Data-driven PT ridership prediction approach

including comfort aspects

dr. N. van Oort
ir. T. Brands
ir. M.P. Drost
ir. M. Yap
Policy questions

- Impact of construction works (rerouting, ridership decrease)
- Simple efficiency improvements (schedule, fares)
- Dealing with budget savings (least damage)

Supporting decision making taking into account:
- Passenger impacts
- Costs (service) and revenues (tickets)
- Societal costs/benefits (value of time)
# Available tools

<table>
<thead>
<tr>
<th></th>
<th>Multimodal model</th>
<th>Quick-Scan model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Modes</strong></td>
<td>Car, public transport, bike</td>
<td>Public transport</td>
</tr>
<tr>
<td><strong>Scale</strong></td>
<td>National, regional, urban</td>
<td>Urban</td>
</tr>
<tr>
<td><strong>Time horizon</strong></td>
<td>10-20 years</td>
<td>&lt; 5 years</td>
</tr>
<tr>
<td><strong>Project type</strong></td>
<td>Strategic, policies, infrastructure changes</td>
<td>Tactical, changing lines, frequencies</td>
</tr>
<tr>
<td><strong>Usage</strong></td>
<td>Modal split, cost-benefit analysis</td>
<td>Route choice effects</td>
</tr>
</tbody>
</table>
New generation of models

<table>
<thead>
<tr>
<th>Traditional (4-step) model</th>
<th>Simple calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multimodal (~PT)</td>
<td>PT only</td>
</tr>
<tr>
<td>Network</td>
<td>Line</td>
</tr>
<tr>
<td>Complex</td>
<td>Transparent</td>
</tr>
<tr>
<td>Long calculation time</td>
<td>Short calculation time</td>
</tr>
<tr>
<td>Visualisation</td>
<td>Only numbers</td>
</tr>
<tr>
<td>Much data</td>
<td>Little data</td>
</tr>
<tr>
<td>Detailed results</td>
<td>Assessments</td>
</tr>
</tbody>
</table>

Short term predictions

Elasticity method based on smartcard data
Smartcard data

Our research focus:

Connecting to transport model
- Evaluating history
- Predicting the future

What if scenario’s
- Stops: removing
- Shorter travel times and higher frequencies
- Route changes

Quick insights into
- Expected Ridership
- Expected costs (coverage)
Smartcard data

The Netherlands
- OV Chipkaart
- Nationwide (since 2012)
- All modes: train, metro, tram, bus
- Tap in and tap out
- Bus and tram: devices are in the vehicle

Issues
- Privacy
- Data accessibility via operators

Data
- 19 million smartcards; 42 million transactions every week
Connecting data to transport model

1) Importing PT networks (GTFS) (Open data)

2) Importing smartcard data (Closed data)

3) Processing, cleaning and matching

4) Route choice and visualization options of transport model
What if?
What if: elasticity approach

Elasticities

Literature (e.g. Balcombe)

“Proven “ rules of thumb

NOTE:
Simple changes
Short term
Primarily LOS changes
Accuracy

$$C_{ij} = \alpha_1 T_{ij} + \alpha_2 WT_{ij} + \alpha_3 NT_{ij} + \alpha_4 F_{ij}$$

With:
- $C_{ij}$: Generalised costs on OD pair $i,j$
- $\alpha_1, \alpha_2, \alpha_3, \alpha_4$: Weight coefficients
- $T_{ij}$: In-vehicle travel time on OD pair $i,j$
- $WT_{ij}$: Waiting time on OD pair $i,j$
- $NT_{ij}$: Number of transfers on OD pair $i,j$
- $F_{ij}$: Fare to be paid by the traveler on OD pair $i,j$
Crowding in PT

- Perception of in-vehicle time of travellers: a crowded vehicle is less attractive
  - Travel time may remain the same
- Dwell time may increase in a crowded vehicle
  - Boarding and alighting of passengers takes more time
- Very crowded vehicles result in denied boarding
  - Additional waiting time of one (or more) entire headway
Crowding model

- (perceived) in-vehicle time depends on crowding level
- Iterative assignment is needed
- Two values indicate capacity:
  - Number of seats
  - Crush capacity: maximum capacity of vehicle: sitting and standing passengers together

\[
VC = \left\{ \begin{array}{l}
\frac{L}{C_{\text{seats}}} \\
\frac{L-C_{\text{seats}}}{1+C_{\text{crush}}-C_{\text{seats}}}
\end{array} \right.
\]

- Distinguish between vehicles with relatively large / small number of seats
Crowding model

\[ T_{ij}^{per} = T_{ij} \times F \]

\[ C_{ij} = \alpha_1 T_{ij}^{per} + \alpha_2 W_{ij} + \alpha_3 N_{ij} + \alpha_4 F_{ij} \]

Douglas Economics (2006)
MVA Consultancy (2008)
Wardman and Whelan (2011)
Crowding model

- Modelled time period
  - Usually an entire peak period of 2 hours is modelled
  - Some vehicles may be busier than other vehicles
  - Evenness of the load distribution over this period
  - → a correction factor may be applied that is lower than 2, to incorporate this effect
Case study: The Hague tram network
Network changes

- Increase frequency of tram line 15 from 6 to 8 times per hour during morning peak and evening peak

- Results:

<table>
<thead>
<tr>
<th></th>
<th>Model without comfort</th>
<th>Model including comfort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average work day</td>
<td>+8%</td>
<td>+10%</td>
</tr>
</tbody>
</table>
Result frequency increase line 15
Network changes

- Transformation of line 25 from bus line to tram line
- Due to larger vehicles, the frequency decreases
Conclusions

- Smartcard data supports ridership predictions
- Combining strengths of both worlds
- Comfort is explicitly taken into account
- Limited computation time needed for real size networks

- Benefits (revenues and societal) of certain measures become larger when comfort is taken into account
- Up to 30% underestimation of the effects when comfort is neglected
- Reduce crowding may compensate frequency reduction
Work in progress

- Validating the model using revealed data (smart card data), including behaviour during disturbances
- Incorporating denied boarding and extended dwell times
- Adding service unreliability costs
- Applying the quantified comfort effect in cost-benefit analysis
Questions

**Related papers:**
[http://nielsvanoort.weblog.tudelft.nl/](http://nielsvanoort.weblog.tudelft.nl/)

- Niels van Oort
  - NvOort@goudappel.nl / N.vanOort@TUDelft.nl

- Ties Brands: TBrands@goudappel.nl
  [http://www.utwente.nl/ctw/vvr/People/brands/](http://www.utwente.nl/ctw/vvr/People/brands/)

- Marc Drost: M.Drost@htm.net

- Menno Yap: MYap@goudappel.nl