Importance sampling for activity path choice

Antonin Danalet

Michel Bierlaire

EPFL

2015

STRC | **15th Swiss Transport Research Conference**

Monte Verità / Ascona, April 15 – 17, 2015

EPFL

Importance sampling for activity path choice

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

phone: +41 21 693 25 32
fax:
{antonin.danalet,michel.bierlaire}@epfl.ch

2015

Abstract

We propose a model for the choice of an activity pattern. Models of activity participation patterns allow to assess the impact of demand management strategies on activity and destination choices.

In particular, we focus on choice set generation of activity patterns using recent developments in route choice modeling. Spatial choices deal with large choice sets. We develop a framework for choice set generation based on path choice. The activity-episode sequence is modeled as a path in an activity network defining the activity type, duration and time of day. The large dimensionality of the choice set is managed through an importance sampling based on Metropolis-Hastings algorithm.

Our model can be used to forecast demand at the urban scale and also in pedestrian facilities, such as transport hubs or mass gathering. Validation of the approach is performed on synthetic data and a case study using WiFi traces on a campus is presented.

Keywords

Activity-based modeling; pedestrians; choice set generation; multimodal transport hub; train station; airport

Illustration on front page: EPFL campus pedestrian network

Contents

Notations

a_{ψ_i}	an activity episode, $a_{\psi_i} = (x, t^-, t^+)$, for individual i
$a_{1:\Psi_i} = (a_1, ..., a_{\Psi_i})$	an observed activity-episode sequence
$att(x, t)$	the attractivity for location $x \in POI$ at time t
A_ψ	an activity, $A_\psi = (\mathcal{A}_k, t^-, t^+)$
$A_{1:\Psi_i} = (A_1, ..., A_\psi, ..., A_{\Psi_i})$	an activity pattern with Ψ_i activities, indexed by ψ
$A(a_{\psi_i})$	a function $a_{\psi_i} \mapsto A(a_{\psi_i}) = \mathcal{A}_k \in \{\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_K\}$
$\mathcal{A}_k$	an activity type, $k \in K$
$\mathcal{A}_{k,\tau}$	a node in the activity network, corresponding to activity type $\mathcal{A}_k$ and unit of time τ
$\mathcal{A}_{1:T}$	an activity path, i.e., a representation of an activity pattern $A_{1:\Psi_i}$ in an activity network
α	the parameter associated to travel time in the schedule delay approach
$b(\Gamma)$	the unnormalized target weights for path Γ
b_k	the preferred time for activity type $\mathcal{A}_k$ in the Cauchy distribution for time-of-day preference in the activity utility
DDR	the domain of data relevance
β	the parameters of the choice model
c_k	the width of the Cauchy distribution for time-of-day preference in the activity utility for activity type $\mathcal{A}_k$
C_i	the choice set for an individual i
d	distance between the iterations in the Metropolis-Hastings algorithm
δ	a cost function
$\delta_v(v)$	the cost of node v
$\delta_\Gamma(\Gamma)$	the non-link-additive cost of path Γ
$\delta(\Gamma)$	the generalized cost of path Γ
e	the end node of an activity network
$\mathcal{E}$	the set of edges in $SERG$
η_k	the parameters for satiation for activity type $\mathcal{A}_k$
f	the labeling function, $f : \mathcal{N} \to \tilde{\mathcal{L}}$, in $SERG$
g	a function associating nodes with coordinates in a coordinate system, $g : \mathcal{N} \to \mathbb{R} \times \mathbb{R} \times \mathbb{R}$
γ_e, γ_l	the parameters for early and late schedule delay
Γ	a path in the activity network

i	an individual
I	the set of all individuals in the period of interest
j	the index of a measurement $\hat{m}_j$
J	the total number of measurements
k	the index of an activity type.
$k_{\Gamma n}$	the number of times activity path Γ is drawn
K	the number of activity types.
L	the number of different activity-episode sequences $a_{1:K_i}$ corresponding to observation i
$\mathcal{L}_i$	the set of all activity patterns corresponding to observation i
$\tilde{\mathcal{L}}$	a set of relevant labels for rooms in $SERG$
$\hat{m}$	a raw measurement, containing location $\hat{x}$ and timestamp $\hat{t}$
$\hat{m}_{1:J}$	a set of measurements $\hat{m}_j$
μ	scale parameter for the Metropolis-Hastings algorithm, $\mu \geqslant 0$, $\mu = \frac{\ln 2}{(\zeta-1)\delta_{SP}}$
n	a node in SERG, $n \in \mathcal{N}$
$\mathcal{N}$	the set of all nodes in $SERG$
POI	the set of points of interest, $POI \in \mathcal{N}$
$\mathcal{P}_i$	the set of all candidate activity paths for observation i
$\mathcal{P}_I$	the set of all candidate activity paths for all individuals $i \in I$ in the period of interest.
$\mathcal{P}_{A_{1:\Psi_i}}$	the set of candidate paths corresponding to the activity pattern $A_{1:\Psi_i}$
ψ_i	the index of a activity episode a_{ψ_i}.
Ψ_i	the total number of episodes a_{ψ_i} in the activity-episode sequence $a_{1:\Psi_i}$ and the total number of activities A in the activity pattern $A_{1:\Psi_i}$. Ψ_i is individual specific.
$q(j)$	the sampling probability
s	the start node of an activity network
$S_{x,i}(t)$	the instantaneous potential attractivity measure in location $x \in POI$ at time t for individual i
$S_{x,i}(t^-, t^+)$	the potential attractivity measure in location $x \in POI$ between start time t^- and end time t^+ for individual i
$S_{\mathcal{A}_k,\tau}$	the potential attractivity measure for activity type $\mathcal{A}_k$ at time interval τ for individual i, corresponding to node $\mathcal{A}_{k,t}$
$sched_{x,i}(t)$	a dummy variable for time constraints in location $x \in POI$ at time t for individual i

SDE, SDL	early and late schedule delay, defined as $SDE = \max(t^* - t^a, 0)$ and $SDL = \max(t^a - t^*, 0)$
$SERG$	a semantically-enriched routing graph, $SERG := (\mathcal{N}, \mathcal{E}, \mathcal{L}, f, g, POI)$
t	time, continuous
τ	a discrete unit of time in the activity network, $\tau \in 1, 2, ..., T$. τ can also be seen as a time interval between τ_{LB} and τ_{UB}
τ_{LB}	the lower bound of the time interval τ
τ_{UB}	the upper bound of the time interval τ
T	the total number of units of time τ
T_{min}	a minimum time threshold for activity episodes (typically 5 minutes in a pedestrian context)
$\hat{t}$	a timestamp of a raw measurement $\hat{m}$
t^a	the actual arrival time in the schedule delay approach
t^-	the start time of an activity episode a, a continuous random variable
t^+	the end time of an activity episode a, a continuous random variable
t^*	the preferred arrival time in the schedule delay approach
tt	the travel time to destination in the schedule delay approach
$tt_{x_\psi, x_{\psi+1}}$	the travel time from x_ψ to $x_{\psi+1}$.
$\mathcal{U}$	the choice set corresponding to the activity network
v	a node in the activity path Γ, $v \in \Gamma$
$V_{\max,k}$	the scale of the Cauchy distribution used for the time-of-day preference part of the activity utility
$w(\Gamma)$	the target weight of Γ
x	the episode location, $x \in POI$
$\hat{x}$	the position of a raw measurement $\hat{m}$. $\hat{x} \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ (x-y coordinates in a coordinate system, and floor or altitude in a multi-floor environment).
ζ	scale-invariant parameter for the Metropolis-Hastings algorithm, $\zeta \geqslant 1$, $\mu = \frac{\ln 2}{(\zeta - 1)\delta_{SP}}$

1 Introduction

In this paper, we propose a model for the choice of an activity pattern. We focus on choice set generation of activity patterns using recent developments in route choice modeling. Our model can be used to forecast demand at the urban scale and also in pedestrian facilities, such as transport hubs or mass gathering.

Activity-based approach is motivated by the fact that the choice of an activity pattern triggers the choice of destinations (Bierlaire and Robin, 2009). It is relevant for transportation policy simulation such as congestion pricing, toll lanes, and changing schedules for work or shops (Davidson *et al.,* 2007). It is also relevant in models of pedestrian movement (Papadimitriou *et al.,* 2009).

Several models accounting for interactions that shape participation in different activities have been proposed. Activity scheduling model over an entire-day framework is a mixture of rule-based algorithm, duration models and discrete choice frameworks. The main drawback of most of these models is the postulated rules: they are structured on home and tours from home, with models applied sequentially according to priorities of activity types. Very often, the large dimensionality of the problem (activity types, continuous time, number of episodes in the day) implies aggregation or hierarchy of dimensions (broad periods of time, mandatory vs non mandatory, primary vs secondary).

Our modeling approach is not tour-based and does not assume any priorities between activities. It can be applied to weekdays, weekends, or even activities in a pedestrian facility, such as an airport or a supermarket. Our only assumption is that people first choose the activity type, trip timing and activity duration in their activity patterns (strategic level). Only then, conditionally on the activity type and time of day, they choose their specific destination (tactical level). Behaviorally, activity choice usually precedes destination choice (e.g., Bowman and Ben-Akiva, 2001, Arentze and Timmermans, 2004, Abou-Zeid and Ben-Akiva, 2012, Kang and Recker, 2013). Dimensionally, considering activity and destination as sequential choices makes the problem tractable. In this paper, we focus on activity type, duration and time-of-day choices. Examples of destination choice conditional on the activity type are Arentze *et al.* (2013) and Horni (2013), and examples in pedestrian facilities can be found in Ton (2014), Kalakou and Moura (2014) and Tinguely *et al.* (2015).

The chosen alternative is one sequence of activity episodes; utility is associated to the full pattern. Represented as a path in a network, the sequence is a single choice, contrary to Pinjari and Bhat (2010) who consider the activity-episode sequence as multiple choices of activity types and duration.

In this context, challenges of activity choice modeling include the definition of the alternative in the choice set, the choice set generation process and the utility structure for the given definition of alternatives. Former approaches of these challenges are presented in Section 2, while our methodology is developed in Section 3. The methodology is further applied to synthetic data (Section 4) and exemplified with WiFi traces on a campus (Section 5).

2 Literature review

This review presents some relevant aspects of activity choice for our approach. Concepts from route choice literature are also used in this paper. A review of choice set generation in route choice literature can be found in Danalet and Bierlaire (2014). General reviews about activity-based travel demand modeling can be found in Ettema (1996), Bhat and Koppelman (1999), Roorda (2005), Habib (2007), Bowman (2009), Feil (2010), Pinjari and Bhat (2011), Miller (2014) and Rasouli and Timmermans (2014).

All models of activity choice are built on the assumption that demand for traveling stems from a demand for activity participation, and understanding activity patterns is important for demand modeling. The goal of these models is to improve the traditional 4-step model by considering activities instead of trips (see for instance Bowman, 2009, Pinjari and Bhat, 2011). They model the interactions between time and space, trying to limit the independence assumption of the different elements of the choice.

This field of research faces challenges such as modeling social interactions (coupling constraints for household activities (e.g., Ho and Mulley, 2013, Gupta and Vovsha, 2013) or group activities from the social network), dynamic modeling demand across several days (e.g., Nijland *et al.*, 2014), or modeling matching of demand and supply (i.e., congestion, as in Bradley *et al.* (2010), Ouyang *et al.* (2011), Habib *et al.* (2013)). We previously reviewed the use of big data for activity modeling (in particular unstructured data) and the application of activity-based models for pedestrians in Danalet and Bierlaire (2014).

The challenge we address in this paper is how the time representation impacts the modeling of choice set generation in activity-based modeling. Different representations of time and activity types drive different strategies to manage the very large number of alternatives. In this review of activity-based travel demand modeling, we focus on the time representation (Section 2.1), the choice set generation (Section 2.2), the formulation of the utility function (Section 2.3) and the correlation between activity patterns (Section 2.4).

2.1 Time representation in activity modeling

Time is the most important dimension in modeling activity-travel behavior (Yamamoto and Kitamura, 1999). Its adequate representation is a prerequisite for accurate forecasting (Hägerstraand, 1970, Pinjari and Bhat, 2011). Compared to trip-based models, scheduling models see time not as a cost, but as a resource being used (Bhat, 2005). Time can be represented as continuous, decomposed in tours from home, or as the chronological activity episodes (e.g., Li and Lee, 2014).

An example of continuous time is the dynamic discrete-continuous choice model by Habib (2011). In this model, the decision maker sequentially chooses an activity type, and then its duration, constrained by a given time budget and the expected utility for the remaining time. This model is myopic, in the sense that it assumes that decision makers choose activities sequentially, without planning later activities. The concept of composite activity covering the rest of the day does not allow to model the choice of an overall pattern of activities with planning behavior. The multiple discrete-continuous extreme value models (Bhat, 2005, Pinjari and Bhat, 2011, Wang and Li, 2011) represent time as continuous and as a constraint (time budget). The decision maker chooses several activity types and allocates to each of them a corresponding time-use, so that it sums to the total time budget. The order in which it is consumed is not modeled.

In the activity-schedule approach (Ben-Akiva *et al.*, 1996, Bowman, 1998, Bowman and Ben-Akiva, 2001, Bradley *et al.*, 2010, Shiftan and Ben-Akiva, 2011, Gupta and Vovsha, 2013), the fundamental unit of time is a tour. A tour is a way to decompose the available time in a day in manageable units with duration. The behavior is modeled as two sequential decisions, the activity pattern and the tours. Only the tour models include information about timing, through (1) the tour time of day choice model, modeling the tour primary destination arrival and departure times, and (2) the trip departure time choice model, modeling some intermediate stop arrival and departure times for trips on the tour (Childress, 2010, Abou-Zeid and Ben-Akiva, 2012). Tours, primary destination of the tours and subtrips on the tours are sequentially modeling the timing of activities in order to simplify the model and reduce the size of the choice set.

The time representation in activity modeling is generally continuous or tour-based. In continuous time, the time representation is close to reality but misses an overall pattern choice (choice of time expenditure per activity type without order in activity episodes (e.g., Pinjari and Bhat, 2011) or per activity episode without pattern utility (Habib, 2011) but not of choice of a sequence of activity episodes). With tours, the choice of pattern is explicit but independent of the timing decision (start time and duration). It assumes a primary activity of the tour and a main mode (e.g., Shiftan and Ben-Akiva, 2011).

2.2 Choice set generation in activity choice

Choice set specification has received attention in different fields such as route choice or residential location choice (Ben-Akiva and Boccara, 1995, Swait, 2001, Başar and Bhat, 2004, de Lapparent, 2009, Rasouli *et al.,* 2013). A general review about choice set generation in spatial context can be found in Pagliara and Timmermans (2009a,b).

Activity scheduling consists of two steps: generating the choice set and making a choice among this set (Liao *et al.,* 2013). The definition of the choice set for each decision is a major weakness of the approach (Kang and Recker, 2013). The choice set generation in the context of activity modeling is complex and fundamental to have unbiased estimates for the parameters of the model, but data about the actual choice-set are usually missing. Moreover, the choice set containing all combinations of activity episodes is large (Bowman, 1998, p.74).

Most research in the field is applied to simpler problems. The size of the choice set is reduced through limitations in what enters the choice set and different assumptions related to the representation of time.

For each activity type $k \in \{1, ..., K\}$ (e.g., shopping), the activity-schedule approach considers two choices: (1) Is there a home-based tour with this activity type as the primary destination? (2) Are there secondary stops on this tour? Thus, at a maximum, the choice set for the activity pattern consists of 2^{2K} alternatives (Abou-Zeid and Ben-Akiva, 2012). It is further reduced using logical rules, such as the impossibility to have secondary stops if no corresponding primary tour is chosen. Different applications of the activity-schedule approach use different sizes of the choice set in the activity pattern model (Shiftan and Ben-Akiva, 2011). Nevertheless, they all use the home-based tour structure to reach a manageable size for the activity pattern choice set (48 elements in the San Francisco model, 114 in the Portland model (Shiftan and Ben-Akiva, 2011)). The timing of the tours and stops is estimated in submodels.

In the case of activity sequences (Li and Lee, 2014), the choice set contains K^M sequences, assuming a maximum number of possible activities in the day M. Timing is not considered and thus it decreases the size of the choice set. Some models consider a single activity pattern, typically home-work-home, with composite activity episodes for before and after work (Ettema *et al.,* 2007, Jenelius *et al.,* 2011). In these cases, the choice is only about timing. In the multiple discrete-continuous extreme value models, the choice set only contains K alternatives, one for each activity type, with decreasing utility with time.

As a general strategy, activity type and scheduling are often not considered simultaneously. By considering tours or sequential order, without timing, the size of the choice set decreases. When considering simultaneously activity type and scheduling, assuming that the period of interest contains T time units, the number of alternatives would be K^T and the choice set size explodes. Two answers are proposed in the literature: (1) Considering time as a continuous variable and using a multiple discrete-continuous approach (Bhat, 2005) or a dynamic discrete-continuous approach (Habib, 2011), or (2) Using the universal choice set containing all possible combinations of time units and activity types and performing importance sampling, as mentioned in Flötteröd and Bierlaire (2013). In this case, Lemp and Kockelman (2012) propose a first estimation using simple random sampling (SRS) and a second estimation drawing alternatives in proportion to the choice-probability estimates (not specifically for activity choice).

2.3 Formulation of utility function for activity sequences

Modeling the activity pattern formation requires a proper definition of the utility of an activity pattern. Bhat (2005) considers the overall utility on an individual as the sum of the utilities of each activity episode. This utility must reflect the time-of-day preferences, the fatigue effects and the scheduling constraints (Ettema *et al.,* 2007). The activity-schedule approach also defines the primary activity as the most important activity of the day (Bowman and Ben-Akiva, 2001).

Regarding one activity episode, Winston (1982) decomposes the utility of an episode as the satisfaction from performing the activity (process utility) and the satisfaction of completing the activity (goal-achievement utility). Nurul Habib and Miller (2009) decompose the activity choice in two parts and consider that the goal-achievement utility impacts activity program generation and process utility is included in the activity scheduling. In the activity-schedule approach, the model is decomposed in three levels: the activity pattern model, the tour-level models and the trip-level models (Shiftan and Ben-Akiva, 2011).

Our goal is to merge the estimation of the activity pattern and the tour-level models. Here, we describe the main variables entering the definition of the utility for sequences of activity episodes.

2.3.1 Time of day preference

The time-of-day element of the utility represents the variation in gain from performing the activity at different periods of time ("when"-dimension). Ettema and Timmermans (2003), Joh *et al.* (2004), Jenelius *et al.* (2011), Fu and Lam (2014) assume that the marginal utility follows an unimodal function, increasing first for a warming up phase and decreasing after reaching a saturation point. Ettema *et al.* (2004) propose to use a Cauchy distribution to express the marginal utility.

2.3.2 Satiation effect

The utility of an activity episode increases with time, while marginal utility of activity participation decreases (Yamamoto and Kitamura, 1999, Ettema *et al.*, 2007, Pinjari and Bhat, 2010, Habib, 2011). Satiation[1] expresses a fatigue effect ("how long"-dimension). In Habib (2011), the utility of time expenditure is multiplied by $\frac{1}{\alpha}(t^{\alpha} - 1)$, with α the satiation parameter, specified as $\alpha = 1 - \exp(-\tau y)$. In practice, α depends on a constant per activity type and on the time-of-day. Ettema *et al.* (2007) use a logarithmic function for the fatigue effect, $\eta_k \ln(t)$. η_k is specific to the activity type. The corresponding marginal utility is $\frac{\eta_k}{t}$.

2.3.3 Schedule constraints

Time of day preference and fatigue effect have been used in several papers (Ettema and Timmermans, 2003, Jenelius *et al.*, 2011). Ettema *et al.* (2007) add schedule constraints in the utility function, inspired by Small (1982). First introduced by Vickrey (1969), scheduling costs explain the choice of departure time (e.g., Arnott *et al.*, 1990, 1993). Small (1982) defines the utility of trip departure time as a function of travel time and schedule delay:

$$V = \alpha tt + \gamma_e \mathrm{SDE} + \gamma_l \mathrm{SDL} \tag{1}$$

where tt is the travel time, SDE is the early schedule delay and SDL is the late schedule delay. Schedule delays are defined as the difference between the preferred arrival time t^* and the actual

[1]Researchers in the transportation research community use "satiation" in order to express the decreasing marginal returns in activity modeling literature. We comply with this usage. Still, it must not be confused with the non-satiation assumption of preferences in standard microeconomics, stating that more is always better, or more precisely that, regardless of the individual's comsumption, an arbitrarily small quantity of a good is generating positive utility. For a proper use of "satiation" in transportation research, see Kockelman (1998, 2001).

arrival time t^a:

$$\begin{aligned} \text{SDE} &= \max(t^* - t^a, 0) \quad (2) \\ \text{SDL} &= \max(t^a - t^*, 0) \quad (3) \end{aligned}$$

Schedule delay approach introduces constraints in the schedule, such as a train departure time, preferred time to start working, or beginning of courses. They are a mathematical expression of coupling constraints as defined by Hägerstraand (1970). Ettema *et al.* (2007) and Hess *et al.* (2007) include schedule delay in models of choice of activity patterns. This approach has some limitations. It assumes a linear, yet asymmetric, effect of being early or late. More recent works propose non-constant marginal utilities for the scheduling preferences Tseng and Verhoef (2008). It needs anchor points t^* to be known.

2.4 Correlation between activity patterns

The correlation among alternatives (i.e., paths) is a well-known issue in route choice modeling when using random utility models. The similarity between two paths is usually measured as physical overlap (Vovsha and Bekhor, 1998). Frejinger and Bierlaire (2007) decompose the strategies to address this issue in two categories of models: deterministic correction (C-Logit models (Cascetta *et al.,* 1996), Path Size Logit (Ben-Akiva and Bierlaire, 1999)) and explicit modeling of correlation in the error term (Cross Nested Logit (Lai and Bierlaire, 2014), Error Component models (Bolduc and Ben-Akiva, 1991)). The first category is the most frequent and the simpler to compute.

The Path Size Logit consists of including an attribute, called Path Size (PS), in the deterministic part of the utility, in order to correct for overlapping paths. It is derived from aggregation of alternatives (Ben-Akiva and Lerman, 1985, ch.9), where the elemental alternatives are the paths and the aggregate alternatives are the links. In the route choice context, the size of the aggregate alternatives, i.e., a link, equals the number of paths using the link (see Frejinger and Bierlaire (2007) for a detailed development).

In activity choice context, utility is related to time of day, satiation and schedule delay (Section 2.3). All these parameters are defined for a given activity type (e.g., shopping). In the Activity Schedule approach, Bowman (1998, p.25) mentions that patterns sharing primary purpose are probably correlated and let it as a future research. Primary activity is defined as the most important activity of the day, either by asking the respondents in a survey (Antonisse *et al.,* 1986) or by counting the number of activity episodes in a tour and using deterministic priority rules from other studies when not available from the survey (Bowman and Ben-Akiva, 2001). Note that when deterministic rules cannot discriminate between different activity purposes, the activity of longer duration is defined primary.

As a general conclusion, this literature review for models of activity choice shows the need for modeling an explicit pattern utility with timing dimension (start time and duration). Considering simultaneously the order, timing and duration of activity-episode patterns generates very large choice sets. Large choice sets have been managed with discrete-continuous approaches. Importance sampling techniques have been recently improved and are an interesting approach, but have not been implemented yet. The most fundamental component of activity pattern utility are the time-of-day preference for activity types and a decreasing marginal utility for a given activity episode. Since 1998, correlation between activity patterns have been detected as an issue but no answer has been given to it to our knowledge.

3 A path choice approach to activity modeling

Our approach to activity-based travel demand modeling decomposes the behavior in two steps. First, a path choice approach models how people choose their activity type in time, taking into account the type of activities, their sequence and their timing/duration. Once the activity type, sequence and timing are chosen, a second step consists in modeling destination choice (e.g., choosing a restaurant, knowing the activity type: eating). We present here only the first step, the choice of activity sequence. A destination choice model applied to the same data set can be found in Tinguely *et al.* (2015).

There are mathematical and behavioral motivations for this decomposition. Mathematically, the problem is complex. The number of possible destinations in a given area is usually large. The number of sequences of destinations is larger. Including the duration spent at destination makes the problem definitively too large and intractable. Behaviorally, the choice of activity type and time of day precedes the choice of destination.

3.1 Representation of activity sequence: activity network and path

An activity network represents the choice set and contains all possible activity sequences. It is discrete with respect to activity types $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_K$ and time $\tau \in 1, 2, ..., T$. It is composed of links and nodes. Nodes $\mathcal{A}_{k,\tau}$ represent the performance by the individual of an activity type k for a unit of time τ. At a given unit of time τ, the number of nodes represent the available activity types K. There are two special nodes, start node s and end node e. They represent the beginning and the end of the observed activity pattern. In total, the activity network contains $KT + 2$ nodes. Edges connect some nodes and represent the fact that they are successively performed. s is connected to all nodes at the first time unit. All nodes at the last time unit are connected to e. All nodes of a given time unit t are connected with the nodes corresponding to the next time unit $t + 1$; it represents the choice of changing activity type or maintaining the activity type for one more time unit. In total, the activity network contains $2K + K^2(T - 1)$ edges.

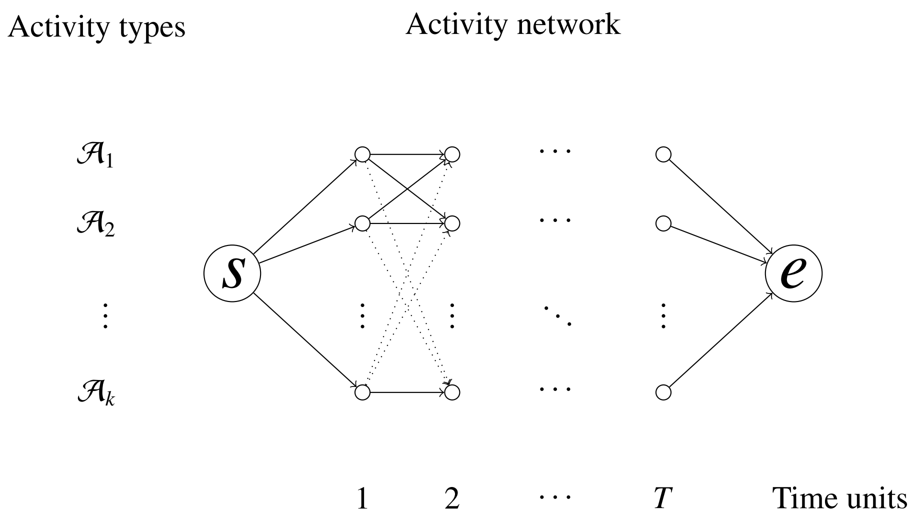

Figure 1: The activity network

The activity network is a representation of the universal choice set. The first and last time units represent the period of time under observation (e.g., 4am and midnight for a day).

Activity paths $\mathcal{A}_{1:T}$ are the representation of activity sequences in an activity network. Activity paths are the alternatives of the choice process. Typical attributes from the tour-based activity-pattern model, such as number of tours per day or number of stops in a tour, can easily be included in the utility of an activity path. Since we model all activity episodes at once when describing them as activity paths, our modeling approach gives the opportunity to interact activity pattern model attributes (e.g., number and purpose of tours, stops per tour) with tour-level models (e.g., timing of main activity).

3.2 Choice set generation

The universal choice set contains K^T alternatives. With a large temporal resolution T, a complete enumeration of all paths and the estimation of a choice model with the universal choice set is infeasible. Importance sampling allows to estimate discrete choice models in the case of an extremely large choice set but the sampling probabilities must be known in order to get consistent estimators.

The Metropolis-Hastings algorithm defined by Flötteröd and Bierlaire (2013) samples paths according to a given distribution without normalizing it. Only target weights are requested. A Markov chain with a predefined stationary distribution is generated by randomly modifying paths between an origin and a destination. Splice and shuffle operations on the paths are made such that the paths appear with the frequency specified by the weights.

Weights can be defined for each link, and then added up, or defined directly for the whole path. As explained in Frejinger and Bierlaire (2010), "*the sample should include attractive alternatives*" in order to provide efficient estimators. In the context of activity choice modeling, we propose to use the frequency of observed nodes and the frequency of activity-episode lengths in the activity network as a measurements of the attractive alternatives.

3.2.1 Node attractivity

The attractivity of a node in the activity network for choice set generation purpose is defined by using all observed paths in the network. In the route choice context, Chen (2013) calls it the observation score. It is defined here as the sum of all observed path through a given node v.

$$\sum_{\mathcal{A}_{1:T} \in \mathcal{P}} I(v \in \mathcal{A}_{1:T}) \tag{4}$$

where $I(v \in \mathcal{A}_{1:T})$ is an indicative function with value 1 if v is a node of activity path $\mathcal{A}_{1:T}$ and 0 otherwise, and $\mathcal{P}$ is the set of all observed activity paths. In order to keep the shortest path formulation from Flötteröd and Bierlaire (2013), node cost δ_v for node v can be defined as:

$$\delta_v(v) = \max_{\substack{k=1,\dots,K \\ \tau=1,\dots,T}} \left(\sum_{\mathcal{A}_{1:T} \in \mathcal{P}} I(\mathcal{A}_{k,\tau} \in \mathcal{A}_{1:T}) \right) - \sum_{\mathcal{A}_{1:T} \in \mathcal{P}} I(v \in \mathcal{A}_{1:T}) + 1 \tag{5}$$

The fact that the Metropolis-Hastings algorithm is based on shortest path is not an issue regarding the length of the activity path, since the number of time units for an activity path $\mathcal{A}_{1:T}$ is always T. The Metropolis-Hastings algorithm simply samples the activity paths with the most attractive nodes, i.e., the most attractive activity types at a given time.

3.2.2 Activity-episode length attractivity

Node attractivity does not suffice to define the most attractive and realistic paths. It defines the most attractive path as the most visited activity type at each time unit. Path sampling using only node attractivity risks to jump too often from one activity type to another, or on the contrary to stay in the same activity type for a long period, neglecting the satiation effect mentioned in the literature (see Section 2.3.2).

We define an activity episode $a = (\mathcal{A}_k, \tau^-, \tau^+)$ in the activity network as an activity type $\mathcal{A}_k$ that the user is performing from a start time τ^- to an end time τ^+. The length of an activity episode a is $|a| = \tau^+ - \tau^- + 1$. An activity path $\mathcal{A}_{1:T}$ contains several activity episodes $a_{\mathcal{A}_{1:T}}$.

The attractivity of an activity episode a is defined as the number of observed activity episodes $a_{\mathcal{A}_{1:T}}$ with same length $|a|$ and same activity type k in all observed paths $\mathcal{A}_{1:T} \in \mathcal{P}$ in the activity network:

$$\sum_{\mathcal{A}_{1:T} \in \mathcal{P}} \sum_{a_{\mathcal{A}_{1:T}} \in \mathcal{A}_{1:T}} I(|a| = |a_{\mathcal{A}_{1:T}}|) I(k = k_{\mathcal{A}_{1:T}}) \tag{6}$$

The activity-episode length attractivity of a generated path Γ is defined as the sum of the attractivity of each activity episode a_Γ in this path. Similarly than in Eq. 5, the shortest path formulation is maintained by defining path cost δ_Γ using the maximum of activity-episode length attractivity:

$$\begin{aligned} \delta_\Gamma(\Gamma) \quad = \quad & max\left(\sum_{a_\Gamma \in \Gamma} \sum_{\mathcal{A}_{1:T} \in \mathcal{P}} \sum_{a_{\mathcal{A}_{1:T}} \in \mathcal{A}_{1:T}} I(|a_\Gamma| = |a_{\mathcal{A}_{1:T}}|) I(k = k_{\mathcal{A}_{1:T}}) \right) \\ & - \sum_{a_\Gamma \in \Gamma} \sum_{\mathcal{A}_{1:T} \in \mathcal{P}} \sum_{a_{\mathcal{A}_{1:T}} \in \mathcal{A}_{1:T}} I(|a_\Gamma| = |a_{\mathcal{A}_{1:T}}|) I(k = k_{\mathcal{A}_{1:T}}) \end{aligned} \tag{7}$$

By adding the attractivity of activity episodes using path cost δ_Γ in the weights of the Metropolis-Hastings algorithm, the generated activity paths in the choice set approximate the observed behavior with respect to satiation. The total cost $\delta(\Gamma)$ is:

$$\delta(\Gamma) = \sum_{v \in \Gamma} \delta_v(v) + r\delta_\Gamma(\Gamma) \tag{8}$$

where r represents the ratio between path cost δ_Γ and node cost δ_v. This ratio is independent of the number of observations and of the number of time units.

The target weight is an geometrically decreasing function $\exp(-\mu\delta(\Gamma))$ of the total cost $\delta(\Gamma)$ of the path, as used in Flötteröd and Bierlaire (2013). μ is a scale parameter controlling for the deviation from the shortest path in the path sampling. If $\mu = 0$, the Metropolis-Hastings algorithm samples paths uniformly, independently of their cost; if $\mu \to \infty$, only the shortest path is sampled (with cost δ_{SP}). μ depends on the level of cost, varying with the number of activity types and of time units in the activity network, and with the total number of observations. As described in Flötteröd and Bierlaire (2013), the Metropolis-Hastings algorithm can sample a path of cost $\zeta\delta_{SP}$ with half the probability of the shortest path, using a single scale-invariant parameter ζ:

$$\mu = \frac{\ln 2}{(\zeta - 1)\delta_{SP}} \tag{9}$$

If $\zeta \to \infty$, the Metropolis-Hastings algorithm samples paths uniformly, independently of their cost; if $\zeta = 1$, only the shortest path is sampled. In this paper, we report the scale-invariant parameterζ.

It needs to be stressed that the activity network is cycle-free and the only path between the upstream node and the downstream node of every link is the link. Thus, it complies with the requirements in Flötteröd and Bierlaire (2013), independently of definition of the target weights. Another requirement in the algorithm implies that if two paths have the same weight, only one is sampled. The total cost for two different activity paths are equal with probability almost surely equal to zero. This situation never appears in the validation and the case study presented below.

For each observation, the choice set C_i contains the output of the Metropolis-Hastings sampling of paths and, if not already included, the chosen alternative.

3.3 Sampling correction in the utility

We propose a choice modeling framework for activity paths. This framework uses the choice set generated in the previous section and corrects for importance sampling. The structure of the utility function is described.

We assume the choice set to be the universal choice set containing all possible paths between s and e in the activity network. The sampling strategy for choice set generation presented in Section 3.2 requires the deterministic part of the utility to be corrected in order to estimate unbiased parameters (McFadden, 1978). According to Frejinger *et al.* (2009), a sampling correction term $\ln \frac{k_{\Gamma n}}{q(\Gamma)}$ must be added to the utility function for activity path Γ, where $k_{\Gamma n}$ is the number of times activity path Γ is drawn in C_i and $q(\Gamma)$ is the sampling probability of path Γ.

The sampling probability $q(\Gamma)$ is available using the unnormalized target weights $b(\Gamma)$ but require full enumeration for normalization: $q(\Gamma) = \frac{b(\Gamma)}{\sum_{\Gamma' \in \mathcal{U}} b(\Gamma')}$. In practice, the normalizing sum cancels out in the logit formulation and $b(\Gamma)$ can be used instead of $q(\Gamma)$.

The utility usually includes the time of day preference, the satiation effect and the schedule delay, as described in Section 2.3. In this case, the path utility depends on the utilities of individual nodes $\mathcal{A}_{k,\tau}$ and on the utilities of the activity episodes a of the activity path. The utility $V(\mathcal{A}_{k,\tau})$ of a node $\mathcal{A}_{k,\tau}$ represents the individual marginal utility from allocating one time unit to a certain activity type. It corresponds to the time-of-day utility and depends on both the activity type k and the time interval τ. It can be generally expressed as $\beta_{k,\tau} I_{k,\tau}$, where $I_{k,\tau}$ is a dummy variable (with value 1 if the activity path include node $\mathcal{A}_{k,\tau}$ and 0 otherwise) and $\beta_{k,\tau}$ is the corresponding parameter. In practice, some β's might be equal. The utility $V(a)$ of an activity episode a includes the satiation effect and the schedule delay.

Our modeling framework also gives the opportunity to add attributes that are not link-additive, such as the repetition of certain activity types in the activity path, or the structure of the activity path.

Parameters can be interacted with socioeconomic variables. For identification purpose, the time-of-day preference parameters $\beta_{k,\tau}$ for one activity type k must be fixed.

The deterministic part of the utility correcting for the sampling of alternatives is:

$$\mu\left(\sum_{k=1}^{K}\sum_{\tau=1}^{T} V(\mathcal{A}_{k,\tau}) + \sum_{a \in \mathcal{A}_{1:T}} V(a) + V(\Gamma)\right) + \ln \frac{k_{\Gamma n}}{b(\Gamma)} \tag{10}$$

Concrete instances of node utility $V(\mathcal{A}_{k,\tau})$, activity-episode utility $V(a)$ and activity path utility $V(\Gamma)$ are presented in the validation section and in the case study below.

3.4 Activity path size for correlation between activity paths

The Multinomial Logit (MNL) model is restricted by the Independent from Irrelevant Alternative (IIA) property. It might not be appropriate in the activity path approach since activity paths share unobserved attributes due to overlaps. Overlaps correspond to performing the same activity type at the same time and might be correlated. In this section, we first show that the traditional Path Size correction term cancels out in the activity path approach and then propose two different deterministic corrections for correlation, the Primary Activity Path Size (PAPS) and the Activity Pattern Path Size (APPS), for the utility of overlapping activity paths.

The choice of an activity path can be seen as an aggregation of alternatives (about aggregation of alternatives, see Ben-Akiva and Lerman, 1985, ch.9). The elemental alternatives are the activity paths, and the aggregate alternatives are the nodes in the activity network. In the route choice context, the Path Size formulation defines the size of an aggregate alternative (i.e., a link) as the number of paths using the link (Frejinger and Bierlaire, 2007). The motivation for this definition of size is related to the assumption that a physical overlap "measures" shared unobserved attributes. In the activity path context, the number of paths using a node is constant, K^{T-1}, due to the structure of the activity network, and so the traditional Path Size formulation from route choice leads to a constant correction term for each alternative and consequently no correction. Fundamentally, this result comes from the symmetry of the activity network and the hypothesis that duration of the same activity type at the same time of day is a measure of similarity, i.e., one replaces physical shared length (overlap) from the route choice context by shared time length (duration) for a given activity type at a given time of day.

Let's assume that similarity for activity patterns is measured through shared primary activity purpose and pattern (Bowman, 1998). Two formulations are possible, the Primary Activity Path Size (PAPS) and the Activity Pattern Path Size (APPS).

3.4.1 The Primary Activity Path Size (PAPS)

Each path has a primary activity $\mathcal{A}_p$, defined as the relative majority of nodes, i.e., the longest duration. Let's define C_i as the choice set of individual i. We denote the subset $C_{p,i} \subset C_i$ of paths with primary activity $\mathcal{A}_p$. We define the size $M_{\mathcal{A}_{k,\tau}}$ of node $\mathcal{A}_{k,\tau}$ as the number of paths using node $\mathcal{A}_{k,\tau}$ with primary activity being $\mathcal{A}_p$:

$$M_{\mathcal{A}_{k,\tau}} = \sum_{\Gamma \in C_p} \delta_{\mathcal{A}_{k,\tau},\Gamma} \tag{11}$$

where $\delta_{\mathcal{A}_{k,\tau},\Gamma}$ is the node-path incidence variable, with value 1 if path Γ contains node $\mathcal{A}_{k,\tau}$ and 0 otherwise.

Computing $M_{\mathcal{A}_{k,\tau}}$ requires first to determine if node $\mathcal{A}_{k,\tau}$ corresponds to the primary activity $\mathcal{A}_p$

(i.e., $k = \mathcal{A}_p$). If this is the case,

$$M_{\mathcal{A}_{k,\tau}} = \left[\frac{x^{T-1}}{(T-1)!}\right] \sum_{j\geq 0} \frac{x^{j-1}}{(j-1)!} \left(1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{j-1}}{(j-1)!}\right)^{K-1} \tag{12}$$

where the notation $[z^n]\, f(z)$ denotes the operation of extracting the coefficient of z^n in the formal power series $f(z) = \sum f_n z^n$ in analytical combinatorics (Flajolet and Sedgewick, 2002, p.19). By the theory of exponential generating functions, the number of paths with a relative majority of nodes corresponding to the primary activity $\mathcal{A}_p$ and a fixed node corresponding to the primary activity $\mathcal{A}_p$ can be determined by choosing $T-1$ nodes with $j-1$ of these nodes being the primary activity $\mathcal{A}_p$ and a maximum of $j-1$ other activity types for the remaining nodes. In this way, there are j nodes corresponding to primary activity $\mathcal{A}_p$ and a maximum of $j-1$ nodes with one of $K-1$ other activity types.

In the case when node $\mathcal{A}_{k,\tau}$ does not correspond to the primary activity $\mathcal{A}_p$, the size $M_{\mathcal{A}_{k,\tau}}$ of node $\mathcal{A}_{k,\tau}$ can be computed using the same reasoning: choosing $T-1$ nodes with j of these nodes being the primary activity $\mathcal{A}_p$, a maximum of $j-2$ for the activity type k of the node $\mathcal{A}_{k,\tau}$ and a maximum of $j-1$ nodes for the $K-2$ remaining activity types:

$$M_{\mathcal{A}_{k,\tau}} = \left[\frac{x^{T-1}}{(T-1)!}\right] \sum_{j\geq 0} \frac{x^{j}}{j!} \left(1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{j-1}}{(j-1)!}\right)^{K-2} \left(1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{j-2}}{(j-2)!}\right) \tag{13}$$

According to the theory on aggregation of alternatives, the utility associated with node $\mathcal{A}_{k,\tau}$ is defined by

$$U_{\mathcal{A}_{k,\tau},i} = \bar{V}_{\mathcal{A}_{k,\tau},i} + \frac{1}{\mu} \ln M_{\mathcal{A}_{k,\tau}} + \varepsilon_{\mathcal{A}_{k,\tau},i} \tag{14}$$

where $\bar{V}_{\mathcal{A}_{k,\tau},i} = \frac{1}{M_{\mathcal{A}_{k,\tau}}} \sum_{\Gamma\in C_p} V_{\Gamma,i}$ is the average deterministic utility of paths using node $\mathcal{A}_{k,\tau}$.

The contribution of a node $\mathcal{A}_{k,\tau}$ for size correction of a path is $\frac{1}{\mu} \ln \frac{1}{M_{\mathcal{A}_{k,\tau}}}$. When assuming that the size is proportional to the duration of the path, we define the Primary Activity Path Size (PAPS) as

$$PAPS_{\mathcal{A}_{1:T}} = \frac{1}{T} \sum_{\mathcal{A}_{k,\tau}\in\mathcal{A}_{1:T}} \frac{1}{M_{\mathcal{A}_{k,\tau}}} \tag{15}$$

Making the size proportional to the duration introduces a term $\frac{1}{T}$ and Eq. 10 becomes

$$\mu\left(\sum_{k=1}^{K}\sum_{\tau=1}^{T} V(\mathcal{A}_{k,\tau}) + \sum_{a\in\mathcal{A}_{1:T}} V(a) + V(\Gamma)\right) + \ln\frac{k_{\Gamma n}}{b(\Gamma)} + \beta_{PAPS}\ln PAPS_{\mathcal{A}_{k,\tau}} \tag{16}$$

with $\beta_{PAPS} > 0$.

3.4.2 The Activity Pattern Path Size (APPS)

Assuming that similarity for activity patterns is measured through shared pattern, correlation between paths can be deterministically corrected by using the number of paths with similar pattern through a node $\mathcal{A}_{k,\tau}$ as the size $M_{\mathcal{A}_{k,\tau}}$ of this node.

We define the pattern p of an activity path as the ordered sequence of activity types of the activity episodes, without taking duration into account, e.g., an activity path Home-Home-Home-Work-Work-Work-Work-Work-Shop-Home-Home corresponds to the activity pattern Home-Work-Shop-Home. Let's further define $|p|$ the number of elements in pattern p, $|p_k|$ the number of times activity type k appears in pattern p,

For the Activity Pattern Path Size, the size of node $\mathcal{A}_{k,\tau}$ is:

$$M_{\mathcal{A}_{k,\tau}} = \sum_{i=1}^{|p_k|}\binom{\tau-1}{L_i-1}\binom{T-\tau}{|p|-L_i} \tag{17}$$

where L_i is the index of the ith occurence of activity type k in pattern p. This can be seen as the count of decomposing the time units before and after the considered node into the subpatterns containing the pattern before and after the considered node.

The corresponding Activity Pattern Path Size is:

$$APPS_{\mathcal{A}_{1:T}} = \sum_{\mathcal{A}\in\mathcal{A}_{1:T}}\frac{1}{M_{\mathcal{A}_{k,\tau}}} \tag{18}$$

4 Validation with synthetic data

We present results with synthetic data, including how we generated the synthetic data, the activity network and the utility function used (Section 4.1), the sampling of paths (Section 4.2) and the results of the estimation using importance sampling (Section 4.3).

4.1 Synthetic data

The activity network used for the validation contains 3 activity types and 6 time units (Fig. 2). It contains 729 possible activity paths.

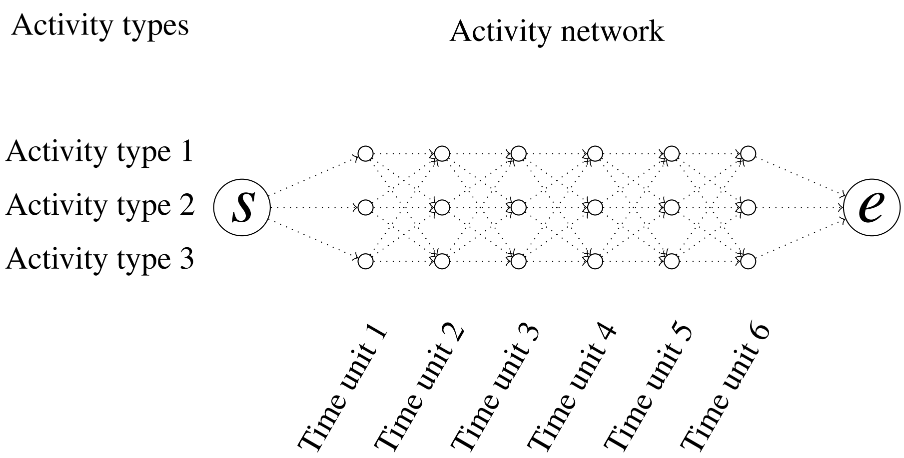

Figure 2: The activity network for validation.

Node utility $V(\mathcal{A}_{k,\tau}) = \sum_{k=1}^{K}\sum_{\tau=1}^{T}\beta_{k,\tau}I_{k,\tau}$ with $\beta_{1,1} = \beta_{1,2} = \beta_{1,5} = -0.5, \beta_{1,3} = \beta_{1,4} = \beta_{1,6} = 1.5$, $\beta_{2,1} = \beta_{2,2} = \beta_{2,3} = -1, \beta_{2,4} = \beta_{2,5} = \beta_{2,6} = 2$ and $\beta_{3,\tau} = 0\ \forall\tau$.

Satiation effect is expressed here as $\eta_k \ln(|a|)$, with $\eta_1 = 2$, $\eta_2 = 1.3$ and $\eta_3 = 0.8$. Early and late schedule delays for activity type 1 are computed with $\gamma_e = -1.2$ and $\gamma_l = -1.8$. The deterministic utility of activity path $\mathcal{A}_{1:T}$ is:

$$\begin{aligned}
V(\mathcal{A}_{1:T}) &= \sum_{k=1}^{K}\sum_{\tau=1}^{T} V(\mathcal{A}_{k,\tau}) + \sum_{a\in\mathcal{A}_{1:T}} V(a) \\
&= \sum_{k=1}^{K}\sum_{\tau=1}^{T}\beta_{k,\tau}I_{k,\tau} + \\
&\quad \sum_{a\in\mathcal{A}_{1:T}} \Big(\eta_k \ln(|a|) + \gamma_e \max(t^* - \tau^-, 0) + \gamma_l \max(\tau^- - t^*, 0)\Big)
\end{aligned} \tag{19}$$

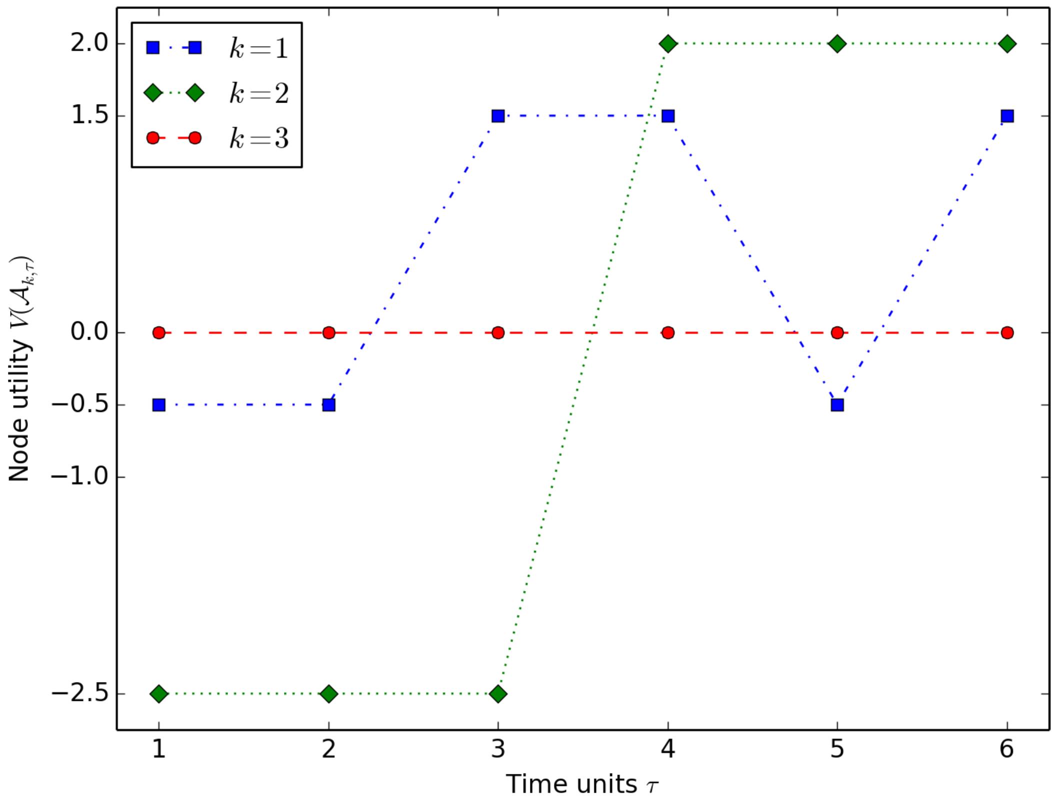

Figure 3: Value of node utility $V(\mathcal{A}_{k,\tau})$ as a function of activity type $k \in \{1, 2, 3\}$ and time unity $\tau \in \{1, 2, ..., 6\}$. Node utility is fixed to 0 for activity type 3 for identification purpose. Node utility is first low and then high for activity type 2 (e.g., morning and afternoon effect). Node utility is varying more for activity type 1 (e.g., lunch and dinner time in a restaurant).

9 parameters are defined in total. The preferred time t^* is randomly generated for each individual between the first and the last time units, $t^* = [1, 6]$. 2'000 synthetic observations of the chosen activity paths are generated assuming a logit model with utility as in Eq. 19 and $\varepsilon_{i\Gamma}$ distributed Extreme Value with scale 1 and location 0.

A model was estimated on these synthetic data with a full choice set (729 alternatives) in order to validate the generated synthetic data. Results are shown in Table 2. The last column is the t-test against the true value. It shows that the results are not significantly different from their true value, and therefore the synthetic data are valid.

Parameter	True value	Estimate	Std. error	t-test
$\beta_{\text{low},1}$	-0.5	-0.681	0.109	1.66
$\beta_{\text{high},1}$	1.5	1.42	0.0928	0.86
$\beta_{\text{low},2}$	-2.5	-2.63	0.103	1.26
$\beta_{\text{high},2}$	2.0	1.88	0.115	1.04
η_1	1.8	1.97	0.189	0.90
η_2	1.3	1.48	0.194	0.93
η_3	0.8	0.688	0.135	0.83
γ_e	-2.2	-2.15	0.0633	0.79
γ_l	-2.8	-2.82	0.141	0.14

Table 2: Estimated values of the parameters with full choice set and synthetic data (2'000 observations).

4.2 Sampling of paths

The full choice set can be used in this example, but in practice, with more activity types and more time units, it is not the case anymore. We sample paths with the Metropolis-Hastings algorithm described in Section 3.2, using total cost $\delta(\Gamma)$ (Eq. 8) to define the target weight, with $\zeta = 1.3$ and $r = 0.3$.

Each successive state of the MH algorithm is very likely to be similar to the previous one. This similarity stabilizes with the distance d between iterations. Stabilization means independence of the sampled paths. A similarity measure is defined in Flötteröd and Bierlaire (2013):

$$\phi(d) = \frac{1}{K} \sum_{\kappa=1}^{K} \frac{|\Gamma^k \cap \Gamma^{k+d}|}{\frac{1}{2}(|\Gamma^k| + |\Gamma^{k+d}|)} \tag{20}$$

where $|\Gamma^k \cap \Gamma^{k+d}|$ is the number of identical nodes in the paths generated in iterations k and $k+d$. Fig. 4 shows that the similarity is around 0.65 after a warming-up period. The similarity is still large once stabilized, because paths overlap very often in such a small network. Similarly to Flötteröd and Bierlaire (2013), we fit a linear regression model on 10 consecutive similarity values $\phi(d), ..., \phi(d+9)$ for consecutive distances $d, ..., d+9$. We assume the sampled paths to be independent when the absolute slope of the linear model is below 10^{-5}. For the synthetic data, independence is reached for a distance $d = 1047$ with $\zeta = 1.3$ and $r = 0.3$.

For the validation with synthetic data, we choose a conservative distance $d = 1200$ between iterations.

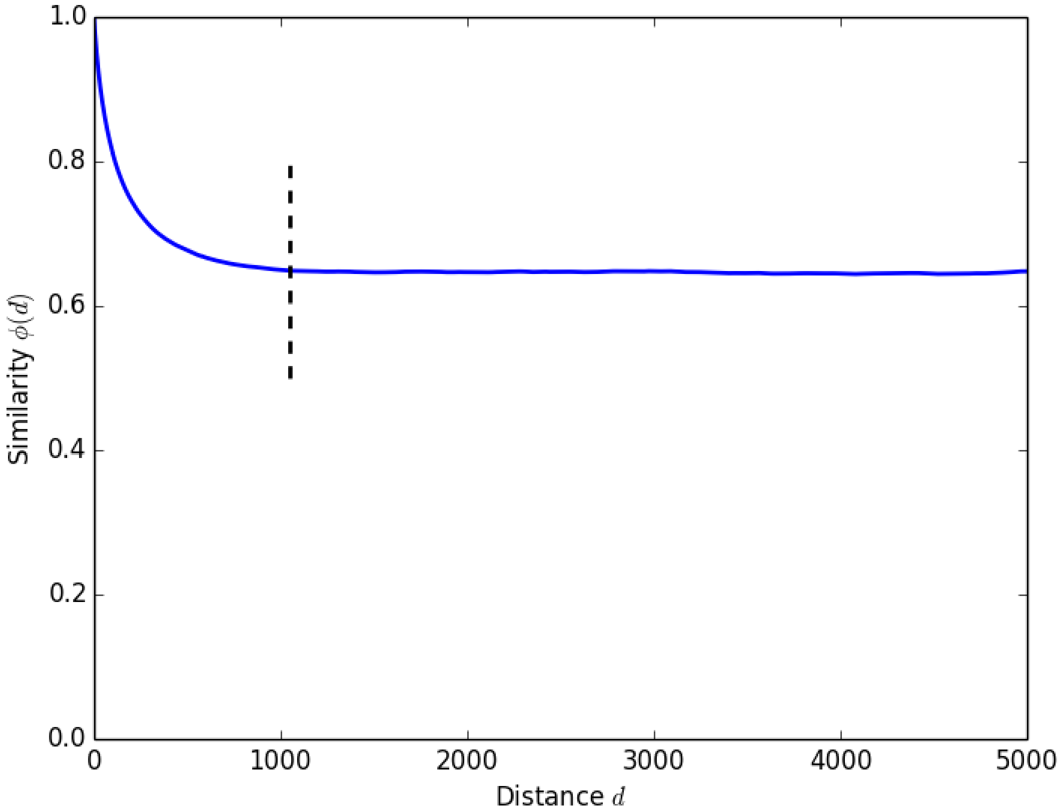

Figure 4: Similarity measure as a function of the distance. 10^6 paths were generated from the synthetic data, with $\zeta = 1.3$ and $r = 0.3$.

4.3 Estimation using importance sampling

We generate 20 elements in the choice set for each observation with $\zeta = 1.3$, $r = 0.3$ and $d = 1200$ and we estimate a logit with the following utility function:

$$\mu\Big(V(\mathcal{A}_{1:T})\Big) + \ln \frac{k_{\Gamma n}}{b(\Gamma)} \tag{21}$$

Note that the sampling correction term $\ln \frac{k_{\Gamma n}}{b(\Gamma)}$ is not multiplied by the scale parameter μ. Sampling correction does not depend on the variance of the data. For identification, we need to fix one parameter to its true value in $V(\mathcal{A}_{1:T})$ (say $\beta_{\text{low},1}$). By estimating μ, the sampling correction term is not scaled and the parameter estimates have the same scale than the true ones for comparison.

Results are evaluated by counting the number of parameters that are not significantly different from their real value. The total number of estimated parameters is 9. Results are shown in Table 3.

Parameter	True value	Estimate	Std. error	t-test
$\beta_{\text{high},1}$	1.5	1.73	0.623	0.37
$\beta_{\text{low},2}$	-2.5	-2.92	0.849	0.49
$\beta_{\text{high},2}$	2.0	1.9	0.633	0.16
η_1	1.8	2.3	0.577	0.87
η_2	1.3	2.01	0.65	1.09
η_3	0.8	0.893	0.402	0.23
γ_e	-2.2	-2.67	0.845	0.56
γ_l	-2.8	-3.54	1.11	0.67
μ	1.0	0.9	0.275	0.36

Table 3: Estimated values of the parameters with importance sampling and 20 generated elements in the choice set ($\zeta = 1.3$, $r = 0.3$).

All parameters are not significantly different from their true value. It shows that parameters can be properly estimated using the Metropolis-Hastings sampling of paths with node and activity-episode length attractivities and correcting for importance sampling.

4.4 Sensitivity analysis

This section shows the sensitivity of the results with respect to the number of elements in the choice set, the scale-invariant parameter for the Metropolis-Hastings sampling of paths ζ, and the ratio r between path cost and node cost.

Section 4.3 shows above that all parameters are not significantly different from their true values when generating 20 elements for the choice set ($\zeta = 1.3$, $r = 0.3$). Fig. 5 presents the results for choice sets with 1, 2, ..., 50 elements. For each number of elements included in the choice set between 1 and 50, the Metropolis-Hastings sampling of paths ran 5 times. With 1 element in the choice set, one of the 5 estimated models has 9 estimates that are significantly different from their true value; with 2 elements in the choice set, 2 of the 5 estimated models has 9 estimates that are significantly different from their true values; with 5 elements in the choice set, 1 estimated model has 9 estimates significantly different from their true values and 1 other model with 1 estimate significantly different from its true value. All 245 other models do not have estimates that are significantly different from their true values.

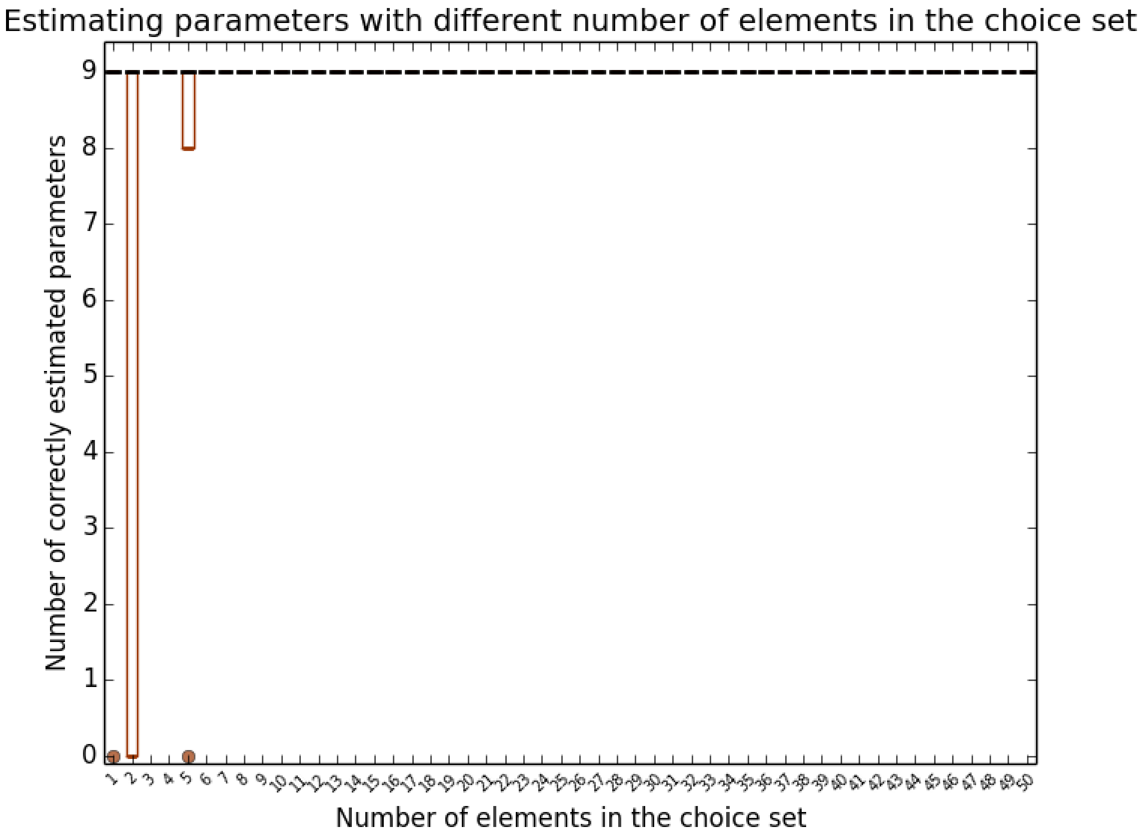

Figure 5: Number of non-significantly different parameters compared to their true values as a function of the number of elements in the choice set ($\zeta = 1.3$, $r = 0.3$). For each number of elements in the choice set per observation, 5 choice models were estimated based on different random outputs from the Metropolis-Hastings sampling of paths.

Figure 6 shows the number of non-significantly different parameters compared to their true values depending on the scale-invariant parameter for the Metropolis-Hastings sampling of paths ζ. For each value of $\zeta \in \{1.1, 1.2, ..., 1.9\}$, 5 models were estimated with 10 elements in the choice set per observation, 5 models were estimated with 30 elements in the choice set and 5 models were estimated with 50 elements in the choice set. When $\zeta \leqslant 1.2$, the sampled paths are very similar to the shortest path. There is not enough variety in the choice set to estimate properly the parameters. In this case, the scale of the choice model μ is underestimated. Oppositely, when $\zeta \geqslant 1.8$, importance sampling do not consider attractivity of the generated paths enough. With $\zeta = 1.3$, all parameters are non-significantly different compared to their true values. With 10 elements in the choice set, all parameters are properly estimated only for $\zeta = 1.3$, while with 20 or 50 elements in the choice set, zeta can vary between 1.3 and 1.5.

Figure 7 shows the number of non-significantly different parameters compared to their true values depending on the ratio r between path cost and node cost. For each value of $r \in \{0.1, 0.2, ..., 0.5\}$,

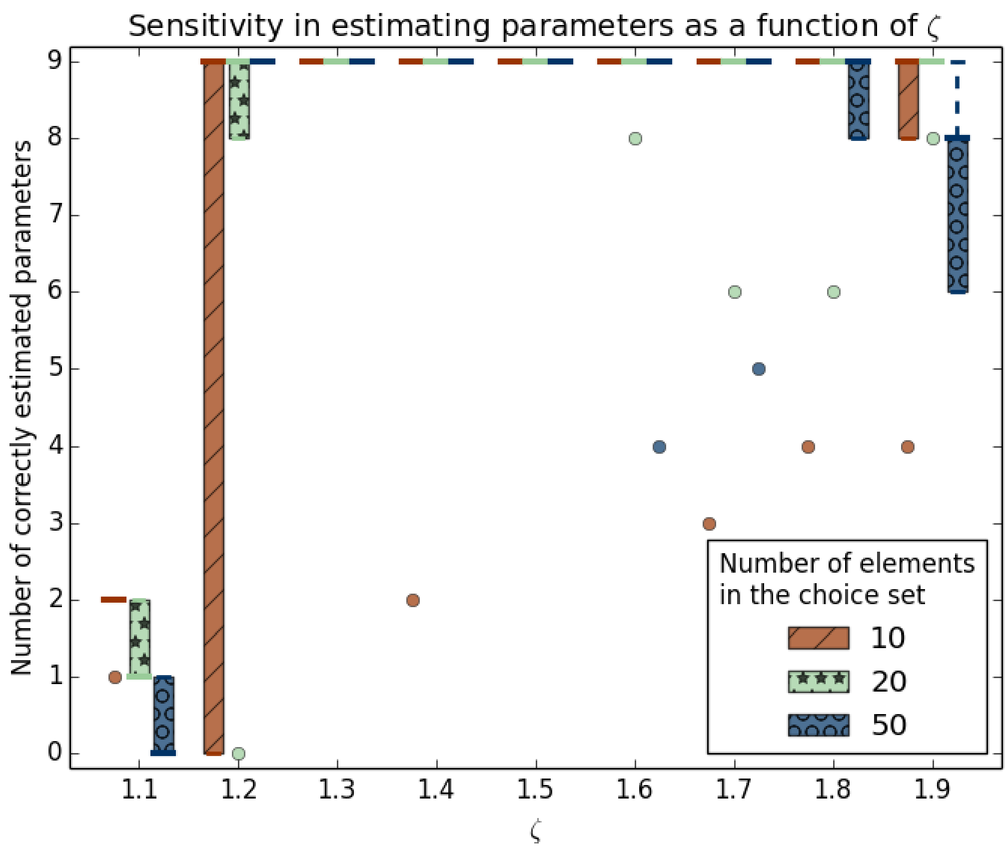

Figure 6: Number of non-significantly different parameters compared to their true values as a function of the scale-invariant parameter of the Metropolis-Hastings sampling of paths ζ ($r = 0.3$). For each value $\zeta \in \{1.1, 1.2, ..., 1.9\}$ and each number of elements in the choice set per observation, 5 choice models were estimated based on different random outputs from the Metropolis-Hastings sampling of paths.

5 models were estimated with choice sets of 10, 20 or 50 elements per observation. When $r = 0.5$, the impact of path cost over node cost is too large. For all other values, all parameters are correctly estimated when there are 10 or 100 elements in the choice set.

Based on this sensitivity analysis, it is recommended to generate more than 16 elements in the choice set per observation. Increasing the number of elements decreases the standard errors of the parameters and increases the chances to estimate parameters that are not significantly different from their true values. With a large enough choice set, the parameters used in the Metropolis-Hastings sampling of paths (ζ) and in the definition of path cost (r) are stable.

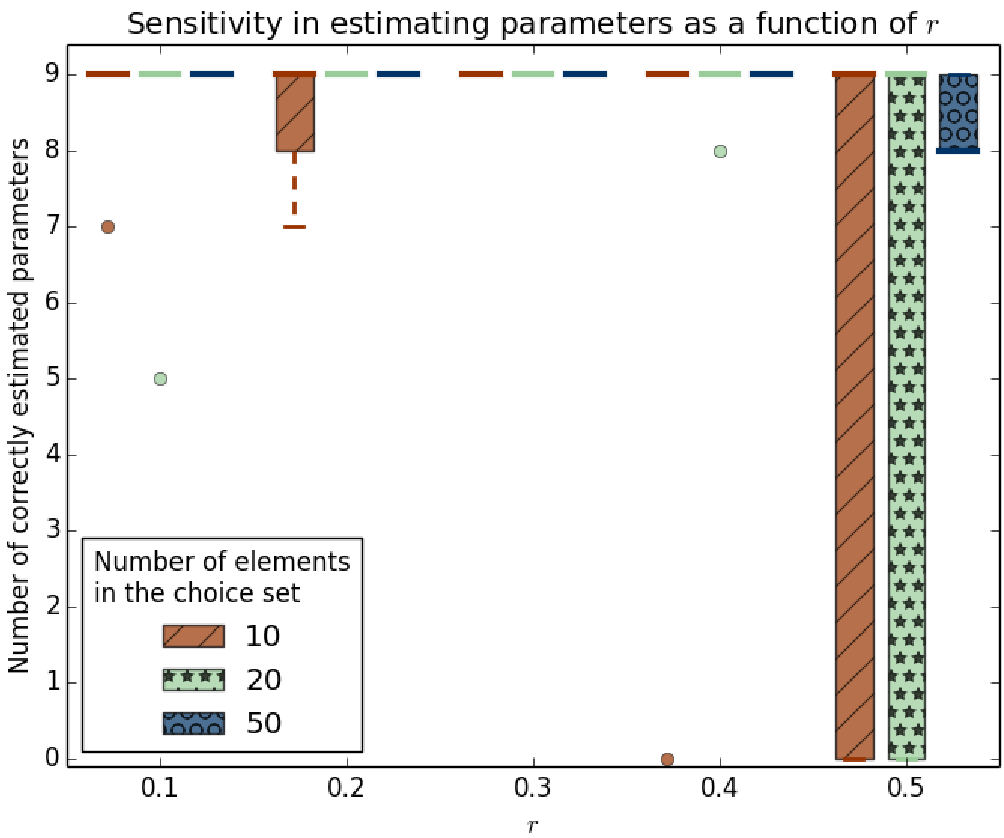

Figure 7: Number of non-significantly different parameters compared to their true values as a function of the ratio r between path cost and node cost ($\zeta = 1.5$). For each value $r \in \{0.1, 0.2, ..., 0.5\}$ and each number of elements in the choice set per observation, 5 choice models were estimated based on different random outputs from the Metropolis-Hastings sampling of paths.

5 Pedestrian case study on a campus

The choice of an activity pattern also trigger people's behavior in pedestrian facilities (Bierlaire and Robin, 2009). Demand management measures can be applied in pedestrian facilities, such as changing schedules (train in train stations, planes in airports, concerts in music festivals or class schedules in universities). Our model is particularly suitable for pedestrian facilities, since it does not assume a tour structure nor a "home" activity type.

In particular, transport hubs (e.g., train stations or airports) are key nodes of a multimodal transport system (with buses, metro, car and bike sharing). Train stations are located in the city centers and include shops and services. All these activities combined with the growth in the number of passengers increase pedestrian flows and threaten the functioning of the train station.

Understanding demand for activities is of utmost importance to define appropriate planning policies.

As a proof of concept, we apply our methodology to WiFi traces collected on the EPFL campus. EPFL campus approximately hosts 13'000 people per day. Similarly to transport hubs, some of them follow schedules (class schedules instead of train schedules) and perform several different activities, such as going to class, having lunch, etc.

Section 5.1 describes the data used in this case study, Section 5.2 describes the choice set generation process and the choice model, and results are presented in Section 5.3.

5.1 Data source and activity network

We collected data as defined in Danalet *et al.* (2014). Campus users authenticate themselves on the WiFi network through WPA using a Radius server. Accounting is one of the process on the Radius server. It allows to associate a MAC address with a username. Each measurement was associated with a unique identifier and a category, such as employee or civil engineering student, bachelor. Data were then anonymized by deleting the MAC address. Details about the data collection campaign and data cleaning can be found in Danalet and Bierlaire (2014), Section 4.1.

The Bayesian approach described in Danalet *et al.* (2014) was then applied to the raw WiFi traces in order to detect activity-episode sequences. It merges WiFi traces with data from the map of the campus, measures of attractivity of each destination and time constraints. The precise definition of these data, and in particular of attractivity measures, can be found in Danalet and Bierlaire (2014).

We assume 8 activity types: classrooms, shops, offices, restaurant, library, lab, other and not being detected. The types "Office", "Classroom" and "Lab" are based on norm DIN 277 defined by the *Deutsches Intitut für Normung*. The types "Shops", "Restaurant", "Library" and "Other" are extracted from a list of points of interest from `http://map.epfl.ch`.

There are $T = 24$ time units in the activity network, from 7am to 7pm, from :00 to :15 and from :15 to :00 at each hour. This structure with 15-minute and 45-minute time intervals is motivated by the time structure of classes at EPFL, mostly starting at :15 and finishing at :00 every hour. It exemplifies the flexibility in the definition of the activity network, in order to focus on a certain time of the day or to adapt to the behavioral time structure.

5.2 Choice set and choice model

Conversely to synthetic data, the full choice set cannot be enumerated (8^{24} paths) in this case study. The Metropolis-Hastings algorithm is used with node and activity-episode length attractivity, as described in Section 3.2, with $\zeta = 1.3$ and $r = 0.3$, the best values according to the synthetic data experiment. A distance $d = 500'000$ between iterations is defined based on the similarity measure presented in Section 4.2 (Fig. 8). 100 paths have been generated per observation as the choice set, sampling paths every d iterations of the Metropolis-Hastings algorithm. The shortest path for connecting the insertion node in the splice operation of the Metropolis-Hastings algorithm uses node attractivity only (since activity-episode length attractivity is not link-additive).

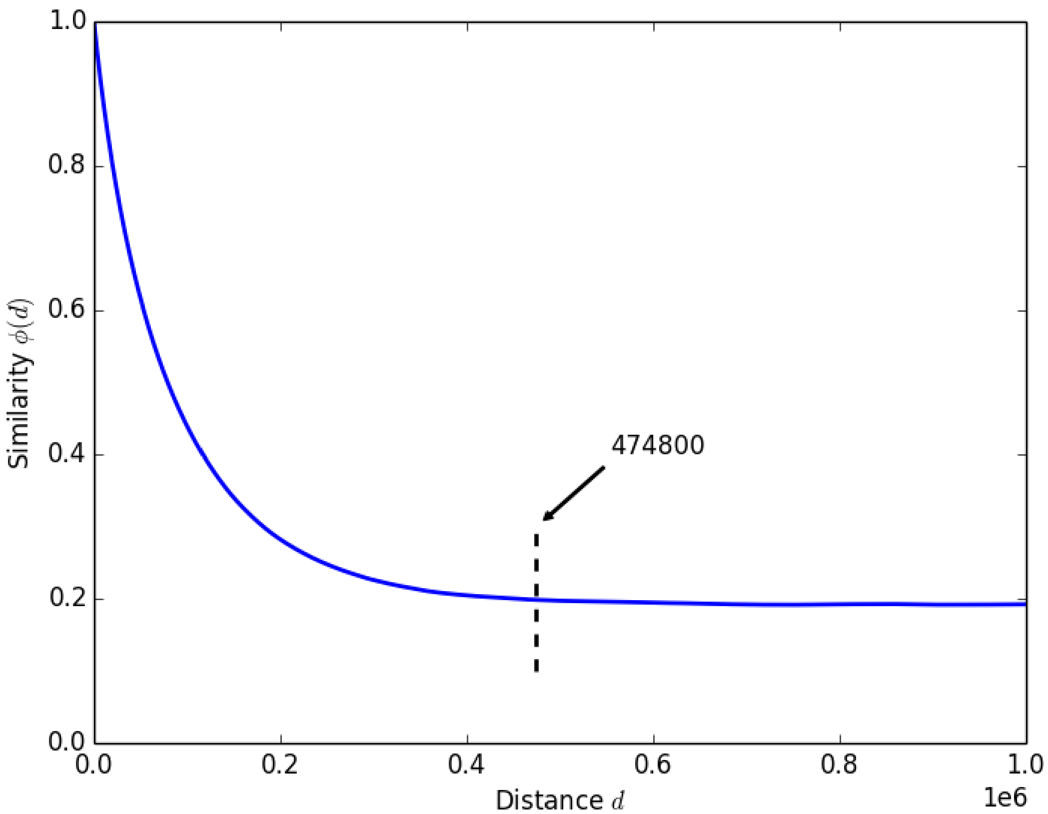

Figure 8: Similarity measure as a function of the distance. 10^6 paths were generated using node and activity-epiosde length attractivity, with $\zeta = 1.3$ and $r = 0.3$.

We estimate a logit model with sampling correction, as described in Section 3.3. No activity path size (Section 3.4) has been included in the model due to lack of data variability, as estimation results show below.

5.3 Estimation results

Estimation results are presented in Table 4 for two different choice sets. The first column presents results with a choice set of 100 elements selected by simple random sampling, similar to the first step in the strategy described by Lemp and Kockelman (2012). The second column presents results with a choice set of 100 elements selected by importance sampling using the Metropolis-Hastings algorithm and node and activity-episode length attractivity., similarly to the methodology presented in Section 3.2 and validated in Section 4.

	Simple random sampling			**Importance sampling**		
Description	Coeff. estimate	Robust Asympt. std. error	*t*-stat	Coeff. estimate	Robust Asympt. std. error	*t*-stat
$\eta_{\text{Classroom, Shop, Library}}$	-0.570	0.206	-2.76	-0.735	0.149	-4.92
$\eta_{\text{Lab, Restaurant, Office, Other}}$	-0.722	0.231	-3.12	-1.02	0.196	-5.20
$\beta_{\text{3+ lab episodes}}$	-0.500	0.106	-4.72	-1.15	0.283	-4.06
$\beta_{\text{3+ office episodes}}$	-0.583	0.167	-3.48	-0.818	0.271	-3.02
$\beta_{\text{3 restaurant episodes}}$	-0.603	0.225	-2.68	-0.720	0.145	-4.98
$\beta_{\text{4+ restaurant episodes}}$	-1.01	0.295	-3.44	-1.57	0.419	-3.74
$\beta_{\text{3+ shop episodes}}$	-0.577	0.175	-3.29	-1.03	0.332	-3.09
$\beta_{\text{NA 7-9, 17-19, employees}}$	0.165	0.0643	2.56	0.349	0.0707	4.93
$\beta_{\text{NA 9-17, employees}}$	-0.424	0.149	-2.85	-1.27	0.251	-5.05
$\beta_{\text{NA 8:15-17, students}}$	-0.649	0.231	-2.81	-0.909	0.192	-4.73
$\beta_{\text{NA 17-19, students}}$	0.296	0.107	2.77	0.550	0.244	2.25
$\beta_{\text{classroom 8:15-12, 13:15-17, employees}}$	-1.46	0.489	-2.99	-1.25	0.400	-3.11
$\beta_{\text{primary activity library, students}}$	0.327	0.131	2.50	0.616	0.247	2.50
Number of parameters			13	13		
$\mathcal{L}(\beta_0)$			-8002.619	-8002.619		
$\mathcal{L}(\hat{\beta})$			-13.244	-10.340		
ρ^2			0.999	0.998		
$\bar{\rho}^2$			0.996	0.997		

Table 4: Estimation results for randomly generated choice set and choice set based on node and length of observations. In both cases, 100 elements were generated per observation. In total, the two models are estimated on 1734 observations. Importance sampling allows for more variables to be included in the model. In both cases, the final log-likelihood, ρ^2 and $\bar{\rho}^2$ are very close to 1. It indicates that the choice set does not contain enough variation to include more variables.

$\eta_{\text{Classroom, Shop, Library}}$ and $\eta_{\text{Lab, Restaurant, Office, Other}}$ represent the satiation parameter for classrooms, shops, libraries and laboratories, restaurants, office, other, respectively; it multiplies $\ln(|a|)$, where $|a|$ is the duration of the activity episode, similarly to Eq. 19; they are both negative, as expected, indicating a decreasing marginal utility when duration increases. $\beta_{\text{4+ classroom episodes}}$ multiplies an indicator variable with value 1 if the activity path contains 4 or more activity episodes corresponding to the "classroom" activity type; the negative sign expresses the preference for a limited number of activity episodes in classes in a day. The following 9 variables represent a similar preference for laboratories, libraries, offices, restaurants and shop activity types. $\beta_{\text{NA 13:15-17, students}}$, $\beta_{\text{NA 7-9, 17-19, employees}}$ and $\beta_{\text{NA 9-17, employees}}$ multiplies the number of quarters of an hour corresponding to the activity type "not being detected" (NA) in the activity path; it express the preference for not being on campus (NA); students are more likely to be on campus in the afternoon, while employees are likely to be on campus during office hours and likely not to be detected on campus early or late in the day. $\beta_{\text{classroom 8:15-12, 13:15-17, employees}}$ multiplies the number of quarters of an hour in the activity path corresponding to the activity type "classroom", for employees; it is negative, showing a preference of employees not to be in classrooms during the day. $\beta_{\text{restaurant, 12-13:15, students}}$ multiplies the number of quarters of an hour in a restaurant in the activity path during lunch break for students in the utility, showing the preference for performing such activities during lunch breaks. $\beta_{\text{office-restaurant-office, 12-13:15, employees}}$ multiplies an indicator variable with value 1 if the activity path contains the pattern office-restaurant-office as a subsequence, with restaurant being during lunch break and possibly with nodes corresponding to "not being detected" in between; it is positive, showing a preference for this pattern for employees in their day. $\beta_{\text{primary activity library, students}}$ multiplies a indicator variable with value 1 when the total number of quarters of an hour is larger for "library" than for any other activity type in the activity path; it is positive for students since data were collected during a revision period before exams. Finally, $\gamma_{\text{late, students}}$ multiplies the late schedule delay SDL $= max(t^a - t^*, 0)$, where t^* is the class start time (based on the class schedules for the category of students) and t^a is the classroom activity-episode start time (if it happens after the class start time and before the class end time); it is negative, showing that students prefer to be on time for classes, if they go to classes.

With our approach, variables that are related to the whole activity path, such as primary activity in the day ($\beta_{\text{primary activity library, students}}$) or patterns ($\beta_{\text{office-restaurant-office, 12-13:15, employees}}$) can be included in the model. However, with both sampling strategies, the final log-likelihood $\mathcal{L}(\hat{\beta})$ is close to zero and the adjusted rho-square $\bar{\rho}^2$ is close to 1. It indicates that the sampled choice sets are dominated by the chosen alternative. While this is intuitive for the simple random sampling, it shows that the importance sampling strategy defined before is not sufficient to define a proper choice set. The current choice set is fully dominated by the observed choice, given the choice model.

Validation has been performed on this model. 80% of the data have been used for estimation, and the model has been applied to the 20% remaining data. For both simple random sampling and importance sampling using node and activity-episode length attractivity, results show that all generated alternatives in the choice set have an almost null probability and the chosen alternative has a very high probability. This results confirms that the generated alternatives in the choice set are clearly dominated and not similar enough to the chosen alternative.

When comparing the two models, we see that simple random sampling (SRS) leads to lower *t*-statistics. In particular, 6 parameters are not significantly different from 0 (p-value > 0.05). Note also that $\beta_{\text{NA 9-17, employees}}$ has a positive sign with simple random sampling, which is not intuitive: it would mean that employees prefer not to be on campus during office hours. Several estimations were made with simple random sampling (only the first one is shown) and in some cases, the value of some parameters explode due to the lack of observations in the data corresponding to the parameter (say $\beta_{\text{3 lab episodes}}$ explodes because there are no or almost no activity path in the SRS-generated choice set with 3 activity episodes corresponding to a "laboratory" activity type).

6 Conclusion and future work

Our approach models the choice of activity sequences of individuals and evaluates the attractivity of the different activity types in time and the satiation effect for each activity type. Our model is not home-based, nor tour-based. It can adapt to different contexts and activity types. The large dimensionality of the problem is managed through importance sampling techniques.

An important feature of our approach is that it allows to add in the utility function variables that are not specific to activity episodes, but are related to the path itself. Patterns can be included in the utility function and their specific utility or satiation parameters can be estimated.

One key issue of this approach concerns the choice set. The choice set is extremely large and the full choice set cannot be used. Consideration choice set is very difficult to define. Importance sampling seems the best strategy to define a choice set based on an *a priori* measure of the attractivity of certain paths. In Danalet and Bierlaire (2014), a measure of attractivity was used to define the weight in importance sampling, based on an aggregate measure of the number of persons at each time performing each activity type; in this paper, we use the observed data, and in particular the node and the activity-episode length attractivity. As we show in Section 5, this strategy does not allow to specify a model predicting activity paths with similar probabilities than the chosen activity path. This is due to the impossibility to define attractivity such that the generated activity paths in the choice set are similar enough to the observed activity paths.

Lemp and Kockelman (2012) propose to use the choice probabilities of a first model as weight in importance sampling. With the impossibility to define exogenously the weight in our context, it seems that their sampling strategy is the best for the activity path choice approach. It will be explored in future research.

7 References

Abou-Zeid, M. and M. Ben-Akiva (2012) Well-being and activity-based models, *Transportation*, **39** (6) 1189–1207, ISSN 00494488.

Antonisse, R., A. Daly and H. Gunn (1986) The primary destination tour approach to modelling trip chains, paper presented at the *Proceedings of Seminar M on Transportation Planning Methods at the 14th PTRC Summer Annual Meeting*, 165–177.

Arentze, T. a., D. Ettema and H. J. Timmermans (2013) Location choice in the context of multi-day activity-travel patterns: model development and empirical results, *Transportmetrica A: Transport Science*, **9** (2) 107–123, ISSN 2324-9935.

Arentze, T. A. and H. J. Timmermans (2004) A learning-based transportation oriented simulation system, *Transportation Research Part B*, **38** (7) 613–633, ISSN 01912615.

Arnott, R., A. de Palma and R. Lindsey (1990) Economics of a bottleneck, *Journal of Urban Economics*, **27**, 111–130, ISSN 00941190.

Arnott, R., A. D. Palma and R. Lindsey (1993) A Structural Model of Peak-Period Congestion: A Traffic Bottleneck with Elastic Demand, *The American Economic Review*, **83** (1) 161–179, ISSN 00028282.

Başar, G. and C. Bhat (2004) A parameterized consideration set model for airport choice: An application to the San Francisco Bay Area, *Transportation Research Part B*, **38**, 889–904, ISSN 01912615.

Ben-Akiva, M. and M. Bierlaire (1999) Discrete choice methods and their applications to short-term travel decisions, in R. Hall (ed.) *Handbook of Transportation Science*, Operations Research and Management Science, 5–34, Kluwer, Netherlands.

Ben-Akiva, M. and B. Boccara (1995) Discrete choice models with latent choice sets, *International Journal of Research in Marketing*, **12** (1) 9–24, ISSN 0167-8116.

Ben-Akiva, M., J. Bowman and D. Gopinath (1996) Travel demand model system for the information era, *Transportation*, **23** (3) 241–266, ISSN 0049-4488.

Ben-Akiva, M. E. and S. R. Lerman (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge, MA, ISBN 9780262022170.

Bhat, C. R. (2005) A multiple discrete-continuous extreme value model: formulation and application to discretionary time-use decisions, *Transportation Research Part B: Methodological*, **39** (8) 679–707, ISSN 01912615.

Bhat, C. R. and F. S. Koppelman (1999) Activity-Based Modelling of Travel Demand, in *Handbook of Transportation Science*, vol. 23, 35–61, Springer US, ISBN 978-1-4613-7370-4.

Bierlaire, M. and T. Robin (2009) Pedestrians Choices, paper presented at the *Pedestrian Behavior. Models, Data Collection and Applications*, 1–26, ISBN 978-1-84855-750-5.

Bolduc, D. and M. Ben-Akiva (1991) A multinomial probit formulation for large choice sets, paper presented at the *Proceedings of the Sixth International Conference on Travel Behaviour*, vol. 2, 243–258.

Bowman, J. and M. Ben-Akiva (2001) Activity-based disaggregate travel demand model system with activity schedules, *Transportation Research Part A*, **35** (1) 1–28, ISSN 09658564.

Bowman, J. L. (1998) The Day Activity Schedule Approach to Travel Demand Analysis, PhD thesis, Massachusetts Institute of Technology.

Bowman, J. L. (2009) Historical Development of Activity Based Model Theory and Practice, *Traffic Engineering and Control*, **50** (7) 314–318.

Bradley, M., J. L. Bowman and B. Griesenbeck (2010) SACSIM: An applied activity-based model system with fine-level spatial and temporal resolution, *Journal of Choice Modelling*, **3** (1) 5–31, ISSN 17555345.

Cascetta, E., A. Nuzzolo, F. Russo and A. Vitetta (1996) A modified logit route choice model overcoming path overlapping problems: Specification and some calibration results for interurban networks, paper presented at the *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, 697–711, Lyon, France, ISBN 0-08-042586-0.

Chen, J. (2013) Modeling route choice behavior using smartphone data, Ph.D. Thesis, Ecole Polytechnique Fédérale de Lausanne, Switzerland.

Childress, S. (2010) Focus Model Calibration 1.0, Denver Regional Council of Governments (DRCOG), Activity-Based Travel Model, *Technical Report*, Denver Regional Council of Governments, Cambridge Systematics.

Danalet, A. and M. Bierlaire (2014) A path choice approach to activity modeling with a pedestrian case study, paper presented at the *14th Swiss Transport Research Conference (STRC)*, Monte Verità, Ascona, Switzerland.

Danalet, A., B. Farooq and M. Bierlaire (2014) A Bayesian approach to detect pedestrian destination-sequences from WiFi signatures, *Transportation Research Part C*, **44**, 146–170.

Davidson, W., R. Donnelly, P. Vovsha, J. Freedman, S. Ruegg, J. Hicks, J. Castiglione and R. Picado (2007) Synthesis of first practices and operational research approaches in activity-based travel demand modeling, *Transportation Research Part A: Policy and Practice*, **41** (5) 464–488, ISSN 09658564.

de Lapparent, M. (2009) Latent class and mixed Logit models with endogenous choice set formation based on compensatory screening rules, in S. Hess and A. Daly (eds.) *Choice Modelling: The State-of-the-Art and the State-of-Practice*, chap. 17, Emerald.

Ettema, D. (1996) Activity based Travel Demand Modelling, Ph.D. Thesis, Eindhoven Technical University, Holland.

Ettema, D., O. Ashiru and J. W. Polak (2004) Modeling timing and duration of activities and trips in response to road-pricing policies, *Transportation Research Record*, **41** (1894) 1–10.

Ettema, D., F. Bastin, J. Polak and O. Ashiru (2007) Modelling the joint choice of activity timing and duration, *Transportation Research Part A*, **41** (9) 827–841, ISSN 09658564.

Ettema, D. and H. Timmermans (2003) Modelling Departure Time Choice in the Context of Activity Scheduling Behavior, *Transportation Research Record*, **1831**, 39–46.

Feil, M. (2010) Choosing the Daily Schedule: Expanding Activity-Based Travel Demand Modelling, Ph.D. Thesis, ETH.

Flajolet, P. and R. Sedgewick (2002) *Analytic combinatorics*, Cambridge University Press, ISBN 9780521898065.

Flötteröd, G. and M. Bierlaire (2013) Metropolis-Hastings sampling of paths, *Transportation Research Part B*, **48**, 53–66, ISSN 01912615.

Frejinger, E. and M. Bierlaire (2007) Capturing correlation with subnetworks in route choice models, *Transportation Research Part B*, **41** (3) 363–378, ISSN 01912615.

Frejinger, E. and M. Bierlaire (2010) On Path Generation Algorithms for Route Choice Models, in S. H. Daly and A. (eds.) *Choice Modelling: The State-of-the-Art and the State-of-Practice*, 307–315, Emerald Group Publishing Limited, ISBN 978-1-84950-772-1.

Frejinger, E., M. Bierlaire and M. Ben-Akiva (2009) Sampling of Alternatives for Route Choice Modeling, *Transportation Research Part B*, **43** (10) 984–994.

Fu, X. and W. H. K. Lam (2014) A network equilibrium approach for modelling activity-travel pattern scheduling problems in multi-modal transit networks with uncertainty, *Transportation*, **41** (1) 37–55, ISSN 0049-4488.

Gupta, S. and P. Vovsha (2013) A model for work activity schedules with synchronization for multiple-worker households, *Transportation*, **40** (4) 827–845, ISSN 0049-4488.

Habib, K. M. N. (2007) Modelling activity generation processes, Ph.D. Thesis, University of Toronto.

Habib, K. M. N. (2011) A random utility maximization (RUM) based dynamic activity scheduling model: Application in weekend activity scheduling, *Transportation*, **38** (1) 123–151, ISSN 0049-4488.

Habib, K. N., A. Sasic, C. Weis and K. Axhausen (2013) Investigating the nonlinear relationship between transportation system performance and daily activity-travel scheduling behaviour, *Transportation Research Part A*, **49**, 342–357, ISSN 09658564.

Hägerstraand, T. (1970) What about people in regional science, *Papers in Regional Science*, **24** (1) 7–24, ISSN 10568190.

Hess, S., J. W. Polak, A. Daly and G. Hyman (2007) Flexible substitution patterns in models of mode and time of day choice: New evidence from the UK and the Netherlands, *Transportation*, **34** (2) 213–238, ISSN 00494488.

Ho, C. and C. Mulley (2013) Tour-based mode choice of joint household travel patterns on weekend and weekday, *Transportation*, **40** (4) 789–811, ISSN 0049-4488.

Horni, A. (2013) Destination choice modeling of disrectionary activiites in transport microsimulations, PhD Thesis, ETH Zurich.

Jenelius, E., L.-G. Mattsson and D. Levinson (2011) Traveler delay costs and value of time with trip chains, flexible activity scheduling and information, *Transportation Research Part B*, **45** (5) 789–807, ISSN 01912615.

Joh, C.-h., T. Arentze and H. Timmermans (2004) Activity-travel scheduling and rescheduling decision processes: empirical estimation of Aurora model, *Transportation Research Record*, **1898**, 10–12.

Kalakou, S. and F. Moura (2014) Effects of terminal planning on passenger choices, paper presented at the *14th Swiss Transport Research Conference (STRC)*, Monte Verità, Ascona, Switzerland.

Kang, J. E. and W. Recker (2013) The location selection problem for the household activity pattern problem, *Transportation Research Part B*, **55**, 75–97, ISSN 01912615.

Kockelman, K. M. (1998) A Utility-Theory-Consistent System-of-Demand-Equations Approach to Household Travel Choice, PhD thesis, University of California, Berkeley.

Kockelman, K. M. (2001) A model for time- and budget-constrained activity demand analysis, `http://dx.doi.org/10.1016/S0191-2615(99)00050-8`.

Lai, X. and M. Bierlaire (2014) Specification of the cross nested logit model with sampling of alternatives for route choice models, *Technical Report*, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne, Lausanne.

Lemp, J. D. and K. M. Kockelman (2012) Strategic sampling for large choice sets in estimation and application, *Transportation Research Part A: Policy and Practice*, **46** (3) 602–613, ISSN 09658564.

Li, S. and D. Lee (2014) Learning Daily Activity Pattern with Probabilistic Grammar, paper presented at the *TRB 93rd Annual Meeting Compendium of Papers*, 17, Washington, DC, USA.

Liao, F., T. Arentze and H. Timmermans (2013) Incorporating space-time constraints and activity-travel time profiles in a multi-state supernetwork approach to individual activity-travel scheduling, *Transportation Research Part B*, **55**, 41–58, ISSN 01912615.

McFadden, D. (1978) Modelling the choice of residential location, in A. Karlqvist, L. Lundqvist, F. Snickars and J. Weibull (eds.) *Spatial interaction theory and residential location*, 75–96, North-Holland, Amsterdam.

Miller, H. J. (2014) Activity-Based Analysis, in M. M. Fischer and P. Nijkamp (eds.) *Handbook of Regional Science*, 705–724, Springer Berlin Heidelberg, Berlin, Heidelberg.

Nijland, L., T. Arentze and H. Timmermans (2014) Multi-day activity scheduling reactions to planned activities and future events in a dynamic model of activity-travel behavior, *Journal of Geographical Systems*, **16** (1) 71–87, ISSN 1435-5930.

Nurul Habib, K. M. and E. J. Miller (2009) Modelling activity generation: a utility-based model for activity-agenda formation, *Transportmetrica*, **5** (1) 3–23, ISSN 1812-8602.

Ouyang, L. Q., W. H. K. Lam, Z. C. Li and D. Huang (2011) Network User Equilibrium Model for Scheduling Daily Activity Travel Patterns in Congested Networks, *Transportation Research Record: Journal of the Transportation Research Board*, **2254** (-1) 131–139, ISSN 0361-1981.

Pagliara, F. and H. J. Timmermans (2009a) Choice set generation in spatial contexts. A retrospective review, paper presented at the *Handbook of Theoretical and Quantitative Geography*, 335–364, Lausanne.

Pagliara, F. and H. J. P. Timmermans (2009b) Choice set generation in spatial contexts: a review, *Transportation Letters: The International Journal of Transportation Research*, **1** (3) 181–196, ISSN 1942-7867.

Papadimitriou, E., G. Yannis and J. Golias (2009) A critical assessment of pedestrian behaviour models, *Transportation Research Part F*, **12** (3) 242–255, ISSN 13698478.

Pinjari, A. R. and C. Bhat (2010) 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, ISSN 01912615.

Pinjari, A. R. and C. R. Bhat (2011) Activity Based Travel Demand Analysis, in A. de Palma, R. Lindsey, E. Quinet and R. Vickerman (eds.) *A Handbook of Transport Economics*, chap. 10, 213–248, Edward Elgar Publishing Ltd.

Rasouli, S., Y. Shiftan and H. Timmermans (2013) Current issues in choice modeling: Choice set specification, non-utility-maximizing behavior and discrete-continuous choice problems, *Journal of Choice Modelling*, **9** (2013) 1–2, ISSN 17555345.

Rasouli, S. and H. Timmermans (2014) Activity-based models of travel demand: promises, progress and prospects, *International Journal of Urban Sciences*, **18** (1) 31–60, ISSN 1226-5934.

Roorda, M. J. (2005) Activity-based modelling of household travel, Ph.D. Thesis, University of Toronto.

Shiftan, Y. and M. Ben-Akiva (2011) A practical policy-sensitive, activity-based, travel-demand model, *Annals of Regional Science*, **47**, 517–541, ISSN 05701864.

Small, B. K. A. (1982) The Scheduling of Consumer Activities: Work Trips, *The American Economic Review*, **72** (3) 467–479.

Swait, J. (2001) A non-compensatory choice model incorporating attribute cutoffs, *Transportation Research Part B*, **35**, 903–928, ISSN 01912615.

Tinguely, L., A. Danalet, M. de Lapparent and M. Bierlaire (2015) Destination Choice Model including a panel effect using WiFi localization in a pedestrian facility, paper presented at the *15th Swiss Transport Research Conference (STRC)*, 44, Monte Verità, Ascona, Switzerland.

Ton, D. (2014) NAVISTATION: a study into the route and activity location choice behaviour of departing pedestrians in train stations, Master thesis, Delft University of Technology.

Tseng, Y.-Y. and E. T. Verhoef (2008) Value of time by time of day: A stated-preference study, *Transportation Research Part B*, **42** (7-8) 607–618, ISSN 01912615.

Vickrey, W. S. (1969) Congestion Theory and Transport Investment, *The American Economic Review*, **59** (2) 251–260.

Vovsha, P. and S. Bekhor (1998) Link-Nested Logit Model of Route Choice: Overcoming Route Overlapping Problem, *Transportation Research Record*, **1645** (1) 133–142, ISSN 0361-1981.

Wang, D. and J. Li (2011) A two-level multiple discrete-continuous model of time allocation to virtual and physical activities, *Transportmetrica*, **7** (6) 395–416, ISSN 1812-8602.

Winston, G. (1982) *The timing of economic activities*, Cambridge University Press, Cambridge, UK, ISBN 9780521070928.

Yamamoto, T. and R. Kitamura (1999) An analysis of time allocation to in-home and out-of-home discretionary activities across working days and non-working days, *Transportation*, **26** (2) 211–230.