A Bayesian Approach to Detect Pedestrian Destination-Sequences from WiFi Signatures

Antonin Danalet
Bilal Farooq
Michel Bierlaire

TRISTAN VIII, San Pedro de Atacama
June 10th, 2013

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Presentation outline

- Motivation
- Data requirement
- Methodology
- A case study on EPFL campus
- Conclusion
- Future work

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MOTIVATION

Walking is the key for efficient multimodal transport systems

Crowd in a railway station in Mumbai, India
Photo: National Geographic

Lake Geneva region

By 2030, 100’000 passengers per day between Geneva and Lausanne

= 2000 travelers/day

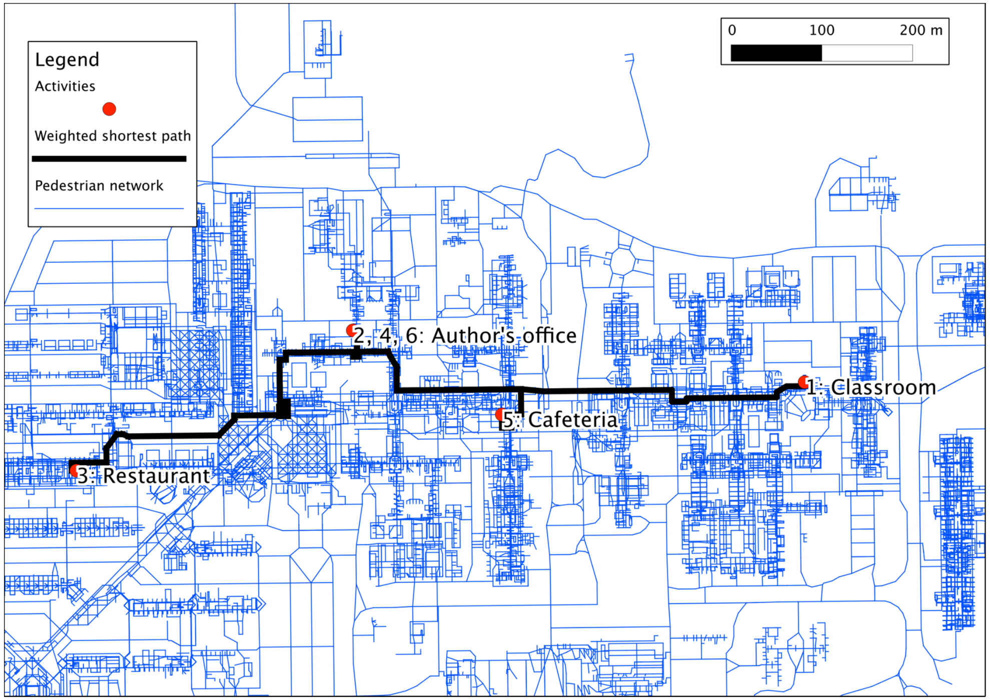
* Forecast by Swiss Railways for the maximum scenario

TRANSP-OR

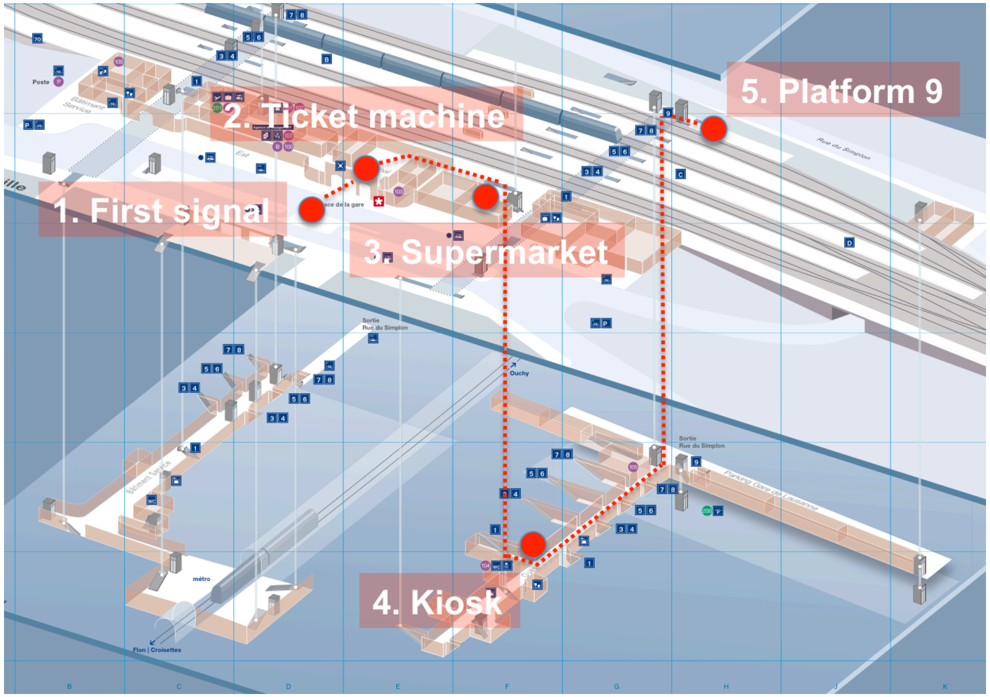
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Understand pedestrian activities

What we are doing: Campus

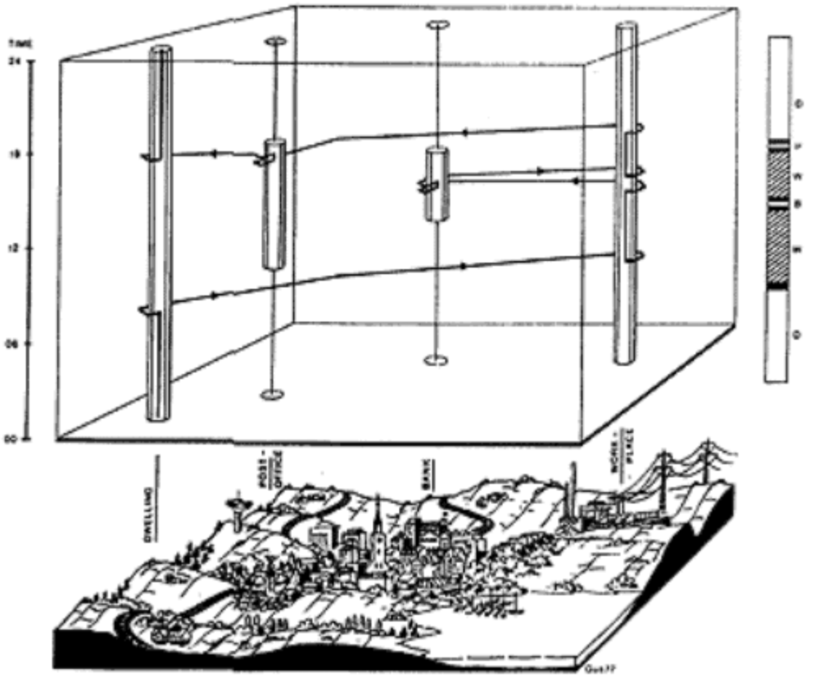


What we want to do: Station



Challenges

- Detect pedestrian destinations
- Model pedestrian activity scheduling behavior
- Forecast the impact of changes in the infrastructure



Carlstein, T. (1978)

DATA REQUIREMENT

Data requirement

- Required
  - Localization data with full coverage of the facility
  - Semantically-enriched routing graph for pedestrians
- Not really required but often available information
  - Prior potential attractivity

Data requirement: Localization

- Data from communication network infrastructure
  - GSM traces (Calabrese et al. 2011, Bekhor et al. 2011)
  - WiFi traces
- Data processing is needed (Rieser-Schüsseler 2012)
  - Detection of stop points
  - Activity purpose detection through land-use information and spatial matching

Data requirement: Pedestrian network

- We need maps
  - With **points of interests** (space)
  - With **shortest path** (time)
- More and more available in airports, malls, museums, campuses, hospitals
  - **Nokia**: 214 shopping malls in 2011, 4605 indoor maps in July 2012, 5100 in December 2012
  - **Microsoft**: 2700 indoor maps
  - **Google**: > 10'000 indoor maps
  - **Start-ups**: Wifarer, Meridian, Point Insider, ByteLight

Data requirement: Potential attractivity

- **Potential attractivity C(x,t)** depends on
  - **destination x**
    - Classroom, platform, scene, …
  - **time t**
    - class schedules, train schedules, opening hours, …
- Examples:
  - 1500 passengers on platform 4 arriving at 16h04
  - 32 students in a classroom from 8h15 to 10h
  - 400 seats in a restaurant open from 11h to 14h30

METHODOLOGY

Methodology

- **Goal:** extract the possible activity-episodes performed by pedestrians from digital traces from communication networks

Input

- Localization measurement
- Potential attractivity
- Semantically-enriched routing graph

Output

- set of candidate activity-episodes sequences associated with the likelihood to be the true one

Definitions / Notations

- Measurement: $\hat{s} = (\hat{x}, \hat{t})$
- Activity-episode: $a = (x, t^-, t^+)$
- Episode location, start time and end time
- Activity-episode sequence: $(a_1, ..., a_m) = a_{1:m}$
- Activity: $A(a)$
- Activity pattern: $(A_1, ..., A_m) = A_{1:m}$

Methodology

- Probabilistic measurement model:
  A Bayesian approach
  - Measurement equation
  - Prior
- Generation of activity-episode sequences
  - Episode location
  - Episode start and end times

Probabilistic measurement model

Measurement likelihood

Prior

$$P(a_{1:m}|\hat{s}_{1:n}) \propto P(\hat{s}_{1:n}|a_{1:m}) \cdot P(a_{1:m})$$

Activity model

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

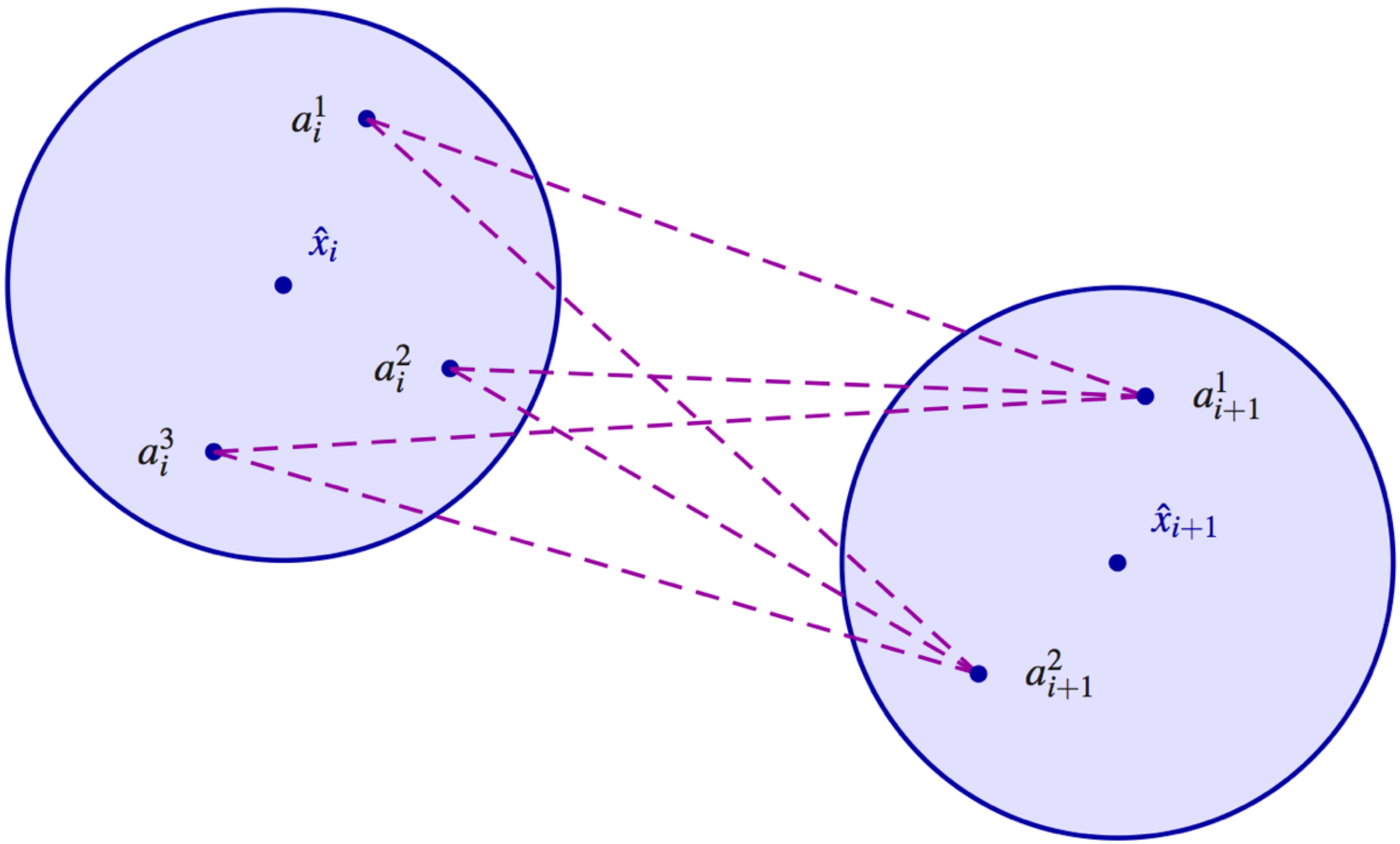
Measurement error

$$
\begin{aligned}
P(\hat{s}_{1:n}|a_{1:m}) &= \prod_{j=1}^{m} P(\hat{s}_{i_{j-1}+1:i_j}|a_j) && \longleftarrow \text{ Independence between activities} \\
&= \prod_{j=1}^{m} \prod_{i=1}^{n} P(\hat{s}_{i_j}|a_j) && \longleftarrow \text{ Independence between signals} \\
&= \prod_{j=1}^{m} \prod_{i=1}^{n} P(\hat{x}_{i_j}|x_j) && \longleftarrow \text{ No time measurement error}
\end{aligned}
$$

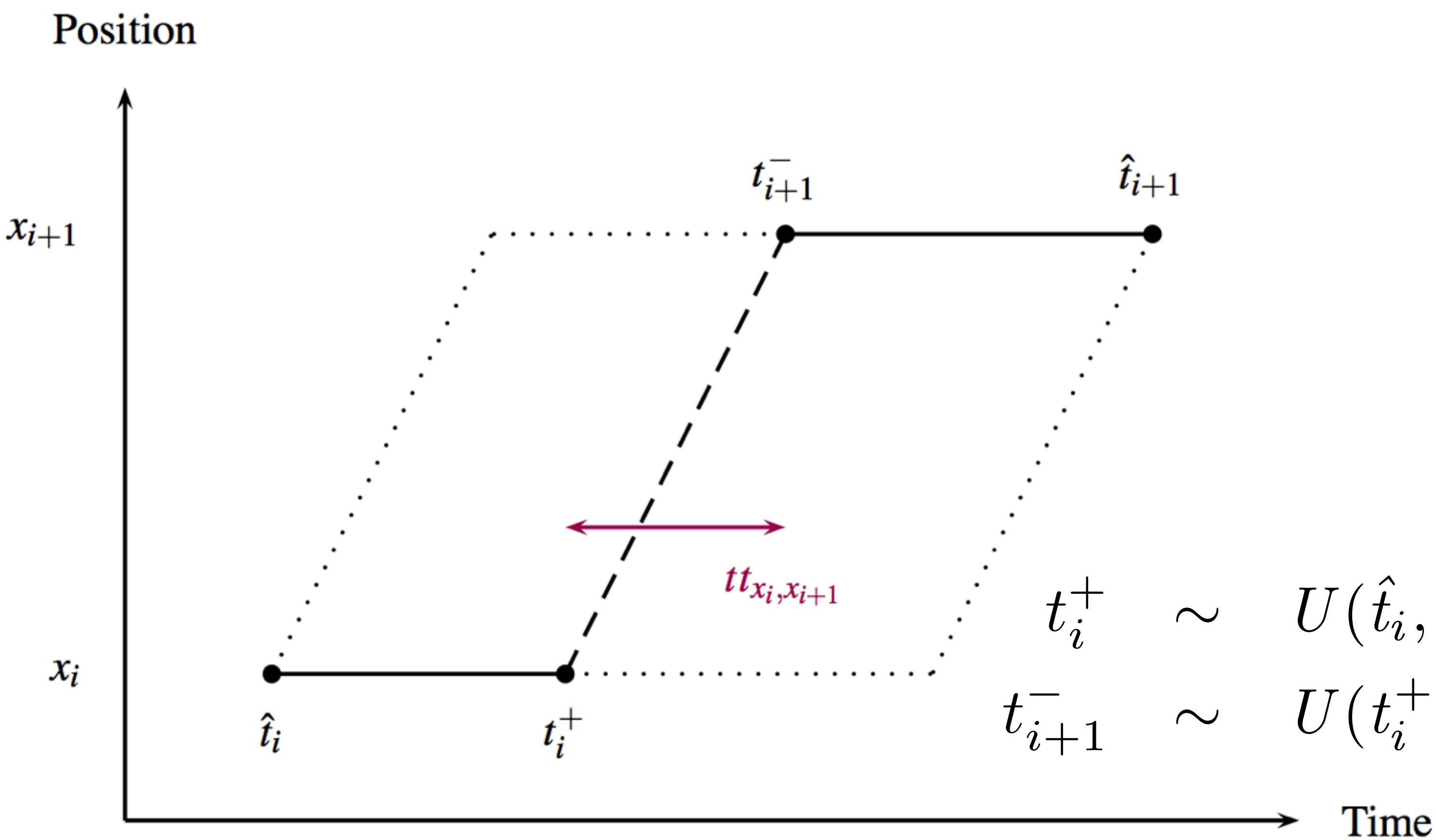
Prior

$$
\begin{aligned}
P(a_{1:m}) &= \prod_{j=1}^{m} P(a_j) && (1)\\
&= \prod_{j=1}^{m} P(x_j, t_j^-, t_j^+) && (2)\\
&= \prod_{j=1}^{m} \frac{C_{x_j}(t_j^-, t_j^+)}{\sum_{x \in X} C_x(t_j^-, t_j^+)} && (3)\\
&= \prod_{j=1}^{m} \int_{t_j = t_j^-}^{t_j^+} \frac{C_{x_j}(t_j)}{\sum_{x \in X} C_x(t_j)} dt && (4)
\end{aligned}
$$

Generation of activity-episode sequences

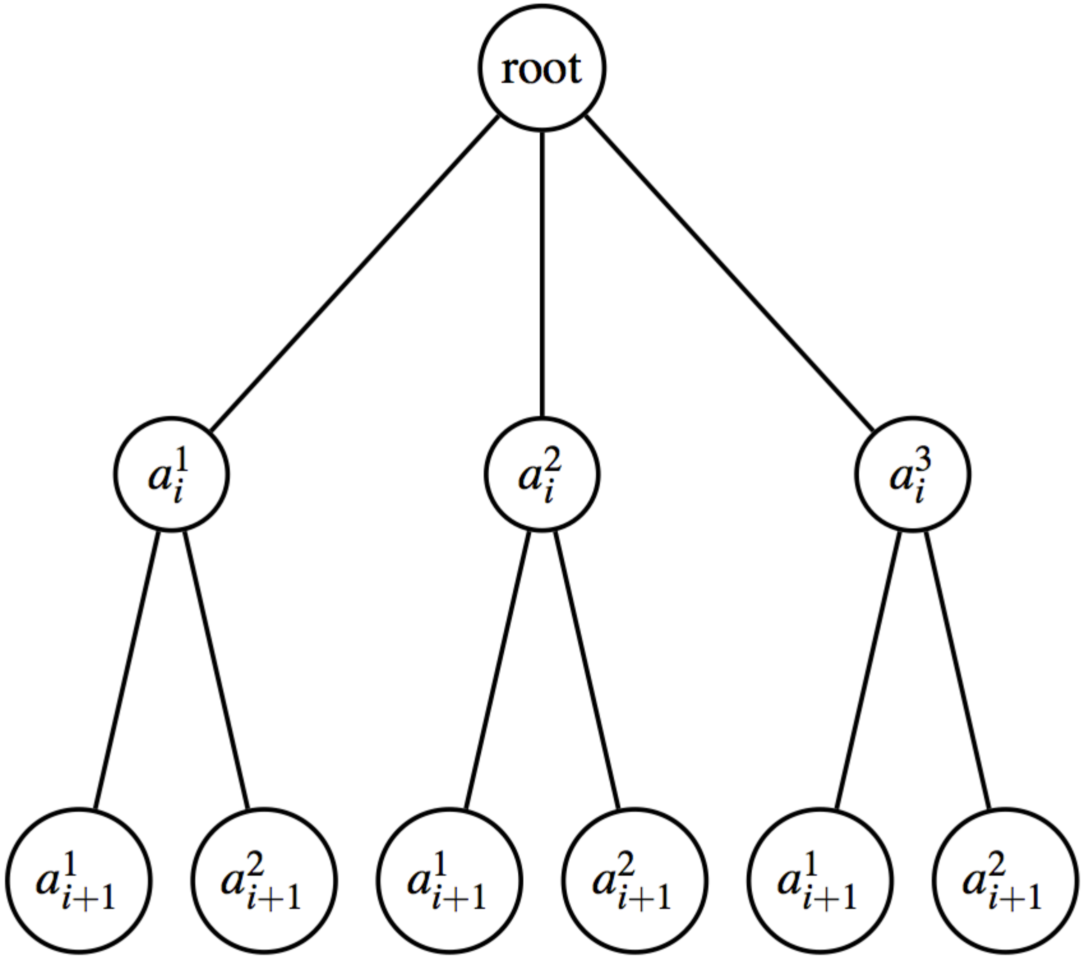


Generation of activity-episode sequences



$$
\begin{aligned}
t_i^+ &\sim U(\hat{t}_i, \hat{t}_{i+1} - tt_{x_i,x_{i+1}}) \\
t_{i+1}^- &\sim U(t_i^+ + tt_{x_i,x_{i+1}}, \hat{t}_{i+1})
\end{aligned}
$$

Generation of activity-episode sequences



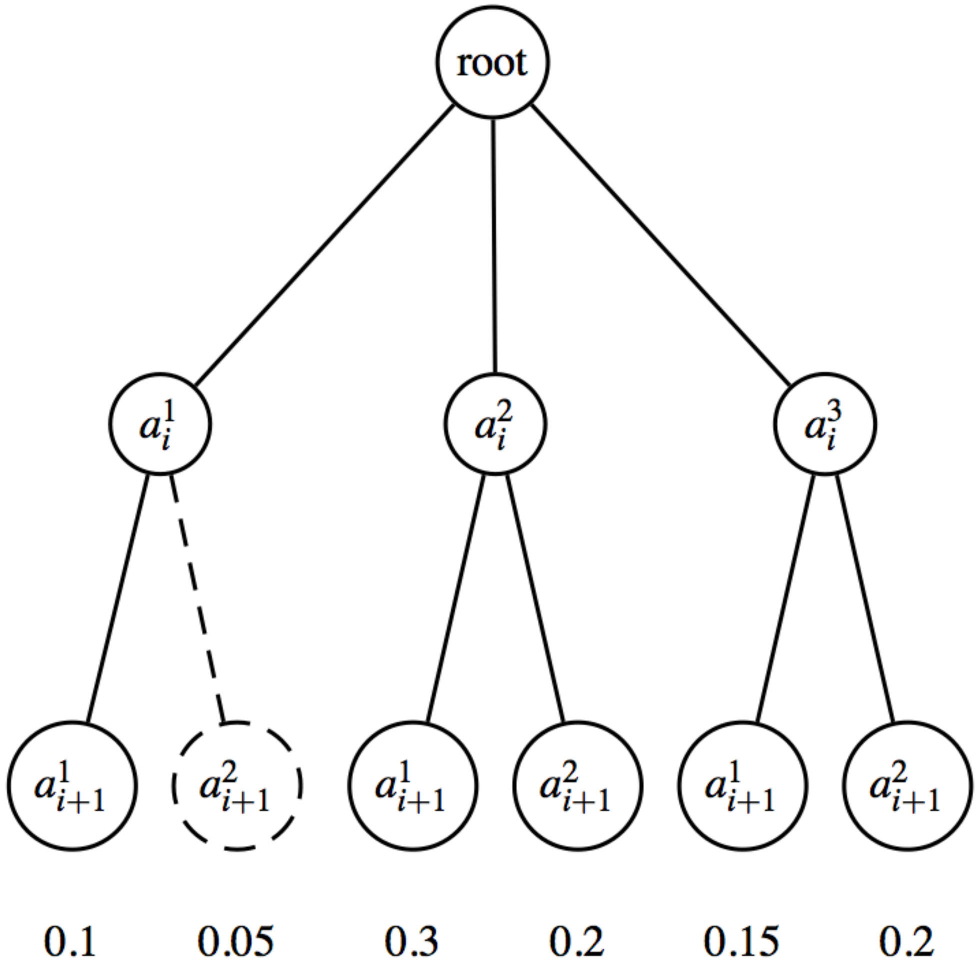
Intermediary signals

- Eliminate intermediary signal if

$$E(t^+) - E(t^-) < T_{min}$$

since we generate an activity episode at each signal.

Sequence elimination



A CASE STUDY ON EPFL CAMPUS

EPFL data: Localization

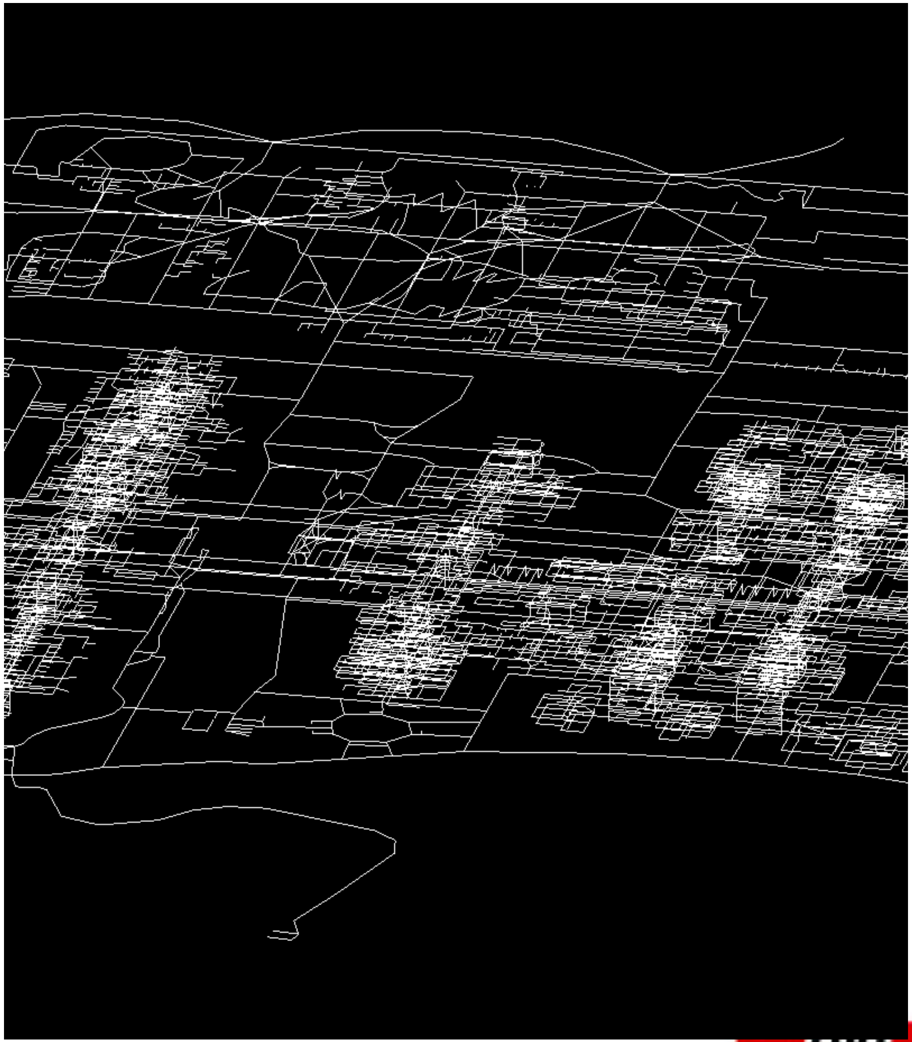
- 8 participants for 2 months with known ID
- Non-participants: 46 days, but only 10 with courses
  - 200 students in 6 different classes
  - 317 employees
  - 700 students from University of Lausanne
- For 151 CE students, 152’598 observations
- Precision: 191 meters

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

EPFL data: Pedestrian network

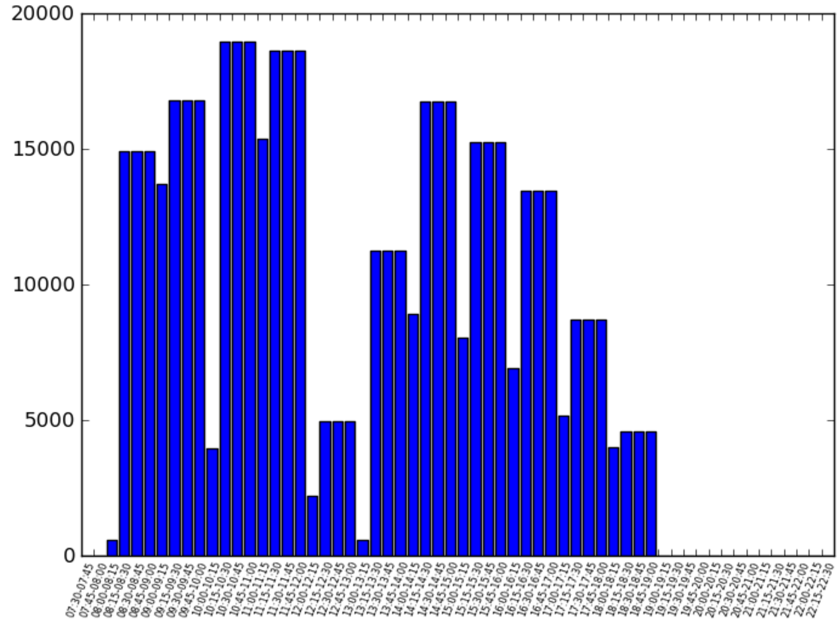
- Source: map.epfl.ch
- 56’655 edges
- 4 different levels of path
  - Major (« highway »)
  - Inter-building
  - Intra-building
  - Access to offices
- Shortest path
- All offices, restaurants, classrooms and other points of interest are coded: X



ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

EPFL data: Potential attractivity C(x,t)

- Class schedules with
  - Number of students
  - Name of the classroom
- Number of employees per office
  - Name of the office
  - Sum of percent of work (e.g, 3 full times = 300%)
- Number of seats in restaurants
  - Localization
  - Opening hours
- Number of seats in library



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Results

Model				Truth			Δx
Arrival time	*Departure time*	*Floor*	*Location*	*Time spent*	*Floor*	*Location*	*(in m.)*
8:33-8:33	10:38-10:38	1	Classroom	8.32am-10.30am	1	Classroom	0
10:40-10:40	11:51-11:51	3	Office	Until 11.47am	3	Author's office	7
11:54-11:54	12:47-12:53	1	Restaurant	From 11.55 am	1	Restaurant	0
12:51-12:58	13:03-13:44	3	Office	Around 1pm	3	Author's office	7
13:06-13:47	13:53-14:02	2	Cafeteria	Around 2pm	2	Cafeteria	0
13:55-14:04	19:45-19:45	3	Office	Until around 7.45pm	3	Author's office	7
19:47-19:47	19:52-19:52	3	Workshop	-	3	Metro stop	366

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

SENSITIVITY ANALYSIS

Sensitivity analysis: prior

- With flat prior
  - # destinations / Start and end time: OK
  - Distance / category of destination: Not OK
- Attractivity of visited destinations should be 3x bigger than of non-visited destinations
- Global capacity creates bias

CONCLUSION

Conclusion

- Prior needed to **overcome low precision**
- **Localization data brings dynamics** in the model
- Pedestrian map gives:
  - Spatial information
  - Temporal information
- Our methodology is **merging** these different types of data

- Robust for **low density data**

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

FUTURE WORK

Future work

- Binary choice model for attendance of scheduled activity
- Actual start and end times of scheduled activity-episodes
- Analysis of the access to the facility
  - First and last destination of the sequence
  - Arrival times, departure times
- Based on class attendance and on available time budget: activity scheduling

THANK YOU

References

- Bekhor, S., Cohen, Y. and Solomon, C. (2011), Evaluating long-distance travel patterns in Israel by tracking cellular phone positions. J. Adv. Transp.. doi: 10.1002/atr.170
- Calabrese, F.; Di Lorenzo, G.; Liang Liu; Ratti, C., "Estimating Origin-Destination Flows Using Mobile Phone Location Data," Pervasive Computing, IEEE , vol.10, no.4, pp.36,44, April 2011 doi: 10.1109/MPRV.2011.41
- Rieser-Schüssler, N. (2012). Capitalising Modern Data Sources for Observing and Modelling Transport Behaviour, Transportation Letters: the International Journal of Transportation Research (2): 115–128. doi: 10.3328/TL.2012.04.02.115-128

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE