Origin-Destination estimation using mobile network probe data

Patrick Bonnel [a], Mariem Fekih [b], Zbigniew Smoreda [b]

[a] *Laboratoire Aménagement Economie Transports, ENTPE, Lyon, France, patrick.bonnel@entpe.fr*
[b] *SENSE, Orange Labs, Paris, France, mariem.fekih@orange.com, zbigniew.smoreda@orange.com*

Objectives

Mobile phone operators produce enormous amounts of data. These data contain information regarding individual positions both in time and space. Several researches have been performed using these data to produce origin-destination (O-D) matrices for planning, modelling…
But few researches **confront mobile phone O-D matrices with "ground truth"**.

Number of users according to number of events (Orange, June 1st, 2017)
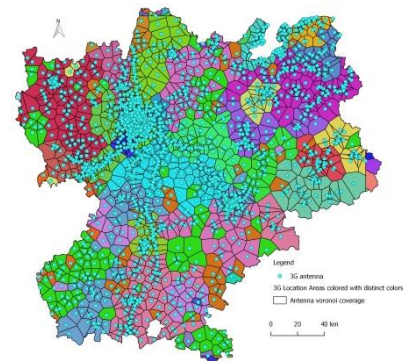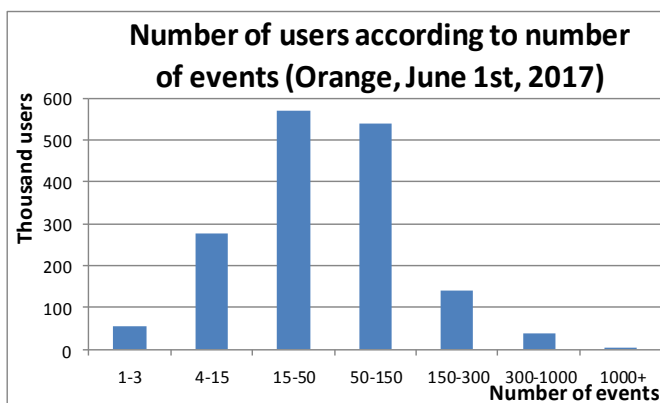
Area covered by the 3G Location Area in the Rhône-Alpes region

Data

- **Orange mobile phone** network "Cigale" probes through 3G base stations collect signaling data (voice, SMS, internet data, handover and locations area update) for June 1st 2017 ("normal" working day);
- Rhône-Alpes region (France) **travel survey** (2012-2015): 37,000 individuals, 143,000 trips.

O-D matrix construction **methodology**

- zoning: travel survey sector;
- trip definition: a movement in space with an activity (stationarity) at origin and at destination;
- Stationarity definition: at least two successive (or a succession of) events with a time interval above a threshold (30mn-1h) and within the same zone;
- Mobile data expansion factor:

$$C_{exp} = \frac{Population\ of\ RA\ region}{Nb\ of\ users\ using\ 3G\ network}$$
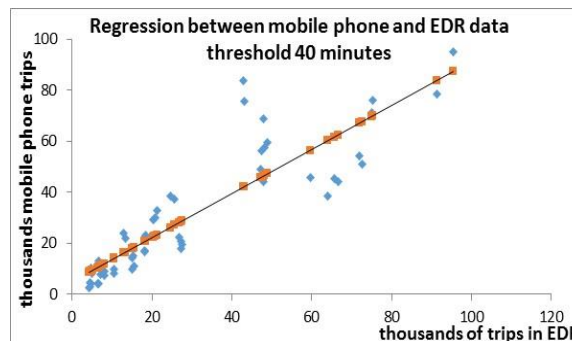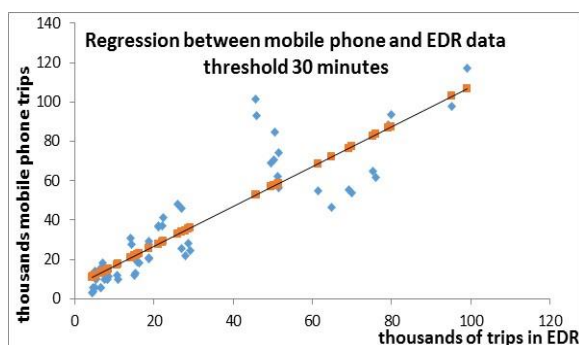
Number of trips: mobile phone data; travel survey (EDR)

Threshold to determine stationary activity	60mn	50mn	40mn	30mn
EDR (in thousand)	2,211	2,260	2,344	2,448
Mobile phones (in thousand)	1,605	1,872	2,226	2,762

Results

30 mn threshold: $y_{ij} = 1.01 * x_{ij} + 6{,}517$, with $R^2=0.79$; student t constant = 2.31 slope = 14.5
40 mn threshold: $y_{ij} = 0.87 * x_{ij} + 4{,}912$, with $R^2=0.79$; student t constant = 2.13 slope = 14.39
With Y number of mobile phone trips, X number of travel survey trips

Regression between mobile phone and EDR data threshold 30 minutes

Regression between mobile phone and EDR data threshold 40 minutes

Synthesis and perspectives

- ❖ origin-destination matrix (volume and structure) are very sensitive to trips ends (stationarity) definition;
- ❖ Correlation between mobile phone data and travel survey rather good (close to 0.8);
- ❖ But at origin-destination level, differences might be quite big for some O-D (up to 100%);
- ❖ Mobile phone overestimations mainly concern adjacent zones which might illustrate zone delimitation problem or ping-pong effects;
- ❖ Strong hypothesis were defined which need research development to overcome:
  - Mobile phone data contain all trips on territory, while travel survey collect resident trips;
  - Orange operator mobile phone users are representative of the whole population;
  - Trip definition, stationarity threshold, mobile base station = Voronoi polygon…