Choice set generation for activities using importance sampling

Antonin Danalet, Michel Bierlaire

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

Ninth Workshop on Discrete Choice Models

June 19, 2014
Outline

1. Motivation: Activity-based model for pedestrian facilities
2. Literature review: from consideration set to importance sampling
3. Importance sampling for activity modeling
4. Case study: A multimodal transport hub
5. Conclusion
Pedestrian demand management strategies

- Pedestrian facilities
  - Transportation hubs (train stations, airports, ...)
  - Mass gathering (music festivals, ...)
  - Shops
  - ...

- Challenges
  - Designing efficient buildings
  - Locating points of interest
  - Modifying schedules
  - ...

⇒ Forecast the impact of pedestrian demand management strategies on activity and destination choices of visitors
Spatial choices in pedestrian infrastructure

1. First signal
2. Ticket machine
3. Supermarket
4. Kiosk
5. Platform 9
The challenges of spatial choices: Large choice sets

In a transport hub

<table>
<thead>
<tr>
<th>Number</th>
<th></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>
<tr>
<td>Number of destinations per activity type</td>
<td>1-5</td>
</tr>
<tr>
<td>Number of destinations per activity-episode sequence</td>
<td>(5^{10})</td>
</tr>
</tbody>
</table>

Without considering time spent at each destination...
Sequential choice:
1. activity type, sequence, time of day and duration
2. destination choice conditional on 1

Motivations:
1. Behavior: precedence of activity choice over destination choice (e.g., Bowman and Ben-Akiva; 2001)
2. Dimensional: destinations × time × position in the sequence is not tractable

Here we focus on 1.
Examples of 2: Ton (2014); Kalakou and Moura (2014).
Choice set generation

- Universal choice set $U$:
  - Computational: Too big, not usable
  - Behavior: Decision makers do not consider all alternatives

- Consideration choice set $C_n$:
  - Not known
    - Manski (1977): $P_n(i) = \sum_{C \in G} P_n(i|C)P_n(C)$
    - Set $G$ of all non-empty subsets of $U$ is exponentially large
    - Usual simplification: $G = \{C_n\}$ and $P(C_n) = 1$
  - Coverage issue: the chosen alternative (supposedly the best) not in $C_n$

- Sampling of alternatives from $U$
  - Contains the chosen alternative and the considered alternatives
  - Assumption about biases:
    - forgetting alternatives $>\$ adding non-considered alternatives

(Frejinger and Bierlaire; 2010)
Choice set generation in route choice

### Consideration choice set

**Shortest-path based algorithms**
- Deterministic algorithm
  - link elimination
  - link penalty
  - labeled paths
  - branch-and-bound (Prato and Bekhor; 2006)
- Monte-Carlo simulation
- Gateway algorithm (e.g., Bierlaire and Frejinger; 2008)

### Sampling of alternatives

- Random walk (Frejinger et al.; 2009)
- Link sampling for recursive logit (Fosgerau et al.; 2013)
- Metropolis-Hastings path sampling (Flötteröd and Bierlaire; 2013; Chen; 2013)

(Frejinger and Bierlaire; 2010; Chen; 2013)
Choice set generation in activity/destination choices

Consideration choice set
- General review before 2009 in Pagliara and Timmermans (2009)
- Dominance-based choice set in destination choice (Cascetta and Papola; 2009)
- Refueling decision (Pramono and Oppewal; 2012)

Sampling of alternatives
- Residential location choice (McFadden; 1978; Ben-Akiva and Bowman; 1998)
- Destination choice (Yagi and Mohammadian; 2008)
Observations: activity patterns in a transport hub

Activity types

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

Activity types

\[ \mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_k \]

Activity network

\[ S \quad \cdots \quad e \]

Time

1 2 \ldots T
Importance sampling for activity modeling

Activity network

Convenience store
Fast food
Cafe
Service
Shop
No activity

Choice set generation for activities

June 19, 2014 12 / 36
Activity network and importance sampling

- Universal choice set is behaviorally meaningful in the activity network: Decision maker can consider all alternatives (consider all activity types and time duration, not all combinations)
- Unattractive paths will be assigned a very small choice probability
Choice set generation: Metropolis-Hastings algorithm

(Flötteröd and Bierlaire; 2013)
Importance sampling for activity modeling

Choice set generation in the activity network

- Sample paths from given distribution, without full enumeration
- With Metropolis-Hastings algorithm, possibility to define non-link additive cost
- Target weight defined as

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

with
- link cost: frequency of observations
- path cost: length of observed paths
Path and link cost for different path lengths.
Sum of path and link cost per length, weight ratio of 1
Case study: A multimodal transport hub

Sum of path and link cost per length, weight ratio of 1

A. Danalet (TRANSP-OR ENAC EPFL)

Choice set generation for activities

June 19, 2014 18 / 36
Activity network: frequency of observations

Convenience store
Fast food
Cafe
Service
Shop
No activity

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
Activity network: frequency of observations: Zoom
Activity network: Length of observations
Activity network: Length of observations with activities
Generated path with $\mu_V = 1$ and $\mu_G = 0$
Generated path with $\mu_v = 0.001$ and $\mu_T = 0$
Generated path with $\mu_V = 0.005$ and $\mu_T = 0$
Similarity measure

- Transition distribution is local, similar states generated in iterations
- Similarity measure:

\[
\frac{1}{K} \sum_{k=1}^{K} \frac{|\Gamma^k \cap \Gamma^{k+d}|}{\frac{1}{2} |\Gamma^k| + |\Gamma^{k+d}|}
\]

with \( |\Gamma^k \cap \Gamma^{k+d}| \) nb of identical nodes, \( k \) nb of iterations
Generated path length with $\mu_v = 0.005$, $\mu_\Gamma = 0$ and sample interval of 200
Generated path length with $\mu_V = 0.005$, $\mu_I = 1$ and sample interval of 200
Generated path length with $\mu_\nu = 0.005$, $\mu_\Gamma = 0.001$ and sample interval of 200
Generated path length with $\mu_v = 0.005$, $\mu_\Gamma = 0.002$ and sample interval of 200
New approach to activity-based modeling
- Importance sampling based on
  - time-of-day/activity attractivity
  - activity-episode duration
- Probability $q(\Gamma)$ of generating path $\Gamma$ can be then used in choice model, as in Danalet and Bierlaire (2014)
Open questions

Are node attractivity and path length the best measure of an “attractive” activity path?
Most common activity path in relative time

Convenience store
Fast food
Cafe
Service
Shop
No activity

0-1' 1-2' 2-3'
Most common activity path in relative time including at least one activity

Convenience store
Fast food
Cafe
Service
Shop
No activity
Different sequences of activities, independent of time

(a) 40'897 obs.  (b) 706 observations  (c) 548 observations  (d) 360 observations

(e) 270 observations  (f) 98 observations  (g) 39 observations  (h) 24 observations
Thank you!

Questions?
References I


URL: http://dx.doi.org/10.1080/0042098984529


URL: http://dx.doi.org/10.1016/j.trc.2007.07.007


URL: http://dx.doi.org/10.1016/S0965-8564(99)00043-9
References II

URL: http://dx.doi.org/10.1016/j.tra.2008.10.003

URL: http://dx.doi.org/10.5075/epfl-thesis-5649


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

**URL:** [http://dx.doi.org/10.1016/j.trb.2013.07.012](http://dx.doi.org/10.1016/j.trb.2013.07.012)


**URL:** [http://dx.doi.org/10.1016/j.trb.2009.03.001](http://dx.doi.org/10.1016/j.trb.2009.03.001)
**URL:** [http://www.strc.ch/conferences/2014/Kalakou_Moura.pdf](http://www.strc.ch/conferences/2014/Kalakou_Moura.pdf)


**URL:** http://dx.doi.org/10.1016/S1755-5345(13)70047-2


**URL:** http://dx.doi.org/10.3141/1985-03

References VI


**URL:** http://trb.metapress.com/index/0W4J63HV72WH3270.pdf