Modeling the dynamics of all-day activity plans

Willem Himpe[1]
Gunnar Flötteröd, Ricardo Hurtubia, and Michel Bierlaire[2]

June 22, 2010

[1]KU Leuven
[2]EPF Lausanne

Motivation

- activity-based travel demand modeling
- personalized services on smartphones

Outline

Methodology

Model estimation

Preliminary validation

Summary, outlook

Outline

Methodology

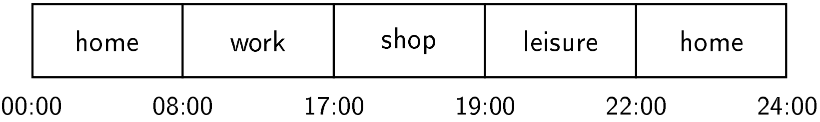
Model estimation

Preliminary validation

Summary, outlook

Terminology

activity type: home, work, education, leisure, shopping, other
activity schedule: temporal sequence of activity types



Dynamics of activity schedules

- timing
  - facility (e.g., shop, office) opening times
  - avoid exhaustive activities early in the day
- duration
  - being at work for 8 h is desirable, just 2 h are not
  - playing tennis for 2 h is fun, playing for 8 h is not
- sequencing
  - don't go shopping twice per day
  - bring kids to kindergarden → pick them up later

Probabilistic activity sequencing model

home	work	???	???	home

- activity schedule consists of N activity slots with fixed timing
- nth activity is $a_n \in$ {home, work, edu, leisure, shop, other}
- M unknown activities with indices x_m, $m = 1 \dots M$
- objective: model probability distribution of full schedule

$$P(a_{\{x_1 \dots x_M\}} | a_{\{1 \dots N\} \setminus \{x_1 \dots x_M\}})$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Single gap model

- problem with full model: combinatorial explosion
- assume that only one activity is unknown

$$P(a_x | a_{\{1 \ldots N\} \setminus x})$$

- one-dimensional distribution, no combinatorial issues
- can be estimated from data

Multiple-gap model

- reconstruct full (multiple-gap) distribution

$$P(a_{\{x_1 \dots x_M\}} | a_{\{1 \dots N\} \setminus \{x_1 \dots x_M\}})$$

from marginals

$$P(a_{x_m} | a_{\{1 \dots N\} \setminus x_m}),\ m = 1 \dots M$$

- computational technique: Gibbs sampling
- only requires the estimation of single-gap models

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Outline

Methodology

Model estimation

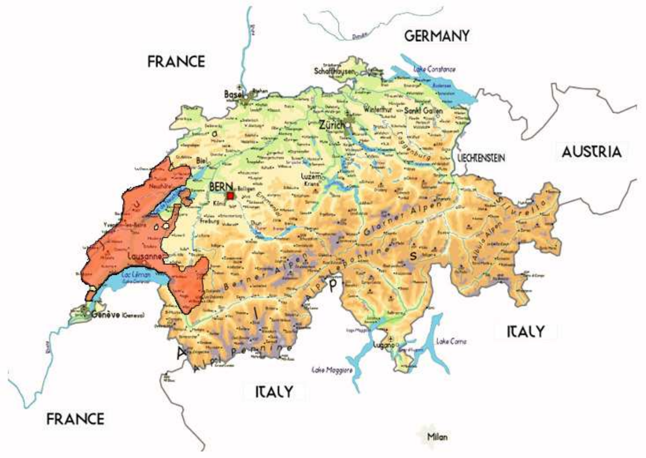
Preliminary validation

Summary, outlook

Data

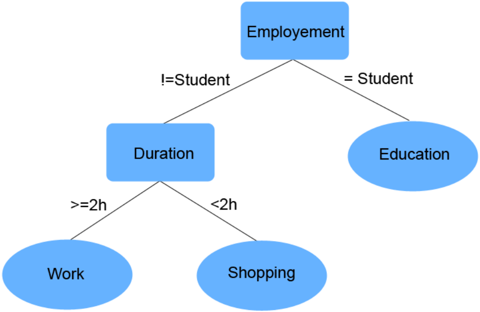
- Swiss microcensus 2005
  - overall 33 390 respondents
  - activity and travel behavior for a single day
  - linked to socio-economics of respondents
- consider only canton of Vaud
  - 2157 persons
  - 8508 activities

Study region: canton of Vaud



Model structure

- represent single-gap model through decision tree
- example:



Explanatory variables

name	values	description
start_day	minutes after midnight	first time at which home is left
activ_start	minutes after midnight	starting time of the activity gap
duration	minutes	duration of the activity gap
{tot_*activity*}	minutes	total duration of *activity* outside gap
employment	full time, part time, student, unemployed	daytime occupation of respondent
weekday	any weekday	considered day of the week
prev_act	an activity type	activity conducted before the gap
next_act	an activity type	activity conducted after the gap

Estimation of single-gap model

- use C4.5 learning algorithm (Weka software package)
- incrementally builds a decision tree, then extracts rules
- selects branching conditions that separate the data well
- technically, maximizes entropy of child node distribution
- yields "crisp rules", no probabilities
  $\Rightarrow$ attach empirical activity distribution to each leaf

Estimation results

- tree with 57 leaves
- $>68\,\%$ of data correctly classified (10-fold cross-validation)
- extracted rules are plausible
- example:

  IF act_start_time $\leq$ *532 min (08:52)*
  AND employment $\notin$ *{student, unemployed}*
  THEN Pr(gap=work) = 0.81

Estimation results

work	shop	leisure	home	edu	other	model/data
1308	57	218	127	22	0	work
89	**568**	486	81	5	6	shop
162	293	**2251**	295	31	2	leisure
114	59	259	**1277**	4	1	home
13	3	47	15	**391**	0	education
32	124	133	23	3	**9**	other

Outline

Methodology

Model estimation

Preliminary validation

Summary, outlook

Smartphone data collection campaign

- joint project with Nokia
- approx. 50 respondents carry smartphones that observe
  - GPS, wireless, Bluetooth
  - actions conduced on phone
  - visual & acoustic samples
- supplementary survey

⇒ behavioral modeling

- preliminary data base
  - one respondent
  - 45 days

Results – performance measures

- single-gap log-likelihood

$$\mathcal{L}_{single} = \sum_{d=1}^{45} \sum_{x=1}^{N_d} \log P(a_x = y_{dx} | a_{\{1 \dots N_d\} \setminus x})$$

where

  - N_d is the number of activities in day d
  - y_{dx} is the reported activity in slot x of day d

- multiple-gap log-likelihood

$$\mathcal{L}_{multiple} = \sum_{d=1}^{45} \sum_{x=1}^{N_d} \log P(a_x = y_{dx} | \cdot)$$

Results – performance measures

- null log-likelihood

$$\mathcal{L}_0 = \sum_{d=1}^{45} \sum_{x=1}^{N_d} \log \frac{1}{6}$$

- results

$\mathcal{L}_0$	-49.802
$\mathcal{L}_{single}$	-28.434
$\mathcal{L}_{multiple}$	-36.224

Outline

Methodology

Model estimation

Preliminary validation

Summary, outlook

Summary

- dynamic, non-behavioral model of activity scheduling
- requires estimation only of models for single activities
- draws full activity schedules with Gibbs sampling

Outlook

- replace known activities by activity distributions
- model more degrees of freedom
  - temporal structure
  - number of activities
  - travel episodes
- replace rule set by behavioral model
  - interpretability
  - extrapolation

Outlook

- replace known activities by activity distributions
- model more degrees of freedom
  - temporal structure
  - number of activities
  - travel episodes
- replace rule set by behavioral model
  - interpretability
  - extrapolation

Thank you for your attention.