STRC

Importance sampling for activity path choice

Antonin Danalet, Michel Bierlaire

Monte Verità, Ascona, Switzerland,
April 15, 2015

Outline

Outline

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)

Outline

Motivation: Activity-based model for pedestrian facilities

Literature review

A path choice approach to activity modeling

Choice set generation

Literature review

- Tour-based approach [BBA01, SBA11, AZBA12]
- Multiple discrete continuous nested extreme value model [PB10]
- Dynamic scheduling process [Hab11]

Literature review

Time representation in activity modeling:

- Time is decomposed in **tours** [BBA01, SBA11, AZBA12]
- Time is allocated to activity types (**no sequence**) [PB10]
- Time is allocated to activity types (**sequentially in time**) [Hab11]

Literature review

Problems

- **tours** [BBA01, SBA11, AZBA12]
- **no sequence** [PB10]
- **no pattern utility** [Hab11]

Outline

Motivation: Activity-based model for pedestrian facilities

Literature review

A path choice approach to activity modeling

Choice set generation

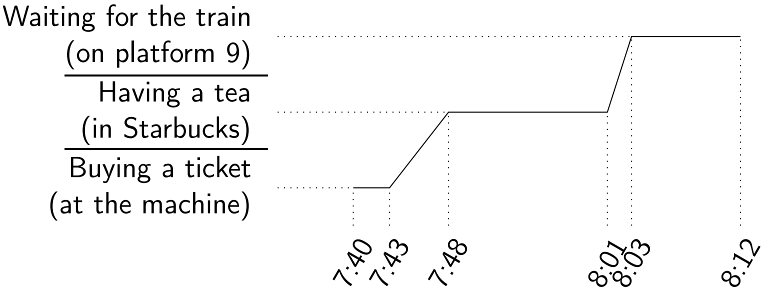
Modeling assumption

- 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

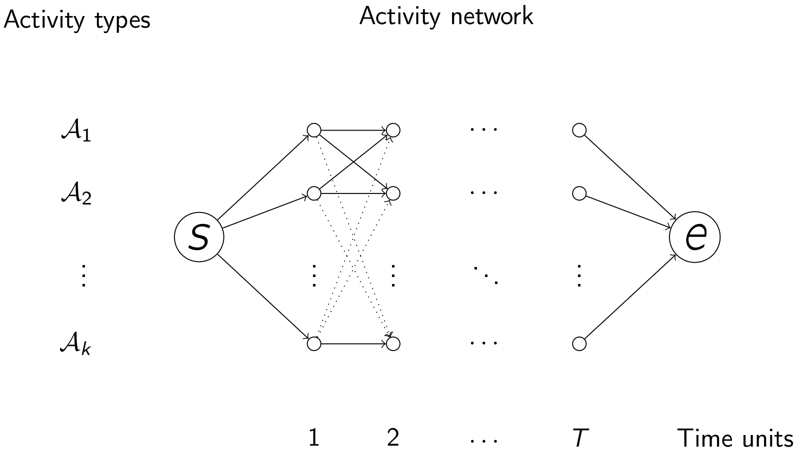
Today, we focus on 1. [DB15].
Tomorrow, 16:20, example of 2. on the same data [TDdLB15].

Observations: activity patterns in a transport hub

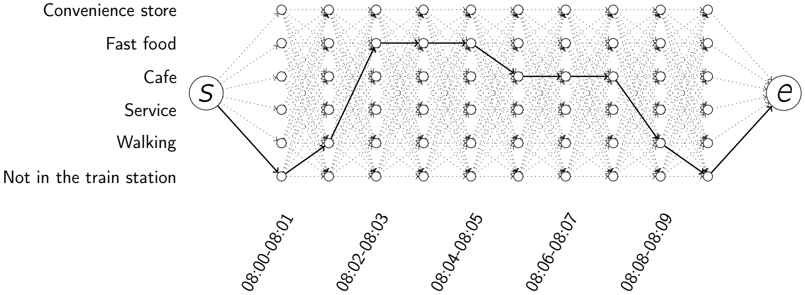
Activity types



Activity network



Activity path

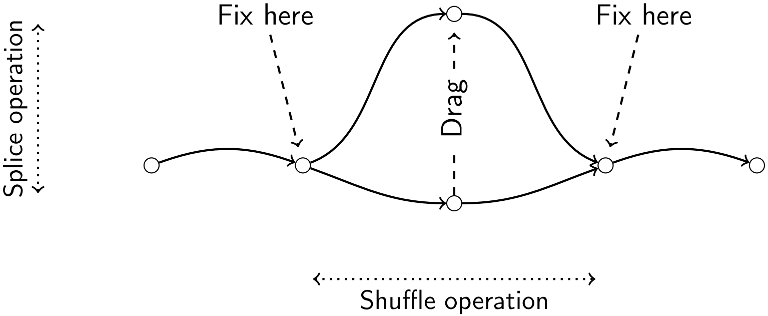


Outline

Sampling strategies

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

Metropolis-Hastings sampling of paths



[FB13]

Metropolis-Hastings sampling of paths

- Sample paths from given distribution, without full enumeration
- To be defined:
  - Target weight:
    $$b(i) = \exp\big(- \mu\delta(\Gamma)\big) \quad (1)$$
    Also with non-node-additive utility
  - Proposal distribution:
    $$P_{\text{insert}} = \frac{e^{-\tilde{\mu}\delta_{\boldsymbol{SP}}(\textit{origin},v)+\delta_{\boldsymbol{SP}}(v,\textit{destination})}}{\sum_w e^{-\tilde{\mu}\delta_{\boldsymbol{SP}}(\textit{origin},w)+\delta_{\boldsymbol{SP}}(w,\textit{destination})}} \quad (2)$$
    Relies on shortest paths, node-additive cost.

Utility structure

- Utility of activity pattern:
  - **time-of-day preferences**
  - **satiation effects: marginal utility decreases with increasing duration**

$$V(\mathit{duration}) = \eta \ln(\mathit{duration})$$

  - scheduling constraints: schedule delay

[EBPA07]

Observation score

- Node attractivity $\delta_v(v)$
- Activity-episode length attractivity $\delta_a(a)$
- Total attractivity:

$$\delta(\Gamma) = \sum_{v \in \Gamma} \delta_v(v) + r \sum_{a \in \Gamma} \delta_a(a) \tag{3}$$

- Scale and r estimated based on synthetic data [DB15].

Strategic sampling

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

Case study

- Activity-episode sequences from WiFi traces on EPFL campus [DFB14]
- Activity network
  - 8 activity types
  - 24 time units (:00 - :15 / :15 - :59 between 7am and 7pm)

Results

- 100 elements in the choice set:
  SRS vs observation score.
- 10 elements in the choice set:
  SRS vs observation score vs strategic sampling.

Results: SRS, 100 el. in choice set

Attributes	Estimates	Std. error	t-stat
$\eta_{\text{Classroom, Shop, Library}}$	-0.492	0.168	-2.93
$\eta_{\text{Lab, Restaurant, Office, Other}}$	-0.638	0.167	-3.81
$\beta_{\text{3 lab episodes}}$	-0.998	0.265	-3.77
$\beta_{\text{4+ lab episodes}}$	-0.100	0.0243	-4.12
$\beta_{\text{3 office episodes}}$	-0.505	0.112	-4.49
$\beta_{\text{4+ office episodes}}$	-0.0494	0.0107	-4.62
$\beta_{\text{3 restaurant episodes}}$	-0.352	0.150	-2.34
$\beta_{\text{4+ restaurant episodes}}$	-0.0945	0.0270	-3.50
$\beta_{\text{3+ shop episodes}}$	-1.21	0.321	-3.77
$\beta_{\text{nb nodes NA afternoon, students}}$	-0.941	0.269	-3.50
$\beta_{\text{nb nodes NA before/after work, employees}}$	0.245	0.0726	3.38
$\beta_{\text{nb nodes NA work, employees}}$	-1.07	0.278	-3.86
$\beta_{\text{nb nodes classroom morning/afternoon, employees}}$	-0.132	0.0296	-4.46
$\beta_{\text{primary activity library, students}}$	0.0404	0.0108	3.73

Number of observations = 1734

Number of estimated parameters = 14

$$\begin{aligned} \mathcal{L}(\beta_0) &= -8002.619 \\ \mathcal{L}(\hat{\beta}) &= -10.234 \\ \rho^2 &= 0.999 \\ \bar{\rho}^2 &= 0.997 \end{aligned}$$

Results: observation score, 100 el. in choice set

Attributes	Estimates	Std. error	t-stat
$\eta_{\text{Classroom, Shop, Library}}$	-0.484	0.0877	-5.52
$\eta_{\text{Lab, Restaurant, Office, Other}}$	-0.687	0.137	-5.02
$\beta_{\text{3 lab episodes}}$	-0.710	0.146	-4.86
$\beta_{\text{4+ lab episodes}}$	-0.0735	0.0241	-3.05
$\beta_{\text{3 office episodes}}$	-0.427	0.139	-3.08
$\beta_{\text{4+ office episodes}}$	-0.0794	0.0265	-3.00
$\beta_{\text{3 restaurant episodes}}$	-0.0535	0.0122	-4.39
$\beta_{\text{4+ restaurant episodes}}$	-0.731	0.199	-3.67
$\beta_{\text{3+ shop episodes}}$	-0.740	0.250	-2.96
$\beta_{\text{nb nodes NA afternoon, students}}$	-1.10	0.347	-3.17
$\beta_{\text{nb nodes NA before/after work, employees}}$	0.231	0.0523	4.42
$\beta_{\text{nb nodes NA work, employees}}$	-0.0762	0.0199	-3.83
$\beta_{\text{nb nodes classroom morning/afternoon, employees}}$	-0.0908	0.0460	-1.97
$\beta_{\text{primary activity library, students}}$	0.0592	0.0260	2.28

Number of observations = 1734

Number of estimated parameters = 14

$$
\begin{aligned}
\mathcal{L}(\beta_0) &= -8002.619 \\
\mathcal{L}(\hat{\beta}) &= -13.293 \\
\rho^2 &= 0.998 \\
\bar{\rho}^2 &= 0.997
\end{aligned}
$$

Results: SRS, 10 el. in choice set

Attributes	Estimates	Std. error	t-stat
$\eta_{\text{Classroom, Shop, Library}}$	-2.48	0.00727	-341.00
$\eta_{\text{Lab, Restaurant, Office, Other}}$	-4.41	1.80e+308	-0.00
$\beta_{\text{3 lab episodes}}$	-3.42	0.00211	-1621.37
$\beta_{\text{4+ lab episodes}}$	-0.372	0.00406	-91.48
$\beta_{\text{3 office episodes}}$	-1.11	1.80e+308	-0.00
$\beta_{\text{4+ office episodes}}$	-0.598	0.00710	-84.27
$\beta_{\text{3 restaurant episodes}}$	-4.54	1.80e+308	-0.00
$\beta_{\text{4+ restaurant episodes}}$	-0.515	0.00418	-123.07
$\beta_{\text{3+ shop episodes}}$	-6.06	0.00167	-3637.41
$\beta_{\text{nb nodes NA afternoon, students}}$	-3.71	1.80e+308	-0.00
$\beta_{\text{nb nodes NA before/after work, employees}}$	0.886	0.00197	449.89
$\beta_{\text{nb nodes NA work, employees}}$	-0.922	0.00555	-166.01
$\beta_{\text{nb nodes classroom morning/afternoon, employees}}$	-0.856	0.00125	-685.45
$\beta_{\text{primary activity library, students}}$	0.267	0.00382	69.75

Number of observations = 1734

Number of estimated parameters = 14

$$\begin{aligned}
\mathcal{L}(\beta_0) &= -4157.950 \\
\mathcal{L}(\hat{\beta}) &= -0.000 \\
\rho^2 &= 1.000 \\
\bar{\rho}^2 &= 0.997
\end{aligned}$$

Results: observation score, 10 el. in choice set

Attributes	Estimates	Std. error	t-stat
$\eta_{\text{Classroom, Shop, Library}}$	-2.83	0.0400	-70.68
$\eta_{\text{Lab, Restaurant, Office, Other}}$	-4.47	1.80e+308	-0.00
$\beta_{\text{3 lab episodes}}$	-3.06	0.0404	-75.63
$\beta_{\text{4+ lab episodes}}$	-0.484	0.0256	-18.96
$\beta_{\text{3 office episodes}}$	-3.66	0.0772	-47.48
$\beta_{\text{4+ office episodes}}$	-0.575	0.00909	-63.30
$\beta_{\text{3 restaurant episodes}}$	-4.82	0.0462	-104.19
$\beta_{\text{4+ restaurant episodes}}$	-0.530	0.0175	-30.26
$\beta_{\text{3+ shop episodes}}$	-4.80	1.80e+308	-0.00
$\beta_{\text{nb nodes NA afternoon, students}}$	-6.06	0.0608	-99.70
$\beta_{\text{nb nodes NA before/after work, employees}}$	0.529	1.80e+308	0.00
$\beta_{\text{nb nodes NA work, employees}}$	-0.893	0.0129	-69.37
$\beta_{\text{nb nodes classroom morning/afternoon, employees}}$	-1.02	0.0129	-79.07
$\beta_{\text{primary activity library, students}}$	0.284	0.0120	23.67

Number of observations = 1734

Number of estimated parameters = 14

$$\mathcal{L}(\beta_0) = -4157.950$$
$$\mathcal{L}(\hat{\beta}) = -0.000$$
$$\rho^2 = 1.000$$
$$\bar{\rho}^2 = 0.997$$

Results: strategic sampling, 10 el. in choice set

Attributes	Estimates	Std. error	t-stat
$\eta_{\text{Classroom, Shop, Library}}$	-1.17	0.0469	-24.99
$\eta_{\text{Lab, Restaurant, Office, Other}}$	-1.64	0.0636	-25.86
$\beta_{\text{3 lab episodes}}$	-3.43	0.133	-25.74
$\beta_{\text{4+ lab episodes}}$	-0.188	0.0156	-12.05
$\beta_{\text{3 office episodes}}$	-1.71	0.0575	-29.80
$\beta_{\text{4+ office episodes}}$	-0.204	0.00723	-28.18
$\beta_{\text{3 restaurant episodes}}$	-1.19	0.0900	-13.17
$\beta_{\text{4+ restaurant episodes}}$	-0.135	0.00492	-27.41
$\beta_{\text{3+ shop episodes}}$	-3.20	0.0885	-36.10
$\beta_{\text{nb nodes NA afternoon, students}}$	-1.50	0.123	-12.23
$\beta_{\text{nb nodes NA before/after work, employees}}$	0.112	0.0185	6.09
$\beta_{\text{nb nodes NA work, employees}}$	-0.502	0.0163	-30.84
$\beta_{\text{nb nodes classroom morning/afternoon, employees}}$	-0.441	0.0193	-22.87
$\beta_{\text{primary activity library, students}}$	0.224	0.00725	30.87

Number of observations = 1734

Number of estimated parameters = 14

$$
\begin{aligned}
\mathcal{L}(\beta_0) &= -4157.950 \\
\mathcal{L}(\hat{\beta}) &= -0.000 \\
\rho^2 &= 1.000 \\
\bar{\rho}^2 &= 0.997
\end{aligned}
$$

Results

- 100 elements in the choice set:
  SRS vs observation score.
  - SRS gives similar results as observation score
- 10 elements in the choice set:
  SRS vs observation score vs strategic sampling.
  - preliminary: strategic sampling performs better than SRS, observation score

Conclusion and future work

- SRS and importance sampling with observation score generate dominated alternatives
- Strategic sampling gives the flexibility needed in activity path choice
- Activity path size for correlation between activity paths
  - Primary Activity Path Size (PAPS)
  - Activity Pattern Path Size (APPS)

Thank you

STRC:
Importance sampling
for activity path choice

Antonin Danalet, Michel Bierlaire

– antonin.danalet@epfl.ch

Bibliography I

Maya Abou-Zeid and Moshe Ben-Akiva.
Well-being and activity-based models.
Transportation, 39(6):1189–1207, January 2012.

J.L Bowman and M.E Ben-Akiva.
Activity-based disaggregate travel demand model system with activity schedules.
Transportation Research Part A, 35(1):1–28, January 2001.

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

Bibliography II

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.
Transportation Research Part C, 44:146–170, 2014.

Dick Ettema, Fabian Bastin, John Polak, and Olu Ashiru.
Modelling the joint choice of activity timing and duration.
Transportation Research Part A, 41(9):827–841, November 2007.

Bibliography III

Gunnar Flötteröd and Michel Bierlaire.
Metropolis-Hastings sampling of paths.
Transportation Research Part B, 48:53–66, February 2013.

Khandker M. Nurul Habib.
A random utility maximization (RUM) based dynamic activity scheduling model: Application in weekend activity scheduling.
Transportation, 38(1):123–151, July 2011.

Jason D. Lemp and Kara M. Kockelman.
Strategic sampling for large choice sets in estimation and application.
Transportation Research Part A: Policy and Practice, 46(3):602–613, March 2012.

Bibliography IV

Abdul Rawoof Pinjari and Chandra Bhat.
A multiple discrete-continuous nested extreme value (MDCNEV) model: Formulation and application to non-worker activity time-use and timing behavior on weekdays.
Transportation Research Part B, 44(4):562–583, May 2010.

Yoram Shiftan and Moshe Ben-Akiva.
A practical policy-sensitive, activity-based, travel-demand model.
Annals of Regional Science, 47:517–541, 2011.

Bibliography V

Loïc Tinguely, Antonin Danalet, Matthieu de Lapparent, and Michel Bierlaire.
Destination Choice Model including a panel effect using WiFi localization in a pedestrian facility.
In *15th Swiss Transport Research Conference (STRC)*, page 44, Monte Verità, Ascona, Switzerland, 2015.