Francesco Viti

University of Luxembourg

Multilingual. Personalised. Connected.

# Dynamic Demand Modelling in the era of Big Data

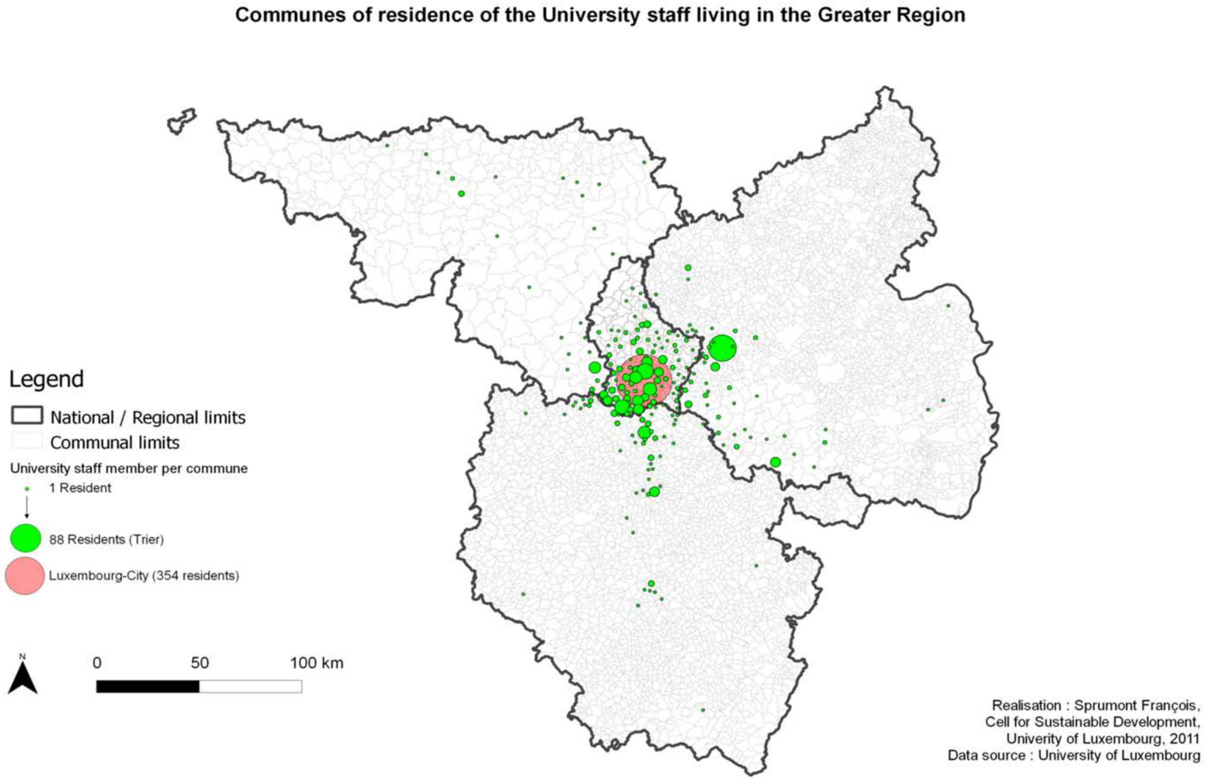
UNIVERSITÉ DU LUXEMBOURG

# Outline

- Introduction
- The demand estimation problem
- Big Data in traffic applications
- Research directions at MobiLab
  - Generation/distribution
  - Route choice and travel times
  - Activity-travel behavior
- Closure

# Mobility in Luxembourg

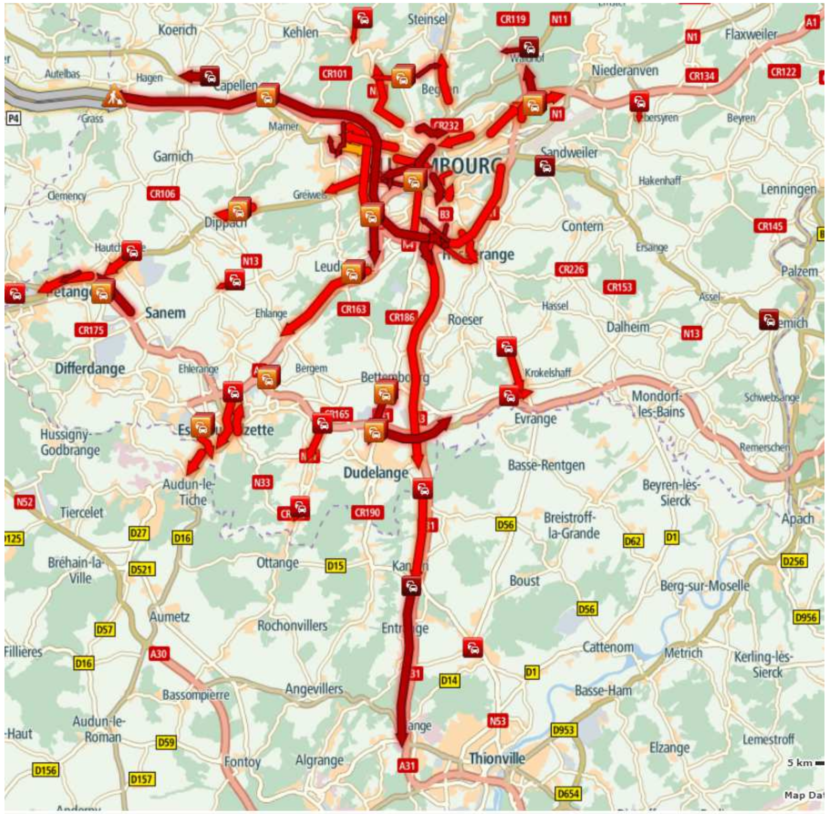
- Luxembourg strong monocentric country, financial and EU-institutional capital
- 76% car users (89% from outside); #1 car ownership rate in EU
- High car-dependency, heavy through traffic flows, truck tourism,…
- 360 000 daily commuters; 180 000 cross-border workers;

# Traffic congestion in Luxembourg

- Frequent regional gridlocks, queues spilling back beyond the borders

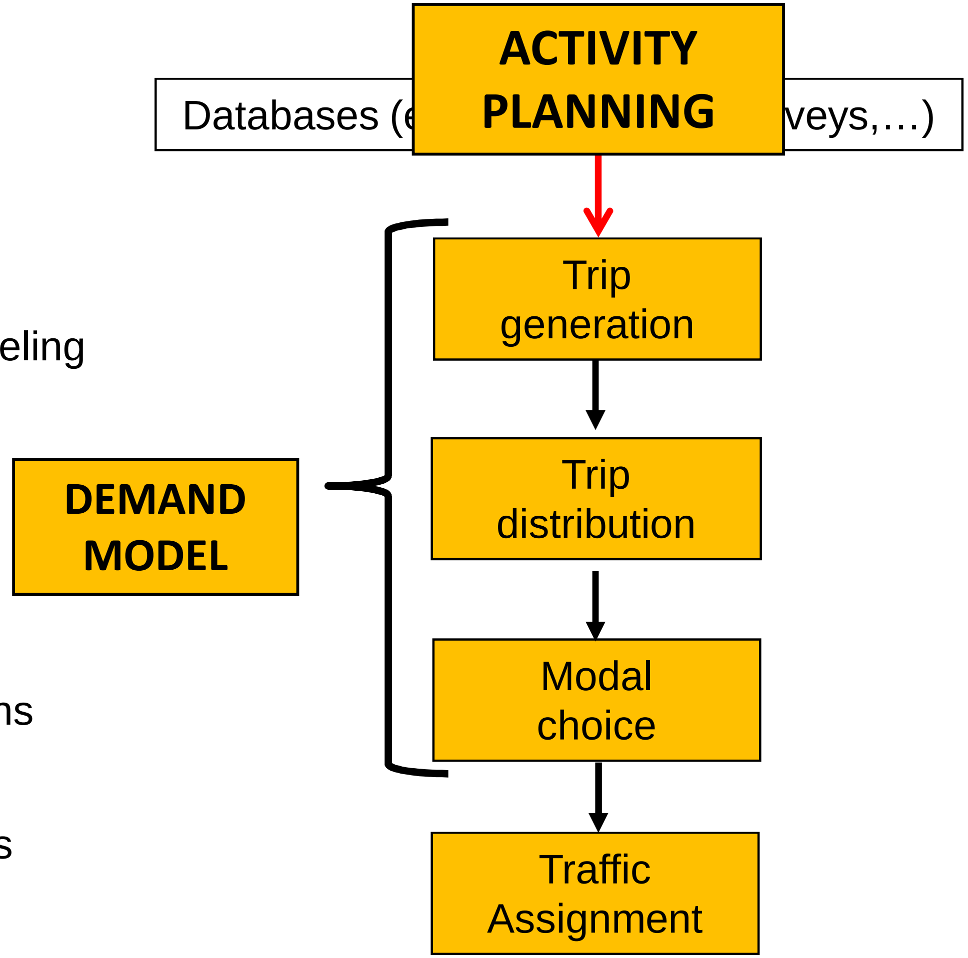
- Need for dynamic traffic models to predict and manage daily complex congestion patterns

# Dynamic modeling approaches

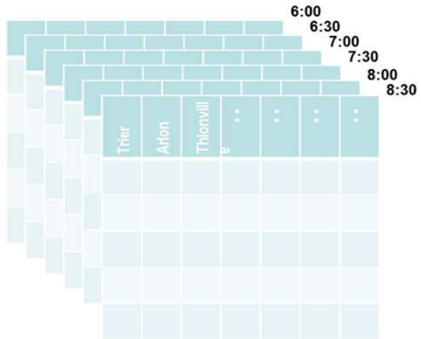
- General lack of a good dynamic demand model

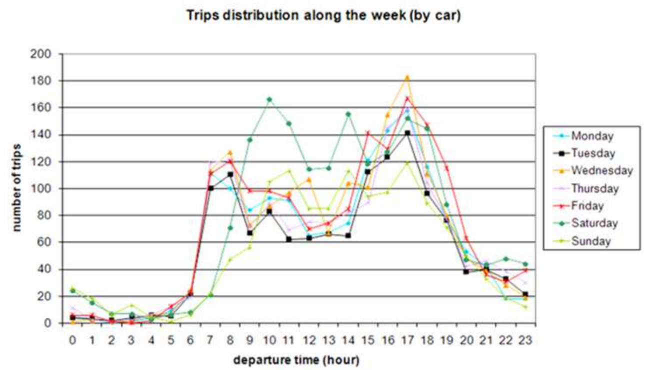
- The 'traditional' 4-stage model
  - Socio-demographic data
  - Travel surveys
  - Trip-based
  - Generally not suited for dynamic demand modeling



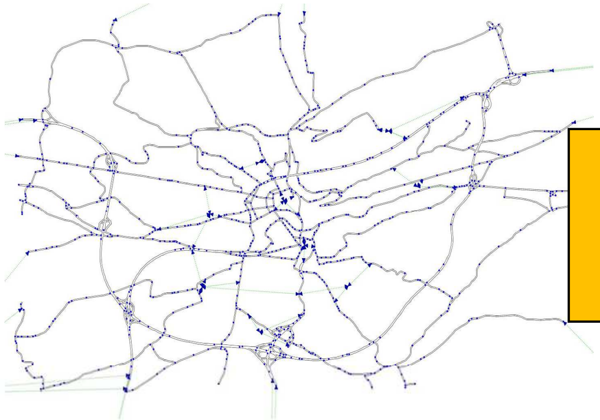

- Activity-based /agent-based models
  - Dynamic schedule-based
  - Able to capture complex daily activity-trip chains
  - Hard to calibrate and to get suitable data
  - Difficult to use to get aggregated demand flows
  - Currently not much used for traffic assignment

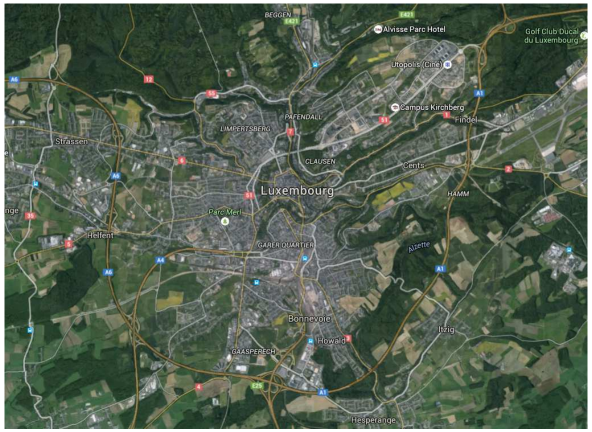
# The estimation/calibration problem

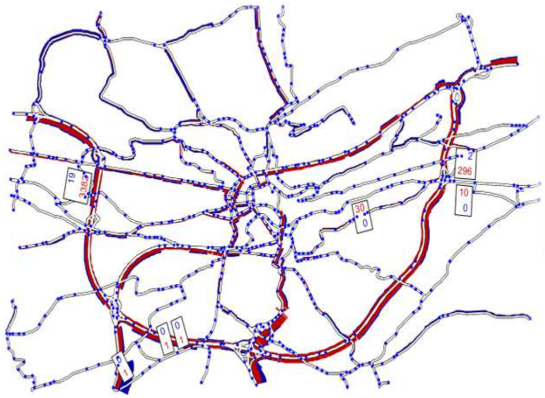



| Model | Reality |
|---|---|
| DEMAND MODEL | TRUE OD FLOWS |
| NETWORK SUPPLY | ACTUAL SUPPLY |
| TRAFFIC STATES | TRAFFIC DATA |





# History of traffic model applications and demand estimation requirements

Planning (static, quasi-static)

- 4-step models, activity-based models (see eg. Ortuzar and Willumsen, 2001, Cascetta, 2008, Timmermans and Arentze, 2010)
- OD matrix correction / adjustments from traffic data (see eg. Van Zuylen and Willumsen, 1980, Maher, 1981, Cascetta, 1984, Hazelton and Watling, 2001)

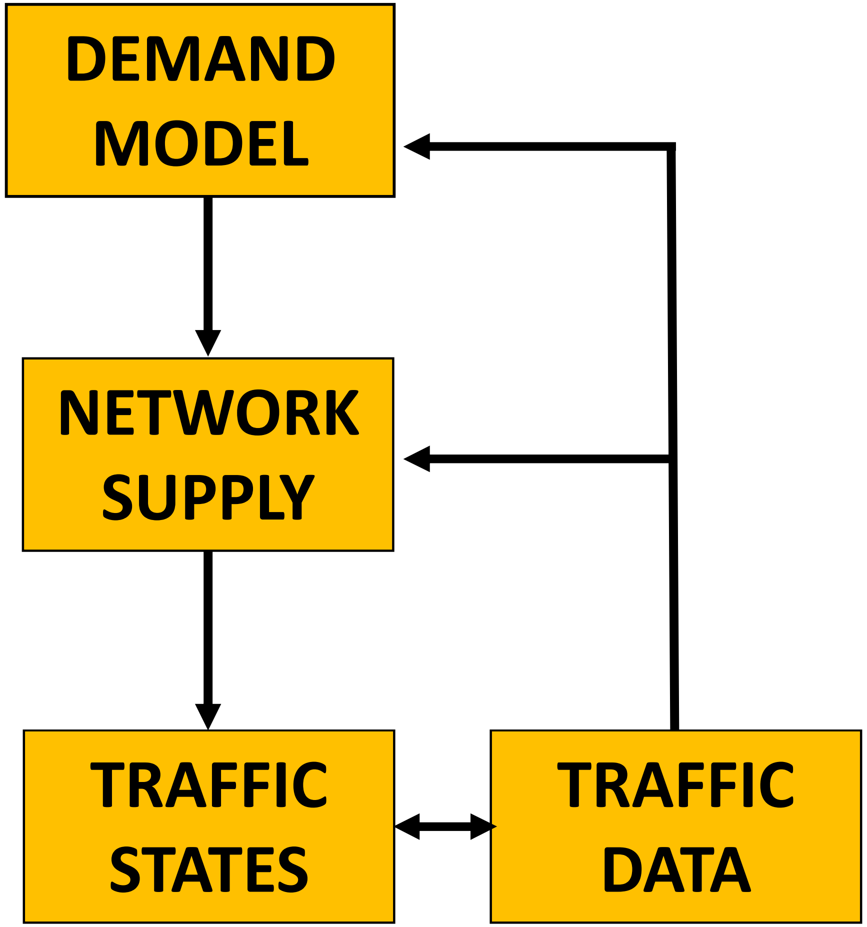
Traffic management (dynamic, offline)

- Quasi-dynamic / sequential / simultaneous (e.g. Cascetta, 2001, Marzano et al., 2012)
- DTA/DNL-based (see e.g. Ziliaksopoulos and Mahmassani, 1999, Tavana, 2001, Frederix, 2013, Cantelmo et al., 2014)

Real-time monitoring, information & control (dynamic, online)

- Data-driven (e.g., Cremer and Keller, 1987, Ashok and Ben-Akiva, 1993, Barcelo et al., 2011)
- Model-driven (e.g., Balakhrishna, 2001, Ashok, 2001, Zhou, 2004)

See Antoniou et al., Trans Res. C (2015) for a comprehensive overview



# The dynamic demand estimation problem

UNIVERSITÉ DU LUXEMBOURG

Goal: find most likely demand and supply characteristics that reproduce the data

$$\mathbf{x} = \arg\min_{x} \left[ \underbrace{\sum_t \sum_j f_1(x_j, \hat{x}_j)}_{\text{Distance btw estimated and seed matrix}} + \underbrace{\sum_t \sum_i f_2(y_i, \hat{y}_i)}_{\text{Distance btw estimated and observed traffic states}} \right]$$

$x$ = demand flows
$y$ = traffic flows

$$\text{s.t.} \quad y_i = \sum_{j \in J_i} A x_j = \sum_{j \in J_i} B(x) P(x) x_j$$

Traffic propagation functions (DNL)

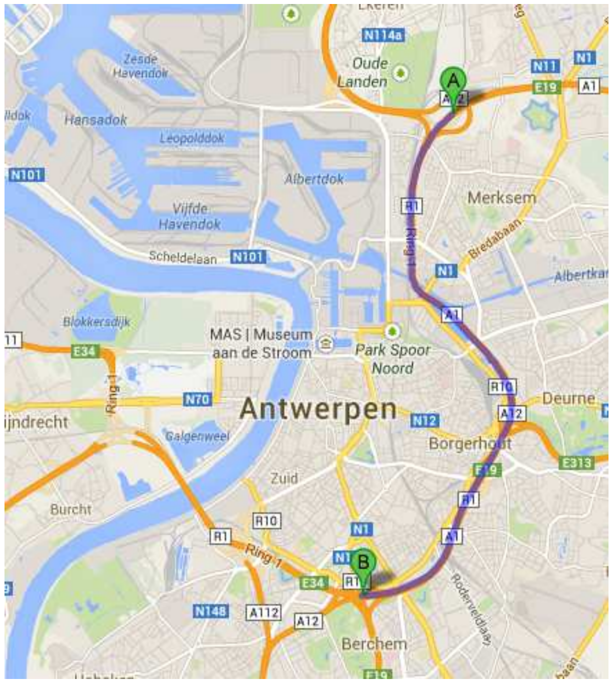
Travel behavior functions (DTA)

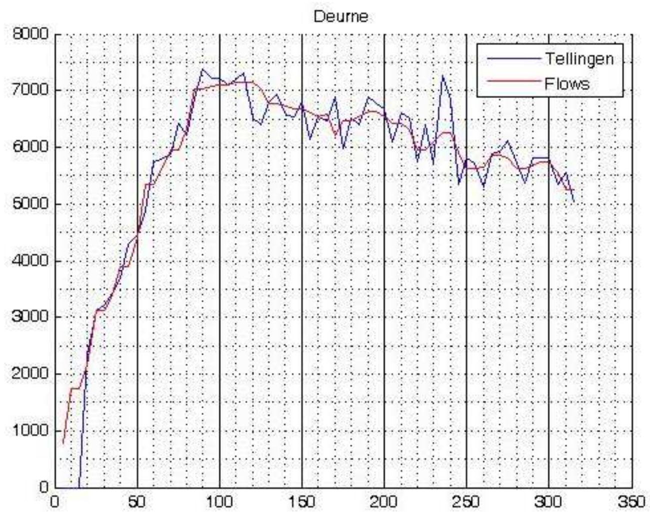
- Some issues
  - Highly combinatorial & non-linear problem
  - Complex dynamics due to travel behavior and traffic propagation phenomena
  - ‘Smart’ combination of demand and traffic information necessary
  - Traffic models (DNL/DTA) must be sufficiently realistic and accurate

# A simple example

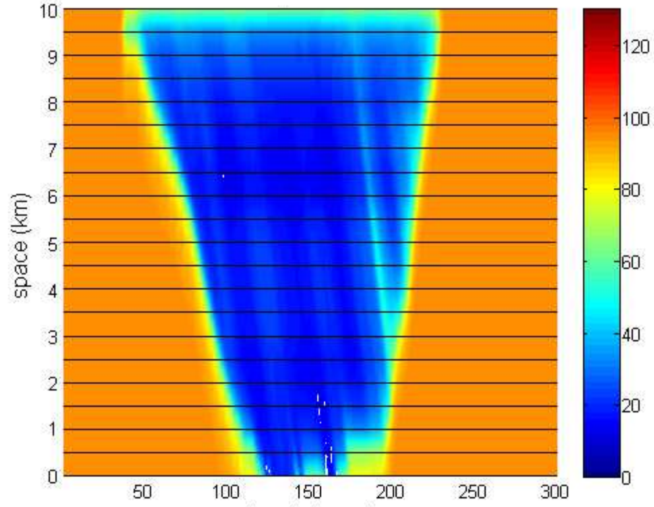
- Few route choice options
- Only traffic counts used for calibration
- Wrong seed matrix

- Better data and/or better seed can solve the issue

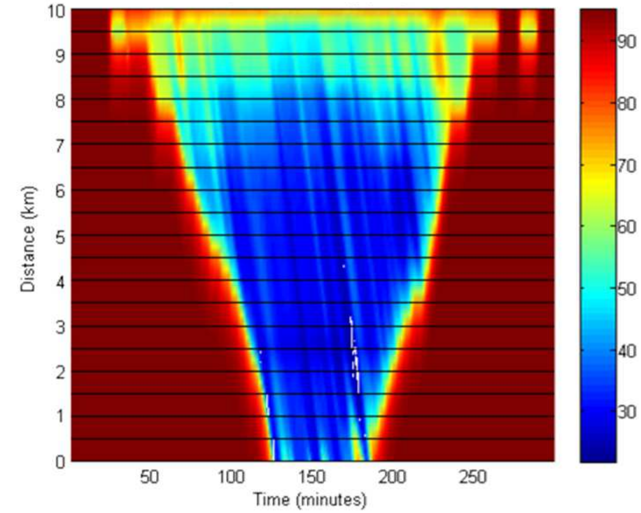



Measured speeds




Estimated speeds




Acknowledgments: Rodric Frederix Chris Tampere (KU Leuven)

# How to obtain a reliable dynamic estimation?

1. Look for demand information
    1. Through better demand data
    2. Through better demand models
2. Look for good traffic data
    1. Through smarter sensor locations
    2. Through exploiting new data types

Reduce solution search space and information reliability

3. Develop accurate dynamic traffic flow models
    1. Able to simulate traffic dynamics
    2. Able to reproduce congestion
4. Develop realistic travel behavior models
    1. Through good route choice models
    2. Through including activity-scheduling info

Reduce the mismatch between model and reality

**Need for smarter data and smarter models!**

# The era of Big (Transport & Mobility) Data and how it can be used to improve demand modeling

UNIVERSITÉ DU LUXEMBOURG

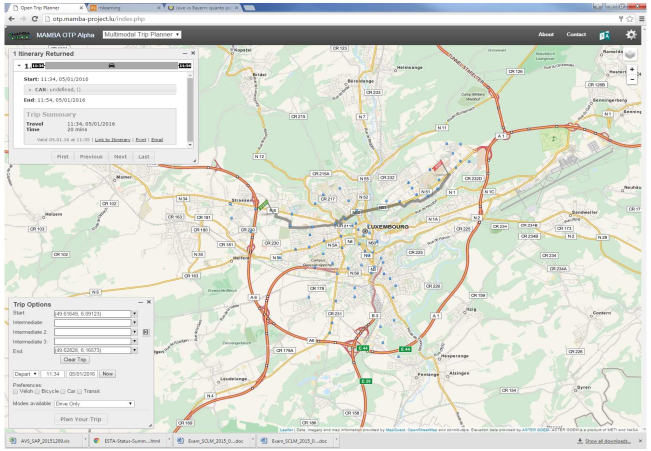
New opportunities, old problems

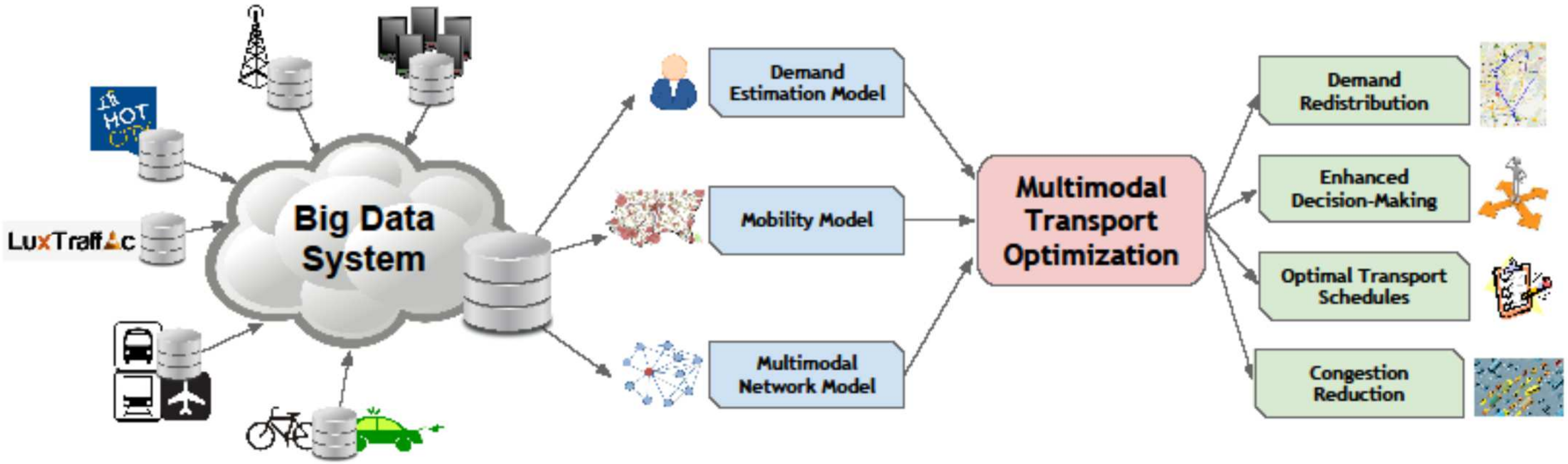
- Traditional traffic Data → many opportunities and threats
- **Big Data ➔ many Big opportunities and threats**

# Current research directions at UL using Big Data

- Big Data-based applications
  - Demand estimation
  - Multimodal modelling
  - Mobility analysis
  - Travel planners
- Real data available of Luxembourg
  - Cellphone data (Post)
  - GPS-based & WiFi-based FCD (Be-mobile, LuxTraffic)
  - Smartphones & smartwatches (go2uni platform)
  - Other ‘more traditional’ (counts, PT info, bikesharing, OSM,…)

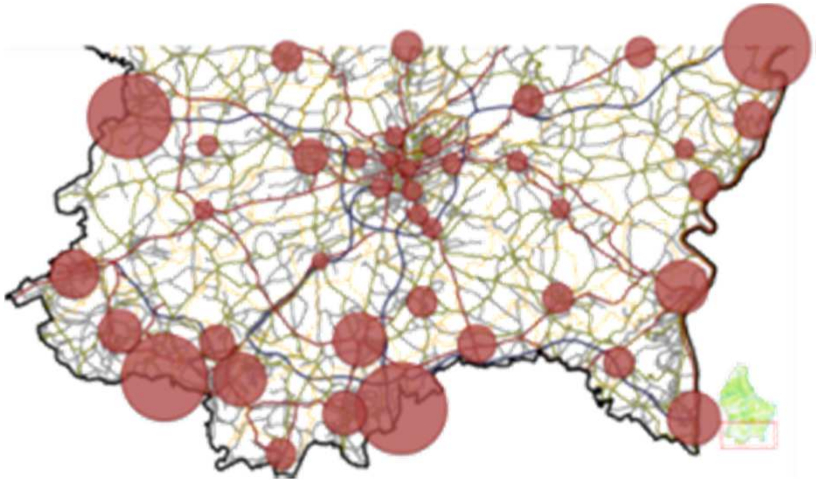
MAMBA MULTIMODAL MOBILITY





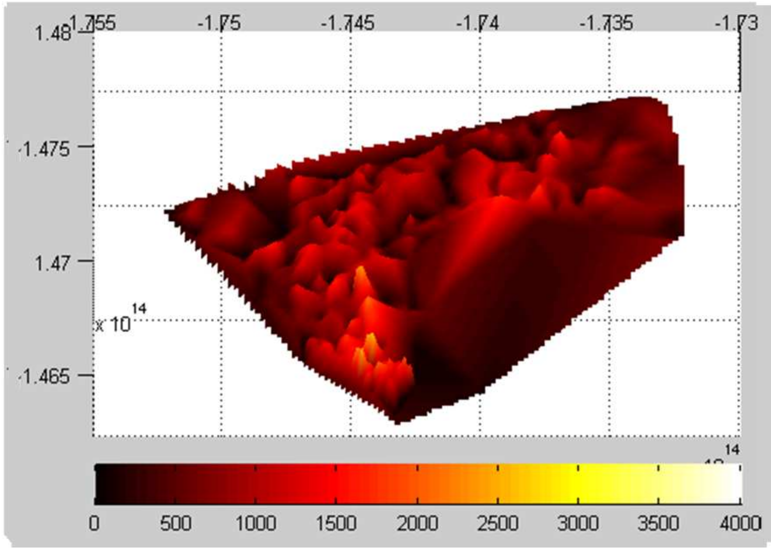
# Using cellphone data for demand modeling

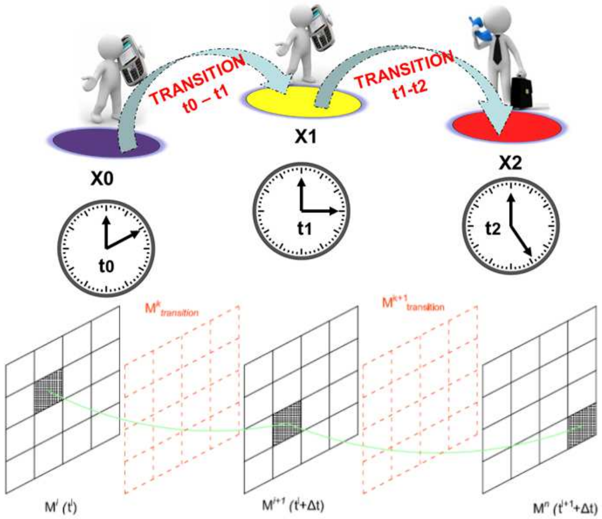
- PRO's
  - Nearly full geographical coverage
  - 93.2% persons subscribed in the World
  - Continuous data flow
  - Information-rich data for demand prediction
- CON's
  - Very hard to get from operators
  - Includes any type of mobility (and not only)
  - Hard to distinguish modes
  - Antenna densities vs. Call Detail Records

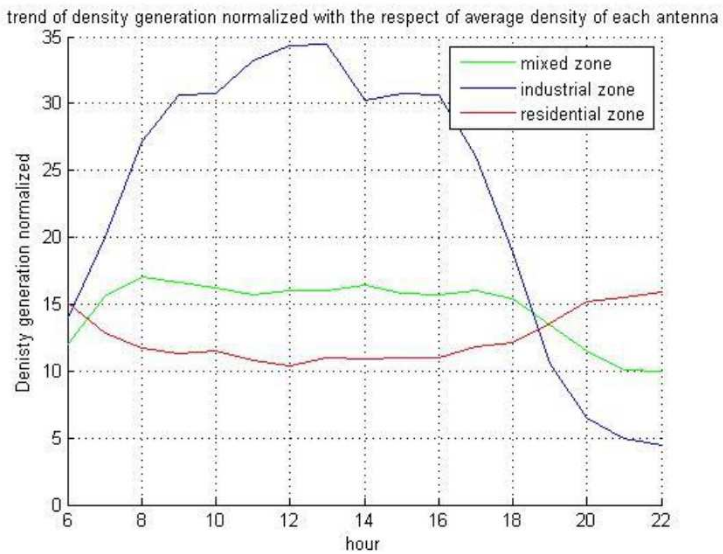


# Ongoing work with cellphone data (1)

- Using GSM data for estimating and predicting dynamic production and distribution of flows
  - Mobility dynamics modelled as Markov-chains
    - Production ~ antenna densities
    - Distribution ~ transitions between antennas
  - Automatic zoning through clustering
  - Semantic interpretation of demand patterns
    - Main zone activity automatically detected



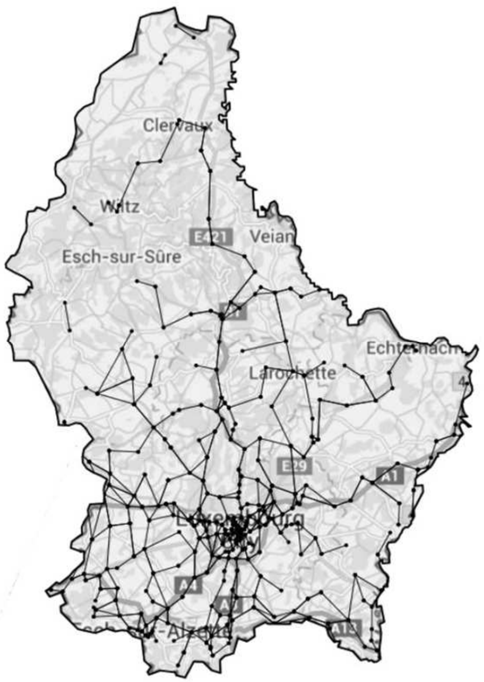



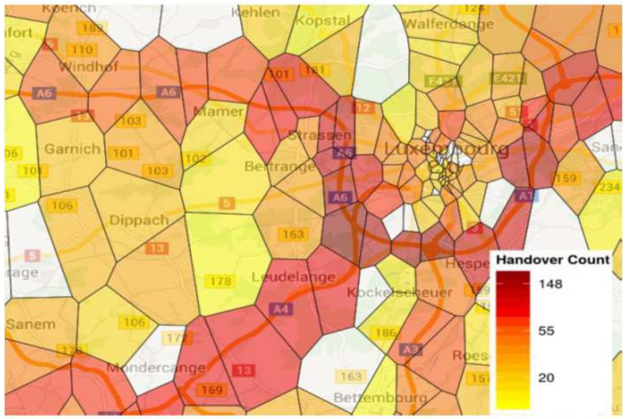




Acknowledgments: Guido Cantelmo, Simone Di Donna (UL)

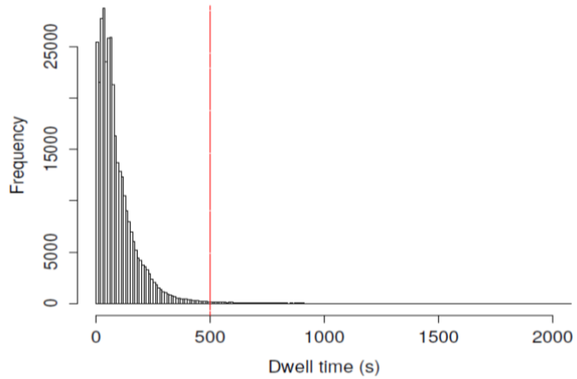
# Ongoing work with cellphone data (2)

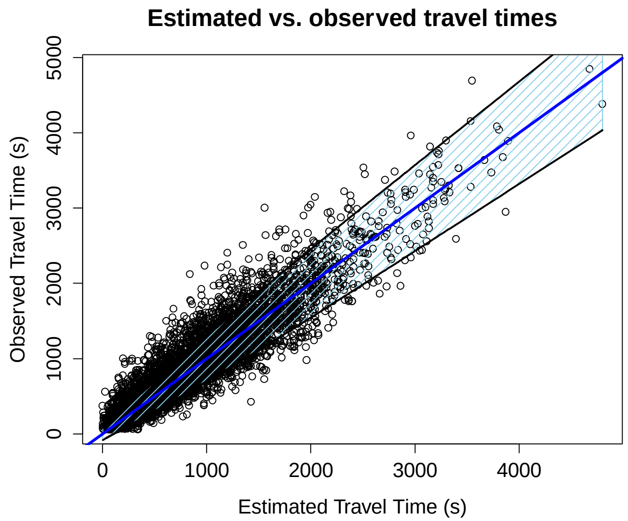
- Estimating travel time and dwell time distributions
  - Correlations between handovers and travel times
  - Using dwell time as predictors of travel times
  - FDCs used as ground truth





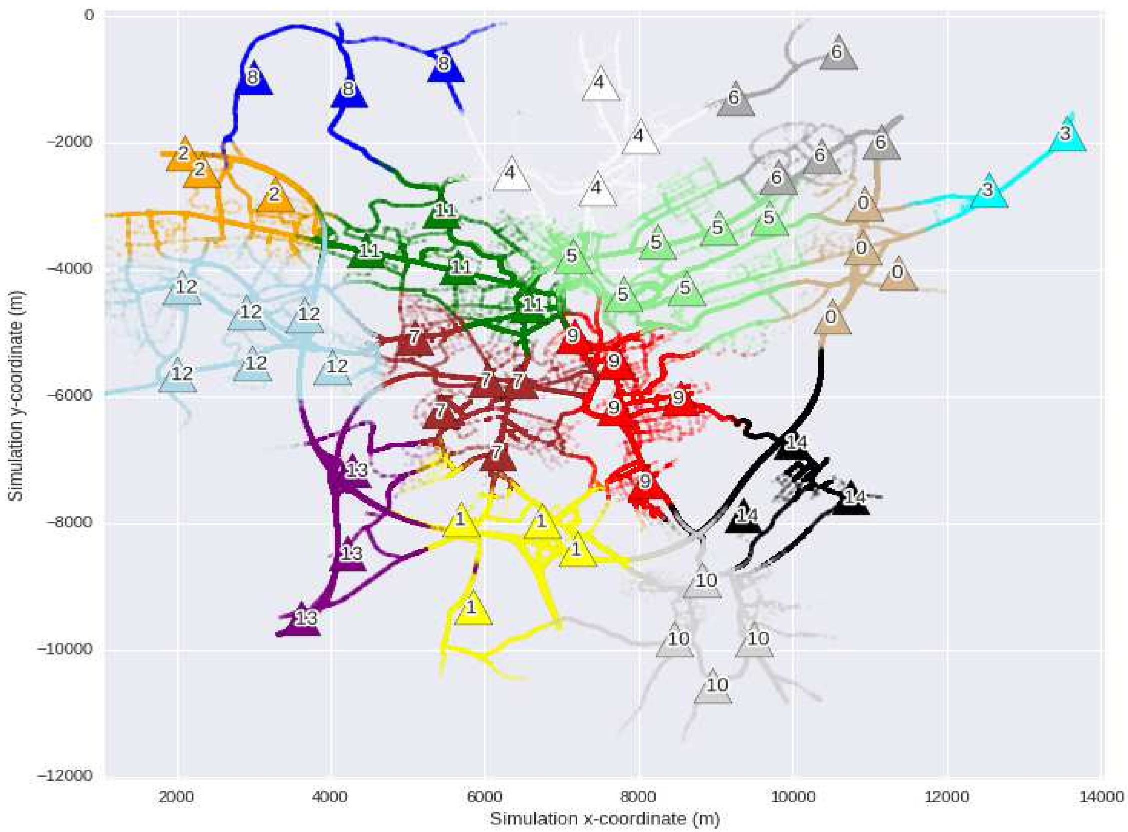


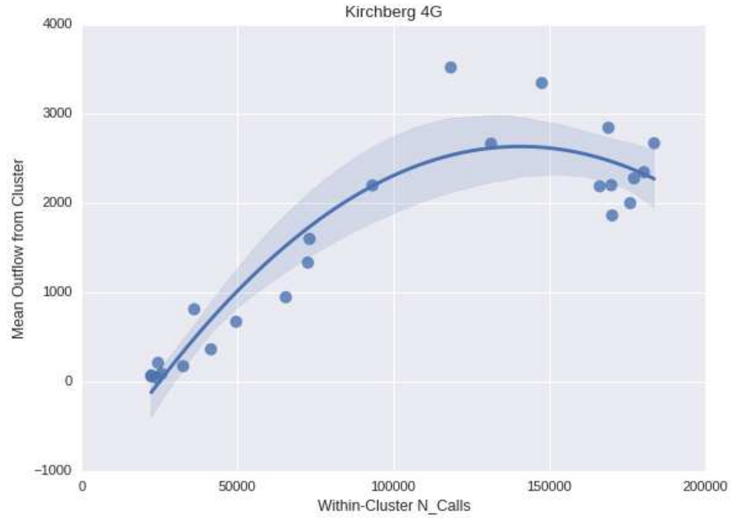



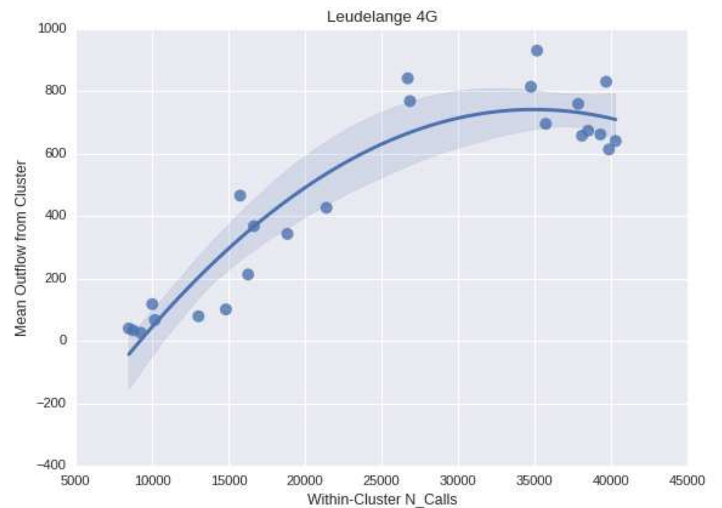

Acknowledgments: Raphael Frank & Thierry Derrmann (SnT)

# Ongoing work with cellphone data (3)

- MFD estimated from mobile phone data*
  - Exploiting the relation between number of handovers and zone outflows
  - Clustering homogeneous zones in MFD terms



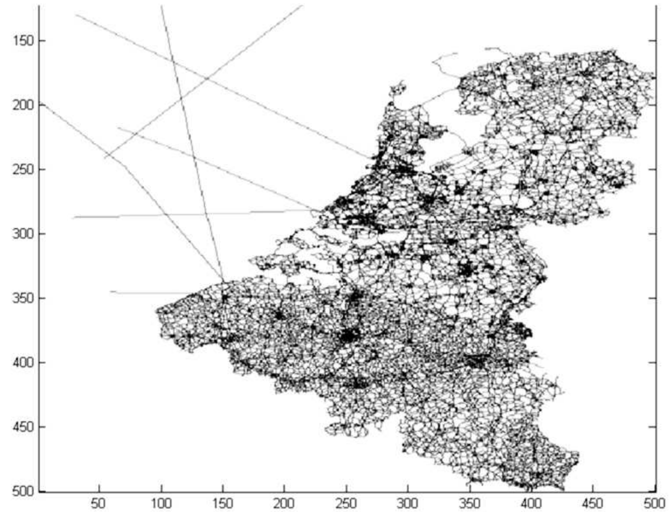

Acknowledgments: Raphael Frank & Thierry Derrmann (SnT)

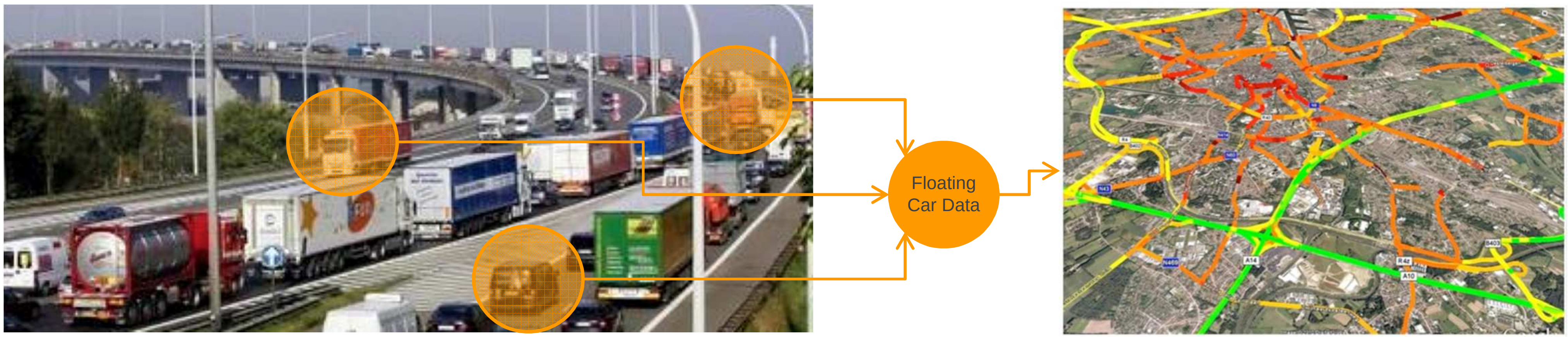
*so far results only done with SUMO simulated data

# Using GPS/FCD data for demand modeling

UNIVERSITÉ DU LUXEMBOURG

- PRO's
  - Flexible and accurate, i.e. data can be 'virtually' available any time and anywhere
  - Provide full routes travel time data, speed on segments
  - Low installation/maintenance costs
  - Sample sizes grow with traffic density
- CON's
  - Not necessarily tracing only vehicles
  - Scalability
  - Biases (e.g. taxis, lorries,…)
  - Coverage

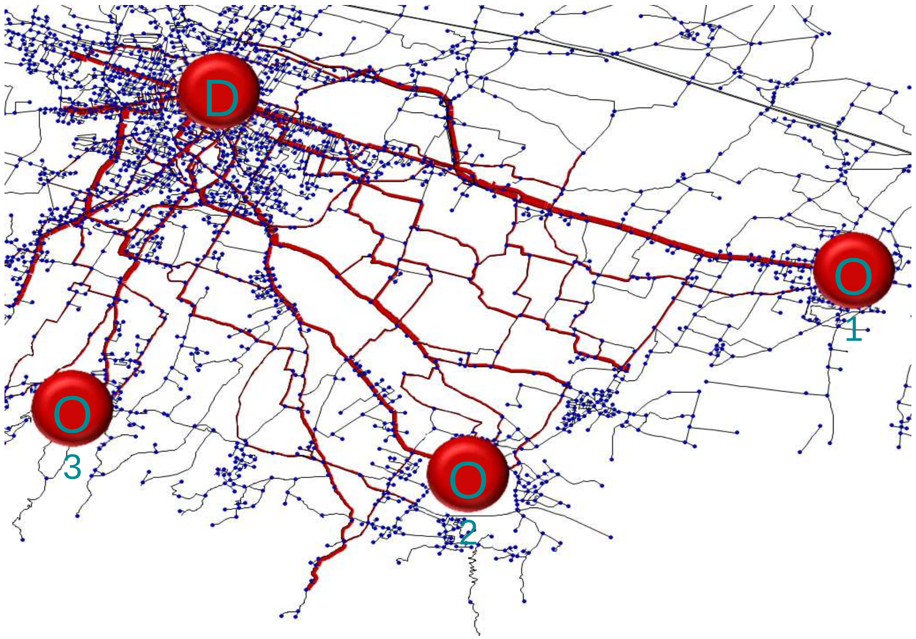


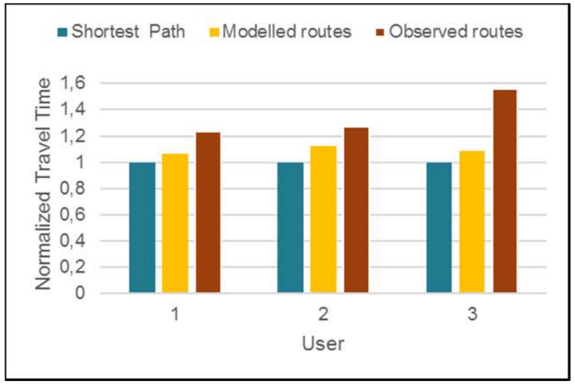



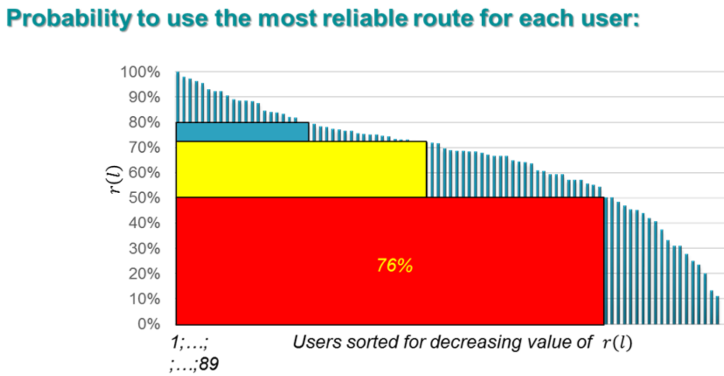

Acknowledgments: Steven Logghe and Isaak Yperman (Be-Mobile)

# Ongoing work with GPS/FCD data (1)

- Improving route choice models
  - Repetitive trips on multiple OD pairs
  - Time-varying travel time information
  - Equilibrium vs. non-equilibrium analysis
  - Reliability analysis



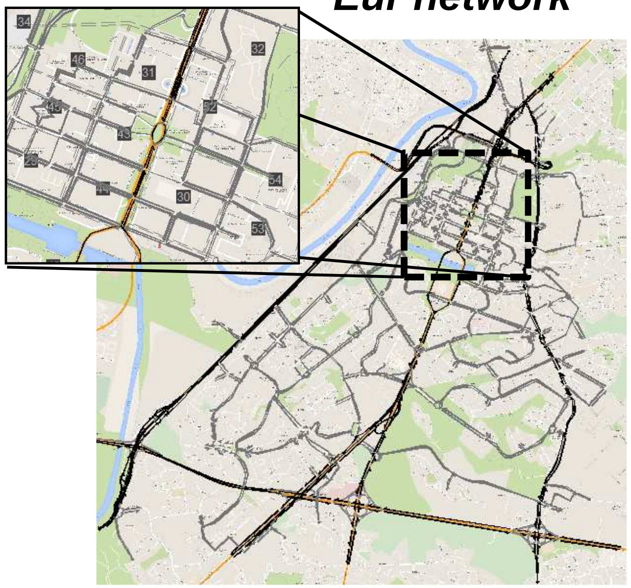




Acknowledgments: Guido Cantelmo (UL) & Selini Hadjidimitriou (Univ. Reggio Emilia)

# Ongoing work with GPS/FCD data (2)

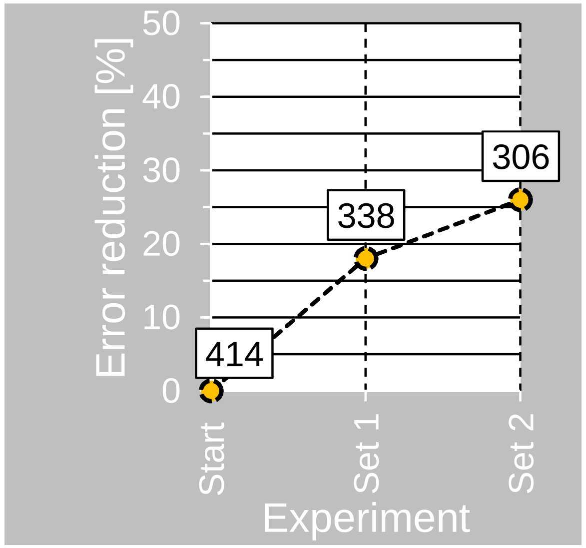
- Using route travel times in demand estimation
  - FCD-based travel times
  - No reliable seed matrix available
  - Counts available on few links

**Eur network**

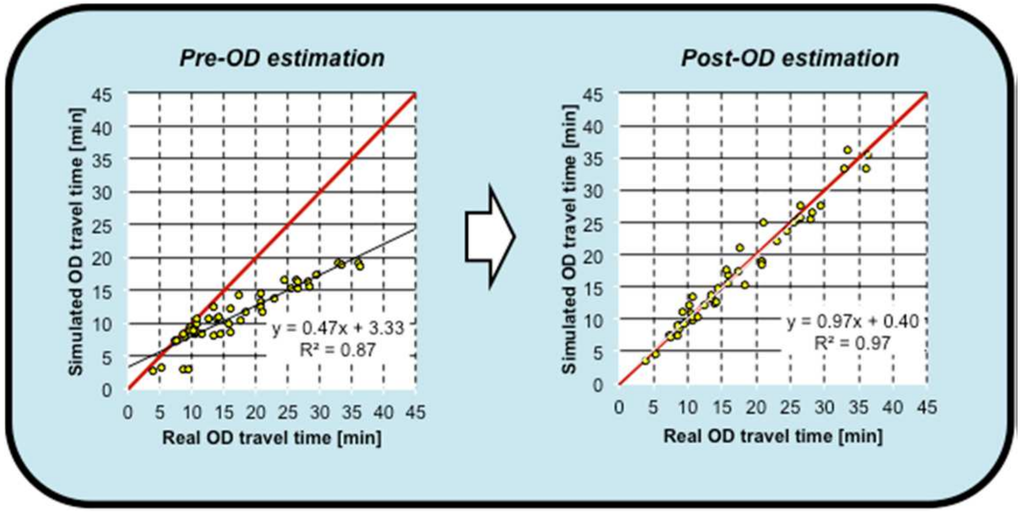


| **Measures\Set** | **Set 1** | **Set 2** |
|---|---|---|
| OD target | + | + |
| Link Flows | + | + |
| OD travel time | | + |

Spatial – Temporal distribution



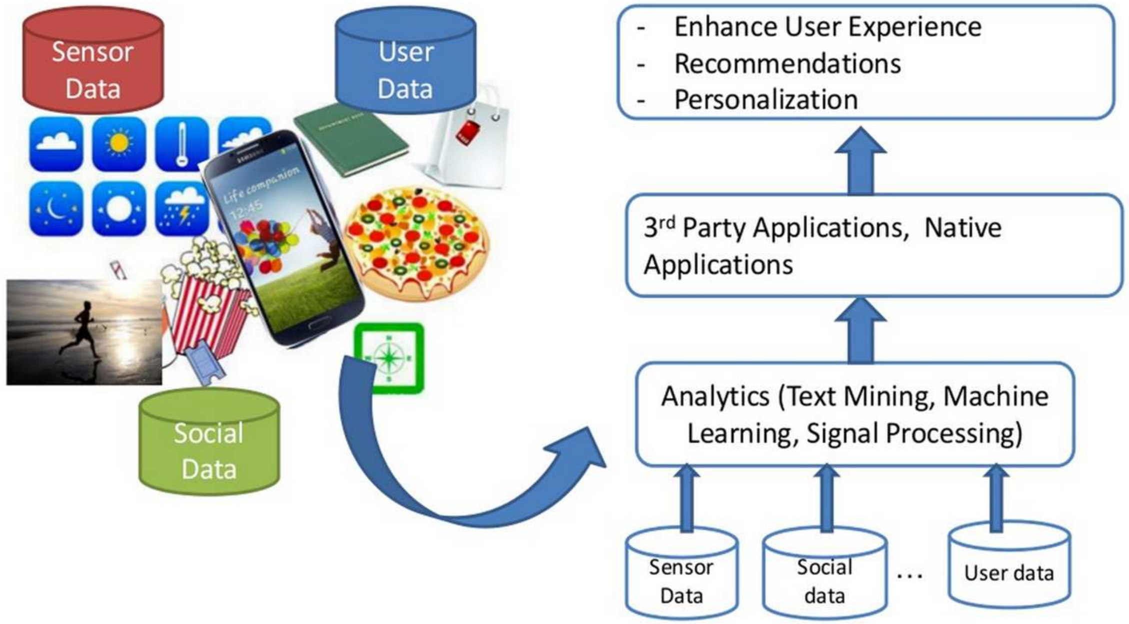

☐ Euclidean Distance




Acknowledgments: Andrea Del Giudice, Ernesto Cipriani & Marialisa Nigro (Roma Tre) & Guido Cantelmo (UL)

# Using smartphone (and smartwatch) data for demand modeling

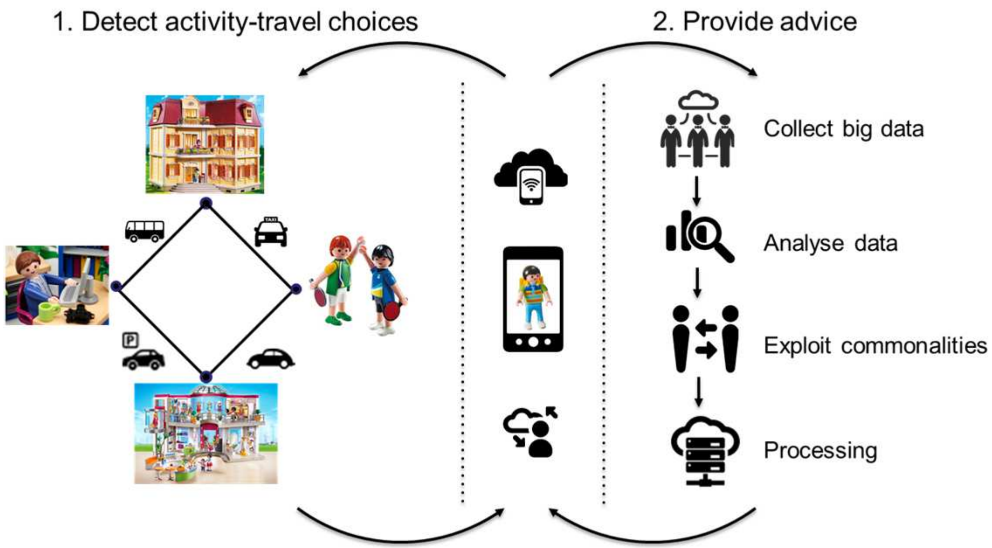
- PRO's
  - Same as GSM & GPS data
  - Many more sensors (accelerometer, gyroscope,…)
  - Closing the information loop
  - Connecting technologies and people (smartcities)
- CON's
  - Data trust and privacy
  - Crowdsourcing issues with reliability
  - Sensing is battery draining



# Ongoing work with smartphone data (1)

- Using smartphones and smartwatches to study activity-travel patterns
  - Automatic mode and purpose recognition
  - Sensor fusion opportunities
  - Activity-travel behavior analysis
  - Analyze transport service usage
  - Correlating service users
  - Provide personalized travel plans



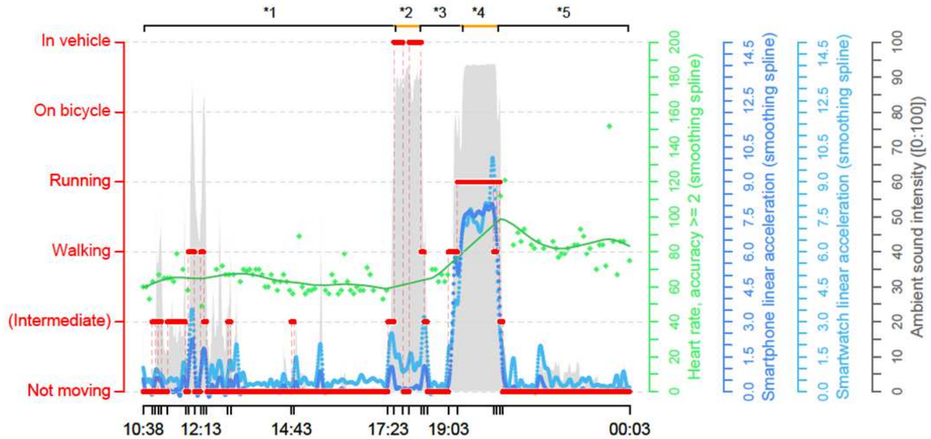




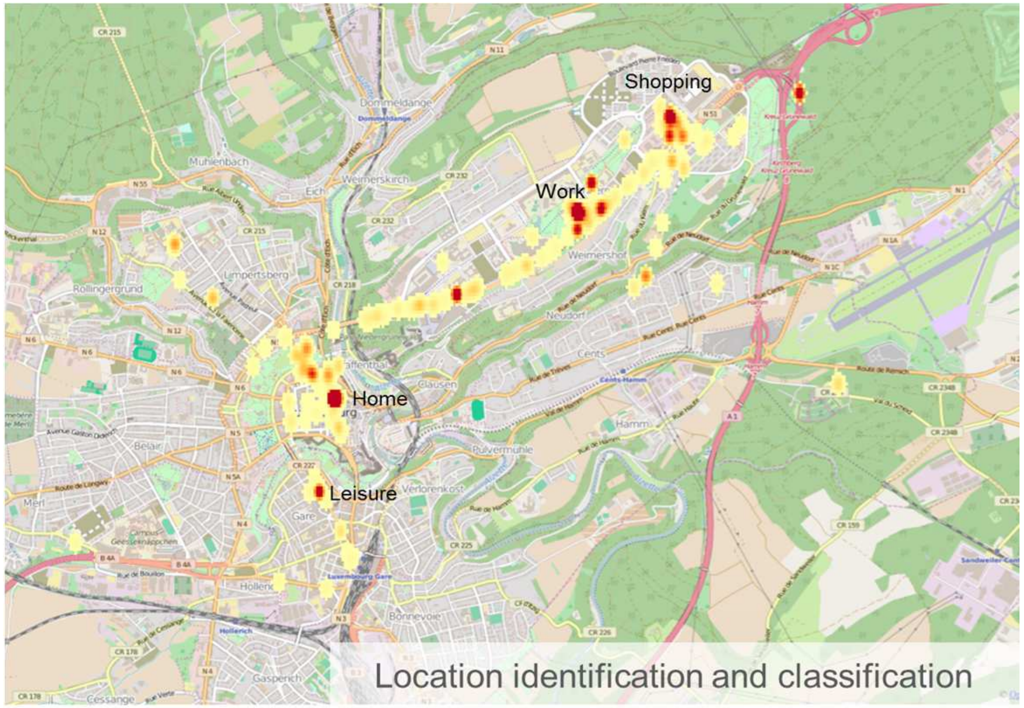

Acknowledgments: Bogdan Toader, Francois Sprumont (UL) & Sebastien Faye (SnT)

# Ongoing work with smartphone data (2)

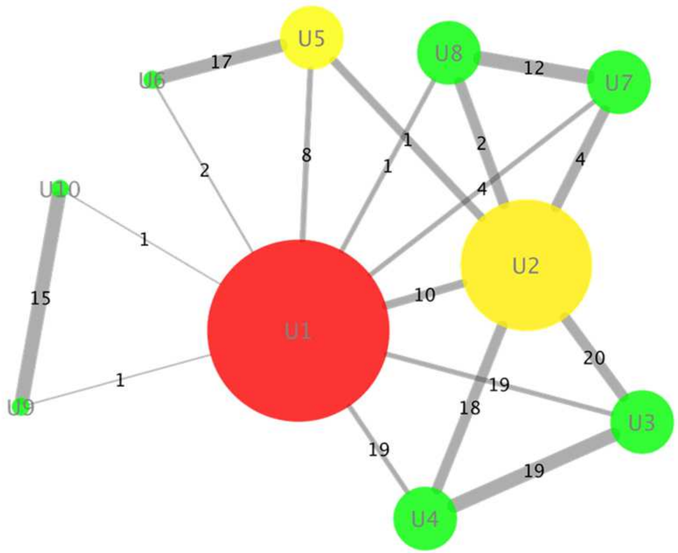
- Advanced data collection methods for activity and mode recognition
  - Data fusion
  - Location analysis
  - Correlating users in social networks
  - Indices for shared service matching opportunities




Travel mode and activity identification




Location identification and classification

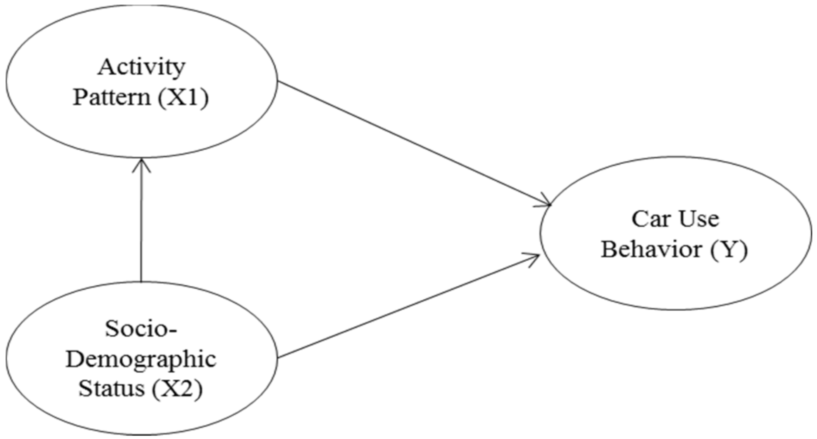



Social network graph by frequency of interactions

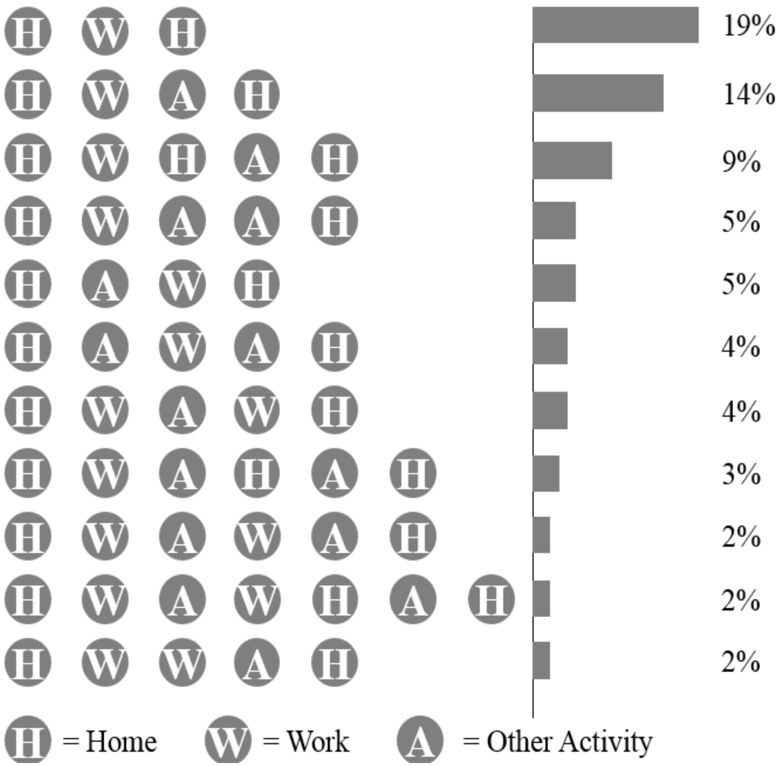
Acknowledgments: Bogdan Toader (UL) & Sebastien Faye (SnT)

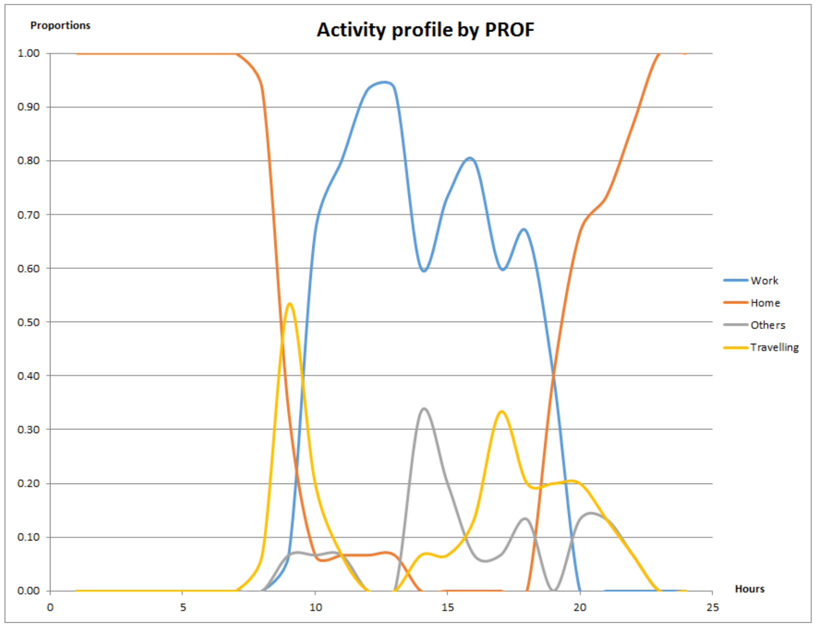
# Ongoing work with smartphone data (3)

- Understanding activity-travel patterns*
  - Repetitivity and regularity in demand
  - Relation btw activity patterns and mode choices
  - Response to policies and TDM




The 11 most recurrent activity chains



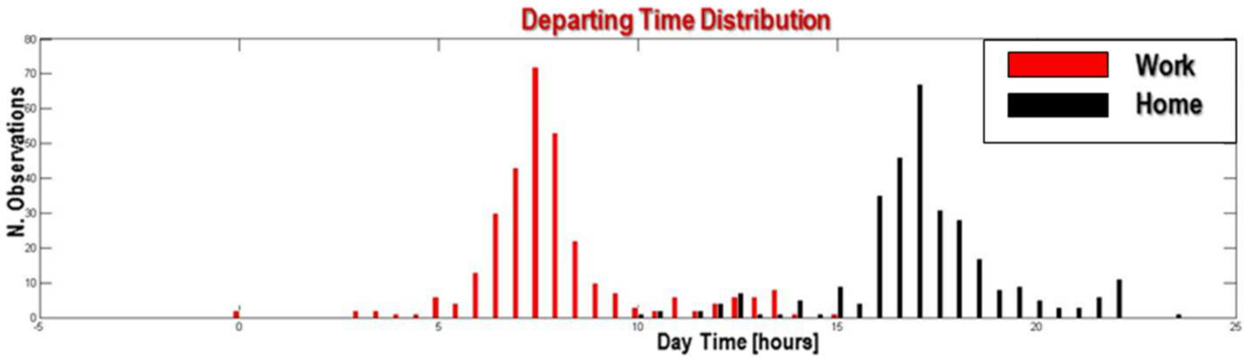



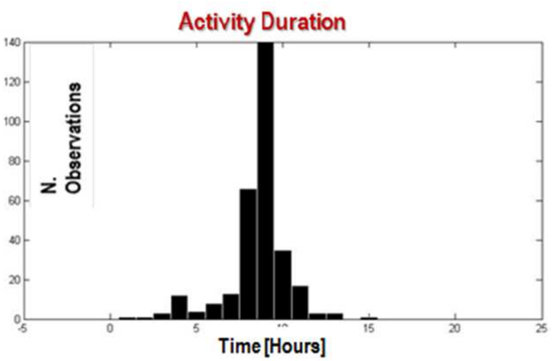

Acknowledgments: Francois Sprumont (UL)

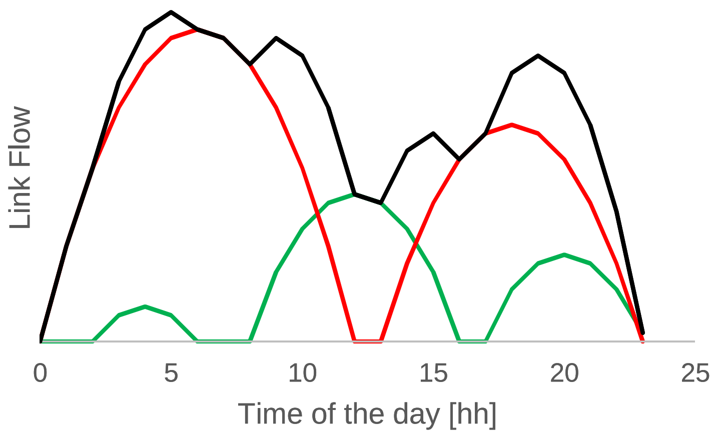
*so far results only with travel diaries data

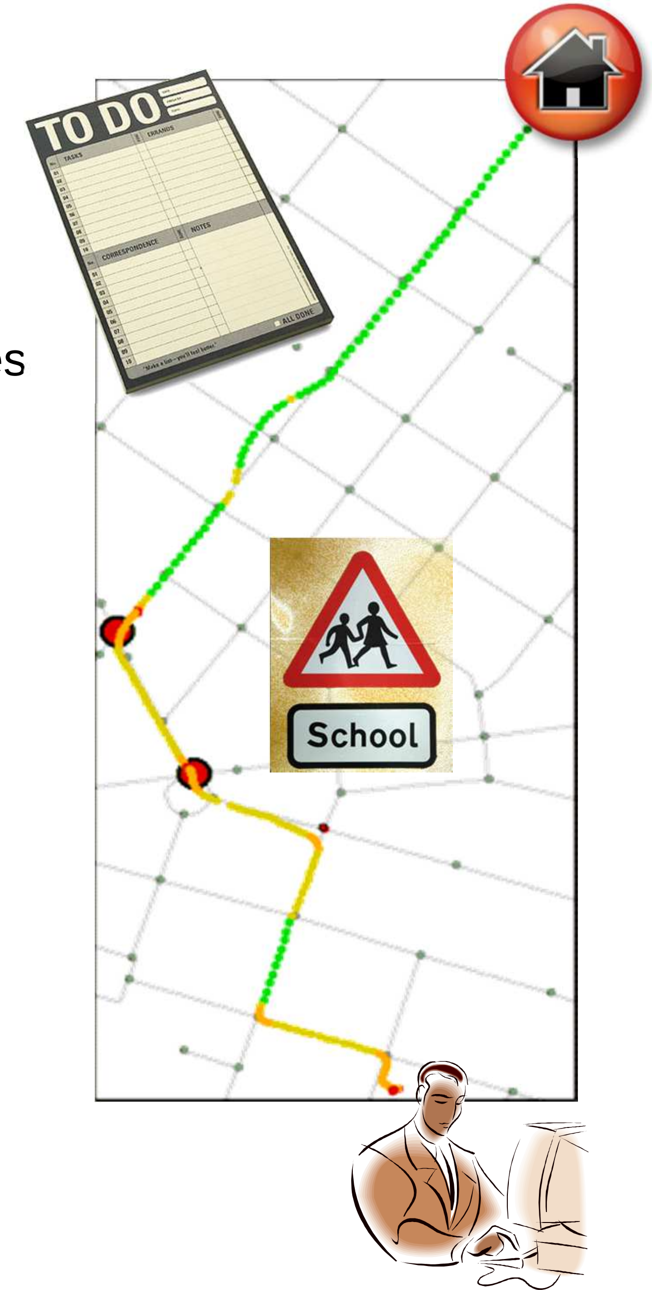
# Connecting activity scheduling and duration with demand estimation models

- Including activity information in demand estimation
  - Mobility needs of travelers → activity-travel patterns → observed trips
  - Trips link activities ←→ activities constrain travel behavior
  - Trip chains → scheduling of activities constrain trip schedules



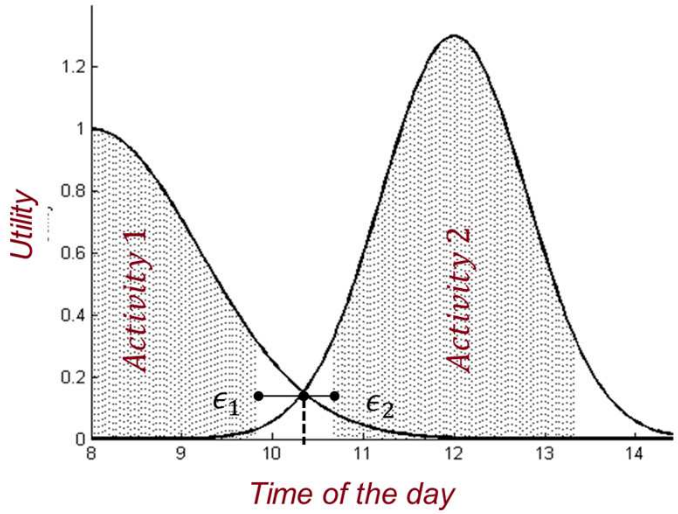



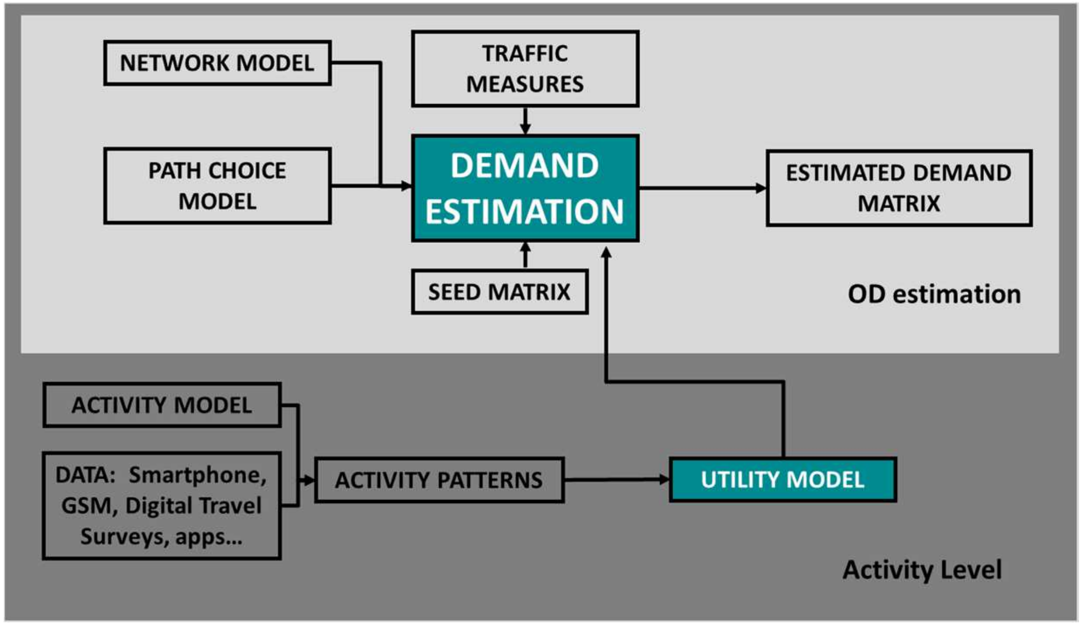



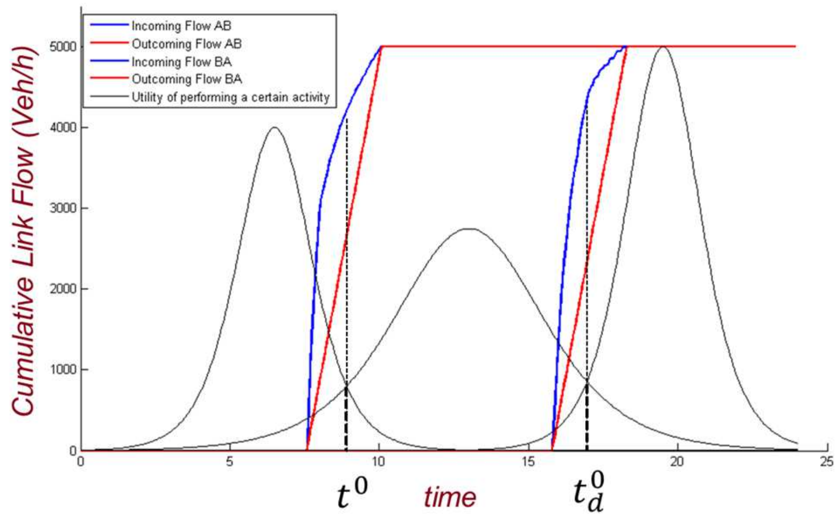
# Including activity information in demand estimation

UNIVERSITÉ DU LUXEMBOURG

- Utility-based demand modeling and estimation
  - Activity sequence and duration as input
  - Deriving activity primitives
  - Activity scheduling as trade-off problem
  - Reducing the localism in the optimization




Acknowledgments: Guido Cantelmo (UL)

# Closure

- New Big Data era: new opportunities and challenges
  - Understanding mobility needs
  - Forecast future activity-travel patterns
  - Enable users with enhanced information
- Examples of transport applications
  - Dynamic traffic modelling
  - Multimodal travel planning
  - Decision support services
  - Transport systems optimisation
  - …

# THANK YOU!

Francesco.viti@uni.lu