Application of the MRI framework to a large network: Québec city

Evanthia Kazagli
Michel Bierlaire

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
École Polytechnique Fédérale de Lausanne

June 22, 2017
1. Preamble
2. Modeling
3. Results
4. Conclusion
Agenda

1. Preamble
2. Modeling
3. Results
4. Conclusion
Identify the route that a traveler would choose to go from the origin (O) to the destination (D).

- Key travel demand model.
- At the core of traffic assignment.
- Off-line and real time services and applications:
  - Decision-aid tools and transportation policies.
  - Real time operations and *route guidance*.

→ **Random utility models**
  - Understand, describe and predict route choice behavior.
Towards aggregate route choice

How can we represent a route in a behaviorally realistic way without increasing the model complexity?

→ Model the strategic decisions of people.

★ Mental Representation Item (MRI)

A trivial example of a MRI model
Objective

Application of the MRI framework to a large network.

1. Additional complexity in the definition of the model due to the size of the city of interest.

2. Lack of a detailed disaggregate network model.
Conceptual model that is realistic and meaningful.

1. Operationalization using simple techniques.
   - Definition and operationalization of the aggregate graph.
   - Concrete specifications: compatible with the standard estimation procedures.

2. Application to aggregate route choice analysis.
   - Québec city.
   - Prediction of flows on the major segments of the network.
Agenda

1. Preamble
2. Modeling
3. Results
4. Conclusion
Considerations

Operational aspects of the aggregate model

1. Path-based specification of utilities
   1. Representative points $\rightarrow$ representative paths.
   2. Sensitivity of model parameters.

2. “Path-free” approach
   1. Elimination of representative points/paths.
   2. Less dependent on detailed network data.
   3. Low structural model complexity and computational times.
Dataset

- Québec city.
- **Montrajet** smartphone application (McGill university)\(^1\),\(^2\)
- GPS trajectories of more around 4000 individuals.
- More than 20000 trips.

- Trip purpose.
- Departure time.

\(^1\) Mirando-Moreno L.F., Chung C., Amyot D., Chappon H. (2014), A system for collecting and mapping traffic congestion in a network using GPS smartphones from regular drivers.

Québec city

Origins and destinations in the data sample (2321 trips)
Observed trajectories
Most visited segments
Modeling

Mobility vs accessibility

Schematic of a Portion of an Urban Street Network

Legend
- Arterial street
- Collector street
- Commercial
- Public

Source: Grant Benjamin, “Grand Reductions: 10 Diagrams That Changed City Planning”, The Urbanist, Issue 518, November 2012, SPUR Ideas + Action for a Better City
Geographical span and \textit{OD} zones
The $G^M$ of Québec city as a dual graph
Québec city: upper side

*Dual graph*
Québec city: lower side

*Dual graph*
## The MRIs attributes

<table>
<thead>
<tr>
<th>MRI</th>
<th>travel time [min]</th>
<th>motorway</th>
<th>primary</th>
<th># of connections $c_m$</th>
<th>measure of entropy $\log_2(c_m)$</th>
<th>measure of entropy $\log_2(c_m - 1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>40E</td>
<td>1.83</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>1.58</td>
<td>1.00</td>
</tr>
<tr>
<td>40W</td>
<td>2.87</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>2.00</td>
<td>1.58</td>
</tr>
<tr>
<td>73N</td>
<td>3.53</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>1.58</td>
<td>1.00</td>
</tr>
<tr>
<td>73S_Upper</td>
<td>2.04</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>3.00</td>
<td>2.81</td>
</tr>
<tr>
<td>4073_40</td>
<td>2.15</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>2.58</td>
<td>2.32</td>
</tr>
<tr>
<td>4073_73</td>
<td>1.46</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>2.81</td>
<td>2.58</td>
</tr>
<tr>
<td>573</td>
<td>2.86</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>440E</td>
<td>1.78</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>2.00</td>
<td>1.58</td>
</tr>
<tr>
<td>440M</td>
<td>1.38</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>2.00</td>
<td>1.58</td>
</tr>
<tr>
<td>440W</td>
<td>1.22</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>2.32</td>
<td>2.00</td>
</tr>
<tr>
<td>740</td>
<td>2.43</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>2.58</td>
<td>2.32</td>
</tr>
<tr>
<td>540</td>
<td>1.89</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>2.81</td>
<td>2.58</td>
</tr>
<tr>
<td>973</td>
<td>2.54</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>3.17</td>
<td>3.00</td>
</tr>
<tr>
<td>175N</td>
<td>5.46</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>2.81</td>
<td>2.58</td>
</tr>
<tr>
<td>138E</td>
<td>3.17</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>138M_E</td>
<td>3.15</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>2.00</td>
<td>1.58</td>
</tr>
<tr>
<td>138M_W</td>
<td>2.63</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>2.00</td>
<td>1.58</td>
</tr>
<tr>
<td>138W</td>
<td>4.49</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1.58</td>
<td>1.00</td>
</tr>
<tr>
<td>136</td>
<td>5.31</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>2.58</td>
<td>2.32</td>
</tr>
<tr>
<td>NB</td>
<td>1.55</td>
<td>1</td>
<td>0</td>
<td>11</td>
<td>3.46</td>
<td>3.32</td>
</tr>
<tr>
<td>OB</td>
<td>2.38</td>
<td>0</td>
<td>1</td>
<td>11</td>
<td>3.46</td>
<td>3.32</td>
</tr>
<tr>
<td>20E</td>
<td>3.52</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>20M</td>
<td>1.39</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>3.17</td>
<td>3.00</td>
</tr>
<tr>
<td>20W</td>
<td>3.65</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>2.58</td>
<td>2.32</td>
</tr>
<tr>
<td>132E</td>
<td>7.06</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>132M</td>
<td>3.78</td>
<td>0</td>
<td>1</td>
<td>9</td>
<td>3.17</td>
<td>3.00</td>
</tr>
<tr>
<td>132W</td>
<td>4.77</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>2.58</td>
<td>2.32</td>
</tr>
<tr>
<td>73S_Lower</td>
<td>3.86</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>3.00</td>
<td>2.81</td>
</tr>
<tr>
<td>175S</td>
<td>4.60</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>3.00</td>
<td>2.81</td>
</tr>
<tr>
<td>116</td>
<td>6.81</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>2.58</td>
<td>2.32</td>
</tr>
</tbody>
</table>
Illustration of a simple \((m_1)\) and a complex \((m_2)\) node

Estimation approach: link-based formulation

**Recursive Logit (RL)** (Fosgerau et al., 2013)

1. Sequential link choice in a dynamic framework.
2. Consistently and efficiently estimated on the full choice set of paths without sampling of alternatives.
3. Equivalent to a multinomial logit.
4. Directly applicable to compute link flows.
Overview of the RL model

A state corresponds to a node $m$ in $G^M$

1. At each state $m$ the traveler chooses the next state $\nu$ that maximizes the sum of the instantaneous utility $u_n(\nu \mid m)$ and the expected downstream utility $V^D'(m)$ to the destination $D$, where:
   1. $u_n(\nu \mid m) = v_n(\nu \mid m) + \mu \varepsilon_n(\nu)$, and
   2. $V^D'(m) = \mu \ln \sum_{\nu \in \mathcal{M}} \delta(\nu \mid m) e^{\frac{1}{\mu}(v_n(\nu \mid m) + V^D'(\nu))} \forall m \in \mathcal{M}$.

2. Output: destination specific state transition probabilities $P^D'(\nu \mid m)$.

3. Choice probabilities given by the logit model.

4. MRI element flows: $F(\nu) = G(\nu) + \sum_{m \in \mathcal{M}} P^D'(\nu \mid m) \cdot F^D'$, computed by solving $(I - P^T)F = G$. 
Observed trajectory and its mapping on $G^M$
Observed number of links in $G$ vs number of nodes in $G^M$
Agenda

1. Preamble
2. Modeling
3. Results
4. Conclusion
Model specifications

\[ v(\nu \mid m) = \beta_{time} \text{Time}(\nu) + \beta_{TransferTime} \text{TransferTime}(\nu \mid m). \quad (1) \]

\[ v(\nu \mid m) = \beta_{time, Motor} \text{Time}(\nu) \text{Motorway}(\nu) + \beta_{time, Prim} \text{Time}(\nu) \text{Primary}(\nu) + \]
\[ + \beta_{TransferTime} \text{TransferTime}(\nu \mid m). \quad (2) \]

\[ v(\nu \mid m) = \beta_{time, Motor} \text{Time}(\nu) \text{Motorway}(\nu) + \beta_{time, Prim} \text{Time}(\nu) \text{Primary}(\nu) + \]
\[ + \beta_{TransferTime} \text{TransferTime}(\nu \mid m) + \beta_{Load} \text{CognitiveLoad}(\nu). \quad (3) \]

\[ v(\nu \mid m) = \beta_{time, Motor, Off} \text{Time}(\nu) \text{Motorway}(\nu) \text{OffPeak}_n + \]
\[ + \beta_{time, Motor, Peak} \text{Time}(\nu) \text{Motorway}(\nu) \text{Peak}_n + \]
\[ + \beta_{time, Prim, Off} \text{Time}(\nu) \text{Primary}(\nu) \text{OffPeak}_n + \]
\[ + \beta_{time, Prim, Peak} \text{Time}(\nu) \text{Primary}(\nu) \text{Peak}_n + \]
\[ + \beta_{TransferTime} \text{TransferTime}(\nu \mid m) + \]
\[ + \beta_{Load} \text{CognitiveLoad}(\nu). \quad (4) \]
Distribution of departure time in the data sample
### Estimation results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>value (t-test)</td>
<td>value (t-test)</td>
<td>value (t-test)</td>
<td>value (t-test)</td>
</tr>
<tr>
<td>$\beta_{\text{TransferTime}}$</td>
<td>$-0.56 (-72.17)$</td>
<td>$-0.58 (-78.62)$</td>
<td>$-0.54 (-72.04)$</td>
<td>$-0.54 (-72.10)$</td>
</tr>
<tr>
<td>$\beta_{\text{time}}$</td>
<td>$-1.12 (-152.25)$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$\beta_{\text{time, Motor}}$</td>
<td>$\times$</td>
<td>$-1.24 (-148.41)$</td>
<td>$-0.50 (-28.94)$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$\beta_{\text{time, Prim}}$</td>
<td>$\times$</td>
<td>$-0.85 (-75.39)$</td>
<td>$-0.52 (-38.37)$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$\beta_{\text{time, Motor, OffPeak}}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$-0.70 (-40.68)$</td>
</tr>
<tr>
<td>$\beta_{\text{time, Motor, Peak}}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$-0.48 (-25.60)$</td>
</tr>
<tr>
<td>$\beta_{\text{time, Prim, OffPeak}}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$-0.51 (-27.78)$</td>
</tr>
<tr>
<td>$\beta_{\text{time, Prim, Peak}}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$-0.54 (-25.28)$</td>
</tr>
<tr>
<td>$\beta_{\text{Load}}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$-0.70 (-40.65)$</td>
</tr>
</tbody>
</table>

| # observations | 2321 | 2321 | 2321 | 2321 |
| # parameters   | 2    | 3    | 4    | 6    |
| $\mathcal{L}(\hat{\beta})$ | $-8145.1$ | $-7593.6$ | $-6935.0$ | $-6930.7$ |
| Estimation time (min) | 2.5  | 3.1  | 4.0  | 7.2  |
Forecasted choice probabilities (Model 4)

20% <= 0.01, 10% <= 0.001

The graph shows the frequency distribution of forecasted choice probabilities. The number of observations is 2321.
State choice probabilities (Model 4)

# of observed states 8755

State choice probability range: 0 to 1

Frequency range: 0 to 700

Kazagi (TRANSP-OR, EPFL)
DCA workshop 2017
June 22, 2017 28 / 37
Model validation

Final log likelihood of the training samples
Model validation

*Predicted log likelihood of the test samples*
Log likelihood loss

Avg. error over the samples \( \overline{\text{err}}_t = \frac{1}{t} \sum_{s=1}^{t} \left[ \frac{1}{|S_s|} \sum_{i \in S_s} \ln P(i, \beta_s) \right] \quad \forall \ t = 1, \ldots, 100 \)
Model application

Aggregate element flows
Model application

Boxplot of aggregate element flows over the samples

Flow

Element

Observed
Full estimation
Agenda

1. Preamble
2. Modeling
3. Results
4. Conclusion
Conclusion

1. Generalization of the applicability of the MRI framework.
2. Operational challenges related to the specification of the utility functions are tackled.
3. Not as simple as the first model, yet of much lower structural complexity in comparison with the disaggregate approach.
4. Insightful understanding and description of the aggregate route choice of individuals.
5. Readily applied to the prediction of flows on the major segments of the network.

* Compatible with route guidance.
* Motivates and can benefit from new data collection approaches.
Future work

1. Application to more cities.
2. Application to pedestrian route choice in major hubs.
3. Investigate extensions to macroscopic traffic assignment.
“In the process of way-finding, the strategic link is the environmental image, the generalized mental picture of the exterior physical world that is held by the individual. This image is the product both of immediate sensation and of the memory of past experience, and it is used to interpret information and to guide action. The need to recognize and pattern our surroundings is so crucial, and has such long roots in the past, that this image has wide practical and emotional importance to the individual.”
Thank you!

evanthia.kazagli@epfl.ch
transp-or.epfl.ch