of coefficients.

For route choice analysis, the characteristics can be divided to three categories: individual-specific (e.g. age and income), O-D pair specific (e.g. distance) and choice situation specific (e.g. weather).

6.1.3 Incorporating unobserved heterogeneity

The mixed logit model is used to incorporate the unobserved taste heterogeneity. In a mixed logit model, the vector of taste coefficient β is assumed to follow a certain random distribution with probability density function $f(\beta|\Omega)$. Ω represents a set of parameters of the distribution of β. Combined with the observed heterogeneity, the structure of the taste coefficient will then be

$$\beta_{n,m,t} = \alpha' y_{n,m,t} + z_{n,m,t} \tag{6.6}$$

where $z_{n,m,t}$ is assumed to be normally distributed with a 0 mean (the mean is incorporated in the constant term in $\alpha' y_{n,m,t}$).

In the context of the mixed logit model, two main specifications exist. The cross-sectional specification is the standard approach for a one-shot choice. With this specification, all observations are treated as independent. In route choice analysis, the observations are usually from different travelers. However, for each traveler, there will often be more than one observation. With the cross-sectional specification, separated observations from the same traveler are treated as if they came from separate travelers. From the perspective of taste heterogeneity, this means that sensitivities vary across choices for a given traveler in the same way they vary across travelers.

Accordingly, another specification (Revelt and Train, 1998) has been designed for the case of repeated choice data. With this specification, separated observations from the same traveler share the same taste coefficients, while observations from separate travelers are assumed to be independent. From a taste heterogeneity perspective, this specification can only allow for inter-traveler heterogeneity.

A more generalized specification was proposed by Hess and Rose (2009) in which intra-traveler heterogeneity can also be considered. With this specification, $z_{n,m,t}$ in Eq. (5) is the sum of two terms, δ_n and $\eta_{n,m,t}$; that is,

$$z_{n,m,t} = \delta_n + \eta_{n,m,t} \tag{6.7}$$

where δ_n is distributed across travelers but not over multiple choice situations for a given traveler and $\eta_{n,m,t}$ varies over all choices as well as travelers. In this specification, δ_n captures the inter-individual variation in tastes while $\eta_{n,m,t}$ captures intra-individual variation.

With this specification, for observations of a given traveler, the intra-individual variation $\eta_{n,m,t}$ is treated as independent. However, in the context of route choice, different observations from the same traveler are often for the same O-D pair. As with individual-specific heterogeneity, correlations over choice situations of a given traveler for the same O-D pair should also be considered. To incorporate the O-D pair specific heterogeneity, we extend the framework proposed by Hess and Rose (2009) and develop a new specification particularly for route choice analysis. In our model, $z_{n,m,t}$ becomes the sum of three normally distributed terms δ_n, $\varphi_{n,m}$ and $\eta_{n,m,t}$; that is,

$$z_{n,m,t} = \delta_n + \varphi_{n,m} + \eta_{n,m,t} \tag{6.8}$$

Where δ_n and $\eta_{n,m,t}$ are defined as for Eq. 6, and $\varphi_{n,m}$ varies over O-D pairs for each traveler and captures the inter-OD variation in tastes. All of these three terms are assumed to be normally distributed with mean 0 and variance 1.

It should be noted that, for simplification, correlations of observations for the same O-

D pair but from different travelers are not considered in this specification.

6.1.4 A discussion on the error components

To give a further discussion on the error components, the utility function in the proposed model is rewritten with a flexible error specification:

$$
\begin{aligned}
U_{i,n,m,t} &= \alpha' y_{n,m,t} \cdot x_{i,n,m,t} + \zeta_{i,n,m,t} \\
\zeta_{i,n,m,t} &= z'_{n,m,t} \cdot x_{i,n,m,t} + \varepsilon_{i,n,m,t}
\end{aligned}
\tag{6.9}
$$

where $\zeta_{i,n,m,t}$ is the unobserved (random) portion of utility. As shown in Eq. (7), $z_{n,m,t}$ is the sum of three terms: δ_n , $\varphi_{n,m}$ and $\eta_{n,m,t}$. δ_n and $\varphi_{n,m}$ are not observation specific. Observations from the same individual will share the same δ_n . Observations between the same OD pair from the same individual will have the same $\varphi_{n,m}$. Therefore utility should be correlated over observations from panel data.

It should be noted that, for all the alternatives of one observation, the random term $z_{n,m,t}$ will be same. Therefore, with this specification, the unobserved term $\zeta_{i,n,m,t}$ is also correlated over alternatives. For two different routes *i* and *j*, the covariance of unobserved terms will be

$$
\begin{aligned}
\mathrm{Cov}(\zeta_{i,n,m,t}, \zeta_{j,n,m,t}) &= E[(z'_{n,m,t} \cdot x_{i,n,m,t} + \varepsilon_{i,n,m,t})(z'_{n,m,t} \cdot x_{j,n,m,t} + \varepsilon_{j,n,m,t})] \\
&= x'_{i,n,m,t} W x_{j,n,m,t}
\end{aligned}
\tag{6.10}
$$

where W is the covariance of $z_{n,m,t}$.

We can find that $\mathrm{Cov}(\zeta_{i,n,m,t}, \zeta_{j,n,m,t})$ is nonzero and related to the observed route attributes (e.g. route travel time and route length).

The specification shown in Eq. (6.9) seems similar to the Logit Kernel model proposed by Bekhor et al. (2002). However, in their specification, alternatives are assumed to be correlated because of the overlap problem rather than the correlated random tastes, and the covariance of unobserved terms is related to the overlap length of routes rather than the absolute values of observed route attributes. Therefore, for nonoverlapping routes, the covariance of unobserved terms will be zero in Logit Kernel model, but non-zero in the model proposed in this study.

It should be noted that, it is feasible in methodology to add another random term in the proposed model to consider the correlations caused by route overlaps, as that in Logit Kernel model. However, this will significantly increase the model estimation time. Since this study mainly focus on the heterogeneity, to keep an affordable computational expenditure, we only consider the overlap problem in the deterministic part with the Path-size specification rather than giving a more complicate specification of error term.

6.1.5 Simulation-based estimation

Using the same notation as presented in above, the log-likelihood function of a mixed logit model is given by

$$LL(\Omega) = \ln\left[E\left(\prod_{n=1}^{N} \prod_{m=1}^{M_n} \prod_{t=1}^{T_{n,m}} P_{n,m,t}(i_{n,m,t} \mid \beta_{n,m,t}) \right) \right] \tag{6.11}$$

where N is the number of travelers, M_n is the number of O-D pairs for traveler n, and $T_{n,m}$ is the number of observations for traveler n for O-D pair m.

With the cross-sectional specification, it is assumed that each choice situation is independent of all other choice situations, even if two choice situations relate to the same respondent. With this assumption, and replacing expectation with integration, then Eq. (6.11) is rewritten as

$$LL(\Omega) = \sum_{n=1}^{N} \sum_{m=1}^{M_n} \sum_{t=1}^{T_{n,m}} \ln\left[\int_{\beta_{n,m,t}} P_{n,m,t}(i_{n,m,t} \mid \beta_{n,m,t}) f(\beta_{n,m,t} \mid \Omega) d\beta_{n,m,t} \right] \tag{6.12}$$

With the specification proposed by Revelt and Train (1998), the independent assumption of observations from the same traveler is relaxed. Only the inter-traveler observations are assumed to be independent, then Eq. (10) is rewritten as

$$LL(\Omega) = \sum_{n=1}^{N} \ln\left[\int_{\beta_n} \prod_{m=1}^{M_n} \prod_{t=1}^{T_{n,m}} P_{n,m,t}(i_{n,m,t} \mid \beta_n) f(\beta_{n,m,t} \mid \Omega) d\beta_{n,m,t} \right] \tag{6.13}$$

With the specification proposed by Hess and Rose (2009), the log-likelihood function is given by

$$LL(\Omega)=\sum_{n=1}^{N}\ln\left\{\int_{\delta_n}\left[\left(\prod_{m=1}^{M_n}\prod_{t=1}^{T_{n,m}}\int_{\eta_{n,m,t}}\begin{matrix}P_{n,m,t}(i_{n,m,t}\mid\beta_{n,m,t})\\ k(\eta_{n,m,t}\mid\Omega_\eta)d\eta_{n,m,t}\end{matrix}\right)\right]g(\delta_n\mid\Omega_\delta)d\delta_n\right\} \tag{6.14}$$

where $k(\eta_{n,m,t}\mid\Omega_\eta)$ and $g(\delta_n\mid\Omega_\delta)$ are the probability density functions of $\eta_{n,m,t}$ and δ_n with parameters Ω_η and Ω_δ, respectively.

$$LL(\Omega)=\sum_{n=1}^{N}\ln\left\{\begin{matrix}\int_{\delta_n}\left[\prod_{m=1}^{M_n}\int_{\varphi_{n,m}}\begin{matrix}\left(\prod_{t=1}^{T_{n,m}}\int_{\eta_{n,m,t}}\begin{matrix}P_{n,m,t}(i_{n,m,t}\mid\beta_{n,m,t})\\ \cdot k(\eta_{n,m,t}\mid\Omega_\eta)d\eta_{n,m,t}\end{matrix}\right)\\ \cdot h(\varphi_{n,m}\mid\Omega_\varphi)d\varphi_{n,m}\end{matrix}\right]\\ \cdot g(\delta_n\mid\Omega_\delta)d\delta_n\end{matrix}\right\}- \tag{6.15}$$

With the specification proposed in this study, the log-likelihood function is given by

where $h(\varphi_{n,m}\mid\Omega_\varphi)$ is the probability density functions of $\varphi_{n,m}$ with parameters Ω_φ.

Because the integrals do not take a closed form, the log-likelihood functions are approximated by simulation. Then Eqs. (6.12-15) are approximated by Eqs. (6.16-19), respectively:

$$SLL(\Omega)=\sum_{n=1}^{N}\sum_{m=1}^{M_n}\sum_{t=1}^{T_{n,m}}\ln\left[\frac{1}{R}\sum_{r=1}^{R}P_{n,m,t}(i_{n,m,t}\mid\beta_{r,n,m,t})\right] \tag{6.16}$$

$$SLL(\Omega)=\sum_{n=1}^{N}\ln\left[\frac{1}{R}\sum_{r=1}^{R}\prod_{m=1}^{M_n}\prod_{t=1}^{T_{n,m}}P_{n,m,t}(i_{n,m,t}\mid\beta_{r,n})\right] \tag{6.17}$$

$$SLL(\Omega) = \sum_{n=1}^{N} \ln \left\{ \frac{1}{R} \sum_{r=1}^{R} \left[\prod_{m=1}^{M_n} \prod_{t=1}^{T_{n,m}} \frac{1}{H} \sum_{h=1}^{H} P_{n,m,t}(i_{n,m,t} \mid \delta_{r,n}, \eta_{h,m,n,t}) \right] \right\} \tag{6.18}$$

$$SLL(\Omega) = \sum_{n=1}^{N} \ln \left\{ \frac{1}{R} \sum_{r=1}^{R} \left[\prod_{m=1}^{M_n} \frac{1}{L} \sum_{l=1}^{L} \left(\prod_{t=1}^{T_{n,m}} \frac{1}{H} \sum_{h=1}^{H} \begin{matrix} P_{n,m,t}(i_{n,m,t} \mid \delta_{r,n}, \\ \varphi_{l,m,n}, \eta_{h,m,n,t}) \end{matrix} \right) \right] \right\} \tag{6.19}$$

For Eq. (6.17), there is an alternative that utilizes the cross-sectional formulation but, instead of taking different draws for each choice by a given traveler, uses the same draws in all the choice situations for the same person. Under this approach (Paag et al., 2000), SLL is written

$$SLL(\Omega) = \sum_{n=1}^{N} \sum_{m=1}^{M_n} \sum_{t=1}^{T_{n,m}} \ln \left[\frac{1}{R} \sum_{r=1}^{R} P_{n,m,t}(i_{n,m,t} \mid \beta_{r,n}) \right] \tag{6.20}$$

The only difference in comparison with Eq. (15) is the subscript of β. In Eq. (19), the same set of R draws is reused in the simulation of all choices for traveler n, thus leading to a requirement for NR draws, which is different from the $R \cdot \sum_{n=1}^{N} \sum_{m=1}^{M_n} T_{n,m}$ draws for Eq. (6.16).

This approach attempts to accommodate the panel nature of the data by reusing the same draws across choices for a given traveler. Similarly, an approximation of Eq. (19) can be given as

$$SLL(\Omega) = \sum_{n=1}^{N} \sum_{m=1}^{M_n} \sum_{t=1}^{T_{n,m}} \ln \left\{ \frac{1}{R} \sum_{r=1}^{R} P_{n,m,t}(i_{n,m,t} \mid \delta_{r,n}, \varphi_{r,m,n}, \eta_{r,m,n,t}) \right\} \tag{6.21}$$

Eq. (6.21) incorporates all simulations at the level of individual choices, but the same draws of δ are reused across choices for the same traveler. For φ, draws are reused across choices between the same O-D pair for the same traveler. For η, new draws are used in each choice situation.

Hess and Train (2011) concluded that, although this approach is computationally attractive, it is unable to recover the true patterns of heterogeneity because, as the number of draws increases, it becomes simply a cross-sectional estimator in which the three forms of heterogeneity will not be distinguished. In this study, following their suggestion, we will use the correct specification of the simulated log-likelihood function for the empirical analysis, although the accuracy of Eq. (6.21) also will be tested from the aspect of model fit.

6.2Data

6.2.1 Observations

The GPS data used in this study is the same data that used in last chapter. However, in this study, since there is a more complete description on taste heterogeneity, more travelers are selected as the subjects for this study. 95 drivers who made trips every month in the period 2011.3 to 2011.12 are selected as the subjects for this study. After a basic data cleaning process, a data set with 52,330 trips was constructed.

The raw GPS coordinates are map-matched to a sequence of links. There is an advanced processing method of GPS data that avoids ambiguity in map matching (Bierlaire and Frejinger, 2008). However, its application remains a task for future studies.

For each trip, the outflow node of the first link and the inflow node of the last link are treated as the origin and destination, respectively. Many of these trips start or/and finish outside the target area. In those cases, only the part of the trip within the target area is considered and that part is treated as a complete trip. For each driver, the trips with the same or adjacent first and last links are considered to share the same O-D pairs.

Unlike some choice situations, such as travel mode choice, route choice presents a large choice set. Because of computational issues, it is very difficult to use all of the 52,330 extracted observations to estimate the complicated mixed logit models. Therefore, we extract a small data set for the empirical analysis.

The small data set is constructed as follows. First, for every driver, O-D pairs are random selected. At most 10 unique O-D pairs are selected for each driver. Then, for each O-D pair, at most 10 unique observations are selected at random. The small data set resulting from this process consist of 2,182 observations.

6.2.2 Choice set generation

For each trip, before analyzing route choice behavior, a choice set must be generated. Several choice set generation algorithms have been proposed and evaluated for route choice analysis (Bekhor et al., 2006). As the same as in last chapter, the random walk is adopted in this study. However, we make some modifications on the original random walk method. In this study, different from that in last chapter, we assume that the consideration choice set is the sampled set of paths, but not the universal set of all the available paths.

Given an origin-destination pair (s_o, s_d), a path consisting of an ordered set of links denoted by Γ is generated using the following algorithm:

Initialization: Set current node $v{=}s_o$, $\Gamma = \varnothing$;

For every link l in the network, set $C(l) = C_0(l)$, where $C(l)$ is the cost of link l used in the **Loop** step; $C_0(l)$ is the original cost of link l.

Loop: While $v \neq s_d$ perform the following:

Weights For every link $l = (v,w) \in E_v$, where E_v is the set of outgoing links from v. Then, for each link, the weight is calculated:

$$w(l \mid b_1, b_2) = 1 - (1 - x_l^{b_1})^{b_2} \tag{6.22}$$

$$x_l = \frac{SP(v, s_d)}{C(l) + SP(w, s_d)} \tag{6.23}$$

$C(l)$ is the cost of link l. $SP(\cdot)$ is the length of the shortest path between the two nodes. In this study, b_1=5, b_2=1.

If $\sum_{k\in\Gamma} C_0(k) + C_0(l) + SP(w, s_d) > 1.5 \cdot SP(s_o, s_d)$, then set $w(l \mid b_1, b_2) = 0$.

Set the cost of every link $l = (v, w) \in E_v$ to be $C(l)$=∞.

Probability For each link $l = (v, w) \in E_v$, the probability of choosing a link is

$$q(l \mid E_v, b_1, b_2) = \frac{w(l \mid b_1, b_2)}{\sum_{l^{'}\in E_v} w(l^{'} \mid b_1, b_2)} \tag{6.24}$$

Draw Randomly select a link $(v, w)^*$ in E_v based on the above probability.

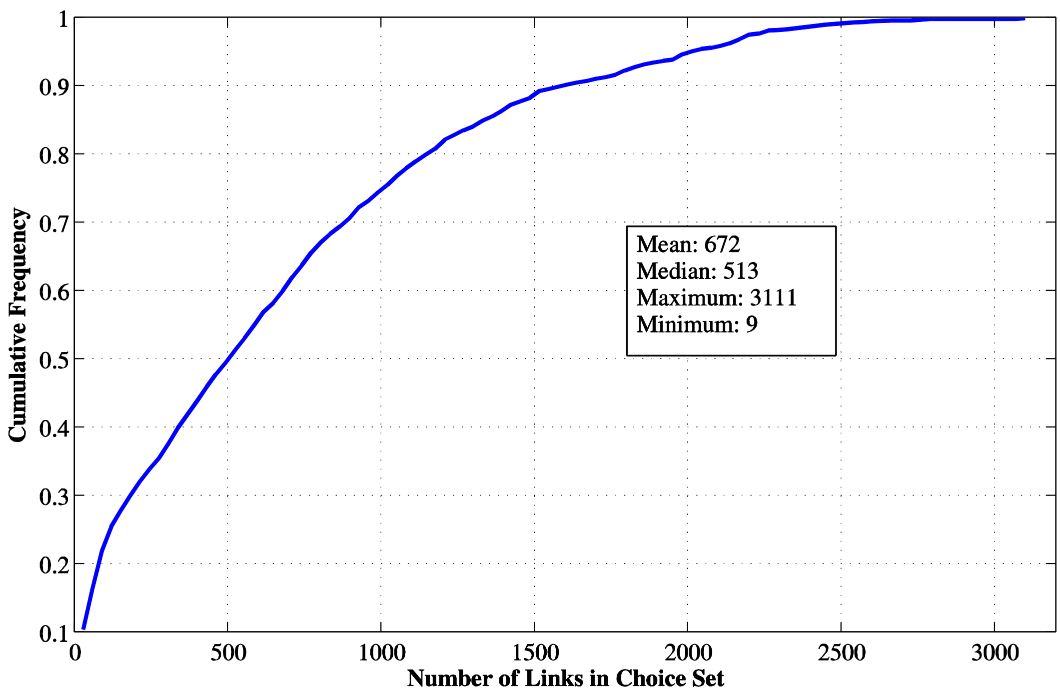
Update path $\Gamma = \Gamma \cup (v, w)^*$.

For each trip, 50 paths are randomly drawn. Repeated draws are discarded. In contrast with the original version of the random walk method, two constraints are set in the **Weights** step. In order to avoid cyclic routes, the cost of outgoing links is set to ∞ at the last of **Weights** step. Only paths satisfying the detour constraint are used to construct a choice set for use in route choice analysis, which means that the sampled path must be shorter than 1.5 times of the shortest path. Figure 6.1 gives a detail description on the generated choice set.

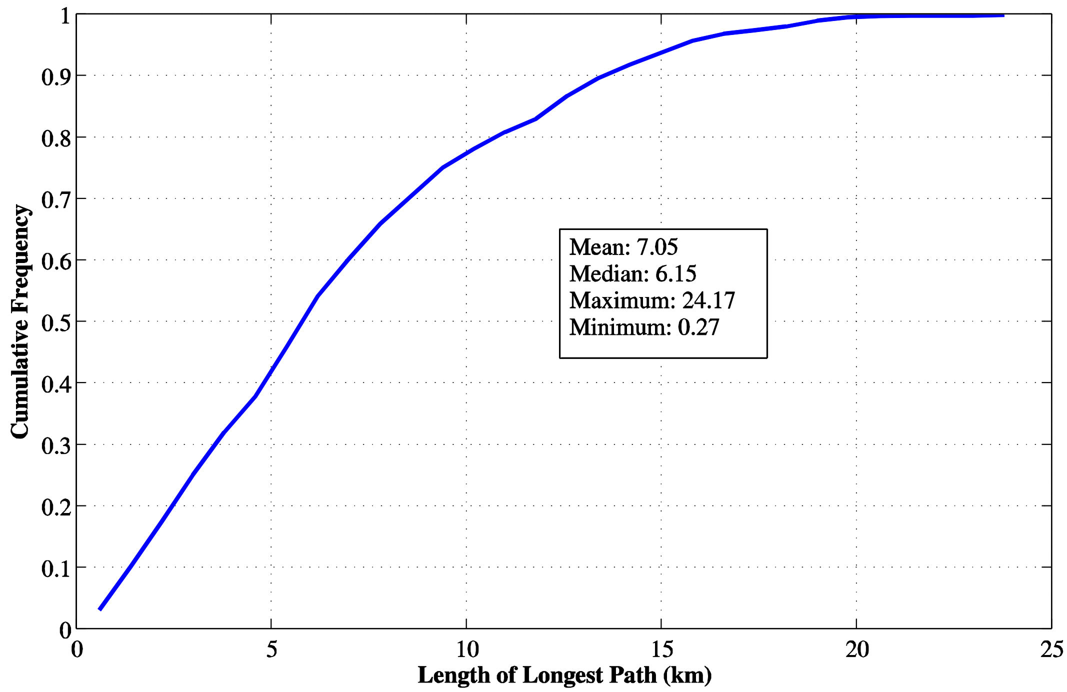
6.2.3 Data set

The route characteristics, as well as the characteristics of drivers and O-D pairs, to be used in route choice analysis are shown in Table 6.1.

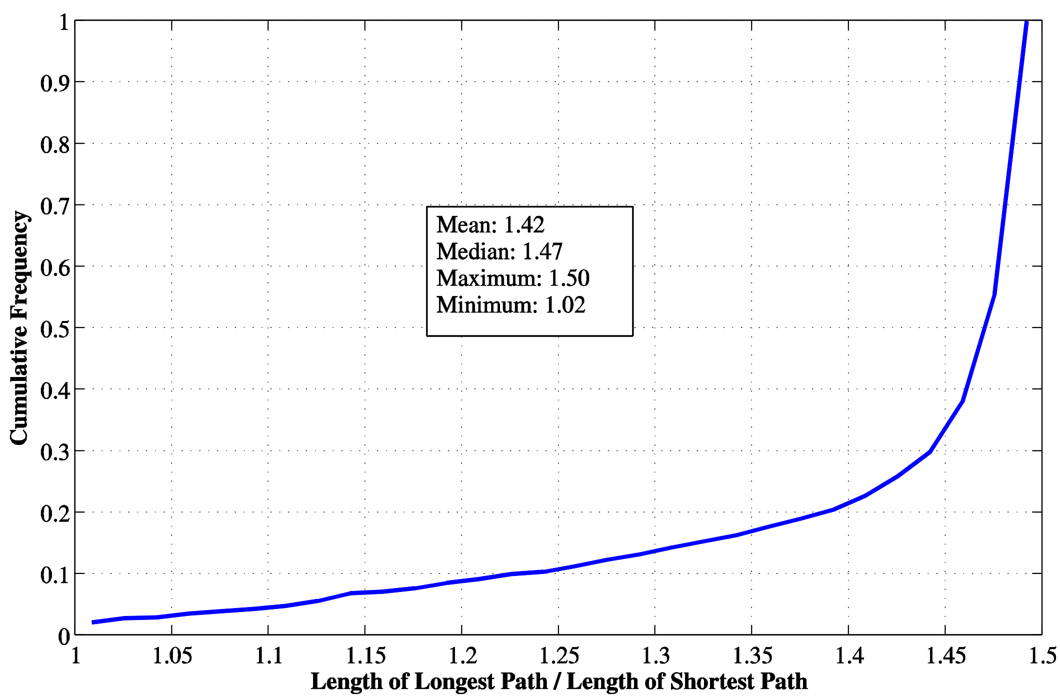
It should be noted that, as always a problem with studies on real large road network, the data about real time or congested link travel time are not available for this study. Drivers are often assumed to choose route according to the route travel time perceived at the time that they make the choice. In some previous studies, the probe data are collected by a large number of taxies; therefore they can estimate real time link travel time for the whole networks and then include the real time route travel time in route choice analysis(Li et al., 2013; Morikawa and Miwa, 2006). However, in this study, the GPS data are collected by only about one hundred private cars, which is not enough for the estimation of real time link travel time. In some other studies, the estimated peak hour link travel time is included to the effect of congestion (Ramming, 2001). However, the network data available for us only include some basic attributes of links, without the information of congested travel time. Therefore, due to lack of data, like some previous research(Bierlaire and Frejinger, 2008), only free travel time is considered in this study.



(a) The Distribution of Link Number.



(b) The Distribution of Longest Path Length



(c) The Distribution of Longest Path Length/Shortest Path Length

Figure 6.1 Descriptions on the Choice Set

Table 6.1 Description of Characteristics Used for Route Choice Analysis

Characteristic	Description	Value
Route specific		
Free_Travel_Time	Free travel time for the route* (unit: minute)	-
Number_of_Intersections	The number of intersections along the route	-
Traveler specific		
Gender	Male (90%)	1
	Female (10%)	0
Age	<= 35 years old (17%)	0
	> 35 years old (83%)	1
Car displacement	<=1.79 L (55%)	0
	>1.79 L (45%)	1
O-D specific		
Distance	The distance from O to D (unit: km)	0.44~34.22
Familiarity	1 + the number of times that the driver has traveled between the O-D pair (unit: 100 times)	0.01~1.9

* calculated according to the free-flow speed of the link, which is determined by the characteristics of the street layout.

6.3Model Specifications

A total of 13 models are estimated in this analysis, ranging from a basic path-size model to the specification given in Eq. (6.8). A summary of the different structures is provided in Table 6.2.

Table 6.2 Summary of Estimated Models

Model	Observed heterogeneity		Unobserved heterogeneity		
	Indv. specific	OD specific	Indv. specific	OD specific	Choice specific
0	×	×	×	×	×
O1	○	×	×	×	×
O2	○	○	×	×	×
U1	×	×	○	×	×
U2	×	×	×	○	×
U3	×	×	×	×	○
U13	×	×	○	×	○
U123	×	×	○	○	○
UO1	○	○	○	×	×
UO2	○	○	×	○	×
UO3	○	○	×	×	○
UO13	○	○	○	×	○
UO123	○	○	○	○	○

The systematic part of the utility without unobserved heterogeneity (i.e. the specification of Model O2) is given as:

$$\begin{aligned} V = {} & (\text{Constant} + \alpha_{11} \cdot \text{Age} + \alpha_{12} \cdot \text{Gender} + \alpha_{13} \cdot \text{Displacement} \\ & + \alpha_{14} \cdot \text{Distance} + \alpha_{15} \cdot \text{Familiarity}) \cdot \text{Free\_Travel\_Time} \\ & + (\text{Constant} + \alpha_{21} \cdot \text{Age} + \alpha_{22} \cdot \text{Gender} + \alpha_{23} \cdot \text{Displacement} \\ & + \alpha_{24} \cdot \text{Distance} + \alpha_{25} \cdot \text{Familiarity}) \cdot \text{Number\_of\_Intersections} \end{aligned} \tag{6.25}$$

Then the specifications of the estimated models can be determined according to Section 6.1.

6.4Estimation Results and Analysis

6.4.1 Findings from the estimation results

Using the extracted observations, all of the 13 models are estimated. Given the high number of models, it is not possible to present detailed estimation results for each single model. Therefore, we only give an overview of the results across models, as shown in Table 6.3. The last row of Table 3 (UO123*) is the performance of Model UO123 estimated using Eq. (6.21). Detailed estimation results of several selected models are shown in Table 6.4.

Several findings can be obtained from Table 6.3, which evaluates the models with respect to goodness of fit.

The first significant observation to be made is the significant improvement in model fit that occurs when O-D pair characteristics are considered in the observed heterogeneity. As noted in the introduction section, most previous research concerning taste heterogeneity in route choice takes into account travelers' characteristics only, because it was based on experiments using a single O-D pair. Few previous investigations concern heterogeneity that is O-D pair specific. However, this finding implies that O-D pair specific heterogeneity may have a much greater effect on route choice behavior than individual characteristics. One very important application of route choice models is as a core part of traffic assignment. Previous research (Chen et al., 2012; Miwa et al., 2010) has demonstrated the effect of O-D specific route choice models on traffic assignment.

This finding provides empirical evidence for the necessity of applying O-D characteristics in structured route choice models used for traffic assignment.

Table 6.3 Summary of Model Performance

Model	Log-likelihood at estimates	Count of parameters	Adjusted Rho^2
0	-5588.888	3	0.277
O1	-5238.737	9	0.321
O2	-4361.250	13	0.434
U1	-4878.550	5	0.368
U2	-5161.488	5	0.332
U3	-5569.032	5	0.279
U13	-4877.656	7	0.368
U123	-4704.112	9	0.390
UO1	-4013.565	15	0.479
UO2	-4134.722	15	0.463
UO3	-4360.373	15	0.434
UO13	-4000.761	17	0.480
UO123	-3920.780	19	0.490
UO123*	-4293.357	19	0.442

Table 6.4 Estimation Results of Selected Models

	Parameter	Estimation (t-stat.)			
		0	O2	U123	UO123
Free travel time	Constant	-0.104 (-6.409)	-1.498 (-12.715)	-0.254 (-8.667)	-1.967 (12.868)
	Age (α_{11})		-0.655 (-12.831)		-0.456 (-6.732)
	Gender (α_{12})		0.723 (7.229)		0.817 (6.386)
	Displacement (α_{13})		0.140 (3.590)		0.283 (4.618)
	Distance (α_{14})		0.104 (21.316)		0.150 (22.012)
	Familiarity (α_{15})		-0.271 (-5.880)		-0.552 (-8.771)
	Std. d(δ)			0.762 (19.534)	0.636 (13.149)
	Std. d(φ)			0.621 (19.013)	0.659 (15.277)
	Std. d(η)			0.021 (0.454)	0.023 (0.657)
Number of intersections	Constant	-0.067 (-24.329)	-0.098 (-5.552)	0.088 (-17.892)	-0.053 (-2.447)
	Age (α_{21})		0.033 (3.282)		-0.005 (-0.373)
	Gender (α_{22})		-0.073 (-4.866)		-0.119 (-6.741)
	Displacement (α_{23})		0.005 (0.724)		0.009 (1.036)
	Distance (α_{24})		0.007 (9.567)		0.003 (3.759)
	Familiarity (α_{25})		0.022 (3.252)		0.062 (6.932)
	Std. d(δ)			0.054 (18.320)	0.051 (13.848)
	Std. d(φ)			0.110 (16.736)	0.070 (11.088)
	Std. d(η)			0.002 (0.318)	0.003 (0.467)
Ln(PS)		2.105 (50.934)	2.199 (47.074)	2.100 (44.355)	2.195 (43.311)

The next observation is that all models in which only unobserved heterogeneity is considered have a comfortably higher log-likelihood than Model 0 (the largest p-value in the likelihood ratio test is only 2.38×10^{-9}). However, they all have much lower log-likelihood than Model O2 (where the largest p-value in the likelihood ratio test approaches 0), which considers only observed heterogeneity. This finding confirms that the random coefficients specification can enhance the performance of the logit model significantly. However, it also indicates that incorporating random taste heterogeneity cannot replace the observed characteristics of travelers and O-D pairs. Considering the difficulties involved in the estimation and implementation of mixed logit models, this suggests that the more efficient way to enhance the performance of route choice models is to add more observed characteristics of travelers and O-D pairs rather than increasing the complexity of the specification.

It also should be noted that, compared with Model O2, Model UO3 has a slightly higher log-likelihood. However, the p-value in the likelihood ratio test is 0.416. This means that although there is an improvement in model fit, it is only significant at the 58.4% level. But all other models incorporating both unobserved and observed taste heterogeneity have a considerably higher log-likelihood than Model UO2 (where the largest p-value in the likelihood ratio test value approaches 0). Models U3 and UO3 are in fact the mixed logit models that treat the repeat choice observations as inter-sectional data. Recalling that Model U3 also has the worst performance among the models incorporating only unobserved heterogeneity, this suggests that when dealing with panel data, it is not appropriate to assume each choice situation is independent of all other choice situations.

We now proceed to look at the structure of random taste heterogeneity. Compared with Model U1, Model U13 with the specification suggested by Hess and Rose (2009) has improved model fit. However, this improvement is not statistically significant (the p-value is 0.409). Model U123, which also considers intra-OD heterogeneity, has a considerably higher log-likelihood than Model U13 (where the p-value approaches 0). A similar finding is obtained from the estimation results of the UO series of models. This suggests that it is desirable to add inter-OD heterogeneity while incorporating intra-traveler heterogeneity.

Model U123 has the best fit among the U series of models. This indicates that it is necessary to combine inter-traveler, inter-OD, and inter-choice heterogeneity when specifying random taste heterogeneity. Looking at goodness of fit, Model UO123 is significantly better than all of the other models. This suggests that it is desirable to incorporate both observed and unobserved heterogeneity.

The fact that Model UO123 has the best fit also indicates that all other models are mis-specified. Concerning the estimation results of Model UO123*, the log-likelihood is much lower than Model UO123, and only Model UO3 has a lower value among the UO series of models. This is consistent with the finding of Hess and Train (2011), and does not support the use of the simplified simulated log-likelihood function.

The estimated models shown in Table 6.4 are applied to a specific scenario to illustrate the potential bias introduced by over-simplifying the assumptions of homogeneity. In this scenario, we assume that a driver is to make choice among three routes without overlap:

Route 1: free travel time is 10 minutes, number of intersections is 20 (lower free travel time, more intersections);

Route 2: free travel time is 15 minutes, number of intersections is 15 (greater free travel time, fewer intersections);

Route 3: free travel time is 10 minutes, number of intersections is 15 (lower free travel time, fewer intersections).

This driver is assumed to be male and older than 35 years old. His car has a displacement of 1.6 L. The distance between origin and destination is 5km. This is the first time that he has traveled between this O-D pair.

Figure 6.2 presents the expected route shares on the three routes for each of the three models. It is clear that the predictions obtained using these models are quite different. According to the likelihood ratio test, Model UO123 is considered the most accurate model. The significant variation in route shares given by the other models suggests that any simplifying of heterogeneity considerations will result in incorrect predictions.

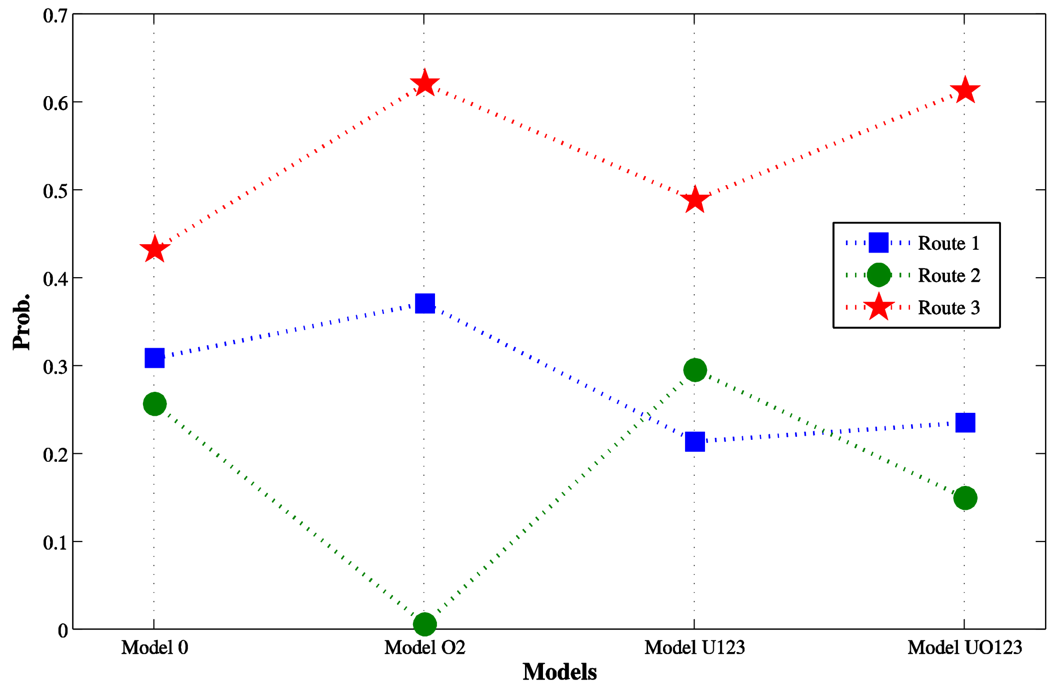
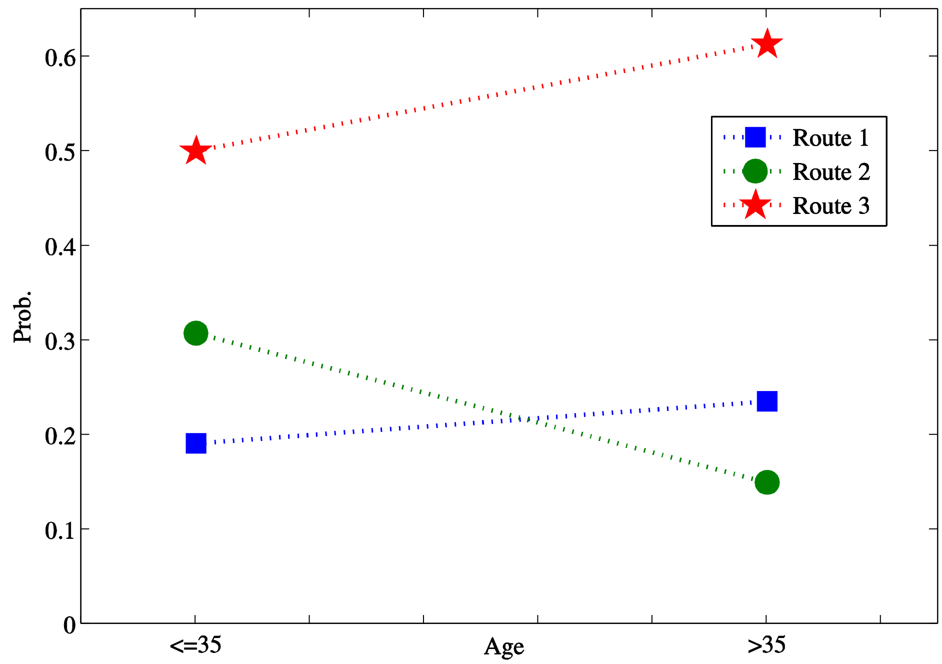


Figure 6.2 Estimated Route Shares

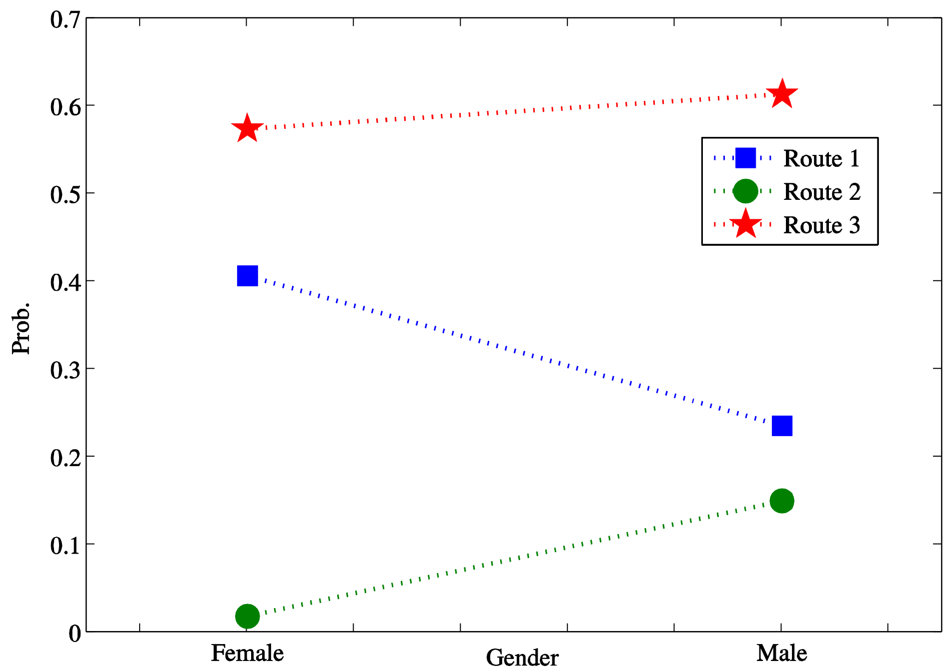
Moving on to details of the estimated parameters, several behavioral findings can be obtained. Figure 6.3 shows the effect of observed heterogeneity on route choice prediction when applying Model UO123. This shows that, for the same route characteristics, different drivers will make significantly different choices.

First, as expected, the constant part of the estimated parameter has a negative sign in each case. Then, we look at the sign of the observed individual and O-D pair specific characteristics.

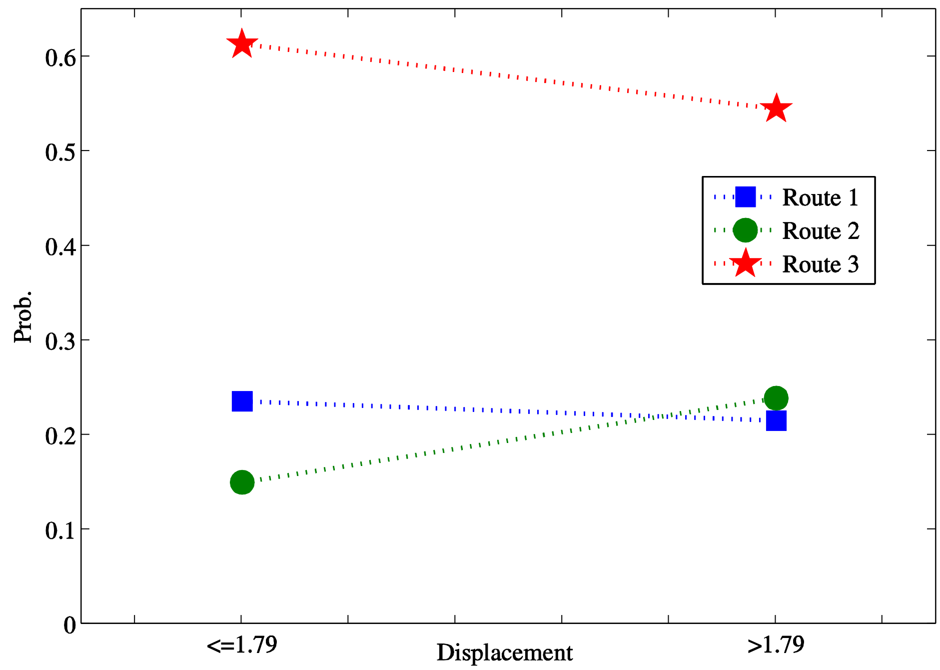
The negative sign and t-statistic of α_{11} suggest that age will affect the taste for free travel time significantly, with older people being more sensitive to free travel time. The low t-statistic of α_{21} implies that age does not have a significant effect on the taste for the number of intersections at the level of 95%.



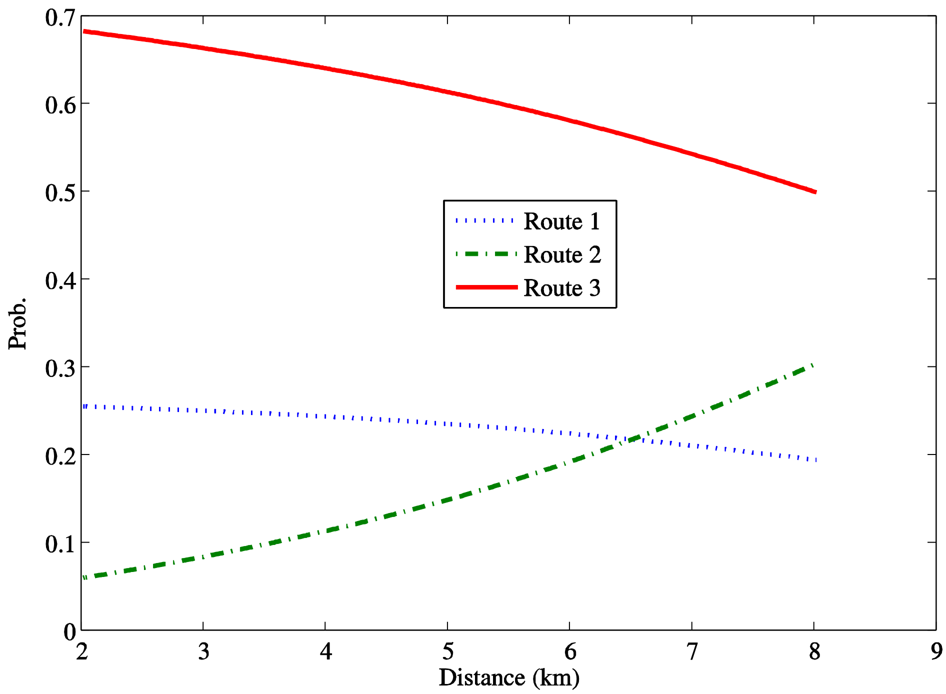
(a) Age



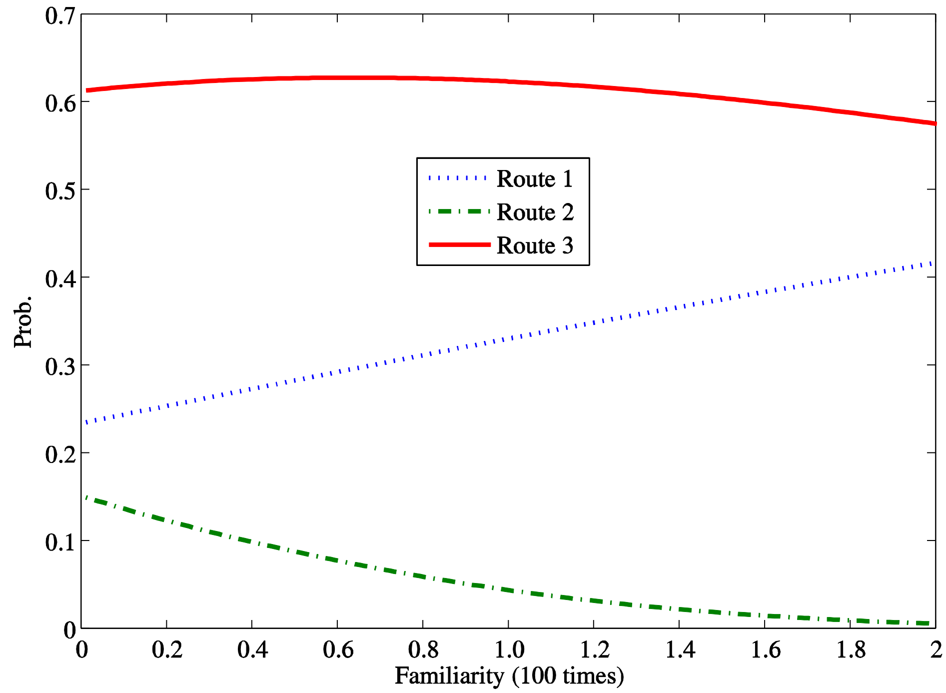
(b) Gender



(c) Displacement



(d) Distance



(e) Familiarity

Figure 6.3 Estimated route shares

As shown in Figure 6.3 (a), as a driver's age moves from younger than 35 to older than 35, the differences between route shares are larger. This finding is a little counterintuitive, since it is often assumed that the younger driver will be more sensitive to the observed cost. Here is a possible explanation: Since the scale parameter cannot be identified from taste coefficients in Logit models, all the taste coefficients have less absolute value can be caused by a smaller scale parameter, which means a larger variance of the error term. Therefore the negative signs of α_{11} and α_{21} can be interpreted as meaning that there are more unobserved factors affect younger drivers' route choice behavior than the older drivers.

According to the t-test, gender has a significant effect on both free travel time and number of intersections. Male drivers are more sensitive to the number of intersections, but less sensitive to free travel time. As found in Figure 6.3 (b), since Route 1 offers lower free travel time but more intersections, for male drivers, the expected route share of

Route 1 is much lower than for female drivers.

Car displacement also has a significant effect on free travel time. However its effect on the number of intersections is not significant. The positive signs of α_{13} and α_{23} mean that drivers of vehicles with larger displacement will be less sensitive to route cost. From Figure 6.3 (c), we find that the expected share of Route 2 is higher than that of Route 1 for drivers with larger displacement vehicles, while for drivers with lower displacement vehicles, the expected share of Route 2 is the lowest. This indicates that the effect of displacement on taste for free travel time is larger than that on the taste for intersection count.

Both α_{14} and α_{24} are positive and statistically significant. This confirms that as the distance between the O-D pair increases, drivers will be less sensitive to the observed attributes. This also can be interpreted as an increasing variance in the unobserved utility, which is consistent with the assumptions made in previous research on traffic assignment (Chen et al., 2012; Miwa et al., 2010).

α_{15} and α_{25} are negative and positive, respectively, and both are significant. This implies that when driving between more familiar O-D pairs, drivers will be more sensitive to free travel time but less sensitive to the number of intersections. We find in Figure 4(e) that, as the familiarity of the O-D pair increases, the expected share of Route 1 increases as compared with Route 2.

Finally, it should be noted that the absolute value of the parameters in Model U123 is much larger that of Model 0. This is because Model U123 decomposes the unobserved portion of utility into taste heterogeneity.

6.4.2 The stability of behavioral findings

In this study, Path-size model and Random Walk method are chosen to overcome the overlap and choice set generation problems, respectively. Because of the computational issues, we only use a small part of the all extracted observations.

In this section we will check the stability of the behavioral findings in the above analysis for different route choice models, choice set generation methods and datasets.

As shown in Table 6.5, with the heterogeneity specification same as Model O2, Generalized Nested Logit (GNL) model (Bekhor and Prashker, 2001), C-logit model (Cascetta et al., 1996), Logit Kernel Model and Logit Kernel model with Path-size (Bekhor et al., 2002) are estimated using the some dataset described in Section 3. Model UO123, the most complicate model proposed in this study is not used here because of the unaffordable computational burden caused by combining the GNL or Logit Kernel Model with the multi-level random parameter specification.

At first, we can find that, although the estimates are different with different models, the signs are consistent. Therefore the behavioral findings, which are mainly based on the signs of estimates, are not dependent on the choice models. Second, from aspect of model fit, Path-size Logit model is superior to other models except the Logit Kernel model with path-size. This is consistent to the findings in Ramming (2001). At last, from the computational aspect, GNL model and Logit Kernel model take much longer time for estimation than Path-size model and C-logit model.

Therefore, the behavioral findings are stable for different route choice models. Path-size Logit model is a reasonable choice when consider both the model fit and computational efficiency.

The route choice model estimates are found to be sensitive to the choice set compositions (Bliemer and Bovy, 2008; Prato and Bekhor, 2007). Therefore, we generate two more choice sets for the dataset described in Section 3 (denoted as Data 0), using link penalty method (Park and Rilett, 1997) and simulation method (Ramming, 2001) respectively. Using the random walk method, we generated two datasets with different random selected observations (denoted as Data 1 and Data 2). Observations in Data 0, Data 1 and Data 2 do not have any intersections because the already selected observations are excluded when random extract a new small dataset. The Labeling method (Ben-Akiva et al., 1984) is not applied because there is no enough data about “labels” to generate routes. The branch-and-bound method (Prato and Bekhor, 2006) is not applied because it is too time consuming in such a dense network (Rieser-Schüssler et al., 2012). The estimation results are shown in Table 6.6.

Table 6.5 Estimation Results of Model O2 with Different Specifications

	Parameter	Estimation (t-stat.)				
		GNL	Logit Kernel	C-logit	Path-size Logit	Logit Kernel with Ln(PS)
Free travel time	Constant	-0.961 (-8.240)	-2.396 (-9.980)	-1.214 (-10.521)	-1.498 (-12.715)	-1.660 (-12.234)
	Age (α_{11})	-0.556 (-10.539)	-1.497 (-12.487)	-0.608 (-11.811)	-0.655 (-12.831)	-0.734 (-11.911)
	Gender (α_{12})	0.855 (8.402)	1.736 (9.363)	0.810 (8.116)	0.723 (7.229)	0.797 (7.219)
	Displacement (α_{13})	0.157 (4.108)	0.257 (2.909)	0.155 (4.142)	0.140 (3.590)	0.141 (3.141)
	Distance (α_{14})	0.065 (14.875)	0.166 (15.067)	0.080 (17.412)	0.104 (21.316)	0.118 (18.890)
	Familiarity (α_{15})	-0.203 (-4.458)	-0.677 (-6.372)	-0.247 (-5.652)	-0.271 (-5.880)	-0.310 (-5.553)
Number of intersections	Constant	-0.080 (-4.384)	-0.184 (-5.206)	-0.083 (-4.729)	-0.098 (-5.552)	-0.103 (-5.103)
	Age (α_{21})	0.030 (2.912)	0.082 (4.274)	0.037 (3.664)	0.033 (3.282)	0.041 (3.510)
	Gender (α_{22})	-0.115 (-7.352)	-0.228 (-8.048)	-0.094 (-6.394)	-0.073 (-4.866)	-0.089 (-5.297)
	Displacement (α_{23})	0.000 (0.045)	0.017 (1.177)	0.001 (0.090)	0.005 (0.724)	0.006 (0.735)
	Distance (α_{24})	0.008 (12.843)	0.016 (10.315)	0.007 (9.968)	0.007 (9.567)	0.007 (8.017)
	Familiarity (α_{25})	0.026 (3.400)	0.073 (5.041)	0.028 (3.991)	0.022 (3.252)	0.026 (3.103)
Ln(PS)					2.199 (47.074)	2.188 (43.230)
Nesting Parameter		2.602 (4.23)				
Gaussian Covariance Parameter			2.052 (19.838)			0.454 (10.098)
Commonality Factor				1.793 (33.444)		
Log-likelihood at estimates		-4670.18	-5321.35	-5015.56	-4361.25	4316.35
Adjusted Rho^2		0.394	0.310	0.349	0.434	0.440
Estimation time		13 hours	7 hours	<5 mins	<5 mins	7 hours

Table 6.6 Estimation Results of Model UO123 with Different Choice Sets and Datasets

	Parameter	LP (Data 0)	SIM (Data 0)	RW (Data 1)	RW (Data 2)	RW (Data0)
		Estimation (t-stat.)				
Free travel time	Constant	-1.378 (-9.847)	-0.375 (-3.675)	-2.467 (-16.928)	-1.991 (-14.471)	-1.967 (12.868)
	Age (α_{11})	-0.465 (-9.892)	-0.483 (-9.381)	-0.359 (-5.123)	-0.165 (-2.348)	-0.456 (-6.732)
	Gender (α_{12})	1.078 (9.979)	0.794 (8.253)	0.735 (6.474)	0.653 (6.859)	0.817 (6.386)
	Displacement (α_{13})	0.307 (7.183)	0.158 (3.900)	0.117 (1.890)	0.094 (1.956)	0.283 (4.618)
	Distance (α_{14})	0.040 (4.608)	0.005 (5.140)	0.175 (23.066)	0.145 (20.729)	0.150 (22.012)
	Familiarity (α_{15})	-0.216 (-5.028)	-0.188 (-4.964)	-0.260 (-1.906)	-0.970 (-7.299)	-0.552 (-8.771)
	Std. d(δ)	0.397 (18.323)	0.369 (16.875)	0.506 (11.923)	0.439 (13.170)	0.636 (13.149)
	Std. d(φ)	0.515 (21.696)	0.402 (18.926)	0.658 (16.294)	0.510 (10.366)	0.659 (15.277)
	Std. d(η)	0.084 (2.613)	0.058 (3.380)	0.156 (3.707)	0.012 (0.430)	0.023 (0.657)
Number of intersections	Constant	-0.026 (-9.600)	-0.013 (-1.006)	-0.074 (-3.841)	-0.091 (-4.739)	-0.053 (-2.447)
	Age (α_{21})	-0.004 (-4.513)	**0.001 (0.488)**	**0.020 (1.134)**	**0.005 (0.407)**	-0.005 (-0.373)
	Gender (α_{22})	-0.150 (-12.164)	-0.091 (-8.517)	-0.041 (-2.594)	-0.050 (-3.449)	-0.119 (-6.741)
	Displacement (α_{23})	0.002 (2.548)	**-0.000 (-0.600)**	**-0.013 (-1.264)**	0.019 (2.225)	0.009 (1.036)
	Distance (α_{24})	0.002 (5.880)	0.000 (1.781)	0.001 (1.392)	0.004 (3.972)	0.003 (3.759)
	Familiarity (α_{25})	0.057 (7.866)	0.037 (6.250)	0.115 (3.109)	0.109 (5.020)	0.062 (6.932)
	Std. d(δ)	0.088 (16.697)	0.043 (13.077)	0.032 (8.176)	0.042 (10.692)	0.051 (13.848)
	Std. d(φ)	0.081 (19.595)	0.049 (14.347)	0.076 (12.956)	0.044 (2.968)	0.070 (11.088)
	Std. d(η)	0.008 (1.537)	0.008 (2.378)	0.021 (3.574)	0.001 (0.311)	0.003 (0.467)
Ln(PS)		2.262 (29.217)	1.904 (34.747)	2.160 (41.853)	2.074 (43.497)	2.195 (43.311)

According to the results shown in Table 6.6, it is confirmed again that, the estimation of route choice models is sensitive to the choice set generation methods. With different path sets, the estimates are significantly different.

However, the behavioral interpretations are mainly based on the signs of estimated parameters. Most of the estimates in Table 6 have consistent signs. Only the effect of *age* and *displacement* on the taste of *number of intersection*s is not stable. But according to the t-values, these two parameters are not significant.

From the tests and analysis above, it can be found that the behavioral findings in this study are stable for different route choice models, choice set generation methods and datasets.

6.5Conclusions and Future Directions

In this research, we explore taste heterogeneity in route choice behavior. Taste heterogeneity may be incorporated in route choice analysis by introducing observed individual socio-economic characteristics. However, it is very likely that some taste heterogeneity will remain even after accounting for any observed characteristics. To incorporate both observed and unobserved characteristics, a mixed logit based method is proposed, in which the taste coefficients are treated as random and structured as observed characteristics.

It is not appropriate to consider route choice as a one-shot choice problem, so the use of panel data to deal with the random element is also discussed. Random taste may consist of three components: traveler specific, O-D pair specific and choice situation specific. In most previous research, only traveler specific heterogeneity is considered when dealing with panel data. However, as discussed by Hess and Rose (2009), it is necessary to also consider intra-traveler heterogeneity. In the generalized model they proposed, the random coefficients are divided to two parts: traveler specific for consideration of inter-traveler heterogeneity and choice situation specific for consideration of intra-traveler heterogeneity. However, in the particular application of route choice, it is declared in this paper that intra-traveler heterogeneity should be divided into an O-D pair specific part and a choice situation specific part.

The data used in this study is GPS location data obtained from private vehicles in Toyota city, Japan. Route choice observations between multiple O-D pairs are extracted from map-matched trajectory data. The random walk method is used for choice set generation. Models with various assumptions about heterogeneity are estimated and compared. Empirical analysis suggests that, to enhance the performance of route choice models, it is more efficient to add more observed characteristics relating to travelers and O-D pairs than to increase the complexity of the specification. It is found inappropriate from the aspect of model fit to assume independence between choice situations when dealing with panel data. It is also proved that the incorporation of O-D pair specific unobserved taste heterogeneity can enhance the performance of a route choice model significantly.

Further, the empirical analysis supports the conclusion of Hess and Train (2011) that to guarantee recovery of the true patterns of heterogeneity, analysts should make use of the correct specification of the simulated log-likelihood function and avoid any shortcuts.

Looking at the estimations in more detail, it is found that age affects the taste for free travel time significantly, with older people being more sensitive to free travel time. However age does not have a significant effect on taste for the number of intersections. Gender significantly affects taste for both free travel time and number of intersections. Male drivers are more sensitive to the number of intersections, while they are less sensitive to free travel time. Car displacement has a significant effect on taste for free travel time, while its effect on the number of intersections is not significant. As distance between O-D pair increases, drivers become less sensitive to the observed attributes. When driving between more familiar O-D pairs, drivers are more sensitive to free travel time but less sensitive to the number of intersections. The stability of the behavioral findings is also checked. It is found that the findings in this study are stable for different route choice models, choice set generation methods and datasets.

The number of driver characteristics considered in this research is limited because of a lack of data. In further research, the GPS data might be combined with questionnaire data so as to take into account a greater number of behavioral terms. Further, in this study, the random coefficients are formed linearly so as not to make the model too complicated.

However, this assumption should be discussed in future work because, for example, with increasing distance, the mean taste coefficients of costs will be positive, which is not appropriate. Recently, Fosgerau et al. (2013) propose a link based route choice model with unrestricted choice set. It is also possible to discuss the heterogeneity in that framework.

References

Bekhor, S., Ben-Akiva, M., Scott Ramming, M. (2002) Adaptation of logit kernel to route choice situation. *Transportation Research Record: Journal of the Transportation Research Board* 1805, 78-85.

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

Bekhor, S., Prashker, J. (2001) Stochastic user equilibrium formulation for generalized nested logit model. *Transportation Research Record: Journal of the Transportation Research Board* 1752, 84-90.

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

Bhat, C.R. (1998) Accommodating variations in responsiveness to level-of-service measures in travel mode choice modeling. *Transportation Research Part A: Policy and Practice* 32, 495-507.

Bhat, C.R. (2000) Incorporating observed and unobserved heterogeneity in urban work travel mode choice modeling. *Transportation Science* 34, 228-238.

Bierlaire, M., Frejinger, E. (2008) Route choice modeling with network-free data. *Transportation Research Part C: Emerging Technologies* 16, 187-198.

Bliemer, M.C., Bovy, P.H. (2008) Impact of route choice set on route choice probabilities. *Transportation Research Record: Journal of the Transportation Research Board* 2076, 10-19.

Bogers, E.A.I. (2009) Traffic information and learning in day-to-day route choice. Proefschrift TU Delft TRAIL Thesis Series.

Cascetta, E., Nuzzolo, A., Russo, F., Vitetta, A. (1996) A modified logit route choice model overcoming path overlapping problems. Specification and some calibration results for interurban networks. *Proceedings of Internaional symposium on transportation and traffic theory*, pp. 697-711.

Chen, A., Pravinvongvuth, S., Xu, X., Ryu, S., Chootinan, P. (2012) Examining the scaling effect and overlapping problem in logit-based stochastic user equilibrium models. *Transportation Research Part A: Policy and Practice*.

Fosgerau, M., Frejinger, E., Karlstrom, A. (2013) A link based network route choice model with unrestricted choice set. *Transportation Research Part B: Methodological* 56, 70-80.

Hess, S., Rose, J.M. (2009) Allowing for intra-respondent variations in coefficients estimated on repeated choice data. *Transportation Research Part B: Methodological* 43, 708-719.

Hess, S., Train, K.E. (2011) Recovery of inter- and intra-personal heterogeneity using mixed logit models. *Transportation Research Part B: Methodological* 45, 973-990.

Li, D., Miwa, T., Morikawa, T. (2013) Dynamic route choice behavior analysis considering en route learning and choices (forthcoming). *Transportation Research Record: Journal of the Transportation Research Board*.

Miwa, T., Okada, Y., Morikawa, T. (2010) Applying a structured dispersion parameter to multiclass stochastic user equilibrium assignment model. *Transportation Research Record: Journal of the Transportation Research Board* 2196, 142-149.

Morikawa, T., Miwa, T. (2006) Preliminary analysis on dynamic route choice behavior: Using probe‐vehicle data. *Journal of advanced transportation* 40, 140-163.

Paag, H., Daly, A., Rohr, C. (2000) Predicting use of the copenhagen harbour tunnel. *Proceedings of INTERNATIONAL CONFERENCE ON TRAVEL BEHAVIOUR RESEARCH, 9TH, 2000, GOLD COAST, QUEENSLAND, AUSTRALIA, VOL 12*.

Park, D., Rilett, L.R. (1997) Identifying multiple and reasonable paths in transportation networks: A heuristic approach. *Transportation Research Record: Journal of the Transportation Research Board* 1607, 31-37.

Prato, C.G., Bekhor, S. (2006) Applying branch-and-bound technique to route choice set generation. *Transportation Research Record: Journal of the Transportation Research Board* 1985, 19-28.

Prato, C.G., Bekhor, S. (2007) Modeling route choice behavior: How relevant is the composition of choice set? *Transportation Research Record: Journal of the Transportation Research Board* 2003, 64-73.

Ramming, M.S. (2001) Network knowledge and route choice. Massachusetts Institute of Technology.

Revelt, D., Train, K. (1998) Mixed logit with repeated choices: Households' choices of appliance efficiency level. *Review of Economics and Statistics* 80, 647-657.

Rieser-Schüssler, N., Balmer, M., Axhausen, K.W. (2012) Route choice sets for very high-resolution data. *Transportmetrica*, 1-21.

Train, K.E. (2003) *Discrete choice methods with simulation*. Cambridge university press.

Chapter 7

Conclusions and Future Directions

This thesis deals with some issues in the context of route choice modeling. More particularly, these issues are concluded as three kinds of heterogeneity: heterogeneity in perceptions, heterogeneity in processes and heterogeneity in tastes. GPS data are used to illustrate the estimation of proposed models. Some behavioral findings can be obtained from the estimation results. This chapter provides a concluding overview of this thesis. Section 7.1 discusses the contributions of this effort. Section 7.2 describes the future research directions.

7.1Contributions

- A Bayesian network based method to model the dynamic perceptions of travel time;

Travelers' travel time perceptions are important inputs of route choice models. However, there is no previous study to model drivers' dynamic perceptions of travel time considering the en-route experiences, according to our limited knowledge. In Chapter 3, a Bayesian networks (BN) based approach is proposed to model travelers' travel time perceptions for route choice analysis. In this method, drivers' dynamic travel time perception process is described by the inference problem of BN. Their heterogeneity in perceptions can be considered by inputting different evidences to the BN for inferences. The evidences can consider drivers' en-route experiences explicitly.

A BN about a small part of the road network of Beijing is estimated using the probe data. Using the estimated BN and the mean travel time estimated under different traffic states on each road link, a dynamic route choice process is described with a simple route choice rule.

- A process based model to consider the en-route choices;

In Chapter 4, we propose a process-based method for analyzing dynamic route choice behavior. In this model, for each trip, the observed chosen route is actually the result of the driver's dynamic choice process. The dynamic choice process is defined as the sequence of choices during a trip, including the route choices (both pre-trip and en-route choices) and the choices of making a route choice again at every decision node. As compared with previous static models, two additional problems are considered in this process based route-choice model: the en-route updating of driver knowledge and the en-route choice to make a new route decision at a decision node.

The en-route updating is based on the BN model proposed in Chapter 3. The en-route choice problem is solved by a utility based model. It is assumed that, there is a binary choice at each decision node about whether make route choice again. Since drivers are assumed to prefer to use current route, a term indicate the cost of making decision is added in the utility function.

Using the proposed dynamic route model, a case study over a small network is carried out. The model is estimated and compared with conventional models using probe vehicle data. The results confirm that drivers do not tend to make route choice decisions at all decision nodes. The probability of making an en-route choice is related to a driver's sensitivity to benefit and the cost of making the decision. The absolute value of the decision cost is positively correlated with distance to the origin and negative correlated with the spatial scale of the intersection at the decision node.

- Application of several methods to explore the effect of observed attributes on heterogeneity in tastes;

Chapter 5 is an application of several methods to explore the taste heterogeneity which is related to some observed attributes. Particularly, we explore the effect of familiarity on route choice behavior. Familiarity considered here is both individual and OD pair specific, different from previous researches which only consider the individual specific familiarity.

Three methods are applied in this chapter. At first, we divide the observations to several classes, and estimate the parameters for each class. We are the first in the literature to use the two-stage variant of Chow test to explore the significance of taste heterogeneity, in the field of travel behavior modeling. Using this test, the effect of familiarity to O-D pairs on route choice behavior is proved to be statistically significant using the data collected in Toyota by private cars.

Then 2 specifications of choice models are proposed: a model with structured scale parameter and a model with structured parameters of explanation variables. These two models are estimated and apply to predict the route shares under a specific setting of choice situation. According to the estimation and prediction results, the proposed models can capture the effect of familiarity. The model with structured parameters of explanation variables has a better performance.

Behavioral finding from the results, drivers will be more easy affected by some unobserved factors (to the analyst), such as en-route information and perception, when travel between more familiar OD pairs. The estimated parameters also imply that drivers will be less sensitive to the count of intersections than free travel time when travel between more familiar OD pairs. From an application perspective, this study can be used in agent based travel demand prediction system and advanced navigation system.

- A multi-level mixed logit model to consider both observed and unobserved heterogeneity;

Chapter 6 is an extension of Chapter 5. In chapter 5, taste heterogeneity is incorporated in route choice analysis by introducing observed individual socio-economic characteristics. However, it is very likely that some taste heterogeneity will remain even after accounting for any observed characteristics. Chapter 6 proposed a multi-level mixed logit model to incorporate both observed and unobserved characteristics.

In the proposed model, the taste coefficients are treated as random and structured as observed characteristics. Further, to deal with the panel data problem, random taste is divided to three components: traveler specific, O-D pair specific and choice situation specific.

Using the GPS data collected in Toyota, Models with various assumptions about heterogeneity are estimated and compared. Empirical analysis suggests that, to enhance the performance of route choice models, it is more efficient to add more observed characteristics relating to travelers and O-D pairs than to increase the complexity of the specification. It is also proved that the incorporation of O-D pair specific unobserved taste heterogeneity can enhance the performance of a route choice model significantly.

From the estimation results, we also get several behavioral findings: age affects the taste for free travel time significantly, with older people being more sensitive to free travel time. However age does not have a significant effect on taste for the number of intersections. Gender significantly affects taste for both free travel time and number of intersections. Male drivers are more sensitive to the number of intersections, while they are less sensitive to free travel time. Car displacement has a significant effect on taste for free travel time, while its effect on the number of intersections is not significant. As distance between O-D pair increases, drivers become less sensitive to the observed attributes. When driving between more familiar O-D pairs, drivers are more sensitive to free travel time but less sensitive to the number of intersections. It is also found that these findings are stable for different route choice models, choice set generation methods and datasets.

7.2Future Research

Modeling route choice is an important part of modeling travel demand. In the macroscopic models, such as the classical 4-steps model (Karlaftis, 2001), we only need some simple route choice models. However, in some more advanced travel demand modeling system, such as the activity based models (Ben-Akiva and Bowman, 1998), we must consider the heterogeneity because these models are agent-based. On the other hand, in this “big data” age, benefited from the advanced information techniques, it also becomes affordable to explore the heterogeneity in route choice.

In this thesis, we consider 3 kinds of heterogeneity: the heterogeneity in perceptions, the heterogeneity in processes and the heterogeneity in tastes. The first two kinds of heterogeneity are few discussed in the previous studies. Therefore, our studies are only a beginning and need more development in the future studies. And the first issue is to make the proposed methods to be feasible for applying on a large network. About the heterogeneity in tastes, although there are abundant studies about it in the field of choice modeling, it is still difficult to apply the most advance methods in the context of route choice, because of its complicate nature. In the future studies, we can try to combine the latent class model and mixed logit model to reduce the complexity of the proposed multlevel mixed logit model. We can also try to apply the proposed models in some simulation tools.

Beside the three kinds of heterogeneity in this thesis, there are still many other kinds of heterogeneity should be considered in the route choice modeling: the heterogeneity of consideration sets(Yamamoto et al., 2012), the heterogeneity of attributes attendance (Collins et al., 2013) and the heterogeneity of behavioral protocols (Tian et al., 2012).

References

Ben-Akiva, M., Bowman, J.L. (1998) Activity based travel demand model systems. *Equilibrium and advanced transportation modeling*, 27-46.

Collins, A.T., Rose, J.M., Hensher, D.A. (2013) Specification issues in a generalised random parameters attribute nonattendance model. *Transportation Research Part B: Methodological* 56, 234-253.

Karlaftis, M.G. (2001) Handbook of transport modelling. *Journal of Transportation Engineering* 127, 546-546.

Tian, H., Gao, S., Fisher, D.L., Post, B. (2012) A mixed-logit latent-class model of strategic route choice behavior with real-time information. *Proceedings of Transportation Research Board 91st Annual Meeting*.

Yamamoto, T., Takamura, S., Morikawa, T. (2012) Analysis of mode and walk-route choice in a downtown area considering heterogeneity in trip distance. *Proceedings of The 91st Annual meeting of Transportation Research Board*, Washington DC, USA.

List of Publications

- *Li, D., Miwa, T. and Morikawa, T. (2012) Modeling travelers' perception of travel time for dynamic route choice behavior analysis (scientific paper), Proceedings of the 19th ITS World Congress.*
- *Li, D., Miwa, T., Morikawa, T. (2013) Dynamic Route Choice Behavior Analysis Considering En Route Learning and Choices. Transportation Research Record: Journal of the Transportation Research Board XXXX, XX-XX.*
- *Li, D., Miwa, T., Morikawa, T. (2013) Use of private probe data in route choice analysis to explore heterogeneity in drivers' familiarity with origin-destination pairs. Transportation Research Record: Journal of the Transportation Research Board 2338, 20-28.*
- *(Accepted for proceeding, under re-review for publication) Li, D., Miwa, T., Morikawa, T. Analysis of Car Usage Time Frontiers incorporating both Inter- and Intra-Individual Variation using GPS Data. Transportation Research Record: Journal of the Transportation Research Board.*
- *(Under Review) Li, D., Miwa, T., Morikawa, T. Incorporating observed and unobserved heterogeneity in route choice analysis. Transportation Research Part B.*
- *(Under Review) Li, D., Miwa, T., Morikawa, T. Considering En-Route Choices in Utility-Based Route Choice Modelling. Networks and Spatial Economics.*
- *(Revising for the re-review) Li, D., Miwa, T., Morikawa, T.* Modeling Time-of-Day Car Use Behavior: A Bayesian Network Approach. *Transportation*.