A path choice approach to activity modeling with a pedestrian case study

Antonin Danalet

Statistique, transport et activités, Grenoble
November 5, 2013

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Presentation outline

- **Motivation**: Why pedestrian activities?
- **Detection**: Where are pedestrians?
- **Modeling** pedestrian behavior:
  - Activity-episode sequences and activity patterns
  - Activity network
  - Activity paths
  - Choice set generation
  - Activity path choice model for WiFi traces
- **Conclusion**
- **Future work**

MOTIVATION

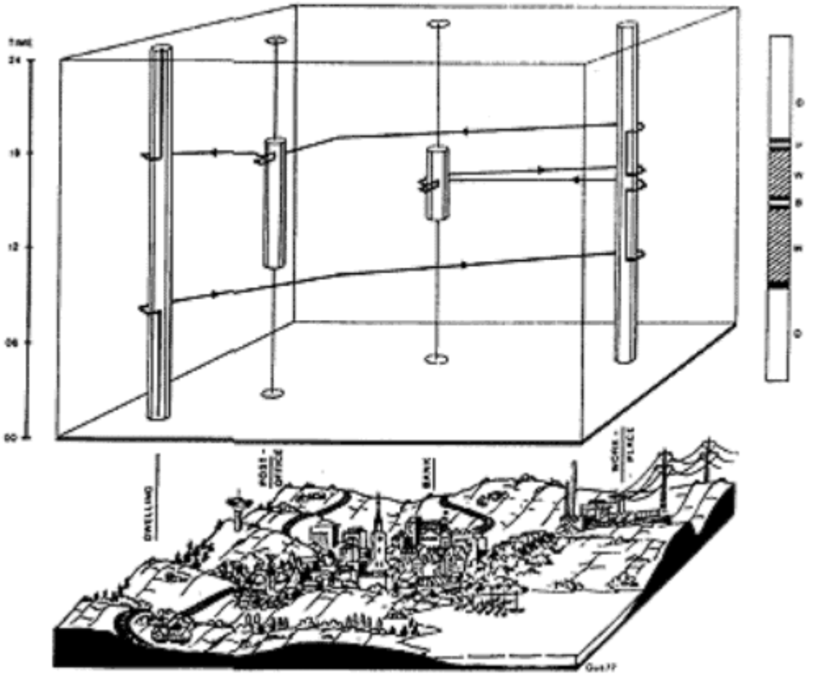
Activity modeling for pedestrian infrastructure

Goal

- Adapt traditional activity modeling framework for pedestrian activities

Challenges

- Detect pedestrians
- Model activity patterns
- Forecast scenarios



Carlstein, T. (1978)

3 examples

- **Multimodal transport hubs**:
  Lausanne railway station
- **Mass gathering**:
  Paléo music festival
- **Campus**:
  EPFL new “Quartier Nord”

Walking is the key for efficient multimodal transport systems

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

Lake Geneva region: Léman 2030

> **25’000** travellers/day between Geneva and Lausanne*

> **50’000** travellers/day between Geneva and Lausanne*

> **100’000** travellers/day between Geneva and Lausanne*

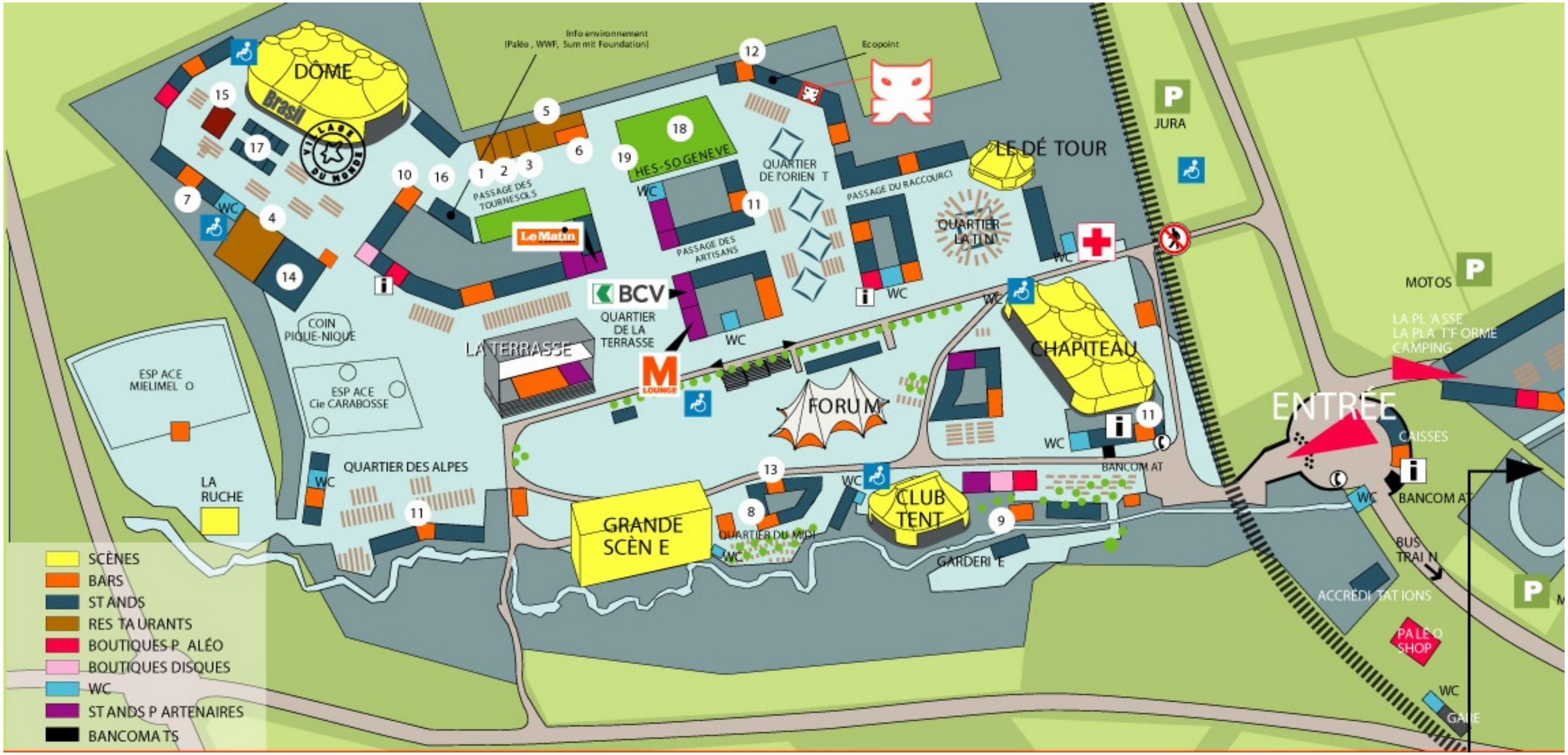
By 2030, 100’000 passengers per day between Geneva and Lausanne

= 2000 travelers/day

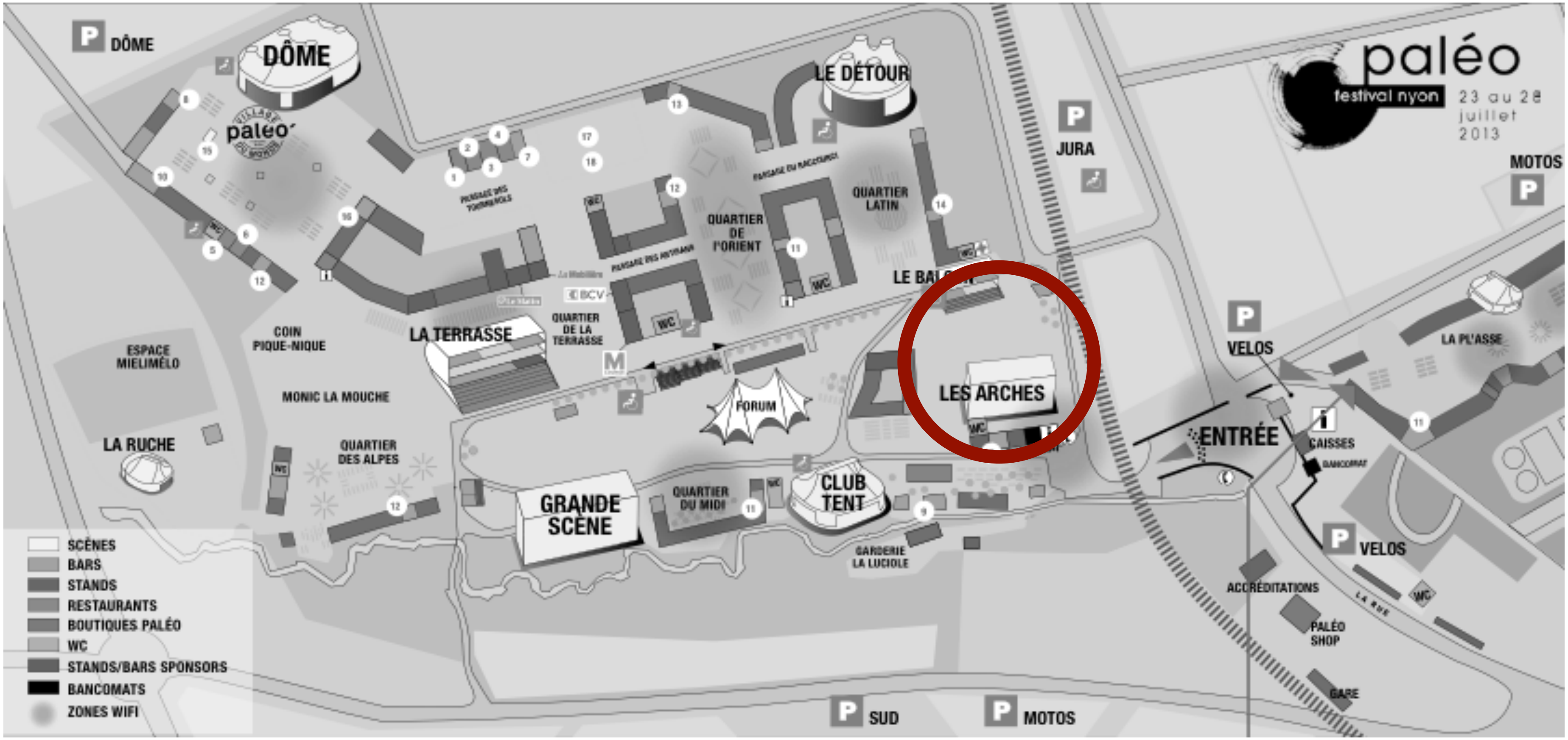
* Forecast by Swiss Railways for the maximum scenario

Mass gathering

Paléo 2012


DÔME
Brasil
VILLAGE DU MONDE
Info environnement (Paléo , WWF, Sum mit Foundation)
Ecopoint
LE DÉ TOUR
JURA
HES-SO GENEVE
QUARTIER DE l'ORIEN T
PASSAGE DES TOURNESOLS
PASSAGE DU RACCOURCI
PASSAGE DES ARTISANS
QUARTIER LATIN
Le Matin
BCV
QUARTIER DE LA TERRASSE
LA TERRASSE
COIN PIQUE-NIQUE
ESP ACE MIELIMEL O
ESP ACE Cie CARABOSSE
LA RUCHE
QUARTIER DES ALPES
FORU M
CHAPITEAU
MOTOS
LA PL ASSE
LA PLA T F ORME
CAMPING
ENTRÉE
CAISSES
BANCOM AT
BUS TRAI N
ACCRÉDI TAT IONS
PALÉ O SHOP
GARE
GRANDE SCÈN E
QUARTIER DU MIDI
CLUB TENT
GARDERI E
WC
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
🟨 SCÈNES
🟧 BARS
■ ST ANDS
🟫 RES TA URANTS
🟥 BOUTIQUES P ALÉO
■ BOUTIQUES DISQUES
🟦 WC
🟪 ST ANDS P ARTENAIRES
⬛ BANCOMA TS

Paléo 2013


DÔME
LE DÉTOUR
JURA
paléo festival nyon
23 au 28 juillet 2013
MOTOS
QUARTIER LATIN
QUARTIER DE L'ORIENT
LE BALCON
LES ARCHES
LA TERRASSE
QUARTIER DE LA TERRASSE
FORUM
ESPACE MIELIMÉLO
COIN PIQUE-NIQUE
MONIC LA MOUCHE
LA RUCHE
QUARTIER DES ALPES
GRANDE SCÈNE
QUARTIER DU MIDI
CLUB TENT
GARDERIE LA LUCIOLE
ENTRÉE
CAISSES
VELOS
LA PL'ASSE
ACCRÉDITATIONS
PALÉO SHOP
GARE
SUD
MOTOS
SCÈNES
BARS
STANDS
RESTAURANTS
BOUTIQUES PALÉO
WC
STANDS/BARS SPONSORS
BANCOMATS
ZONES WIFI

Campus

EPFL
Quartier Nord

Campus


Entrée du parking
LES ARCADES
tl Transports publics de la région lausannoise
Metro M1 Arrêt EPFL
Accès parking

DETECTION

Data input

- Localization data with full coverage of the facility
- Semantically-enriched routing graph for pedestrians
- Potential attractivity measure

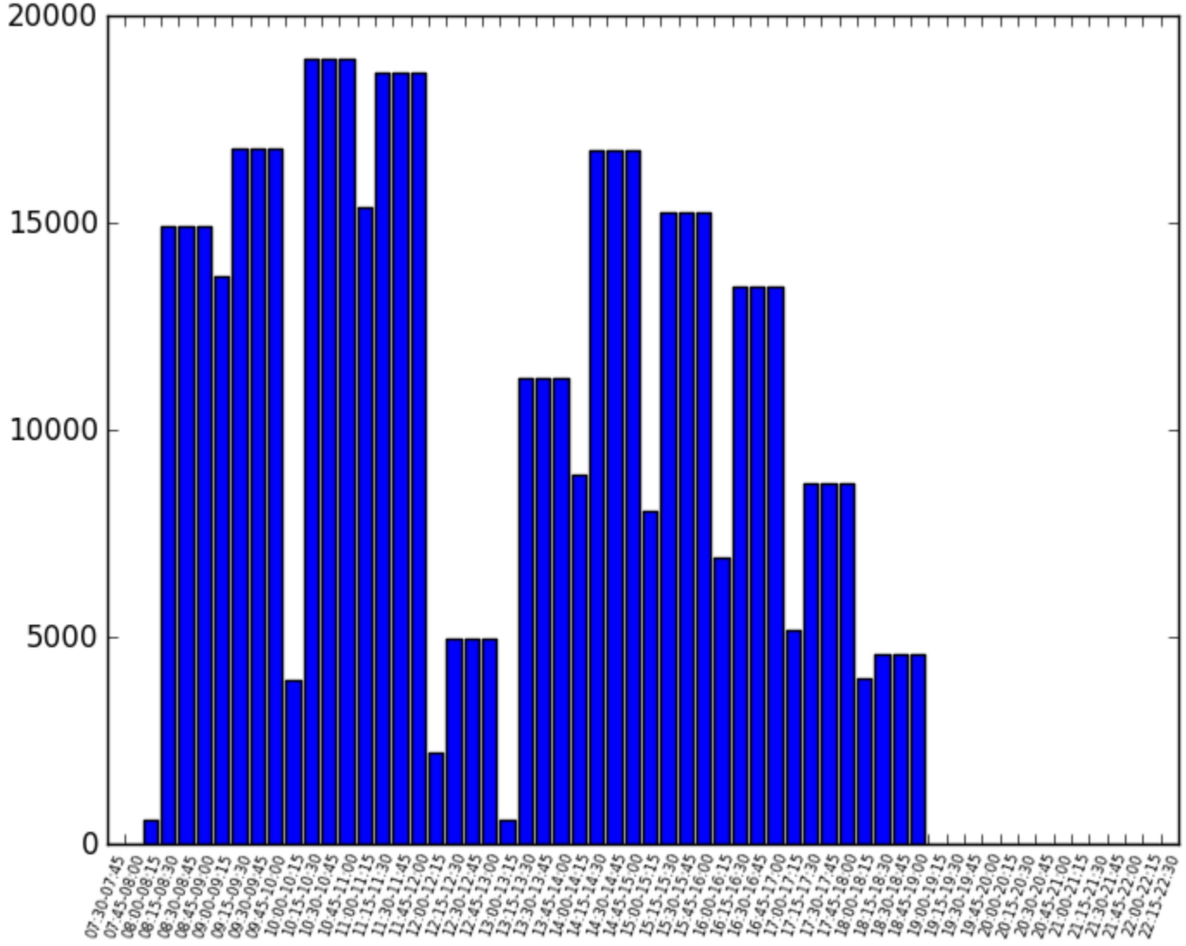
Data requirement: Potential attractivity

- **Potential attractivity measure (PAM)** depends on
  - **Destination attractivity** $att(x,t)$
    - Classroom, platform, scene, …
  - **Time-constraints** $\delta_{x,i}(t)$
    - Class schedules, train schedules, opening hours, …

$$PAM_{x,i}(t^-,t^+) = \int_{t=t^-}^{t^+} \delta_{x,i}(t) \cdot att(x,t)$$

- 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

Data requirement: Potential attractivity



Methodology

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

Input

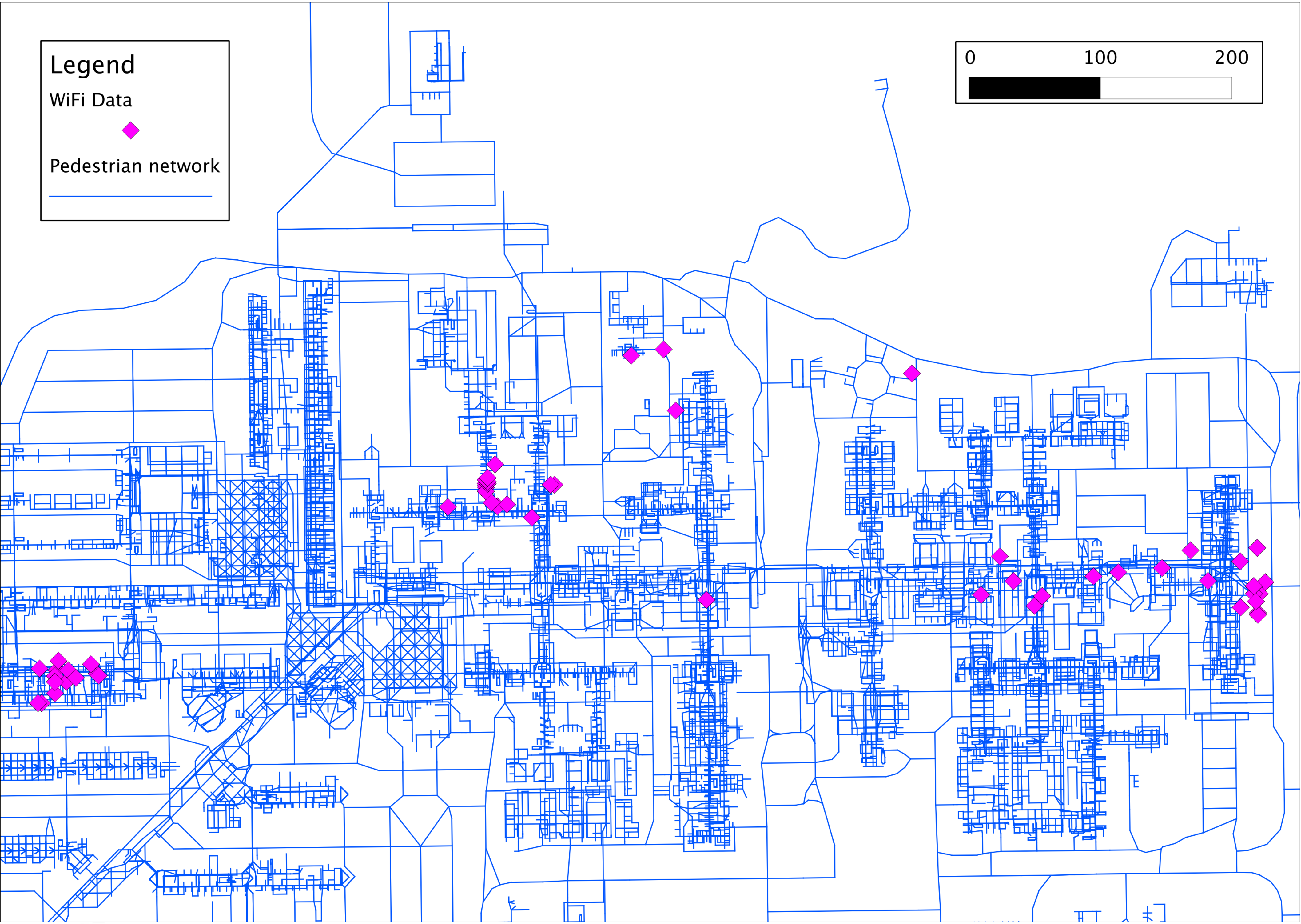
- Localization measurement
- Semantically-enriched routing graph
- PAM

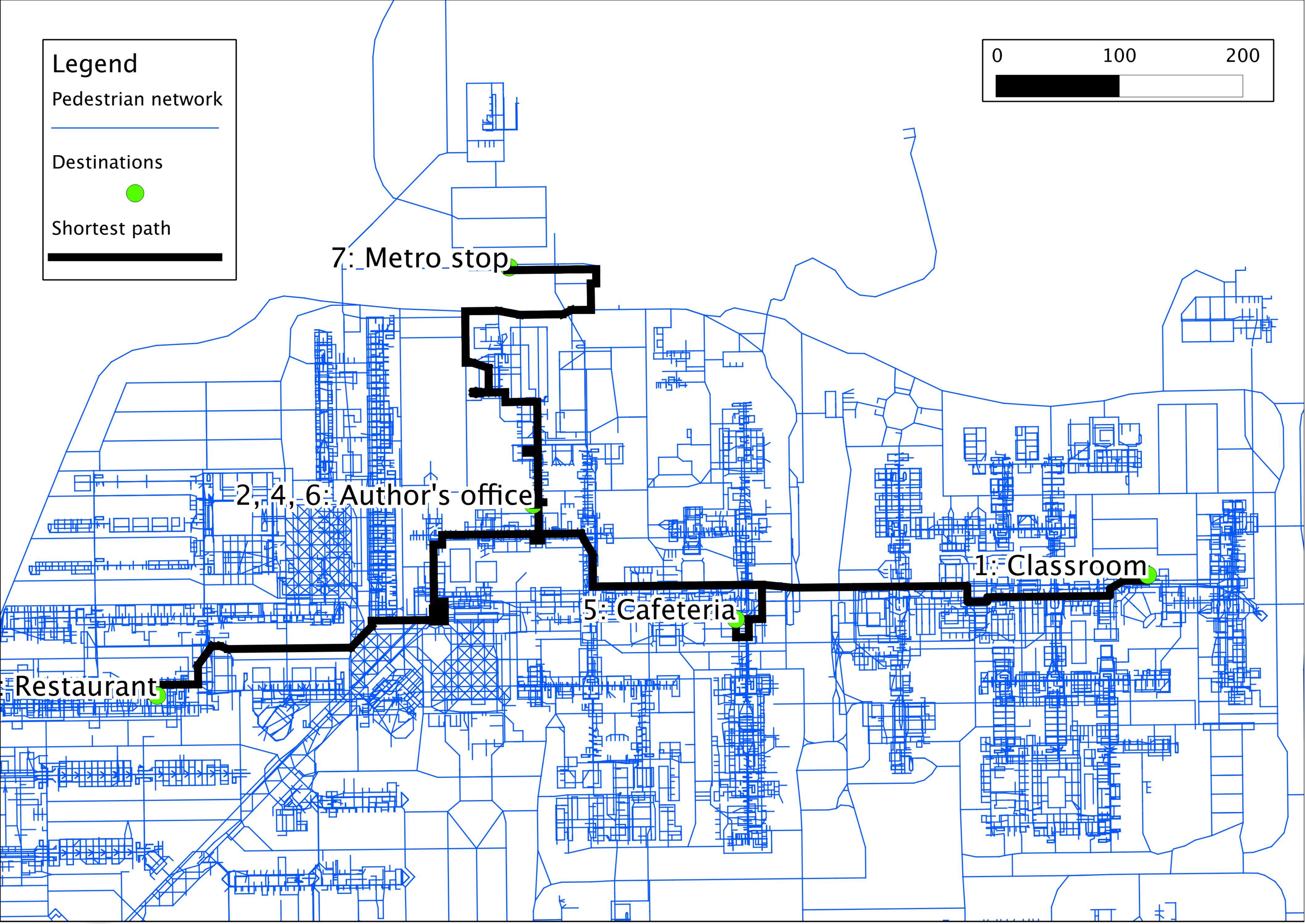
Output

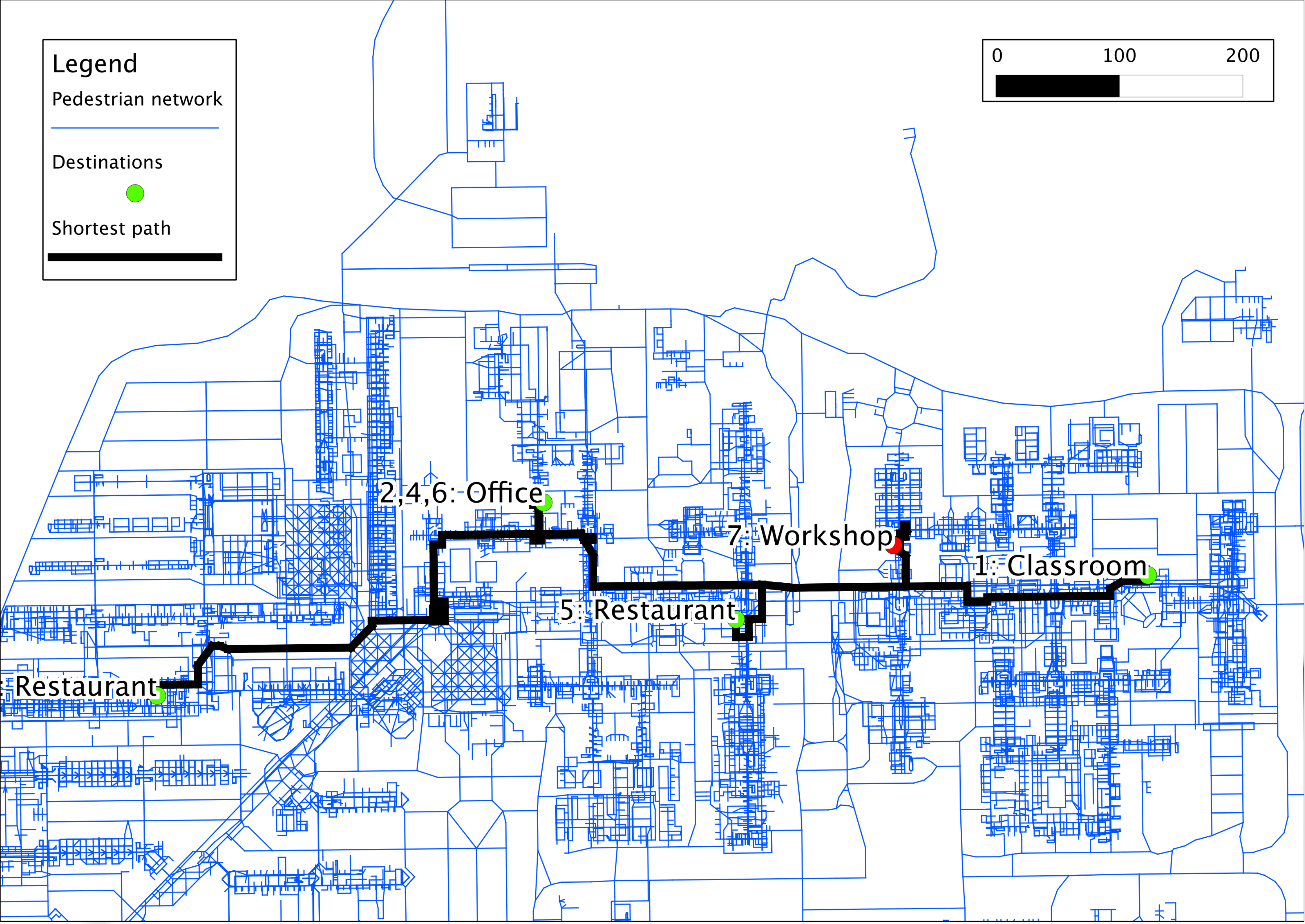
- set of candidate activity-episodes sequences associated with the likelihood to be the true one

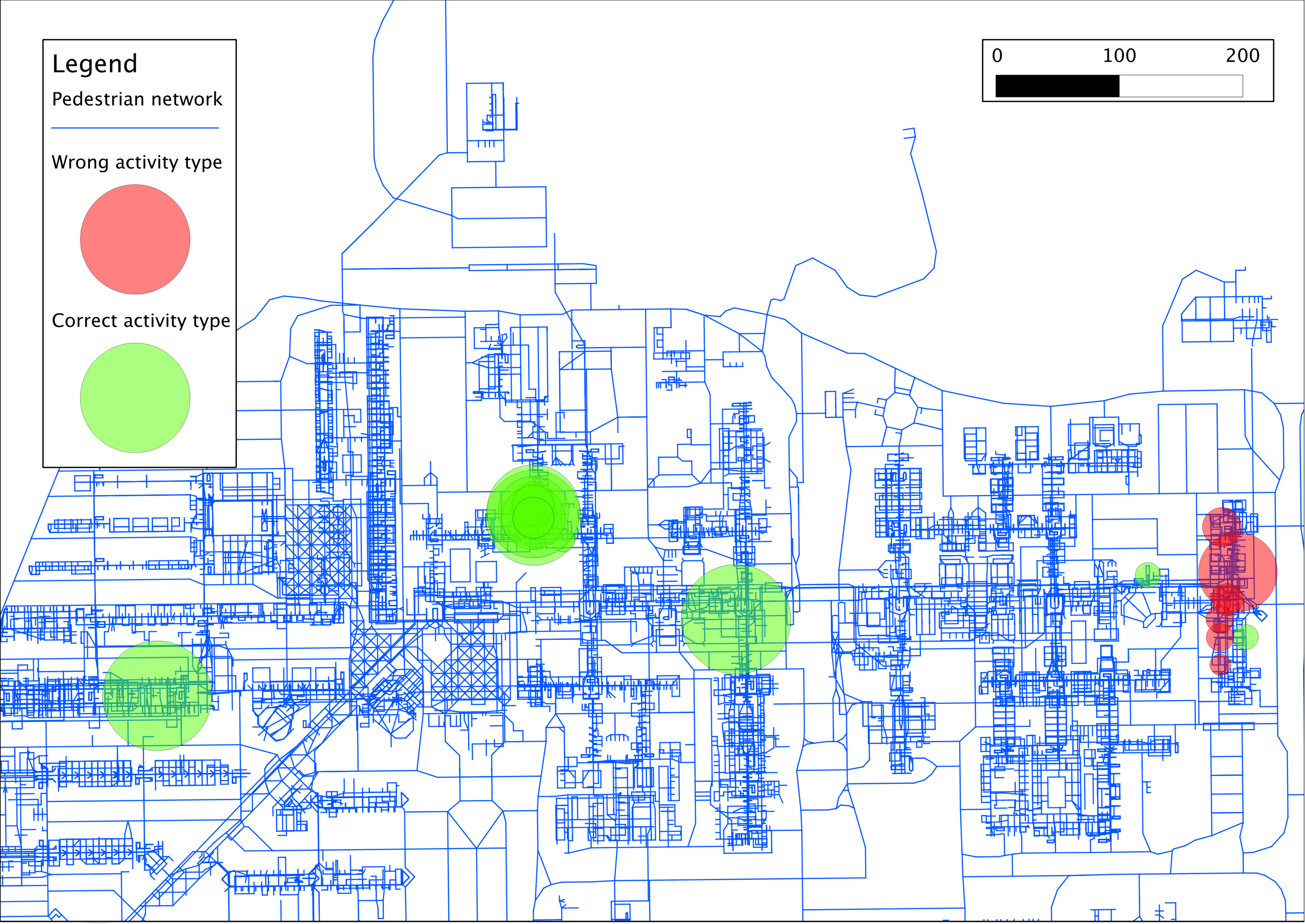
Probabilistic measurement model

$$\underbrace{P(a_{1:m}|\hat{s}_{1:n})}_{\text{Activity probability}} \propto \overbrace{P(\hat{s}_{1:n}|a_{1:m})}^{\text{Measurement likelihood}} \cdot \overbrace{P(a_{1:m})}^{\text{Prior}}$$


Legend
WiFi Data
◆
Pedestrian network
0
100
200


Legend
Pedestrian network
Destinations
Shortest path
0
100
200
7: Metro stop
2, 4, 6: Author's office
1: Classroom
5: Cafeteria
Restaurant


Legend
Pedestrian network
Destinations
Shortest path
0
100
200
2,4,6: Office
7: Workshop
1: Classroom
5: Restaurant
: Restaurant


Legend
Pedestrian network
Wrong activity type
Correct activity type
0
100
200

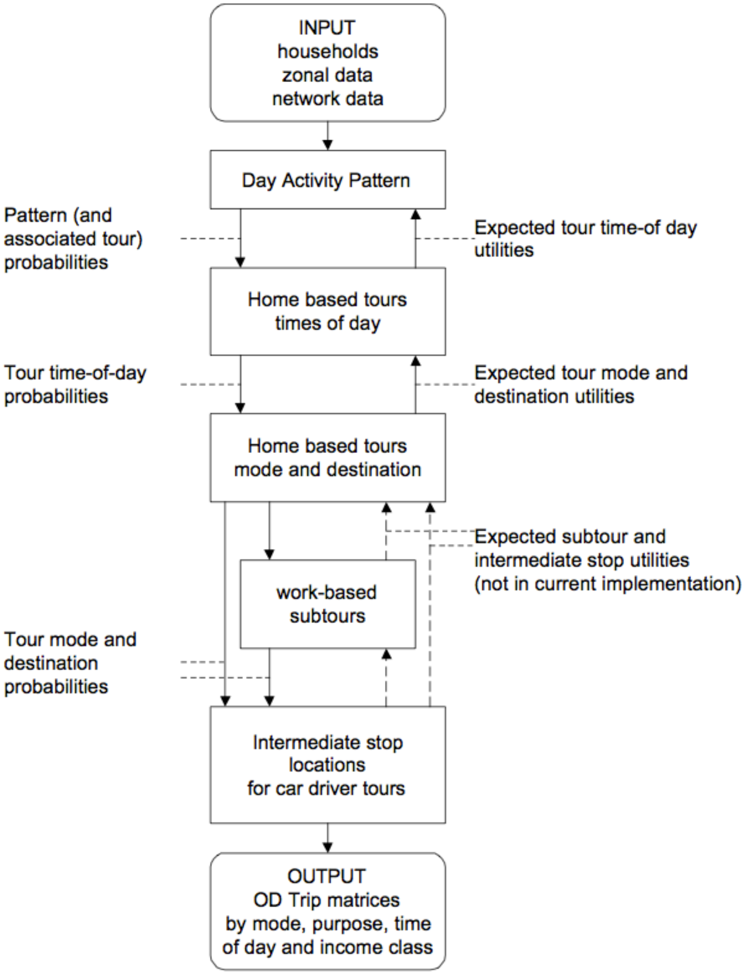
More on detection

- Technical report:
  A. Danalet, B. Farooq and M. Bierlaire. A Bayesian Approach to Detect Pedestrian Destination-Sequences from WiFi Signatures, 2013.

A PATH CHOICE APPROACH TO ACTIVITY MODELING

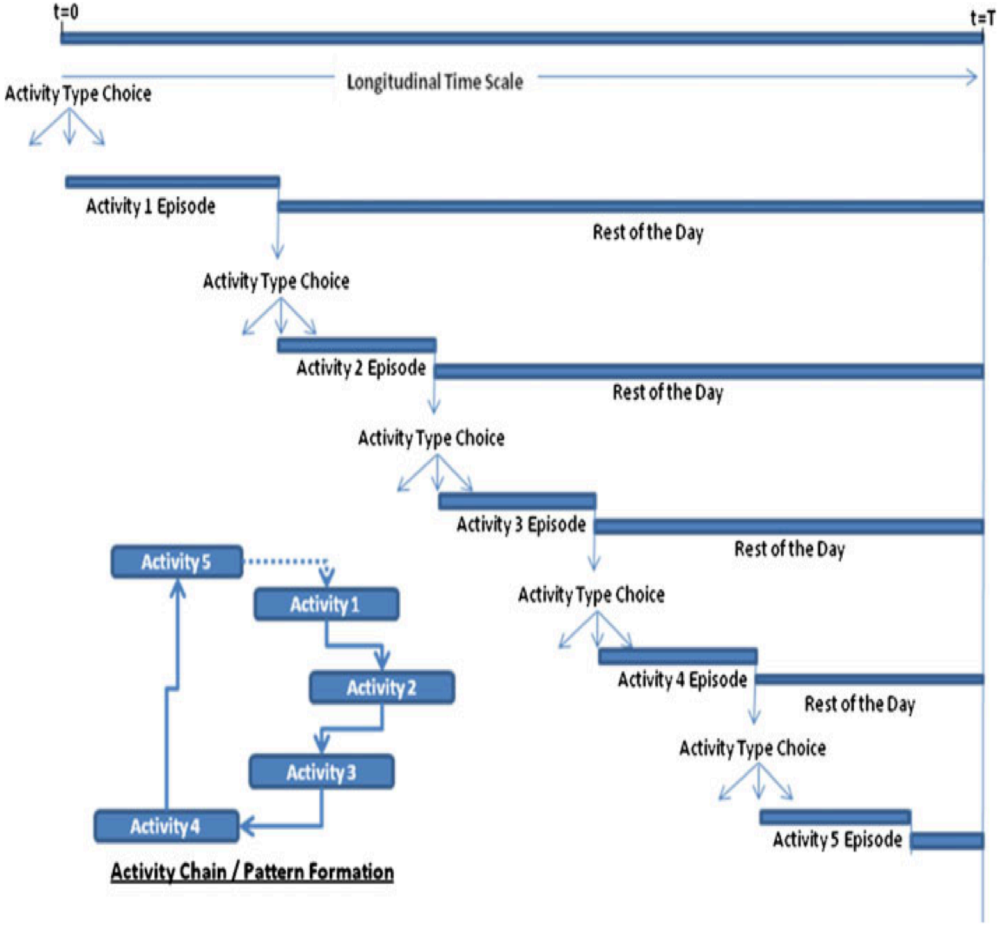
Activity travel pattern

- System of choice models (Bowman, 1998)
  - Activity pattern choice model
  - Tour choice model
    - primary destination
    - mode choice
    - time of day
    - number of stops in the tour
  - Trip choice model
    - idem, for secondary destinations



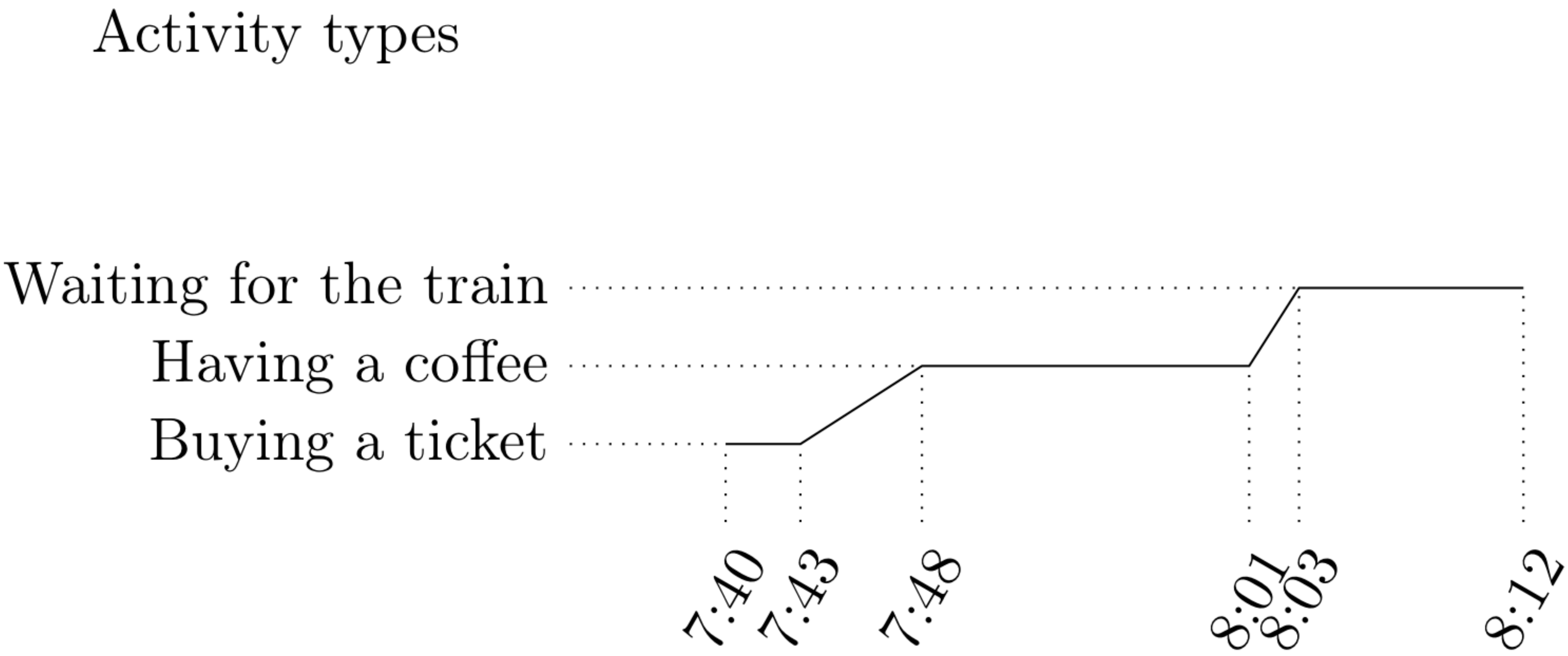
Dynamic discrete continuous models

- Continuous time
- Composite activity integrating all activities in the rest of the day
- Maximization of utility between the current specific activity and the composite future activity
- Activity pattern built sequentially



Habib, K. M. N. (2011)

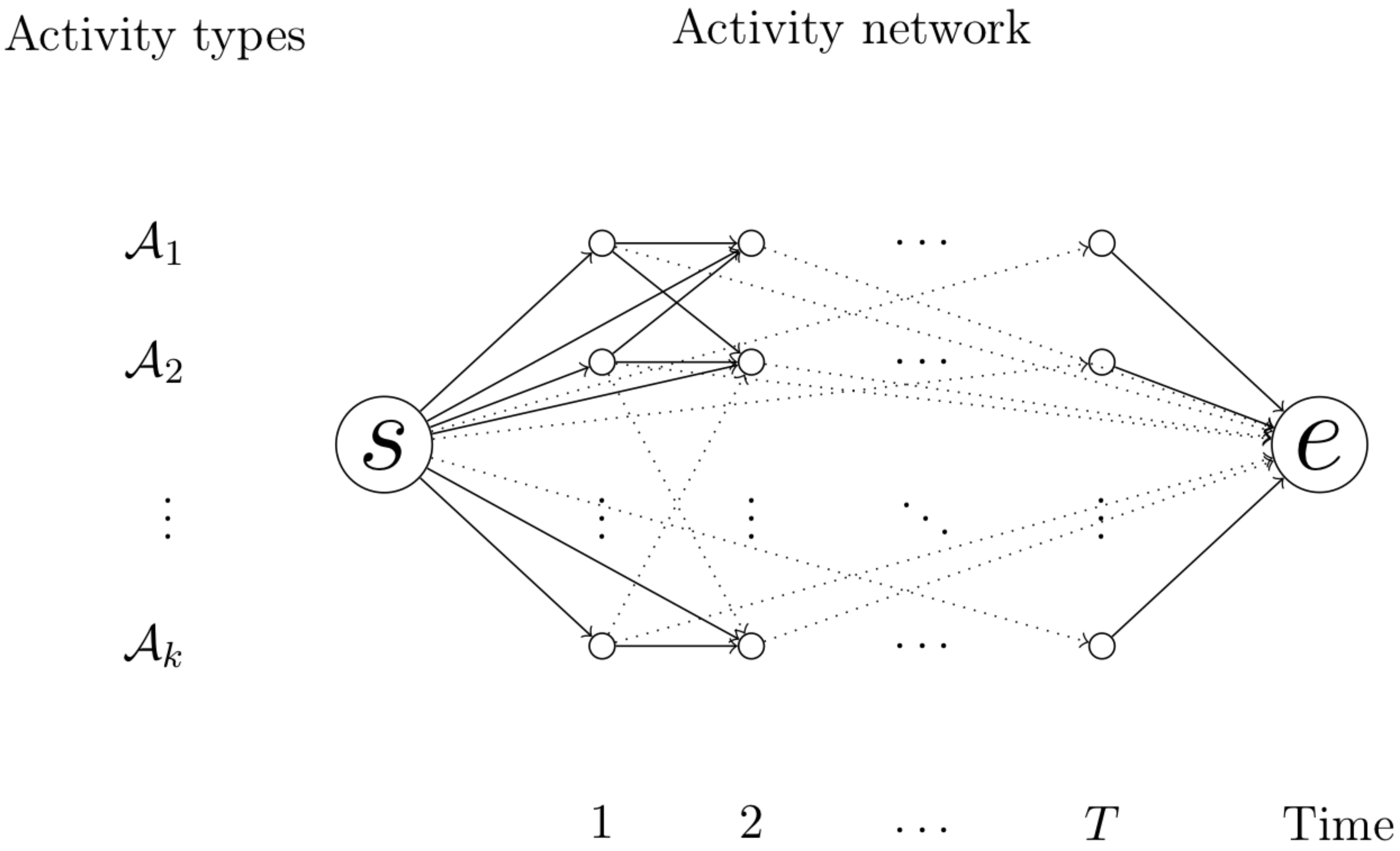
Activity-episode sequences and activity patterns



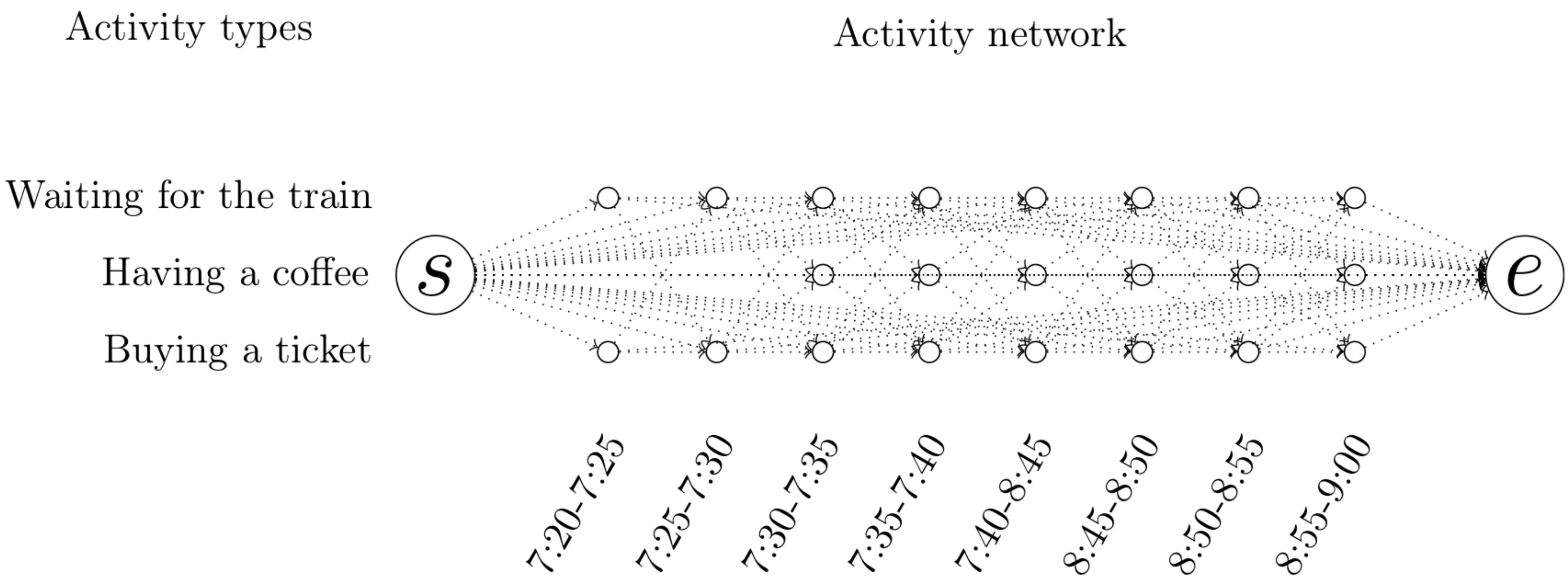
Activity-episode sequences and activity patterns

- Activity episode $a_n = (x, t^-, t^+)$
  - Start and end times are continuous random variables
  - Activity-episode sequences $(a_1, ..., a_{M_n}) = a_{1:M_n}$
- Activity types $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_K$
- Activity $\mathcal{A}_n = \left(\mathcal{A}(a_n), t^-, t^+\right)$
  - Activity pattern $(\mathcal{A}_1, ..., \mathcal{A}_{M_n}) = \mathcal{A}_{1:M_n}$
  - Set of all activity patterns corresponding to an observation $\mathcal{L}_i$
- All activity patterns are associated to a measurement likelihood
- Activity patterns are the behavior we observe

Activity network



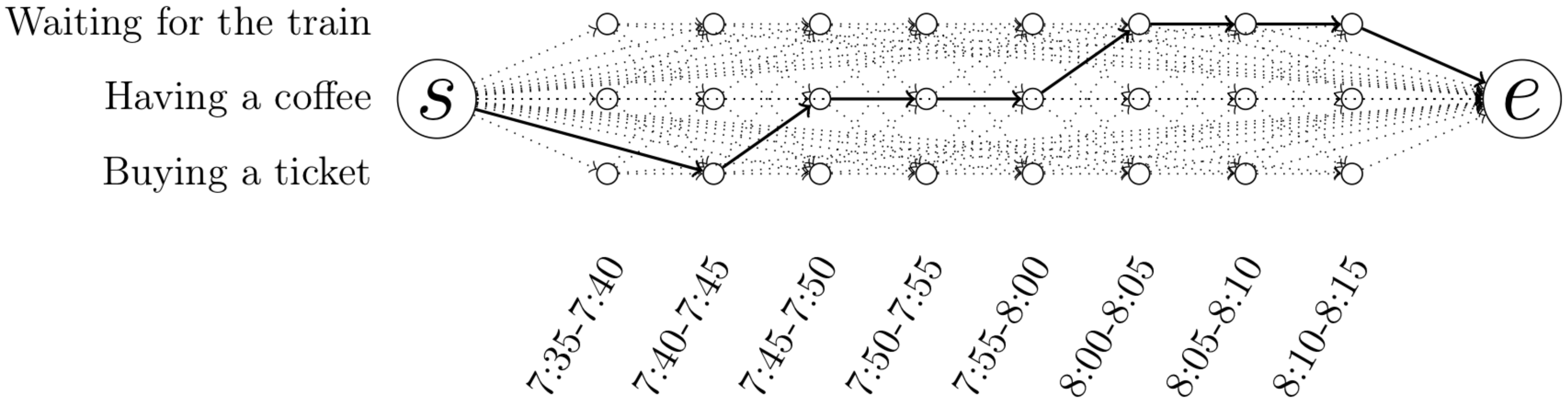
Activity network



Activity network

- Contains all possible activity patterns
  - Universal choice set
- Discretization of time $\tau \in 1, 2, ..., T$
- Nodes $\mathcal{A}_{k,\tau}$
  - represent the performance by an individual of an activity type k for a unit of time τ
  - Beginning and end of the observed activity pattern: s, e
  - Max number of nodes: $KT + 2$
- Edges
  - Max number of edges: $2KT + K^2T$

Activity paths

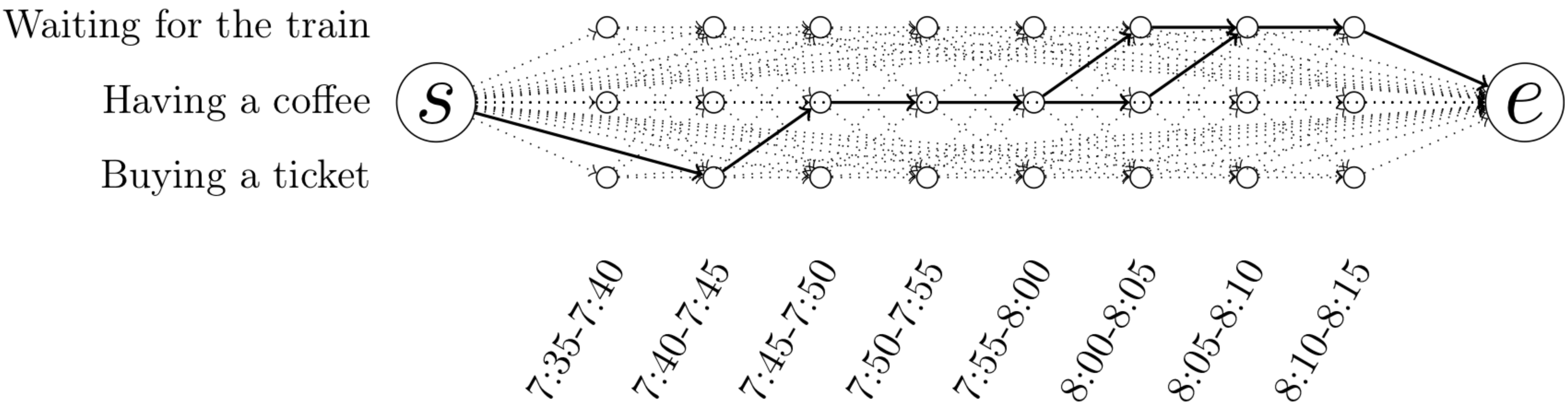


Activity paths $\mathcal{A}_{1:T}$

- Representation of $A_{1:M_n}$ in an activity network
- All activity paths are associated to a measurement likelihood
- Activity of the time unit in the activity network = longest activity in the activity pattern for this time interval

Activity paths

- Time is a random variable. If support is larger than a time unit, one activity pattern can be represented by several activity paths



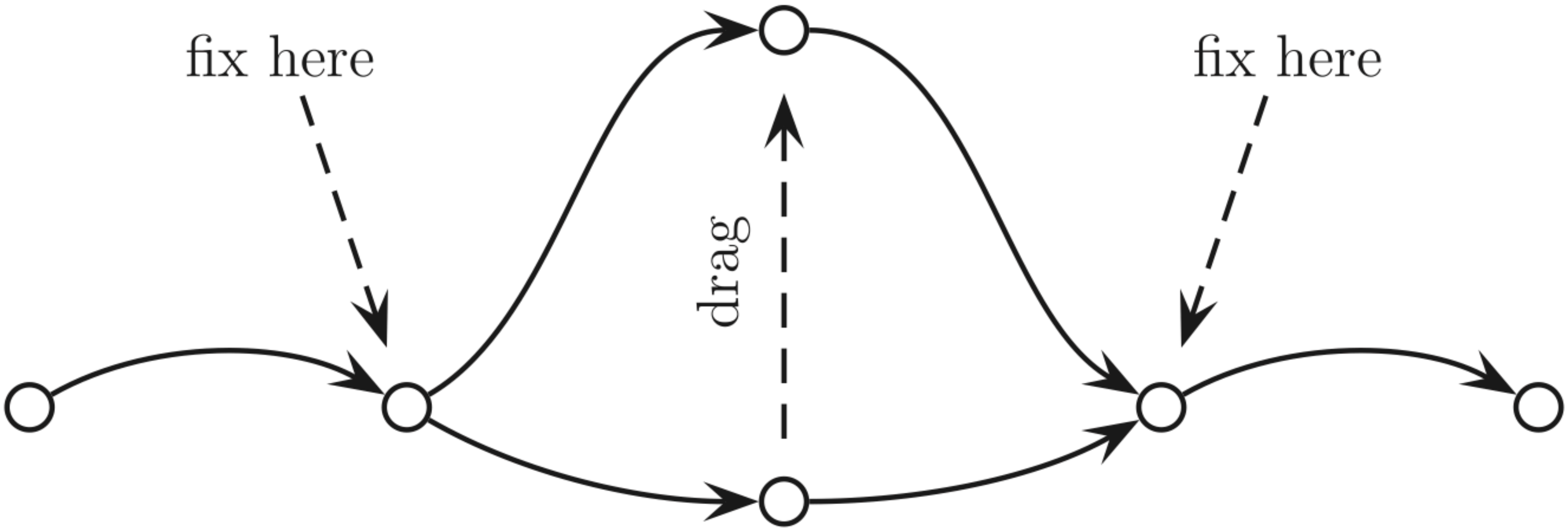
Choice set generation

- In route choice context, universal choice set is big
- Decision maker doesn't consider all of them
- Consideration choice set not available or too small
- Consideration choice set modeling
  - Latent class choice model
  - Repeated shortest path search
  - Branch-and-bound
- Sampling of alternatives from the universal choice set
  - Frejinger, Bierlaire and Ben-Akiva (2009)
  - Fosgerau, Frejinger and Karlstrom (2013)

Choice set generation: Metropolis-Hastings algorithm

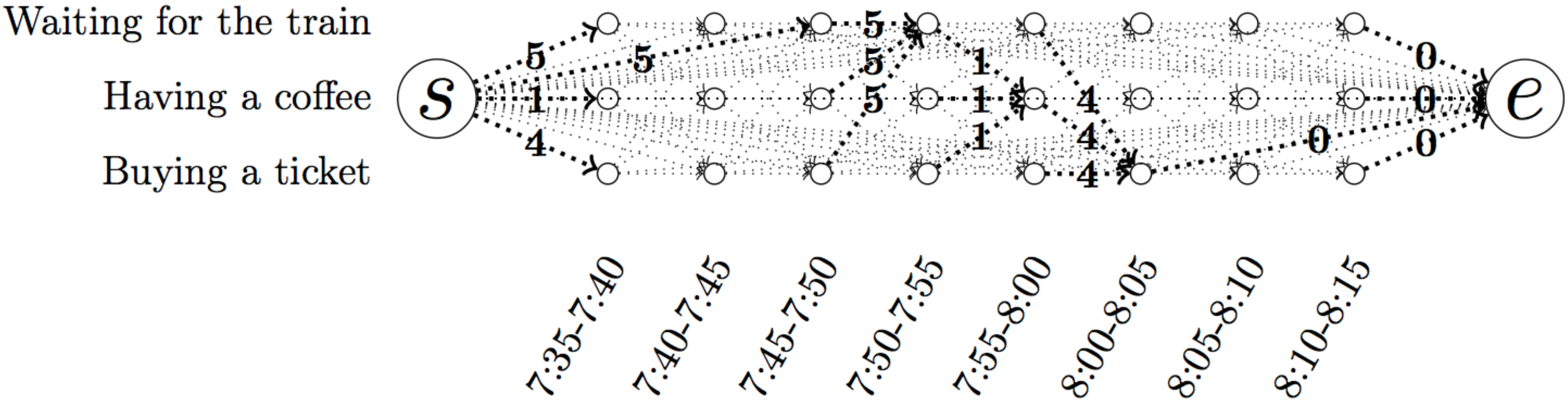
- Flötteröd and Bierlaire (2013)
- Paths are sampled according to an arbitrary distribution, avoiding complete enumeration
- Sampling probabilities do not need to be defined by link, but can be defined directly for the whole path
- Chen (2013) in Ch.5: weight function is composed of the length and frequency of observation
- Frejinger and Bierlaire (2010): « sample should include attractive alternatives »

Choice set generation: Metropolis-Hastings algorithm



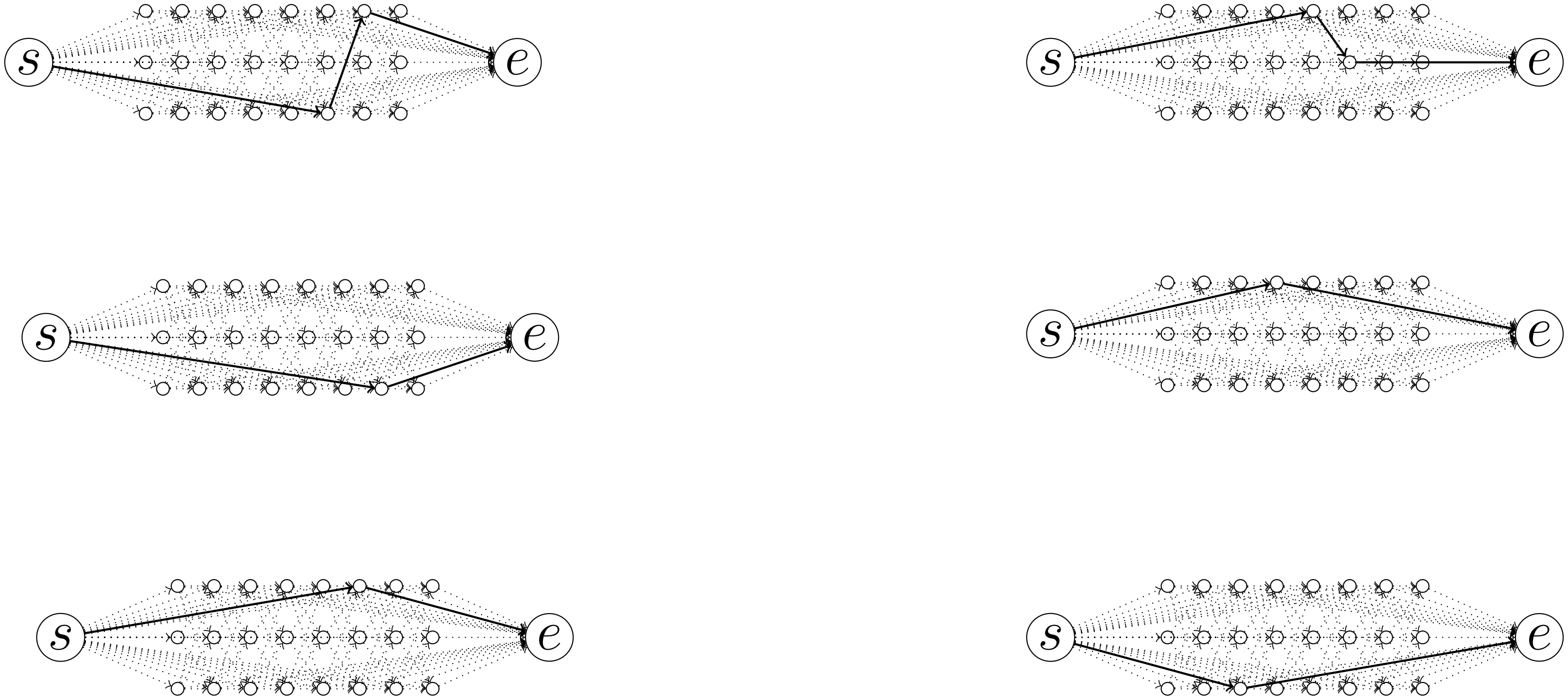
Choice set generation in the activity network

- We propose to use potential attractivity measure



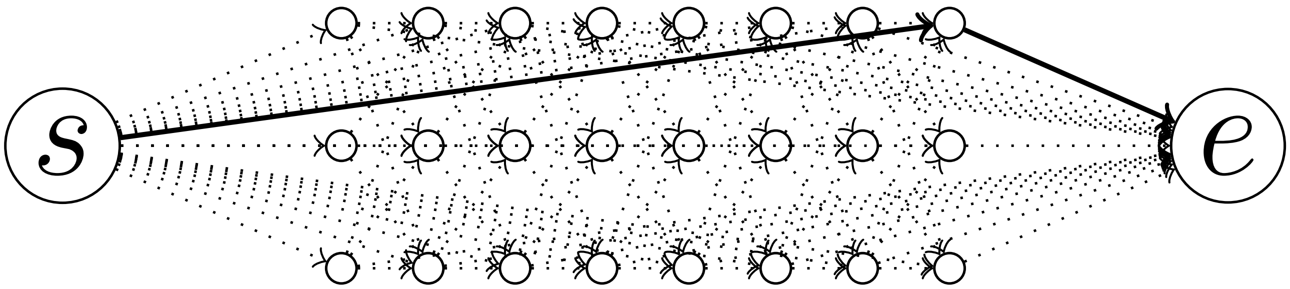
Choice set generation in the activity network

- Based only on potential attractivity measure: too short (based on shortest path)



Choice set generation in the activity network

- Force it to end on the platform: same problem
- Most likely output is



Choice set generation in the activity network

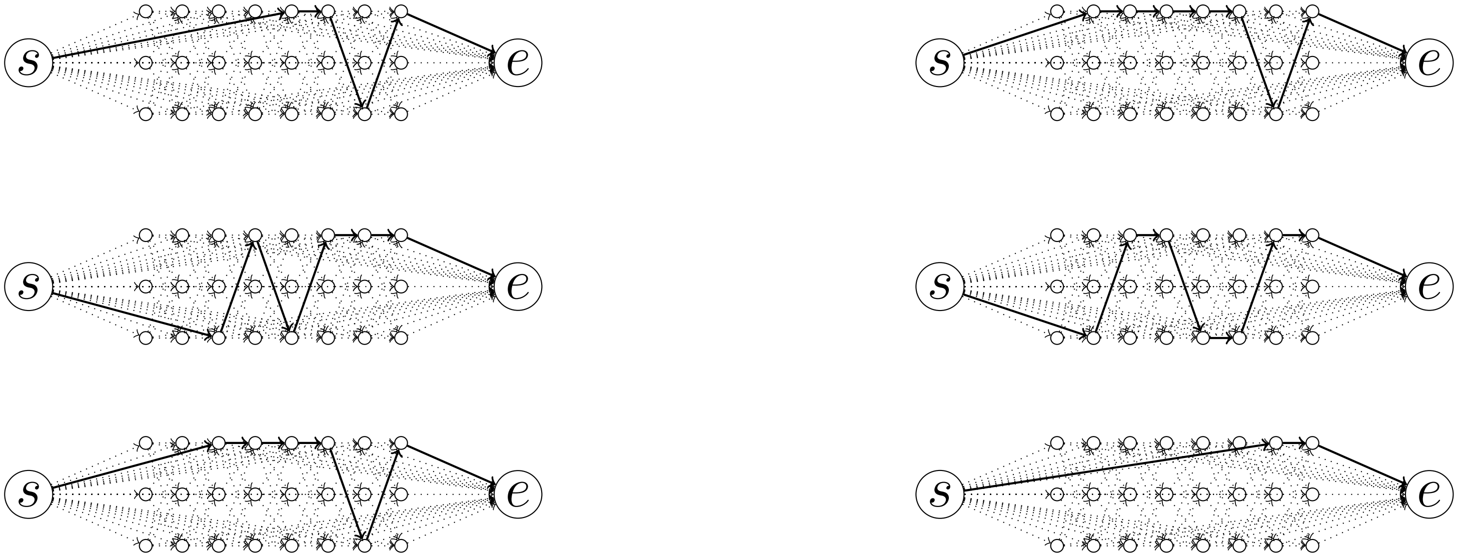
- Attractivity is link additive
- With the Metropolis-Hastings algorithm, possibility to define non-link-additive cost
- Penality for path length different from the observed ones

$$\delta_\Gamma(\Gamma) = \prod_{k=1}^{K} \left(\frac{1}{N} \sum_{\mathcal{A}_{1:T} \in N} \mathcal{I}(|\mathcal{A}_{1:T,k}| = |\Gamma_k|) \right)$$

Choice set generation in the activity network

- With target weight defined as

$$\delta(\Gamma) = -\mu_v \cdot \sum_{v \in \Gamma} \delta_v(v) - \delta_\Gamma(\Gamma)$$



Activity path choice model for WiFi traces

- Inspired by Bierlaire and Frejinger (2008) and Chen (2013): each individual n generates J network-free signals $\hat{s}_{1:J}$

$$P(\hat{s}_{1:J}) = \sum_{\mathcal{A}_{1:T} \in \mathcal{U}} \underbrace{P(\hat{s}_{1:J} | \mathcal{A}_{1:T})}_{\text{Measurement likelihood}} \cdot \underbrace{P(\mathcal{A}_{1:T} | \mathcal{U}; \beta)}_{\text{Choice model}}$$

Measurement likelihood

Choice model

Activity path choice model for WiFi traces: route choice model

- To be operationalized, the model must correct
  - for the sampling of alternatives
  - for the correlation structure of a route choice

Activity path choice model for WiFi traces: sampling of alternatives

- Frejinger et al. (2009): a sampling correction term must be added

Nb of occurrences

$$\ln q(\mathcal{C}_n|\Gamma) = \ln \frac{k_{\Gamma n}}{q(\Gamma)}$$

Sampling probability

- Sampling probability requires full enumeration

$$q(\Gamma) = \frac{b(\Gamma)}{\sum_{\Gamma' \in \mathcal{U}} b(\Gamma')}$$

but cancels out in logit

Activity path choice model for WiFi traces: sampling of alternatives

$$
\begin{aligned}
P(\Gamma|\mathcal{C}_n) &= \frac{e^{\mu V_{\Gamma n} + \ln \frac{k_{\Gamma n}}{q(\Gamma)}}}{\sum_{\Gamma' \in \mathcal{C}_n} e^{\mu V_{\Gamma n} + \ln \frac{k_{\Gamma' n}}{q(\Gamma')}}} \\
&= \frac{\sum_{\Gamma' \in \mathcal{U}} b(\Gamma') \cdot e^{\mu V_{\Gamma n}} \cdot \frac{k_{\Gamma n}}{b(\Gamma)}}{\sum_{\Gamma' \in \mathcal{U}} b(\Gamma') \cdot \sum_{\Gamma' \in \mathcal{C}_n} e^{\mu V_{\Gamma n}} \cdot \frac{k_{\Gamma' n}}{b(\Gamma')}} \\
&= \frac{e^{\mu V_{\Gamma n}} \cdot \frac{k_{\Gamma n}}{b(\Gamma)}}{\sum_{\Gamma' \in \mathcal{C}_n} e^{\mu V_{\Gamma n}} \cdot \frac{k_{\Gamma' n}}{b(\Gamma')}}
\end{aligned}
$$

Activity path choice model for WiFi traces: path size

- Ben-Akiva and Bierlaire (1999): path size logit
- Path size attribute PS_p corrects the utility for the correlation related to overlapping segments

$$PS_\Gamma = \sum_{a \in \Gamma} \frac{1}{M_a} \frac{L_a \leftarrow}{L_\Gamma \leftarrow}$$

Arcs and path length

- When using universal choice set, full enumeration

$$M_a = \sum_{\Gamma' \in \mathcal{U}} \delta_{a\Gamma'} \leftarrow$$

link-path incidence variable

- Frejinger et al. (2009): use a large set of paths

Activity path choice model for WiFi traces: activity path size

- Due to the structure of the activity network, the activity path size is:

$$APS_\Gamma = \frac{1}{K^{\tau-1}}$$

- The deterministic part of the utility function

$$V_{\Gamma n} = \beta x + \ln \frac{k_{\Gamma n}}{b(\Gamma)} + \beta_{PS} \ln APS_\Gamma$$

CONCLUSION

Conclusion

- Discretization of time: loosing information but
  - Easier to specify
  - Integration of measurement error in the model
- Modeling framework
  - allows to define choice attributes related to the whole activity pattern (e.g., number of episodes, number of times in the shopping mall or restaurant, etc.)
  - does not need an *a priori* definition of primary activity
  - does not need the definition of *home*

FUTURE WORK

Future work

- Implementation of the model on campus data
- Currently: activity = location category (e.g. "studying = classroom")
  - If the share of activities per location category does not change over time: prediction still all right
  - If they change? Stated preference survey about activities per location category?
- Once activity/location category chosen: destination choice conditional on it

$$P(a_{1:M_n}) = P(\mathcal{A}_{1:T}) \cdot P(x|\mathcal{A}_{1:T})$$

THANK YOU

References

- Ben-Akiva, M. and Bierlaire, M. (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, Kluwer, Netherlands, pp. 5–34.
- Bowman, J. L. (1998). The Day Activity Schedule Approach to Travel Demand Analysis, PhD thesis, Massachusetts Institute of Technology.
- Chen, J. (2013). Modeling route choice behavior using smartphone data, PhD thesis, Ecole Polytechnique Fédérale de Lausanne, Switzerland. http://dx.doi.org/10.5075/epfl-thesis-5649
- Flötteröd, G., and Bierlaire, M. (2013). Metropolis-Hastings sampling of paths, Transportation Research Part B: Methodological 48:53-66. http://dx.doi.org/10.1016/j.trb.2012.11.002
- Frejinger, E. and Bierlaire, M. (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, Emerald Group Publishing Limited, pp. 307–315.

References

- Frejinger, E., Bierlaire, M., and Ben-Akiva, M. (2009). Sampling of Alternatives for Route Choice Modeling, Transportation Research Part B: Methodological 43 (10):984-994.
  http://dx.doi.org/10.1016/j.trb.2009.03.001
- Habib, K. M. N. (2010). A random utility maximization (RUM) based dynamic activity scheduling model: Application in weekend activity scheduling, Transportation 38(1): 123–151.
  http://dx.doi.org/10.1007/s11116-010-9294-9