Enriching the transport model of the Rotterdam region by cell phone data

MT-ITS Budapest, June 3 2015

Klaas Friso
New data technology. New capabilities

- Using traditional methods (like Household Travel Surveys) can be expensive and time consuming
- No adequately differentiation between residents and visitors (if they measure visitors at all)
- New passive technology, such as collection of anonymous signals from mobile devices, provides new options to planners
Benefits of passive data collection

- **Lower costs**: in comparison with household surveys and license plate surveys
- **More comprehensive**: Sample size can be much larger and time periods longer; data is more timely and relevant than other methodologies
- **More flexible**: Data can be collected 24/7/365 or for any time period of a day
- **Data linking**: population movements linking to behaviour, e.g. household income, age, # vehicles per household, etc.
- **Less data bias**: almost every adult owns a mobile device → all socioeconomic groups are proportionally represented
Our vision

• Travel patterns in the base model will be measured (instead of estimated) using Big Data (like mobile phone data)
• Transport models to determine the effect of change (delta’s) and forecasting

• ‘Data-driven’ models:
  – Quality boost at the core: the OD-matrix
  – Other periods than standard are easily possible
  – Day to day variation
  – Events
  – Real-time modelling
  – …
First case in Abijan (Ivory Coast) (D4D Challenge)

- From mobile phone activity …to…linear movements … to …traffic flows
The Dutch situation

• The case of Ivory Coast and Senegal showed it is possible: building OD-matrices from mobile phone data
• But…what about the quality? Ivory Coast did not have a test framework
• In Holland we have many transport models and empirical data

• We used mobile phone data for the region of Rotterdam
  – **Short term: testing the distribution**
  – **Long term: first step to data-driven modelling**
Mobile phone data & definitions

Data
- Data delivered Mezuro: 1/3 share of the market
- Data set of November 2014: > 100 million of trips

Definitions
- Framing: if mobile device more than 30 minutes at a certain location → destination
- Privacy issue: all OD-relations < 15 are skipped
- Trips with origin and destination in the Netherlands
Trip definition explanation

- **OD-data**: Deventer-Utrecht, Utrecht-Amsterdam, Amsterdam-Deventer (and not any relation with Amersfoort)
First analysis mobile phone data

4 Day Marches Nijmegen
Mobile phone data in transport model environment
Specific relations clearly visible in the data

Almere – Amsterdam

Zoetermeer – Den Haag
Procedure enrichment model Rotterdam region

- Adding distribution of GSM data in transport model
- Trip ends and Modal Split conform unchanged
- 24h period: scaling over all modes, purposes and periods

Result:
- Spatially corrected a priori OD-matrix
Comparison with household travel data

- Short trips (< 8 km) and long trips (> 60 km) underestimated in GSM data → skipped in procedure
- Explanation:
  - Short distance trips: Less phone usage / left at home / 30 minutes definition
  - Long distance trips: privacy rule (< 15 trips not in data set)
Assignment comparison (enriched model vs. a priori model)
Selected area comparison: Hellevoetsluis
Conclusions: Where are we?

- For the first time a Dutch transport model is enriched with mobile phone data
- Results are plausible and show perspective for further optimisation

- Improvement of quality (according to the region of Rotterdam:
  - Distribution of specific OD-relations
  - Through traffic in Rotterdam decreased (according to number plate recognition research)
  - Better fit with counts (+5%) $\rightarrow$ less calibration effect needed $\rightarrow$ improvement of forecasts

- Big Data in transport models: we are just at the beginning, but a lot can already be done!!
Further research

- Usage of information of modes (train vs. non-train)
  - Train is part of the Public Transport OD-matrix. How to deal with this?

- Usage of information of time periods (e.g. peak periods)

- Traditional model describes average working day. Mobile phone data shows information of day to day variation:
  - Model per day of the week?
  - Sunny (beach) day model?
  - Modelling of special events
Thanks for your attention

Köszönöm a figyelmet

Any questions or remarks?

Bármilyen kérdése vagy észrevétele?