IATBR

The activity path approach to activity pattern modeling

Antonin Danalet, Michel Bierlaire

Windsor,
July 20, 2015
Outline

Motivation: Activity-based model for pedestrian facilities

An activity path approach to activity modeling

Choice set generation

Case study: pedestrians on EPFL campus
Outline

Motivation: Activity-based model for pedestrian facilities

An activity path approach to activity modeling

Choice set generation

Case study: pedestrians on EPFL campus
Motivation

- **Activity-based approach**: modeling the activity participation patterns
- **Not tour-based** (no “home” location in pedestrian facilities)
- **No hierarchy** of dimensions or aggregation (high temporal precision)
Activity-schedule approach vs Activity path approach

Pre-processing

Activity-episode sequence detection [DFB14]

Modeling

Activity path choice model [DB15]

Destination choice model with panel effect [Tin15]

SBA11
Goal

• Simultaneously model the choice of:
  – activity types,
  – order,
  – start times and
  – durations
  of activity episodes in a sequence.

• No mode choice.
Outline

Motivation: Activity-based model for pedestrian facilities

An activity path approach to activity modeling

Choice set generation

Case study: pedestrians on EPFL campus
Modeling assumption

• Sequential choice:
  1. activity type, sequence, time of day and duration
  2. destination choice conditional on 1.

• Motivations:
  – Behavior: precedence of activity choice over destination choice
  – Dimensional: destinations $\times$ time $\times$ position in the sequence is not tractable

Today, we focus on 1. [DB15].
Example of 2. on the same data: [Tin15].
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)
Activity network

Activity types

\[ \mathcal{A}_1 \]
\[ \mathcal{A}_2 \]
\[ \vdots \]
\[ \mathcal{A}_k \]

Activity network

\[ S \]
\[ T \]
\[ e \]

1 2 \[ \ldots \] T

Time units
Activity path

Convenience store
Fast food
Cafe
Service
Walking
Not in the train station
Outline

Motivation: Activity-based model for pedestrian facilities

An activity path approach to activity modeling

Choice set generation

Case study: pedestrians on EPFL campus
Sampling strategies

- Simple random sampling (SRS)
- Importance sampling using Metropolis-Hastings algorithm [FB13]
  - Observation score [Che13, DB15]
  - Strategic sampling [LK12]
Metropolis-Hastings algorithm

[FB13]
Metropolis-Hastings sampling of paths

• Sample paths from given distribution, without full enumeration
• To be defined:
  – Target weight:
    \[ b(i) = \exp(-\mu \delta(\Gamma)) \tag{1} \]
    Also with non-node-additive utility
  – Proposal distribution:
    \[
P_{\text{insert}} = \frac{e^{-\bar{\mu} \delta_{SP}(\text{origin},v) + \delta_{SP}(v,\text{destination})}}{\sum_{w} e^{-\bar{\mu} \delta_{SP}(\text{origin},w) + \delta_{SP}(w,\text{destination})}} \tag{2}
\]
    Relies on shortest paths, node-additive cost.
Strategic sampling

- Target weight: previously estimated model
- Proposal distribution: previously estimated model using only time-of-day preferences (node-additive)
Utility structure

- Utility of activity pattern:
  - Node utility \( V(\mathcal{A}_{k,t}) \)
    - time-of-day preferences
  - Activity-episode utility \( V(a) \)
    - satiation effects: decreasing marginal utility, \( \eta \ln(\text{duration}) \)
    - scheduling constraints: schedule delay
  - Activity path utility \( V(\Gamma) \)
    - primary activity

- Sampling correction

\[
\mu \left( \sum_{k=1}^{K} \sum_{\tau=1}^{T} V(\mathcal{A}_{k,\tau}) + \sum_{a \in \mathcal{A}_{1:T}} V(a) + V(\Gamma) \right) + \ln \frac{k_{\Gamma n}}{b(\Gamma)}
\]
Outline

Motivation: Activity-based model for pedestrian facilities

An activity path approach to activity modeling

Choice set generation

Case study: pedestrians on EPFL campus
Case study: pedestrians on EPFL campus

- 13’000 people per day
- 8 activity types:
  - classrooms,
  - shops,
  - offices,
  - restaurant,
  - library,
  - lab,
  - other and
  - not being detected
- 12 time units in the activity network, from 7am to 7pm
- WiFi traces [DFB14]
### Proposal distribution (using simple random sampling)

<table>
<thead>
<tr>
<th>Description</th>
<th>Coeff. estimate</th>
<th>Robust Asympt. std. error</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{NA, 17-19, employees}$</td>
<td>0.263</td>
<td>0.0302</td>
<td>8.70</td>
</tr>
<tr>
<td>$\beta_{NA, 14-17, students}$</td>
<td>-0.222</td>
<td>0.191</td>
<td>-1.16</td>
</tr>
<tr>
<td>$\beta_{NA, 7-8, students}$</td>
<td>0.349</td>
<td>0.0281</td>
<td>12.44</td>
</tr>
<tr>
<td>$\beta_{NA, 7-9, employees}$</td>
<td>0.326</td>
<td>0.0262</td>
<td>12.43</td>
</tr>
<tr>
<td>$\beta_{NA, 17-19, students}$</td>
<td>1.14</td>
<td>0.187</td>
<td>6.09</td>
</tr>
<tr>
<td>$\beta_{classroom, 12-14, students}$</td>
<td>-0.336</td>
<td>0.337</td>
<td>-1.00</td>
</tr>
<tr>
<td>$\beta_{classroom, 7-12, employees}$</td>
<td>-0.723</td>
<td>0.397</td>
<td>-1.82</td>
</tr>
<tr>
<td>$\beta_{classroom, 7-12, students}$</td>
<td>0.598</td>
<td>0.262</td>
<td>2.28</td>
</tr>
<tr>
<td>$\beta_{library, 14-19, employees}$</td>
<td>-0.624</td>
<td>0.553</td>
<td>-1.13</td>
</tr>
<tr>
<td>$\beta_{library, 12-14, employees}$</td>
<td>-0.575</td>
<td>0.481</td>
<td>-1.20</td>
</tr>
<tr>
<td>$\beta_{library, 7-12, employees}$</td>
<td>-1.57</td>
<td>0.508</td>
<td>-3.09</td>
</tr>
<tr>
<td>$\beta_{office, 14-19, employees}$</td>
<td>1.41</td>
<td>0.246</td>
<td>5.73</td>
</tr>
<tr>
<td>$\beta_{office, 7-12, employees}$</td>
<td>1.12</td>
<td>0.228</td>
<td>4.92</td>
</tr>
<tr>
<td>$\beta_{restaurant, 14-19, students}$</td>
<td>-0.410</td>
<td>0.185</td>
<td>-2.21</td>
</tr>
<tr>
<td>$\beta_{restaurant, 12-14, employees}$</td>
<td>0.136</td>
<td>0.0259</td>
<td>5.26</td>
</tr>
<tr>
<td>$\beta_{restaurant, 12-14, students}$</td>
<td>0.665</td>
<td>0.286</td>
<td>2.32</td>
</tr>
</tbody>
</table>

Number of observations = 1087
Number of estimated parameters = 43

\[ \mathcal{L} (\beta_0) = -5016.636 \]
\[ \mathcal{L} (\hat{\beta}) = -453.225 \]
\[ \rho^2 = 0.910 \]
\[ \bar{\rho}^2 = 0.901 \]
Target weight (using simple random sampling)

<table>
<thead>
<tr>
<th>Description</th>
<th>Coeff. estimate</th>
<th>Robust Asympt. std. error</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{\text{library 7-12, employees}}$</td>
<td>-2.08</td>
<td>0.422</td>
<td>-4.93</td>
</tr>
<tr>
<td>$\beta_{\text{office 7-12, 14-19, employees}}$</td>
<td>1.69</td>
<td>0.393</td>
<td>4.30</td>
</tr>
<tr>
<td>$\beta_{\text{restaurant 12-14, employees}}$</td>
<td>1.22</td>
<td>0.502</td>
<td>2.43</td>
</tr>
<tr>
<td>$\beta_{\text{shop 12-14, students}}$</td>
<td>-7.36</td>
<td>1.24</td>
<td>-5.92</td>
</tr>
<tr>
<td>$\beta_{\text{shop 7-12, 14-19, students}}$</td>
<td>-1.16</td>
<td>0.538</td>
<td>-2.16</td>
</tr>
<tr>
<td>$\beta_{\text{NA 7-8, students}}$</td>
<td>4.27</td>
<td>0.995</td>
<td>4.29</td>
</tr>
<tr>
<td>$\beta_{\text{NA 8-12, students}}$</td>
<td>1.40</td>
<td>0.498</td>
<td>2.82</td>
</tr>
<tr>
<td>$\beta_{\text{NA 17-19, students}}$</td>
<td>1.75</td>
<td>0.568</td>
<td>3.08</td>
</tr>
<tr>
<td>$\beta_{\text{NA 9-17, employees}}$</td>
<td>1.43</td>
<td>0.296</td>
<td>4.84</td>
</tr>
<tr>
<td>$\beta_{\text{NA 7-9, 17-19, employees}}$</td>
<td>3.34</td>
<td>0.554</td>
<td>6.02</td>
</tr>
<tr>
<td>$\eta_{\text{Office, Lab, Classroom}}$</td>
<td>5.22</td>
<td>0.764</td>
<td>6.83</td>
</tr>
<tr>
<td>$\eta_{\text{Restaurant, Library, Other}}$</td>
<td>7.85</td>
<td>1.11</td>
<td>7.10</td>
</tr>
<tr>
<td>$\eta_{\text{Shop}}$</td>
<td>7.33</td>
<td>0.894</td>
<td>8.20</td>
</tr>
<tr>
<td>$\eta_{\text{NA}}$</td>
<td>2.75</td>
<td>0.393</td>
<td>7.00</td>
</tr>
<tr>
<td>$\beta_{3+ \text{ lab episodes}}$</td>
<td>-5.03</td>
<td>0.952</td>
<td>-5.28</td>
</tr>
<tr>
<td>$\beta_{3+ \text{ resto episodes}}$</td>
<td>-2.50</td>
<td>0.759</td>
<td>-3.29</td>
</tr>
</tbody>
</table>

Number of observations = 1087
Number of estimated parameters = 22

$L(\beta_0) = -5016.636$

$L(\hat{\beta}) = -47.218$

$\rho^2 = 0.991$

$\bar{\rho}^2 = 0.986$
## Model using strategic sampling

<table>
<thead>
<tr>
<th>Description</th>
<th>Coeff. estimate</th>
<th>Robust Asympt. std. error</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{\text{Library 12-14, students}}$</td>
<td>-1.45</td>
<td>0.235</td>
<td>-6.15</td>
</tr>
<tr>
<td>$\beta_{\text{Restaurant 12-14, students}}$</td>
<td>0.769</td>
<td>0.106</td>
<td>7.26</td>
</tr>
<tr>
<td>$\beta_{\text{Shop 14-19, students}}$</td>
<td>1.14</td>
<td>0.160</td>
<td>7.16</td>
</tr>
<tr>
<td>$\beta_{\text{NA 7-8, students}}$</td>
<td>2.15</td>
<td>0.223</td>
<td>9.63</td>
</tr>
<tr>
<td>$\beta_{\text{NA 8-12, students}}$</td>
<td>1.39</td>
<td>0.0792</td>
<td>17.52</td>
</tr>
<tr>
<td>$\beta_{\text{NA 17-19, students}}$</td>
<td>1.80</td>
<td>0.108</td>
<td>16.69</td>
</tr>
<tr>
<td>$\beta_{\text{NA 7-9, 17-19, employees}}$</td>
<td>1.59</td>
<td>0.0793</td>
<td>20.07</td>
</tr>
<tr>
<td>$\eta_{\text{Office, Lab, Classroom, Library, Other, NA}}$</td>
<td>-2.07</td>
<td>0.110</td>
<td>-18.81</td>
</tr>
<tr>
<td>$\eta_{\text{Restaurant}}$</td>
<td>-3.41</td>
<td>0.284</td>
<td>-11.98</td>
</tr>
<tr>
<td>$\eta_{\text{Shop, Library}}$</td>
<td>-1.35</td>
<td>0.120</td>
<td>-11.23</td>
</tr>
<tr>
<td>$\beta_{1 \text{ Restaurant episode}}$</td>
<td>1.83</td>
<td>0.148</td>
<td>12.35</td>
</tr>
<tr>
<td>$\beta_{2+ \text{ Classroom episodes, employees}}$</td>
<td>-0.736</td>
<td>0.0669</td>
<td>-11.00</td>
</tr>
<tr>
<td>$\beta_{2+ \text{ Shop episodes}}$</td>
<td>-2.79</td>
<td>0.417</td>
<td>-6.69</td>
</tr>
<tr>
<td>$\beta_{0 \text{ Library episode, employees}}$</td>
<td>2.60</td>
<td>0.252</td>
<td>10.32</td>
</tr>
<tr>
<td>$\beta_{\text{primary activity Library, students}}$</td>
<td>0.128</td>
<td>0.0474</td>
<td>2.71</td>
</tr>
<tr>
<td>$\beta_{\text{schedule delay, students}}$</td>
<td>-0.509</td>
<td>0.184</td>
<td>-2.77</td>
</tr>
</tbody>
</table>

Number of observations = 1087  
Number of estimated parameters = 31

\[
\mathcal{L}(\beta_0) = -5016.636
\]

\[
\mathcal{L}(\tilde{\beta}) = -1411.965
\]

\[\rho^2 = 0.719\]

\[\tilde{\rho}^2 = 0.712\]
Validation

Predicted probabilities for the chosen alternative

Probabilities for the chosen alternative

Simple random sampling

Strategic sampling
Forecasting

Number of activity episodes

Number of observations

0  0  14  211  497  285  71  9
1  2  3  4  5  6  7  8
Conclusion

- Simultaneous choice of
  - activity type,
  - time of day,
  - duration,
  - order.

- Not home-based, nor tour-based.

- Importance sampling: allows to include more parameters.

- Important feature: allows to add variables related to the path.
  - patterns (e.g., a office-restaurant-office pattern for lunch) or
  - primary activity.
Thank you

IATBR:
The activity path approach to activity pattern modeling
Antonin Danalet, Michel Bierlaire

– antonin.danalet@epfl.ch
Bibliography I

Jingmin Chen.  
*Modeling route choice behavior using smartphone data.*  
PhD thesis, Ecole Polytechnique Fédérale de Lausanne,  
Switzerland, 2013.

Antonin Danalet and Michel Bierlaire.  
Importance sampling for activity path choice.  
In *15th Swiss Transport Research Conference (STRC)*,  
page 42, Monte Verità, Ascona, Switzerland, 2015.

Antonin Danalet, Bilal Farooq, and Michel Bierlaire.  
A Bayesian approach to detect pedestrian  
destination-sequences from WiFi signatures.  
