A Pedestrian Destination-Chain Choice Model from Bayesian Estimation of Pedestrian Activities using Sensors Data

Antonin Danalet
Michel Bierlaire
Bilal Farooq

hEART 2013, Stockholm
September 4-6
Presentation outline

- Motivation
- Data requirement
- Methodology
- A case study on EPFL campus
- Conclusion
- Future work
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 between Geneva and Lausanne
- 25’000 travelers/day between Geneva and Lausanne
- 50’000 travelers/day between Geneva and Lausanne
- 100’000 travelers/day between Geneva and Lausanne

* Forecast by Swiss Railways for the maximum scenario
Understand pedestrian activities

What we are doing: Campus

What we want to do: Station
Challenges

- Detect pedestrian activity-episode locations
- 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**
  - Potential attractivity measure
Data requirement: Localization
Data requirement: Pedestrian network

J. Lopez-Montenegro Ramil, Antonin Danalet (Dir.), Michel Bierlaire (Dir.), Visualization of pedestrian demand in a 3D graph, semester project, Spring semester, 2013
Data requirement: Potential attractivity

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

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

- 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

- Localization measurement
- Semantically-enriched routing graph
- PAM

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: \( (a_1, \ldots, a_m) = a_{1:m} \)
- Activity: \( A(a) \)
- Activity pattern: \( (A_1, \ldots, A_m) = A_{1:m} \)
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

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

Measurement likelihood
Prior
Activity probability
Measurement error

\[ P(\hat{s}_{1:n}|a_{1:m}) = \prod_{j=1}^{m} P(\hat{s}_{1:n}^j|a_j) \quad \leftarrow \text{Independence between activities} \]

\[ = \prod_{j=1}^{m} \prod_{i=1}^{n} P(\hat{s}_{i}^j|a_j) \quad \leftarrow \text{Independence between signals} \]

\[ = \prod_{j=1}^{m} \prod_{i=1}^{n} P(\hat{x}_{i}^j|x_j) \quad \leftarrow \text{No time measurement error} \]
\[ P(a_{1:m}) = \prod_{j=1}^{m} P(a_j) \]  \hspace{1cm} (1)

\[ = \prod_{j=1}^{m} P(x_j, t_j^-, t_j^+) \]  \hspace{1cm} (2)

\[ = \prod_{j=1}^{m} \frac{PAM_{x_j,i}(t_j^-, t_j^+)}{\sum_{x \in POI} PAM_{x,i}(t_j^-, t_j^+)} \]  \hspace{1cm} (3)
Probabilistic measurement model

\[ P(a_{1:m} \mid \hat{s}_{1:n}) \propto P(\hat{s}_{1:n} \mid a_{1:m}) \cdot P(a_{1:m}) \]

- Measurement likelihood
- Prior
- Activity model
Generation of activity-episode sequences
Generation of activity-episode sequences

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

\[ t_{i+1}^+ \sim U(\hat{t}_i + tt_{x_i,x_{i+1}}, \hat{t}_{i+1}) \]
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

\[ E(t+1) - E(t-1) < T_{min} \]
A CASE STUDY ON EPFL CAMPUS
Results

A.

Danalet

TRANSP-OR ENAC EPFL

Legend
Pedestrian network

Wrong activity type

Correct activity type
SENSITIVITY ANALYSIS
Sensitivity analysis: prior

- Uniform PAM
  - # destinations / Start and end time: OK
  - Distance / category of destination: Not OK
- Aggregate PAM creates bias
- Disaggregate / individual PAM must be used

- PAM of visited destinations should be 3x bigger than of non-visited destinations
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
  - Explicitely showing the ambiguity of the signal
- Robust for **low density data**
FUTURE WORK
Future work

- Model of choice of activity sequence as a path in a decision network
  - Discretization of time
  - Including the ambiguity in the “chosen” path
  - Sampling path from PAM as sampling probabilities (Flötteröd and Bierlaire, 2013)
- Railway station case study
  - Shorter activity episodes
THANK YOU