Validation of a discrete choice model of walking behavior

Th. Robin, M. Bierlai, G. Antonini and J. Cruz

TRANSP-OR Transport and Mobility Laboratory
Ecole Polytechnique Fédérale de Lausanne
transp-or.epfl.ch
Context

- Pedestrian walking behavior
- in normal conditions
- as a function of other pedestrians

*Click here for an example*
Context

Objectives:

- Specify a mathematical model to forecast the walking behavior
- Estimate the model parameters on real data
- Validate the model with real data, not involved in the estimation

Applications:

- Pedestrian simulation [Click here][With background]
- Pedestrian tracking [Click here]
  - Tracking without model [Click here]
  - Tracking with model [Click here]
Outline

• Modeling elements
• Model specification
• Estimation data
• Estimation results
• Validation
• Conclusion
Modeling elements

Pedestrian movement based on a hierarchical framework (Daamen, 2004)

- **Strategical**: list of activities
- **Tactical**: activity schedule
  - Time and location of activities
  - Choice of itinerary
- **Operational**: short-term walking behavior
  - The “next step” decision
  - Direction, speed
  - Collision avoidance
  - Leader-follower
Modeling elements

In our context

- strategical and tactical decisions are exogenous
- current intermediary destination is known (“next door”)
- we focus of a “myopic” behavior
- reactions to the immediate environment, mainly other pedestrians
Modeling elements

Given

- the current position $p_n = (x_n, y_n)$
- the current speed $v_n$ (m/sec)
- the current direction $d_n$, $d_n \in \mathbb{R}^2$, $\|d_n\| = 1$
- a visual angle $\theta_n = 170^\circ$
Modeling elements

Choice set $C_n$: individual-specific discretization
Modeling elements

- 11 directions, relative to $d_n$
- 3 speed regimes: $0.5v_n$, $v_n$, $1.5v_n$
- 33 alternatives
- Each alternative is a combination of a direction $d$ and a speed regime $v$
- Each alternative corresponds to the physical position of the next step

$$c_{vd} = p_n + vtd,$$
Modeling elements

Behavioral elements

- Pedestrian walking behavior
  - Unconstrained
    - Keep direction
    - Toward destination
    - Free flow acc/dec
  - Constrained
    - Collision avoidance
    - Leader follower
Outline

✓ Modeling elements
  • Model specification
  • Estimation data
  • Estimation results
  • Validation
  • Conclusion
Model specification

Utility:

\[ U_{in} = V_{in} + \varepsilon_{in} \]
\[ U_{vdn} = V_{vdn} + \varepsilon_{vdn} \]

1. Specification of \( V_{vdn} \) to capture the behavioral elements
2. Specification of \( \varepsilon_{vdn} \) to capture the spatial correlation
Model specification: keep direction
Model specification: keep direction

- The greater the angle, the lower the utility
- Not necessarily in a pure proportional way
- We include the following terms in the utility function

$$\beta_{\text{dir central}} \text{dir}_{dn} I_{\text{central}} + \beta_{\text{dir side}} \text{dir}_{dn} I_{\text{side}} + \beta_{\text{dir extreme}} \text{dir}_{dn} I_{\text{extreme}}$$

- Only one of the terms is non zero
- $I_k$ is 1 if the alternative belongs to zone $k$
- $\beta$ are unknown parameters to be estimated from the data
- We expect them to be negative
Model specification: keep direction

Estimated contributions to the utility

-2.364  -1.630  -1.693  -1.042-0.252  0  -0.252-1.042  -1.693  -1.630  -2.364
Model specification: toward destination

![Diagram with variables and labels]

- Destination
- $d_n$
- $d_{dist_{vdn}}$
- $c_{vdn}$
- $d_{dir_{dn}}$
- $d$

Validation of a discrete choice model of walking behavior – p.16/50
Model specification: toward destination

- The destination is exogenously given
- Two effects: the distance and the angle
- We include the following terms in the utility function

$$\beta_{ddist}ddist_{vn} + \beta_{ddir}ddir_{dn}$$

- $\beta$ are unknown parameters to be estimated from the data
- We expect them to be negative
- Results:

$$-1.55 \text{ ddist}_{vn} - 0.079 \text{ ddir}_{dn}$$
Model specification: free flow acceleration

- Constant speed is assumed to be the most comfortable
- However, pedestrians accelerate and decelerate to achieve a desired speed
- The desired speed is unknown to the analyst
- Alternatives corresponding to acceleration and deceleration are penalized
- The penalty varies with the current speed
- If the speed is already low, deceleration is less likely
- If the speed is already high, acceleration is less likely
Model specification: free flow acceleration

Penalty for alternatives corresponding to deceleration

![Graph showing the penalty contribution to the utility as a function of speed of the pedestrian.]
Model specification: free flow acceleration

Penalty for alternatives corresponding to acceleration

![Graph showing the contribution to the utility as a function of speed of the pedestrian. The graph indicates a decrease in utility as the speed of the pedestrian increases.]
Model specification: free flow acceleration

- We include the following terms in the utility function

\[
\beta_{\text{dec}} I_{v,\text{dec}} \left( \frac{v_n}{v_{\text{max}}} \right)^{\lambda_{\text{dec}}} + \\
\beta_{\text{accLS}} I_{LS} I_{v,\text{acc}} \left( \frac{v_n}{v_{\text{maxLS}}} \right)^{\lambda_{\text{accLS}}} + \\
\beta_{\text{accHS}} I_{HS} I_{v,\text{acc}} \left( \frac{v_n}{v_{\text{max}}} \right)^{\lambda_{\text{accHS}}}
\]

- Maximum one term is not zero for each alternative
- \(I_{v,\text{dec}}\) and \(I_{v,\text{dec}}\) indicates if the alternative corresponds to acceleration or deceleration
- \(I_{LS}\) and \(I_{HS}\) indicates low speed (\(\leq 1.39\)) and high speed
- \(\beta\) and \(\lambda\) are unknown parameters to be estimated from the data
- Normalization: \(v_{\text{max}} = 4.84, v_{\text{maxLS}} = 1.39\) (1.39 m/s = 5km/h)
Pedestrian walking behavior

- Unconstrained
  - Keep direction
  - Toward destination
  - Free flow acc/dec

- Constrained
  - Collision avoidance
  - Leader follower

Validation of a discrete choice model of walking behavior – p.22/50
Model specification: leader-follower

- Tendency to follow individuals going in about the same direction
- In each cone, we identify potential leaders
- Individual $k$ is a potential leader

\[
\begin{align*}
\text{if } & d_l \leq d_k \leq d_r \quad \text{(is in the cone)}, \\
\text{and } & 0 < D_k \leq D_{th} \quad \text{(not too far)}, \\
\text{and } & 0 < |\Delta \theta_k| \leq \Delta \theta_{th} \quad \text{(walking in almost the same direction)},
\end{align*}
\]

- Among them, the individual $k$ who is the closest is selected as the leader
- Her speed and direction are recorded
- Her presence may trigger a change of speed
- ...with different effects for acceleration and deceleration
Model specification: leader-follower

\[ D_{th} = 5D_{max} \]
Model specification: leader-follower
Model specification: leader-follower

- We include the following terms in the utility function

\[
I_v, acc I_{acc}^L \alpha_{acc}^L D_{acc}^L \Delta v_{acc}^L \Delta \theta_{acc}^L + I_v, dec I_{dec}^L \alpha_{dec}^L D_{dec}^L \Delta v_{dec}^L \Delta \theta_{dec}^L.
\]

- Indicators
- Sensitivity
- Stimulus
- \(\alpha^L, \rho^L, \gamma^L\) and \(\delta^L\) are unknown parameters to be estimated from the data
Model specification: collision avoidance

- Tendency to avoid individuals coming in the opposite direction
- In each cone, we identify potential “colliders”
- Individual $k$ is a potential collider

\[
\begin{align*}
\text{if } d_l & \leq d_k \leq d_r \quad \text{(is in the cone),} \\
\text{and } 0 & < D_k \leq D'_{th} \quad \text{(not too far),} \\
\text{and } \frac{\pi}{2} & \leq |\Delta \theta_k| \leq \pi \quad \text{(walking in the other direction).}
\end{align*}
\]

- Among them, the collider is identified as the individual $k$ whose walking direction is the closest to the opposite direction, that is the one with $|\Delta \theta_k|$ closest to $\pi$. 
Model specification: collision avoidance

- Potential colliders
- Collider
- \( D'k \)
- \( D''k \)
Model specification: collision avoidance

We include the following terms in the utility function

\[ I_d, d_n I_C \alpha C e^{\rho C D C} \Delta v^C C \Delta \theta^C C. \]

- Indicators
- Sensitivity
- Stimulus

- \( I_d, d_n = 1 \) if \( d \neq d_n \), otherwise, that is the term is zero for alternatives corresponding to walking straight ahead
- \( I_C = 1 \) if there is a collider in the cone, 0 otherwise.
Model specification

\[ V_{vdn} = \beta_{\text{dir}_\text{central}} \text{dir}_{dn} I_{\text{central}} + \beta_{\text{dir}_\text{side}} \text{dir}_{dn} I_{\text{side}} + \beta_{\text{dir}_\text{extreme}} \text{dir}_{dn} I_{\text{extreme}} + \beta_{\text{ddist}} \text{ddist}_{vdn} + \beta_{\text{ddir}} \text{ddir}_{dn} \]

keep direction

\[ \beta_{\text{dec}} I_{v,\text{dec}} \left( \frac{v_n}{v_{\text{max}}} \right)^{\lambda_{\text{dec}}} \]

toward destination

\[ \beta_{\text{accLS}} I_{\text{LS}} I_{v,\text{acc}} \left( \frac{v_n}{v_{\text{maxLS}}} \right)^{\lambda_{\text{accLS}}} \]

free flow acceleration

\[ \beta_{\text{accHS}} I_{\text{HS}} I_{v,\text{acc}} \left( \frac{v_n}{v_{\text{max}}} \right)^{\lambda_{\text{accHS}}} \]

leader-follower

\[ I_{v,\text{acc}} I_{\text{acc}} I_{\text{L}} \alpha_{\text{acc}}^L D_L^{\rho_{\text{acc}}} \Delta v_{\text{L}}^{\gamma_{\text{acc}}} \Delta \theta_{L}^{\delta_{\text{acc}}} \]

\[ I_{v,\text{dec}} I_{\text{dec}} I_{\text{L}} \alpha_{\text{dec}}^L D_L^{\rho_{\text{dec}}} \Delta v_{\text{L}}^{\gamma_{\text{dec}}} \Delta \theta_{L}^{\delta_{\text{dec}}} \]

\[ I_{\text{d}_n} I_{\text{C}} \alpha_{C} e^{-\rho_{\text{C}} D_C} \Delta v_{C}^{\gamma_{\text{C}}} \Delta \theta_{C}^{\delta_{\text{C}}} \]

collision avoidance
Model specification

Utility:

\[
U_{in} = V_{in} + \varepsilon_{in}
\]

\[
U_{vdn} = V_{vdn} + \varepsilon_{vdn}
\]

√ Specification of \( V_{vdn} \) to capture the behavioral elements
  • Specification of \( \varepsilon_{vdn} \) to capture the spatial correlation

The cross-nested logit model

• Nests for directions
• Nests for speed regimes
• Each alternative belongs to two nests
Model specification

- Choice model:

\[ P(i|C) = \frac{\left(\sum_{j\in C} \alpha_{jm}^{\mu_m} e^{\mu_m V_j}\right)^{1/\mu_m}}{\sum_{n=1}^{M} \left(\sum_{j\in C} \alpha_{jn}^{\mu_n} e^{\mu_n V_j}\right)^{1/\mu_n}} \]

- \( \mu \) are unknown parameters to be estimated from data
- \( \alpha \) are all fixed to 0.5 in this context.
Model specification

In summary:

- The context is described by various variables
- The variables are used to associate a utility with each cell
- The utilities are used to associate a probability with each cell
Outline

√ Modeling elements
√ Model specification
  • Estimation data
  • Estimation results
  • Validation
  • Conclusion
Estimation data

Sendaï, Japan, August 2000 (K. Teknomo)
Estimation data

- 190 pedestrian trajectories
- manually tracked, frame by frame
- 10200 positions
- Two frames per second
- Data from Arsenal Research

Frame used to compute speed and direction

Current frame

Observed choice

0.5s 1s
Example of a trajectory with some choice sets
Estimation results

- Model estimated with biogeme
- Number of estimated parameters: 24
- Signs of the parameters consistent with our expectation
Outline

√ Modeling elements
√ Model specification
√ Estimation data
√ Estimation results
  • Validation
  • Conclusion
Validation

- Two data sets
  - Japanese: used for model estimation
  - Dutch: not used for model estimation
- Cross-calibration
- Compare predicted and observed choices
Japanese data: predicted probabilities

Predicted probabilities for japanese data

Hazard = 1/33
Japanese data: predicted probabilities

Predicted

Observed
Japanese data: cross-validation

- Verify the robustness of the specification
- Re-estimate the model on 80% of the data
- Apply it on the remaining 20%
- Do the same with a simple model which exactly replicates the shares in the data
- Outliers with full model: 7.13%
- Outliers with constant-only model: 19.90%

<table>
<thead>
<tr>
<th>Model</th>
<th>Exp. 1</th>
<th>Exp. 2</th>
<th>Exp. 3</th>
<th>Exp. 4</th>
<th>Exp. 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed spec.</td>
<td>8.78%</td>
<td>6.36%</td>
<td>7.60%</td>
<td>7.87%</td>
<td>5.87%</td>
</tr>
<tr>
<td>Constant only</td>
<td>20.79%</td>
<td>20.70%</td>
<td>17.13%</td>
<td>19.88%</td>
<td>18.64%</td>
</tr>
</tbody>
</table>

Validation of a discrete choice model of walking behavior – p.43/50
Dutch data

- Collected at TU Delft, 2000-2001 (Hoogendoorn & Daamen)
- Controlled experiment with volunteer pedestrians
Dutch data

Predicted probabilities for dutch data

Hazard = 1/33
Dutch data

Comparison with a constant-only model
Dutch data

Predicted

Observed

Validation of a discrete choice model of walking behavior – p.47/50
Dutch data

<table>
<thead>
<tr>
<th>Cone</th>
<th>$\Gamma$</th>
<th>$M_\Gamma$</th>
<th>$R_\Gamma$</th>
<th>$(M_\Gamma - R_\Gamma)/R_\Gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front</td>
<td>5 – 7, 16 – 18, 27 – 29</td>
<td>43619.98</td>
<td>43374</td>
<td>0.0057</td>
</tr>
<tr>
<td>Left</td>
<td>3, 4, 14, 15, 25, 26</td>
<td>1968.79</td>
<td>2089</td>
<td>−0.0575</td>
</tr>
<tr>
<td>Right</td>
<td>8, 9, 19, 20, 30, 31</td>
<td>1764.39</td>
<td>1972</td>
<td>−0.1053</td>
</tr>
<tr>
<td>Extreme left</td>
<td>1, 2, 12, 13, 23, 24</td>
<td>45.86</td>
<td>27</td>
<td>0.6985</td>
</tr>
<tr>
<td>Extreme right</td>
<td>10, 11, 21, 22, 32, 33</td>
<td>81.97</td>
<td>19</td>
<td>3.3144</td>
</tr>
</tbody>
</table>

Predicted ($M_\Gamma$) and observed ($R_\Gamma$) shares for alternatives grouped by directions with the Dutch data set.
Predicted ($M_\Gamma$) and observed ($R_\Gamma$) shares for alternatives grouped by speed regime with the Dutch data set.

<table>
<thead>
<tr>
<th>Area</th>
<th>$\Gamma$</th>
<th>$M_\Gamma$</th>
<th>$R_\Gamma$</th>
<th>($M_\Gamma - R_\Gamma$) / $R_\Gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>acceleration</td>
<td>1 – 11</td>
<td>3892.35</td>
<td>1273</td>
<td>2.0576</td>
</tr>
<tr>
<td>constant speed</td>
<td>12 – 22</td>
<td>40733.53</td>
<td>45869</td>
<td>−0.112</td>
</tr>
<tr>
<td>deceleration</td>
<td>23 – 33</td>
<td>2855.12</td>
<td>339</td>
<td>7.4222</td>
</tr>
</tbody>
</table>
Conclusion

- Model for pedestrian walking behavior
- New methodological framework
- Discrete choice model – random utility model
- Specification of the utility to capture key behavioral aspects
- Parameters estimated on real data
- Model has been successfully validated on experimental data collected in TU Delft (The Netherlands)