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

Antonin Danalet

Statistique, transport et activités, Grenoble
November 5, 2013
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
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
Mass gathering
Paléo 2012
Paléo 2013
Campus
EPFL Quartier Nord
Campus
DETECTION
Data input

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

**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

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

Measurement likelihood \quad Prior

Activity probability
More on detection

- Technical report: 
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 types

Waiting for the train
Having a coffee
Buying a ticket

7:40 7:43 7:48 8:01 8:03 8:12
Activity-episode sequences and activity patterns

- Activity episode $a_n = (x, t^-, t^+)$
  - Start and end times are continuous random variables
  - Activity-episode sequences $(a_1, \ldots, a_{M_n}) = a_{1:M_n}$
- Activity types $A_1, A_2, \ldots, A_K$
- Activity $A_n = (A(a_n), t^-, t^+)$
  - Activity pattern $(A_1, \ldots, A_{M_n}) = A_{1:M_n}$
  - Set of all activity patterns corresponding to an observation $L_i$
- All activity patterns are associated to a measurement likelihood
- Activity patterns are the behavior we observe
Activity network

Activity types

$A_1$

$A_2$

$\vdots$

$A_k$

Activity network

1  2  …  $T$  Time
Activity network

Activity types

- Waiting for the train
- Having a coffee
- Buying a ticket

Activity network diagram with time intervals 7:20-7:25, 7:25-7:30, 7:30-7:35, 7:35-7:40, 7:40-8:45, 8:45-8:50, 8:50-8:55, 8:55-9:00.
Activity network

- Contains all possible activity patterns
  - Universal choice set
- Discretization of time $\tau \in 1, 2, \ldots, T$
- Nodes $A_{k,\tau}$
  - represent the performance by an individual of an activity type $k$ for a unit of time $\tau$
  - 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

Waiting for the train
Having a coffee
Buying a ticket
Activity paths \( A_{1:T} \)

- 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.

Waiting for the train
Having a coffee
Buying a ticket
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 attractiveness 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
- Penalty for path length different from the observed ones

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

$$P(\hat{s}_{1:J}) = \sum_{A_{1:T} \in \mathcal{U}} P(\hat{s}_{1:J} | A_{1:T}) \cdot P(A_{1:T} | \mathcal{U}; \beta)$$

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 and 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

\[ \ln q(C_n | \Gamma) = \ln \frac{k_{\Gamma n}}{q(\Gamma)} \]

- 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

\[ P(\Gamma|C_n) = \frac{e^{\mu V_{\Gamma_n}} + \ln \frac{k_{\Gamma_n}}{q(\Gamma')}}{\sum_{\Gamma' \in C_n} e^{\mu V_{\Gamma_n}} + \ln \frac{k_{\Gamma'_n}}{q(\Gamma')}} \]

\[ = \frac{\sum_{\Gamma' \in U} b(\Gamma') \cdot e^{\mu V_{\Gamma_n}} \cdot \frac{k_{\Gamma_n}}{b(\Gamma')}}{\sum_{\Gamma' \in U} b(\Gamma') \cdot \sum_{\Gamma' \in C_n} e^{\mu V_{\Gamma_n}} \cdot \frac{k_{\Gamma'_n}}{b(\Gamma')}} \]

\[ = \frac{e^{\mu V_{\Gamma_n}} \cdot \frac{k_{\Gamma_n}}{b(\Gamma')}}{\sum_{\Gamma' \in C_n} e^{\mu V_{\Gamma_n}} \cdot \frac{k_{\Gamma'_n}}{b(\Gamma')}} \]
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

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

- Arcs and path length

- When using universal choice set, full enumeration

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

- 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: losing 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

\[ P(a_{1:M_n}) = P(A_{1:T}) \cdot P(x|A_{1:T}) \]
THANK YOU
References


References
