ASSESSING THE INFLUENCE OF OBJECTIVE CHOICE SETS ON ROUTE CHOICE MODEL ESTIMATES AND PREDICTIONS

Carlo Giacomo PRATO
Research Fellow
Transportation Research Institute
Technion – Israel Institute of Technology
Technion City, Haifa 32000, Israel
Phone: +972.4.8295134, Fax: +972.4. 8225716
e-mail: prato@technion.ac.il

Abstract:
In the context of modeling route choice behavior in transport networks, travelers are presumed to choose the best alternative from a given set of routes connecting origin and destination of their trip. Although most equilibrium assignment techniques implicitly suppose that all existing paths between origin and destination are available to all travelers, path-based assignment methods assume that route choice sets are generated prior to the choice modeling by implementing some path generation technique.

Theoretically, model estimation and prediction rely on either generated objective choice sets, which include relevant and feasible alternatives or a representative subset, or observed subjective choice sets, which contain all the alternatives considered by travelers. Practically, modelers make probabilistic predictions from objective choice sets because subjective choice sets are unobservable. Modelers select a path generation technique, which determines composition and size of the objective choice sets and consequently affects model estimates, model predictions, and thus assignment of traffic flows.

Recently, the impact of choice set composition and size on model performances has received increasing attention. Likelihood values were compared to evaluate model performances at the aggregate level, even though without providing insight into the estimation quality at the alternative level. Choice probabilities were calculated to assess prediction abilities at the individual route level, even though considering an unrealistic universal realm consisting of only 12 alternatives. Objective function values and convergence times were compared for several choice set sizes to measure their influence on convergence properties of stochastic equilibrium assignment methods. These studies mainly focused on the analysis of the robustness of route choice models, rather than on the exam of the effects of path generation techniques, as none actually proposed the implementation of different methods for enumerating paths.

This paper intends to assess the effects of path generation techniques on model estimates and prediction abilities. The proposed experiment relies on the design of a synthetic network, the definition of subjective choice sets, the selection of a postulated model, and the implementation of deterministic and stochastic path generation techniques. Hypotheses about subjective choice sets and postulated model allow defining “true parameters” and observing “actual choices”. Model estimation and prediction considering different objective choice sets allow evaluating the effects on model estimates and prediction values.
of the implementation of different path generation techniques.

The design of the synthetic network defines a 5×5 grid with 25 nodes and 40 links. This network structure enables to control for composition and size of the objective choice sets, while offering a realistically large universal realm of hundreds of alternative routes, in particular considering all the acyclic paths connecting the farthest nodes on the grid. Furthermore, the synthetic network allows enumerating the universal choice set that, combined with the definition of the utility according to the postulated model, enables to evaluate the generated objective choice sets with respect to the ability to produce relevant or irrelevant routes.

The definition of the subjective choice sets hypothesizes four types of decision-makers, specifically travelers who consider (i) few similar alternatives, (ii) several similar paths, (iii) few heterogeneous options, or (iv) several heterogeneous routes. This definition agrees with existing observations of subjective route choice sets, which underlined that decision-makers were different in terms of either number of routes or similarity among the alternatives considered. Moreover, this definition enables the experiment to be unbiased with respect to the composition of the travelers’ population. Assuming homogeneous population (i.e., all travelers considering few heterogeneous paths) could have favored techniques that tend to generate a certain type of alternatives (i.e., heterogeneous alternatives), while existing observations depict a heterogeneous population.

The specification of the postulated model considers a Path Size Correction Logit model for the generation of the observations. For each subjective choice set, the utility of each alternative is calculated as a function of (i) route length, (ii) number of turns, (iii) path size correction, and (iv) an error term distributed Extreme Value with scale equal to one and location equal to zero. Assuming the “true parameters” in the described function, the alternative with the highest utility is associated to the choice to generate 4000 observations for estimation purposes and 1000 for prediction purposes. The postulated model allows accounting for similarities among alternatives while maintaining the simple Logit structure, which enables to accommodate a correction term when instances of importance sampling are implemented (e.g., simulation or random walk). Furthermore, considering one postulated model enables the experiment to be unbiased with respect to the model specification. Considering non-robust models with respect to the choice set size (i.e., CNL and Logit Kernel) could have complicated the discernment of the effects of path generation techniques from the effects of model specifications.

The implementation of path generation techniques allows evaluating their effect on model estimates and prediction abilities. Three deterministic techniques are considered: (i) k-shortest paths; (ii) link penalty approach, which at each iteration penalizes the impedance of the shortest path links; (iii) branch and bound algorithm, which relies on behavioral constraints common to all the travelers’ population. Three stochastic techniques are considered: (iv) simulation approach, which at each iteration calculates the shortest path after extracting links impedances from a random distribution; (v) simulation approach with correction term added to the utility function to account for the unequal probability of selecting routes in the network; (vi) the random walk algorithm, also with the aforementioned correction term.

Results illustrate the bias in model estimates and prediction values when the six path generation techniques are implemented to the synthetic network, and then models are
estimated and outcomes are predicted considering the generated objective choice sets. Results also show the degree of relevance of each technique, by calculating the proportion of relevant and irrelevant routes according to a Monte Carlo simulation based on the postulated model and the designed synthetic network. The relationship between the relevance of the technique and the model performances allows suggesting empirical guidelines to the generation of route choice sets.

**Keywords:**
Path generation techniques, Route choice modeling, Objective choice sets, Subjective choice sets, Path Size Correction Logit.