Pedestrian multi-class speed-density relationship: evaluation of integrated and sequential approach

Marija Nikolić, Michel Bierlaire, Iliya Markov, Romain Bondo

Haifa, September 12, 2017
Urbanization

- 1950: 30% of the population lives in cities
- 2014: 54% of the population lives in cities

Challenges
- Energy consumption, pollution, climate change
- Increased traffic and congestion

Source: UN World Urbanization Prospects: 2011 Revision
Pedestrian movements: Congestion

Research challenges

- Understand, describe and predict
- Optimization of current infrastructure and operations
- Efficient planning and management of future pedestrian facilities
Fundamentals

Quantities
- Density $k$ (ped/m$^2$)
- Speed $v$ (m/s)
- Flow $q$ (ped/m·s)

Daamen (2004), Duives et al. (2015)
Speed-density relationship

Models

- Fruin
- Novin & Wheeler
- Older
- Virkler & Elayadith
- Weidmann

Empirical observations

Daamen (2004), Zhang (2012)
What affects the speed of pedestrians?

Modeling framework

Objective: Accounting for pedestrian heterogeneity
Approach: Probabilistic multi-class speed-density relationship
Assumptions

- Population is partitioned into classes
- The speed of pedestrians is a random variable
- The speed-density relationship varies across classes
Model specifications

Probabilistic multi-class speed-density relationship

- **Sequential (two-stage) approach**
  - Population segmentation and movement behavior modeled sequentially

- **Integrated approach**
  - Population segmentation and movement behavior modeled simultaneously
Sequential approach

**1. stage**
- Behavioral profiles
- Trajectories
- Clustering

**2. stage**
- Traffic condition data
- CSM

**Population segmentation: clustering**
- Similarity measures: feature-based, shape-based
- Algorithm: K-means clustering

**Class-specific model (CSM)**
- Class-specific speed-density relationship: $f_j(v_i|k_i,j; \theta_j(k_i))$
Feature-based clustering

General

Context specific

Motivation
Modeling framework
Application
Two-stage approach
Integrated approach
Conclusion
References
Shape-based clustering

Dynamic time warping

Hausdorff distance
Integrated approach

Population segmentation:

- Class membership model (CMM): 
  \( \Pr(j|X_i; \beta_j) \)

- Fitness function:
  \( U_{i,j} = V_{i,j} + \varepsilon_{i,j} = CSC_j + \beta_j X_i + \varepsilon_{i,j} \)

Class-specific model (CSM)

- Class-specific speed-density relationship: 
  \( f_j(v_i|k_i, j; \theta_j(k_i)) \)

Multi-class speed-density (MC-vk)

\[
\sum_{j=1}^{J} f_j(v_i|k_i, j; \theta_j(k_i)) \Pr(j|X_i; \beta_j)
\]

Traffic condition data

Behavioral profiles

CMM

CSM

MC-vk
Application
Case study: Lausanne railway station
Data sets

Pedestrian type
- Arriving
- Departing
- Transferring
- Non-passengers

Period
- Peak
- Off-peak

Walking pattern
- Group
- Alone

Time to departure

OD distance
Two-stage approach
## Clustering: Performance analysis

<table>
<thead>
<tr>
<th>Clustering</th>
<th>R²</th>
<th>CH</th>
<th>DI</th>
<th>DB</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature-based</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pedestrian type</td>
<td>433</td>
<td>1</td>
<td>3.23</td>
<td>5.83</td>
<td>443.06</td>
</tr>
<tr>
<td>Peak - Pedestrian type</td>
<td>515</td>
<td>58.1</td>
<td>3.77</td>
<td>1</td>
<td>578.33</td>
</tr>
<tr>
<td>OD distance</td>
<td>60.8</td>
<td>101</td>
<td>8.27</td>
<td>2.94</td>
<td>173.01</td>
</tr>
<tr>
<td>Time spent</td>
<td>176</td>
<td>6.41</td>
<td>3.03</td>
<td>3.09</td>
<td>188.53</td>
</tr>
<tr>
<td>Shape-based</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DTW</td>
<td>1</td>
<td>52.2</td>
<td>42</td>
<td>69.3</td>
<td>164.5</td>
</tr>
<tr>
<td>HD</td>
<td>28.3</td>
<td>24.6</td>
<td>12.1</td>
<td>4.65</td>
<td>69.65</td>
</tr>
</tbody>
</table>

R²: R-squared, CH: Calinski-Harabasz, DI: Dunn index, DB: Davies-Bouldin

DTW: Dynamic time warping, HD: Hausdorff distance
Class profiling: OD distance

C₁: pedestrians walking longer distances
C₂: pedestrians walking shorter distances
Class profiling: HD

- $C_1$: main stream
- $C_2$: minor stream
- $C_3$: incomplete trajectories
Class-specific model

Rayleigh distribution

\[
f_j(v_i|k_i, j, \mu_j(k_i)) = \frac{v_i}{2\mu_j^2(k_i)/\pi} \exp\left(-\frac{v_i^2}{4\mu_j^2(k_i)/\pi}\right)
\]

\[
\mu_j(k_i) = v_{f,j} - \gamma_j k_i
\]

Alternative specifications

- Weibull, Log-normal, Exponential
- Lower performance (BIC)
Class-specific behavior: OD distance

![Graph showing class-specific behavior of OD distance](image)

- $C_1$: longer OD distance
- $C_2$: shorter OD distance

<table>
<thead>
<tr>
<th>Density [ped/m²]</th>
<th>Average Speed [m/s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6</td>
<td>1.4</td>
</tr>
<tr>
<td>0.8</td>
<td>1.3</td>
</tr>
<tr>
<td>1.0</td>
<td>1.2</td>
</tr>
<tr>
<td>1.2</td>
<td>1.1</td>
</tr>
<tr>
<td>1.4</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Motivation
Modeling framework
Application
Two-stage approach
Integrated approach
Conclusion
References
Class-specific behavior: HD

![Graph showing average speed vs. density for different classes]

- $C_1$: main stream
- $C_2$: minor stream
- $C_3$: incomplete trajectories
Integrated approach
MC-vk: Model specification

Two classes: class $C_1$ and class $C_2$

CMM: Logit model

$$\Pr(j|X_i; \beta_j) = \frac{e^{V_{i,j}}}{\sum_{j=1}^{2} e^{V_{i,j}}}$$

$$U_{i,j} = V_{i,j} + \varepsilon_{i,j} = CSC_j + \beta_j X_i + \varepsilon_{i,j}$$

$$V_{i,1} = CSC_1 + \beta_{DP,1} DP_i + \beta_{TP,1} TP_i + \beta_{NP,1} NP_i$$

$$V_{i,2} = \beta_{TTD,2} TTD_i + \beta_{PP,2} PP_i + \beta_{OD,2} OD_i$$

CSM: Rayleigh distribution

$$f_j(v_i|k_i, j, \mu_j(k_i)) = \frac{v_i}{2\mu_j^2(k_i)/\pi} \exp\left(-\frac{v_i^2}{4\mu_j^2(k_i)/\pi}\right)$$

$$\mu_j(k_i) = v_{f,j} - \gamma_j k_i$$
Class-specific behavior

- $C_1$: less sensitive to congestion
- $C_2$: more sensitive to congestion
Class profiling

- Motivation
- Modeling framework
- Application
- Two-stage approach
- Integrated approach
- Conclusion
- References

Class profiling diagrams showing share [%] distributions for different classes and pedestrian types, as well as average time to departure [s] and average OD distance [m].
Scenario analysis: train timetable modification

- Instrument for policy making and daily operations
- Impact of different scenarios on the movement behavior and LoS
- Augmentation by posterior analysis
Comparison of two-stage and integrated approach

**Cohesion:** how closely related observations in a cluster are

**Separation:** how distinct a cluster is from other clusters

<table>
<thead>
<tr>
<th>Approach</th>
<th>Avg cohesion</th>
<th>Avg separation</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated</td>
<td>0.242 (1)</td>
<td>0.258 (1.67)</td>
<td>2.67</td>
</tr>
<tr>
<td>Two-stage: HD</td>
<td>0.423 (1.75)</td>
<td>0.431 (1)</td>
<td>2.75</td>
</tr>
<tr>
<td>Two-stage: OD</td>
<td>0.242 (1)</td>
<td>0.243 (1.77)</td>
<td>2.77</td>
</tr>
</tbody>
</table>
Conclusion
Main findings

- Probabilistic multi-class models for pedestrian movements
  - Account for population heterogeneity
  - Integrated and two-stage approach
  - Insightful, flexible and fairly general

- Two-stage approach
  - Shape-based clustering better suited to discover behavior of interest
  - Imprecise parameter estimates due to potentially small sample sizes
  - Segmentation may introduce errors in the second stage

- Integrated approach
  - Avoids measurement errors
  - Uses the behavior of interest to define segmentation
  - Suitable for forecasting analysis
Future directions

- Additional criteria for the evaluation of the approaches
- Other feature-based and shape-based similarity measures
- Accounting for dynamics
Thank you

hEART 2017 - 6th Symposium of the European Association for Research in Transportation

Pedestrian multi-class speed-density relationship: evaluation of integrated and sequential approach

- marija.nikolic@epfl.ch
References I


