Human mobility study: using mobile phone data for simulation and transportation research

FuturMob17 workshop , 5-7th September 2017

Mariem Fekih, Orange Labs, Hasselt University

Zbigniew Smoreda, Orange Labs

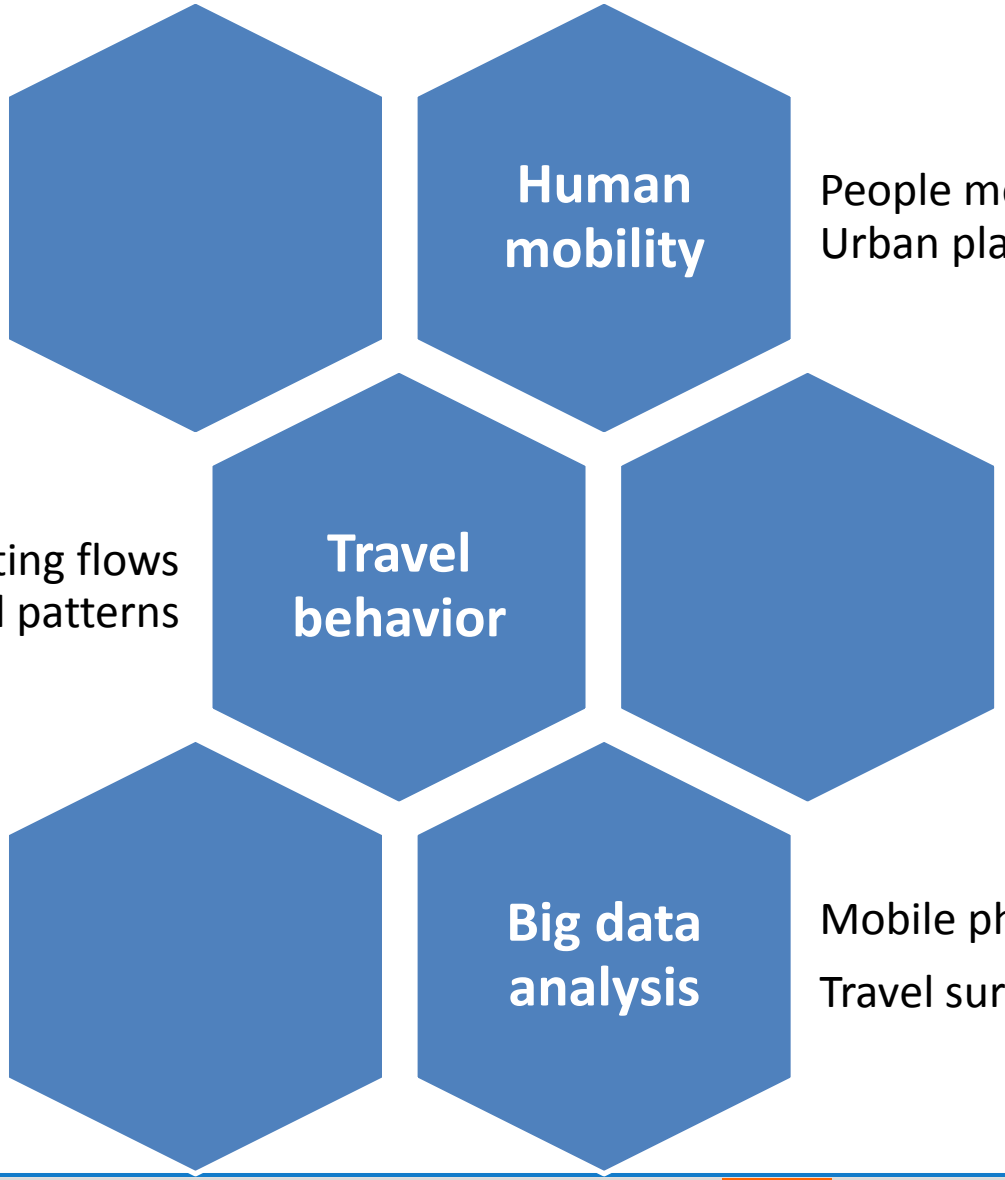
Stéphane Galland, UBFC

Tom Bellemans, Hasselt University

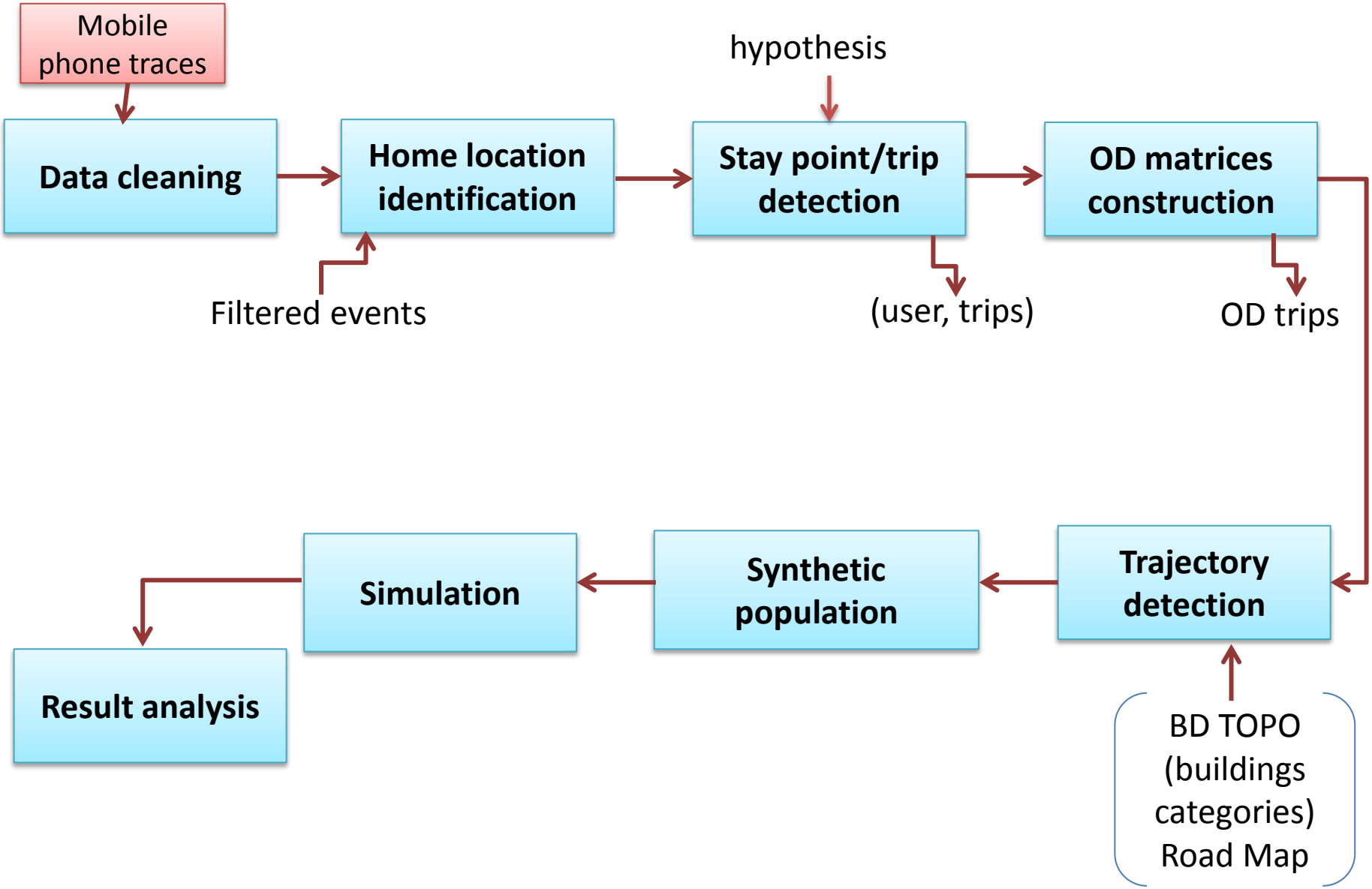
PLAN

- Overview of the PhD project
- Methodology and Objective
- Data description and analysis
- Results
- Conclusion and future work

OVERVIEW


Human mobility
People movement
Urban planning
Travel behavior
Commuting flows
Travel patterns
Big data analysis
Mobile phone data
Travel survey data

METHODOLOGY



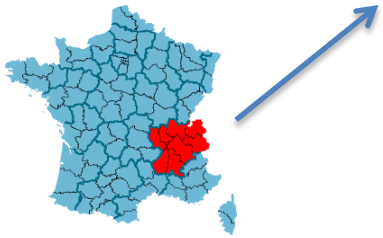
OBJECTIVE

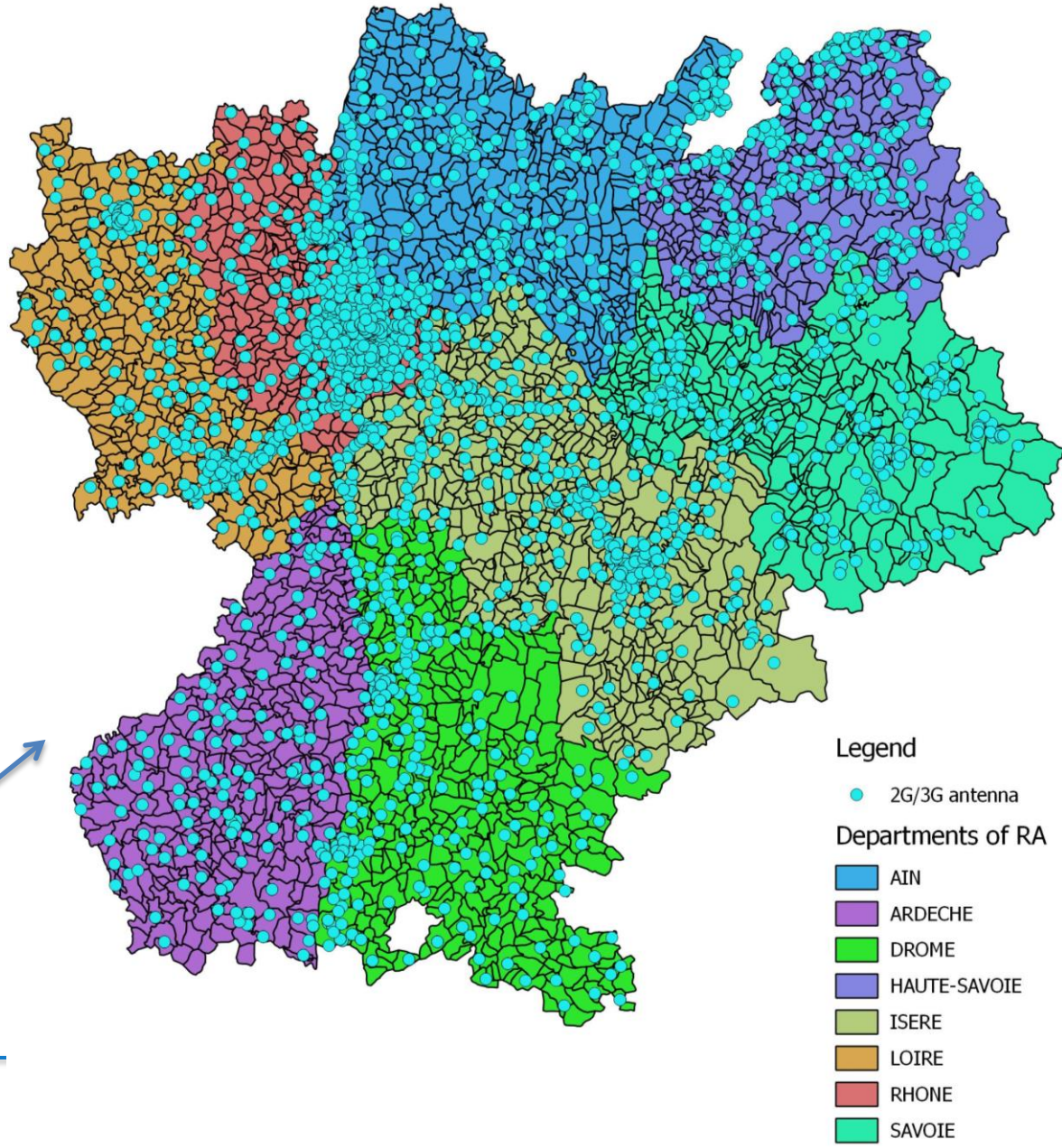
- Generate Origin-Destination matrices from mobile phone data
  - Home location identification
  - Trip detection
- Compare with O-D matrices generated from the household travel survey conducted in the Rhône-Alpes region (**EDR 2015**)

Origin\destination	D1	D2	D3
O1	X_{1-1}	X_{1-2}	X_{1-3}
O2	X_{2-1}	X_{2-2}	X_{2-3}
O3	X_{3-1}	X_{3-2}	X_{3-3}

DATA DESCRIPTION : COVERAGE AREA

- Mobile phone data are collected from Orange mobile network **2G/3G** probes and cover the territory of all the Rhône-Alpes region .





DATA DESCRIPTION

- 2 waves of mobile phones traces :
  - **First wave**: dataset of **19** days (from 1st to 19th of September 2016) of over **1.3 million** mobile phone users/ day.
  - **Second wave**: dataset of **15** days (from 1st to 15th of June 2017) of over **1.6 million** mobile phone users /day.

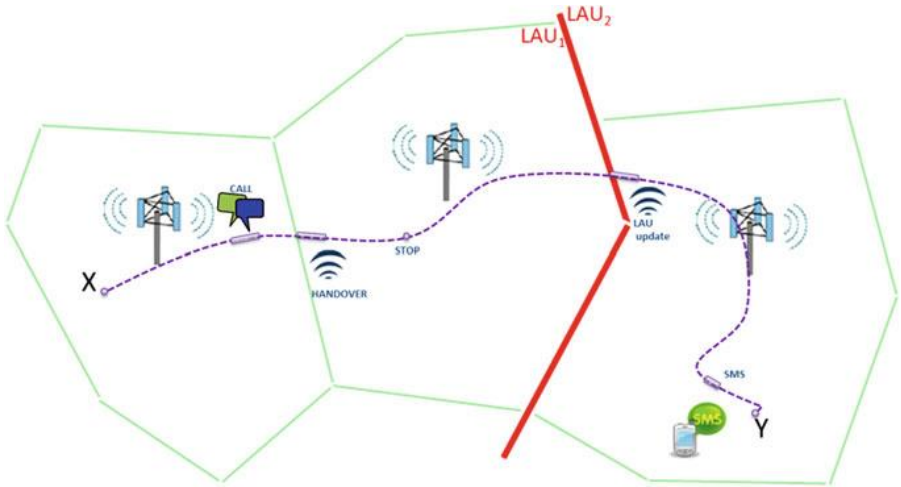
- Each record includes :

Timestamp	IMSI (ID)	LAC	CELL_ID	event
2017-06-01 11:53:33	201803567834	104	20865	CALL

Location = Location Area Code (LAC)+ CELL_ID

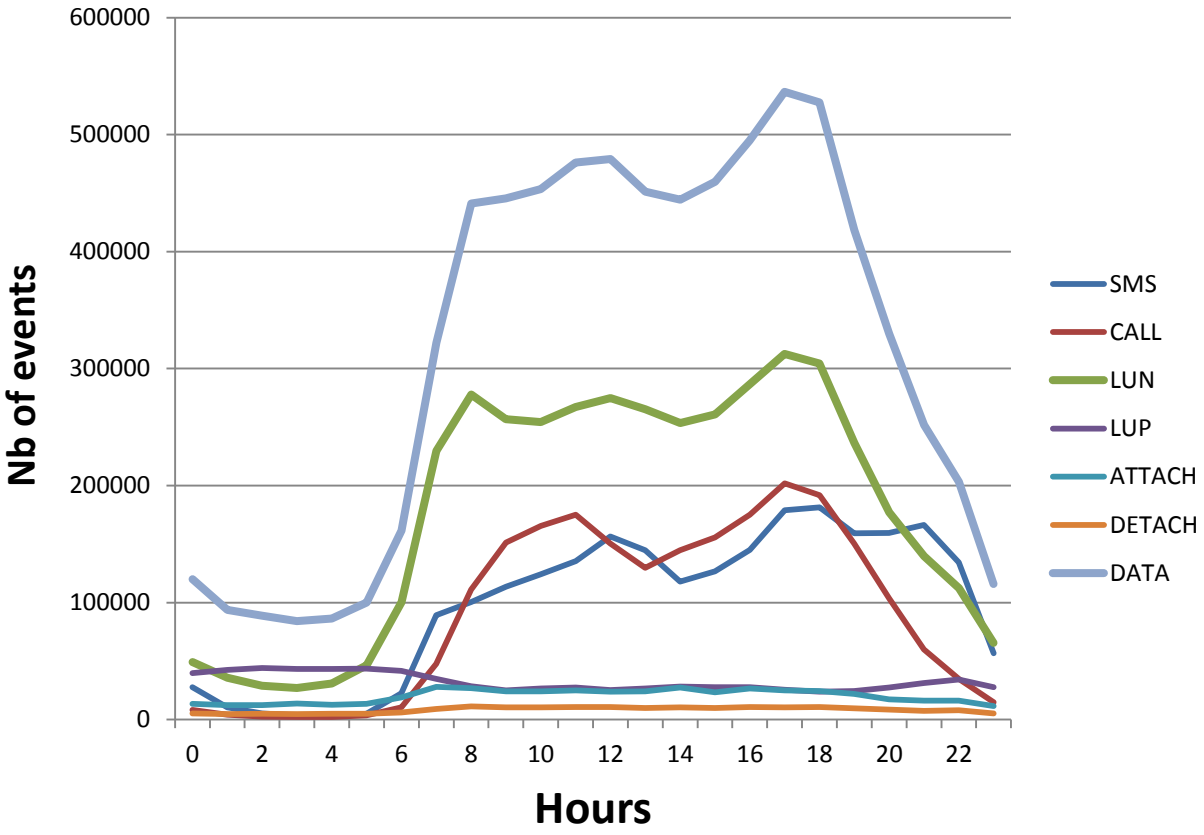
DATA ANALYSIS : EVENTS

- **Type of events:**
  - Communication events (Call , SMS)
  - Itinerancy events
    - Normal Location Area Update (LUN)
    - Periodical Location Area Update (LUP) : after 3 hours of inactivity
  - Attach and Detach events
  - Data events
  - Other events



DATA ANALYSIS : events' distribution

- **Temporal distribution:**



Home location identification : Method

- **Applied Method**

1. Filter events : Call, SMS, LUP, Attach, Detach and Data
2. Consider only traces in night time interval:

3. For each user , extract all locations and label the most frequent cell as home location

user → home location (LAC, Cell ID)

4. Derive the geographical correspondence between cellular towers and municipalities

Cell ID →municipality

5. Aggregate results at municipality level

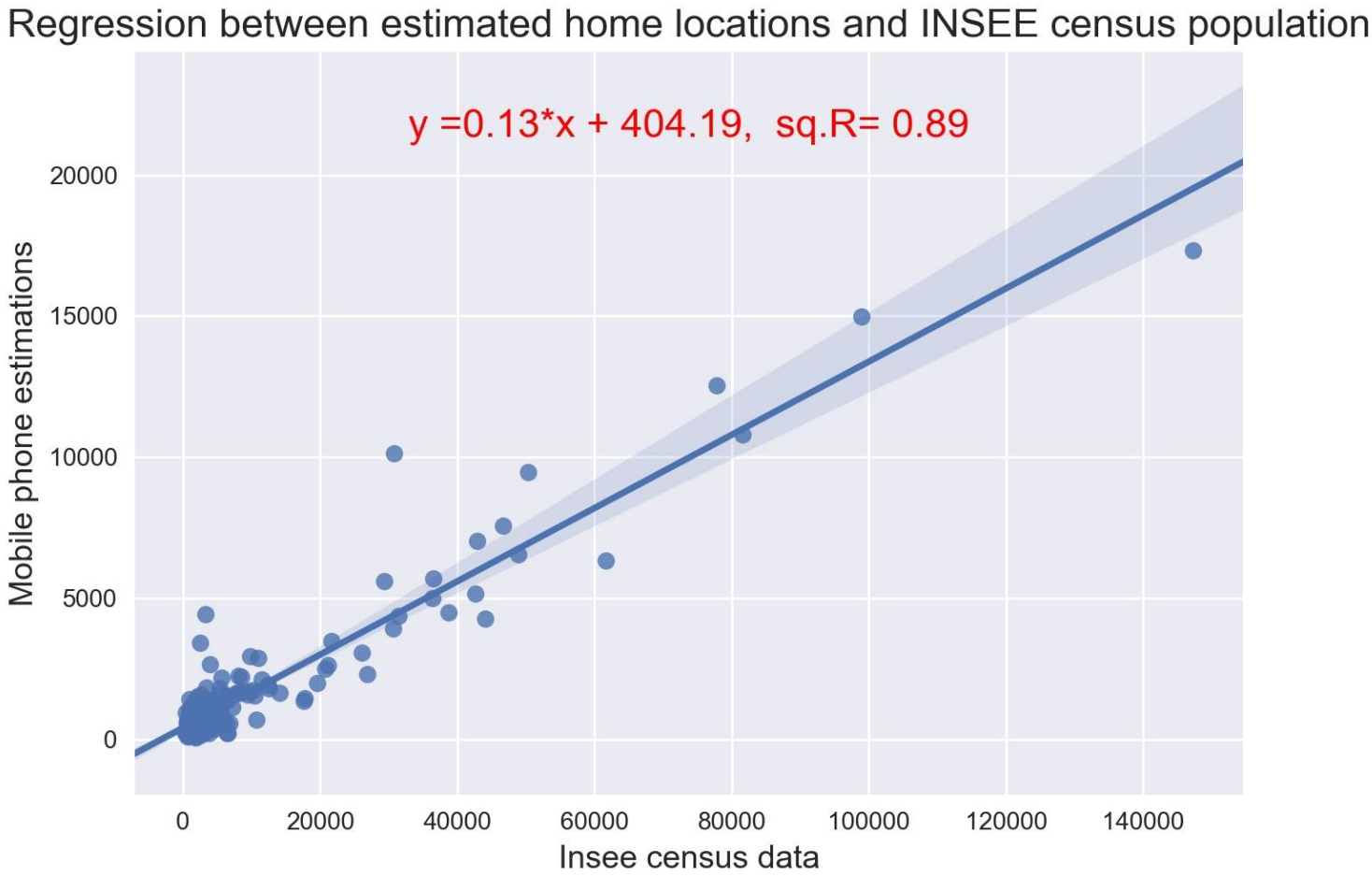
Home location identification : results

- Comparison with INSEE census data
  - Case study : **Rhone department**

Nb of considered municipalities	Nb of subscribers	Population (INSEE2013)	Market share (%)	Mobile phone estimation	MP Vs subscribers (%)	MP Vs INSEE (%)
161	375 226	1 645 658	22,8	**278 717**	**74,28**	**16,94**

Home location identification : results

- **Rhone department: municipality level**



OD Matrices construction: Method

- **Study area : Rhône-Alpes region territory**

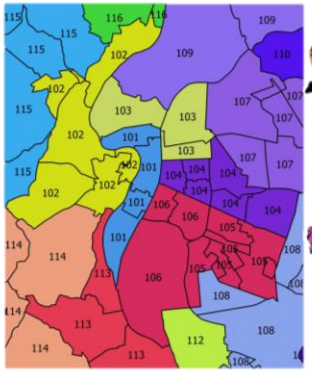
1. **Trips determination : Sector level**
   i. Detect Stop points with minimum stationary time : depends on the mean speed of motorised trips in each "Sector".
   ii. Filter users with 3 events or less
   iii. Determine trips as paths between user's consecutive stop locations : *trip (u, o,d)*

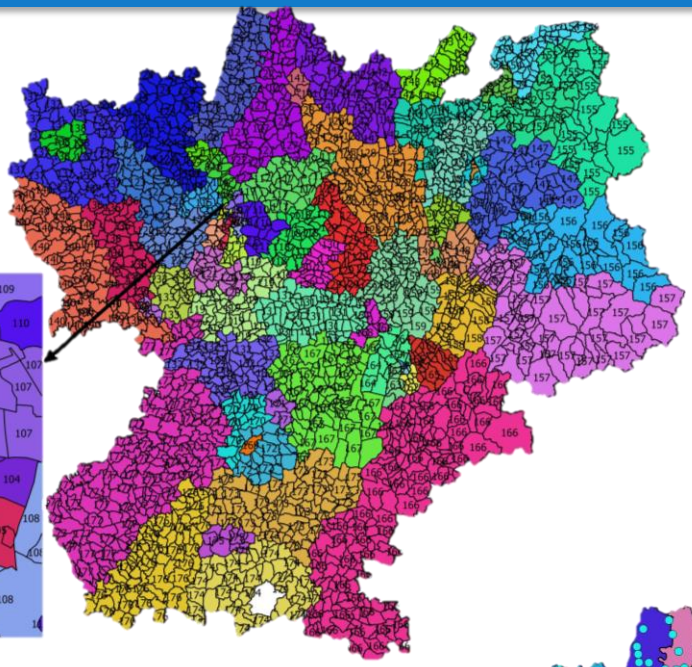
2. **OD estimation**
   i. Origin and destination sectors are extracted for each trip *trip (u, o,d)* of each user
   ii. Trips with same origin and destination sectors are grouped together.

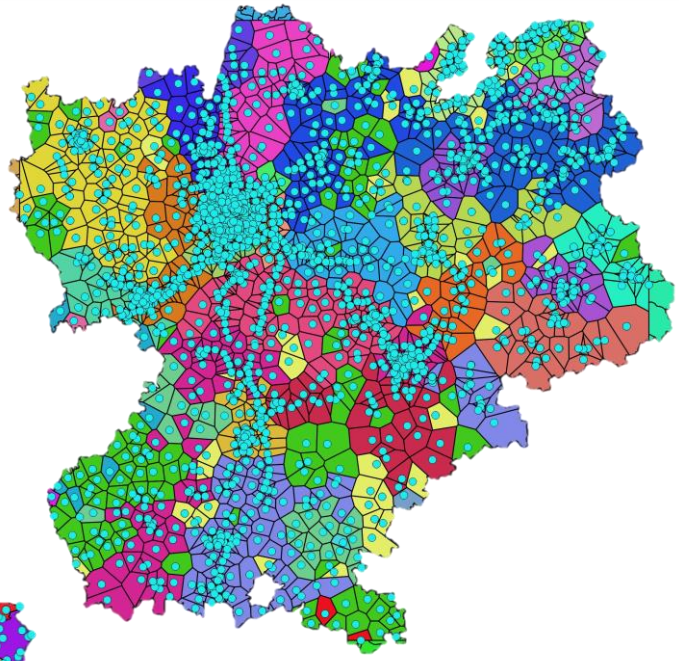
DATA ANALYSIS : geographic zoning systems

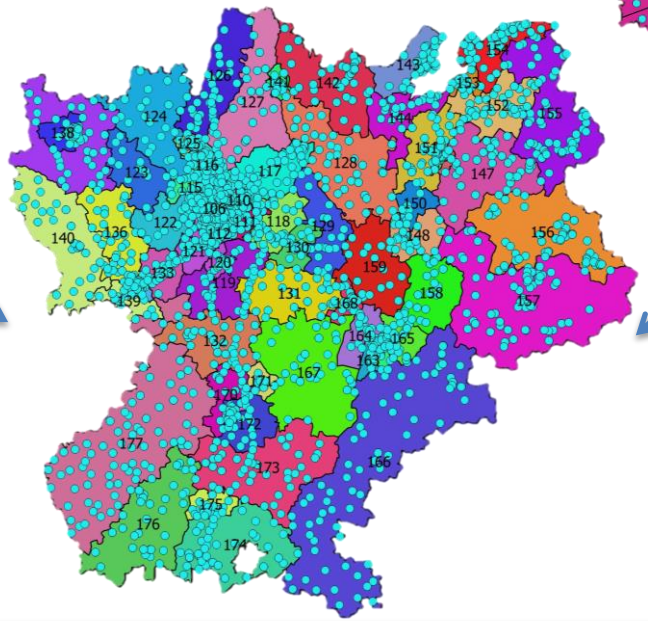
- **Matching between the different geographic zoning systems :**
  - Cellular network of Mobile phone data : Voronoi (cell) partitioning
  - Household Travel survey data (EDR) : different zoning levels (Sectors , aggregated zones ,…)
  - Census data : administrative partitioning

DATA ANALYSIS : geographic zoning systems









EDR survey :
77 zones (“sectors”)

Cellular network

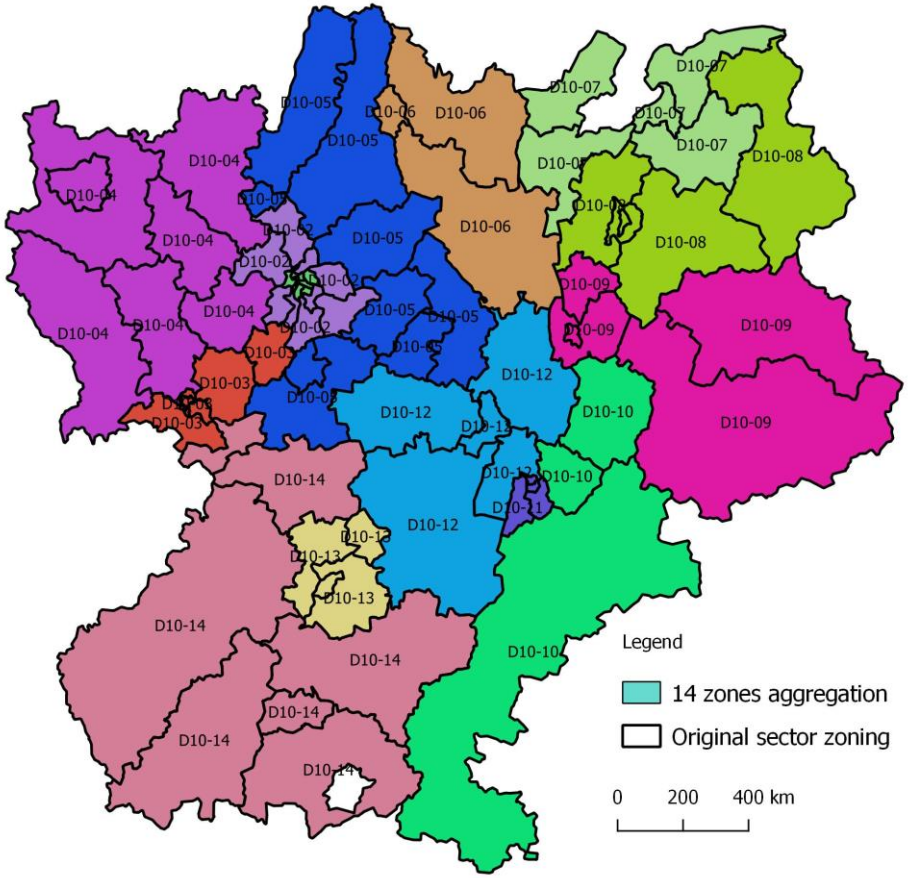
OD Matrices construction: Method

- **EDR Sector aggregation**

i. Construct a conversion matrix to convert 77 zoning system to 14 zoning system

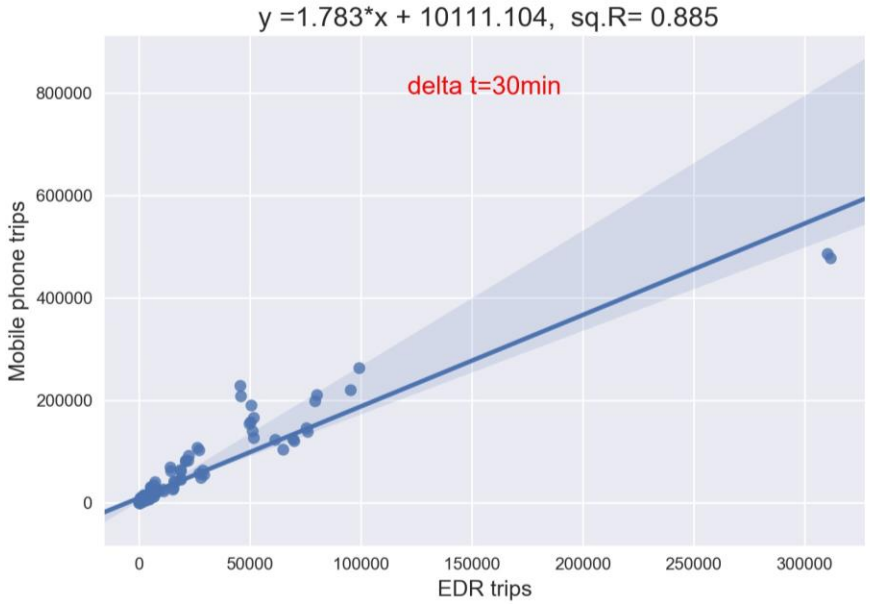
ii. Aggregate trips and generate OD matrix (14 *14 OD pairs)

iii. Compare with OD estimations of the travel survey

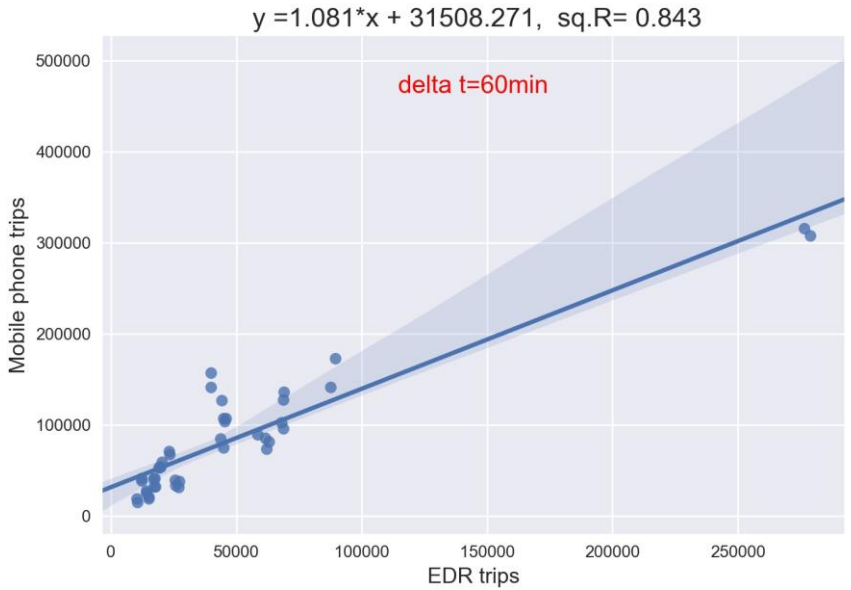


OD Matrices : Results

- **Linear regression models: Comparison between mobile phone and EDR Survey OD estimates**



delta t : minimum stationary time



Expansion factor:

$$C_{expansion} = \frac{Population\ of\ RA\ region}{Nb\ of\ moving\ users\ using\ network}$$

Conclusion and Future work

❑ Proposals

- ➢ Generation of the Origin-Destination matrices from mobile phone data
- ➢ Comparison with O-D matrices generated from the household travel survey conducted in the Rhône-Alpes region (**EDR 2015**)

❑ Futur work

- ➢ Improve OD estimation by including the number of home locations per sector in the expansion factor formula
- ➢ Study and extract individual trajectories at fine-grained scale (eg. city level)

Thank you !

Questions?