University of Luxembourg

Multilingual. Personalised. Connected.

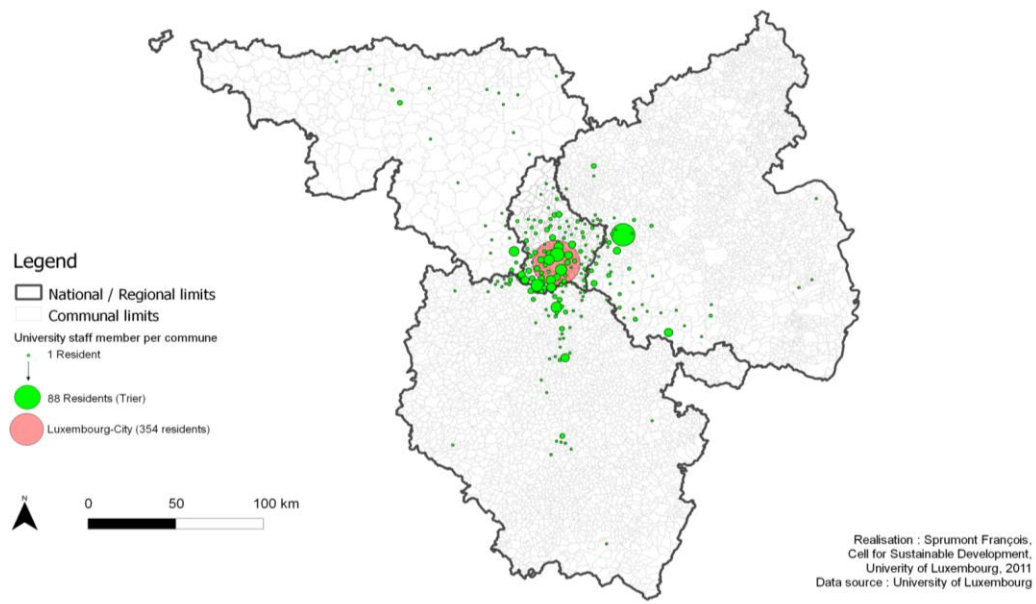
MobiLab@uni.lu

UNIVERSITÉ DU LUXEMBOURG

Transport and Mobility in Luxembourg

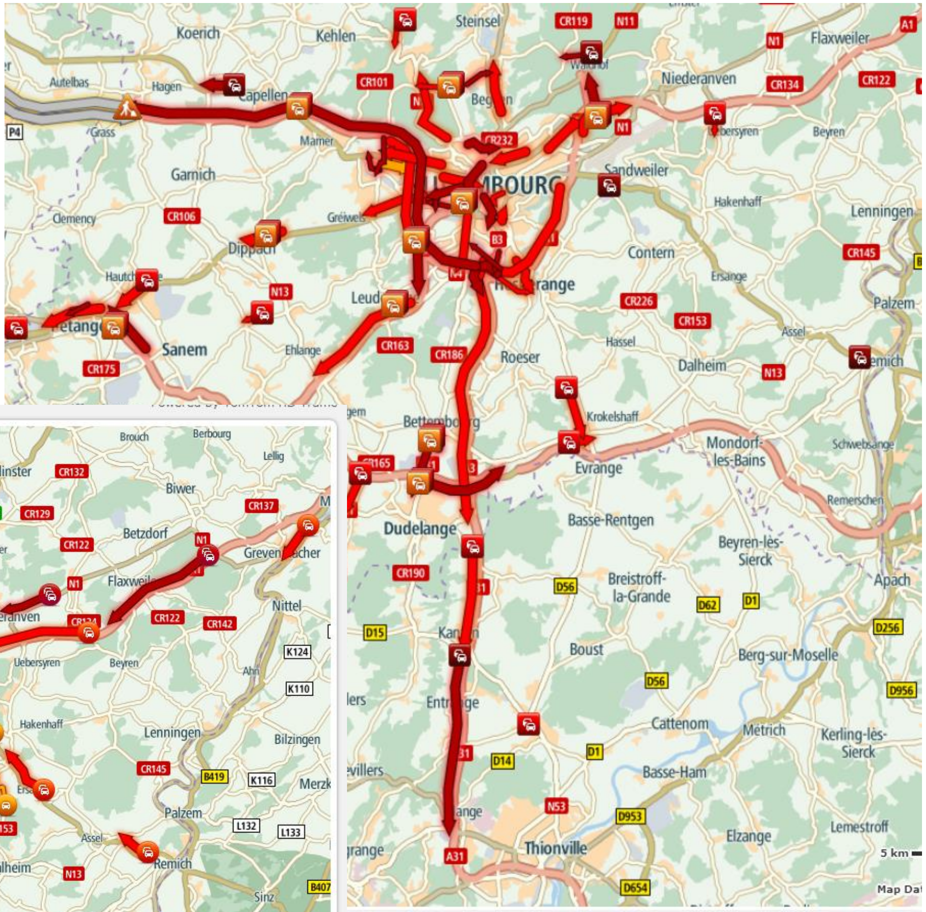
- **Luxembourg strong monocentric country**
- **320 000 commuters; 160 000 cross-borders;**
- **76% car users (89% from outside); #1 car ownership rate in EU**

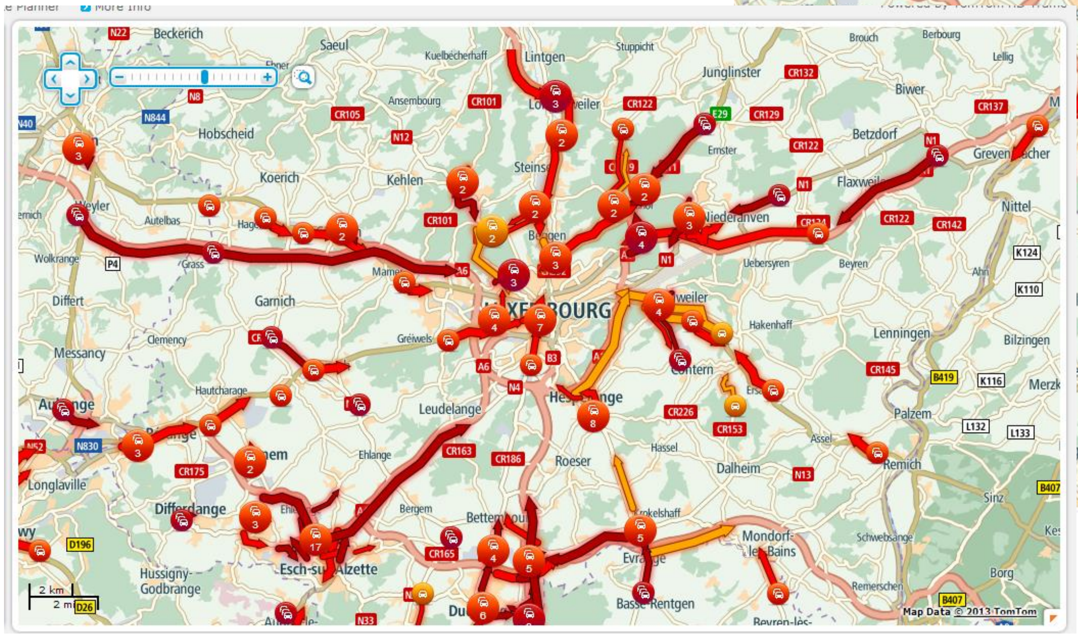
Communes of residence of the University staff living in the Greater Region



Traffic congestion in Luxembourg

- Regional gridlocks
- Queues spilling back beyond the country borders!



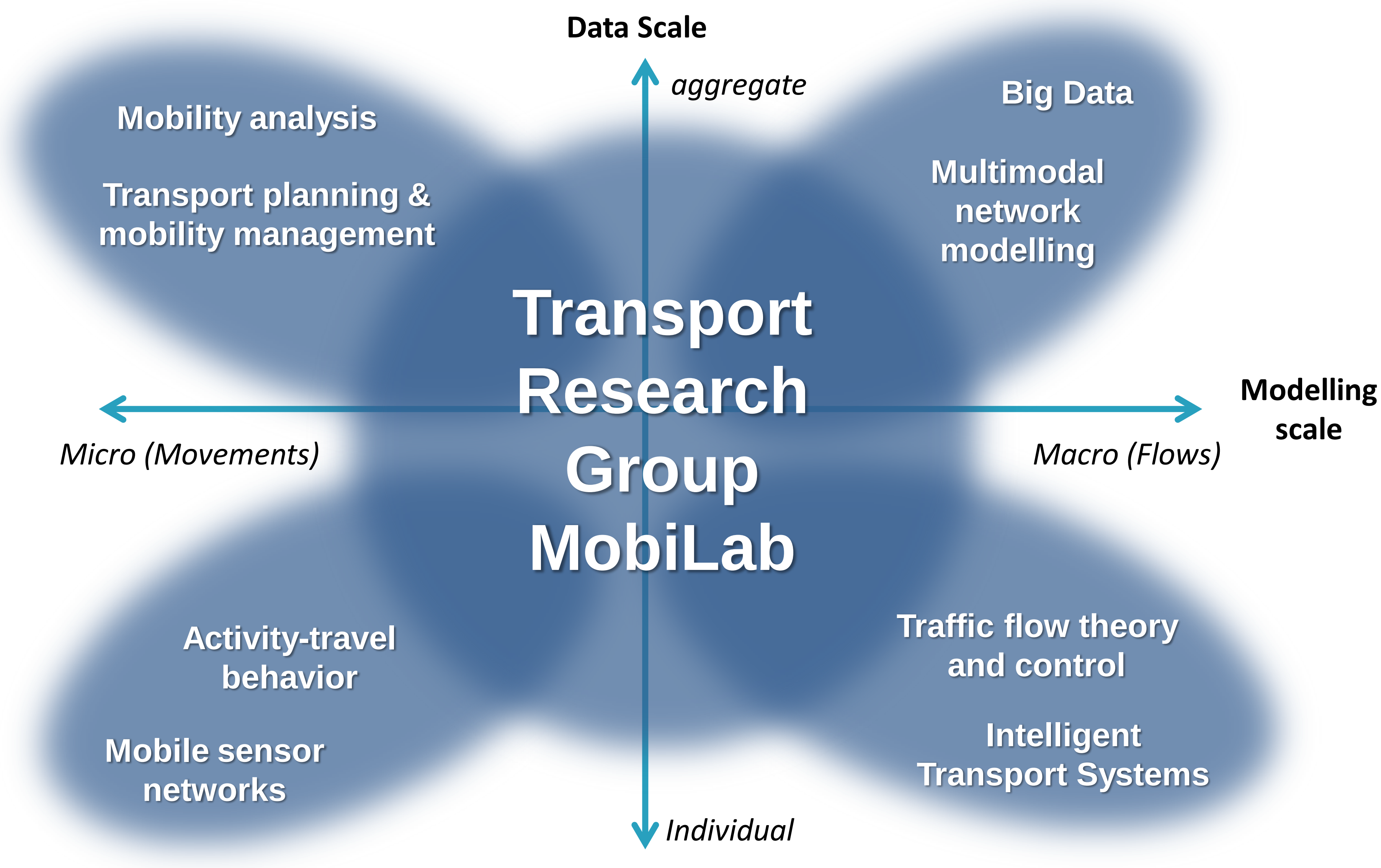


MobiLab @ University of Luxembourg

UNIVERSITÉ DU LUXEMBOURG

- Transport Research Group within Engineering Unit since 2012
- International and interdisciplinary team
  - Head: Prof. dr. Ing. Francesco Viti
    - MSc – Univ. of Naples ‘Federico II’, Civil Engineering degree
    - PhD – TU Delft, PhD in transportation planning and management
    - Post-doc – TU Delft (2007-2008) & Ku Leuven (2007 – 2012)
  - 1 (part time) post doc
    - Sebastien Faye, computer scientist
  - 3 PhD students
    - Francois Sprumont, spatial planner
    - Guido Cantelmo, transport engineer
    - Bogdan Toader, computer scientist
  - Incoming
    - PhD position 1: Giorgos Laskaris, traffic engineer (Jan. ’16)
    - Post doc – Marco Rinaldi, automation and control (Mar 2016)
    - PhD position 2 – to be filled, transport engineer (Summer ’16)

Research at MobiLab: multimodal, multiscale, multi-data



Collaborations

- Within UL
  - Computer Science & SnT
    - NetLab (VehicularLab, IGNITE) – travel assistance systems, (Big) data and mobility, Gaming
    - AutomationLab – autonomous driving, vision and image processing
  - Social Sciences (FLSHASE)
    - IPSE – Activity-travel behavior, mobility planning & management, transport policy research
    - HCI-usability Lab – Human Factors, Human Computer Interface

- Outside UL
  - Luxembourg
    - Luxembourg Institute of Science and Technologies (LIST)
    - Luxembourg Institute of Socio-Economic Research (LISER)
    - Stakeholders (Ministries of Economy, Infrastructure, Sustainable Development, PT operators, Infomobility,…)
  - International
    - KU Leuven
    - TU Delft
    - Universities of Rome ‘La Sapienza’ & ‘Tre’
    - KTH

1. Speed and travel time profiles and distributions from mobile sensors

For Route Planning

UNIVERSITÉ DU LUXEMBOURG

Overview

- Introduction
  - IBBT Project - MobiRoute
- Basic info of FCD technology adopted
  - Coverage requirements
  - Quality requirements
  - Converting point-based data into travel times
- Data conversion and cleansing
  - Detecting and removing biases and errors
  - Handling data correlations
- Data analysis
- Closure

Multimodal Route Planning

UNIVERSITÉ DU LUXEMBOURG

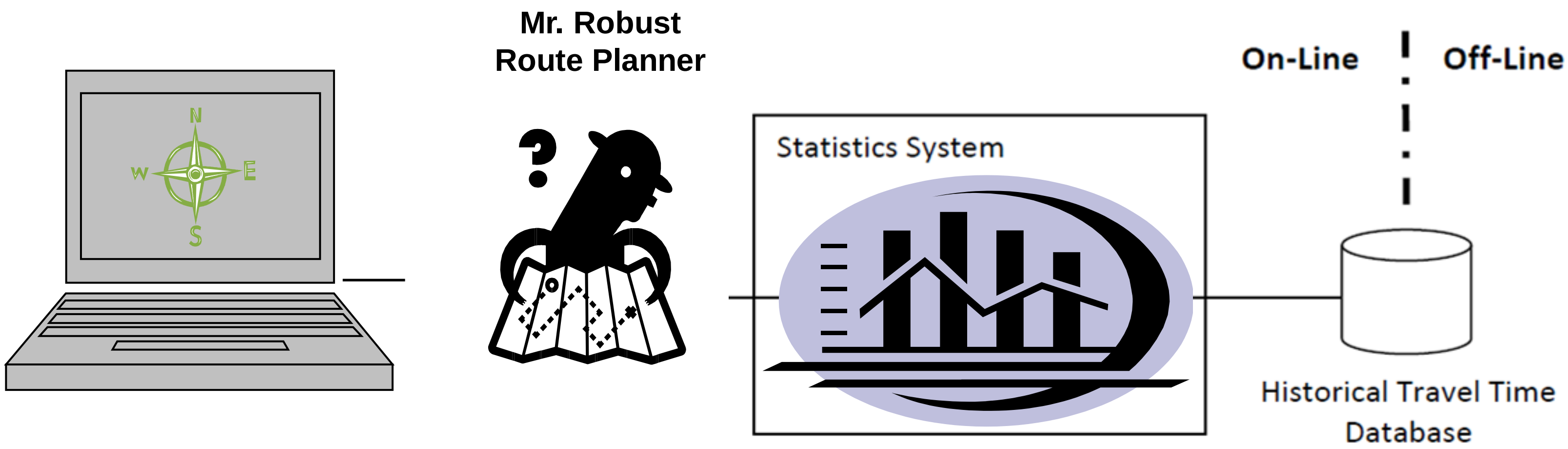
- **MobiRoute - Mobility & Routing**
- IBBT ICON funding for June 2009 - June 2011
- Aim: develop a dynamic and robust route planner using historical traffic data and other metadata (eg weather) for multimodal (car+train) trips.
- Achievements:
  - Unique dataset containing both floating car data (Be-Mobile) and real time train data (DUO);
  - High-performance web-based multimodal route-planner with robust routing (UGent);
  - **Use of advanced statistics to obtain reliable predictions of speeds and travel times (KUL)**
  - Spin-off company Go-Mobile as mean between info providers and services (Be-Mobile, SNCB, De Lijn, …)

Robust routing

- <u>Proposed definition:</u>

  Given an origin-destination pair, and a certain arrival (or departure) time period, display the *k* best routes (if exist) such that:

  1. The mean travel time does not exceed *a*-times the average travel time of the shortest route during the same time interval
  2. Given a certain α probability value, the travel time of the route is at maximum *b*-times its free flow travel time with α probability.



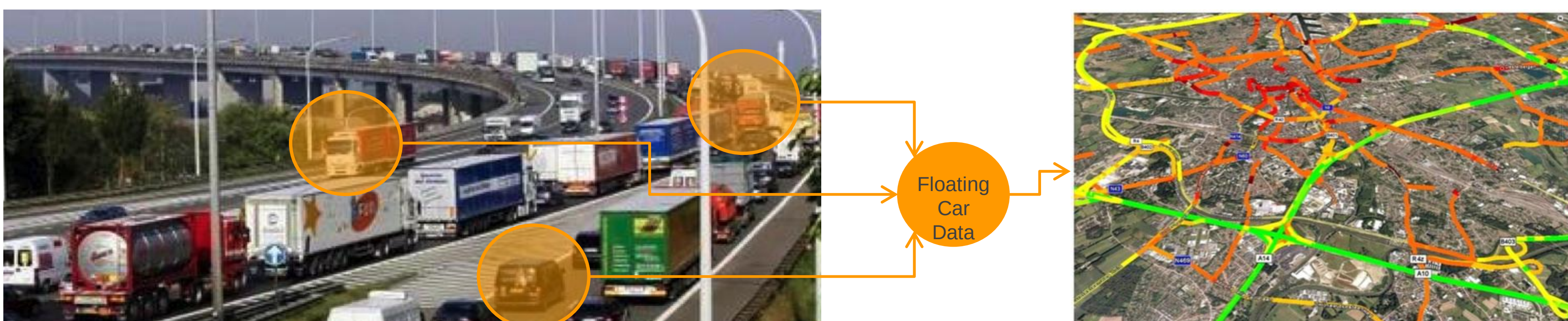
→ Route travel time histograms needed!

Floating Car Data technology

- Advantages
  - Provides full routes travel time data
  - Low installation/maintenance costs
  - Sample sizes grow with density and congestion levels

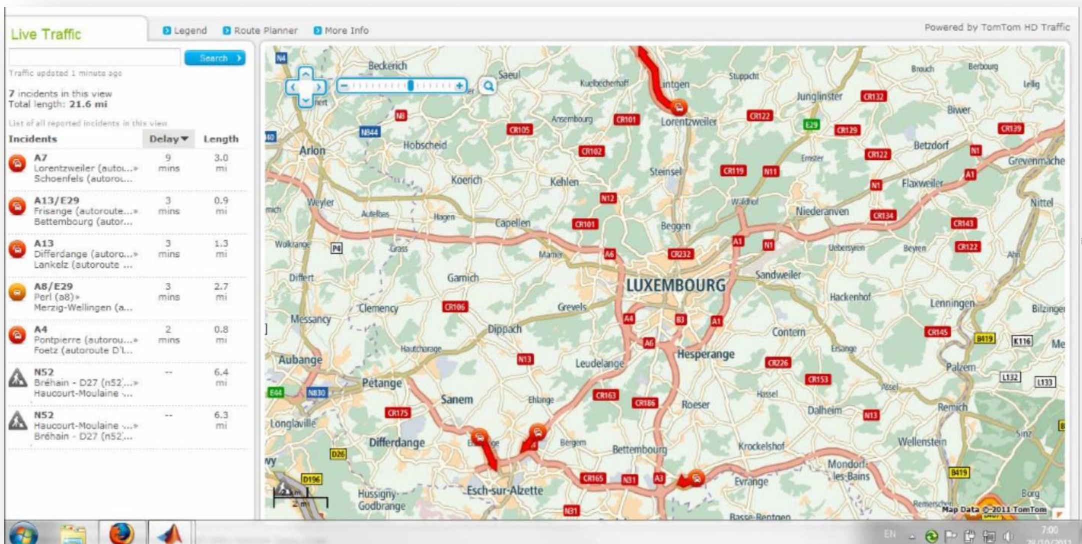
- Disadvantages
  - Scalability, coverage
  - Biases, not necessarily tracing vehicles (e.g., GSM of pedestrians, bikers)

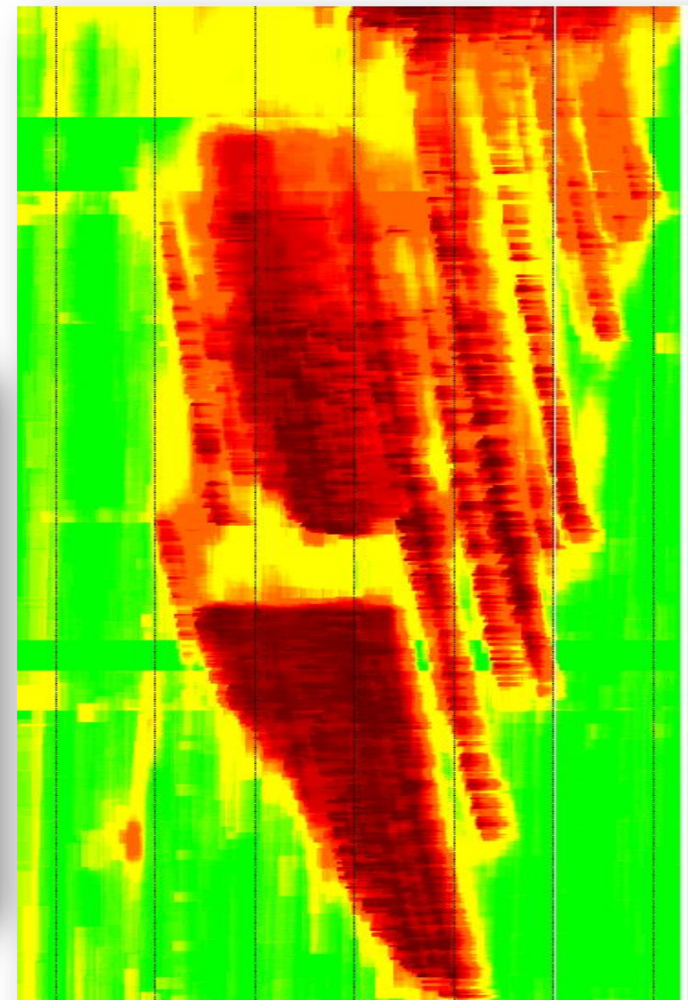
- Typical fleets
  - Taxis, Busses, Commercial vehicles (lorries, trucks,….)
  - Miscellaneous of different types (eg. Be-Mobile’s)



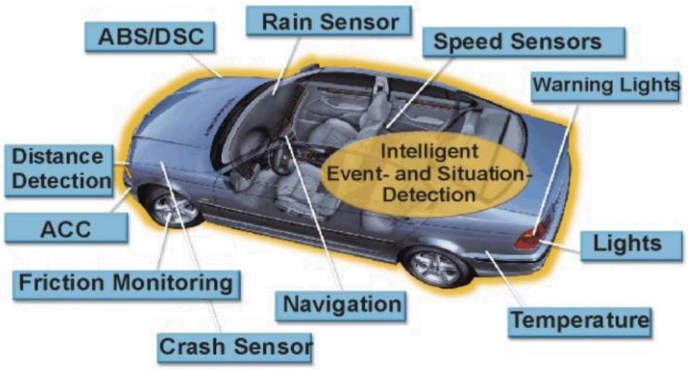
From FCD to traffic info & routing applications

- Accurate space-time plots reproduced from individual trajectories
- Congested sections traced in small time updates





FCD for traffic estimation

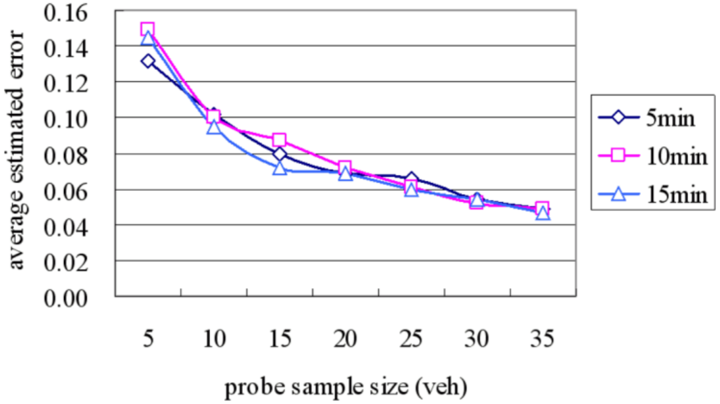


- Floating Car Data
  - By nature trip/route based information
  - Can cover ‘virtually’ all links
- State estimation highly sensitive to sample sizes
- Travel time more representative wrt speed, density, flow,…
- Interpretation issues
  - Low speeds can be interpreted as congestion, parking maneuvers, etc.
  - Tracing activity patterns not possible (e.g., pickup & delivery operations)
  - New generation -> X-FCD

Data coverage issues

- Spatial and temporal coverage:
  - The discrete nature affects the completeness of travel time statistics.
  - Route data might not be available at the time requested because
    - it is insufficient in number or,
    - it does not cover all links or routes in the network, or
    - Data may be available only for parts of the route;
    - Part of the data may be missing (e.g. tunnels).

- Minimum number of probes needed depends on
  - Application (real time info, data analysis, traffic management,…)
  - Aggregation time (1 -> 5 mins)
  - Sampling frequency (10Hz -> 1 min)

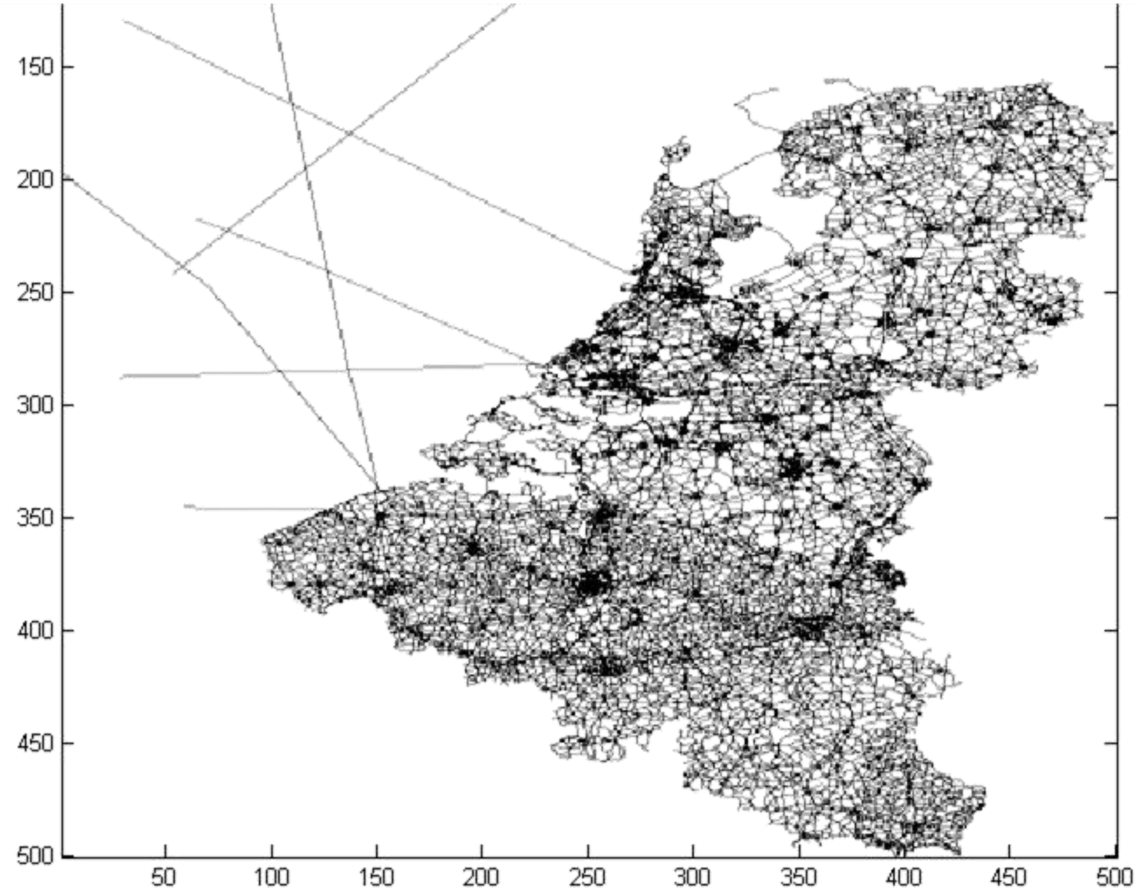


Relationship between link travel time estimation error, aggregation time and sample size (from Jiang et al., 2006)

Data quality issues

- Quality depends on three main aspects:
  - Data acquisition and formatting operations
    - Aggregation
    - Interpolation
    - Conversion to link/path statistics
  - Data completion and smoothing
    - missing data both in time and space,
    - remove (white) noise;
  - Data cleansing
    - remove or correct corrupted or systematic errors
    - identify biases (observation biases, sampling biases, detection lags,…)

MobiRoute coverage



Providing robust info & routing

UNIVERSITÉ DU LUXEMBOURG

FCD positions → estimated link tt → predicted link tt → predict route travel time

- Travel time prediction algorithm based on Floating Car Data
- Data-driven approach for mid-term forecast
- Statistics based on historical data
- Test accuracy of predictions when extracting data by
  - Daily patterns
  - Weekly patterns
  - Seasonal patterns
  - …
- Other metadata included in further improved versions (weather, working zones,…)

Travel time prediction issues

- Combining historical data:
  - How far in the past should we look back to keep a high degree of actuality and to preserve the currently observed traffic patterns?
  - What type of historical data do we need?
    - traffic conditions,
    - time-of-day,
    - day of the week,
    - weather,
    - …
  - How do we deal with historical data correlations and obtain unbiased travel time estimates?
  - Which measures should be adopted?
    - Average, median, average+/-SD,…

Using link travel times for route travel time statistics

Links

Route

TT_1

TT_2

⋮

TT_L

∑ travel times

TT_R

Histogram
Variance
Skewness
Quantiles
...

1 Variance

$$\frac{\sigma^2_{X+Y}}{\sigma^2_X + \sigma^2_Y} = 1 + \frac{2\sigma_X \sigma_Y}{\sigma^2_X + \sigma^2_Y}\rho_{X,Y}$$

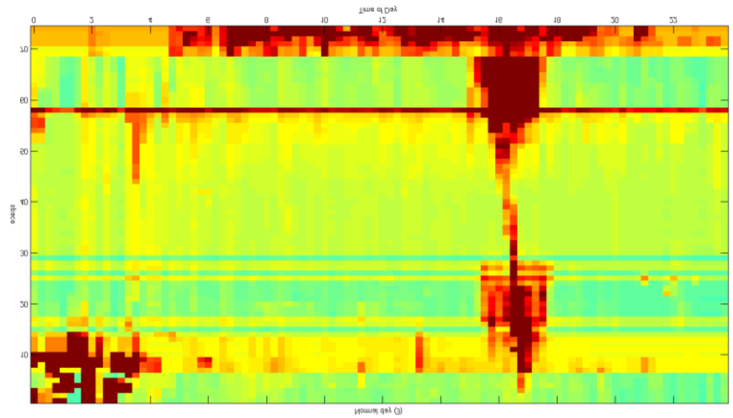
2 Histograms

$$f_{X+Y}(z) = \sum_{k=-\infty}^{\infty} f_X(k)\, f_Y(z-k)$$

X and Y must be statistically independent

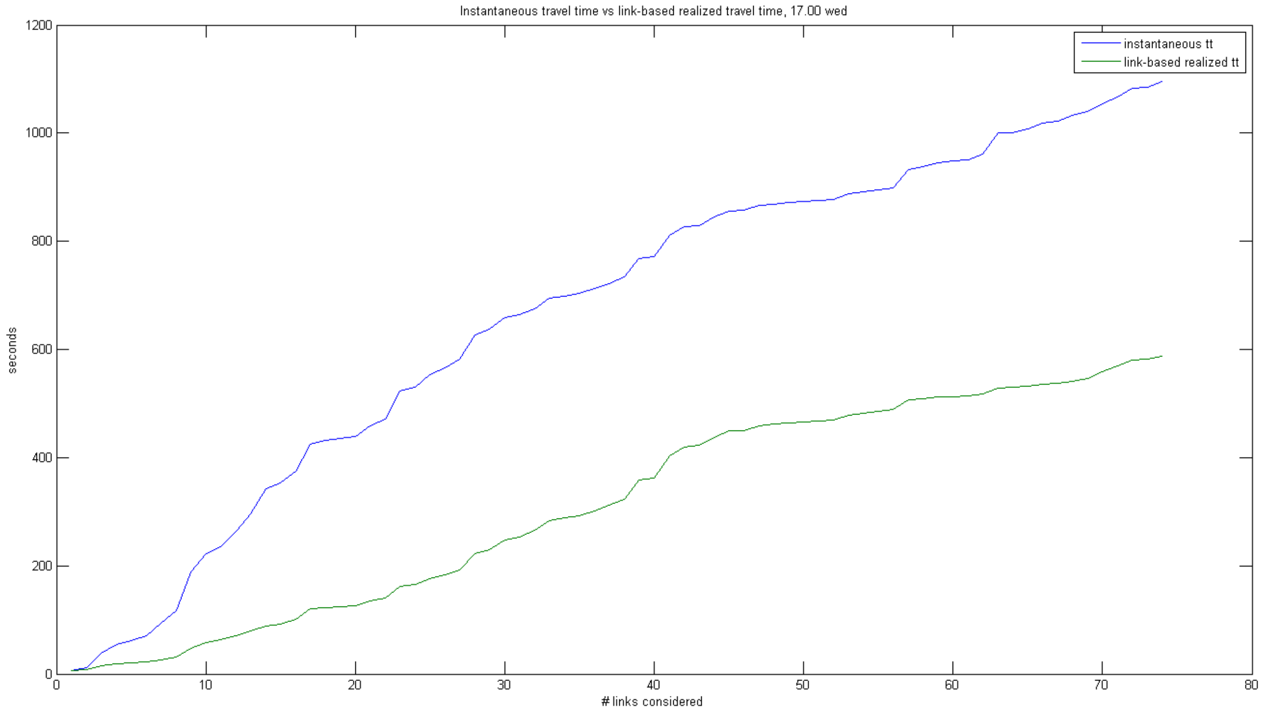
link-based vs route-based predictions

- Spatial and temporal clustering due to link data covariance (delay propagation, spillback, weather,…)
- Route vs. link aggregation
  - Route based distribution represents “reality”
  - Link based distribution neglects covariance in travel time but are easy to calculate and use
- If covariance is fully regarded saved data explodes!



Instantaneous vs. Realized travel time

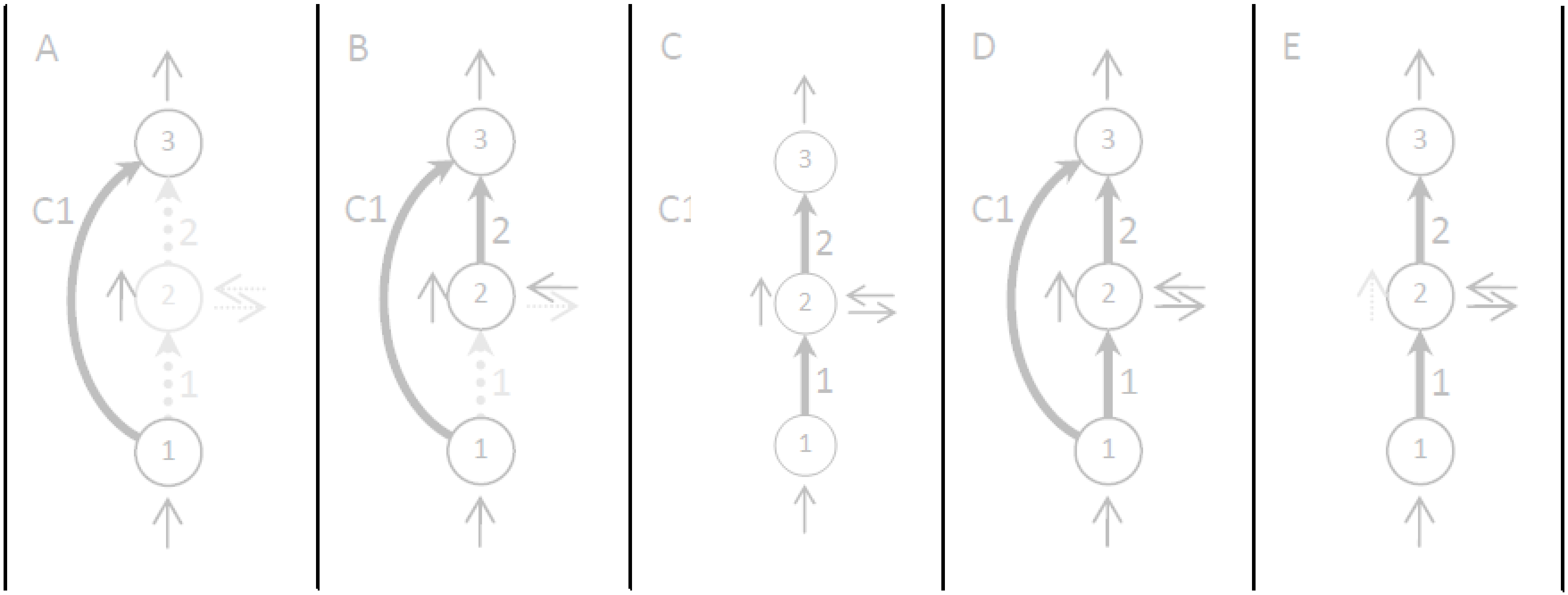
- Link-based instantaneous travel time vs. route-based estimated predicted travel time;
  - Instantaneous travel time ok during off-peak,
  - For congested routes/times realized travel time deviates significantly
  - Solution → use different percentiles for predictions



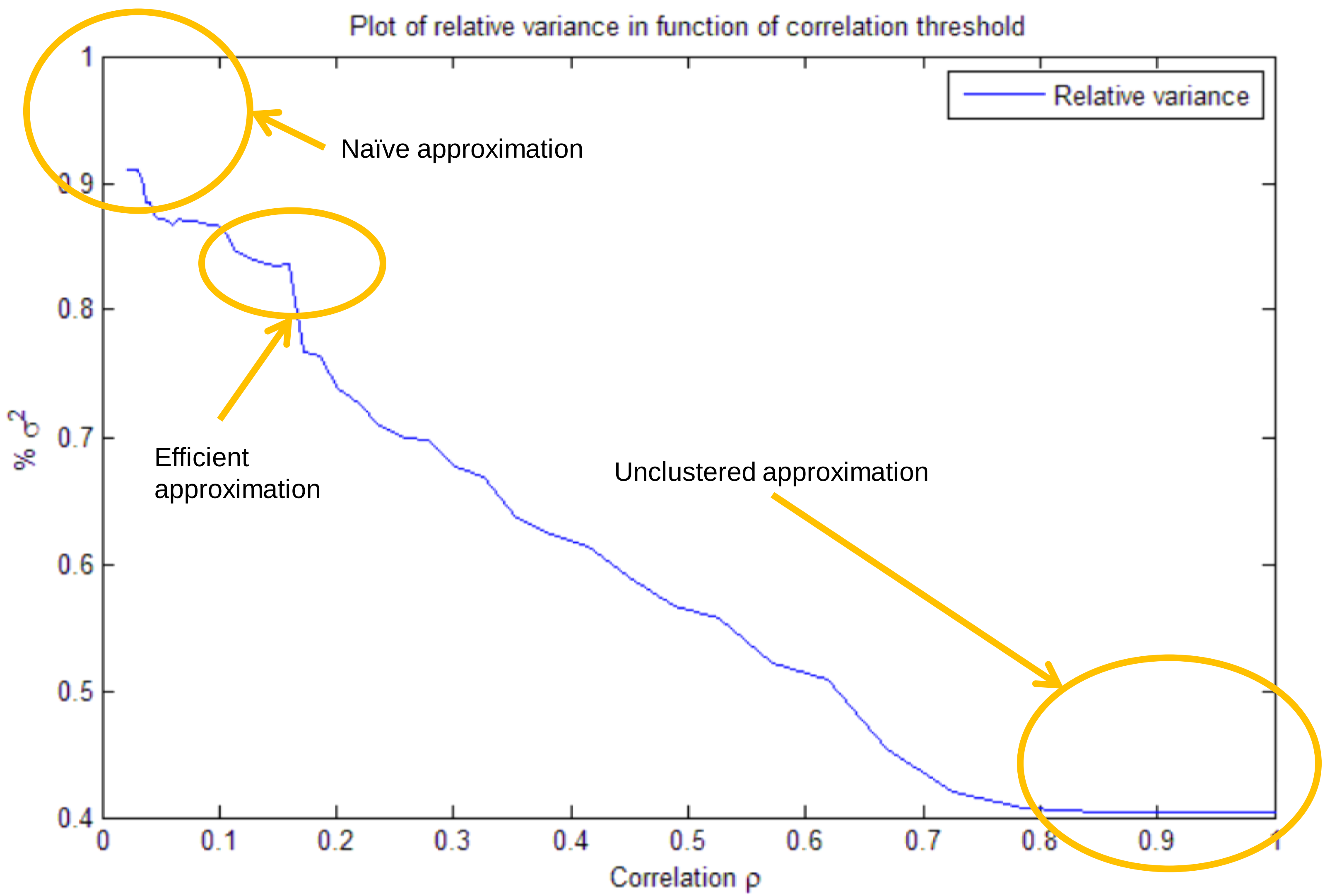
Spatial Clustering approach

Redefine links in the network
- Based on correlations
- Based on node function

Network containing cluster-links and possibly fewer links

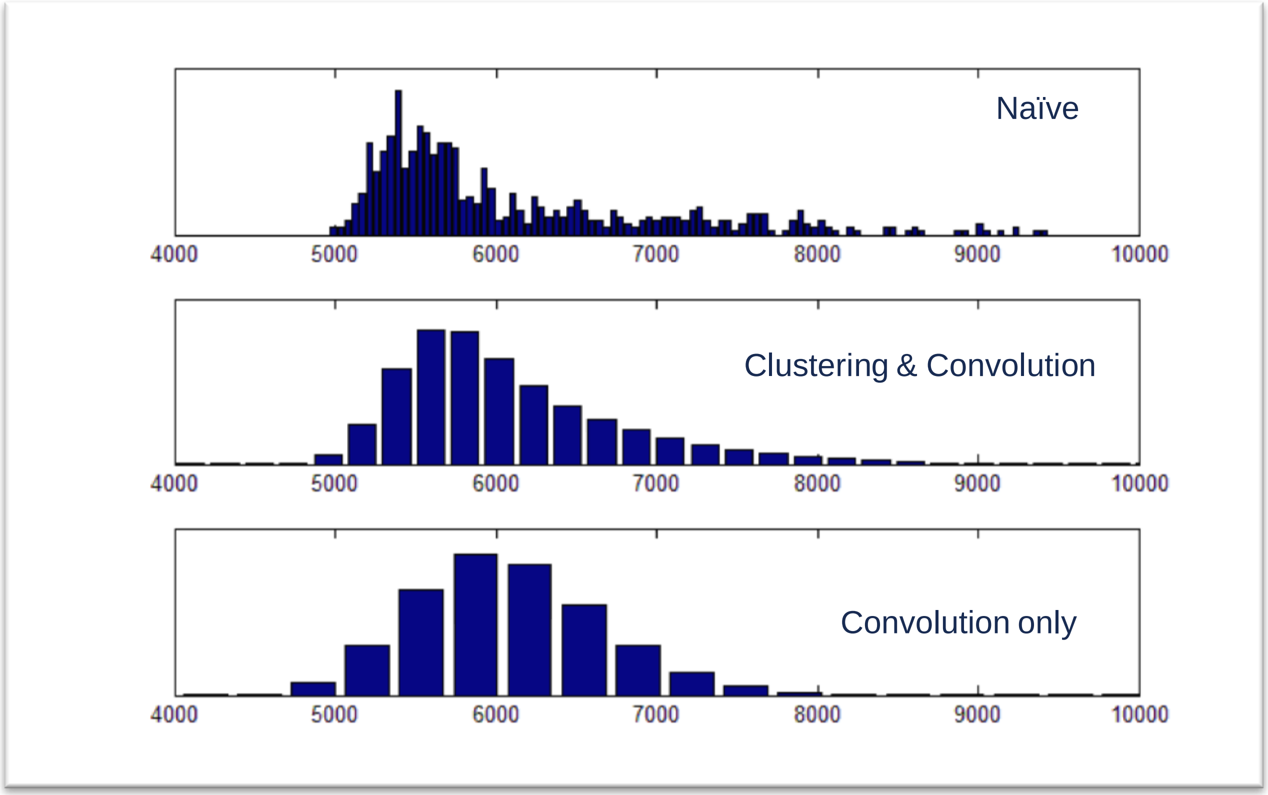


Route advice: Accuracy

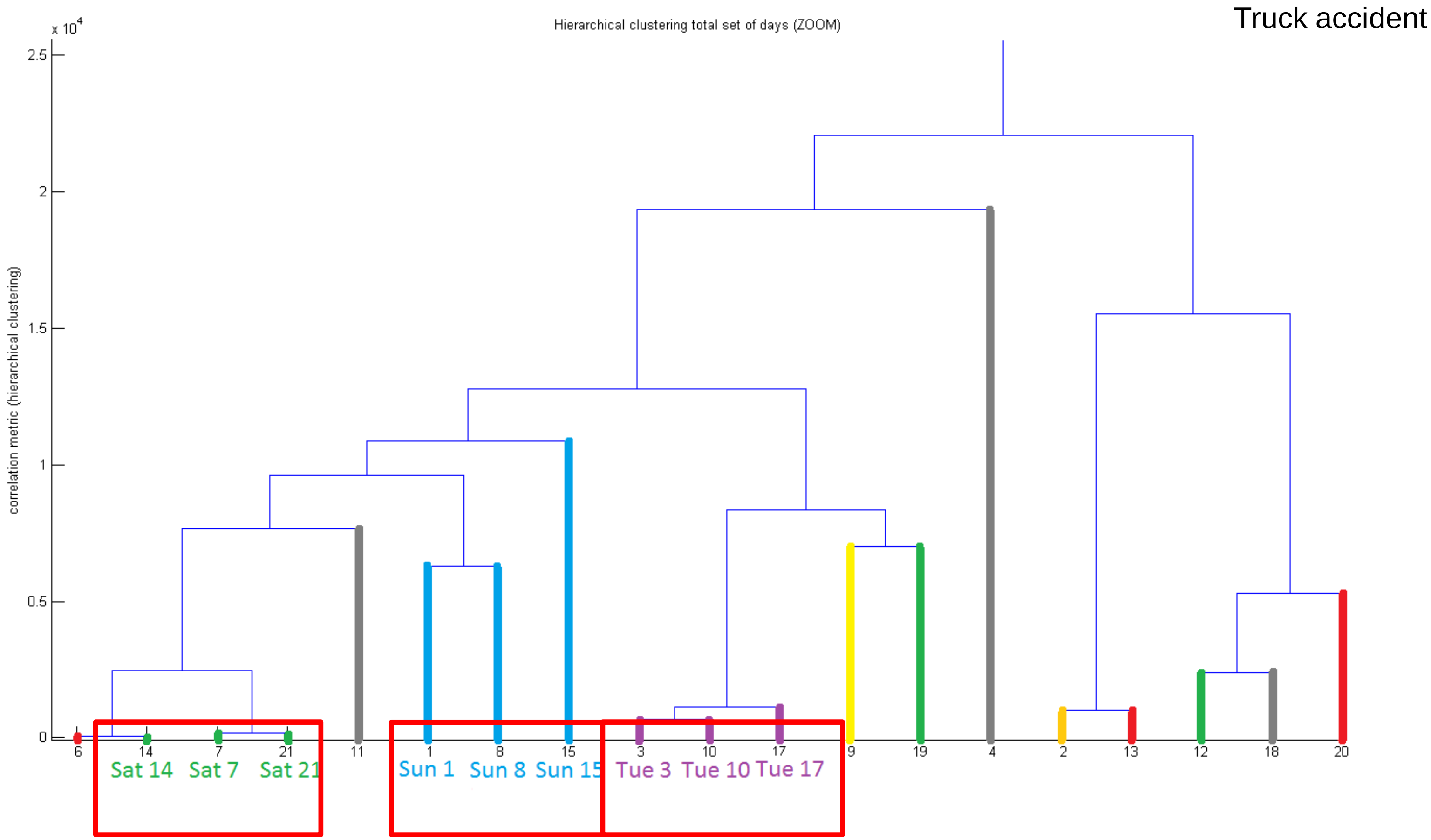


Spatial Clustering

- Route travel time distributions calculated as convolution of link travel time distributions → impact of link TT correlations

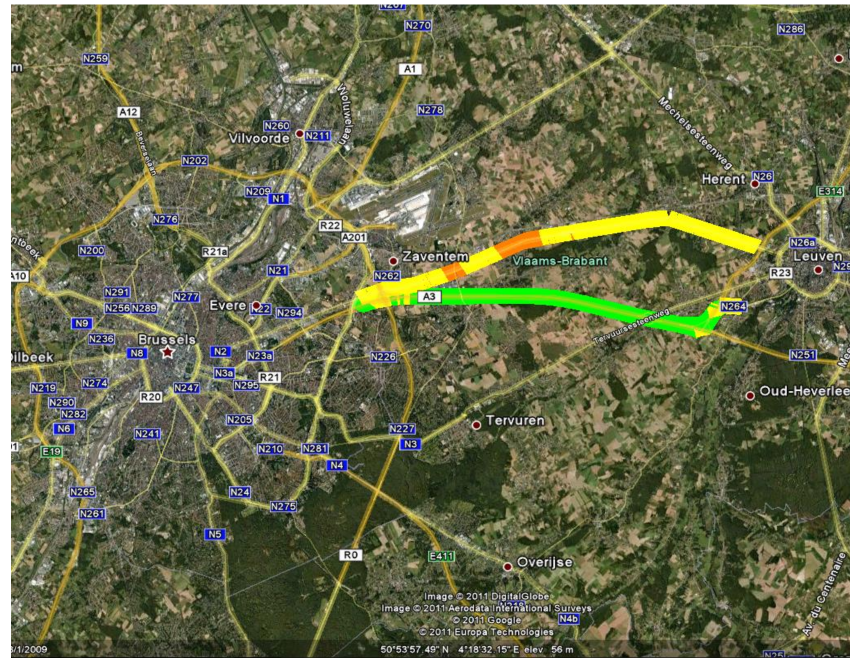


Temporal Clustering – Hierarchical approach



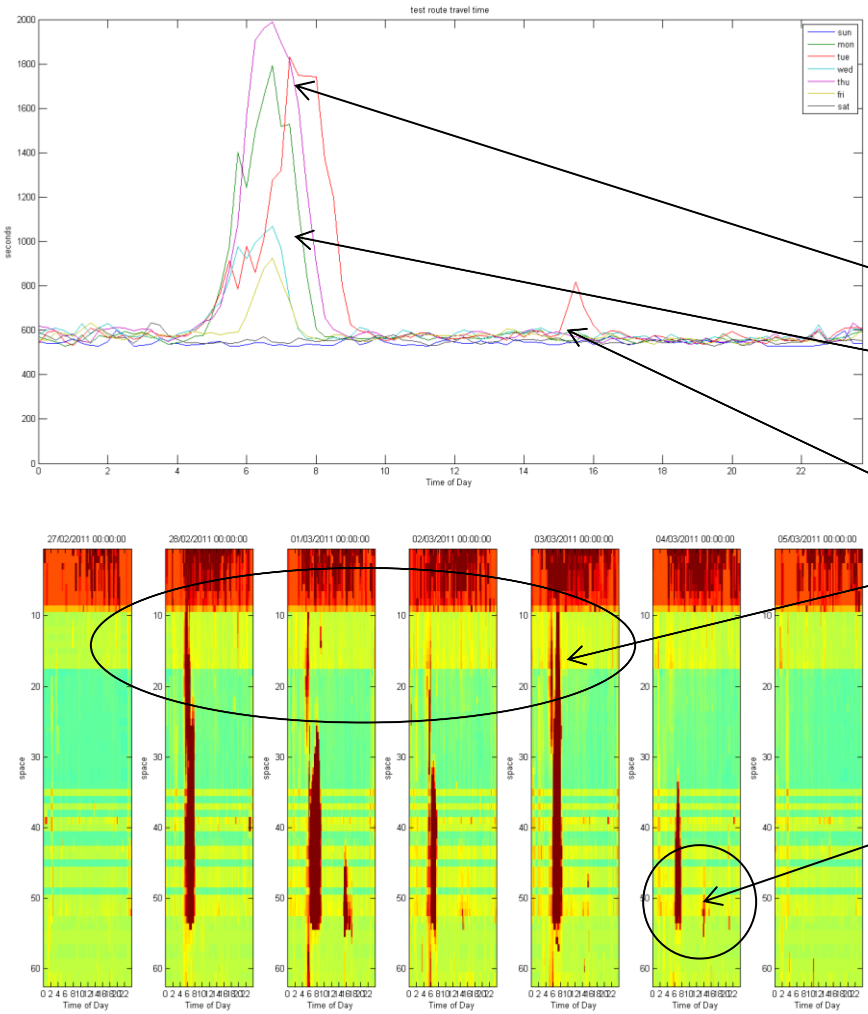
Case study

- Test on 3 routes:
  1. Brussels-Leuven via E40-E314; evening peak
  2. Leuven-Brussels via E40-E314; morning peak
  3. Brussels-Leuven via the Leuvensesteenweg; traffic lights, shops, ...



- Motorway route ~16 km, 75 links, FF travel time ~10 min
- Leuvensesteenweg ~16 km, 90 links, FF travel time ~20 min.

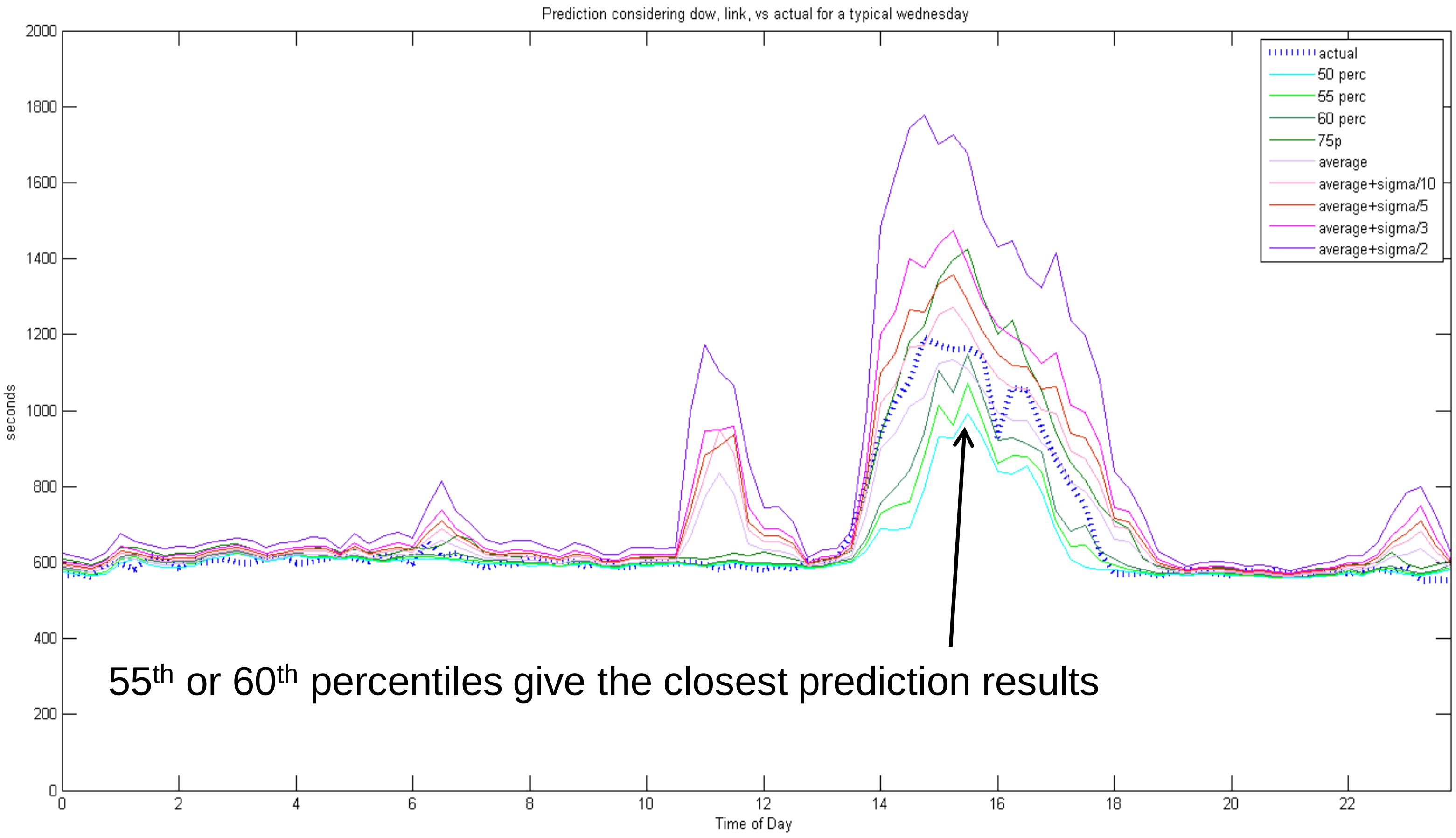
Leuven-Brussels



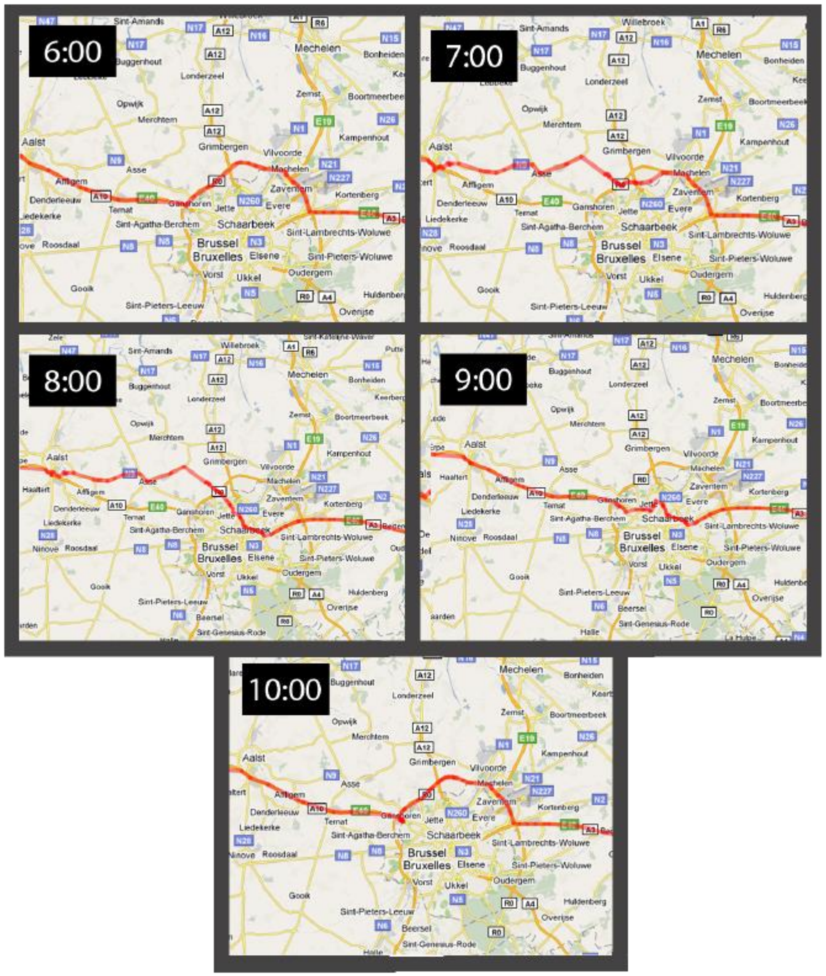
Bottleneck not always active

Congestion spillback

Prediction difference



Dynamic stochastic routing application



Robust routing example

- Comparing 6 routes between Leuven and Brugge:

95%	2	3	6	5	4	1
90%	3	2	6	4	5	1
75%	3	6	4	2	5	1
50%	4	6	3	2	5	1
25%	4	6	3	2	5	1
10%	4	6	3	2	5	1
5%	4	6	3	2	5	1

Summary & Recommendations

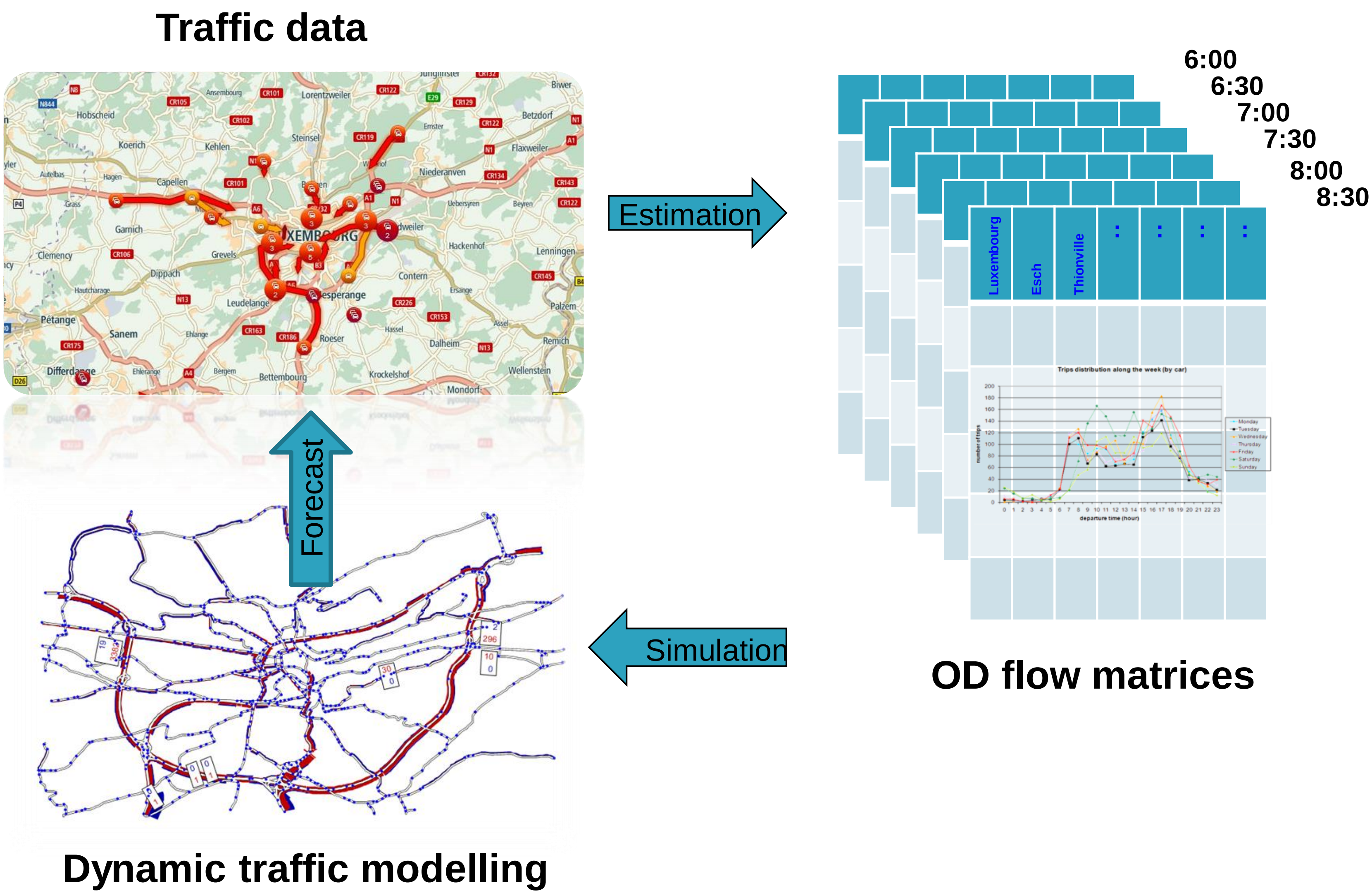
UNIVERSITÉ DU LUXEMBOURG

- FCD has great potentials for information and routing applications;
  - Flexible
  - Cheaper and cheaper
  - Higher and higher coverage
- MobiRoute: Mobility and Routing project
  - Prediction method proposed based on historical data
    - Spatial correlation through link clustering
    - Temporal correlation using hierarchical clustering
- Better predictions using percentiles wrt average-based approaches

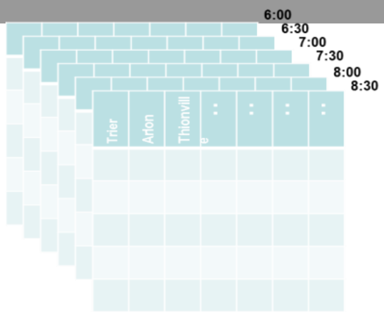
2. Speed and travel time profiles and distributions from mobile sensors

For Demand Estimation

Dynamic demand modeling



The dynamic demand estimation problem



Model

Data

OD MATRIX DEMAND

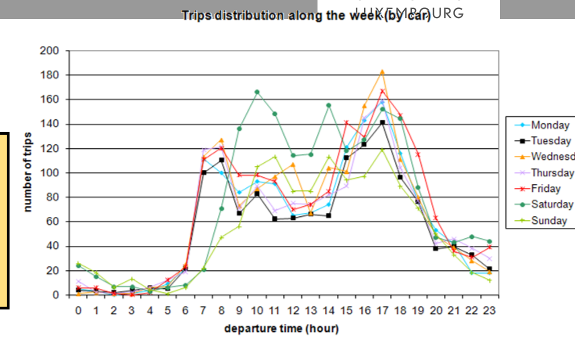
TRUE OD FLOWS

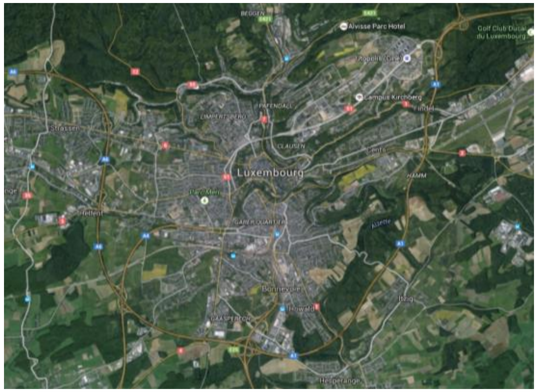
NETWORK MODEL

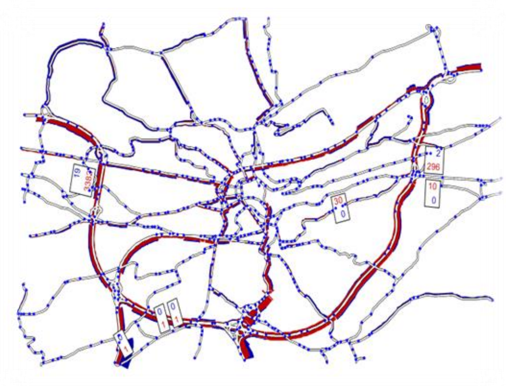
ACTUAL SUPPLY

TRAFFIC STATES

TRAFFIC DATA







History of OD estimation approaches

Planning (static)

- Mobility surveys, 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)

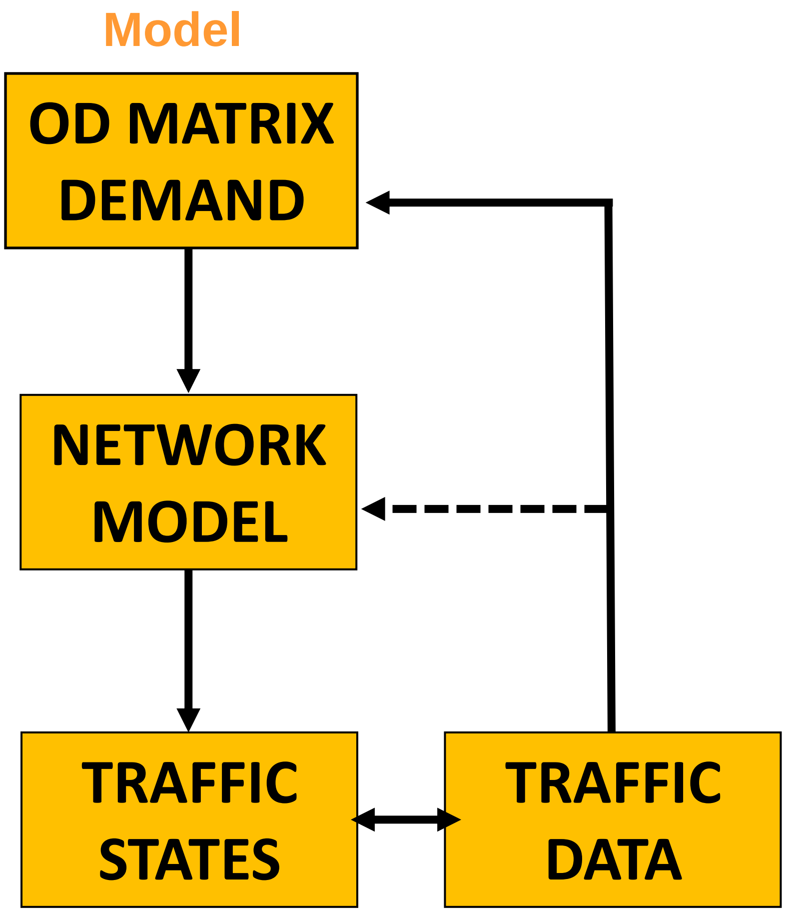
Management (dynamic, offline)

- 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 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 good overview



The OD estimation problem formulation

Goal: find most likely OD matrices that best reproduce the data

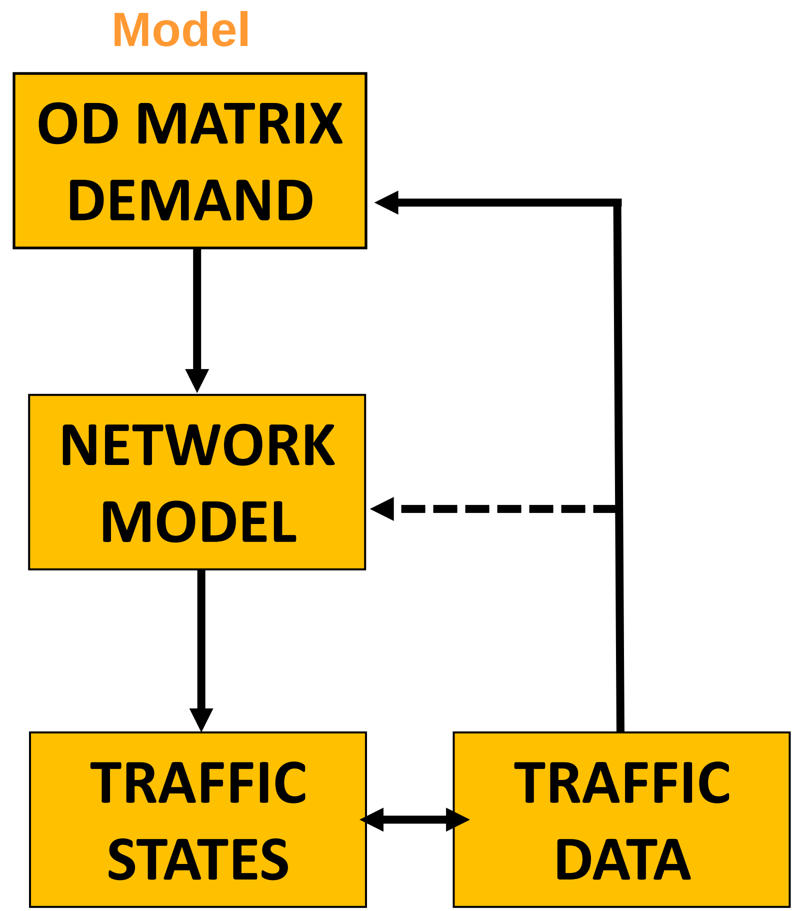
- Highly combinatorial & non-linear problem
- 'Smart' combination of demand and traffic information necessary
- Traffic model should be sufficiently accurate

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

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

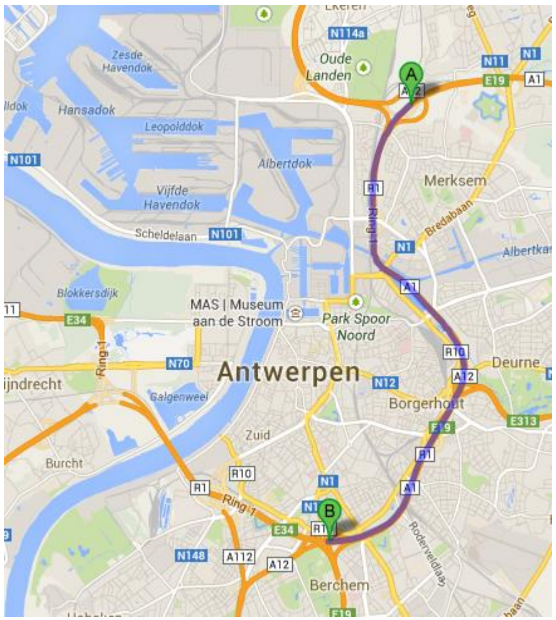
Traffic propagation functions

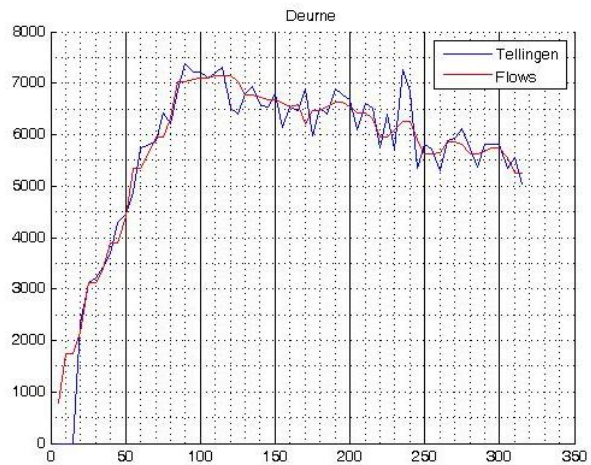
Route choice functions



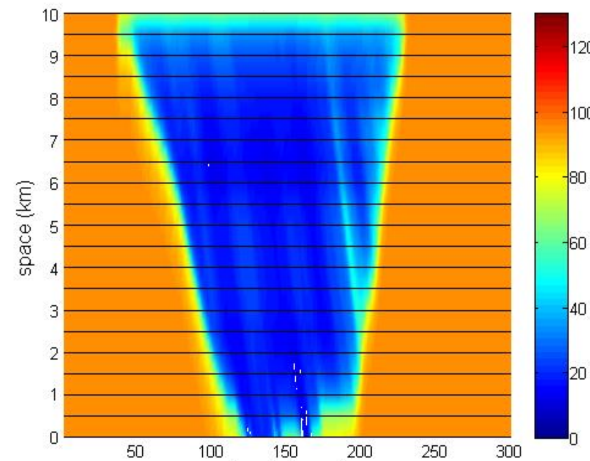
A simple example

The ambiguity of traffic data: supply or demand information?

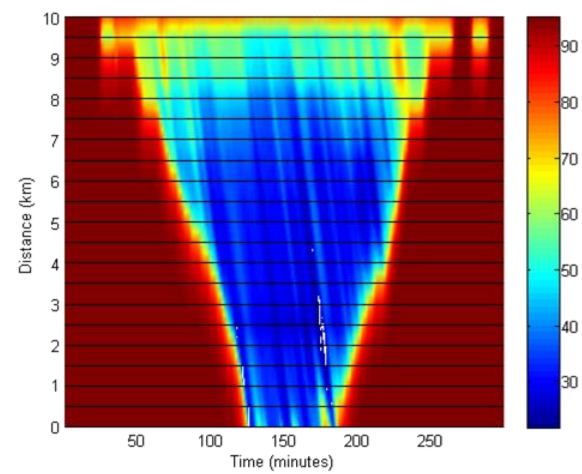


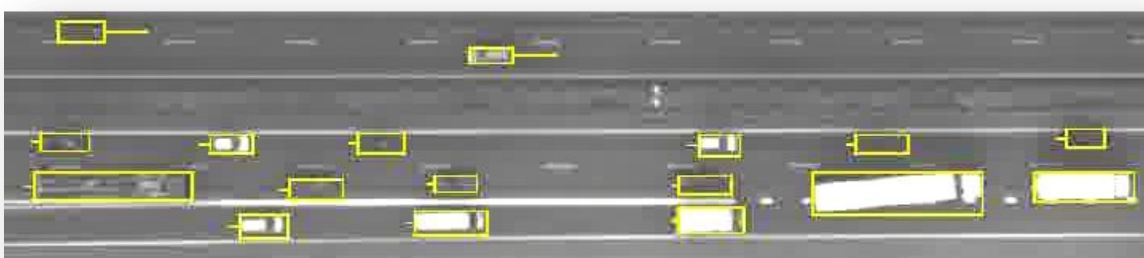


Measured speeds



Estimated speeds





Acknowledgments: Rodric Frederix (KU Leuven)

Using floating car data for dynamic demand modeling

- Analysis of route choice models
- Including path information from floating car data in demand estimation

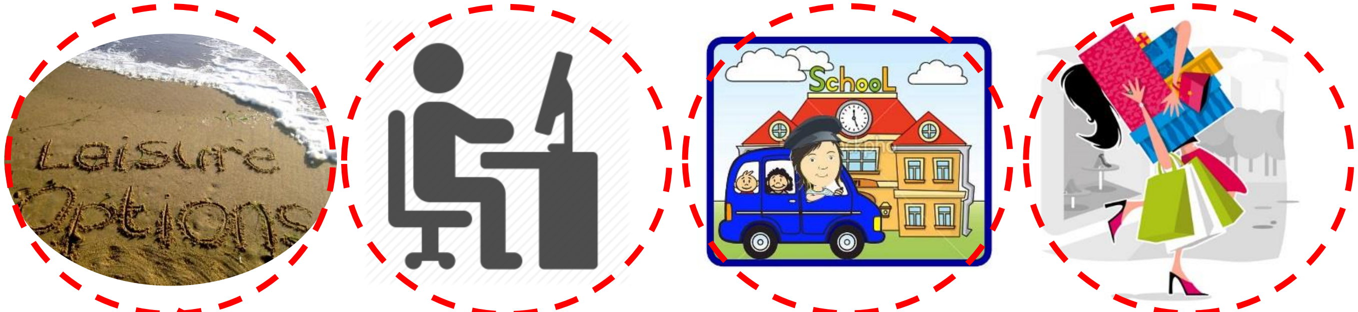
Analysis of route choices using GPS information

- Contribution: (real) shortest path and observed path;
  - Discrepancy in term of overlapping;
  - Discrepancy in term of travel time;

- Innovative elements:
  - Influence of the reliability
  - Average velocities obtained with low-frequency GPS coordinates
  - Congested network

Data Set and Methodology (1)

- **Low-frequency GPS coordinates*:**
  - **89 drivers**
  - **September 2010 – 31 January 2012 (17 months)**
  - **More than 52.458 observed paths (Monday-Friday)**



Systematic Trips

✓ 119 Clusters
✓ 13.766 paths

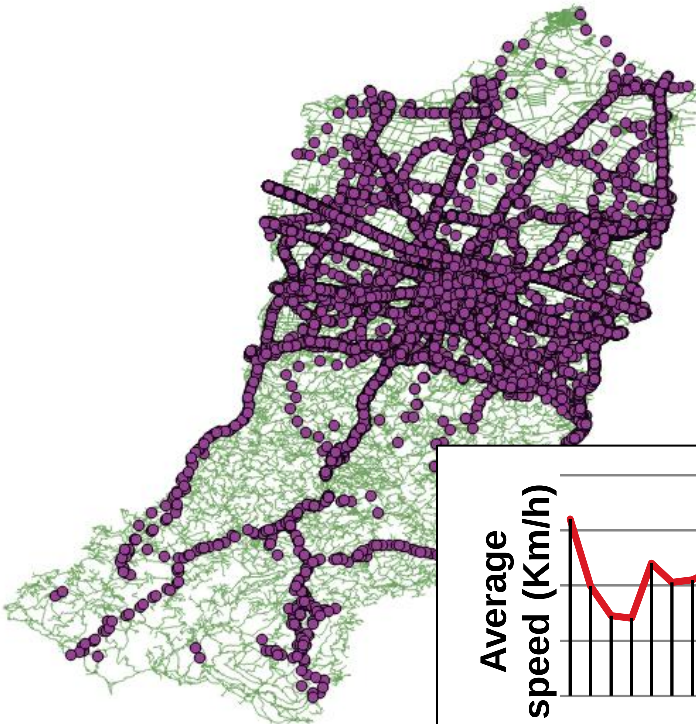
- **Clustering technique:**
  - **Single linkage method**
  - **Euclidean distance as dissimilarity measurement during the clustering**
  - **Cophonetic correlation to identify outliers in the clusters**

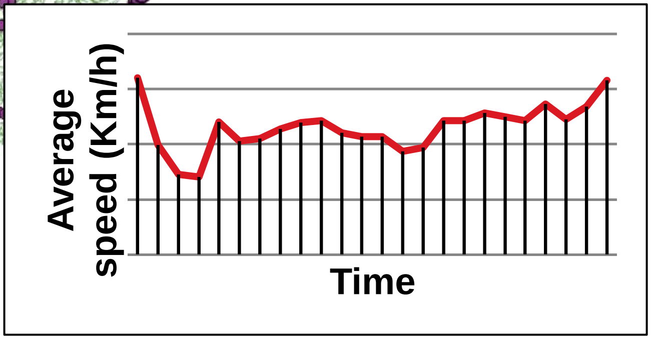
Data Set and Methodology (2)

- **Average velocities:**

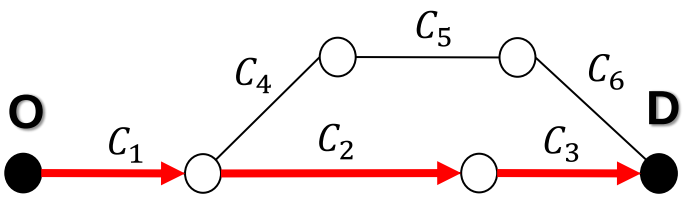
$$C_i = \frac{\sum_{j=1}^{N} V_j^i}{N}$$

  - $j \in N$ is the observed path
  - i is the link id
  - V_j^i is the speed for the j–th observation on the i-th link





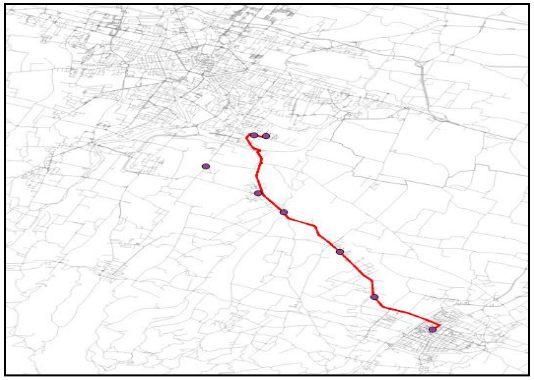
- **Average velocities:**
  - A* Shortest path Algorithm



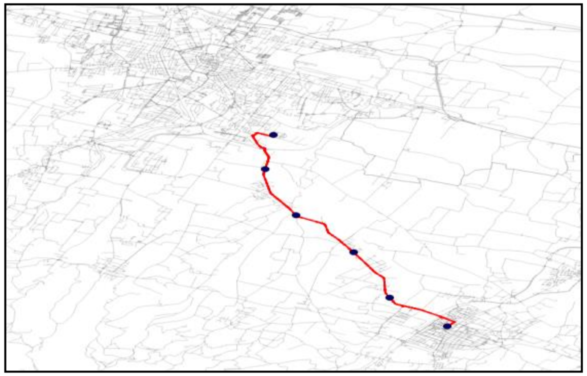
Data Set and Methodology (3)

1) Overlapping

- Shortest path has been represented as a polyline
- Overlapping percentage: the number of GPS coordinates which interpolate the shortest path



75% overlapping



100% overlapping

2) Travel time: *Normalized Average Travel Time*

$$NATT_i = \frac{ATT}{STT} = \left(\frac{\sum_{i=1}^{Np} TT}{N_p}\right) \cdot \frac{1}{STT_i}$$

≈ 1

$\approx STT_i$

- ATT= Average Travel Time
- STT= Shortest path travel time
- I = User

Results (1)

UNIVERSITÉ DU LUXEMBOURG

1) Overlapping: 13.766 observed paths/shortest one

26.62%

51.71%

Overlap	*Percentage*
100%	15.07%
90-99%	1.46%
80-89%	9.62%
70-79%	9.57%
60-69%	11.10%
50-59%	4.89%
40-49%	12.17%
30-39%	13.63%
20-29%	11.11%
10-19%	4.52%
0-9%	0.03%

Results reported in literature:

40% of the observations overlap the shortest paths (≥90%)

[4] [5]

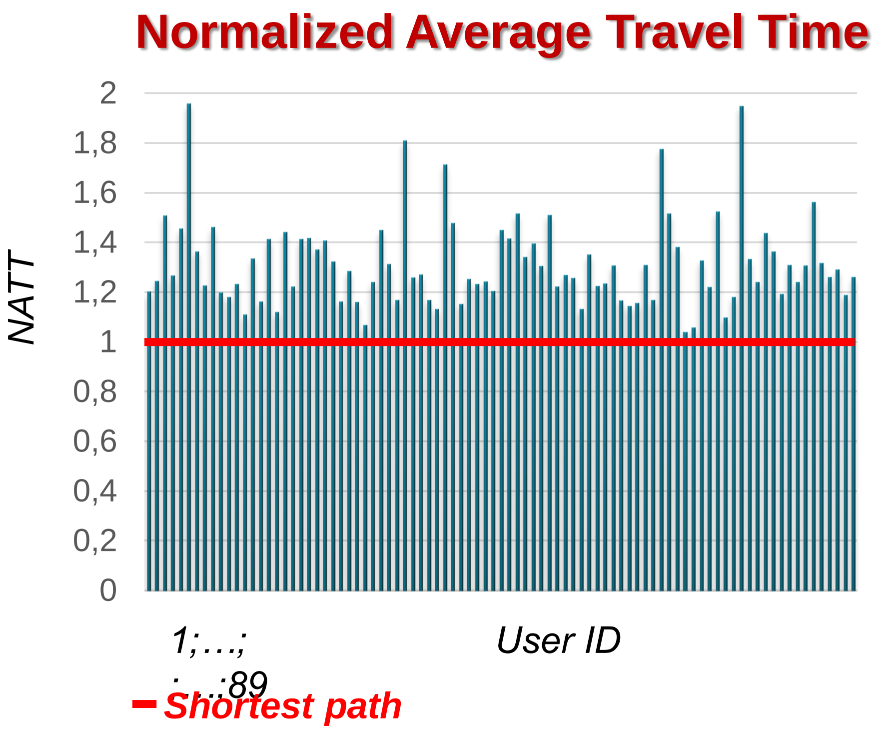
Differences:

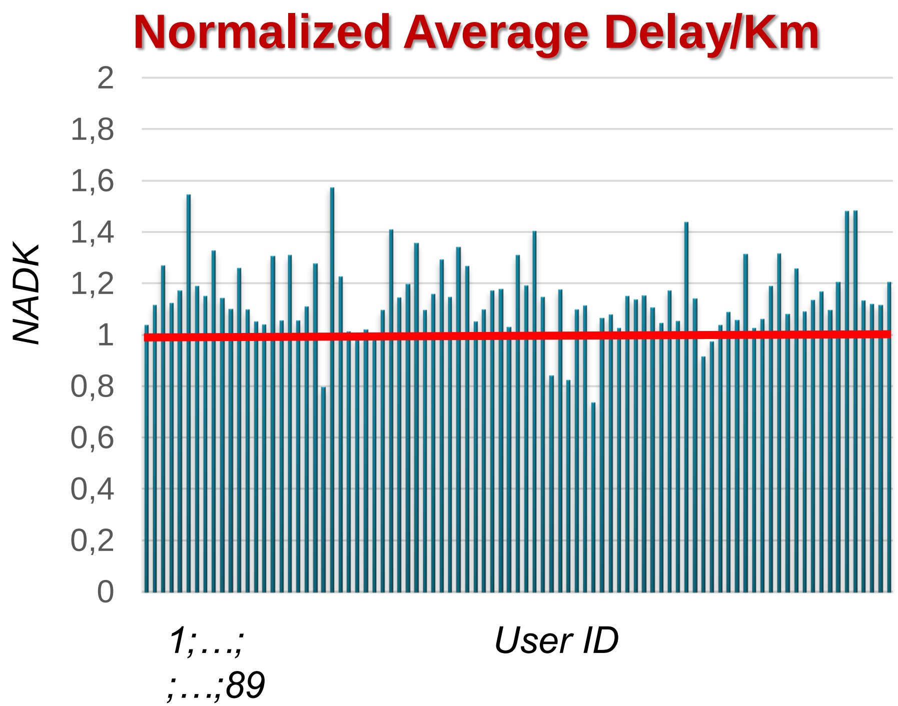
1. **Shortest path computed using the real-actual speed;**
2. **Congested Network;**

Since measured speeds are used, exist at least one path which presents a lower travel time with respect to the observed one, for the specific time interval !!!

Results (2)

2) Travel Time Discrepancy:





- ***On average people have the tendency to use routes 1.3 times longer;***
- ***On average people have more delay with respect to the shortest path (1.15 times longer);***

Results (3)

3) Reliability: *lateness reliability factor*

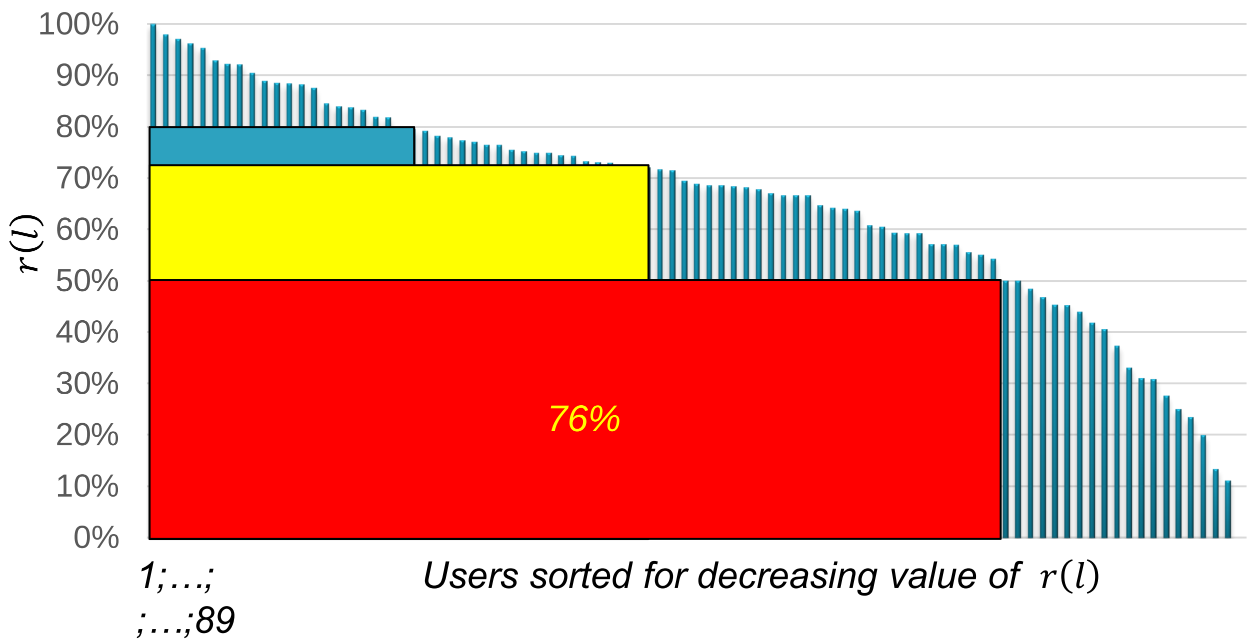
$$r(l) = \exp\left[\frac{1}{2} \cdot T_{log}(l) - z_{\alpha/2} \cdot \sqrt{T_{log}(l)}\right]$$

l = route

T_{log}= variation logarithm – day to day variance in travel time

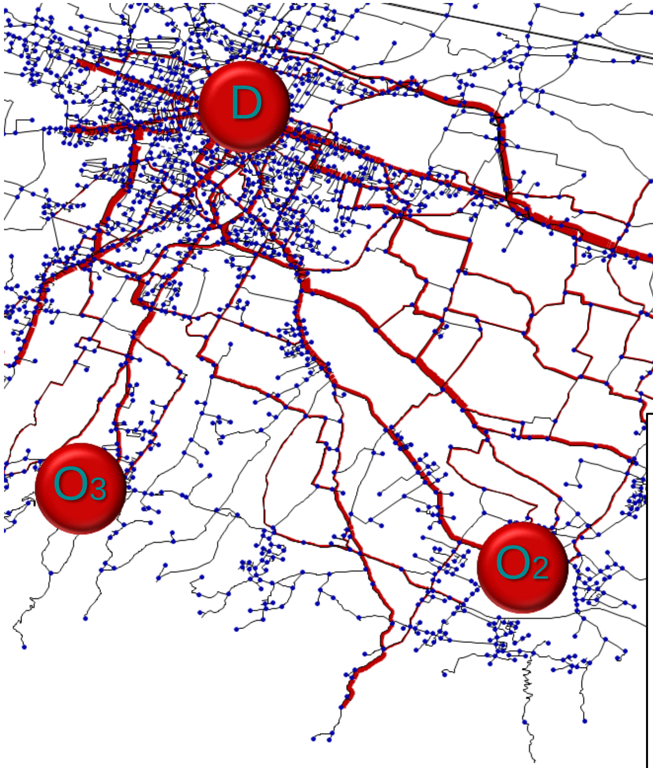
$z_{\alpha/2}$= standard normal distribution tail

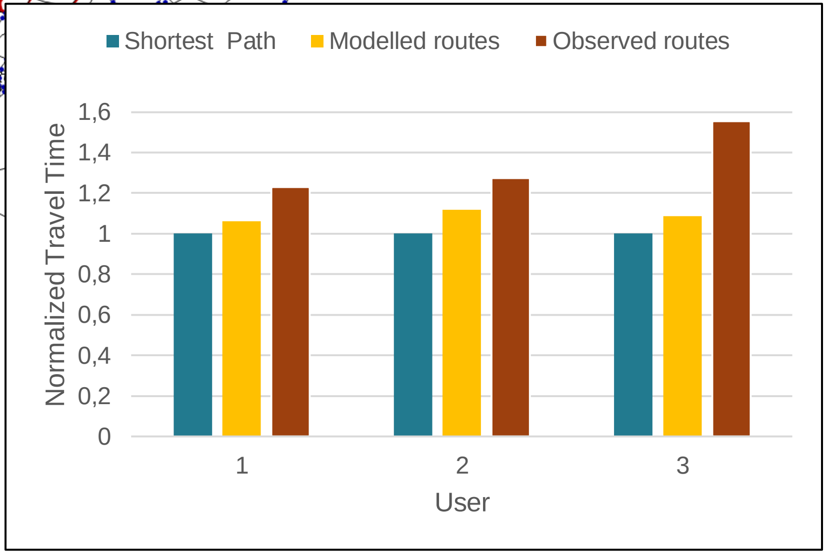
Probability to use the most reliable route for each user:



Results (4)

4) An illustrative example:

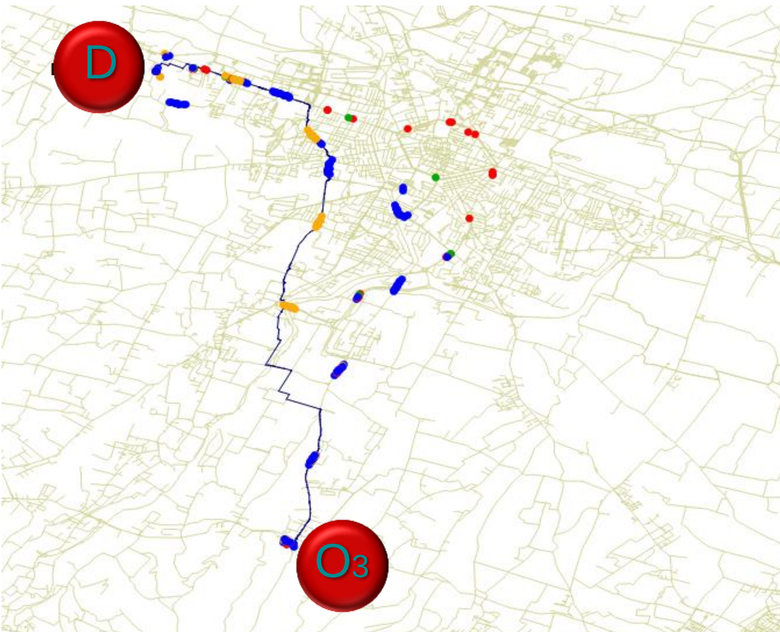


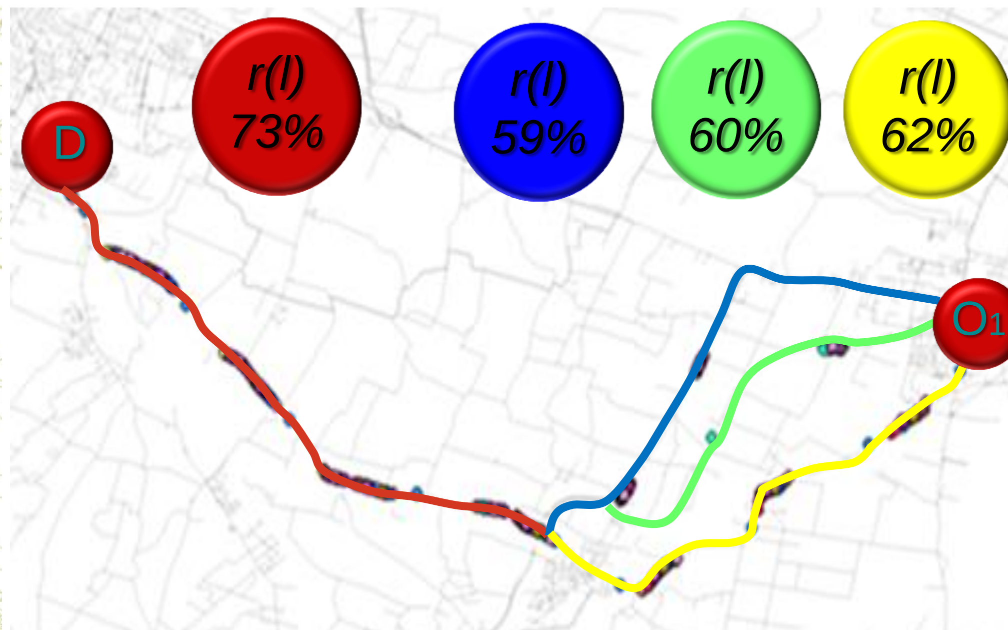


- ***Sub-Network:***
  - *18632 Links, 7455 Nodes*
- ***Realistic traffic conditions:***
  - *RMSE Simulated and observed speeds < 6%*
  - *Simulated and observed shortest path are the same*
- ***Behavior of 3 user is analyzed:***
  - *Only morning peak*
  - *320 observed paths*

Results (5)

Examples of discrepancy between best/modelled and observed alternatives:





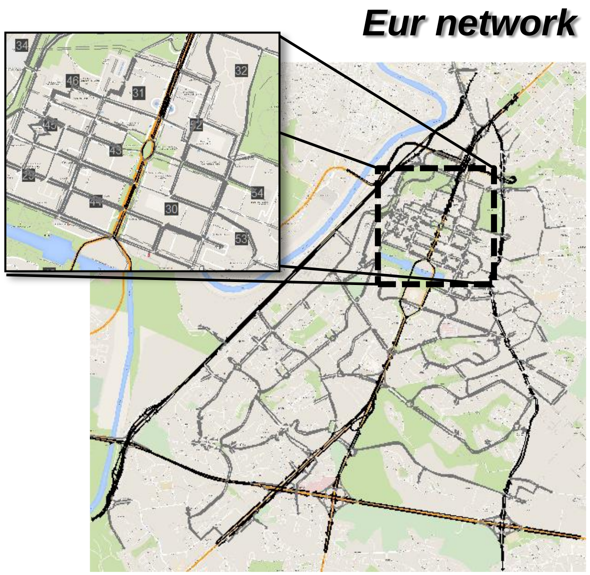
- ***User 3 prefers a longer path, driving around the city center rather than a direct route.***
- ***User 1: The three routes overlap where the reliability is higher***

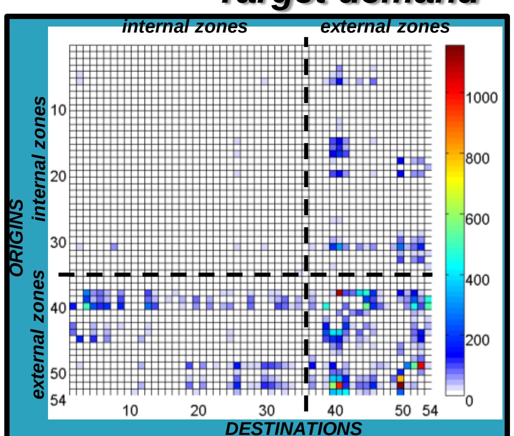
Conclusions

- Do people really use the shortest (time) path?
- Are Wardrop's principles a realistic approximation?
  - On average, +30% travel time with respect to the shortest path
  - On average, +15% delay/km with respect to the shortest path

- Route reliability:
  - Is a relevant aspect in user's route choice
- Observed paths are not similar to the shortest one (i.e. direct one)

Using floating car data for dynamic demand estimation (1)





Total travel demand: 45,000 veh/h

- **Network model**
  400 nodes;
  812 links;
  54 traffic zones.
- Resolution algorithm: SPSA AD-PI
- **Demand**
  Real demand of 45,000 veh/h equally distributed in four time slices;
  Seed demand: 67,800 veh/h
- **Simulated measurements**
  32 count section
  12 monitored ODs

- Additional to OD formulation a term referring to route choice measures: OD's travel time
- Two scenarios have been adopted to perform Dynamic OD estimation

Misures\Set	Set 1	Set 2
OD target	+	+
Links Flow	+	+
OD travel time		+

Error on measures reproduction

Indicator used: percentage average error

Link flows

OD travel time

Error on demand reproduction

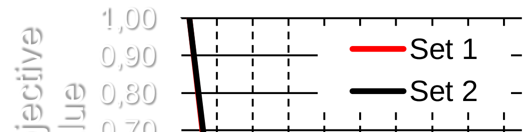
Space features

Temporal features

Indicator used: Euclidean distance

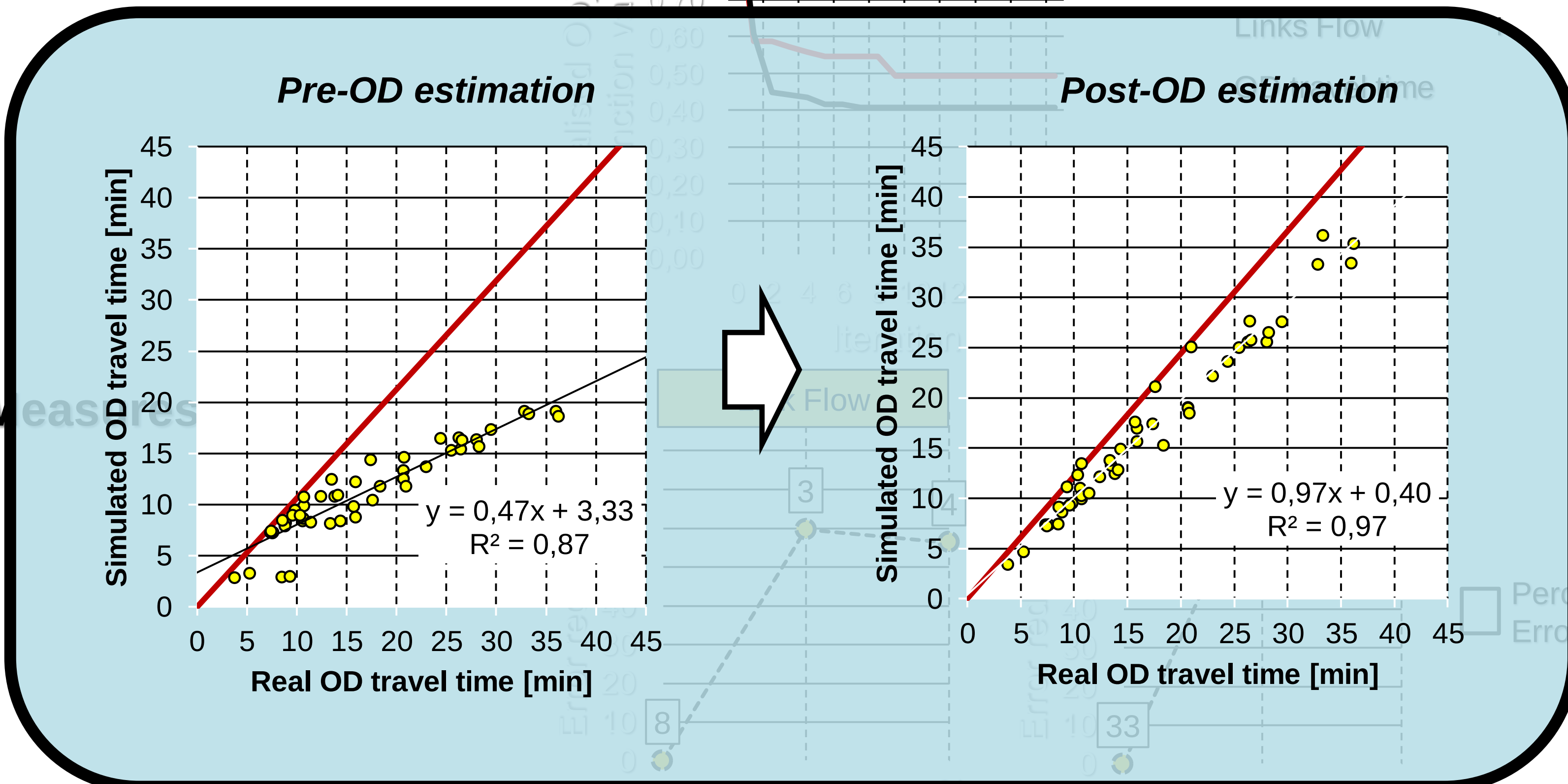
Using floating car data for dynamic demand estimation (2)

- **Objective function**



Misures\Set	Set 1	Set 2
OD target	+	+
Links Flow		+
OD travel time		+

- **Measures**

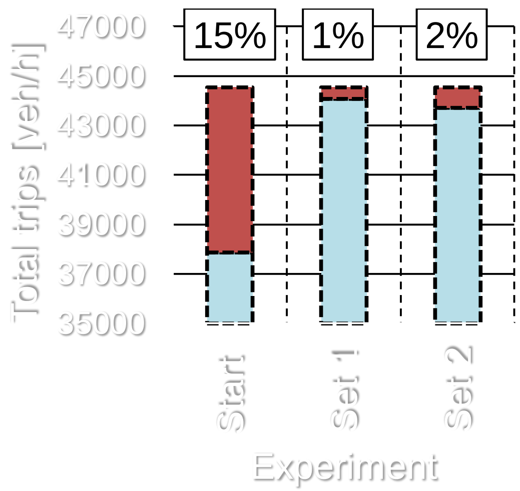


Experiment

Experiment

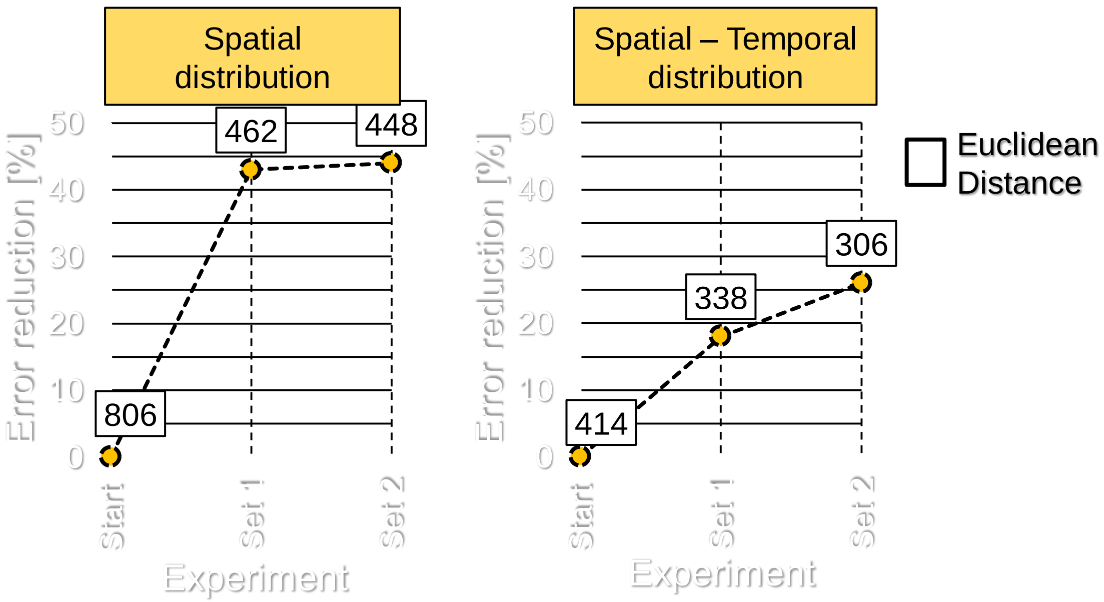
Using floating car data for dynamic demand estimation (3)

- **Total demand**



Misures\Set	Set 1	Set 2
OD target	+	+
Links Flow	+	+
OD travel time		+

- **Demand reproduction**



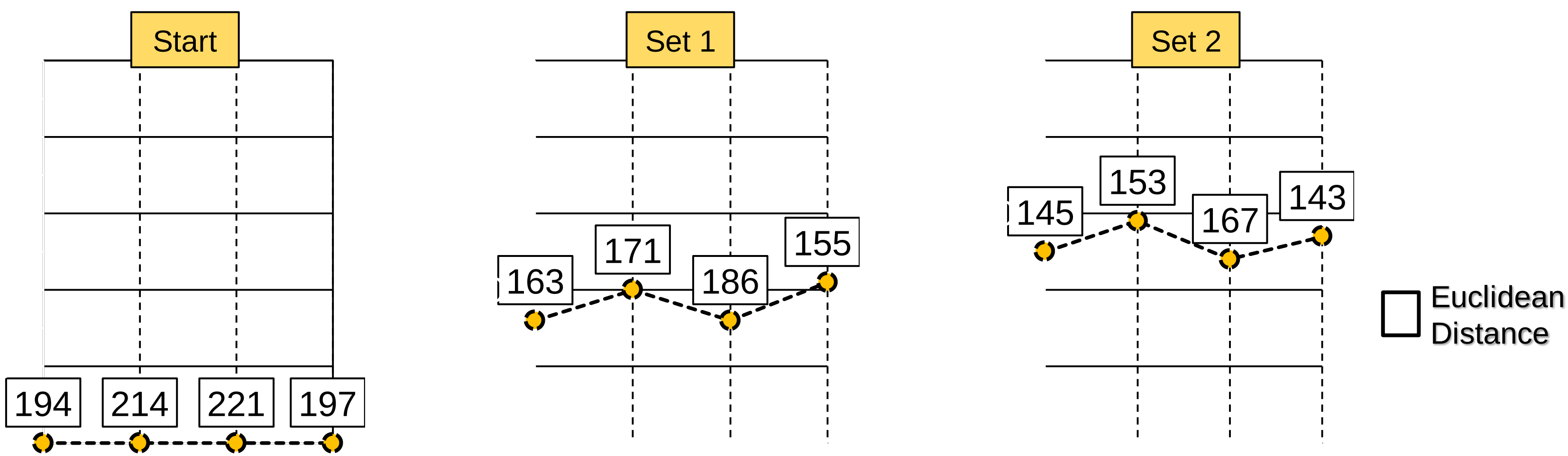
Using floating car data for dynamic demand estimation (3)

Misures\Set	Set 1	Set 2
OD target	+	+
Links Flow	+	+
OD travel time		+

- **Improvement on estimation and correlation of adopted information**

Error – Demand Reproduction	Set 1	Set 2	only OD travel time
Intercepted ODs [%]	67	67	10
Euclidean distance reduction [%] (monitored Ods)	-19	-25	-28
Euclidean distance reduction [%] (not monitored Ods)	-16	-32	-13

- **Distribution for each time interval**



Conclusions

- Floating car data used to improve demand estimation
- Inconsistency of modelled and actual route choices amplifies error in the estimation
- Adding path information helps at finding more reliable results in real sized networks

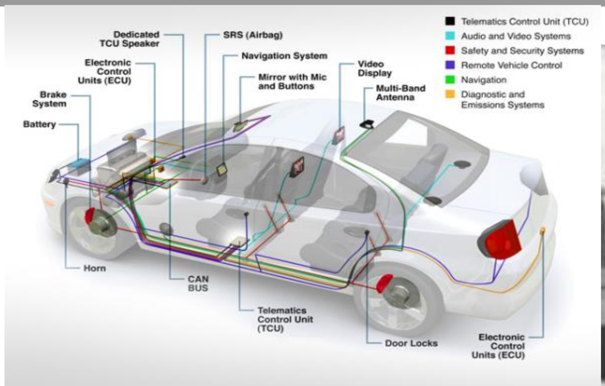
3. Speed and travel time profiles and distributions from mobile sensors

For Mobility Analysis

Mobility analysis

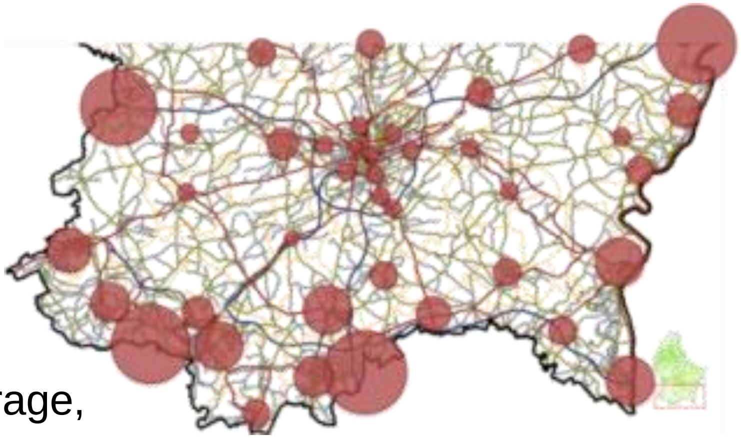
- Activity-travel behavior dynamics
- Travel demand management and transport policy
- Multimodal transportation modelling
- ICT for travel planning and advisory systems

New and more advanced mobile sensors



Model challenges and where data helps (1)

The potentials of (Big) data



New opportunities, old problems

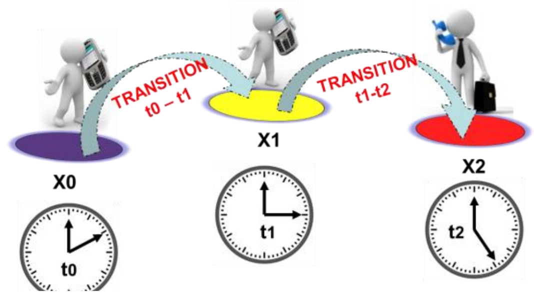
- Data → multiple solutions
- **Big Data → plethora of solutions!**

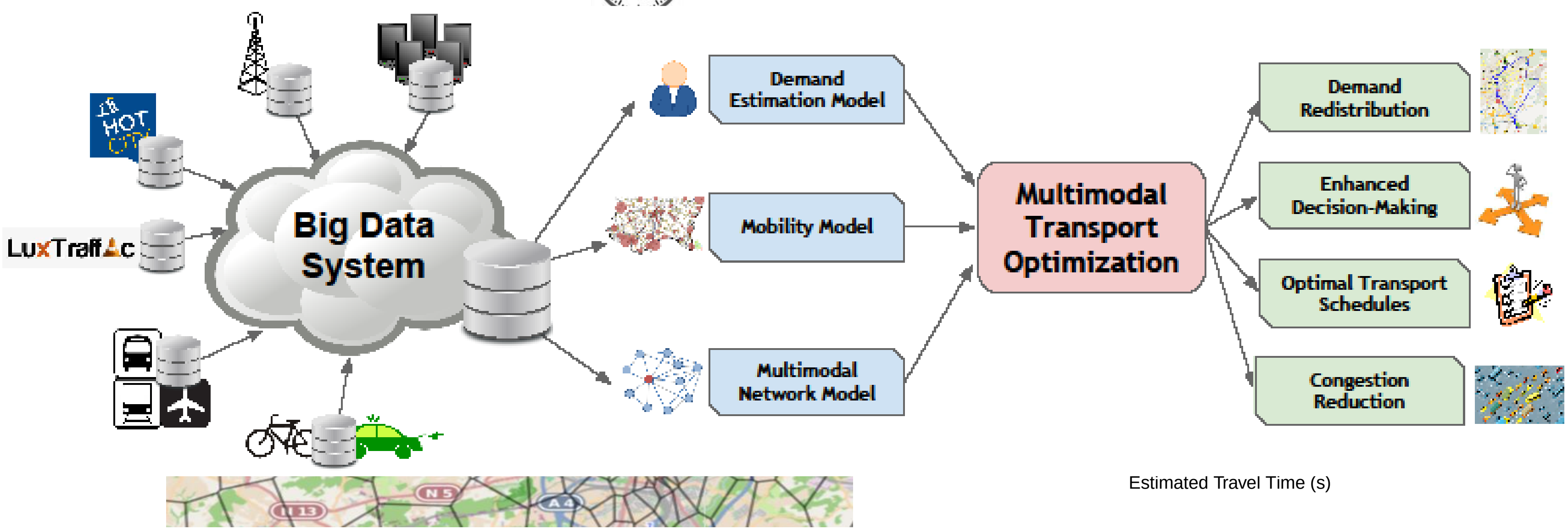
Traffic counts → ambiguity of flows
Mobile sensors→ ambiguity of flows, modes, coverage, biased users, discontinuous in time and space…

Model challenges and where data helps (2)

UNIVERSITÉ DU LUXEMBOURG

- New location-based datasets
  - GSM data
  - WiFi connections
  - Smartphone data



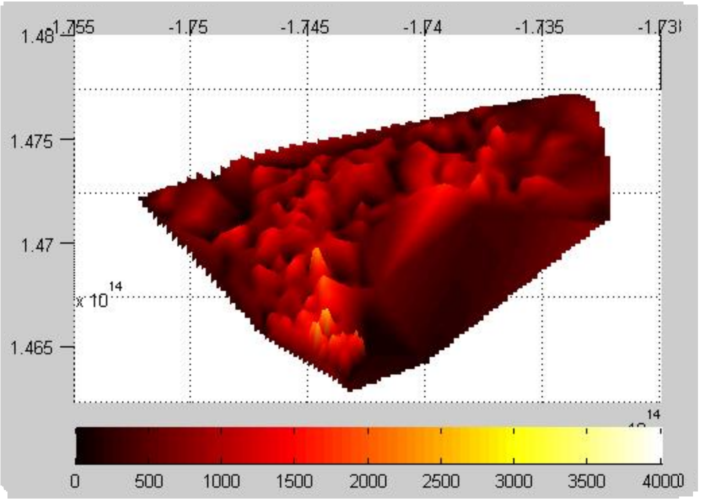


Estimated Travel Time (s)

Acknowledgments: Raphael Frank & Thierry Derrman (SnT)

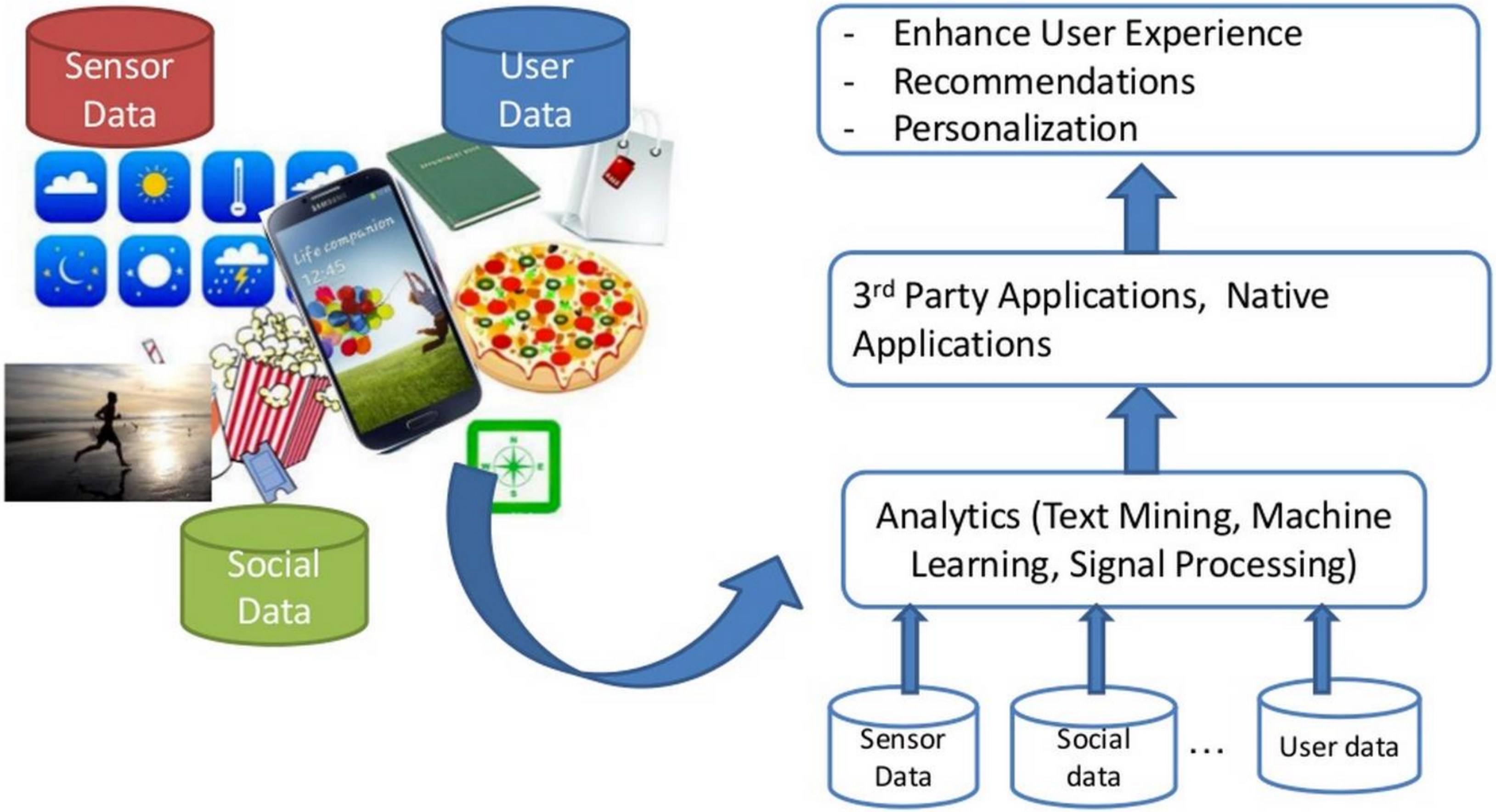
Big Data approach

- Research opportunities
  - Multimodal modeling
  - Demand estimation
  - Travel Assistance systems



The new frontier of mobility analysis: Big data analytics

Collecting personal mobile sensor data: opportunities for decision support services



GO$_2$UNI platform: website and mobile application

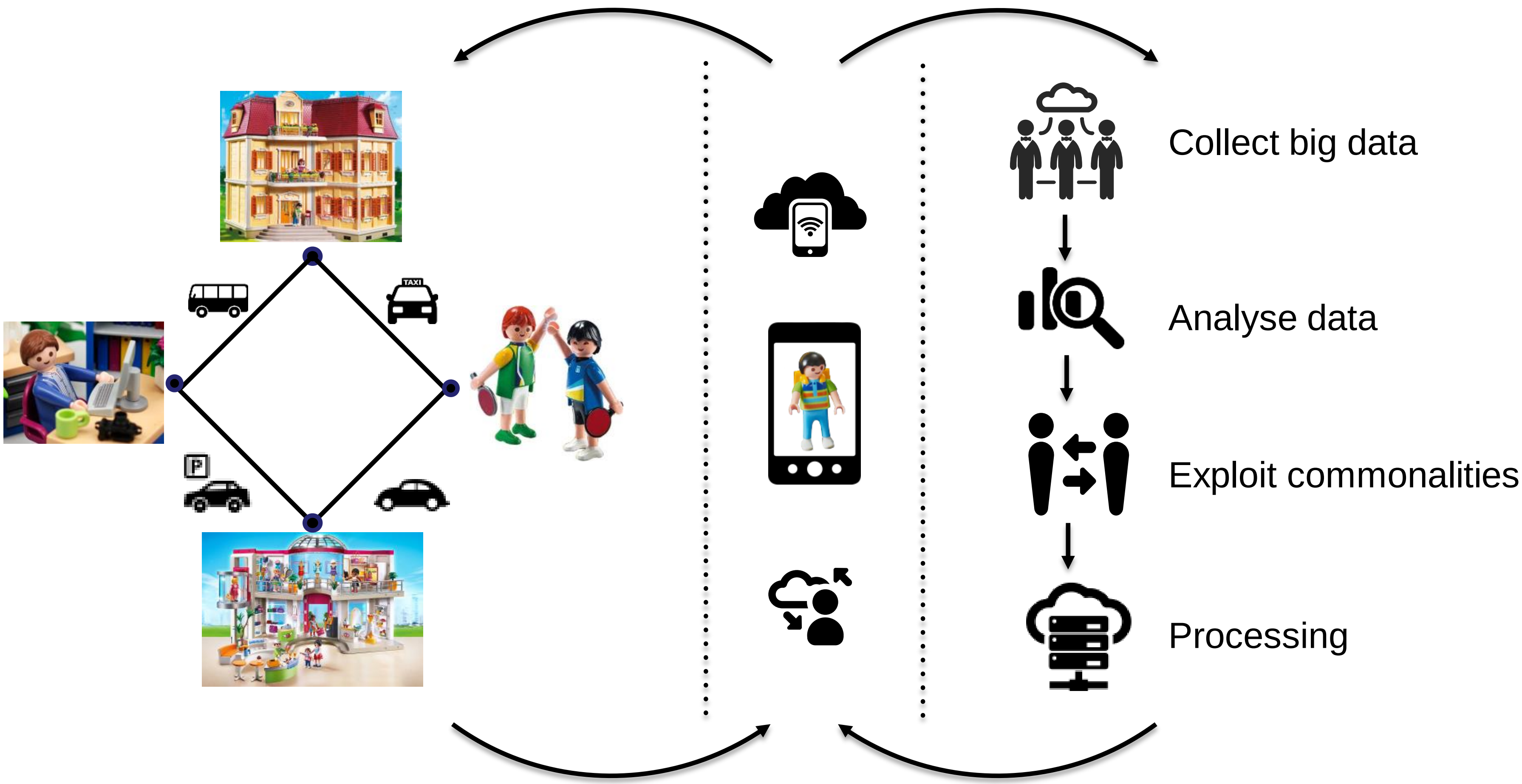
- Carpooling
- Car-sharing
- Intercampus bus shuttle
- Public transport
- Real time information
- Traffic status
- Parking management
- Recommendations

Activity-travel data collection

Closing the loop: user needs and mobility habits fed into transport service optimisation

1. Detect activity-travel choices

2. Provide advice



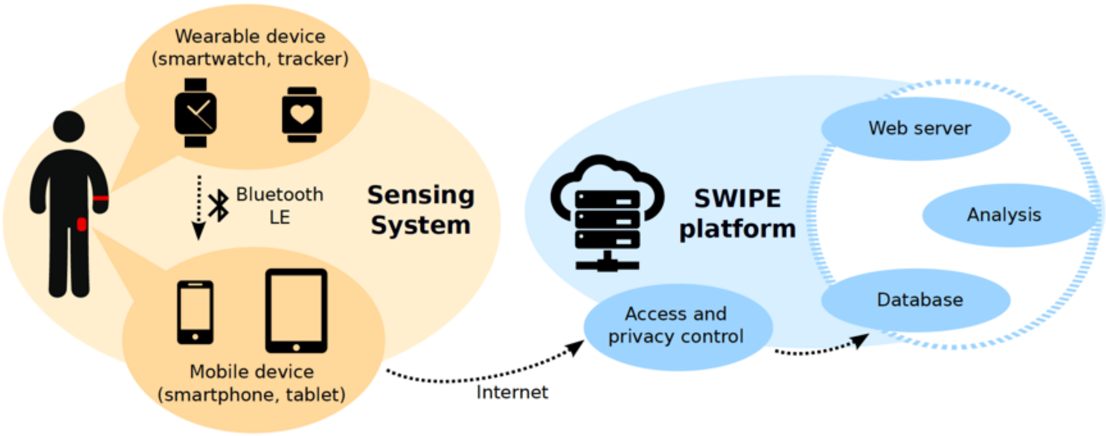
Methodologies for mobility data collection using smartphones and smartwatches

Activity analysis

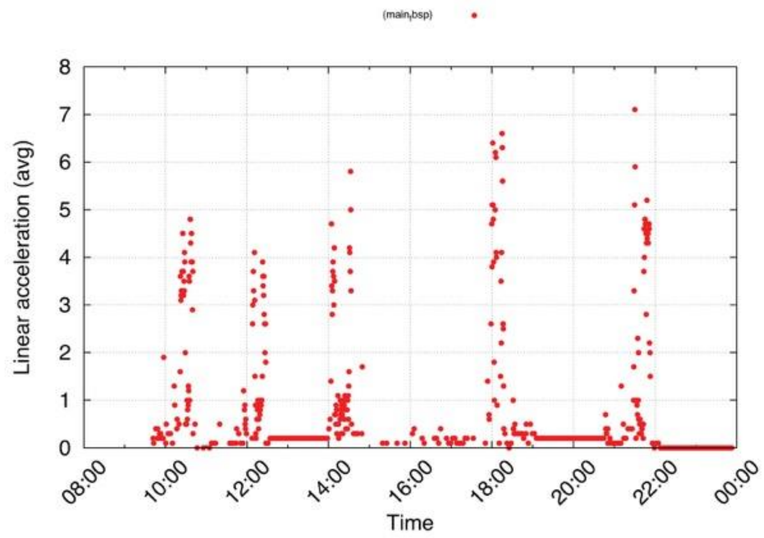
- Accelerometer
- Gyroscope
- Pedometer
- Proximity sensor
- Light sensor
- Sound sensor
- Heart rate monitor

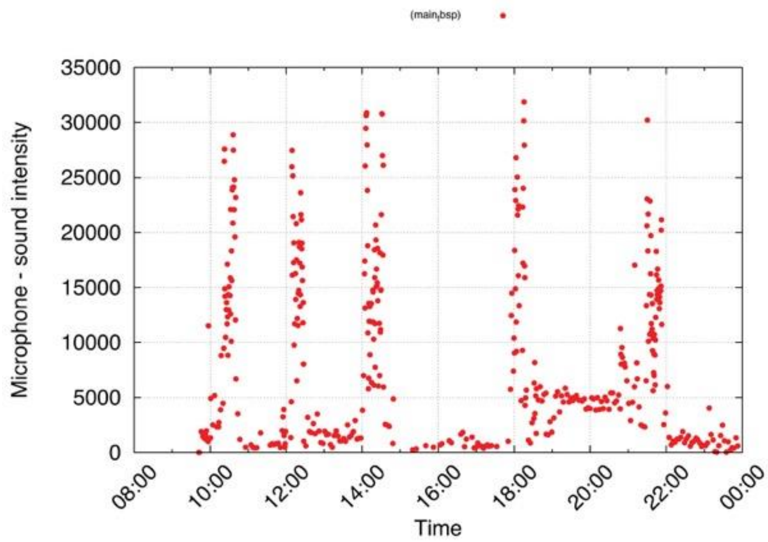
Position and social interaction

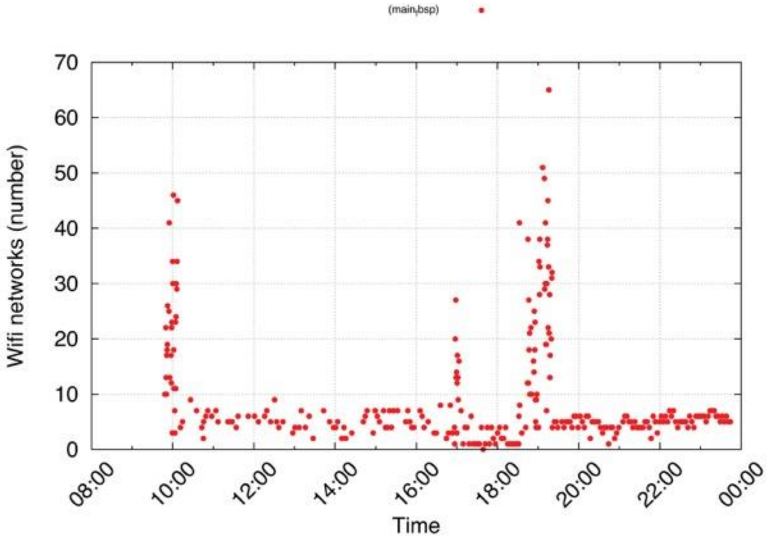
- GPS
- Wi-Fi
- Bluetooth

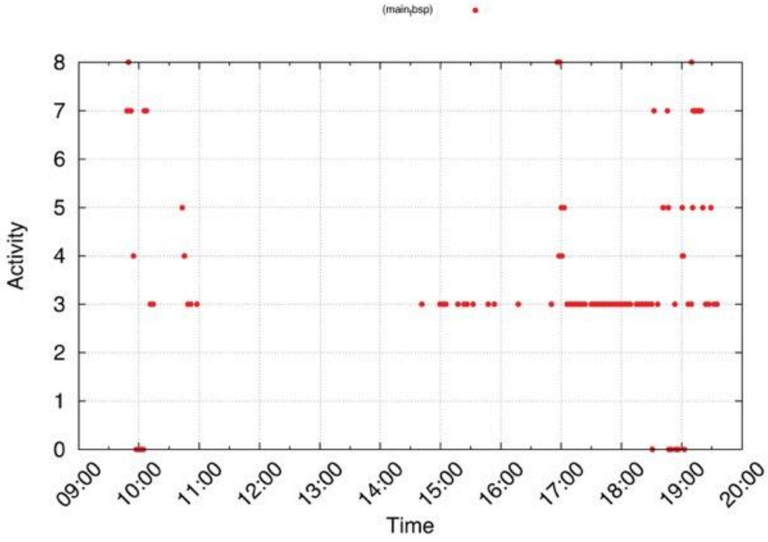


Activity-travel recognition

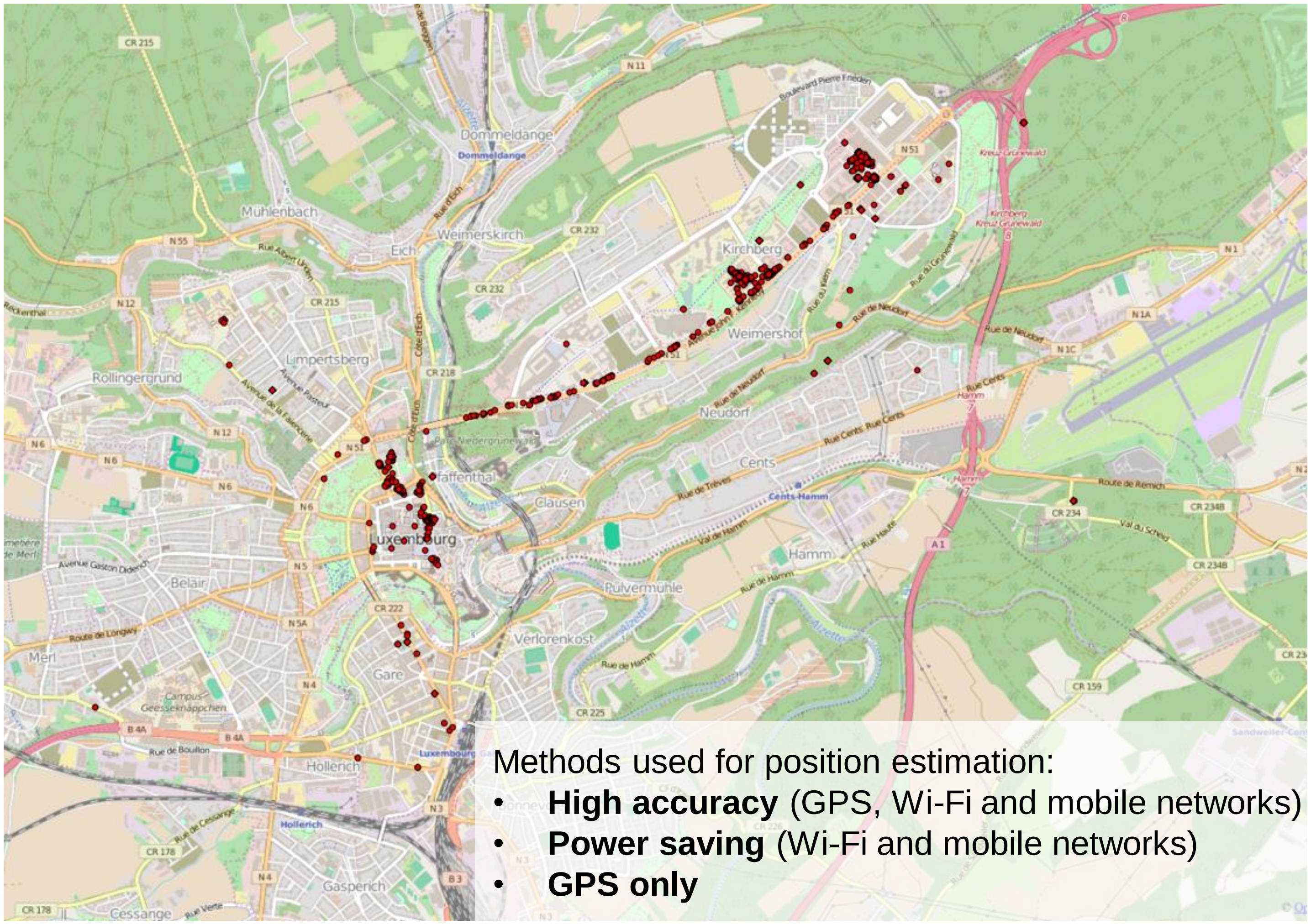








Position estimation

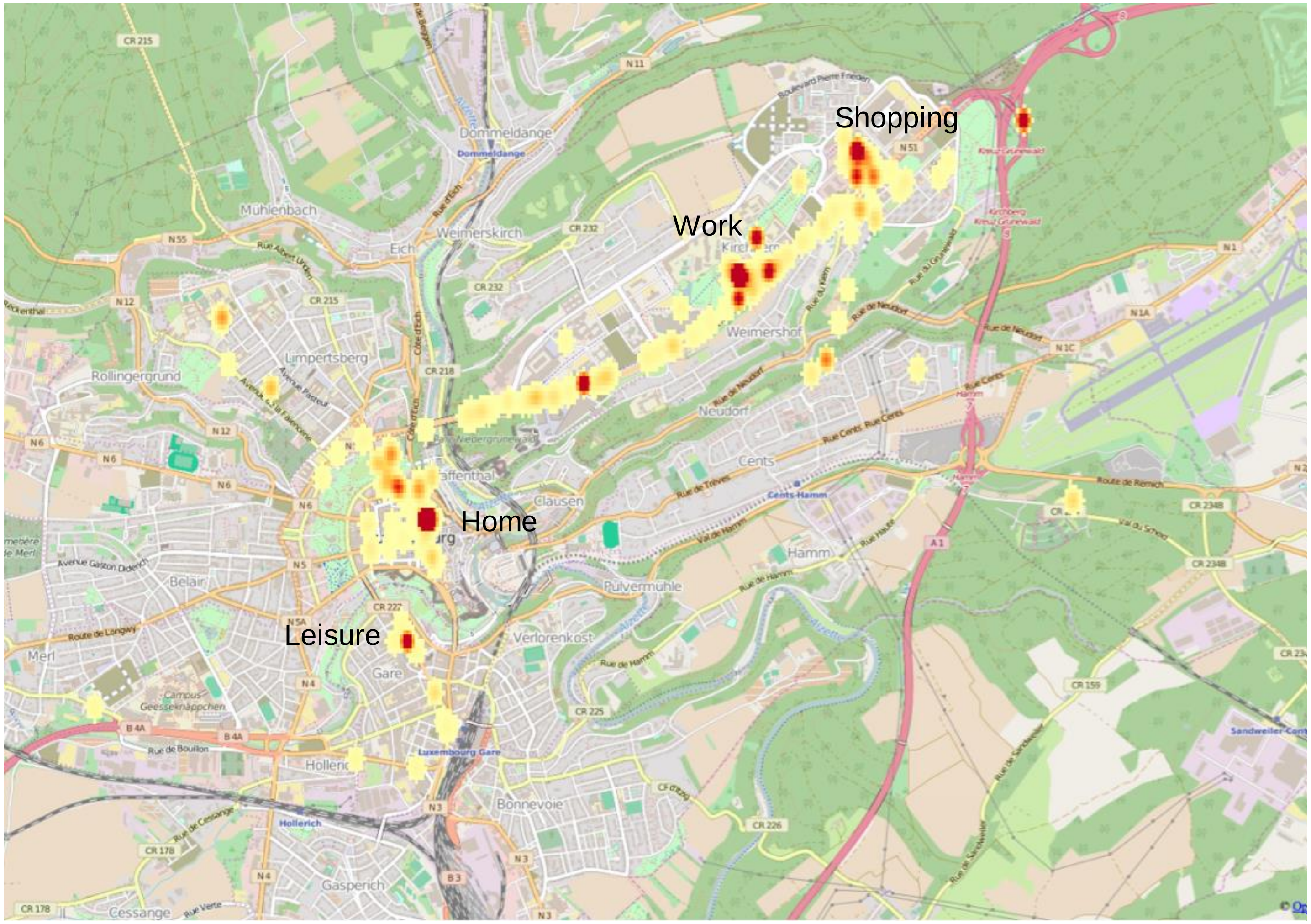


Methods used for position estimation:

- **High accuracy** (GPS, Wi-Fi and mobile networks)
- **Power saving** (Wi-Fi and mobile networks)
- **GPS only**

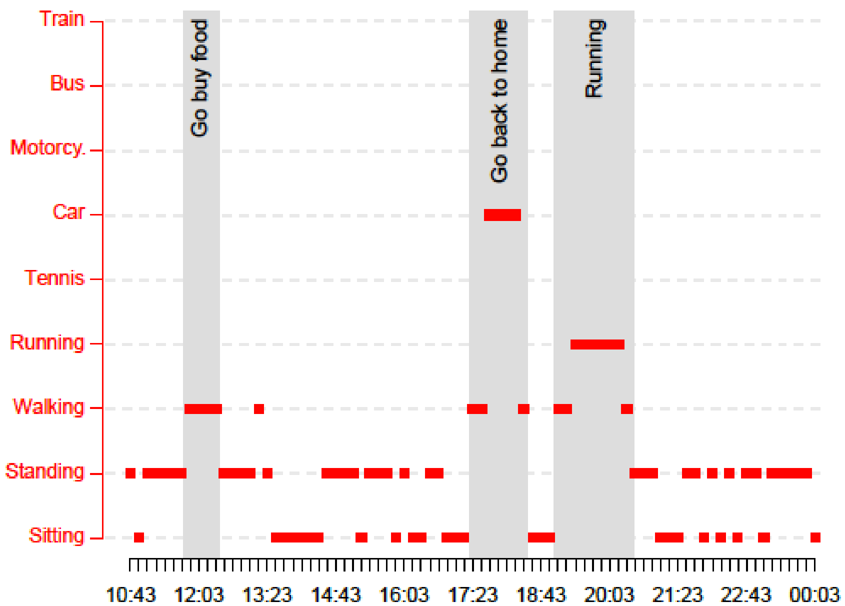
- Position estimation

Location estimation

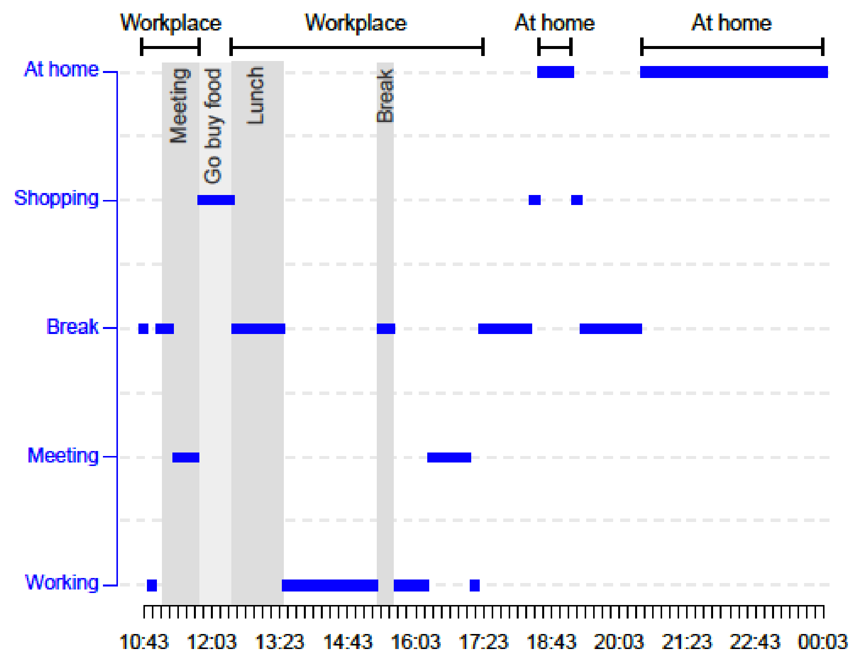


- Location identification and classification

Identifying activities and mode

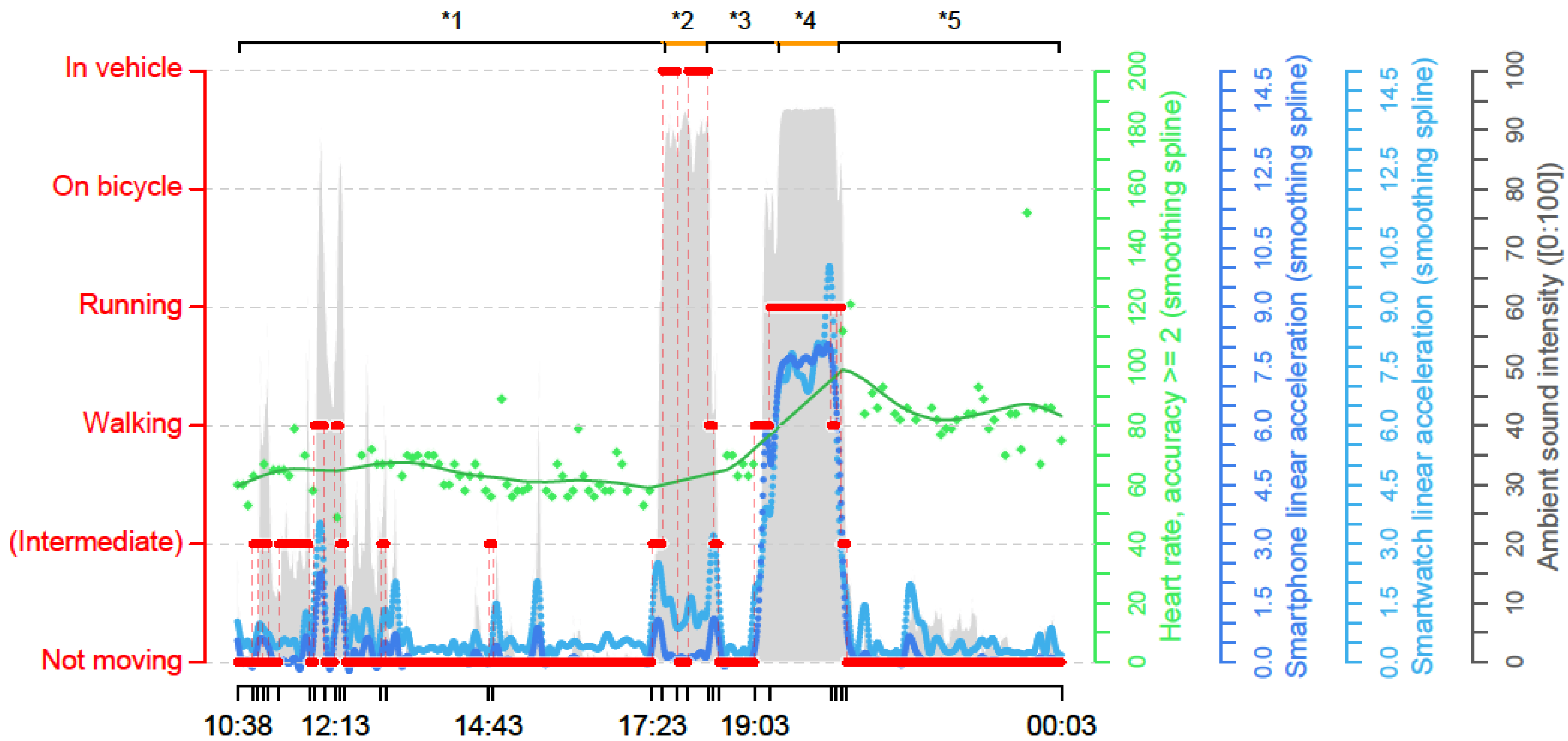


Activities



Contexts

Data fusion and machine learning



*1) Speed: 3km/h (avg), 18km/h (max); steps: 2206 (phone), 2542 (watch)

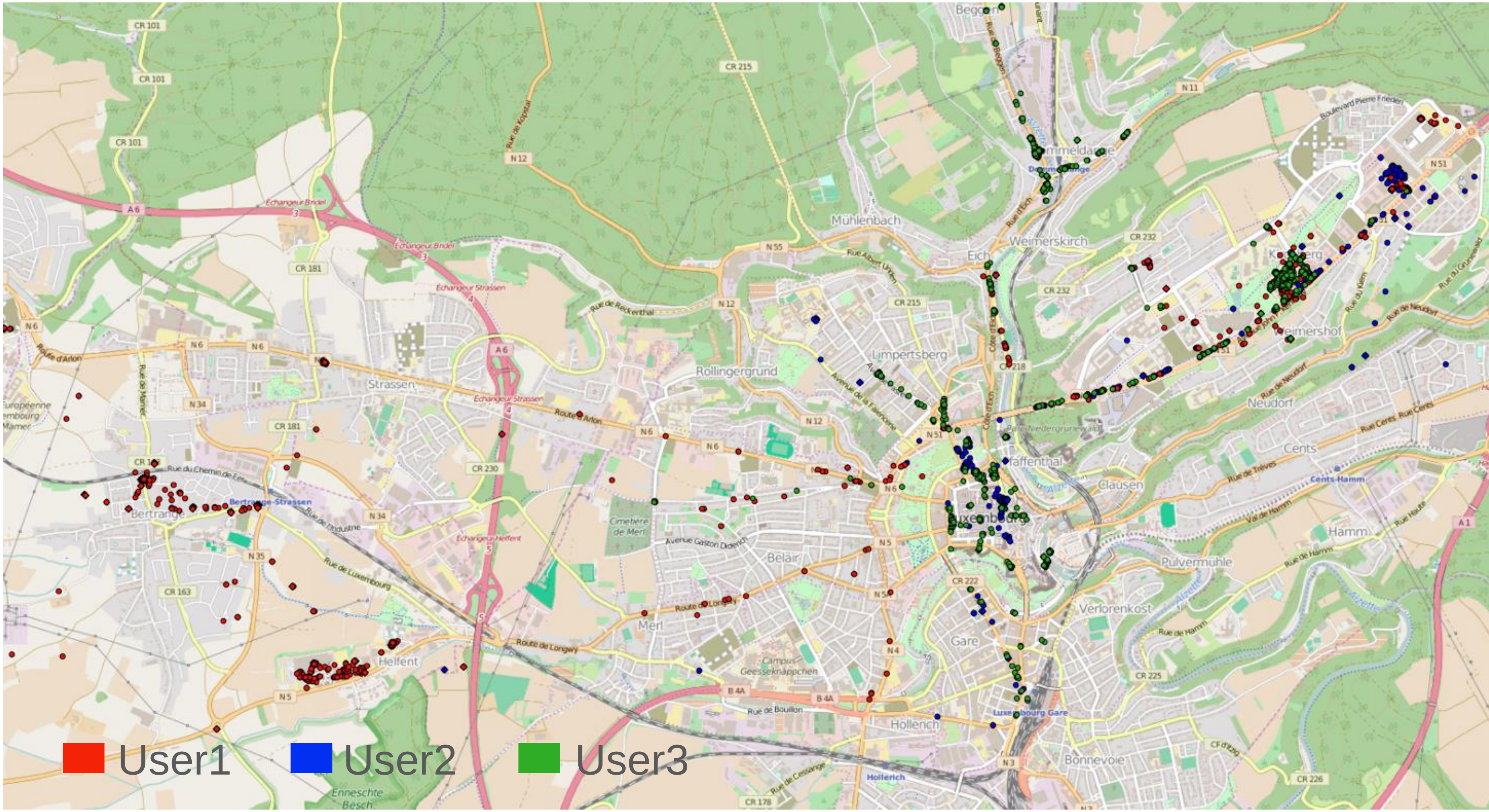
*2) Speed: 32km/h (avg), 107km/h (max); steps: 1 (phone), 0 (watch)

*3) Speed: 6km/h (avg), 20km/h (max); steps: 2394 (phone), 2404 (watch)

*4) Speed: 8km/h (avg), 14km/h (max); steps: 9573 (phone), 9574 (watch)

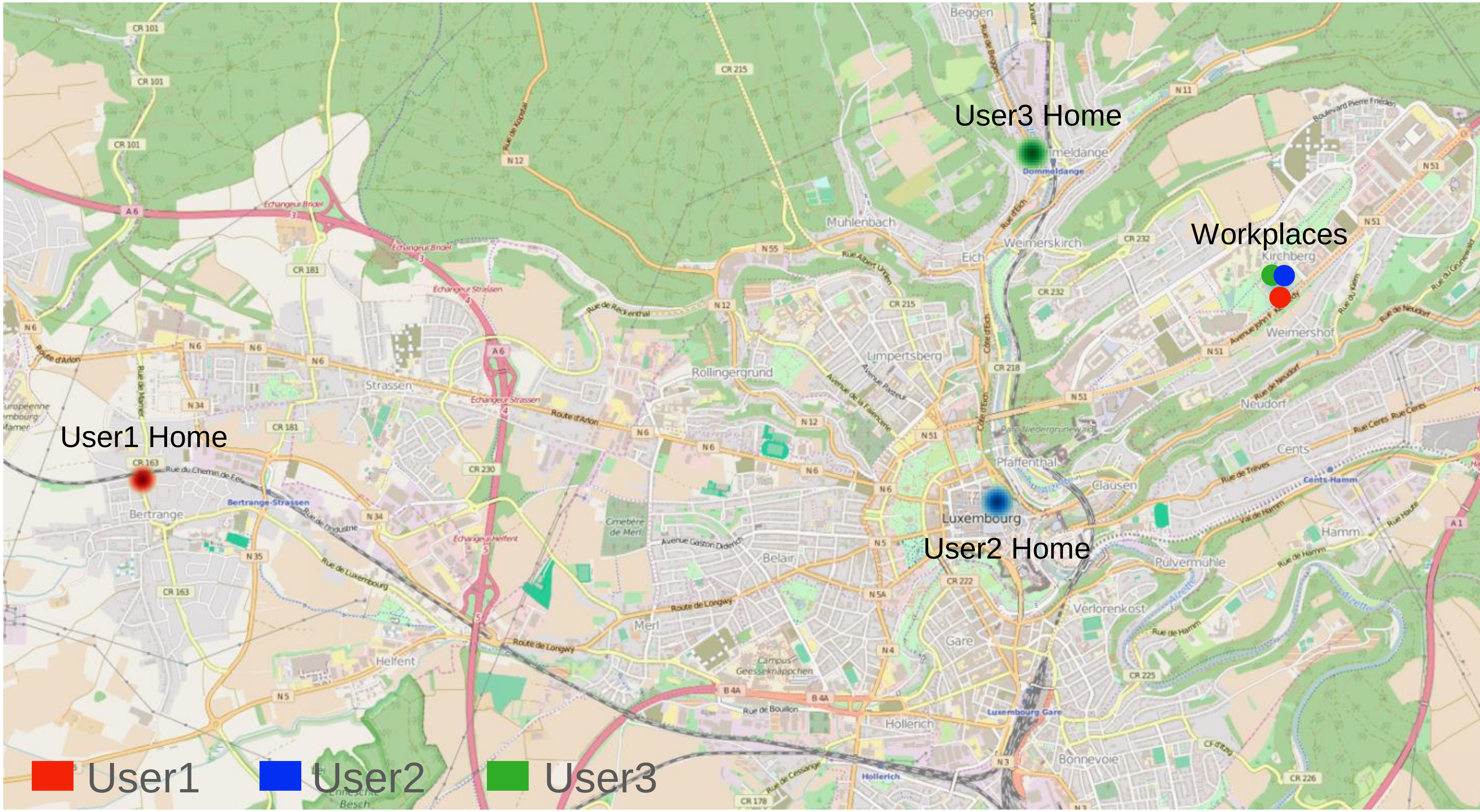
*5) Speed: 1km/h (avg), 5km/h (max); steps: 456 (phone), 623 (watch)

Group activity analysis



GPS data for 3 users

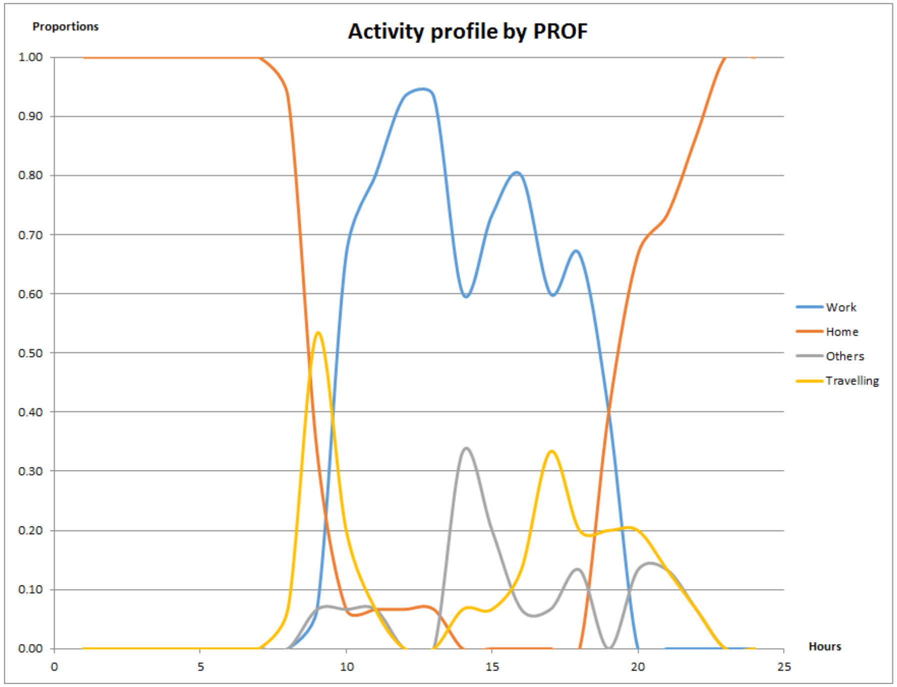
Identifying OD patterns

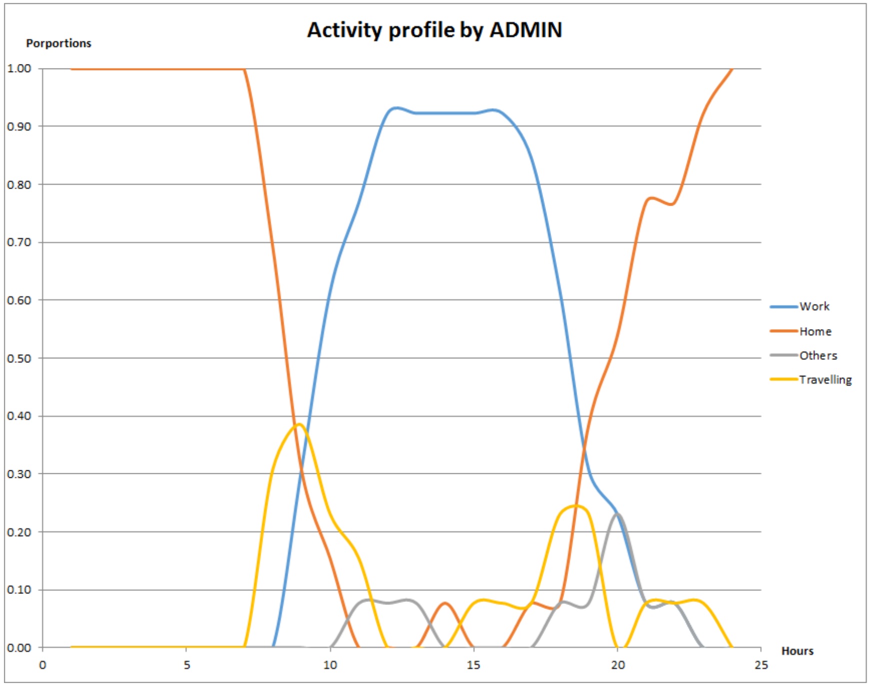


Home and workplace clustering estimation

Activity-travel patterns

- Example of derived daily activity-travel patterns
  - Different arrival/departure times by category
  - Different duration and scheduling of activities





Closure

- Enormous potentials offered by mobile sensors and floating car data technologies

- Applications investigated
  - Robust routing
  - Multimodal route planning
  - Dynamic traffic modelling

- New Big Data era: new opportunities and challenges
  - Understanding mobility needs
  - Forecast future activity-travel patterns
  - Enable users with enhanced information

THANK YOU!

Francesco.viti@uni.lu

uni.lu
UNIVERSITÉ DU LUXEMBOURG