Route Choice Models with Subpath Components

Michel Bierlaire, EPFL
Emma Frejinger, EPFL

Conference paper STRC 2005
Session Transport Modelling II

STRC | **5th Swiss Transport Research Conference**
Monte Verità / Ascona, March 9-11, 2005

Route Choice Models with Subpath Components

Michel Bierlaire
Institute of Mathematics, EPFL
Lausanne

Phone: +41 21 693 25 37
Fax: +41 21 693 55 70
Email: michel.bierlaire@epfl.ch

Emma Frejinger
Institute of Mathematics, EPFL
Lausanne

Phone: +41 21 693 81 00
Fax: +41 21 693 55 70
Email: emma.frejinger@epfl.ch

February 2005

Abstract

The problem of route choice is critical in many contexts, for example in intelligent transport systems, GPS navigation and transportation planning. In order to capture the complexity of the decision process, disaggregate models, such as discrete choice models are required. In the Multinomial Logit (MNL) model, the alternatives are assumed to be independent. This assumption is not valid in a route choice context due to overlapping paths. Several adaptations of the MNL model have therefore been proposed in the literature, thereof the Path Size Logit model.

In this paper we show that, except the original formulation, all Path Size formulations presented in the literature show counter intuitive results regarding the correction of the independence assumption. Furthermore, the generalized Path Size formulation fails its original purpose of penalizing longer paths in favor of shorter ones. There is however an interesting behavioural interpretation of the Path Size attribute. Namely, overlapping paths are attractive since travellers have the possibility of switching between routes. A Path Size attribute (original formulation) could therefore be included in the deterministic part of the utility with a behavioural interpretation, but is not sufficient for correcting the independence assumption.

Considering subpaths instead of links in route choice modelling has two main advantages. First, it is behaviourally more realistic and second, it reduces the complexity of the models. In this paper we have proposed different definitions of subpaths, each with its own specific purpose. Moreover, we have presented a factor analytic specification of the Logit Kernel (LK) model including subpath components. Model estimations show very promising results of the LK model combined with a Path Size attribute. The increase in model fit is remarkable, and the covariance parameter estimates suggest that this formulation captures an important correlation structure. Furthermore, using subpath components compared to links in the correlation structure considerably decreases the complexity of the model, while its capacity of capturing the correlation structure seems as promising the formulation including links.

Keywords

Route Choice Modeling – Subpath Components – Path Size Logit – Logit Kernel – Probabilistic Choice Set Generation – GPS data – Swiss Transport Research Conference – STRC 2005 – Monte Verita

1 Introduction

The problem of route choice is critical in many contexts, for example in intelligent transport systems, GPS navigation and transportation planning. The efficiency of shortest path algorithms has been a strong motivation of many researchers to assume that travelers use the shortest (with regard to any arbitrary generalized cost) route among all. Clearly, the poor behavioural realism of the shortest path assumption motivates the use of more sophisticated models such as discrete choice models.

Designed to forecast how individuals behave in a choice context, discrete choice models (more specifically, random utility models) have motivated a tremendous amount of research in recent years. In the specific context of route choice, the definition of the choice set, and the significant correlation among alternatives are the two main difficulties.

In this paper we first present a literature review (section 2) and then analyse the Path Size Logit model (section 3), an adaptation of the Multinomial Logit model to a route choice situation. In section 4, we introduce the notion of subpath components in route choice and probabilistic choice set generation models. A Logit Kernel model with a factor analytic specification including subpaths is presented in section 4.2. Finally, preliminary model estimation results based on GPS data are presented (section 5.3).

2 Route Choice Models

Given a transport network composed of links, nodes, origin and destinations, what is the chosen route between an origin and destination for a specific transportation mode? This is the route choice problem; a discrete choice problem with specific characteristics. First, the universal choice set is usually very large. Second, the decision-maker does not consider all physically feasible alternatives. Third, some alternatives are usually highly correlated, due to overlapping paths. The choice set generation model is thus very important in order to ensure that only alternatives that an individual would actually consider are included in the choice set. The correlation structure must be captured within the route choice model. A literature review on choice set generation and route choice models will be given in the following two sections.

2.1 Choice Set Generation

Identifying the choice set in a route choice context is a difficult task. The choice set generation can be deterministic or stochastic, depending on the analyst's knowledge of the problem. In the context of deterministic choice set generation, two main approaches can be considered. First, it may be assumed that each individual can potentially choose any path between her/his origin and destination. The choice set is then easy to identify, but the number of alternatives can be very large, causing operational problems in estimating and applying the model. Moreover, this assumption is behaviourally unrealistic. Second, a restricted number of paths may be considered. Dial (1971) proposed to include in the choice set "reasonable" paths composed of links that would not move the traveler further away from her/his destination. The labeling approach (proposed by Ben-Akiva et al., 1984) includes paths meeting specific criteria, such as shortest paths, fastest paths, most scenic paths, etc. Azevedo et al. (1993) propose the link elimination approach, where the shortest path (according to a given impedance) is first calculated and introduced in the choice set. Then, some links belonging to the shortest path are removed, and a the shortest path in the modified network is computed and introduced in the choice set. Instead of eliminating links from the shortest path, the impedances on the links belonging to the shortest path can be increased. This link penalty approach was first proposed by de la Barra et al. (1993) and has the advantage of allowing further use of essential links, while discouraging the use of already identified links. Park and Rilett (1997) and Scott et al. (1997) have further developed this method by proposing different approaches for increasing the link impedances. Cascetta and Papola (2001) propose an implicit probabilistic choice set generation model, where the availability of an alternative is modeled as a Binomial Logit model.

Ramming (2001) used a deterministic simulation method that produces alternative paths by drawing link impedances from different probability distributions. The shortest path according to the randomly distributed impedance is calculated and introduced in the choice set. This approach is adopted here, and is further detailed in section 5.1.

2.1.1 Probabilistic Choice Set Generation Models

Manski (1977) proposed a probabilistic choice set model representing two stages of choice behaviour

$$P_n(i|\mathcal{C}) = \sum_{C \subseteq \mathcal{C}} P_n(i|C) Q_n(C|\mathcal{C}), \tag{1}$$

where $P_n(i|\mathcal{C})$ is the probability that individual n chooses alternative i, $\mathcal{C}$ is the set of all non-empty subsets of the universal choice set $\mathcal{U}$, and the sum is over all subsets C of $\mathcal{C}$. The first stage is thus the choice set formation process modeled by $Q_n(C|\mathcal{C})$; the probability that C is individual n's choice set. The second stage is the choice behaviour given the choice set C ($P_n(i|C)$) which can be modeled by a discrete choice model.

Two issues must be addressed here. First, the number of subsets C involved in the model must be reduced to obtain a tractable formulation. Second, the probability law Q_n of each individual must be defined. The use of deterministic rules to generate the choice set, that is $Q_n(C|\mathcal{C}) = 0$ or 1, has been shown to be unsatisfactory in the context of route choice (Han, 2001 and Ramming, 2001). Indeed, the choice set is not only formed by observable restrictions, but also by psychological restrictions. When there is a large number of alternatives in the universal choice set, as in a route choice context, it is unrealistic to assume that an individual examines and compares numerous alternatives in order to choose the best one. A pure probabilistic approach should therefore be preferred.

A random constraints model was proposed by Swait and Ben-Akiva (1987) that build on a non-compensatory approach. That is, when one constraint is not satisfied for one alternative, then the alternative cannot be included in the choice set. For instance, alternatives that are too expensive, or too long, will be more likely to be rejected. The probability that alternative i is included in the choice set of individual n is

$$q_n(i) = \prod_{k=1}^{K} q_{kn}(i), \tag{2}$$

where $q_{kn}(i)$ is the probability that alternative i satisfies the k-th constraint of individual n, and there is a total of K constraints.

A problem with equation (1) is the combinatorially large number of potential choice sets; if M is the number of paths in $\mathcal{U}$, then there are $2^J - 1$ non empty subsets of $\mathcal{U}$. This problem is addressed by Morikawa (1996) who proposes to reduce the complexity of the model by using pairwise comparison of alternatives. There are two possibilities for alternative i being preferred to alternative j. Either both alternatives are included in the choice set and alternative i has a greater utility than j. Or, alternative i is included in the choice set but alternative j is eliminated at the first stage of the choice set formation. Morikawa (1996) obtained the following model

$$P_n(i|\mathcal{C}) = \frac{1}{1 - Q_n(\varnothing)} q_n(i) P \left(\bigcap_{j \in M, j \neq i} \{(j \in C_n) \cap (U_{in} \geq U_{jn})\} \cup \{j \notin C_n\} \right) \tag{3}$$

where C_n is the (latent) true choice set of individual n and $Q_n(\varnothing)$ is the probability that the random constraint model yields the empty choice set. Note that this model does not involve a sum with an exponential number of terms, as in equation (1).

Models based on latent choice sets (Ben-Akiva and Boccara, 1995, Gopinath, 1995) have also been proposed in the literature, but very few instances of such approaches have been published in the context of route choice models. Ramming (2001) proposed however the latent variable model of "network knowledge", captured by the so-called Multiple Indicator-Multiple Cause specification in the context of route choice. It is a direct application of the concept of "spatial knowledge" proposed by Ben-Akiva et al. (1999).

2.2 Behavioural Rules

The simplest route choice model is deterministic, and assumes that an individual chooses the shortest path between an origin-destination pair. This is however a behaviourally unrealistic assumption. This is the motivation for using random utility models. Several different models have been proposed in the literature. The Multinomial Logit (MNL) model, is simple but restricted by the Independence from Irrelevant Alternatives (IIA) property, which does not hold in the context of route choice due to overlapping paths in the choice set. Efforts have been made to overcome this restriction by making a deterministic correction of the utility for overlapping paths. Two different corrections have been proposed in the literature: Commonality Factor (Cascetta et al., 1996) and Path Size (Ben-Akiva and Bierlaire, 1999). The Path Size Logit (PSL) model will be further discussed in section 3.

Given the shortcomings of the MNL model, Probit models have been proposed in the context of stochastic network loading by Burrell (1968), Dagazo and Sheffi (1977) and Yai et al. (1997). While MNL suffers from its simplicity, Probit models suffers from their complexity. Indeed, there is no analytical formulation for the probabilities and the variance-covariance matrix is complex. The Multinomial Probit with Logit Kernel (LK), introduced by Bolduc and Ben-Akiva (1991), was designed to combine the advantages of both Logit and Probit models. This type of model are also referred as Hybrid Logit or Mixed Logit. Ramming (2001) (also discussed in Bekhor et al., 2002) estimated a route choice model based on the Logit Kernel model of Ben-Akiva and Bolduc (1996) that combines the Logit and Probit models by adding normal error components to a core MNL model to account for correlation. This LK model will be further detailed in section 2.2.1 and adapted to include subpath components in section 4.2.

Vovsha and Bekhor (1998) have proposed the Link-Nested Logit model, which is a Cross-Nested Logit (CNL) formulation (see Bierlaire, forthcoming, for an analysis of the CNL model) where each link of the network corresponds to a nest, and each path to an alternative.

The Paired Combinatorial Logit (PCL) model was proposed by Chu (1989) and further developed by Koppelman and Wen (2000). The model was adapted to the route choice problem by Prashker and Bekhor (1998) and Gliebe et al. (1999) who proposed two different relationships between the similarity parameter and network topology.

Han (2001) (see also Han et al., 2001) used a Mixed Logit model to investigate taste heterogeneity across drivers and the possible correlation between repeated choices. Stated preferences (SP) data was used for the model estimation which limited the number of alternatives in the choice set. He concludes, as Ramming (2001), that the number of draws for the simulated maximum likelihood estimation must be carefully chosen.

Paag et al. (2002) and Nielsen et al. (2002) used a Mixed Logit model with both a random coefficient and error component structure to estimate route choice models for the harbor tunnel project in Copenhagen.

Recently, Marzano and Papola (2004) have proposed a Link-Based Path-Multilevel Logit model, which model path choice as a sequence of choices. Mimicking the structure of a CNL model, where the nests are not designed to capture correlation, but rather to capture the sequence of choices, their model avoids the need for explicit path enumeration.

2.2.1 Logit Kernel Model

A Logit Kernel (LK) model is a combination of a Probit and Logit model and was first introduced by Bolduc and Ben-Akiva (1991). The utility function for individual n and alternative i is

$$U_{in} = V_{in} + \xi_{in} + \nu_{in}$$

where ξ_{in} are normally distributed and capture correlation between alternatives, and ν_{in} are independent and identically distributed Gumbel.

The LK model can be combined with a factor analytic specification, meaning that some structure is explicitly specified in the model, and its complexity is therefore decreased. Ramming (2001) (see also Bekhor et al., 2002) adapted the LK model to a route choice situation with a factor analytic specification explicitly capturing the interdependencies among alternatives. The utility vector U_n (Mx1, where M is the number of paths) is then defined by

$$U_n = V_n + \varepsilon_n = V_n + F_n T \zeta_n + \nu_n, \tag{4}$$

where

- V_n (Mx1) is the vector of deterministic utilities,
- F_n (MxJ) is the matrix factor loading matrix,
- T (JxJ) is a diagonal matrix, and
- ζ_n (Jx1) is the vector of i.i.d. normal variables with zero mean and unit variance.

The following assumptions are specified in Bekhor et al. (2002):

- Link-specific factors are i.i.d. normal,

- the F matrix is the link-path incidence matrix,
- variance is proportional to link length, and
- the T matrix is the link factors variance matrix (diagonal):

$$T = \sigma \text{ diag} \left(\sqrt{l_1}, \sqrt{l_2}, \dots, \sqrt{l_J} \right).$$

where l_a is the length of link a, and σ is the only parameter to estimate.

The covariance matrix can then defined as follows:

$$F_n T T^T F_n^T = \sigma^2 \begin{bmatrix} L_1 & L_{1,2} & \dots & L_{1,M} \\ L_{1,2} & L_2 & \dots & L_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ L_{1,M} & L_{2,M} & \dots & L_M \end{bmatrix}$$

where $L_{i,j}$ is length by which path i overlaps with path j.

In section 4.2 we build on this formulation when we include subpaths in the LK model.

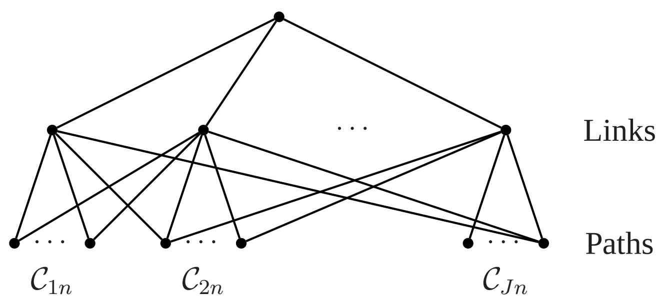
3 Path Size Logit

In this section we analyse the PSL model (MNL with Path Size attribute), and discuss its capacity of correcting the IIA assumption of the MNL model. First, we derive the original formulation of the Path Size attribute and second, analyse the other existing formulations.

3.1 Original Path Size Formulation

The original Path Size formulation was proposed by Ben-Akiva and Bierlaire (1999) (first version, later the formulation was changed for the one presented in section 3.2), and is an application of discrete choice theory for aggregate alternatives (see chapter 9 in Ben-Akiva and Lerman, 1985). In this section, we present the theory for aggregate alternatives in the context of route choice, and derive the original Path Size formulation.

Figure 1: Example Cross-Nested Structure of Aggregate and Elemental Alternatives



Ben-Akiva and Lerman (1985) assumed a nested structure where each nest corresponds to an aggregate alternative that groups a certain number of elemental alternatives together. The elemental alternatives correspond to actual alternatives that individuals are choosing. In a route choice context the elemental alternatives correspond to the paths and the aggregate alternatives to the links. For the derivation of the Path Size formulation we are only interested in the choice of elemental alternative (route choice) and the size of the aggregate alternatives. Therefore, we assume a more general cross-nested structure (example shown in figure 1) where each path can belong to more than one nest. It is the same structure as the one proposed by Vovsha and Bekhor (1998) in their Link-Nested Logit model.

We denote by $\mathcal{C}_n$ the set of paths considered by individual n, which is divided into subsets,

$$\mathcal{C}_{an} \subseteq \mathcal{C}_n, \quad a = 1, \ldots, J,$$

where J is the number of links.

The utility U_{in} associated with path i is

$$U_{in} = V_{in} + \varepsilon_{in}, \quad i \in \mathcal{C}_n,$$

where V_{in} represents the deterministic part of the utility and ε_{in} the random part. Since we assume utility maximization and that an individual chooses one path among all, the utility U_{an} of a link a is defined by

$$U_{an} = \max_{i \in \mathcal{C}_{an}} (V_{in} + \varepsilon_{in}), \quad a = 1, \ldots, J.$$

U_{an} can also be expressed as the sum of its expectation V_{an} and its random term ε_{an}, that is,

$$U_{an} = V_{an} + \varepsilon_{an}, \quad a = 1, \ldots, J,$$

where

$$V_{an} = E[\max_{i \in \mathcal{C}_{an}} (V_{in} + \varepsilon_{in})].$$

Furthermore, the average utility of the paths including link a is defined by

$$\overline{V}_{in} = \frac{1}{M_a} \sum_{i \in \mathcal{C}_{an}} V_{in} \quad a = 1, \ldots, J$$

where M_a is the number of paths including link a. That is, $M_a = \sum_{i \in \mathcal{C}_n} \delta_{ai}$, where δ_{ai} is the link-path incidence variable that is one if link a is on path i and zero otherwise.

If a large number of paths includes a link, and if the random terms of the path utilities ε_{in} are IID Gumbel, then the distribution of the utility of link a is also Gumbel with the same positive scale parameter μ and a location parameter

$$\eta = \frac{1}{\mu} \ln \sum_{i \in \mathcal{C}_{an}} e^{\mu V_{in}} = \overline{V}_{an} + \frac{1}{\mu} \ln \left[\frac{1}{M_a} \sum_{i \in \mathcal{C}_{an}} e^{\mu (V_{in} - \overline{V}_{an})} \right] + \frac{1}{\mu} \ln M_a.$$

The utility for a link a can thus be modeled by

$$U_{an} = \overline{V}_{an} + \frac{1}{\mu} \ln M_a + \frac{1}{\mu} \ln B_{an} + \varepsilon_{an}, \quad a = 1, \ldots, J,$$

where

$$B_{an} = \frac{1}{M_a} \sum_{i \in \mathcal{C}_{an}} e^{\mu(V_{in} - \overline{V}_{an})}.$$

B_{an} can be interpreted as a measure of the variability of the path utilities including link a. It can be shown (see Ben-Akiva and Lerman, 1985) that $B_{an} \geq 1$. The correction for heterogeneity among paths will therefore always be non-negative, and zero if the paths are homogeneous.

It can also be shown (Ben-Akiva and Lerman, 1985) that if we assume the links to have equal variances, then $\frac{1}{\mu} \ln B_{an}$ can be omitted from the utility U_{an} which consequently can be defined by

$$U_{an} = \overline{V}_{an} + \frac{1}{\mu} \ln M_a + \varepsilon_{an}, \quad a = 1, \ldots, J.$$

The original Path Size formulation, correcting the path utility U_{in}, is based on the definition of the link utility U_{an}. Accordingly, the positive correction for the size of an aggregate alternative, results in a negative correction of the utility of an elemental alternative. Moreover, there is no correction of an elemental alternative which belongs to a nest with size one. The size correction for an elemental alternative can therefore be defined as the inverse of the size of an aggregate alternative, that is, $\mu \ln \frac{1}{M_a}$. The contribution of a link a (size of link a) is then

$$\mu \ln \frac{1}{\sum_{j \in \mathcal{C}_n} \delta_{aj}},$$

where δ_{aj} is the link-path incidence variable. Furthermore, we assume that the size of a path is proportional to the size of its links. If l_i denotes the length of link a and L_i the length of path i, the original Path Size attribute can be expressed as

$$\text{PS}_{in} = \sum_{a \in \Gamma_i} \frac{l_a}{L_i} \frac{1}{\sum_{j \in \mathcal{C}_n} \delta_{aj}}, \tag{5}$$

where Γ_i is the set of links in path i. Including a Path Size correction in the utility U_{in} an individual n associated with a path i is then

$$U_{in} = V_{in} + \beta_{\text{PS}} \ln \text{PS}_{in} + \varepsilon_{in}, \quad i \in \mathcal{C}_n,$$

where β_{PS} corresponds to the scale parameter μ and should thus always be included in the utility and be strictly positive. The probability that an individual n chooses a path i is

$$P_n(i) = \frac{e^{V_{in} + \beta_{\text{PS}} \ln \text{PS}_{in}}}{\sum_{j \in C_n} e^{V_{jn} + \beta_{\text{PS}} \ln \text{PS}_{jn}}}.$$

Note that when a path does not share any link with an other path in the choice set (we refer to these paths as *distinct* paths), the correction is $\ln 1 = 0$ and the correction for two identical paths is $-\ln 2$. Furthermore, the Path Size only depends on the number of paths in the choice set sharing the same links, independently of the length of the paths.

Several different formulations have been proposed in the literature based on the original Path Size (equation (5)). Their capacity of correcting the IIA assumption on the random terms will be analysed in the following sections.

3.2 Path Size including Shortest Path

Ben-Akiva and Bierlaire (1999) included the shortest path of the choice set, denoted $L^*_{\mathcal{C}_n}$, in their second version of the Path Size formulation:

$$\text{PS}_{in} = \sum_{a\in\Gamma_i} \frac{l_a}{L_i} \frac{1}{\displaystyle\sum_{j\in\mathcal{C}_n} \frac{L^*_{\mathcal{C}_n}}{L_j}\delta_{aj}}. \tag{6}$$

In order to analyse this formulation we can write the Path Size correction $\ln \text{PS}_{in}$ as follows:

$$\ln \text{PS}_{in} = -\ln L_i - \ln L^*_{\mathcal{C}_n} + \ln \sum_{a\in\Gamma_i} \frac{l_a}{\displaystyle\sum_{j\in\mathcal{C}_n} \frac{1}{L_j}\delta_{aj}}.$$

Note that the term $\ln L^*_{\mathcal{C}_n}$ can be omitted since it has the same value for all paths in the choice set and consequently, does not change their relative utility. Moreover, the correction of a distinct path i where $L_i > L^*_{C_n}$ is not zero. Indeed, the Path Size correction is then defined by

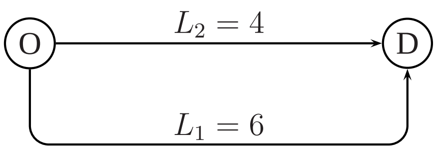
$$\ln \text{PS}_{in} = \ln\left(\frac{L_i}{L^*_{\mathcal{C}_n}}\right) > 0,$$

which is counter intuitive. An example is shown in figure 2 where $\text{PS}_1 = \frac{3}{2}$ and $\text{PS}_2 = 1$, even though both alternatives are distinct. If the deterministic part of the utility only includes the path length and Path Size (the Path Size scale parameter is set to one), it can be expressed as

$$U_i = -L_i + \ln \text{PS}_i.$$

The resulting probabilities are then $P(1|\{1,2\}) = 0.17$ and $P(2|\{1,2\}) = 0.83$, compared to the choice probabilities without Path Size correction; $P(1|\{1,2\}) = 0.12$ and $P(2|\{1,2\}) = 0.88$. We can therefore conclude that this Path Size formulation shows counter intuitive results. Indeed, there is a positive correction of all distinct paths which are longer than the shortest path, although it should be zero.

Figure 2: Counter Example - Path Size including Shortest Path



3.3 Generalized Path Size

We discuss in this section the generalized Path Size formulation proposed by Ramming (2001), and show that is should not be used. This generalized formulation is defined by

$$\text{PS}_{in} = \sum_{a\in\Gamma_i} \frac{l_a}{L_i} \frac{1}{\displaystyle\sum_{j\in C_n} \frac{G(L_i;\gamma)}{G(L_j;\gamma)}},$$

where $G(L_i;\gamma)$ is a function of L_i with parameter γ. He proposed to use $G(L_i;\gamma) = (L_i)^\gamma$,

$$\text{PS}_{in} = \sum_{a\in\Gamma_i} \frac{l_a}{L_i} \frac{1}{\displaystyle\sum_{j\in C_n} \left(\frac{L_i}{L_j}\right)^\gamma \delta_{aj}}, \tag{7}$$

with the following motivation: "This formulation [original Path Size] can therefore suffer when arbitrarily long paths are included in the choice set."(quoting Ramming (2001), page 49). He also comments the previous formulation: "In this way, arbitrarily long paths, which would likely not be considered by travelers, do not reduce the size of other, more reasonable paths that use the same link." (quoting Ramming (2001), page 93). We do not recommend this approach for two reasons. First, there is no issue of the length of paths sharing the same links since only the number of paths influences the Path Size in the original formulation. Second, as we show below, using a $\gamma > 0$ yields counter intuitive results.

Note that $\gamma = 0$ in equation (7) corresponds to the original Path Size formulation (equation (5)). This generalized formulation has been used in several route choice applications (see for example Ramming, 2001 and Hoogendoorn-Lanser, 2005) with different values of γ. In order to analyse the influence of the γ parameter, we write $\ln \text{PS}_{in}$ as follows:

$$\ln \text{PS}_{in} = -(\gamma+1)\ln L_i + \ln \sum_{a\in\Gamma_i} l_a \frac{1}{\displaystyle\sum_{j\in \mathcal{C}_n} \left(\frac{1}{L_j}\right)^\gamma \delta_{aj}}. \tag{8}$$

Independently of the value of the γ parameter, this formulation yields a zero correction when path i is distinct from other paths. Furthermore, it yields the same correction as the original Path Size formulation ($-\ln 2$) when two paths are identical. However, it is theoretically difficult to give an interpretation as well as motivation of the γ parameter, especially when $\gamma \to +\infty$. Indeed, if we assume that $L_i \geq 1 \quad \forall\, i \in \mathcal{C}_n$, then the limits of the two terms in equation (8) are

$$\lim_{\gamma\to+\infty} -(\gamma+1)\ln L_i = -\infty \qquad \lim_{\gamma\to+\infty} \ln \sum_{a\in\Gamma_i} l_a \frac{1}{\displaystyle\sum_{j\in C_n} \left(\frac{1}{L_j}\right)^\gamma \delta_{aj}} = +\infty.$$

Ramming (2001) argues that low values of γ could yield counter intuitive results and using $\gamma = +\infty$ shows best results in terms of model fit. He remains however sceptical to use $\gamma = +\infty$ and suggests to use a large finite value of γ. On the contrary, we argue that $\gamma > 0$ show counter intuitive results. We illustrate this statement with two examples (figure 3 and 4). First, we consider the same example as in Ramming (2001) (also used in Hoogendoorn-Lanser, 2005), shown in figure 3.

The Path Size values for different γ are shown in table 1. The example clearly shows that $\gamma = +\infty$ only corrects the utility of long alternatives and no correction at all of short alternatives. Moreover, the correction of path 3 is then the same, as the correction that the original Path Size formulation gives for two identical paths ($-\ln 2$).

In order to illustrate the counter intuitive results of using the generalized Path Size formulation with $\gamma > 0$ we provide an other example (figure 4). Indeed, from equation (7) we can see that the influence of the γ parameter is highly dependent on the value of $\frac{L_i}{L_j}$ being greater or equal to one, or less than one. In the following example we consider three correlated alternatives instead

Figure 3: Ramming's Path Size Example

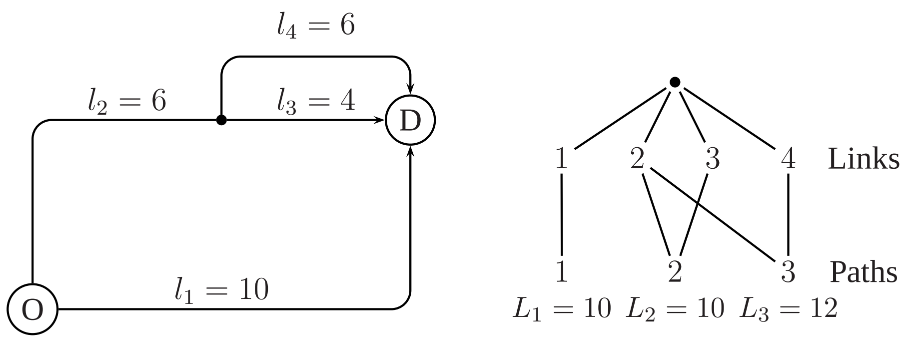


Table 1: Ramming's Path Size Example

$\gamma:$	0	1	2	4	14	∞
PS_1	1.00	1.00	1.00	1.00	1.00	1.00
PS_2	0.70	0.73	0.75	0.81	0.96	1.00
PS_3	0.75	0.73	0.71	0.66	0.54	0.50

of two. The Path Size values for different γ can be found in table 2. Note that $\frac{L_3}{L_2} > 1$ but $\frac{L_3}{L_4} < 1$ which explains the counter intuitive results for values of $\gamma > 0$. Indeed, path 3 is more penalized than path 4, even though $L_3 < L_4$. (Recall that the Path Size correction is $\ln \mathrm{PS}_i$, meaning that the lower the Path Size value, the more severe the correction.)

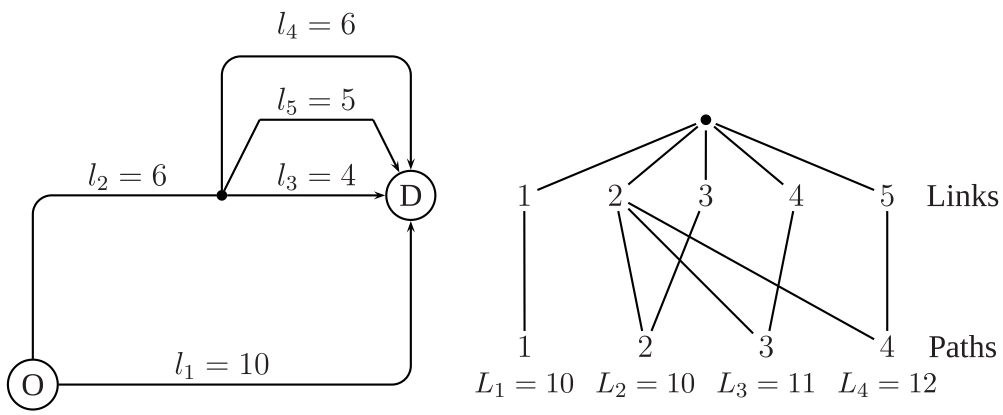
The observed improvement in model fit (Ramming, 2001, Hoogendoorn-Lanser, 2005), by increasing the value of γ suggests that the Path Size attribute has an behavioural interpretation (this will be discussed further in section 5.3). Nevertheless, based on the discussion above it is difficult to interpret the generalized Path Size when $\gamma > 0$.

We conclude that not only does the generalized Path Size formulation show counter intuitive results for $\gamma > 0$ regarding correction of the IIA assumption on the random terms. But also, does not serve its original purpose namely, penalising long paths in favor of shorter ones.

Table 2: Counter Example Generalized Path Size

$\gamma:$	0	1	2	4	14	∞
PS_1	1.00	1.00	1.00	1.00	1.00	1.00
PS_2	0.60	0.62	0.64	0.68	0.85	1.00
PS_3	0.63	0.59	0.56	0.53	0.49	0.46
PS_4	0.67	0.65	0.64	0.61	0.53	0.50

Figure 4: Counter Example Generalized Path Size



4 Subpath Components

It is reasonable to believe that travellers do not evaluate their route choice in every network junction (for each link), but rather for a sequence of links, that is for subpaths. The choice process then consists in choosing a certain number of subpaths. We propose three definitions of subpaths:

- **Behavioural subpath** is used to refer to the concept used by individuals when describing an itinerary. It may be a specific location in an environment (commercial center, church, river, etc.), an actual node in the transportation network (roundabout, crossroad, parking, etc.) or an actual subpath within the transportation network (highway I95, Beverly Hills, beltway, etc.). A distinction can be made between individual specific locations such as home, school or work from more generic ones.
- **Topological subpath** is a list of consecutive links designed to simplify the path representation. It is a common concept in the design of data structures for path enumeration.
- **Strategic subpath** refer to the strategic choice of an individual. For example, taking the highway or passing by the city center.

Each one of these definitions has its own specific purpose. The *behavioural subpath* is a central concept in the data collection process. It can be used to design stated preferences survey, and to extract relevant information from actual path descriptions. The *topological subpath* is necessary to characterize the choice set. Finally, the *strategic subpath* is the central paradigm used for the model definition, namely the assumptions about the error structure.

A two-step modelling framework can be defined where the final choice of an itinerary is a combination of *strategic* choices (e.g. take the highway) and *implementation* decisions (e.g. how to get to the highway). This idea has been originally suggested by Ben-Akiva (private communication, 1997). In this context, the elements of the choice set for the strategic choices are the *strategic subpaths*. It is legitimate to believe that the error structure for the strategic decisions is significantly different from the error structure of implementation decisions. Moreover, the issue of structural correlation may be better captured in this framework, where Generalized Extreme Value (GEV) models such as Cross-Nested Logit (CNL) or Network GEV models (Bierlaire, 2002, Daly and Bierlaire, 2003) will be naturally appropriate.

4.1 Subpaths in Probabilistic Choice Set Models

An extension of the ideas presented on probabilistic choice set generation in section 2.1.1 (page 4) can be made by integrating the concept of subpaths. If S_n is the (deterministic) set of strategic subpaths considered by individual n, we have

$$P_n(i|S_n) = \sum_{C_n \subseteq \mathcal{C}_n} \sum_{\mathcal{C}_n \supseteq s} \sum_{s \subseteq S_n} P(i|C_n)P(C_n|\mathcal{C}_n)P(\mathcal{C}_n|s)P(s|S_n) \tag{9}$$

where

- s are the possible strategic subpaths,
- $\mathcal{C}_n$ is the potential choice set for individual n, and
- C_n is the actual choice set for individual n.

The selection of subpaths $P(s|S_n)$ considered in this framework is the main strategic decisions performed by decision-makers. Therefore, it is critical for the validity of this approach that the definition of *strategic* subpaths used in the model is consistent with the *behavioural* subpaths used by decision-makers.

The rest of the model is similar to the approach by Morikawa (1996), where the universal choice set $\mathcal{C}_n$ is now conditional on the selected subpaths and, therefore, significantly smaller than in the classical route choice context. Namely, the choice conditional on $\mathcal{C}_n$, that is

$$P_n(i|\mathcal{C}_n) = \sum_{C_n \subseteq \mathcal{C}_n} P(i|C_n)P(C_n|\mathcal{C}_n),$$

can be modeled by equation (3).

4.2 Logit Kernel Model with Subpaths

For the inclusion of subpaths in the LK model, we build on the work by Ramming (2001) and Bekhor et al. (2002) presented in section 2.2.1. Instead of including links in the factor analytic specification, we include subpaths. Since the number of subpaths in the choice set is considerably smaller than the number of links, the complexity of the model is reduced.

The assumptions presented in section 2.2.1 remain the same but the F_n matrix in equation (4) is now the subpath-path incidence matrix (instead of link-path incidence matrix), and the T matrix is defined as follows:

$$T = \sigma \text{ diag}\left(\sqrt{l_{S_1}}, \sqrt{l_{S_2}}, \ldots, \sqrt{l_{S_K}}\right).$$

where l_{S_q} is the length on subpath q (K is the number of subpaths in the choice set).

Results from LK estimations based on the specification above are presented in section 5.3.

5 Preliminary results

The estimation results presented in this section are based on a GPS data set collected during a traffic safety study in the Swedish city of Borlänge. About 200 vehicles were equipped with a GPS device and the vehicles were monitored within a radius of about 25 km around the city center. Since the data set was not originally collected for route choice analysis an extensive amount of data processing has been performed in order to clean the data and obtain coherent routes. The data processing for obtaining data for route choice analysis was mainly performed by J. Wolf and M. Oliveira at GeoStats, Atlanta. Data of 24 vehicles, 16035 observations, are now available for route choice analysis. (See Axhausen et al., 2003, Schönfelder and Samaga, 2003 and Schönfelder et al., 2002 for more details on the Borlänge GPS data set.)

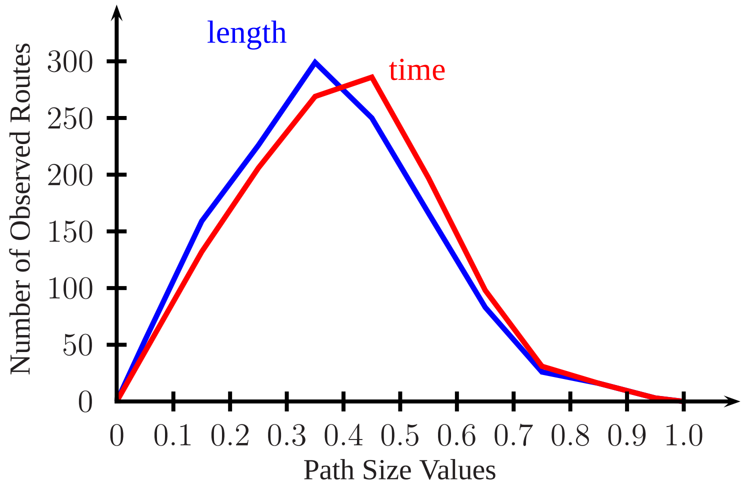
Borlänge is situated in the middle of Sweden and has about 47000 inhabitants. The road network contains 3077 nodes and 3843 links. Here, we consider a total of 1461 observations (1282 observed routes) of one vehicle observed during 214 days. Note that this results in an average of 6.8 trips per day, to be compared with the Swedish average of 1.7 trips per day (SIKA, 2001). The average of the 24 vehicles available is 5.4 trips per day, which is less than 6.8 but considerably higher than the Swedish average. It can also be noted that for the vehicle considered here, there are 927 origin destination pairs and thus an average of 1.6 observations per pair. There are several possible explanations for the high number of trips per day and high number of origin destination pairs, all related to how the data was collected (for example, logging frequency, point filtering, GPS device on and off events), sampling bias, as well as data processing issues. A detailed discussion on the source of these problems is out of the scope of this paper. Nevertheless, the route choice behaviour is still captured by the data.

5.1 Choice Set Generation

In order to estimate route choice models, the choice set for each origin destination pair needs to be defined. For these preliminary results, we have used a deterministic simulation approach, also used by Ramming (2001). This simulation method produces alternative paths by drawing link impedances from a probability distribution. The shortest path according to the randomly distributed impedance is calculated and introduced in the choice set. We used the estimated time link impedance and a truncated normal distribution with 20 draws (with mean and standard error based on the observations). As discussed earlier, deterministic choice set generation methods are not sufficient in a route choice context. We know for instance that the shortest path assumption is not behaviourally realistic. Given the known limitations of this approach, the observed routes are added to the choice set, if not already present. The resulting choice sets include an average of 9.3 routes (maximum 22 and minimum 2 routes). More details on the choice set generation approach can be found in Ramming (2001) or Bekhor et al. (2001).

Figure 5 shows the Path Size values for the observed routes based on time and length for the original Path Size formulation (that is, the generalized formulation with $\gamma = 0$). Note that the observed routes have, in general, a rather high overlap with the other routes in the choice set. The Path Size values based on length and time are comparable. We will therefore use the Path Size based on length in the model estimation since the link length is known with certainty whereas the estimated time is an approximation.

Figure 5: Path Size Values for Observed Routes (Generalized Path Size with $\gamma = 0$)



5.2 Subpaths in the Borlänge Network

In section 4 we hypothesized the error structure for strategic decisions to be significantly different from the error structure of implementation decisions. Since only data collected through passive monitoring is available here, we cannot know the strategic choices of the individual. However, we can identify some intuitive definitions of strategic subpaths in the Borlänge network. Namely, there are some main roads for traversing the city center. In figure 6 these main roads are shown. There are two parallel roads to traverse the city from north-west to south-east (or the opposite) named S_1 and S_2 in the figure. Moreover, we have identified five parallel roads to traverse the city center from the north-east to the south-west (or the opposite), these subpaths are named S_3, S_4, S_5, S_6 and S_7 respectively. All the subpaths have been identified based on a city map.

The identification of subpaths is somehow subjective. In order to analyse the importance of the identification process we have included a "test" subpath (see figure 7), named S_8, that is arbitrarily chosen such that it is likely not to have any behavioural impact.

5.3 Model Estimation

The results of the model estimations are shown in table 4. For the comparison of the parameter estimates of the different models we have provided a scaled parameter estimate. The scaling is based on $\beta_{\text{left turns}}$ and the magnitude of this parameter is the same for all the models. In addition to the Path Size, two other attributes are included in the deterministic part of the utility (equation (10)). First, the number of left turns in non controlled crossings, that is with out traffic lights (computed with the geographical information system TransCAD) are included. This attribute is expected to have negative influence on the utility since left turns are more dangerous, and take more time than right turns. Second, a non linear formulation modelling how the length parameter varies with average speed is included. A linear formulation would

Figure 6: Subpaths in Borlänge

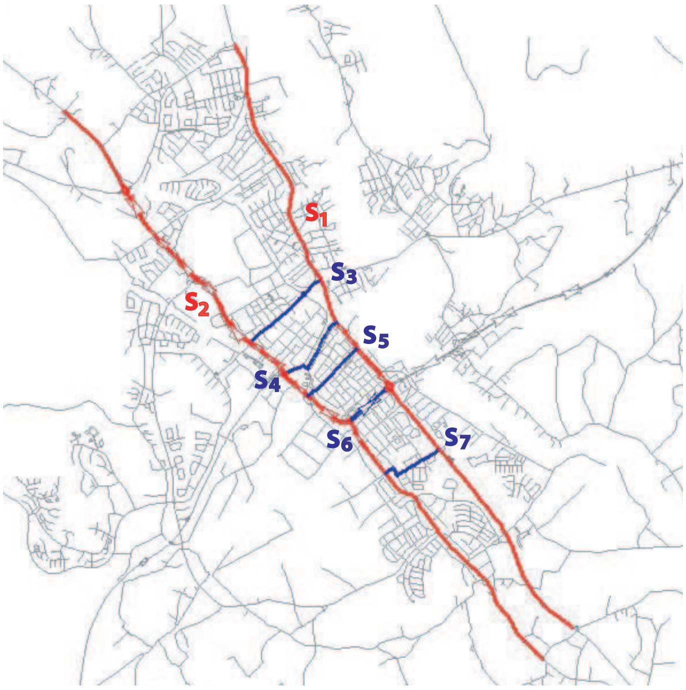
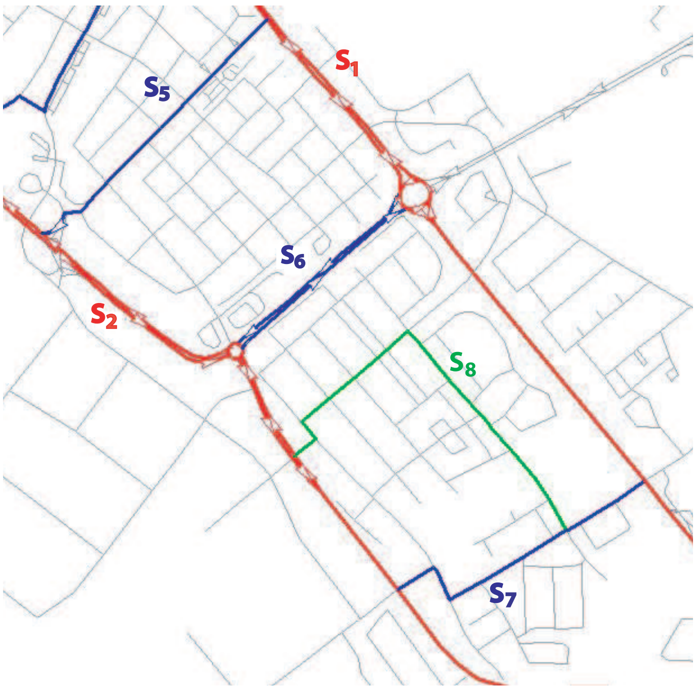


Figure 7: Test Subpath Definition



have been $\beta_{\text{length}}\text{length}_i$. Here we have

$$\beta_{\text{length}} = \hat{\beta}_{\text{length}} \left(\frac{\text{avg speed}_i}{\text{avg speed}_{\text{ref}}} \right)^{\text{elasticity}}$$

where the elasticity is the actual elasticity of β_{length} regarding average speed, that is,

$$\text{elasticity} = \frac{\text{d}\beta_{\text{length}}}{\text{davg speed}_i} \left(\frac{(\text{avg speed}_i)}{\beta_{\text{length}}} \right).$$

This attribute formulation has been included in order to model that travellers can prefer longer routes in order to keep a high average speed. Making a linear formulation including only length, would intuitively yield a negative correction of the utility while this nonlinear formulation captures the positive effect of longer routes under condition that a high average speed can be kept. We therefore expect positive sign for $\hat{\beta}_{\text{length}}$ as well as for the elasticity parameter. The deterministic part for the utility of alternative i and the one considered individual is then

$$V_i = \beta_{\text{PS}}\text{PS}_i + \beta_{\text{left turns}}\text{left turns}_i + \hat{\beta}_{\text{length}}\text{length}_i \left(\frac{\text{avg speed}_i}{\text{avg speed}_{\text{ref}}} \right)^{\text{elasticity}}, \quad (10)$$

where the reference average speed is 50 km/h. A summary of the attributes is given in table 3. The Logit Kernel models have the structure presented in section 4.2.

Table 3: Model Attributes

Attribute name	Description
Path Size	γ = 0, based on link length.
Left turns	Number of left turns in crossings without traffic lights.
Length	How length varies with average speed.
Elasticity	Elasticity of length regarding average speed

All the parameter estimates except β_{PS} have their expected signs and are highly significant in the PSL model. The signs remain unchanged for the Logit Kernel models and the general interpretation stays therefore the same. Note however that the magnitude of the parameter estimates (comparison of the scaled estimates) as well as their significance change from the PSL model to the LK models. This has also been observed by Ramming (2001) when he compared PSL with LK estimations.

The β_{PS} is negative in all the models and highly significant, meaning that overlapping routes get increased utility. This suggests that the Path Size attribute has a behavioural interpretation; overlapping routes are preferred since a traveller has the possibility of switching between routes. This interpretation is intuitive, especially in an urban areas where accidents or congestion can block a part of a route. The same discussion is held in Hoogendoorn-Lanser (2005) (and Hoogendoorn-Lanser et al., 2005) in the context of multi-modal networks. However, a negative β_{PS} also means that the Path Size attribute does not correct the assumption on the random terms. The Path Size attribute captures two effects where one has a positive effect on the utility, and the other one a negative effect. Consequently, while explicitly specifying some of the correlation structure in the model we would expect that the Path Size attribute captures less of the negative

Table 4: Estimation Results

Beta parameters		PSL	LK_1	LK_2	LK_3	LK_4
Path Size, $\gamma = 0$	Estimate	-1.90	-2.94	-2.95	-3.06	-3.06
	Scaled Estimate	*-1.90*	*-4.52*	*-4.54*	*-4.71*	*-4.71*
	(Std. Error)	(0.12)	(0.22)	(0.22)	(0.28)	(0.28)
	T-Statistic	-15.56	-13.54	-13.55	-11.02	-11.01
Left turns		-0.20	-0.13	-0.13	-0.13	-0.13
		-0.20	*-0.20*	*-0.20*	*-0.20*	*-0.20*
		(0.01)	(0.03)	(0.03)	(0.03)	(0.03)
		-15.65	-4.84	-4.85	-4.18	-4.16
Length (regarding avg speed)		0.46	0.52	0.52	0.44	0.44
		0.46	*0.80*	*0.80*	*0.68*	*0.68*
		(0.04)	(0.09)	(0.09)	(0.10)	(0.10)
		11.47	5.58	5.61	4.22	4.14
Elasticity		0.18	0.18	0.17	0.49	0.49
		(0.12)	(0.40)	(0.39)	(0.29)	(0.29)
		6.10	0.44	0.43	1.71	1.67
σ			0.36	0.36	0.41	0.40
			(0.03)	(0.03)	(0.03)	(0.03)
			10.71	10.71	11.83	11.85
Nb of simulation draws		-	1000	1000	1000	1000
Subpaths		-	S_1, S_2	S_1, S_2, S_8	$S_1 - S_7$	$S_1 - S_8$
Nb of estimated parameters		-	5	5	5	5
Final Log-Likelihood		-2796.70	-2262.50	-2263.78	-2111.83	-2113.72
Rho-square		0.18	0.34	0.33	0.38	0.38
Sample size : 1461						

Biogeme (Bierlaire, 2003b, Bierlaire, 2003a) has been used for all model estimations.

effect (thus less of the correlation among alternatives) and the magnitude of the positive effect should thus increase. Indeed, the scaled magnitude of the β_{PS} estimate is highly increased in the LK models compared to the PSL model. Moreover, when seven subpaths are included in the correlation structure instead of two (LK_1 compared to LK_3) an increase in magnitude can also be observed.

These results suggest that the Path Size attribute should be included in the utility, but does not correct the IIA assumption on the random terms. We do however not recommend to use $\gamma > 0$ in the generalized Path Size formulation since we have shown (see section 3.3) that this formulation shows counter intuitive results even when it comes to penalizing longer paths in favor of shorter ones.

The σ estimate is highly significant for all the Logit Kernel models, suggesting that this factor analytic specification captures an important correlation structure. The LK_1 model has a significantly higher final log-likelihood value than the PSL model (from -2796.70 to -2262.50) and there is a remarkable increase in rho-square value (from 0.18 to 0.33). Moreover, when all the subpaths are considered in the correlation structure there is a further increase in model fit. Ramming (2001) also noted in his estimations that the Logit Kernel model combined with a Path Size attribute had a better model fit than the PSL model. Hoogendoorn-Lanser et al. (2005) discussed the factor analytic specification used by Ramming (2001), and found it promising but pointed out that the estimation process is very long and that the stability of the solutions is a concern. This is due to the high number of Gaussian variates. Ramming (2001) used one for each link in the choice set which resulted in 856 Gaussian variates and up to 100'000 simulation draws were used. The advantage of the subpath approach is clear. The number of Gaussian variates is decreased while the results seem as promising as the one presented by Ramming (2001).

Finally, note that the results seem robust to the inclusion of the test subpath (S_8). When comparing LK_1 with LK_2 and LK_3 with LK_4 no significant change in parameter estimates neither in model fit can be observed.

6 Conclusion

In this paper we have shown that all Path Size formulations presented in the literature, except the original one, show counter intuitive results regarding the correction of the IIA assumption on the random terms in the MNL model. Furthermore, the generalized Path Size does not achieve its original goal namely, penalizing longer paths in favor of shorter ones. There is however, an interesting behavioural interpretation of the Path Size attribute. Indeed, overlap can be attractive for travellers since it provides the possibility of switching between different routes. Model estimations results presented here, and also in the literature (Ramming, 2001, Hoogendoorn-Lanser, 2005, Hoogendoorn-Lanser et al., 2005) suggest that this behavioural aspect is very important. We therefore conclude that a Path Size attribute (original formulation) should be included in the deterministic part of the utility. It is however clear that including such an attribute does not correct the IIA assumption on the random terms, finding an appropriate formulation for this purpose is an interesting issue for future research.

An adaptation of the Logit Kernel model to include a correlation structure of subpaths has been presented. The model estimations show very promising results for the Logit Kernel model combined with a Path Size attribute. The increase in model fit is remarkable, and the covariance parameter estimates suggest that this formulation captures an important correlation structure. The same observations has been made in Bekhor et al. (2002), who presented a factor analytic specification of the Logit Kernel model based on links. Including subpaths instead of links in the correlation structure considerably decreases the complexity of the model in terms of number of Gaussian variates, but seems to still capture the correlation structure.

These preliminary results show the potential of including subpath components in route choice modelling. Further research will be dedicated to the robustness of the subpath definition. The deterministic choice set generation used here, will also be replaced by the probabilistic approach presented in section 4.1.

References

Axhausen, K., Schönfelder, S., Wolf, J., Oliveira, M. and Samaga, U. (2003). 80 weeks of GPS traces: Approaches to enriching the trip information, *Technical report*, Institut für Verkehrsplanung and Transportsysteme, ETH Zürich.

Azevedo, J., Costa, M. S., Madeira, J. S. and Martins, E. V. (1993). An algorithm for the ranking of shortest paths, *European Journal of Operational Research* **69**: 97–106.

Bekhor, S., Ben-Akiva, M. and Ramming, M. (2001). Estimating route choice models for large urban networks, 9th World Conference on Transport Research, Seoul, Korea.

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

Ben-Akiva, M., Bergman, M., Daly, A. and Ramaswamy, R. (1984). Modeling inter-urban route choice behaviour, *in* J. Vollmuller and R. Hamerslag (eds), *Proceedings of the 9th international symposium on transportation and traffic theory*, VNU Science Press, Utrecht, Netherlands, pp. 299–330.

Ben-Akiva, M. and Bierlaire, M. (1999). *Discrete choice models with applications to departure time and route choice*, Vol. 26 of *Handbook of Transportation Science*, Kluwer, pp. 5–34.

Ben-Akiva, M. and Boccara, B. (1995). Discrete choice models with latent choice sets, *International Journal of Research in Marketing* **12**: 9–24.

Ben-Akiva, M. and Bolduc, D. (1996). Multinomial probit with a logit kernel and a general parametric specification of the covariance structure, *Paper presented at the 3rd Invitational Choice Symposium*, Columbia University.

Ben-Akiva, M. and Lerman, S. (1985). *Discrete choice analysis: Theroy and application to travel demand*, MIT Press, Cambridge, Massachusetts.

Ben-Akiva, M., Ramming, M. and Walker, J. (1999). Improving behvioral realism of urban transportation models through explicit treatment of individuals' spatial ability. Submitted for presentation at the conference on Social Change and Sustainable Transport, Berkely, California, March 10-13.

Bierlaire, M. (2002). The network GEV model, *Proceedings of the 2nd Swiss Transportation Research Conference*, Ascona, Switzerland. www.strc.ch/pdf_2002/bierlaire2.zip.

Bierlaire, M. (2003a). Biogeme: a free package for the estimation of discrete choice models, *3rd Swiss Transport Research Conference, Ascona*.

Bierlaire, M. (2003b). An introduction to biogeme. http://roso.epfl.ch/biogeme.

Bierlaire, M. (forthcoming). A theoretical analysis of the cross-nested logit model, *Annals of operations research* . Accepted for publication.

Bolduc, D. and Ben-Akiva, M. (1991). A multinomial probit formulation for large choice sets, *Proceedings of the 6th International Conference on Travel Behaviour*, Vol. 2, pp. 243–258.

Burrell, J. (1968). Multipath route assignment and its application to capacity restraint, *Proceedings of the 4th International Symposium on the Theory of Road and Traffic Flow*.

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

Cascetta, E. and Papola, A. (2001). Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand, *Transportation Research Part C* **9**: 249–263.

Chu, C. (1989). A paired combinatorial logit model for travel demand analysis, *Proceedings of the fifth World Conference on Transportation Researach*, Vol. 4, Ventura, CA, pp. 295–309.

Dagazo, C. and Sheffi, Y. (1977). On stochastic models of traffic assignment, *Transportation Science* **11**(3): 253–274.

Daly, A. and Bierlaire, M. (2003). A general and operational representation of GEV models, *Technical Report RO-030502*, Institute of Mathematics, Operations Research Group ROSO, EPFL, Lausanne, Switzerland.

de la Barra, T., Pérez, B. and Añez, J. (1993). Mutidimensional path search and assignment, *Proceedings of the 21st PTRC Summer Meeting*, pp. 307–319.

Dial, R. (1971). A probabilitstic multipath traffic assignment algorithm which obviates path enumeration, *Transportation Research* **5**(2): 83–111.

Gliebe, J., Koppelman, F. and Ziliaskopoulos, A. (1999). Route choice using a paired combinatorial logit model, *78th Annual Meeting of the Transportation Research Board*, Washington, DC.

Gopinath, D. (1995). *Modeling heterogeneity in discrete choice processes: Application to travel demand*, PhD thesis, MIT.

Han, B. (2001). *Analyzing car ownership and route choices using discrete choice models*, PhD thesis, Department of infrastructure and planning, KTH, Stockholm.

Han, B., Algers, S. and Engelson, L. (2001). Accommodating drivers' taste variation and repetead choice correlation in route choice modeling by using the mixed logit model, *80th Annual Meeting of the Transportation Research Board*.

Hoogendoorn-Lanser, S. (2005). *Modelling Travel Behaviour in Multi-modal Networks*, PhD thesis, Deft University of Technology.

Hoogendoorn-Lanser, S., van Nes, R. and Bovy, P. (2005). Path size and overlap in multi-model transport networks, a new interpretation, 16th International Symposium on Transportation and Traffic Theory. Accepted for publication.

Koppelman, F. and Wen, C.-H. (2000). The paired combinatorial logit model: properties, estimation and application, *Transportation Research Part B* **34**: 75–89.

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

Marzano, V. and Papola, A. (2004). A link based path-multilevel logit model for route choice which allows implicit path enumeration, *Proceedings of the European Transport Conference*.

Morikawa, T. (1996). A hybrid probabilistic choice set model with compensatory and noncompensatory choice rules, *Proceedings of the 7th World Conference on Transport Research*, Vol. 1, pp. 317–325.

Nielsen, O., Daly, A. and Frederiksen, R. (2002). A stochastic route choice model for car travellers in the Copenhagen region, *Networks and Spatial Economics* **2**: 327–346.

Paag, H., Daly, A. and Rohr, C. (2002). Predicting use of the copenhagen harbour tunnel, *in* D. Hensher (ed.), *Travel Behaviour Research: The Leading Edge*, Pergamon Press, pp. 627–646.

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

Prashker, J. and Bekhor, S. (1998). Investigation of stochastic network loading procedures, *77th Annual Meeting of the Transportation Research Board* .

Ramming, M. (2001). *Network Knowledge and Route Choice*, PhD thesis, Massachusetts Institute of Technology.

Schönfelder, S., Axhausen, K., Antille, N. and Bierlaire, M. (2002). Exploring the potentials of automatically collected GPS data for travel behaviour analysis - a swedish data source, *in* J. Möltgen and A. Wytzisk (eds), *GI-Technologien für Verkehr und Logistik*, number 13, pp. 155–179.

Schönfelder, S. and Samaga, U. (2003). Where do you want to go today? - More observations on daily mobility, *3rd Swiss Transport Research Conference, Ascona*.

Scott, K., Pabón-Jiménez, G. and Bernstein, D. (1997). Finding alternatives to the best path, *76th Annual Meeting of the Transportation Research Board*.

SIKA (2001). *RES 2001. Den nationella resundersökningen*, Statistiska Centralbyrån. www.sika-institute.se.

Swait, J. and Ben-Akiva, M. (1987). Incorporating random constraints in discrete models of choice set generation, *Transportation Research Part B* **21**(2): 91–102.

Vovsha, P. and Bekhor, S. (1998). Link-nested logit model of route choice overcoming route overlapping problem, *Transportation Research Record* **1645**: 133–142.

Yai, T., Iwakura, S. and Morichi, S. (1997). Multinomial probit with structured covariance for route choice behavior, *Transportation Research Part B* **31**(3): 195–207.