Effects of Choice Set Size and Route Choice Models on Path-Based Traffic Assignment

Shlomo Bekhor*, Tomer Toledo and Joseph N. Prashker

Faculty of Civil and Environmental Engineering
Technion – Israel Institute of Technology
Haifa, 32000, Israel

Published in Transportmetrica 4(2), pp. 117-133, 2008

* Corresponding Author:
Phone: +972-4-8292460
Fax: +972-4-8295708
Email: sbekhor@technion.ac.il

ABSTRACT

Few of the recently developed route choice models have actually been applied in traffic assignment problems. This paper discusses the implementation of selected route choice models in stochastic user equilibrium algorithms. The focus of the paper is on path-based assignment, which is essential in the implementation of route choice models. The paper analyzes the effect of choice set size and selected choice models on problem convergence, running time and selected results. The results presented in the paper indicate that for real-size networks, generation of a large number of alternative routes is needed. Furthermore, convergence properties greatly improve if the generated routes are sufficiently disjointed.

KEYWORDS

Route Choice, Traffic Assignment, Stochastic User Equilibrium, Path-based Algorithms

1. INTRODUCTION

A number of discrete choice model structures have been adapted to model route choice behavior in recent years. These models attempt to capture the impact of similarity among various routes on drivers' perceptions and decisions. They range from modifications of the multinomial logit (MNL) model, such as C-logit and path-size logit (PSL) that capture similarities through additional terms in the systematic utilities of the various routes, to more complex models based on the generalized extreme value (GEV) theory, such as paired combinatorial logit (PCL), cross-nested logit (CNL) and generalized nested logit (GNL), and error components and logit kernel (LK) models, which capture these similarities by allowing more general correlation structures.

Few of these route choice models have actually been applied in traffic assignment procedures. This paper discusses the potential implementation of selected route choice models as loading procedures in stochastic user equilibrium (SUE) assignment and presents results to illustrate the combined effect of congestion and the representation of similarity among routes on assignment results.

Two main issues are investigated in this paper: The first issue is the influence of the path set on the equilibrium solution. The approach of the paper is to generate routes prior to the assignment, perform quantitative changes in the number of routes in the path set and subsequently calculate the SUE solution for different networks. For real-sized networks, it is impossible to enumerate all routes, and so heuristic rules are needed to confine the path set to a reasonable size. The paper uses a combination of two well-known generation methods to produce the choice set.

The second issue discussed in the paper is the influence of the route choice model on the equilibrium flow pattern. It is well known in the literature that for very high demand levels SUE flow patterns approaches those given by the deterministic user equilibrium solution. However, depending on the route choice model SUE flows may be quite different from the deterministic UE flows for moderate to high levels of demand. Using real size networks, the paper compares link flows and path flows obtained from path-based SUE assignments for two choice models: the MNL model as a representative of a simple choice structure, and the CNL model as a representative of a more complex choice structure.

The paper is organized as follows. First, a brief review of route choice models, choice set generation methods and SUE formulations and path-based algorithms are presented. The definition and generation of the choice set is discussed next. The SUE solutions for different choice sets and model specifications are illustrated in a subsequent section. The final section discusses the potential use of the different route choice models in traffic assignment problems.

2. LITERATURE REVIEW

2.1 Route Choice Models

Route choice models can be classified according to the model structure, following Prashker and Bekhor (2004). The first class of models relates to the MNL and its modifications, such as C-Logit (Cascetta et al., 1996) and Path Size Logit (Ben-Akiva and Bierlaire, 1999). These models were developed in an attempt to overcome the overlapping problem, while still retaining the MNL structure. In these models, the similarity among routes is modeled by including commonality measures in the deterministic component of the utility function.

The second class encompasses models derived from the GEV theorem of McFadden (1978). The fundamental difference between route choice models in this class and the former class is that the similarity among routes is captured in the structure of the error component of the utility function. By suitably defining a generator function that satisfies the properties required by the GEV theorem, more general logit functions can be obtained. Examples of such models are the PCL model (Chu 1989), the CNL model (Vovsha 1997) and the GNL model (Wen and Koppelman 2001). These models were adapted to route choice situation by several researchers: Prashker and Bekhor (1998), Gliebe et al. (1999), Bekhor and Prashker (2001), Pravinvongvuth and Chen (2005).

The third class of models includes the Multinomial Probit (MNP) and LK models. These models are more general than the closed-form GEV models. The similarity among routes is also captured in the error component of the utility function. The MNP model was first proposed by Daganzo and Sheffi (1977) to model route choice as an alternative to the MNL model. The choice model is based on the assumption of normal distribution for the random utility.

The MNP and LK models require explicit evaluation of the covariance matrix in order to compute the choice probabilities. The key question is how to relate these covariance to measurable network parameters. Sheffi and Powell (1982) proposed to assume that the variance is proportional to a fixed characteristic of the link as e.g. free flow travel time or length. Yai et al. (1997) proposed an MNP model which relates the covariance matrix to measurable overlapping measures, such as the common length of the routes. This method was also used by Bekhor et al. (2002) to adapt the LK to route choice situation. To our knowledge, path-based algorithms for MNP and LK are yet to be developed. For this reason, this paper compares path-based results for closed-form route choice models such as MNL (as a representative of the first class of models) and CNL (as a representative of the second class of models).

2.2 Path Set Generation Methods

The actual large number of routes in a network poses a problem for specifying the alternative routes and in defining choice sets. Even for small networks, the enumeration of all possible alternative routes is not a trivial matter. The number of possible routes increases exponentially with the dimensions of the network. When routes are explicitly considered, a selective approach is needed to approximate the choice set. In this subsection, we briefly review methods to explicitly generate choice sets.

The most straightforward path generation approach searches for the first “K” shortest paths that minimize the generalized path costs. Shortest path algorithms assume implicitly awareness of all the link attributes. Van der Zijpp and Fiorenzo-Catalano (2001) present a constrained method to generate routes by finding directly feasible K-shortest paths and exploiting a wide class of constraints.

Ben-Akiva et al. (1984) propose an approach for generating possible paths by labeling each route according to a criterion for which the path is optimum. This approach assumes that travelers may have different objective functions. Each criterion yields a different preferred route. Dial (2000) generalizes the labeling method by constructing a set of efficient paths, where being efficient means minimizing a linear combination of label costs.

Azevedo et al. (1993) define a link elimination approach where the shortest path is calculated and then all the links on this shortest path are removed from the network. A new shortest path is then calculated and so on. This approach guarantees generation of disjoint paths each time. The main problem with this approach is that it may quickly lead to network disconnection due to the removal of centroid connectors and major junctions. A variant to this approach obviates the problem by eliminating individual links or a combination of links from the shortest path rather than all the links on the shortest path simultaneously.

De la Barra et al. (1993) proposed the link penalty approach, where the impedances of the links on the shortest path are increased before the next best path is calculated. The process continues until no more new paths are produced. Park and Rilett (1997) use a variation of this approach in which the impedance on links within a certain distance from the origin or the destination is not increased. Scott et al. (1997) optimize the penalizing factor for impedances in order to generate a next best path that overlaps with the shortest path by no more than a given number of links.

Simulation methods assume that travelers erroneously perceive link attributes. Therefore extracting random draws from a distribution that might represent drivers' perceptions appears suitable. Fiorenzo-Catalano and Van der Zijpp (2001) implement a version of the Monte Carlo technique by gradually increasing the variance of the random components in the model in order to keep the rate with which new paths are found constant. Bovy and Fiorenzo-Catalano (2007) discuss route choice set generation in uni-modal and multi-modal networks. Bekhor et al. (2006) verify the suitability of simulation methods to produce paths similar to those observed for a case network in Boston. A combination of methods was found to yield the highest coverage of traveled routes, with the simulation method being the best single generation method.

Another class of constrained enumeration procedure was introduced by Friedrich et al. (2001), who applied a branch and bound assignment procedure for transit networks using a timetable-based search algorithm. Hoogendoorn-Lanser (2005) adapted the same procedure for choice set generation in the analysis of multimodal transport networks. In the transit applications mentioned above, the method exploits predefined route sections. Prato and Bekhor (2006) designed a different approach for road networks. The algorithm constructs a connection tree between origin and destination of a trip, by processing sequences of links according to a branching rule that accounts for logical constraints formulated to increase route likelihood and heterogeneity. Initial results presented in Prato and Bekhor (2006) indicate that the branch and bound method outperforms the simulation method, however more research is needed to validate both the simulation and branch and bound methods against large datasets. In the present paper, we use simpler techniques such as link penalty and link elimination methods, because of the relative ease and fast computational implementation. According to Bekhor et al. (2006), many routes need to be generated with these methods in order to obtain an acceptable coverage.

Recently, Bar-Gera and Boyce (2005) analyzed the characteristics of route sets in the solution of a deterministic user equilibrium model. The network used in the analysis represents the Chicago area, composed of 1,790 zones, 12,982 nodes and 39,018 links. An average of 2.686

routes per OD pair was obtained, 56% of the OD pairs were connected by only one route, and more than 95% of the OD pairs were connected by up to 10 routes. The authors found that numbers of routes used largely depends on the level of congestion in the network.

In the context of path set generation in stochastic traffic assignment problems, there is relatively few evidence in the literature. Most models use implicit methods such as Dial’s (1971) method. In Huang (1995) the set of paths is determined in a preliminary phase which combines the paths found by Dial's approach and those obtained from solving a deterministic user equilibrium assignment. In Damberg et al. (1996), the choice set was based on the set of shortest paths obtained during the iterations of deterministic assignment. The authors recognized that the choice set affects the solution properties.

2.3 Stochastic User Equilibrium Models and Algorithms

Daganzo and Sheffi (1977) defined the concept of SUE as a state in which no driver can improve his/her perceived travel time by unilaterally changing routes. The “stochastic” term is related to a probabilistic route choice model, instead of simply assuming the shortest path as in the deterministic user equilibrium model.

Sheffi and Powell (1982) formulated the SUE assignment problem as an unconstrained mathematical program. This formulation is appropriate for a wide range of probability distributions for the path costs that satisfy certain conditions. The well-known Method of Successive Averages (MSA) developed by Powell and Sheffi (1982) was the first algorithm applied to solve the SUE problem. This algorithm can be applied with any stochastic network loading method. The step size is predetermined by a descent sequence with respect to the iterations.

For GEV-type route choice models, alternative optimization formulations can be derived. Fisk (1980) developed a path-based SUE formulation for the MNL model, using an entropy term. Similar entropy-based formulations have been proposed for SUE assignment based on other GEV models, including CNL and PCL (Bekhor and Prashker 1999), and GNL (Bekhor and Prashker 2001). For brevity, we present the SUE optimization formulation for the CNL model.

$$
\begin{aligned}
&Min \quad Z = Z_1 + Z_2 + Z_3 \\
&Z_1 = \sum_a \int_0^{x_a} c_a(w)dw \\
&Z_2 = \frac{\mu}{\theta} \sum_{rs} \sum_m \sum_k f_k^{rs} \ln \frac{f_{mk}^{rs}}{\left(\alpha_{mk}^{rs}\right)^{1/\mu}} \\
&Z_3 = \frac{1-\mu}{\theta} \sum_{rs} \sum_m \left(\sum_k f_{mk}^{rs} \right) \ln \left(\sum_k f_{mk}^{rs} \right) \\
&s.t. \ \sum_{\mathrm{m}} \sum_k f_{mk}^{rs} = q^{rs}, \ \forall \ \mathrm{r,s} \\
&\qquad f_{mk}^{rs} \geq 0, \qquad \forall \ \mathrm{m,k,r,s}
\end{aligned}
\tag{1}
$$

Where:

f_{mk}^{rs}: Flow on path k between origin r and destination s that is attributed to the choice of link m. The path flow f_k^{rs} is related to these flows as: $f_k^{rs} = \sum_m f_{mk}^{rs}$

q^{rs}: travel demand between r and s

c_a: travel cost on link a

θ: dispersion coefficient (travel time parameter)

μ: nesting coefficient of the CNL function

α_{mk}^{rs}: inclusion coefficient of link m in path k between r and s.

Bekhor and Prashker (1999) showed that the above formulation collapses to Fisk's formulation for the MNL based SUE assignment when the nesting coefficient is equal to 1. The above CNL-SUE formulation assumes a single nesting coefficient μ. If this coefficient is nest-specific, the resulting formulation is equivalent to the GNL-SUE model. To simplify interpretation of results, we will assume a single nesting coefficient in this paper.

Maher (1998) developed link-based algorithms for the logit SUE problem, using Sheffi and Powell's formulation. He proposed an algorithm that uses the same search direction as the MSA algorithm, but calculates an approximately optimal step size in this direction, thus improving overall convergence. Dial (2001) also developed an algorithm for the logit route choice model, which exploits the property of the MNL model that allows a unique mapping of path flows from link flows (and vice versa).

The algorithms discussed above solve the assignment problem in the space of link flows, and so do not require explicit enumeration of the path choice set. Instead, these algorithms assume an implicit choice set, such as the use of all efficient paths (Maher 1998, Dial 2001), or all cyclic and acyclic paths (Bell 1995, Akamatsu 1996). Huang and Bell (1998) proposed a method to exclude all cyclic flows in the logit-based SUE assignment. However, from a behavioral standpoint it may be more realistic to assume that drivers choose from more limited route sets. Path-based formulations allow a more flexible definition of the choice set.

Damberg et al. (1996) extended the path-based disaggregated simplicial decomposition (DSD) algorithm (Larsson and Patriksson 1992) to solve the SUE problem for the MNL model. They show that given a current (feasible) path flow pattern, an auxiliary path flow pattern can be obtained by calculating the MNL model with current travel times. A line search is then performed along the direction of the difference between the current flow and auxiliary flow patterns. The authors show that this procedure converges to a unique solution for a given set of routes.

The above algorithm can be extended to other logit-type models, such as CNL and PCL. Bekhor and Toledo (2005) developed the GP2 algorithm and compared the performance with respect to the DSD algorithm. Their conclusion was that both algorithms have similar convergence properties, and both outperform the path-based MSA algorithm. Chen and al. (2003) developed an algorithm based on the partial linearization method for solving the PCL stochastic user equilibrium problem. Recently, Bekhor et al. (2007) presents an extension of the DSD algorithm for the CNL model, which is used in the present paper.

3. INFLUENCE OF THE PATH CHOICE SET

Two networks were used to examine the impact of the size of the path choice set on the equilibrium solution reached. The first network is the well-known Sioux Falls network (Leblanc, 1973). The second is the Winnipeg network provided in the EMME/2 software. We first analyze the results for the Sioux Falls network.

3.1 The Sioux Falls Network

The network is composed of 24 nodes, 76 links and 528 OD pairs with positive demand. The paths were generated prior to the assignment, using a combination of the link elimination method of Azevedo et al. (1993) and the penalty method of De La Barra et al. (1993), with a penalty of 5% on the travel times on the shortest path links. Only acyclic paths were considered in these methods. On average, 6.3 paths were generated for each OD pair. The maximum number of routes generated for an OD pair was 13.

For the purposes of the present analysis, we assumed a value of 0.5 units/min for the dispersion parameter θ. This value is consistent with empirical evidence (Ramming 2001), which estimated values in the range from –0.4 to –0.6 units/min, depending on the model structure and other parameters in the model. This value indicates that given a 5-minute difference between two paths, about 8% of the drivers will chose the route with the higher cost.

To test the impact of the choice set on the equilibrium results, we varied the maximum numbers of paths allowed for each OD pair and performed path-based SUE assignments using MNL as the route choice model. In all assignments, the DSD algorithm of Damberg et al. (1996) was used, using the Armijo rule for step size calculation. The stopping criterion for the assignment is 1% for the root mean square percentage error (RMSPE) of the path flows between consecutive iterations. Figure 1 below shows the equilibrium objective function values reached.

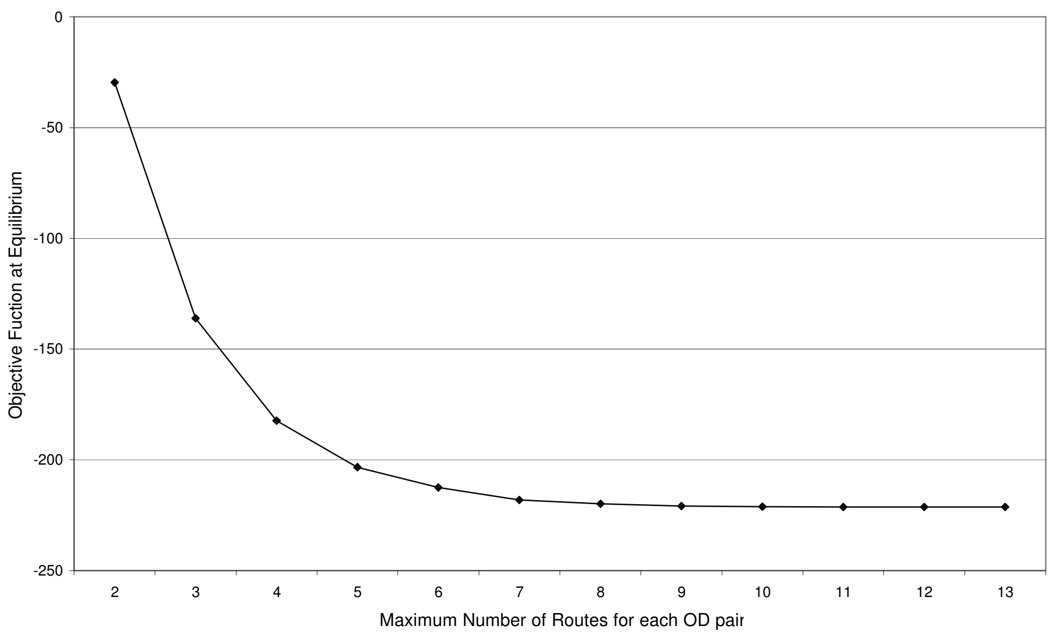

Figure 1. Influence of the Choice Set Size on the Equilibrium Solution – Sioux Falls Network.

The results indicate that for the Sioux Falls network, relatively few paths are needed to achieve equilibrium at a reasonable level. This result suggests that path-based SUE assignment may be performed with few routes in the path set.

3.2 The Winnipeg Network

This network is composed of 948 nodes (154 of them corresponding to traffic area zones), 2,535 links and 4,345 OD pairs with positive demand in the matrix. The total demand on the network is for 54,459 trips. The volume-delay function for each link is based on the BPR formula with link-specific parameters, calculated from the original Emme/2 data.

Similar to the case of the Sioux Falls network, travel routes were generated prior to the assignment using a combination of the link elimination method and the penalty method. A total of 174,491 unique paths were generated for all OD pairs (average of 40.1 paths per OD pair). The maximum number of routes generated for any OD pair was 50. More routes can be generated, but as we will show below for the purposes of this paper this maximum number is sufficient. Figure 2 illustrates the Winnipeg network. The dots indicate the centroids of the traffic area zones.

Figure 2. The Winnipeg network

Inspection of the generated routes reveals that the choice set used for the analysis includes both completely disjointed routes and very similar routes. This was expected due to the methods (link penalty and link elimination) selected to generate the routes: the link elimination method produces disjoint routes (because of the removal of all links belonging to the shortest path) and the link penalty method produces similar routes because of the low penalty (5% increase link travel time) used to find the subsequent routes.

The methodology used for the comparison was similar to the Sioux Falls example: each path-based SUE assignment was performed with a different maximum number of routes in the path set, starting from 2 routes up to the maximum of 50, using the MNL as the route choice probability function. In all assignments, the DSD algorithm of Damberg et al. (1996) was used, using the Armijo rule for step size calculation. The stopping criterion for the assignment is 1% for the RMSPE of the path flows. Table 1 summarizes the results, and Figure 3 displays the results of the ratio Z / Z* for both the MNL and CNL route choice models. The CNL-SUE results are further presented in Table 3.

Table 1. Summary of SUE results for the Winnipeg Network – MNL choice function.

Maximum paths generated per OD	Total number of paths	Number of paths per OD	Minimum Objective function value (Z)	Ratio Z / Z* using 50 paths	Number of Iterations to converge	CPU time (sec)	CPU time per iteration (sec)
2	8690	2.0	1111964	1.21	35	493	14.1
3	13035	3.0	1051540	1.15	58	1335	23.0
4	17380	4.0	1023947	1.12	63	1943	30.8
5	21723	5.0	1007347	1.10	73	2895	39.7
6	26057	6.0	994993	1.08	76	3865	50.9
7	30386	7.0	985460	1.07	64	3569	55.8
8	34708	8.0	978177	1.07	56	3571	63.8
9	39025	9.0	972004	1.06	56	4101	73.2
10	43332	10.0	966810	1.05	51	4012	78.7
15	64575	14.9	949299	1.03	41	5152	125.7
20	85278	19.6	939014	1.02	36	5767	160.2
25	105198	24.2	932290	1.02	37	7539	203.8
30	124066	28.6	927121	1.01	29	6952	239.7
40	157104	36.2	920440	1.00	29	9360	322.8
50	174491	40.2	917976	1.00	28	9967	356.0

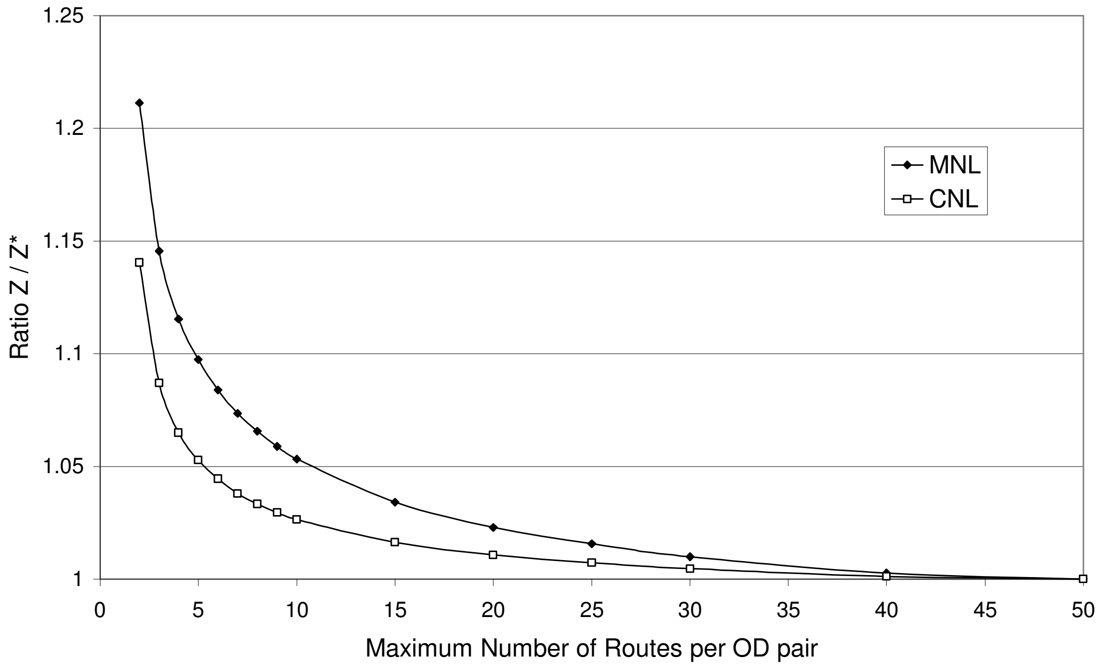

Figure 3. Influence of the Choice Set Size on the Equilibrium Solution – Winnipeg Network.

The results displayed above show a pattern similar to the one observed in the Sioux Falls result (Figure 1): as the number of routes increases, the minimum value of the objective function at convergence decreases. However, compared to the Sioux Falls network, more routes are needed in order to obtain equilibrium at a reasonable level. The objective function using only 2 routes for each OD pair is 21% higher than the one using 50 routes. If 10 routes are used, the objective function is 5% higher compared to the full set of 50 routes. The CNL-SUE pattern is similar to the MNL-SUE pattern.

Note that the CPU time per iteration increases linearly with the size of the path set, as can be seen in Figure 4 below. This result is expected because a predefined path set is used in the assignment procedure, and the increase of the path set mainly affects the probability calculations. The average time to perform an iteration for the maximum path set used (50) was close to 6 minutes, using a Pentium IV computer with 3.0 GHz CPU and 512 MB RAM.

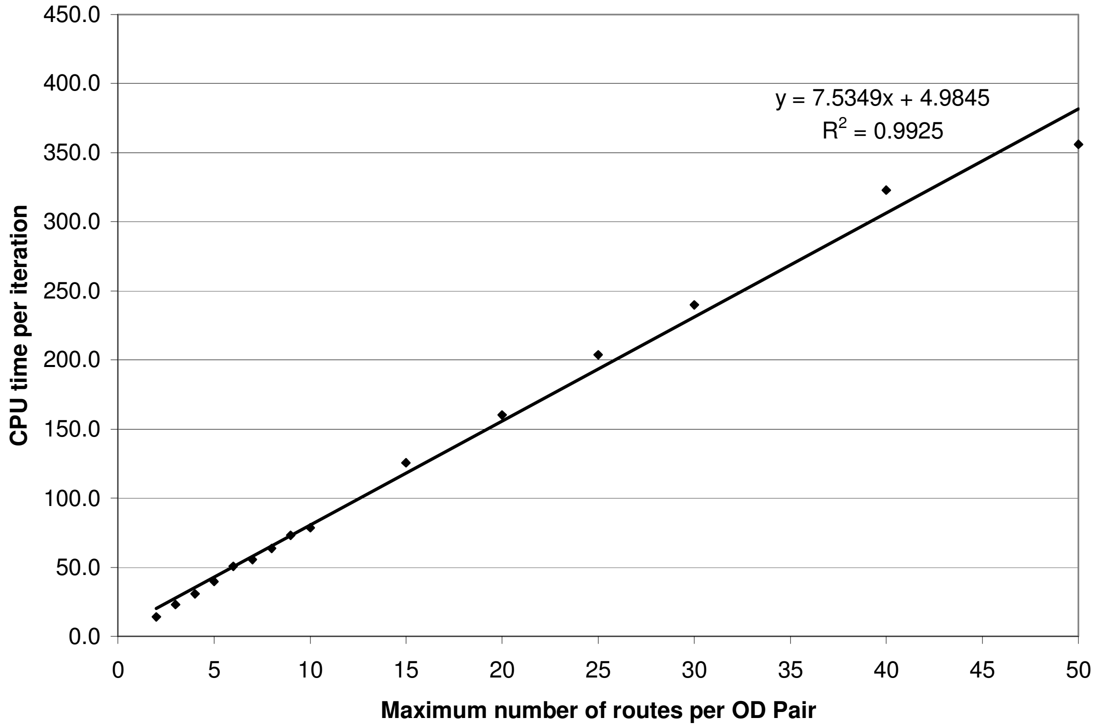

Figure 4. Influence of the Choice Set Size on Algorithmic Performance – Winnipeg Network

4. INFLUENCE OF THE CHOICE MODEL

This section analyzes the influence of the route choice model on the equilibrium flow pattern. The section focuses on the comparison of the CNL-SUE and MNL-SUE path flows for different networks. Subsection 4.1 presents results for the Sioux Falls network, and subsection 4.2 presents results for the Winnipeg network. Subsection 4.3 summarizes selected results for both networks.

4.1 Sioux Falls Network

This subsection compares the assignment results obtained with CNL and MNL route choice models depending on the values of the model parameters and the level of demand for travel. A path-based DSD algorithm was used for all models, and the stopping criterion was set to 0.001 maximum difference between link flows in consecutive iterations. The following delay function was used for all links:

$$c_a = c_{0a} * \left(1 + 0.6\left(\frac{x_a}{1000}\right)^4\right) \qquad (2)$$

Where:

c_a: travel cost on link a;

c_{0a}: free-flow travel cost on link a;

x_a: flow on link a.

Figures 5 to 7 shows the RMSPE between CNL-SUE and MNL-SUE path flows for different values of the parameters θ, μ and the demand. Note that when the nesting coefficient is equal to 1, the CNL-SUE model collapses to the MNL-SUE model and consequently the difference is zero. This theoretical result was verified by the assignment results.

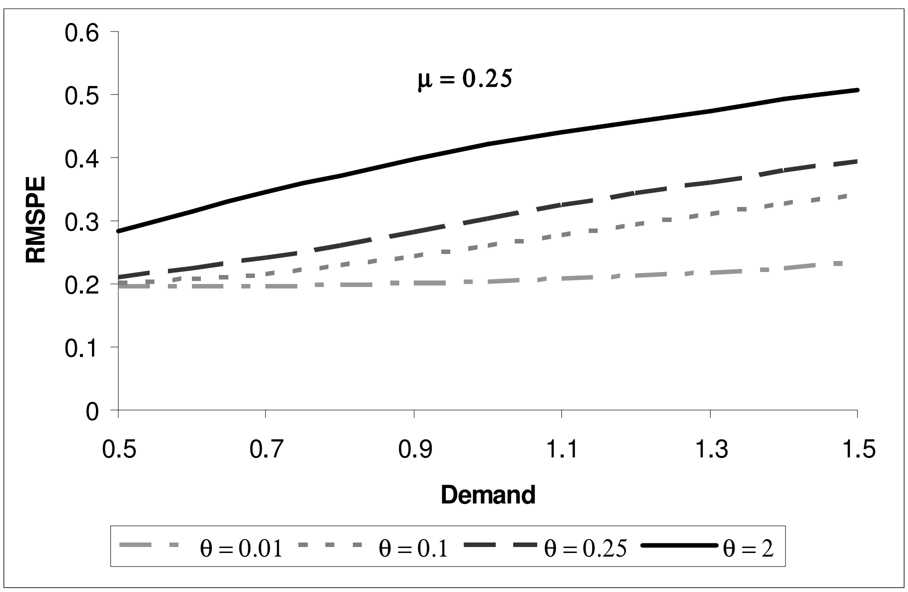

Figure 5. Sioux Falls Network - Influence of the Demand and θ on the difference between CNL-SUE and MNL-SUE path flows (μ=0.25)

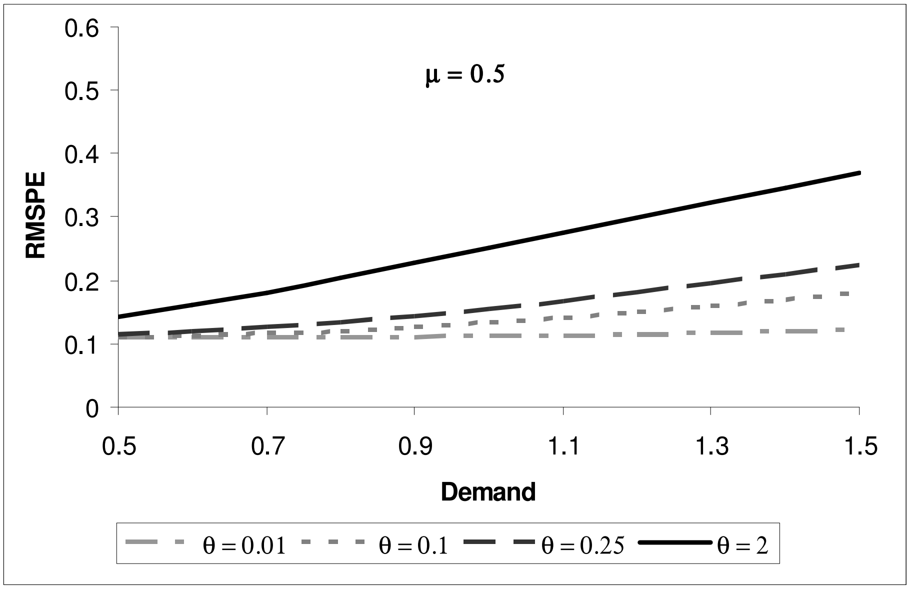

Figure 6. Sioux Falls Network - Influence of the Demand and θ on the difference between CNL-SUE and MNL-SUE path flows (μ=0.5)

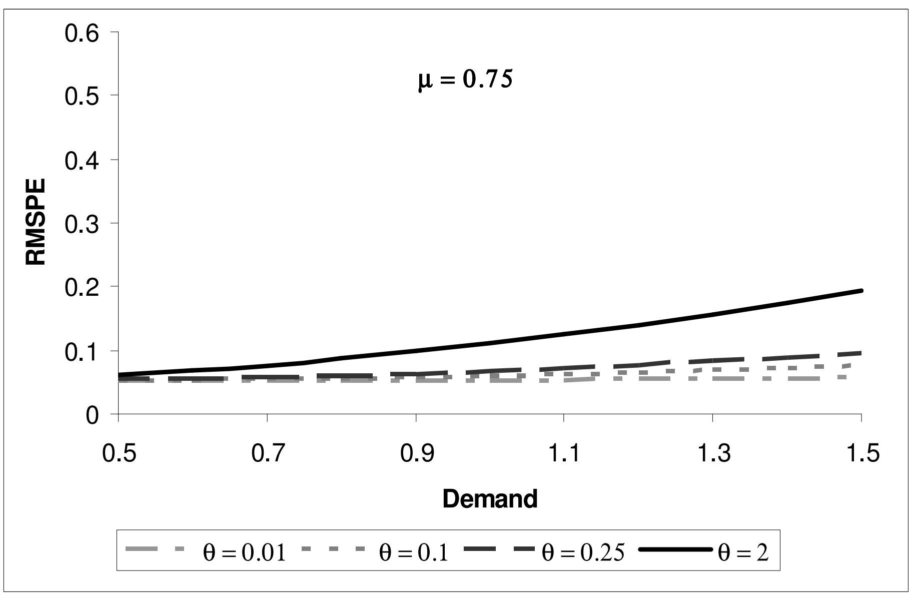

Figure 7. Sioux Falls Network - Influence of the Demand and θ on the difference between CNL-SUE and MNL-SUE path flows (μ=0.75)

The results indicate that only for large values of θ and small values of μ, the differences between CNL-SUE and MNL-SUE path flows are significant. As expected, when μ increases, the differences between CNL-SUE and MNL-SUE path flows decrease.

4.2 Winnipeg Network

This subsection compares results obtained from path-based SUE assignments using the MNL and CNL route choice models for the Winnipeg network. The CNL-SUE assignments were performed using the same methodology described in section 3.2 for the MNL model. The DSD algorithm of Damberg et al. (1996) was modified to compute choice probabilities based on the CNL model instead of the MNL model. The Armijo rule was used for step size calculation. The stopping criterion for the assignment is 1% for the RMSPE of path flows between consecutive iterations. The nesting coefficient of the CNL model was set to 0.5. The dispersion parameter was set equal to 0.5 units/min for both the CNL and MNL models. Table 2 summarizes the results for the CNL model.

Table 2. Summary of SUE results for the Winnipeg Network – CNL choice function.

Maximum paths generated per OD	Total number of paths	Number of path per OD	Minimum Objective function value (Z)	Ratio to Z using 50 paths	Number of Iterations to converge	CPU time (sec)	CPU time per iteration (sec)
2	8690	2.0	1128444	1.14	50	970	19.4
3	13035	3.0	1075475	1.09	45	1321	29.4
4	17380	4.0	1053691	1.06	42	1578	37.6
5	21723	5.0	1041684	1.05	52	2370	45.6
6	26057	6.0	1033424	1.04	55	3051	55.5
7	30386	7.0	1027005	1.04	76	5097	67.1
8	34708	8.0	1022400	1.03	69	5218	75.6
9	39025	9.0	1018658	1.03	70	6129	87.6
10	43332	10.0	1015593	1.03	70	6612	94.5
15	64575	14.9	1005584	1.02	70	10294	147.1
20	85278	19.6	1000073	1.01	57	11418	200.3
25	105198	24.2	996603	1.01	52	12973	249.5
30	124066	28.6	993961	1.00	56	16248	290.1
40	157104	36.2	990655	1.00	56	21412	382.4
50	174491	40.2	989445	1.00	50	21597	431.9

Comparison of the results reported in Table 2 to those reported for the MNL model in Table 1 shows that, as expected, the CNL model takes more time to converge, because of the additional computations needed to calculate the CNL choice probabilities. The CPU time for the run using all 50 routes is almost 6 hours. As with the MNL model, the CPU time per iteration increases linearly with an increase in the number of routes in the path set. Figure 8 compares the number of iterations in each model to reach the same convergence level of 0.001, representing the maximum difference between link flows in consecutive iterations.

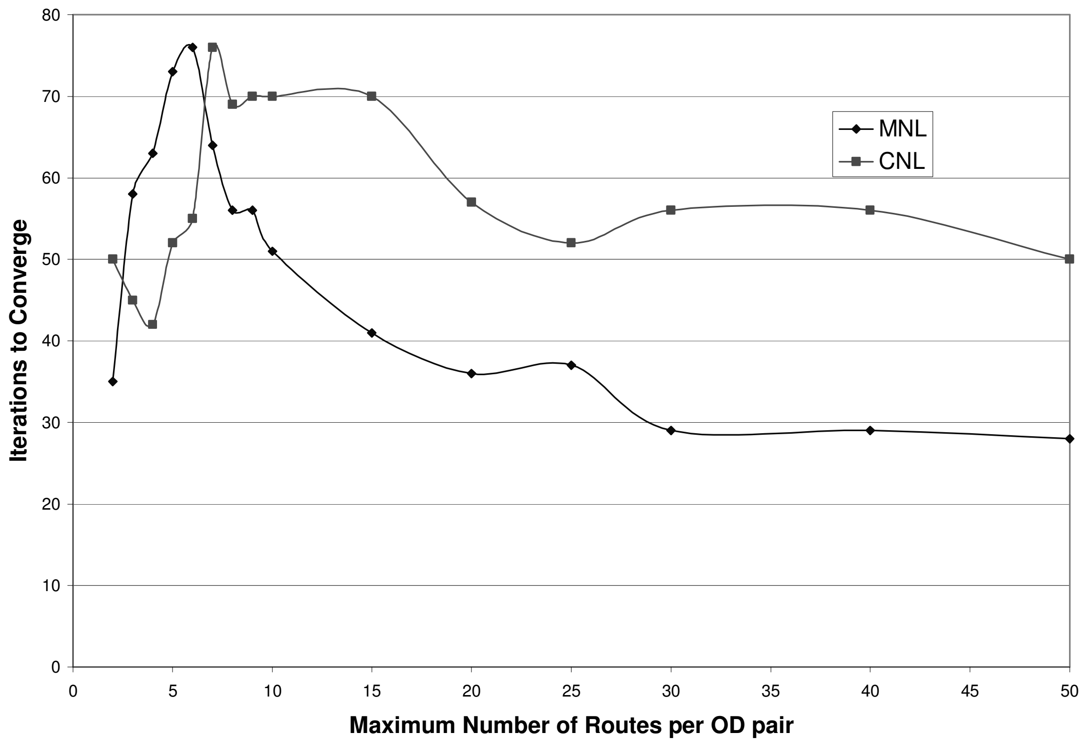

Figure 8. Winnipeg Network – Iterations needed to reach 1% RMSPE of path flows.

The approximately gamma-shaped pattern for both CNL and MNL are similar: when the number of routes in the path set is small (up to 10), more iterations are needed in order to obtain convergence when the path set is increased. When the size of the choice set is further increased the number of iterations required to reach convergence decreases. Interestingly, when few routes are included, the CNL model requires less iterations compared to the MNL. However, for larger choice sets this pattern changes. This result may be explained by the choice set generation method used in both cases, which includes both similar and disjointed routes. Recall that the MNL model is insensitive to similarity between routes, while the CNL takes into account the similarity. This means that the combination of the similarity and congestion effects in the algorithm needs more iterations to converge. If all routes were disjointed, the CNL and MNL model would theoretically need the same number of iterations to converge (since the probability function in this case is the same).

To further study the differences in equilibrium assignments using the MNL and CNL route choice models a comparison of the equilibrium path flows is conducted. The following goodness-of-fit measures were used to quantify the deviations between the results obtained with the two models:

$$RMSE = \sqrt{\frac{1}{K}\sum_{rs}\sum_{k}\left(f_k^{rs(CNL)} - f_k^{rs(MNL)}\right)^2} \tag{3}$$

$$RMSPE_1 = \sqrt{\frac{1}{K}\sum_{rs}\sum_{k}\left(\frac{f_k^{rs(CNL)} - f_k^{rs(MNL)}}{(f_k^{rs(CNL)} + f_k^{rs(MNL)})/2}\right)^2} \tag{4}$$

$$RMSPE_2 = \sqrt{\frac{1}{K}\sum_{rs}\sum_{k}\left(\frac{f_k^{rs(CNL)} - f_k^{rs(MNL)}}{q_{rs} / K_{rs}}\right)^2} \tag{5}$$

Where RMSE is the root mean square error, RMSPE is the root mean square percentage error, K is the total number of routes and K_{rs} is the number of routes for the specific origin-destination pair *rs*. The rationale for using equation (5) is to avoid large contributions to the error by routes which carry a very small fraction of the demand, which may happen when the number of routes in the choice set increases.

Figure 8 presents the results for different choice set sizes. The CNL-SUE model results were computed assuming a nesting coefficient equal to 0.5, similar to the model used in Figure 6. In the present figure, the demand is set equal in both cases, and the differences are computed after reaching convergence level lower than 0.001 (in each model). The left axis in Figure 9 represents RMSPE, and the right axis represents RMSE.

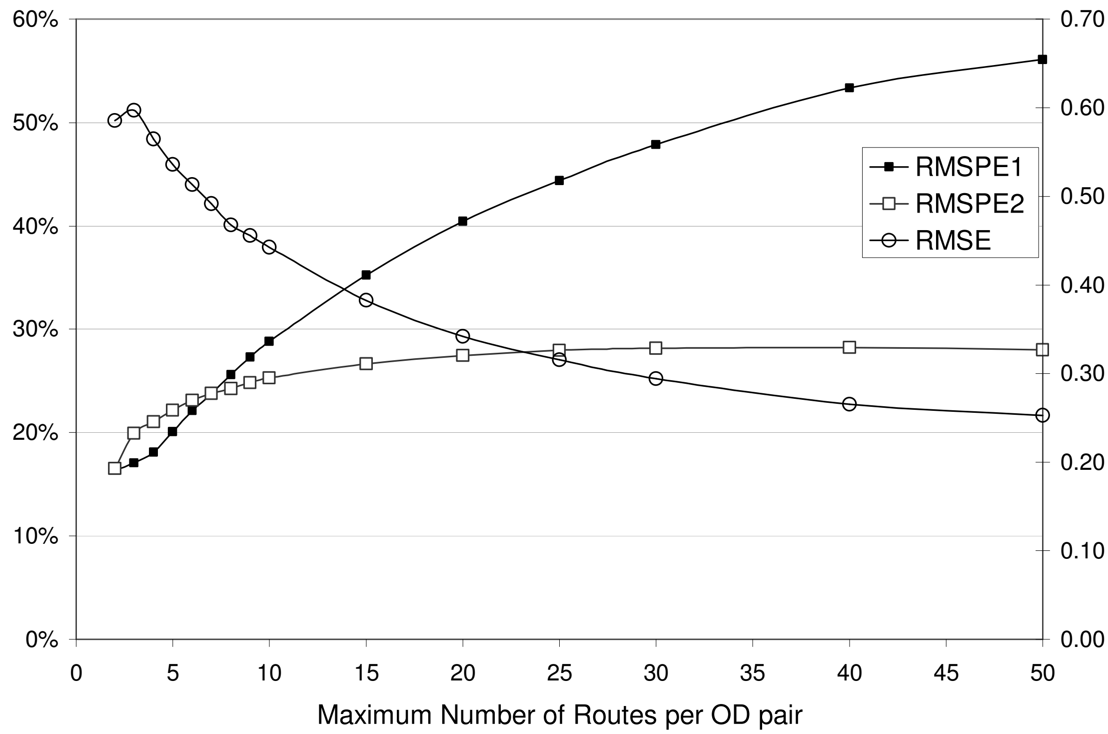

Figure 9. Winnipeg Network – Comparison between CNL and MNL path flows (θ = 0.5 and μ=0.5)

As expected, the RMSE measurement decreases, while the RMSPE measurements increase, relative to the choice set size. To give a numerical interpretation of the results, recall that the average flow for each path in the network for the maximum choice set size (50 routes) is 0.31 (54,459 divided by 174,491). The RMSE is close to 0.25 (flow units), and consequently the relatively high RMSPE1 value found (0.55), and the relatively low RMSPE2 value found (0.28).

Since the fixed routes generated are based on link penalty and link elimination methods, there is a certain degree of similarity among these routes (because of the common links). Therefore, as the number of routes in the choice set increase, the similarity among routes increase. The CNL model accounts for similarity, and consequently there is an increase of the relative difference between CNL and MNL route flows for increasing choice set sizes.

4.3 Network Comparison

Table 3 compares selected MNL-SUE and CNL-SUE results for both networks tested in this paper, assuming θ = 0.5 and μ=0.5 for the CNL model. All results were computed using the same computer platform.

Table 3. Comparison of Selected Network Results (θ = 0.5)

Model	Network Characteristic	Sioux Falls	Winnipeg	Ratio
	Total links	76	2,535	33.4
	Total OD pairs	528	4,345	8.2
	Total generated routes	3,441	174,491	50.7
	Average routes per OD pair	6.5	40.2	6.4
	Average routes per link	45.3	68.8	1.5
	Average OD pair per link	6.9	1.7	0.2
MNL – SUE	Iterations needed to reach 0.001 precision	31	28	0.9
	CPU time (sec)	1.4	9,967	7,119.3
	CPU time per iteration (sec)	0.045	356	7,882.1
CNL – SUE	Iterations needed to reach 0.001 precision	32	50	1.6
	CPU time (sec)	2.8	21,597	7,768.7
	CPU time per iteration (sec)	0.087	432	4,972.0

The CPU time needed to reach the same convergence precision is more than 7,000 times higher in the Winnipeg network, compared to the Sioux Falls network, for both MNL-SUE and CNL-SUE models. The non-linear increase in CPU times may be explained by the increase in network complexity: the Winnipeg network has more routes per OD pair, but fewer routes per link. This means that the Winnipeg network is less dense compared to the Sioux Falls network, and consequently there are more opportunities to swap flows among paths. The 0.001 convergence precision in the Winnipeg network is reached in relatively few iterations because the network is moderately congested, compared to the Sioux Falls network.

5. SUMMARY AND CONCLUSIONS

In practice, route choice models are commonly used as components in large model systems, such as traffic assignment, route guidance, network design, etc. For this reason, the choice models used to capture routing decisions are often very simple, generally consisting of finding the least cost path or the simple MNL model. This paper presents results that illustrate the potential of using more realistic route choice models at affordable resources. This paper analyzed the effect of choice set size and selected choice models on SUE assignment running time and flow patterns.

For small networks, such as the Sioux Falls and the inter-urban network described in Cascetta et al. (1997), few routes are sufficient. This is mainly due to the relatively simple network topology. For larger and more complex networks, such as the Winnipeg network, the results indicate a need to generate a large number of alternative routes.

The problem of choice set generation is still an open problem for SUE models. In the present paper, a heuristic method based on the combination of link penalty and link elimination method was applied for the Winnipeg network. These methods produce both similar and disjointed routes. We agree with previous authors that disjointed routes are essential to increase the convergence properties. However, more research is needed to verify the stability of the results with respect to other choice set generation methods.

REFERENCES

Akamatsu, T. (1996) Cyclic flows, Markov process and stochastic traffic assignment. Transportation Research Part B, 30, 369-386.

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

Bar-Gera, H. and Boyce D. (2005) User-equilibrium route set analysis of a large road network. In H.S. Mahmassani, ed. Transportation and Traffic Theory, Proceedings of the 16th International Symposium on Transportation and Traffic Theory. Oxford: Elsevier Science, pp. 673-692.

Bell, M.G.H. (1995) Alternatives to Dial's Logit assignment algorithm. Transportation Research Part B, 29, 287-295.

Bekhor, S., Ben-Akiva, M.E., and Ramming, M.S. (2002) Adaptation of Logit Kernel to route

choice situation. Transportation Research Record, 1805, 78-85.
Bekhor, S., Ben-Akiva, M.E., and Ramming M.S. (2006) Evaluation of choice set generation algorithms for route choice models. Annals of Operations Research 144, 235-247.
Bekhor, S., and Prashker, J. N. (1999) Formulations of extended logit stochastic user equilibrium assignments. In A. Ceder ed. Transportation and Traffic Theory, Proceedings of the 14th International Symposium on Transportation and Traffic Theory, Pergamon Press, pp. 351-372.
Bekhor, S. and Prashker, J.N. (2001) Stochastic user equilibrium formulation for the generalized nested logit model. Transportation Research Record, 1752, 84-90.
Bekhor, S. and Prato, C.G. (2006) Effects of choice set composition in route choice modeling. 11th International Conference on Travel Behaviour Research, Kyoto, Japan, August 2006.
Bekhor S., and Toledo, T. (2005) Investigation of path-based algorithms for the stochastic user equilibrium problem. Transportation Research Part B, 39, 279-295.
Bekhor, S., Toledo, T., and Reznikova, L. (2007) Application of the cross nested logit route choice model in stochastic user equilibrium traffic assignment. Transportation Research Record, 2003, 41-49.
Ben-Akiva, M.E., Bergman, M.J., Daly, A.J., and Ramaswamy, R. (1984) Modeling inter-urban route choice behaviour. In J. Volmuller and R. Hamerslag, Eds. Proceedings of the Ninth International Symposium on Transportation and Traffic Theory, Utrecht: VNU Science Press, pp. 299-330.
Ben-Akiva, M.E., and Bierlaire, M. (1999) Discrete choice methods and their applications to short term travel decisions. In Randolph W. Hall, ed. Handbook of Transportation Science, Norwell, MA: Kluwer Academic Publishers, Ch. 2.
Bovy, P. and Fiorenzo-Catalano, S. (2007) Stochastic route choice set generation: behavioral and probabilistic foundations. Transportmetrica 3, 173-189.
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.B. Lesort ed. Proceedings of the International Symposium on Transportation and Traffic Theory, Lyon: Elsevier, pp. 697-711.
Chen A., Kasikitwiwat, P., and Ji, Z. (2003) Solving the overlapping problem in route choice with paired combinatorial logit model. Transportation Research Record 1857, 65 – 73.
Chu, C. (1989) A paired combinatorial logit model for travel demand analysis. Proceedings of the Fifth World Conference on Transportation Research, 4, Ventura, CA, pp. 295-309.
Daganzo, C.F., and Sheffi, Y. (1977) On stochastic models of traffic assignment. Transportation Science, 11, 253-274.
Damberg, O., Lundgren, J.T., and Patriksson, M. (1996) An algorithm for the stochastic user equilibrium problem. Transportation Research Part B, 30, 115-131.

De La Barra, T., Perez, B., and Anez, J. (1993) Multidimensional path search and assignment. Proceedings of the 21st PTRC Summer Annual Meeting, Manchester, England, pp. 307-319.
Dial, R.B. (1971) A probabilistic multipath traffic assignment algorithm which obviates path enumeration. Transportation Research, 5, 83-111.
Dial, R.B. (2000) Bicriterion traffic equilibrium: T2 mode, algorithm, software overview. Transportation Research Record, 1725, 54–62.
Dial, R.B. (2001) Equilibrium logit traffic assignment: elementary theory and algorithms. 80th Annual Meeting of the Transportation Research Board, Washington DC, January 2001.
Fiorenzo-Catalano, S., and Van der Zijpp, N.J. (2001) A forecasting model for inland navigation based on route enumeration. Proceedings of the European Transport Conference, PTRC Education and Research Services Ltd., London, pp. 1-11.
Fisk, C. (1980) Some developments in equilibrium traffic assignment. Transportation Research Part B, 14, 243-255.
Friedrich, M., Hofsäß, I, and Wekeck, S. (2001) Timetable-based transit assignment using branch and bound. Transportation Research Record, 1752, 100-107.
Gliebe, J.P., Koppelman, F.S., and Ziliaskopoulos, A. (1999) Route choice using a paired combinatorial logit model. 78th Annual Meeting of the Transportation Research Board, Washington DC, January 1999.
Hoogendoorn-Lanser, S. (2005) Modelling travel behaviour in multi-modal networks. TRAIL Research School, The Netherlands, 2005.
Huang H.-J. (1995) A combined algorithm for solving and calibrating the stochastic user equilibrium assignment model. Journal of the Operational Research Society, 46, 977-987.
Huang H.-J., and Bell, M.G.H. (1998) A study on logit assignment which excludes all cyclic flows. Transportation Research Part B, 32, 401-412.
Larsson, T. and Patriksson, M. (1992) Simplicial decomposition with disaggregated representation for the traffic assignment problem. Transportation Science, 26, 4-17.
Maher, M. (1998) Algorithms for logit-based stochastic user equilibrium assignment. Transportation Research Part B, 32, 539-549.
McFadden, D. (1978) Modeling the choice of residential location. In A. Karlqvist et al. eds. Spatial Interaction Theory and Residential Location, North Holland, Amsterdam, pp. 75-96.
Park, D., and Rilett, L.R. (1997) Identifying multiple and reasonable paths in transportation networks: a heuristic approach. Transportation Research Record, 1607, 31-37.
Powell, W.B. and Sheffi, Y. (1982) The convergence of equilibrium algorithms with predetermined step sizes. Transportation Science, 16, 45-55.
Prashker, J.N., and Bekhor, S. (1998) Investigation of stochastic network loading procedures. Transportation Research Record, 1645, 94-102.
Prashker J. N., and Bekhor, S. (2004) Route choice models used in the stochastic user

equilibrium problem: a review. Transport Reviews 24, 437-463.
Prato, C.G. and Bekhor, S. (2006) Applying branch-and-bound technique to route choice set generation. Transportation Research Record, 1985, 19-28.
Pravinvongvuth, S. and Chen, A. (2005) Adaptation of the paired combinatorial logit model to the route choice problem. Transportmetrica 1, 223-240.
Ramming, M.S. (2001) Network Knowledge and Route Choice. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.
Scott, K., Pabon-Jimenez, G., and Bernstein, D. (1997) Finding alternatives to the best path. 76th Annual Meeting of the Transportation Research Board, Washington, DC, January 1997.
Sheffi, Y., and Powell, W.B. (1982) An algorithm for the equilibrium assignment problem with random link times. Networks, 12, 191-207.
Sheffi, Y. (1985) Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods. Prentice-Hall, Englewood Cliffs, NJ.
Van der Zijpp, N.J. and Fiorenzo-Catalano, S. (2005) Path enumeration by finding the constrained k-shortest paths. Transportation Research Part B, 39, 545-563.
Vovsha, P. (1997). The cross-nested logit model: application to mode choice in the tel-aviv metropolitan area. Transportation Research Record, 1607, 6-15.
Wen, C-H., and Koppelman, F.S. (2001) The generalized nested logit model. Transportation Research Part B, 35, 627-641.
Yai, T., Iwakura, S., and Morichi S. (1997) Multinomial probit with structured covariance for route choice behavior. Transportation Research Part B, 31 195-207.