Technical note: A Stochastic Choice Set Generation Algorithm*

M. Bierlaire E. Frejinger†

February 13, 2007

Report TRANSP-OR 070213
Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne
transp-or.epfl.ch

*This research is supported by the Swiss National Science Foundation grant 200021-107777/1

†École Polytechnique Fédérale de Lausanne, Transport and Mobility Laboratory, CH-1015 Lausanne, Switzerland. E-mail: {michel.bierlaire, emma.frejinger}@epfl.ch

This technical note describes a stochastic choice set generation algorithm that has been used for generating choice sets for data collected in Switzerland. See Vrtic et al. (2006) for project description and Bierlaire et al. (2006) for modeling results.

For a given origin o and destination d, a path is generated in the following way:

1. Selection of a link that the path has to traverse:
   - a probability is associated with each link in the network using the Kumaraswamy distribution (Kumaraswamy, 1980, see discussion below).
   - a link ℓ is randomly selected based on this probability distribution.
2. A path traversing ℓ is generated by computing two shortest paths: one from the origin to the source node i of ℓ ($SP(o,i)$) and one from the sink node j of ℓ to the destination ($SP(j,d)$). The generated path is composed of the segments $SP(o,i)$, ℓ and $SP(j,d)$.
3. The generated path is penalized by updating the costs of its links. The new cost of a link ℓ, $c'(\ell)$, is a function of the previous cost $c(\ell)$ and one additive penalty p_a and a multiplicative penalty p_m:

$$c'(\ell) = p_m c(\ell) + p_a.$$

The probabilities associated with each link in the network are based on the double bounded Kumaraswamy distribution (Kumaraswamy, 1980) whose cumulative distribution function is $F(x_\ell|a,b) = 1-(1-x_\ell{}^a)^b$ for $x_\ell \in [0,1]$. a and b are shape parameters. x_ℓ is defined as

$$x_\ell = \frac{|SP(o,d)|}{|SP(o,i)| + C(\ell) + |SP(j,d)|}$$

where $C(\ell)$ is the cost of link ℓ and $|SP(v_1,v_2)|$ is the cost of the shortest path between nodes v_1 and v_2. The total cost of a path using ℓ, C_ℓ, is the sum in the denominator, that is $C_\ell = |SP(o,i)| + C(\ell) + |SP(j,d)|$. Any

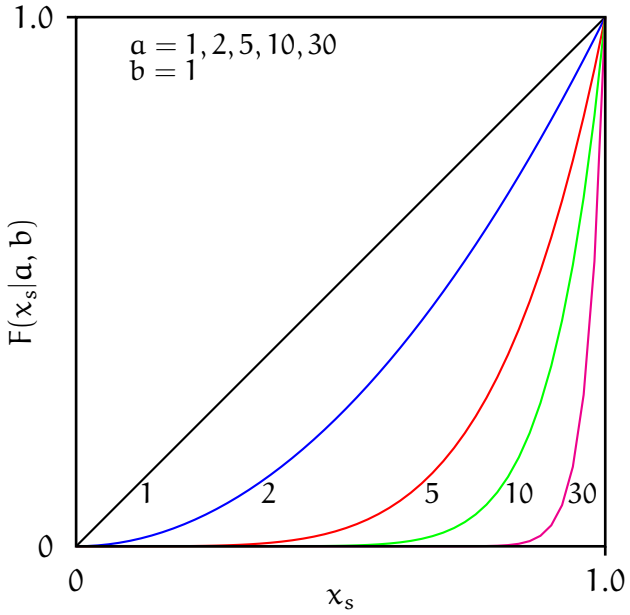

Figure 1: Kumaraswamy distribution - cumulative distribution function

generalized cost can be used in this context. Note that x_ℓ equals one if link ℓ is on the shortest path and $x_\ell \to 0$ as $C_\ell \to \infty$. In Figure 1 we show the cumulative distribution function for different values of a when $b = 1$. The probabilities assigned to the links can be controlled by the definition of the distribution parameters. High values of a when $b = 1$, yield very low probabilities for links with high cost. Low values of a have the opposite effect.

The algorithm is implemented in BioRoute (a route choice modeling tool that generates the input files for BIOGEME), see Bierlaire (2003) and Bierlaire (2005). The BioRoute input file has the following (additional) sections.

- `[GeneralizedCost]`
  The name of the attribute or the definition of the function of attributes that should be minimized in the shortest path computations.

- `[GeneratePaths]`
  Defines the algorithm to be used for the choice set generation. Valid

entries are `$linkElimination` and `$rawPath` (the latter corresponds to this algorithm).

- `[RawPathParameters]`
  This section contains the parameters for the algorithm. Several variants of the algorithms have been implemented where the one described here corresponds to one possibility. The relevant parameters are (with the values used for the Swiss dataset):
  ```
  numberOfThroughLinkPaths = 15
  maxNumberOfIterations = 2
  kumaA = 10.0
  kumaB = 1.0
  multiplicativePenalty = 1.1
  additivePenalty = 0.0
  maxNumberOfGeneratedPaths = 45
  ```
  The other parameters have values so that only the algorithm described here is used:
  ```
  numberOfTrials = 1
  randomWalkDepth = 0
  maxNumberOfGeneratedPathsPerIteration = 0
  ```

References

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

Bierlaire, M. (2005). An introduction to biogeme version 1.4. http://biogeme.epfl.ch.

Bierlaire, M., Frejinger, E. and Stojanovic, J. (2006). A latent route choice model in switzerland, *Proceedings of the European Transport Conference*, Strasbourg, France.

Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes, *Journal of Hydrology* **46**: 79–88.

Vrtic, M., Schüssler, N., Erath, A., Axhausen, K., Frejinger, E., Bierlaire, M., Stojanovic, S., Rudel, R. and Maggi, R. (2006). Einbezug von reisekosten bei der modellierung des mobilittsverhalten. Final report for SVI research program Mobility Pricing: Project B1, on behalf of the Swiss Federal Department of the Environment, Transport, Energy and Communications, IVT ETH Zurich, ROSO EPF Lausanne and USI Lugano.