Route choice models: Introduction and recent developments

Michel Bierlaire

transp-or.epfl.ch

Transport and Mobility Laboratory, EPFL
Route choice model

Given

- a mono- or multi-modal transportation network (nodes, links, origin, destination)
- an origin-destination pair
- link and path attributes

identify the route that a traveler would select.
Choice model

Assumptions about

1. the decision-maker: \( n \)
2. the alternatives
   - Choice set \( C_n \)
   - \( p \in C_n \) is composed of a list of links \((i, j)\)
3. the attributes
   - link-additive: length, travel time, etc.
   - non link-additive: scenic path, usual path, etc.
4. the decision-rules: \( \Pr(p|C_n) \)
Shortest path

Decision-makers  all identical

Alternatives

- all paths between O and D
- $C_n = \mathcal{U} \quad \forall n$
- $\mathcal{U}$ can be unbounded when loops are present

Attributes  one link additive generalized cost

\[
c_p = \sum_{(i,j) \in P} c(i,j)
\]

- traveler independent
- link cost may be negative
- no loop with negative cost must be present so that $c_p > -\infty$ for all $p$
Shortest path

**Decision-rules** path with the minimum cost is selected

\[
\Pr(p) = \begin{cases} 
K & \text{if } c_p \leq c_q \quad \forall c_q \in \mathcal{U} \\
0 & \text{otherwise}
\end{cases}
\]

- \( K \) is a normalizing constant so that \( \sum_{p \in \mathcal{U}} \Pr(p) = 1 \).
- \( K = 1/S \), where \( S \) is the number of shortest paths between \( O \) and \( D \).
- Some methods select one shortest path \( p^* \)

\[
\Pr(p) = \begin{cases} 
1 & \text{if } p = p^* \\
0 & \text{otherwise}
\end{cases}
\]
Shortest path

Advantages:
- well defined
- no need for behavioral data
- efficient algorithms (Dijkstra)

Disadvantages
- behaviorally unrealistic
- instability with respect to variations in cost
- calibration on real data is very difficult
  - inverse shortest path problem is NP complete

Dial’s approach


**Decision-makers** all identical

**Alternatives** efficient paths between O and D

**Attributes** link-additive generalized cost

**Decision-rules** multinomial logit model
Dial’s approach

- Def 1: A path is efficient if every link in it has
  - its initial node closer to the origin than its final node, and
  - its final node closer to the destination than its initial node.
- Def 2: A path is efficient if every link in it has its initial node closer to the origin than its final node.

Efficient path: a path that does not backtrack.
Dial’s approach

- Choice set $C_n = \text{set of efficient paths (finite, no loop)}$
- No explicit enumeration
- Every efficient path has a non zero probability to be selected
- Probability to select a path

$$
Pr(p) = \frac{e^{\theta \left( \sum_{(i,j) \in p^*} c(i,j) - \sum_{(i,j) \in p} c(i,j) \right)}}{\sum_{q \in C_n} e^{\theta \left( \sum_{(i,j) \in p^*} c(i,j) - \sum_{(i,j) \in p} q(i,j) \right)}}
$$

where $p^*$ is the shortest path and $\theta$ is a parameter
Dial’s approach

Note: the length of the shortest path is constant across \( C_n \)

\[
\Pr(p) = \frac{e^{-\theta \sum_{(i,j) \in p} c(i,j)}}{\sum_{q \in C_n} e^{-\theta \sum_{(i,j) \in q} q(i,j)}} = \frac{e^{-\theta c_p}}{\sum_{q \in C_n} e^{-\theta c_q}}
\]

Multinomial logit model with

\[
V_p = -\theta c_p
\]
Dial’s approach

Advantages:

- probabilistic model, more stable
- calibration parameter $\theta$
- avoid path enumeration
- designed for traffic assignment

Disadvantages:

- MNL assumes independence among alternatives
- efficient paths are mathematically convenient but not behaviorally motivated
Dial’s approach

Path 1: $c$

$\Pr(1) = \frac{e^{-\theta c_1}}{\sum_{q \in C} e^{-\theta c_q}} = \frac{e^{-\theta c}}{3e^{-\theta c}} = \frac{1}{3}$ for any $c, \delta, \theta$
Path Size Logit

- With MNL, the utility of overlapping paths is overestimated.
- When $\delta$ is large, there is some sort of “double counting.”
- Idea: include a correction

$$V_p = -\theta c_p + \beta \ln PS_p$$

where

$$PS_p = \sum_{(i, j) \in p} \frac{c(i, j)}{c_p} \frac{1}{\sum_{q \in C} \delta_{i, j}^q}$$

and

$$\delta_{i, j}^q = \begin{cases} 1 & \text{if link } (i, j) \text{ belongs to path } q \\ 0 & \text{otherwise} \end{cases}$$
Path Size Logit

Path 1: \( c \)

\[
PS_1 = \frac{c}{c} = 1
\]

\[
PS_2 = PS_3 = \frac{c-\delta}{c} \frac{1}{2} + \frac{\delta}{c} = \frac{1}{2} + \frac{\delta}{2c}
\]
Path Size Logit

Route choice models: Introduction and recent developments – p.16/40
Path Size Logit

Advantages:
- MNL formulation: simple
- Easy to compute
- Exploits the network topology
- Practical

Disadvantages:
- Derived from the theory on nested logit
- Several formulations have been proposed
- Correlated with observed and unobserved attributes
- May give biased estimates
Path Size Logit: readings


Path Size Logit: readings


Path enumeration

- Dial’s approach avoids path enumeration
- Computationally convenient but behaviorally incorrect
- MNL inappropriate due to significant path overlap
- Generalized cost must be link-additive
- Heterogeneity in terms of behavior, equipments, etc. cannot be accounted for.
- With other DCM models, choice sets must be explicitly defined
- Path enumeration heuristics have been proposed:
  - Deterministic approaches: link elimination (Azevedo et al., 1993), labeled paths (Ben-Akiva et al., 1984)
  - Stochastic approaches: simulation (Ramming, 2001) and doubly stochastic (Bovy and Fiorenzo-Catalano, 2006)
Path enumeration

- Underlying assumption in existing approaches: the actual choice set is generated
- Empirical results suggest that this is not always true
- Our approach:
  - Choice set contains all paths
  - Too large for computation
  - Solution: sampling of alternatives
Sampling of Alternatives

- Multinomial Logit model (e.g. Ben-Akiva and Lerman, 1985):

\[
P(i|C_n) = \frac{q(C_n|i)P(i)}{\sum_{j \in C_n} q(C_n|j)P(j)} = \frac{e^{V_{in} + \ln q(C_n|i)}}{\sum_{j \in C_n} e^{V_{jn} + \ln q(C_n|j)}}
\]

\(C_n\): set of sampled alternatives

\(q(C_n|j)\): probability of sampling \(C_n\) given that \(j\) is the chosen alternative

- If purely random sampling, \(q(C_n|i) = q(C_n|j)\) and

\[
P(i|C_n) = \frac{e^{V_{in} + \ln q(C_n|i)}}{\sum_{j \in C_n} e^{V_{jn} + \ln q(C_n|j)}} = \frac{e^{V_{in}}}{\sum_{j \in C_n} e^{V_{jn}}}
\]
Importance Sampling of Alternatives

- Attractive paths have higher probability of being sampled than unattractive paths
- In this case, $q(C_n|i) \neq q(C_n|j)$

$$P(i|C_n) = \frac{e^{V_{in} + \ln q(C_n|i)}}{\sum_{j \in C_n} e^{V_{jn} + \ln q(C_n|j)}} \neq \frac{e^{V_{in}}}{\sum_{j \in C_n} e^{V_{jn}}}$$

- Path utilities must be corrected in order to obtain unbiased estimation results
Stochastic Path Enumeration

- Key feature: we must be able to compute $q(C_n|i)$
- One possible idea: a biased random walk between $s_o$ and $s_d$ which selects the next link at each node $v$.
- Initialize: $v = s_o$
- Step 1: associate a weight with each outgoing link $\ell = (v, w)$:

$$\omega(\ell|b_1) = 1 - (1 - x_\ell^{b_1})$$

where

$$x_\ell = \frac{SP(v, s_d)}{C(\ell) + SP(w, s_d)},$$

is 1 if $\ell$ is on the shortest path, and decreases when $\ell$ is far from the shortest path.
Stochastic Path Enumeration

$$b_1 = 1, 2, 5, 10, 30$$
Stochastic Path Enumeration

- Step 2: normalize the weights to obtain a probability distribution

\[
q(\ell | \mathcal{E}_v, b_1) = \frac{\omega(\ell | b_1, b_2)}{\sum_{m \in \mathcal{E}_v} \omega(m | b_1)}
\]

- Random draw a link \((v, w^*)\) based on this distribution and add it to the current path

- If \(w^* = s_d\), stop. Else, set \(v = w^*\) and go to step 1.

Probability of generating a path \(j\):

\[
q(j) = \prod_{\ell \in \Gamma_j} q(\ell | \mathcal{E}_v, b_1).
\]
Sampling of Alternatives

• Following Ben-Akiva (1993)
• Sampling protocol
  1. A set $\tilde{C}_n$ is generated by drawing $R$ paths with replacement from the universal set of paths $U$
  2. Add chosen path to $\tilde{C}_n$
• Outcome of sampling: $(\tilde{k}_1, \tilde{k}_2, \ldots, \tilde{k}_J)$ and $\sum_{j=1}^{J} \tilde{k}_j = R$

\[
P(\tilde{k}_1, \tilde{k}_2, \ldots, \tilde{k}_J) = \frac{R!}{\prod_{j \in U} \tilde{k}_j!} \prod_{j \in U} q(j)^{\tilde{k}_j}
\]

• Alternative $j$ appears $k_j = \tilde{k}_j + \delta_{cj}$ in $\tilde{C}_n$
Sampling of Alternatives

- Let $C_n = \{ j \in \mathcal{U} \mid k_j > 0 \}$

$$q(C_n | i) = q(\tilde{C}_n | i) = \frac{R!}{(k_i - 1)!} \prod_{j \in C_n, j \neq i} k_j ! q(i)^{k_i - 1} \prod_{j \in C_n} q(j)^{k_j} = K_{C_n} \frac{k_i}{q(i)}$$

$$K_{C_n} = \frac{R!}{\prod_{j \in C_n} k_j !} \prod_{j \in C_n} q(j)^{k_j}$$

$$P(i | C_n) = \frac{e^{V_i + \ln \left( \frac{k_i}{q(i)} \right)}}{\sum_{j \in C_n} e^{V_j + \ln \left( \frac{k_j}{q(j)} \right)}}$$
Numerical Results

- Estimation of models based on synthetic data generated with a postulated model
- Evaluation of
  - Sampling correction
  - Path Size attribute
  - Biased random walk algorithm parameters
Numerical Results
Numerical Results

- True model: Path Size Logit
  \[ U_j = \beta_{PS} \ln PS_U^j + \beta_L \text{Length}_j + \beta_{SB} \text{SpeedBumps}_j + \varepsilon_j \]
  \[ \beta_{PS} = 1, \beta_L = -0.3, \beta_{SB} = -0.1 \]
  \( \varepsilon_j \) distributed Extreme Value with scale 1 and location 0

- \( PS_U^j = \sum_{\ell \in \Gamma_j} \frac{L_\ell}{L_j} \sum_{p \in U} \frac{1}{\delta_{\ell p}} \)

- 3000 observations
Numerical Results

- Four model specifications

<table>
<thead>
<tr>
<th>Path Size</th>
<th>Sampling</th>
<th>Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Without</td>
<td>With</td>
</tr>
<tr>
<td>$\mathcal{C}$</td>
<td>$M_{PS(\mathcal{C})}^{NoCorr}$</td>
<td>$M_{PS(\mathcal{C})}^{Corr}$</td>
</tr>
<tr>
<td>$\mathcal{U}$</td>
<td>$M_{PS(\mathcal{U})}^{NoCorr}$</td>
<td>$M_{PS(\mathcal{U})}^{Corr}$</td>
</tr>
</tbody>
</table>

$PS_{i}^{\mathcal{U}} = \sum_{\ell \in \Gamma_i} \frac{L_{\ell}}{L_i} \frac{1}{\sum_{j \in \mathcal{U}} \delta_{\ell j}}$

$PS_{in}^{\mathcal{C}} = \sum_{\ell \in \Gamma_i} \frac{L_{\ell}}{L_i} \frac{1}{\sum_{j \in \mathcal{C}_n} \delta_{\ell j}}$
Numerical Results

- Model $M_{PS(C)}^{NoCorr}$:
  $$V_{in} = \mu \left( \beta_{PS} \ln PS^C_{in} - 0.3 \text{Length}_i + \beta_{SB} \text{SpeedBumps}_i \right)$$

- Model $M_{PS(C)}^{Corr}$:
  $$V_{in} = \mu \left( \beta_{PS} \ln PS^C_{in} - 0.3 \text{Length}_i + \beta_{SB} \text{SpeedBumps}_i \right) + \ln \left( \frac{k_i}{q(i)} \right)$$

- Model $M_{PS(U)}^{NoCorr}$:
  $$V_{in} = \mu \left( \beta_{PS} \ln PS^U_{in} - 0.3 \text{Length}_i + \beta_{SB} \text{SpeedBumps}_i \right)$$

- Model $M_{PS(U)}^{Corr}$:
  $$V_{in} = \mu \left( \beta_{PS} \ln PS^U_{in} - 0.3 \text{Length}_i + \beta_{SB} \text{SpeedBumps}_i \right) + \ln \left( \frac{k_i}{q(i)} \right)$$
## Numerical Results

<table>
<thead>
<tr>
<th></th>
<th>True PSL</th>
<th>$M_{PS(C)}^{NoCorr}$ PSL</th>
<th>$M_{PS(C)}^{Corr}$ PSL</th>
<th>$M_{PS(U)}^{NoCorr}$ PSL</th>
<th>$M_{PS(U)}^{Corr}$ PSL</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_L$ fixed</td>
<td>-0.3</td>
<td>-0.3</td>
<td>-0.3</td>
<td>-0.3</td>
<td>-0.3</td>
</tr>
<tr>
<td>$\hat{\mu}$</td>
<td>1</td>
<td>0.182</td>
<td>0.923</td>
<td>0.141</td>
<td>0.977</td>
</tr>
<tr>
<td>standard error</td>
<td></td>
<td>0.0277</td>
<td>0.0246</td>
<td>0.0263</td>
<td>0.0254</td>
</tr>
<tr>
<td>$t$-test w.r.t. 1</td>
<td></td>
<td>-29.54</td>
<td>-3.13</td>
<td>-32.64</td>
<td>-0.91</td>
</tr>
<tr>
<td>$\hat{\beta}_{PS}$</td>
<td>1</td>
<td>1.94</td>
<td>0.308</td>
<td>-1.02</td>
<td>1.02</td>
</tr>
<tr>
<td>standard error</td>
<td></td>
<td>0.428</td>
<td>0.0736</td>
<td>0.383</td>
<td>0.0539</td>
</tr>
<tr>
<td>$t$-test w.r.t. 1</td>
<td></td>
<td>2.20</td>
<td>-9.40</td>
<td>-5.27</td>
<td>0.37</td>
</tr>
<tr>
<td>$\hat{\beta}_{SB}$</td>
<td>-0.1</td>
<td>-1.91</td>
<td>-0.139</td>
<td>-2.82</td>
<td>-0.0951</td>
</tr>
<tr>
<td>standard error</td>
<td></td>
<td>0.25</td>
<td>0.0232</td>
<td>0.428</td>
<td>0.024</td>
</tr>
<tr>
<td>$t$-test w.r.t. -0.1</td>
<td></td>
<td>-7.24</td>
<td>-1.68</td>
<td>-6.36</td>
<td>0.20</td>
</tr>
</tbody>
</table>
### Numerical Results

<table>
<thead>
<tr>
<th></th>
<th>True PSL</th>
<th>$M_{PS(C)}^{NoCorr}$ PSL</th>
<th>$M_{PS(C)}^{Corr}$ PSL</th>
<th>$M_{PS(U)}^{NoCorr}$ PSL</th>
<th>$M_{PS(U)}^{Corr}$ PSL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final log likelihood</td>
<td></td>
<td>-6660.45</td>
<td>-6147.79</td>
<td>-6666.82</td>
<td>-5933.62</td>
</tr>
<tr>
<td>Adj. rho-square</td>
<td>0.018</td>
<td>0.093</td>
<td>0.017</td>
<td>0.125</td>
<td></td>
</tr>
</tbody>
</table>

Null log likelihood: -6784.96, 3000 observations

Algorithm parameters: 10 draws, $b_1 = 5$, $b_2 = 1$, $C(\ell) = L_\ell$

Average size of sampled choice sets: 9.66

BIOGEME (Bierlaire, 2007 and Bierlaire, 2003) has been used for all model estimations
Extended Path Size

- Compute Path Size attribute based on an extended choice set $C_{n}^{extended}$
- Simple random draws from $\mathcal{U} \setminus C_n$ so that $|C_n| \leq |C_{n}^{extended}| \leq |\mathcal{U}|$
Extended Path Size

![Graph showing Extended Path Size](image)

- **Path Size**
- **Speed Bump**
- **Scale Parameter**

$t$-test w.r.t. true value vs. Average number of paths in $C_{n}^{\text{extended}}$
Extended Path Size

- Assume that the true choice set is the set of all paths
- Draw a subset for estimating the choice probability
- Draw a larger subset to compute the path size
- Various heuristics based on the same definition of the link weights can be used
Conclusions

• New point of view on choice set generation and route choice modeling
• Path generation is considered an importance sampling approach
• We present a path generation algorithm and derive the corresponding sampling correction
• Path Size should be computed on largest possible sets
• Numerical results are very promising
Readings

  http://library.epfl.ch/theses/?nr=4009

  Transport and Mobility Laboratory, ENAC, EPFL.