Detecting pedestrian destinations from ubiquitous digital footprints

Antonin Danalet
TRANSP-OR, EPFL

FCL-Talk
January 22nd, 2013

Transport and Mobility lab

- People: Prof. Michel Bierlaire, 4 postdocs, 9 PhD students
- Research streams:
  - Transportation Research,
  - Operations Research,
  - Discrete Choice Models
- 1 multidisciplinary subgroup: pedestrian dynamics
- 2 projects related to pedestrians:
  - Swiss National Science Foundation
  - Swiss Railways (SBB-CFF-FFS)

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Image in the (Swiss) population

4. PERCEPTIONS DES MODES DE TRANSPORT

Pour chacun des moyens de transport suivants, indiquez 3 adjectifs qui, selon vous, les décrivent le mieux.

		Adjectif 1	Adjectif 2	Adjectif 3
1	La VOITURE est :	Comfortable 293	Convenient 294	Fast 295
2	Le TRAIN est :	Comfortable 296	Fast 297	Expensive 298
3	Le BUS, METRO et TRAM sont :	Convenient 299	Fast 300	Crowded 301
4	Le CAR POSTAL est :	Comfortable 302	Convenient 303	Expensive 304
5	Le VELO est :	Healthy 305	Sporty 306	Green 307
6	MARCHER est :	Healthy 308	Slow 309	Pleasant 310

Source: Optima, Projet de recherche sur la mobilité combinée :
Rapport définitif de l'enquête de préférences révélées, EPFL
http://transport.epfl.ch/optima

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Challenge

Presentation outline

- **Activity-based modeling:**
  From car to pedestrians
- **The importance of data:**
  Nur was gezählt wird zählt
- **A Bayesian approach to mix map data, WiFi traces and (train or class) schedules:**
  A case study on campus
- **Toward an activity-based model for pedestrians**
  How and what can we learn?

Activity-based modeling: From car to pedestrians

Travel demand modeling: 1960s

- Rapid increase in **car** ownership and usage
- Need to assess the impact of investments
- Aggregate, gravity type models
- **4-step model**
  - trip generation
  - trip distribution
  - mode choice
  - route choice

Travel demand modeling: 1970s

Shifting paradigm in travel demand modeling (1): from gravity to discrete choice

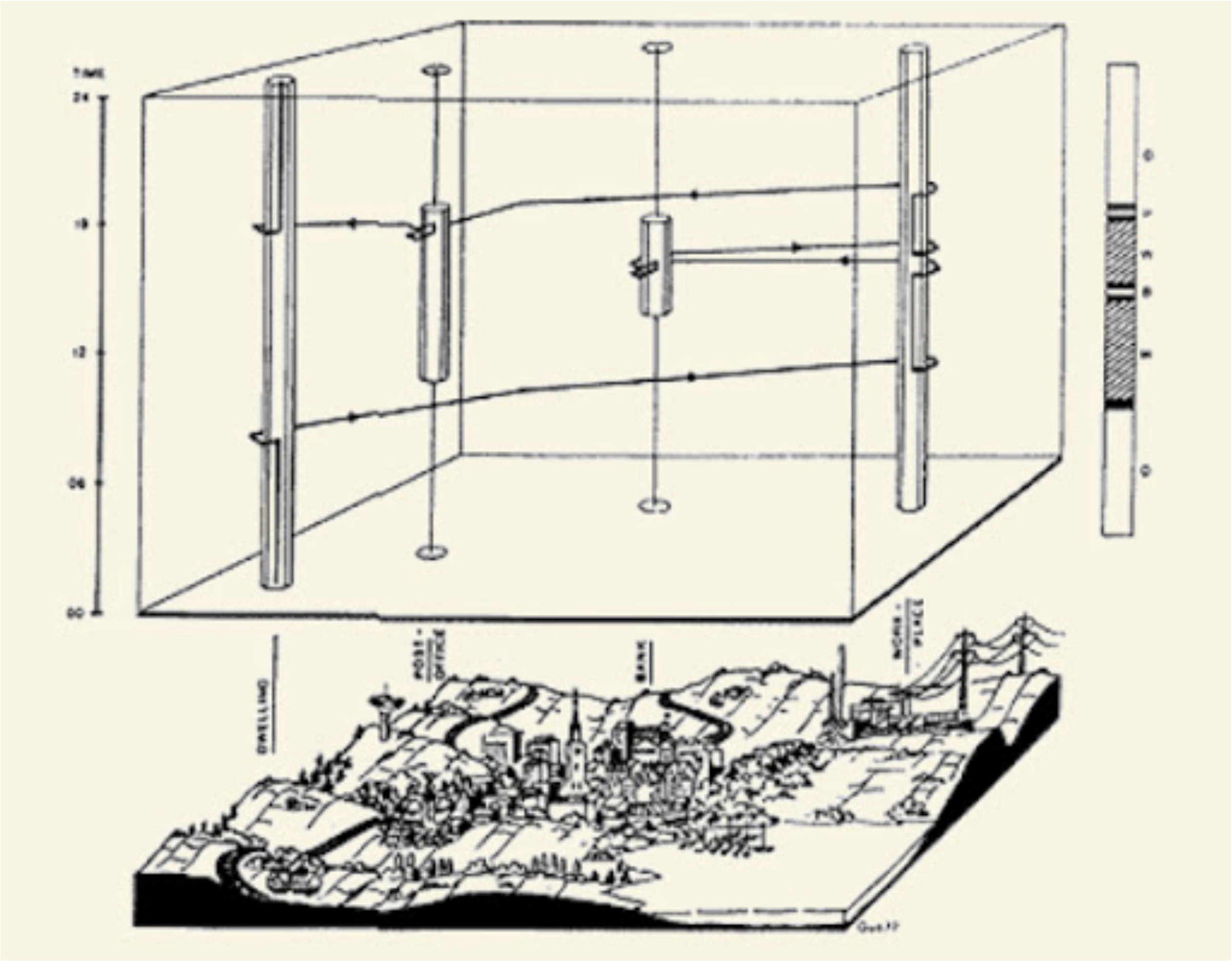
- direct modeling of individual choice behavior, and so more sensitive to policy
- require smaller data sets for calibration
- incorporate more explanatory variables

Travel demand modeling: 1980s

Shifting paradigm in travel demand modeling (2): travel behavior as a derivative of activities

- emphasize on activity scheduling behavior
- compliant with more complex patterns
- compliant with new policies
  (e.g., congestion pricing, ridesharing, ...)

Hägerstrand's Time Geography



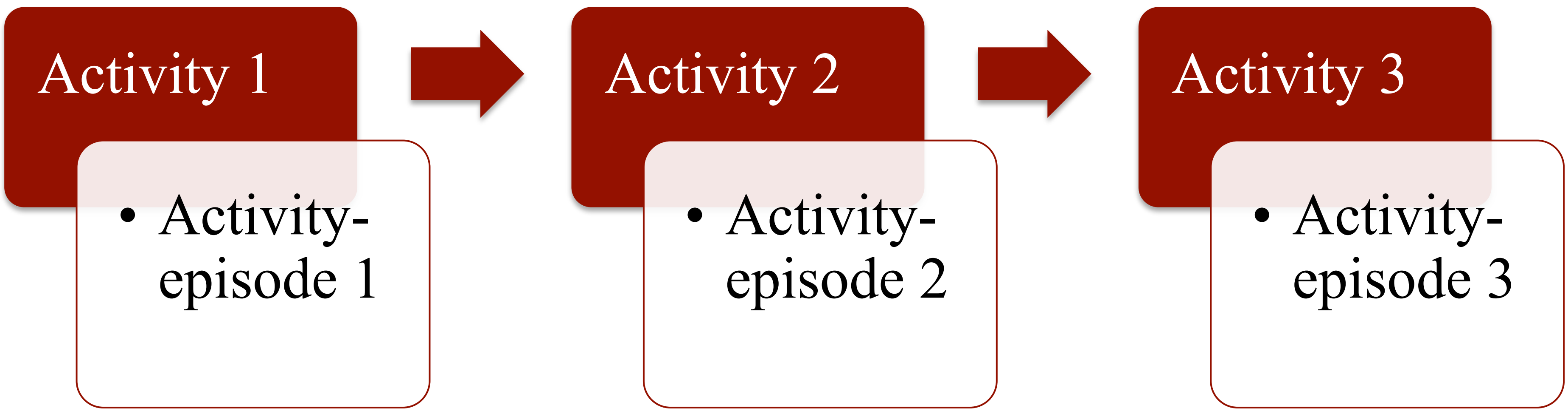
Research on pedestrians

- 3 levels of pedestrian behavior
  - Strategical (destination, activity choices)
  - Tactical (route choice)
  - Operational (walking behavior)
- First developments in mid-90's
- Still a new research area, with a focus on operational level
- TRANSP-OR active in this area, e.g., walking behavior as a choice of next step

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model

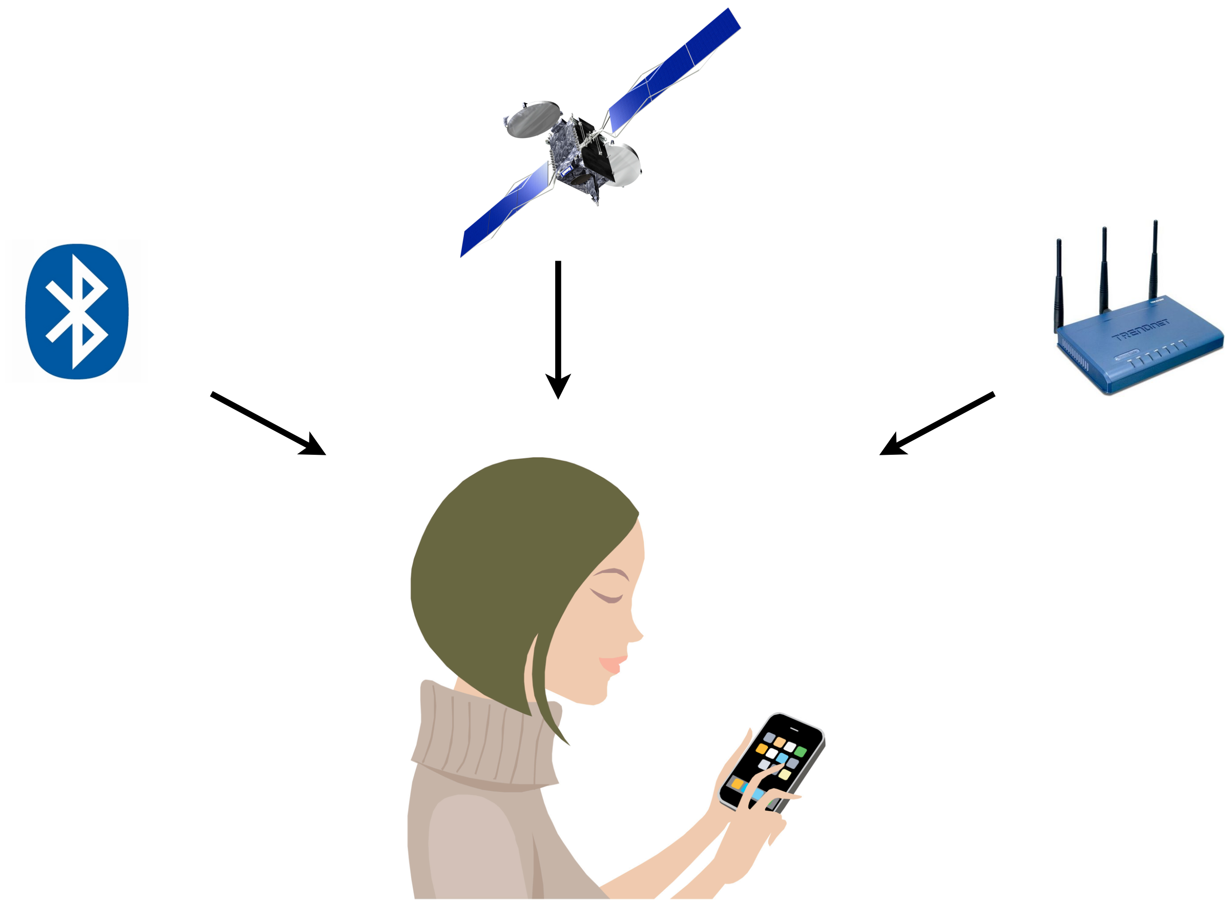

Activity 1
Activity-episode 1
Activity 2
Activity-episode 2
Activity 3
Activity-episode 3

Fachtagung 2010, Fussverkehr Schweiz, Fachverband der FussgängerInnen
(workshop of the Swiss Association of Pedestrians)

(only what you can count counts)

NUR WAS GEZÄHLT WIRD ZÄHLT

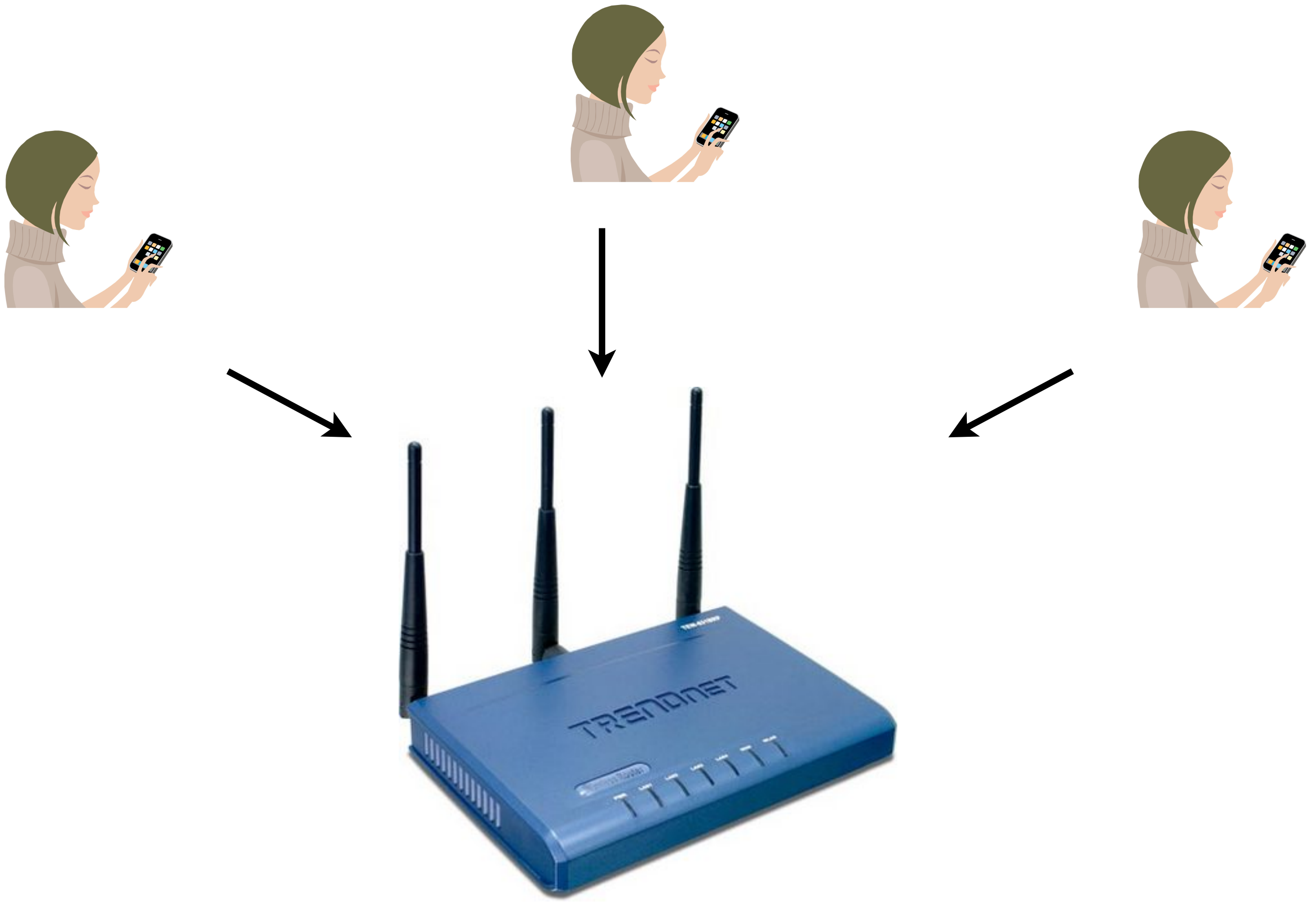
Data: from a device-centric approach…



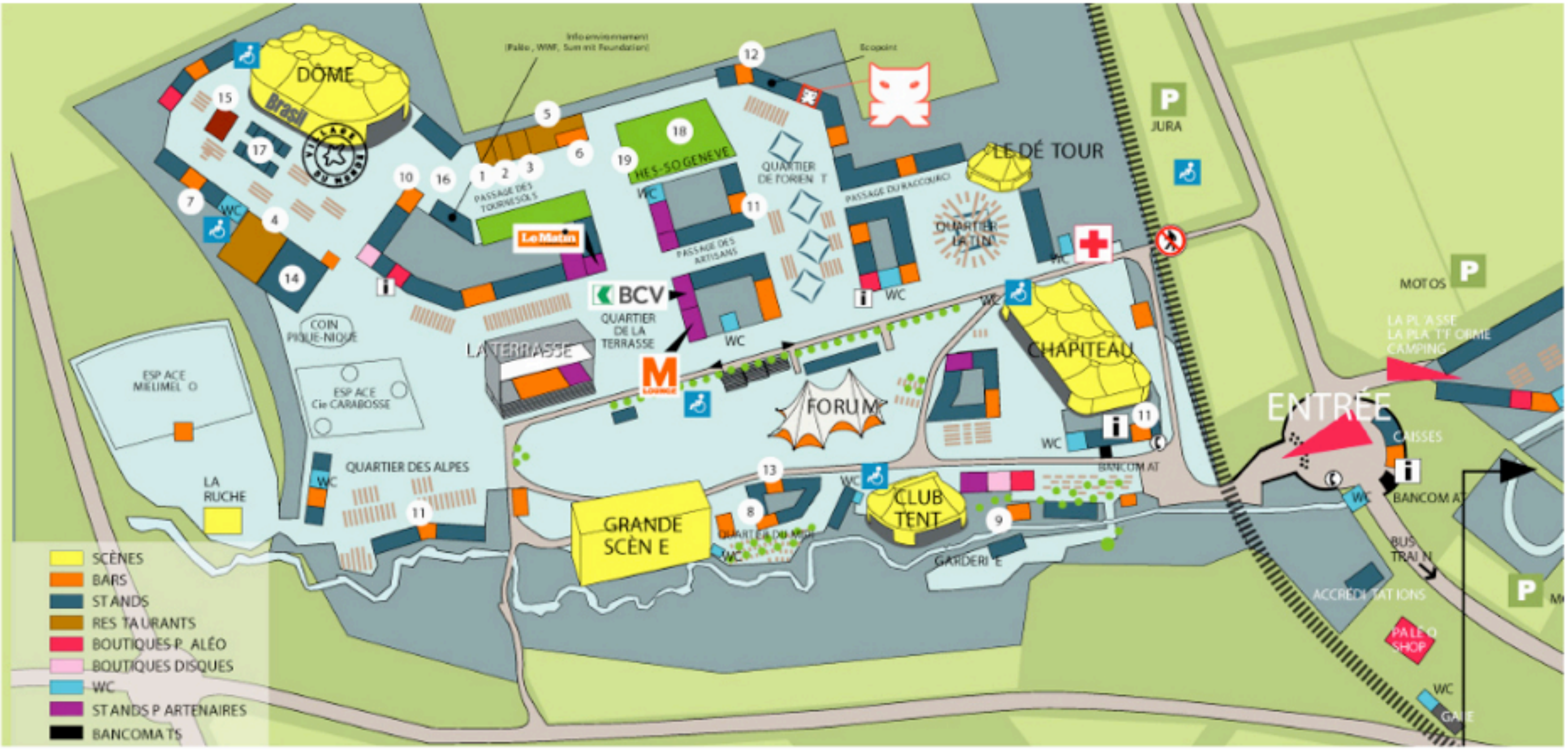
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Data: … to infrastructure traces



Paléo


DÔME
Brasil
LE DÉ TOUR
JURA
HES-SO GENEVE
QUARTIER DE L'ORIENT
PASSAGE DU RACCOURCI
QUARTIER LATIN
PASSAGE DES TOURNESOLS
PASSAGE DES ARTISANS
Le Matin
BCV
QUARTIER DE LA TERRASSE
LA TERRASSE
COIN PIQUE-NIQUE
ESPACE MELIMELO
ESPACE Cie CARABOSSE
LA RUCHE
QUARTIER DES ALPES
FORUM
CHAPITEAU
GRANDE SCÈNE
CLUB TENT
GARDERIE
MOTOS
ENTRÉE
CAISSES
BANCOMAT
BUS TRAIN
ACCREDITATIONS
PALÉO SHOP
GARE
WC
SCÈNES
BARS
STANDS
RESTAURANTS
BOUTIQUES PALÉO
BOUTIQUES DISQUES
WC
STANDS PARTENAIRES
BANCOMATS

Paléo

Paléo

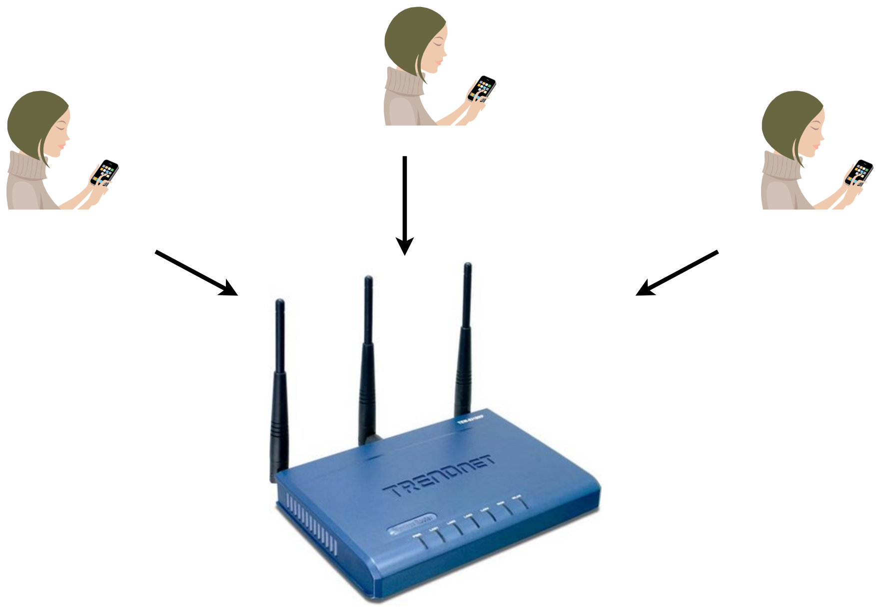
Bluetooth traces from smartphones with GPS

Video not available in PDF format
Please visit:
http://www.youtube.com/watch?v=8Zi45m67jbE

Video: Julien Eberle
http://people.epfl.ch/julien.eberle

Data: F. M. Naini et al., Population Size Estimation Using a Few Individuals as Agents
http://infoscience.epfl.ch/record/169801/files/MovDTV11.pdf

Infrastructure traces: strengths

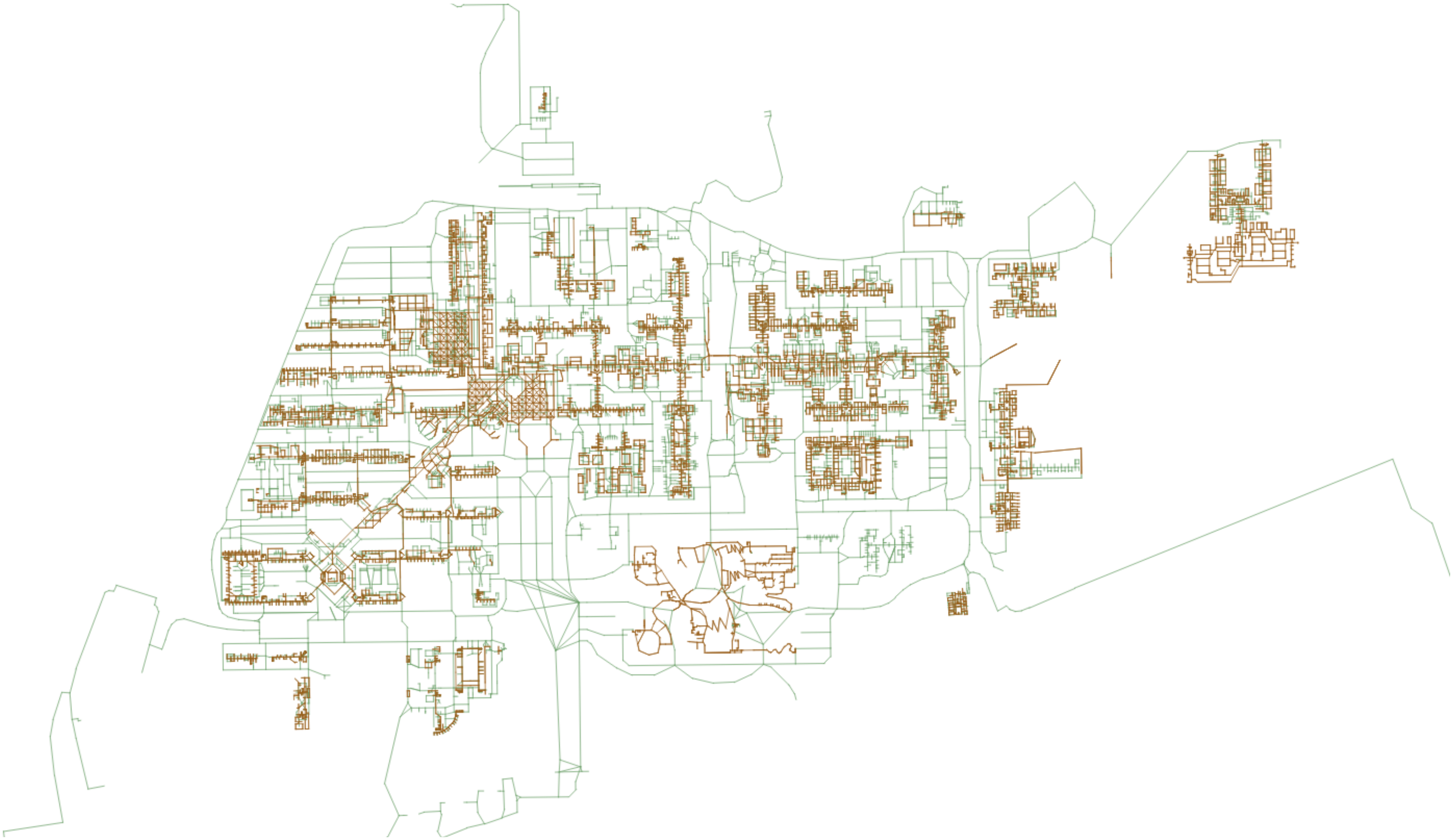


- Data **related to infrastructure**, not to individuals
- Full coverage of the facility is **cheap** and allows for **estimating the overall demand**
- Infrastructure partially **already exists**, and increasing it has **positive side effects**

WiFi Traces on EPFL Campus

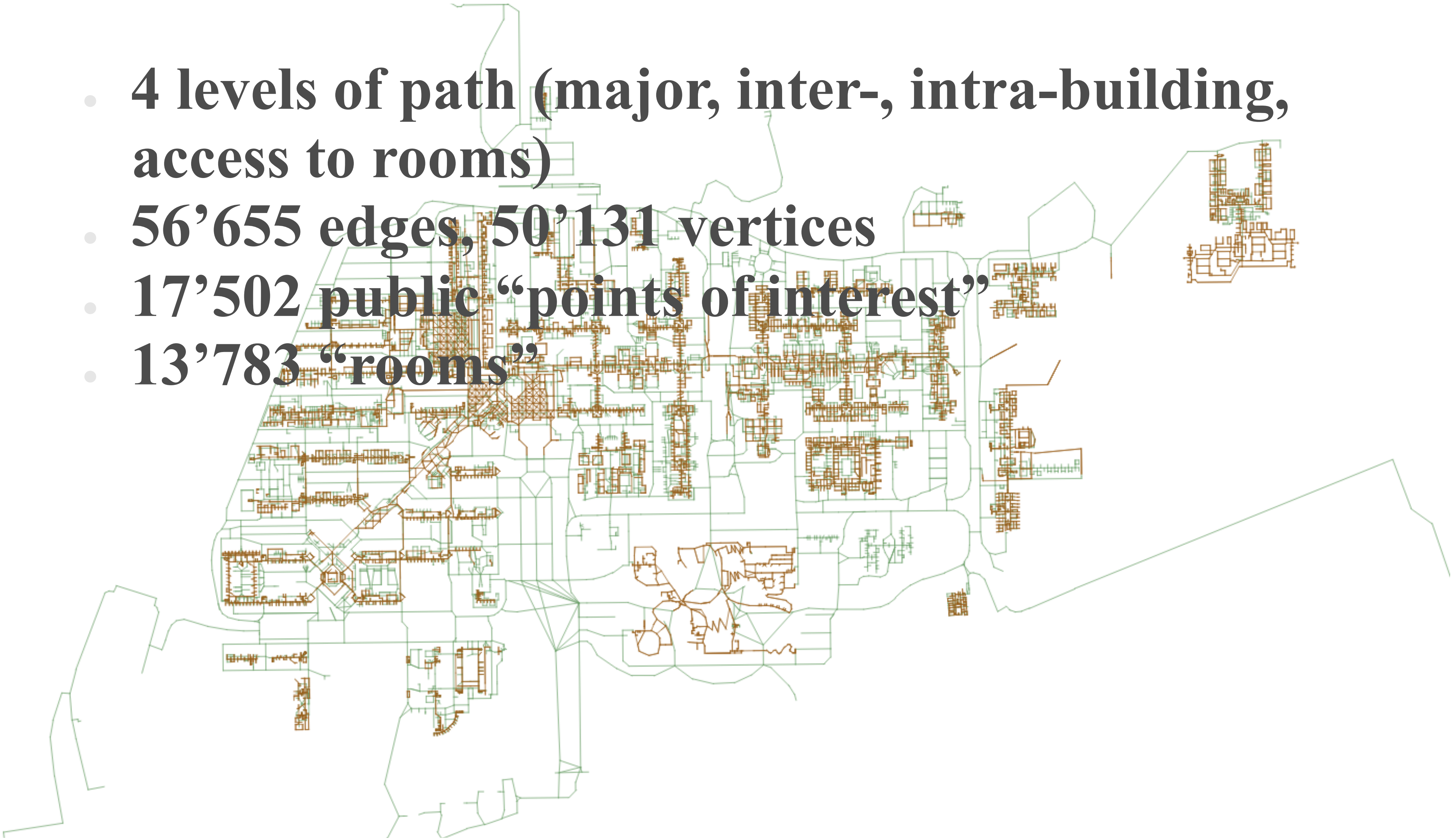
- 789 access points
- 2 datasets:
  - AP to which you are connected
  - Triangulation data (Cisco)
- Pros
  - Already available
  - Covering the full infrastructure
- Cons
  - Low precision

Pedestrian network

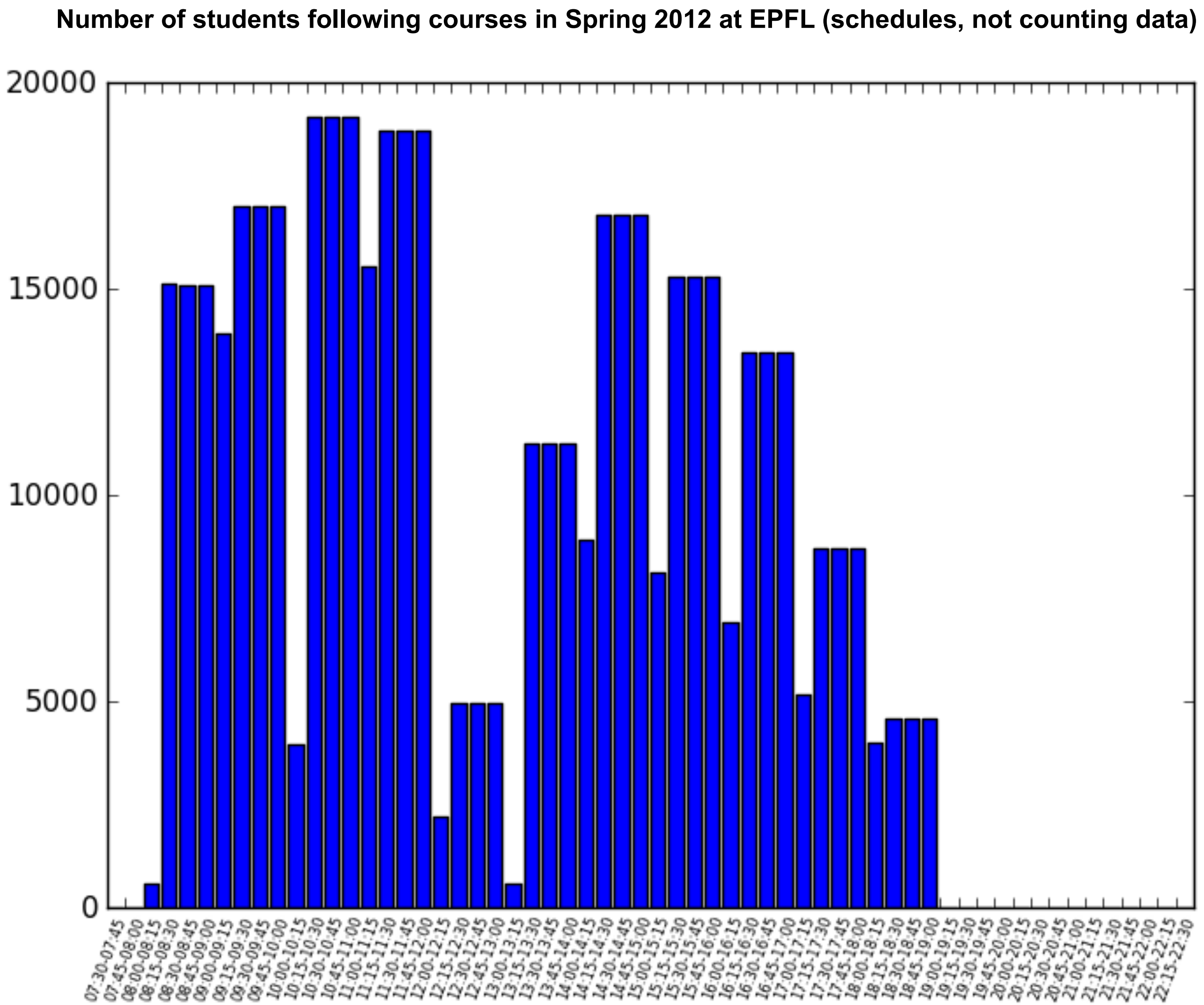


Pedestrian network

- 4 levels of path (major, inter-, intra-building, access to rooms)
- 56’655 edges, 50’131 vertices
- 17’502 public “points of interest”
- 13’783 “rooms”



Class schedules at EPFL



A PROBABILISTIC METHOD FOR ESTIMATING PEDESTRIAN ACTIVITY-EPISODES SEQUENCES

A. Danalet et al., Estimating Pedestrian Destinations using Traces from WiFi Infrastructures
http://infoscience.epfl.ch/record/180079

Probabilistic method

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

Input

- measurement
- prior
- pedestrian map

Output

- set of candidate activity-episodes sequences associated with the probability of being the true one

Probabilistic method

- Probabilistic measurement model
- Generation of activity-episode sequences
- Intermediary signals
- Sequence elimination procedure

Definitions

- 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: $(A_1, ..., A_m) = A_{1:m}$

Probabilistic measurement model

Measurement likelihood

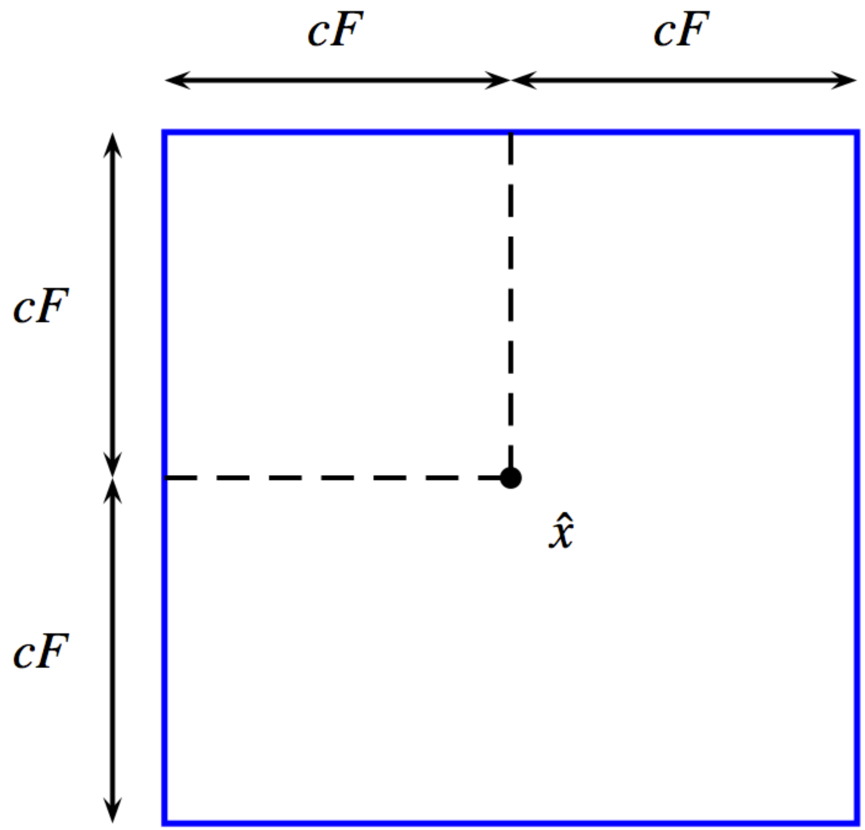
Prior

$$P(a_{1:m}|\hat{s}_{1:n}) = \frac{P(\hat{s}_{1:n}|a_{1:m}) \cdot P(a_{1:m})}{\sum_{a \in A} P(\hat{s}_{1:n}|a_{1:m}) \cdot P(a_{1:m})}$$

Activity model

Probabilistic measurement model

$$P(\hat{s}_{1:n}|a_{1:m}) = \prod_{j=1}^{m} P(\hat{s}_{i_{j-1}+1:i_j}|a_j)$$
$$= \prod_{j=1}^{m}\prod_{i=1}^{n} P(\hat{s}_{i_j}|a_j)$$
$$= \prod_{j=1}^{m}\prod_{i=1}^{n} P(\hat{x}_{i_j}|x_j)$$



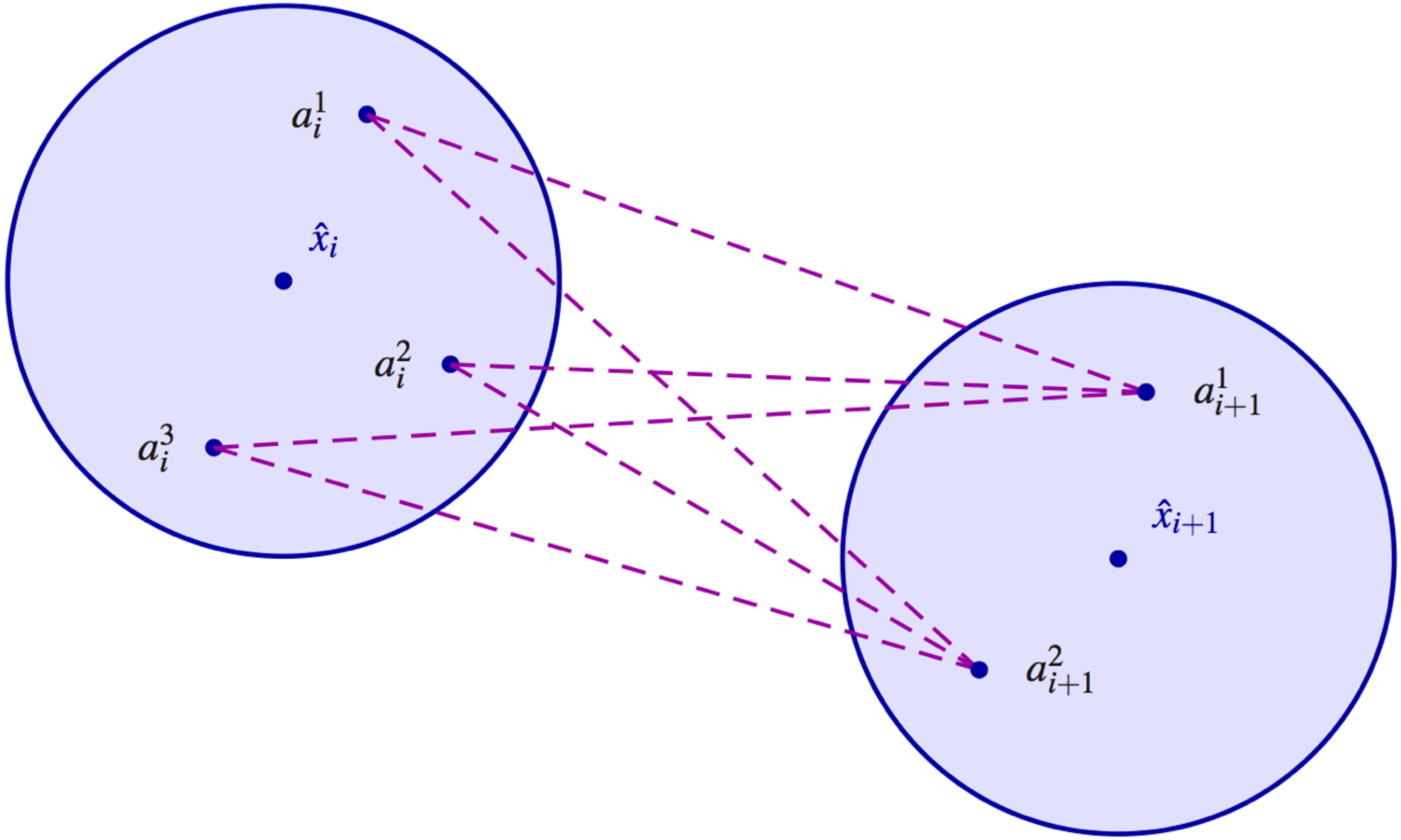
Prior

$$P(x, t^-, t^+)$$

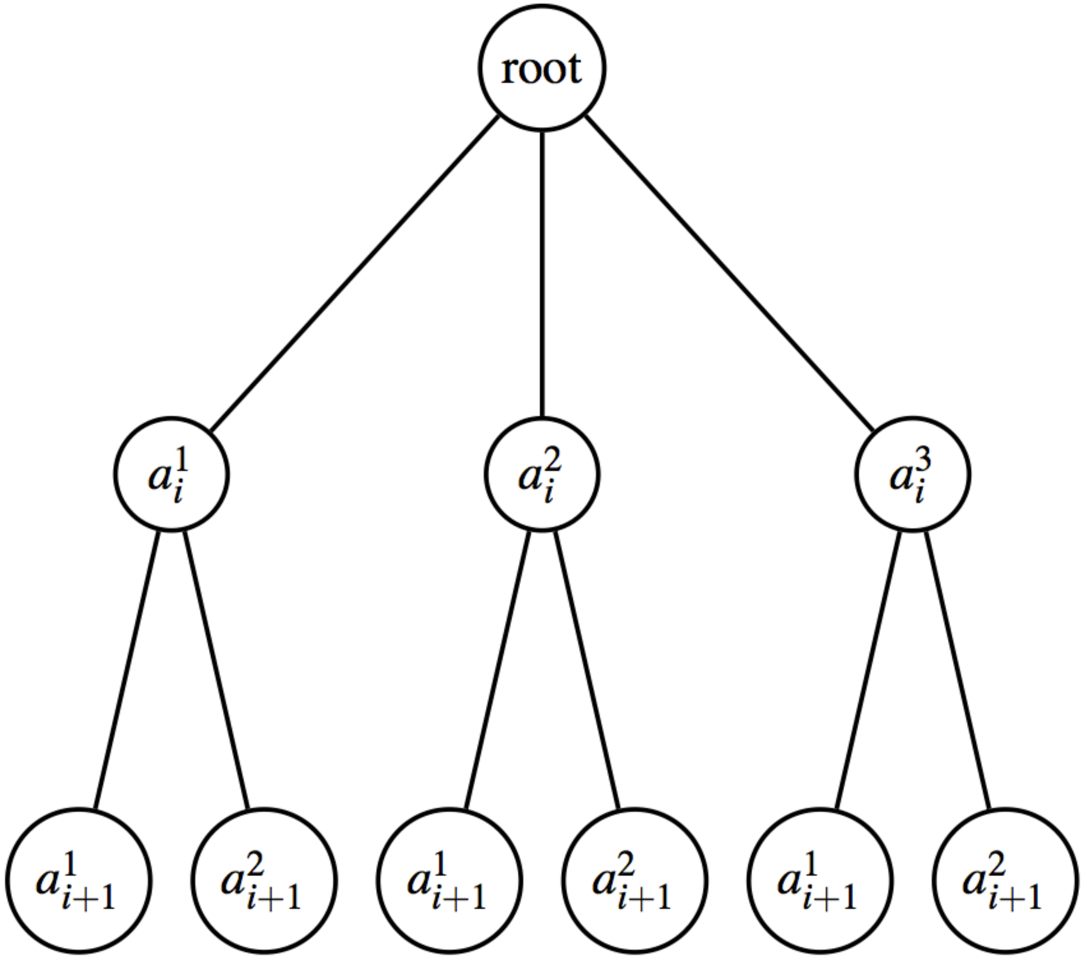
$$P(\text{classroom}) = P(\text{offices}) = P(\text{restaurants}) = P(\text{others}) = \frac{1}{4}$$

$$P(x|\text{classroom}) = \frac{\int_{t=t^-}^{t^+} f_x(t)dt}{\int_{t=t^-}^{t^+} f(t)dt}$$

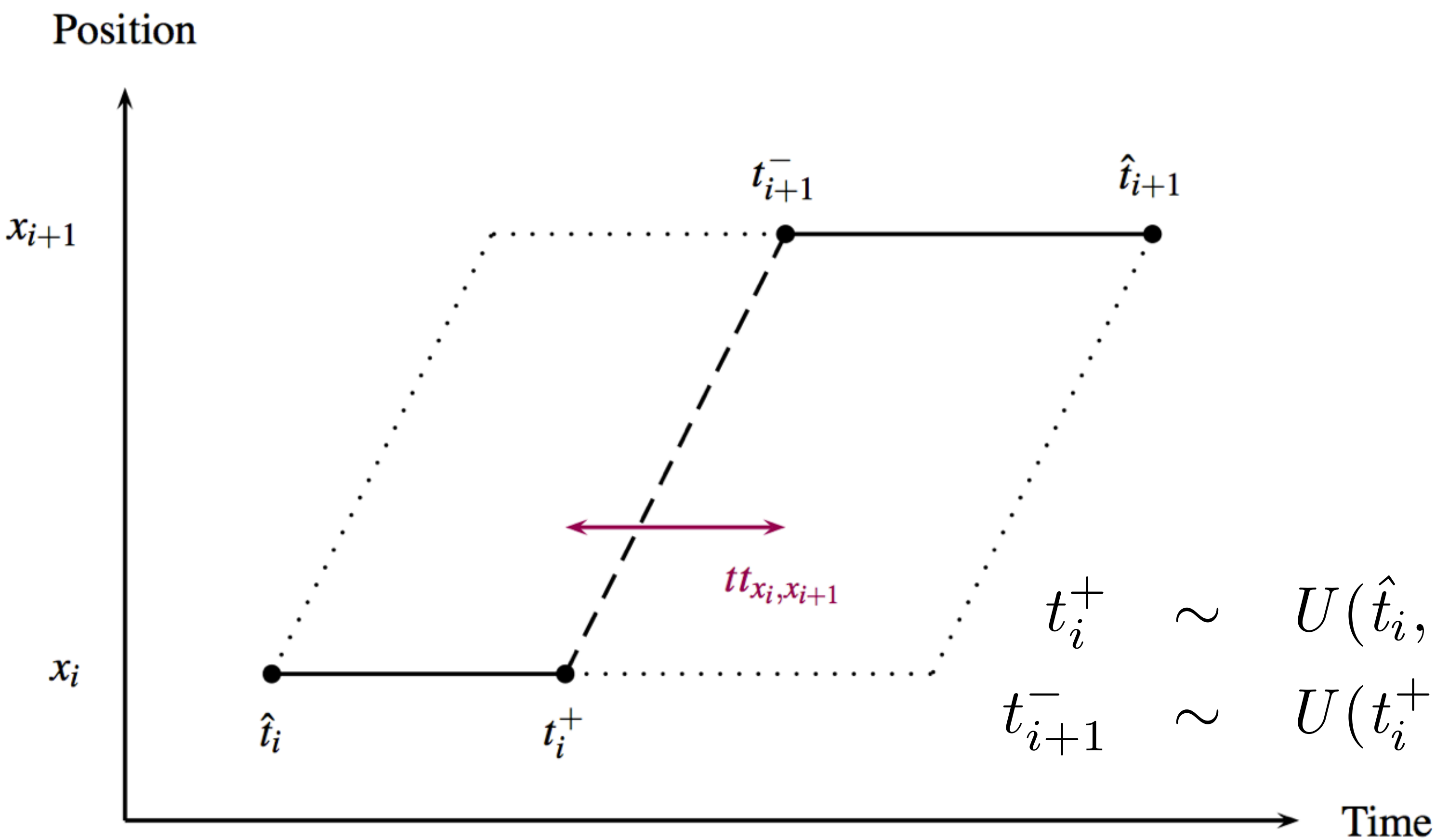
Generation of activity-episode sequences



Generation of activity-episode sequences



Generation of activity-episode sequences



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

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

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Generation of activity-episode sequences

$$f(t_{i+1}^{-}) = \frac{1}{\hat{t}_{i+1} - tt_{x_i, x_{i+1}} - \hat{t}_i} \ln \frac{\hat{t}_{i+1} - tt_{x_i, x_{i+1}} - \hat{t}_i}{\hat{t}_{i+1} - t_{i+1}^{-}}$$

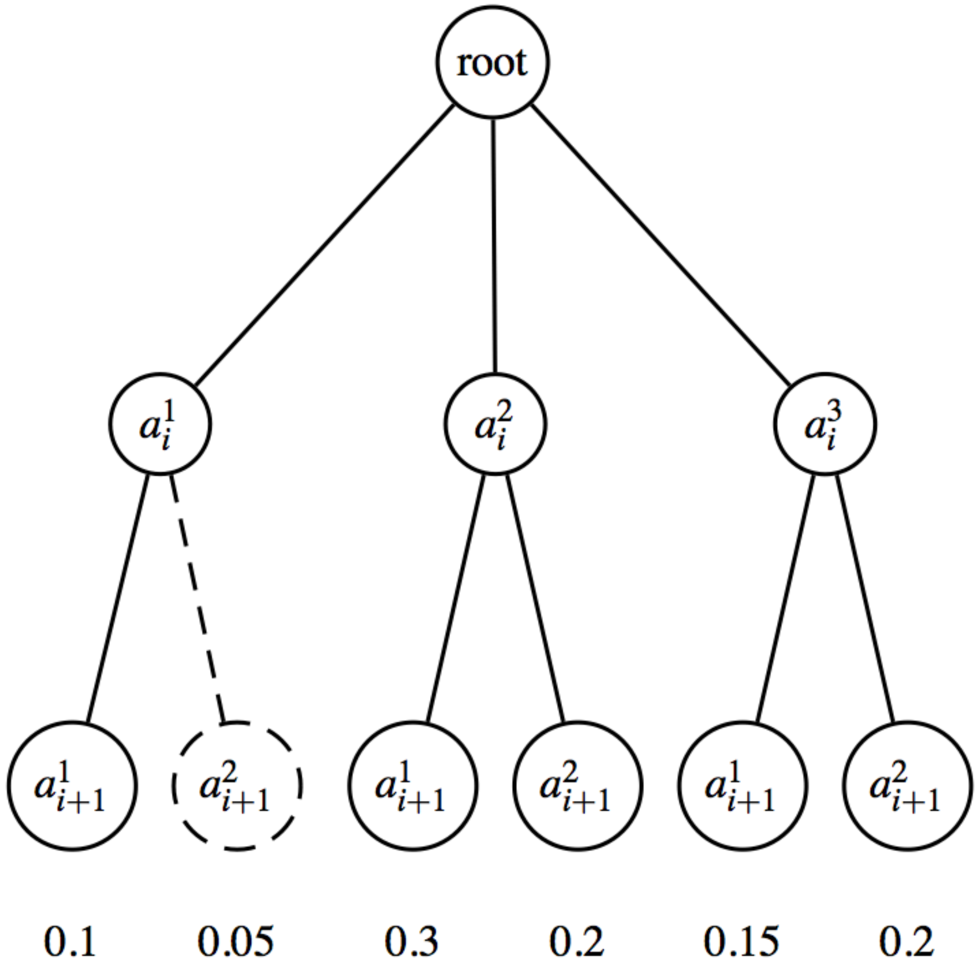
Intermediary signals

- Eliminate intermediary signal if

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

since we generate an activity episode at each signal.

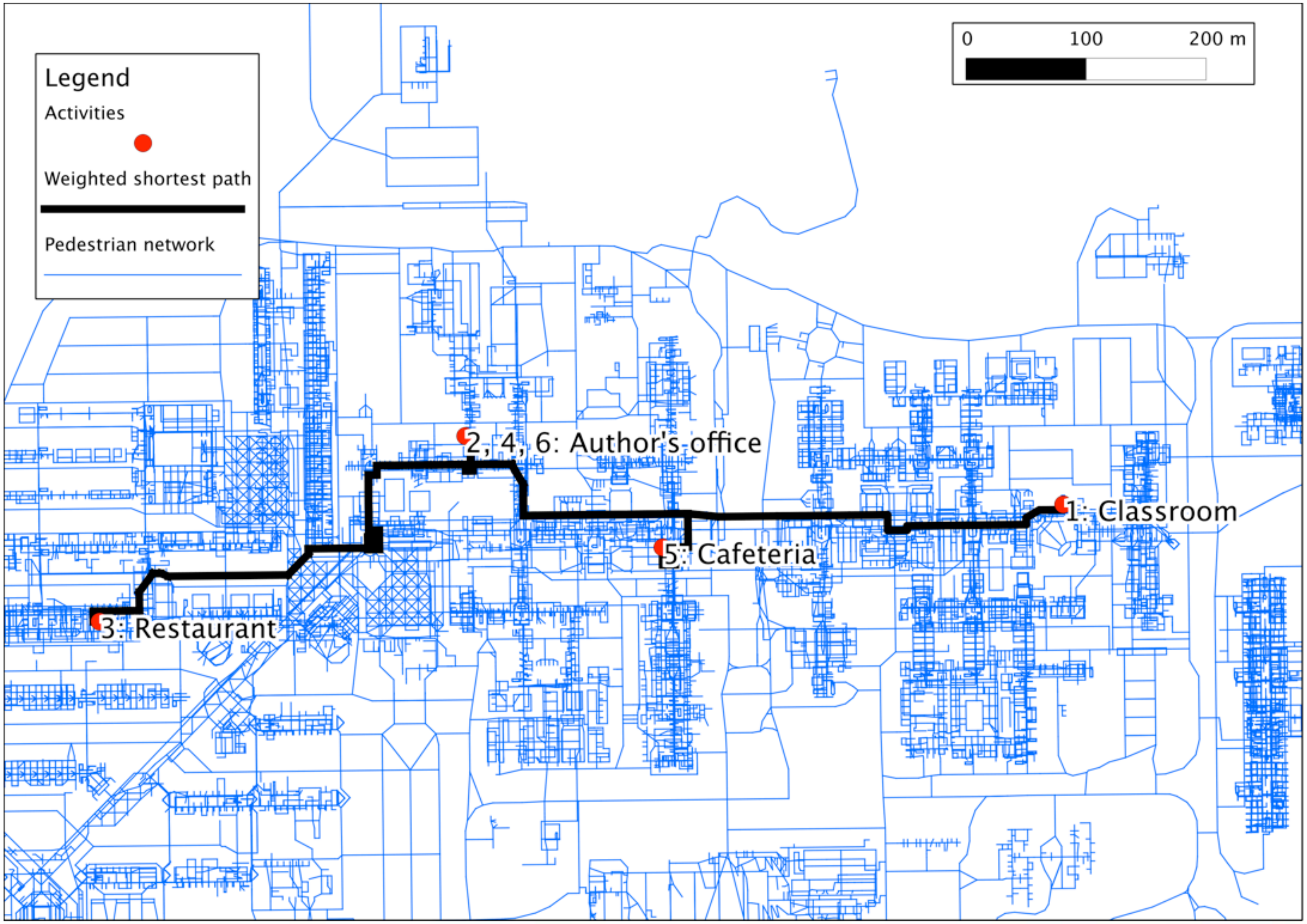
Sequence elimination



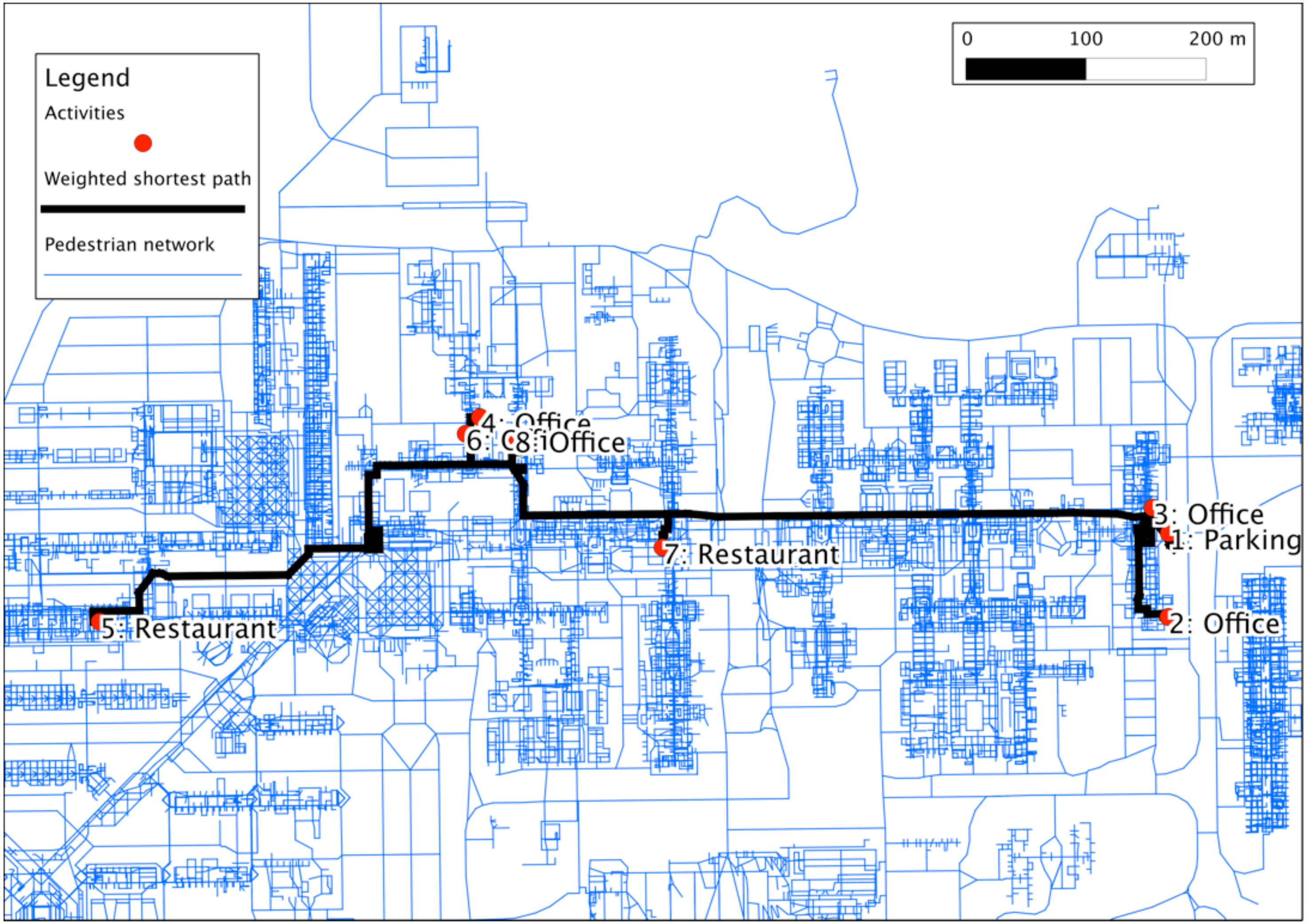
WiFi Traces on EPFL Campus: Results

Model			Truth			Δx
Time spent	*Floor*	*Location*	*Time spent*	*Floor*	*Location*	*(in m.)*
$U(8.40, 8.40)$ - $U(8.40, 9.31)$	0	Classroom	8.32-10.30	1	Classroom	6
$U(8.42, 9.33)$ - $U(10.38, 10.38)$	1	Classroom				72
$U(10.38, 11.18)$ - $U(11.51, 11.51)$	3	Office	Until 11.47	3	Author's office	0
$U(11.52, 12.00)$ - $U(12.47, 12.47)$	2	Classroom	From 11.55	1	Restaurant	16
$U(12.48, 13.03)$ - $U(13.03, 13.44)$	3	Office	Around 13.00	3	Author's office	11
$U(13.06, 13.47)$ - $U(13.53, 13.53)$	2	Restaurant	Around 14.00	2	Cafeteria	0
$U(13.53, 14.10)$ - $U(19.40, 19.42)$	3	Office	Until around 19.45	3	Author's office	8

WiFi Traces on EPFL Campus: Results



WiFi Traces on EPFL Campus: Results



WiFi Traces on EPFL Campus: Results

Video not available in PDF format
Please visit:
http://www.youtube.com/watch?v=SEp-yNXLfUY

How can we learn about activities?

- WiFi infrastructure:
  - Cheap
  - Positive side effects
  - Dense and covers the whole station
- Map knowledge
  - Compensate weakness of localization
- Privacy
  - Data already exist
  - Localization is weak
  - Daily anonymization

What's next?

- Socioeconomic data:
  - Survey?
  - Heterogenous network?
- Develop a location choice model
  - Choice set generation
  - Measurement equation for observations
  - Dynamic of the system
  - Specification of the model
- Move from location choice to activity choice…

What can we learn about activities?

- Overall demand (with calibration)
  - How many people are going on platforms and in shops
- Understanding underlying reasons for activity behavior
  - Why people (don't) go in this shop
- Forecasting demand for activities in case of different scenarios
  - What would happen if this type of shop is moved

Slides and contact information:
http://people.epfl.ch/antonin.danalet