Activity choice modeling for pedestrian facilities: Validation on synthetic data

Antonin Danalet, Michel Bierlaire

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

hEART 2014
3nd Symposium of the European Association for Research in Transportation
University of Leeds

September 11, 2014
Outline

1. Motivation: Activity-based model for pedestrian facilities
2. Importance sampling for activity modeling
3. Validation with synthetic data
Activities in pedestrian infrastructure
Spatial choices in pedestrian infrastructure
The challenges of spatial choices: Large choice sets

In a transport hub

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of activity types</td>
<td>5</td>
</tr>
<tr>
<td>Number of activity-episodes per sequence</td>
<td>0-9</td>
</tr>
<tr>
<td>Number of activity-episode sequences</td>
<td>$5^9$</td>
</tr>
</tbody>
</table>

Without considering destination choice nor time spent at each destination...
Observations: activity patterns in a transport hub

Activity types

- Waiting for the train (on platform 9)
- Having a tea (in Starbucks)
- Buying a ticket (at the machine)

 Timeline:
- 7:40
- 7:43
- 7:48
- 8:01
- 8:03
- 8:12
Activity network

Activity types

$\mathcal{A}_1$

$\mathcal{A}_2$

$\vdots$

$\mathcal{A}_k$

Activity network

$S$

$e$

1 2 \cdots T \quad \text{Time}
Importance sampling for activity modeling

Activity network

Convenience store
Fast food
Cafe
Service
Shop
Walking

08:00-08:01
08:01-08:02
08:02-08:03
08:03-08:04
08:04-08:05
08:05-08:06
08:06-08:07
08:07-08:08
08:08-08:09
08:09-08:10

A. Danalet (EPFL)
Utility structure

- Utility of activity pattern:
  - time-of-day preferences
  - satiation effects: marginal utility decreases with increasing duration
    \[ V(duration) = \eta \ln(duration) \]
  - scheduling constraints: schedule delay

(Ettema et al.; 2007)
Choice set generation: Metropolis-Hastings algorithm

(Flötteröd and Bierlaire; 2013)
Choice set generation in the activity network

- Sample paths from given distribution, without full enumeration
- Possibility to define non-link additive cost
- Path cost defined as

$$
\delta(\Gamma) = - \sum_{v \in \Gamma} \delta_v(v) - \mu_\Gamma \cdot \delta_\Gamma(\Gamma)
$$

with
- link cost: frequency of observations
- path cost: length of observed paths

- Target weight defined as

$$
b(i) = \exp \left( - \mu \delta(\Gamma) \right)
$$

with $\mu$ a scale parameter
Time-invariance

- Different time discretisation and costs $\Rightarrow$ different scale parameters.
- Let’s define the scale parameter as
  \[
  \mu = \frac{\ln 2}{(\zeta - 1)\delta_{SP}}
  \]
- Path of cost $\zeta\delta_{SP}$ sampled twice less than the shortest path.
- $\zeta = 1$ only samples the shortest path;
  $\zeta \to \infty$ sample paths independently of their cost.
Utility function

- Utility of activity path $\Gamma$ with correction term for importance sampling:

$$V_{\Gamma} = \sum_{k,\tau} \beta_k l_{k,\tau} + \sum_{\text{episodes } e} \eta_k \ln(t_{k,e}) + \ln \frac{k_{\Gamma}}{b(\Gamma)}$$

- Fix one $\beta$ to 0 for identification.

- Application to WiFi traces on a campus: Danalet and Bierlaire (2014)

(Frejinger et al.; 2009)
Activity network

Activity types

Activity type 1
Activity type 2
Activity type 3

Activity network

243 alternatives
Activity network

Activity types

Activity type 1
Activity type 2
Activity type 3

Activity network

Time unit 1
Time unit 2
Time unit 3
Time unit 4
Time unit 5
Time unit 6

729 alternatives
Time-of-day preference is Cauchy distributed

- Utility of activity pattern:
  - **Time-of-day preferences**: symmetrical Cauchy distribution
    \[
    V'(\tau) = \frac{V_{max}}{c\pi\left((\frac{\tau-b}{c})^2 + 1\right)}
    \]
  - **Satiation effects**: marginal utility decreases with increasing duration
    \[
    V(duration) = \eta \ln(duration)
    \]
  - Scheduling constraints: schedule delay

(Ettema et al.; 2004)
## True values

<table>
<thead>
<tr>
<th>Parameters</th>
<th>True values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{max,1}$</td>
<td>3.0</td>
</tr>
<tr>
<td>$b_1$</td>
<td>2.5</td>
</tr>
<tr>
<td>$c_1$</td>
<td>2.0</td>
</tr>
<tr>
<td>$V_{max,2}$</td>
<td>4.0</td>
</tr>
<tr>
<td>$b_2$</td>
<td>4.0</td>
</tr>
<tr>
<td>$c_2$</td>
<td>3.0</td>
</tr>
<tr>
<td>$\eta_1$</td>
<td>2.0</td>
</tr>
<tr>
<td>$\eta_2$</td>
<td>1.3</td>
</tr>
<tr>
<td>$\eta_3$</td>
<td>0.8</td>
</tr>
<tr>
<td>$\gamma_e$</td>
<td>-1.2</td>
</tr>
<tr>
<td>$\gamma_l$</td>
<td>-1.8</td>
</tr>
</tbody>
</table>
### Estimation with full choice set

<table>
<thead>
<tr>
<th>Description</th>
<th>Coeff. estimate</th>
<th>Robust Asympt. std. error</th>
<th>t-stat (true value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{max,1}$</td>
<td>3.25</td>
<td>0.322</td>
<td>0.78</td>
</tr>
<tr>
<td>$b_1$</td>
<td>2.42</td>
<td>0.104</td>
<td>0.77</td>
</tr>
<tr>
<td>$c_1$</td>
<td>2.11</td>
<td>0.190</td>
<td>0.58</td>
</tr>
<tr>
<td>$V_{max,2}$</td>
<td>3.91</td>
<td>0.723</td>
<td>0.12</td>
</tr>
<tr>
<td>$b_2$</td>
<td>4.34</td>
<td>0.370</td>
<td>0.92</td>
</tr>
<tr>
<td>$c_2$</td>
<td>3.18</td>
<td>0.646</td>
<td>0.28</td>
</tr>
<tr>
<td>$\eta_1$</td>
<td>1.98</td>
<td>0.0512</td>
<td>0.39</td>
</tr>
<tr>
<td>$\eta_2$</td>
<td>1.38</td>
<td>0.0477</td>
<td>1.68</td>
</tr>
<tr>
<td>$\eta_3$</td>
<td>0.792</td>
<td>0.0522</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Number of observations = 10,000
Importance sampling

• Utility function:

\[ V_G = \mu \cdot \left( \sum_{k,\tau} \frac{V_{\text{max}}}{c \pi \left( \left( \frac{\tau - b}{c} \right)^2 + 1 \right)} + \sum_{\text{episodes } e} \eta_k \ln(t_{k,e}) \right) + \ln \frac{k_G}{b(G)} \]

with the true value for one node fixed, and the scale \( \mu \) estimated.
Number of distinct paths generated
(only time of day, 5 time units)
Validation with synthetic data

Estimation with importance sampling
(only time of day, 5 time units)
Future work

- Sensitivity analysis / 6 time units
- Define clear rules for how to define
  - Cost function in the Metropolis-Hastings algorithm
  - The scale parameter $\zeta$
  - The size of the choice set
- Gunnar’s idea: Define the scale parameter $\zeta$ sequentially (Lemp and Kockelman; 2012): draw alternatives in proportion to updated choice-probability estimates
Thank you
Questions / suggestions?
**URL:** http://www.strc.ch/conferences/2014/Danalet_Bierlaire.pdf


**URL:** http://dx.doi.org/10.1016/j.tra.2007.03.001
References II

**URL:** http://dx.doi.org/10.1016/j.trb.2012.11.002

**URL:** http://dx.doi.org/10.1016/j.trb.2009.03.001

**URL:** http://dx.doi.org/10.1016/j.tra.2011.11.004