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

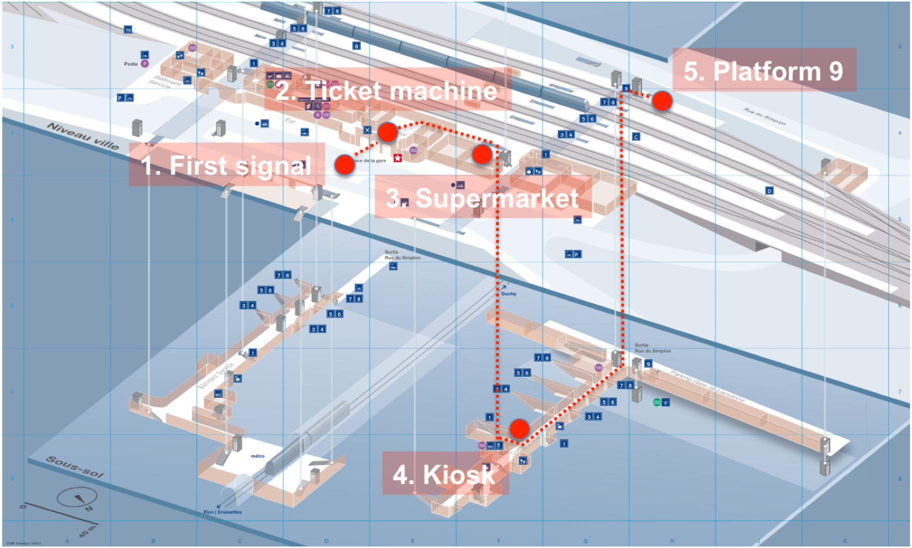
September 11, 2014

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

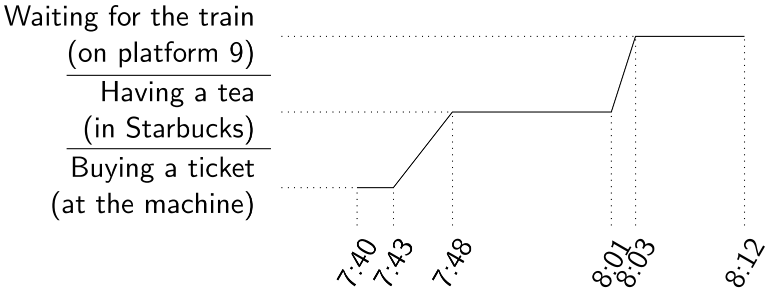
In a transport hub

Number of activity types	5
Number of activity-episodes per sequence	0-9
Number of activity-episode sequences	5^9

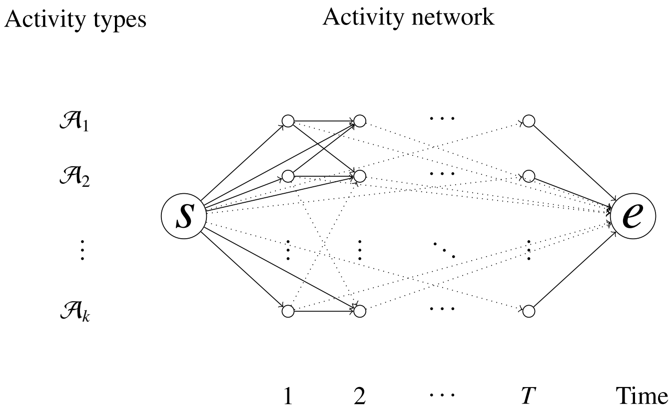
Without considering destination choice nor time spent at each destination...

Observations: activity patterns in a transport hub

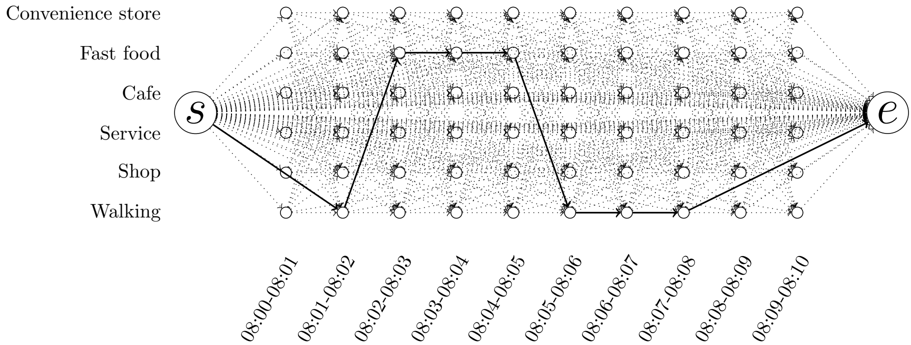
Activity types



Activity network



Activity network



Utility structure

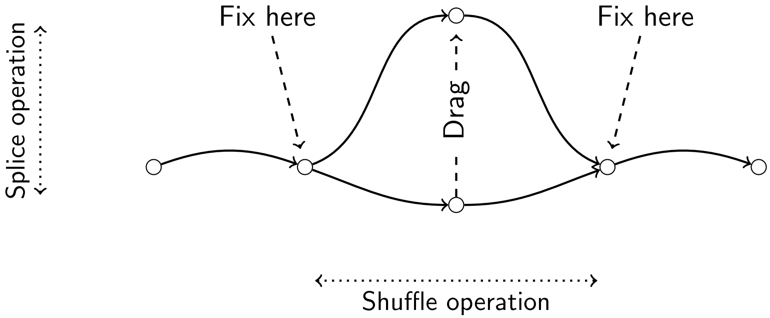
- Utility of activity pattern:
  - **time-of-day preferences**
  - **satiation effects**: marginal utility decreases with increasing duration

$$V(duration) = \eta \ln(duration)$$

  - 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

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

  with μ 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;
  $\zeta \rightarrow \infty$ sample paths independently of their cost.

Utility function

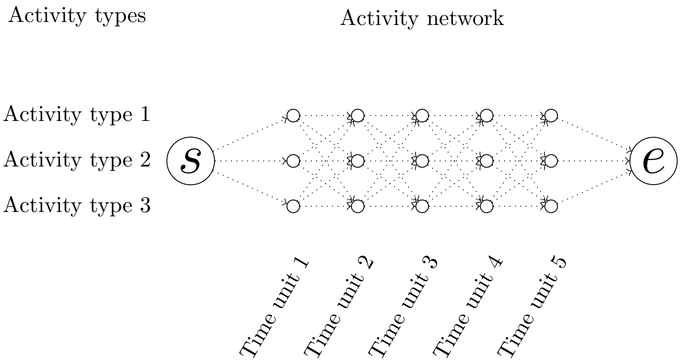
- Utility of activity path Γ with correction term for importance sampling:

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

- Fix one β to 0 for identification.
- Application to WiFi traces on a campus: Danalet and Bierlaire (2014)

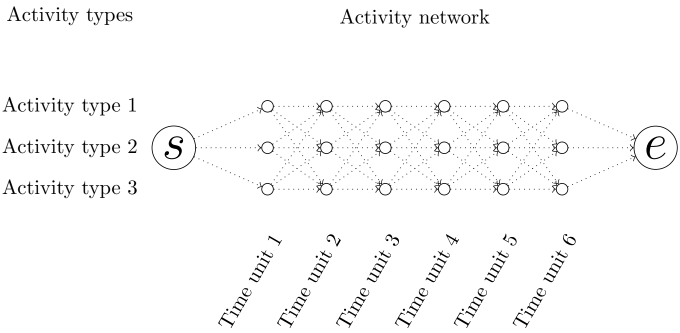
(Frejinger et al.; 2009)

Activity network



243 alternatives

Activity network



729 alternatives

Time-of-day preference is Cauchy distributed

- Utility of activity pattern:
  - **time-of-day preferences**: symmetrical Cauchy distribution

$$V'(\tau) = \frac{V_{max}}{c\pi\left(\left(\frac{\tau - b}{c}\right)^2 + 1\right)}$$

  - **satiation effects**: marginal utility decreases with increasing duration

$$V(duration) = \eta \ln(duration)$$

  - scheduling constraints: schedule delay

(Ettema et al.; 2004)

True values

Parameters	True values
$V_{max,1}$	3.0
b_1	2.5
c_1	2.0
$V_{max,2}$	4.0
b_2	4.0
c_2	3.0
η_1	2.0
η_2	1.3
η_3	0.8
γ_e	-1.2
γ_l	-1.8

Estimation with full choice set

Description	Coeff. estimate	Robust Asympt. std. error	t-stat (true value)
$V_{max,1}$	3.25	0.322	0.78
b_1	2.42	0.104	0.77
c_1	2.11	0.190	0.58
$V_{max,2}$	3.91	0.723	0.12
b_2	4.34	0.370	0.92
c_2	3.18	0.646	0.28
η_1	1.98	0.0512	0.39
η_2	1.38	0.0477	1.68
η_3	0.792	0.0522	0.15

Number of observations = $10'000$

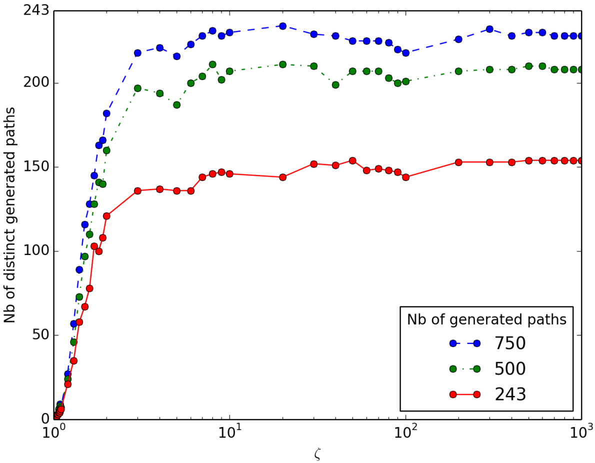
Importance sampling

- Utility function:

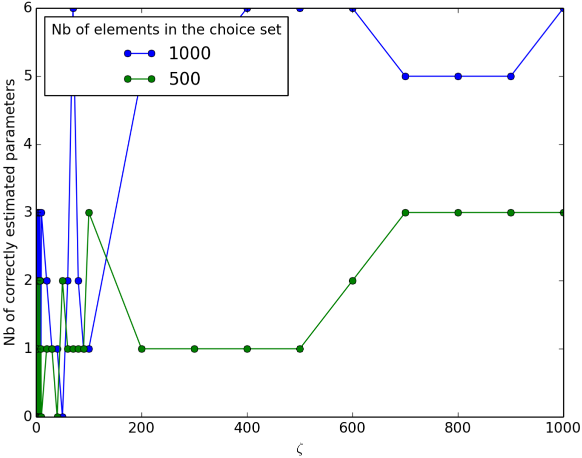
$$V_\Gamma = \mu \cdot \left(\sum_{k,\tau} \frac{V_{max}}{c\pi\left(\left(\frac{\tau-b}{c}\right)^2+1\right)} + \sum_{\text{episodes } e} \eta_k \ln(t_{k,e}) \right) + \ln \frac{k_\Gamma}{b(\Gamma)}$$

with the true value for one node fixed, and the scale μ estimated.

Number of distinct paths generated (only time of day, 5 time units)



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 ζ
    - The size of the choice set
- Gunnar's idea: Define the scale parameter ζ sequentially (Lemp and Kockelman; 2012): draw alternatives in proportion to updated choice-probability estimates

Thank you

Questions / suggestions?

References I

Danalet, A. and Bierlaire, M. (2014). A path choice approach to activity modeling with a pedestrian case study, *14th Swiss Transport Research Conference (STRC)*, Monte Verità, Ascona, Switzerland.
URL: *http://www.strc.ch/conferences/2014/Danalet_Bierlaire.pdf*

Ettema, D., Ashiru, O. and Polak, J. W. (2004). Modeling timing and duration of activities and trips in response to road-pricing policies, *Transportation Research Record* (1894): 1–10.

Ettema, D., Bastin, F., Polak, J. and Ashiru, O. (2007). Modelling the joint choice of activity timing and duration, *Transportation Research Part A* **41**(9): 827–841.
URL: *http://dx.doi.org/10.1016/j.tra.2007.03.001*

References II

Flötteröd, G. and Bierlaire, M. (2013). Metropolis–Hastings sampling of paths, *Transportation Research Part B* **48**: 53–66.
URL: *http://dx.doi.org/10.1016/j.trb.2012.11.002*

Frejinger, E., Bierlaire, M. and Ben-Akiva, M. (2009). Sampling of Alternatives for Route Choice Modeling, *Transportation Research Part B* **43**(10): 984–994.
URL: *http://dx.doi.org/10.1016/j.trb.2009.03.001*

Lemp, J. D. and Kockelman, K. M. (2012). Strategic sampling for large choice sets in estimation and application, *Transportation Research Part A: Policy and Practice* **46**(3): 602–613.
URL: *http://dx.doi.org/10.1016/j.tra.2011.11.004*