Vehicular Crowd-Sensing on Complex Urban Road Networks: a Case Study in the City of Porto

Sergio Di Martino, **Luigi Libero Lucio Starace**

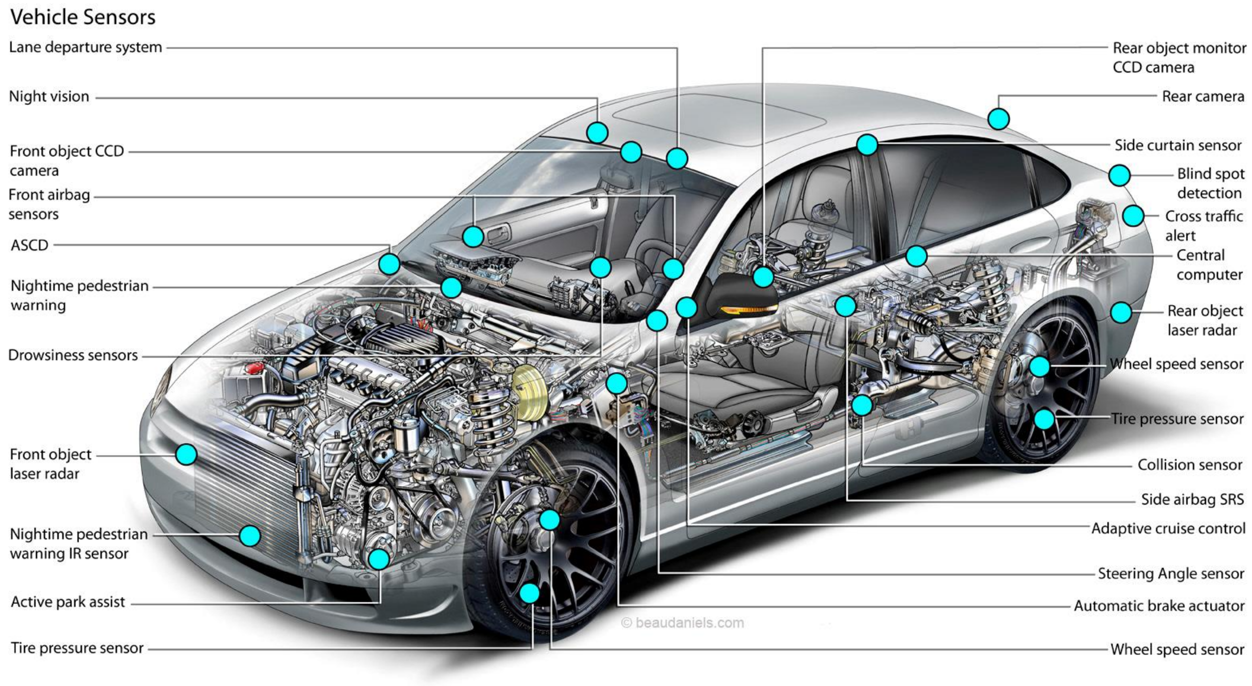
Università degli Studi di Napoli Federico II, Naples, Italy

luigiliberolucio.starace@unina.it

September 9, 2021

Modern Connected Vehicles

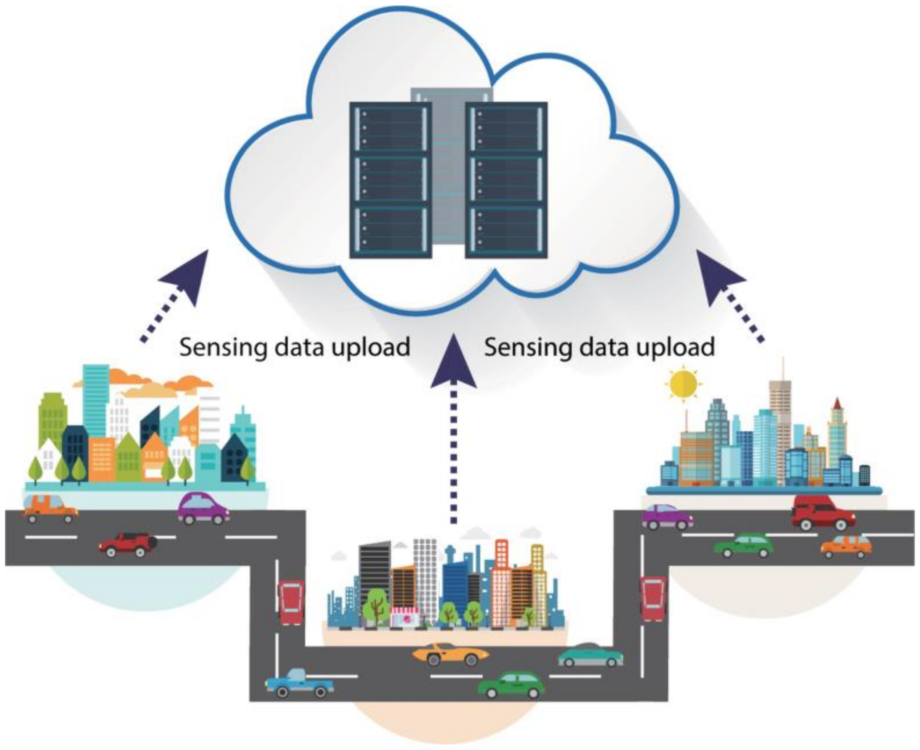
- Modern vehicles are equipped with a number of high-quality **sensors**
- In 2030, **95%** of new vehicles sold globally will be **connected**



Source: beaudaniels.com

Vehicular Crowd-Sensing (VCS)

- Use vehicles as **probes** to collect contextual information on phenomena of interest:
  - Availability of on-street parking
  - Amount of pollutants in a city area
  - Presence of potholes on a street
  - And much more!
- **Economically efficient**, especially when exploiting high-mileage vehicles such as taxis



Yu TY., Zhu X., Maheswaran M. (2018) Vehicular Crowdsensing for Smart Cities. In: Handbook of Smart Cities. Springer, Cham. https://doi.org/10.1007/978-3-319-97271-8_7

Value of VCS-collected data

- A number of use cases and exciting **smart mobility** solutions can be developed on top of VCS data
- Proper exploitation of these data *could deliver up to **$400 billion** in annual incremental value for players across the ecosystem in 2030.*

McKinsey & Company

Automotive & Assembly

Unlocking the full life-cycle value from connected-car data

February 11, 2021 | Article

https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/unlocking-the-full-life-cycle-value-from-connected-car-data.

Road Network Coverage in VCS

- Different use cases require different **sensing distributions**.
- Air quality monitoring requires that a given area is visited once every few hours.
- On-street parking monitoring requires way more frequent and road-specific sensing.
- The achievable **spatio-temporal sensing coverage** is a **Key Performance Indicator** for VCS solutions

Motivations

- Studies investigating the spatio-temporal coverage achievable by a fleet of vehicles:
  - Focus on urban road networks featuring a regular, grid-like topology (e.g.: [1])
  - Consider coverage only at a coarse-grained level of city areas (e.g.: [2])
- Generalizability on different road network topologies has not yet been deeply investigated

Road network of San Francisco, studied in [1]

[1] Bock, F., Di Martino, S., Origlia, A., 2020. Smart parking: Using a crowd of taxis to sense on-street parking space availability. IEEE Transactions on Intelligent Transportation Systems 21, 496–508. doi: 10.1109/TITS.2019.2899149.
[2] Masutani, O., 2015. A sensing coverage analysis of a route control method for vehicular crowd sensing, in: Pervasive Computing and Communication Workshops (PerCom Workshops), 2015 IEEE International Conference on, IEEE. pp. 396–401.

A Case Study in the City of Porto

The Case Study

Goal: evaluating feasibility of using taxis as probes in VCS in cities with complex road networks.

The Case Study

- Leveraged a **massive dataset** of **real** taxi trajectories recorded in the **City of Porto**
- Measured the spatio-temporal coverage achieved by those taxis at a **fine-grained** level of **road segment**

Dataset

- We leveraged the massive dataset presented in [3]
- **1,710,671 trajectories** from **441 taxis**, collected over one year
- Each trajectory is represented by a sequence of GPS points.
- As for the road network, we used data from the well-known **OpenStreetMap** project.

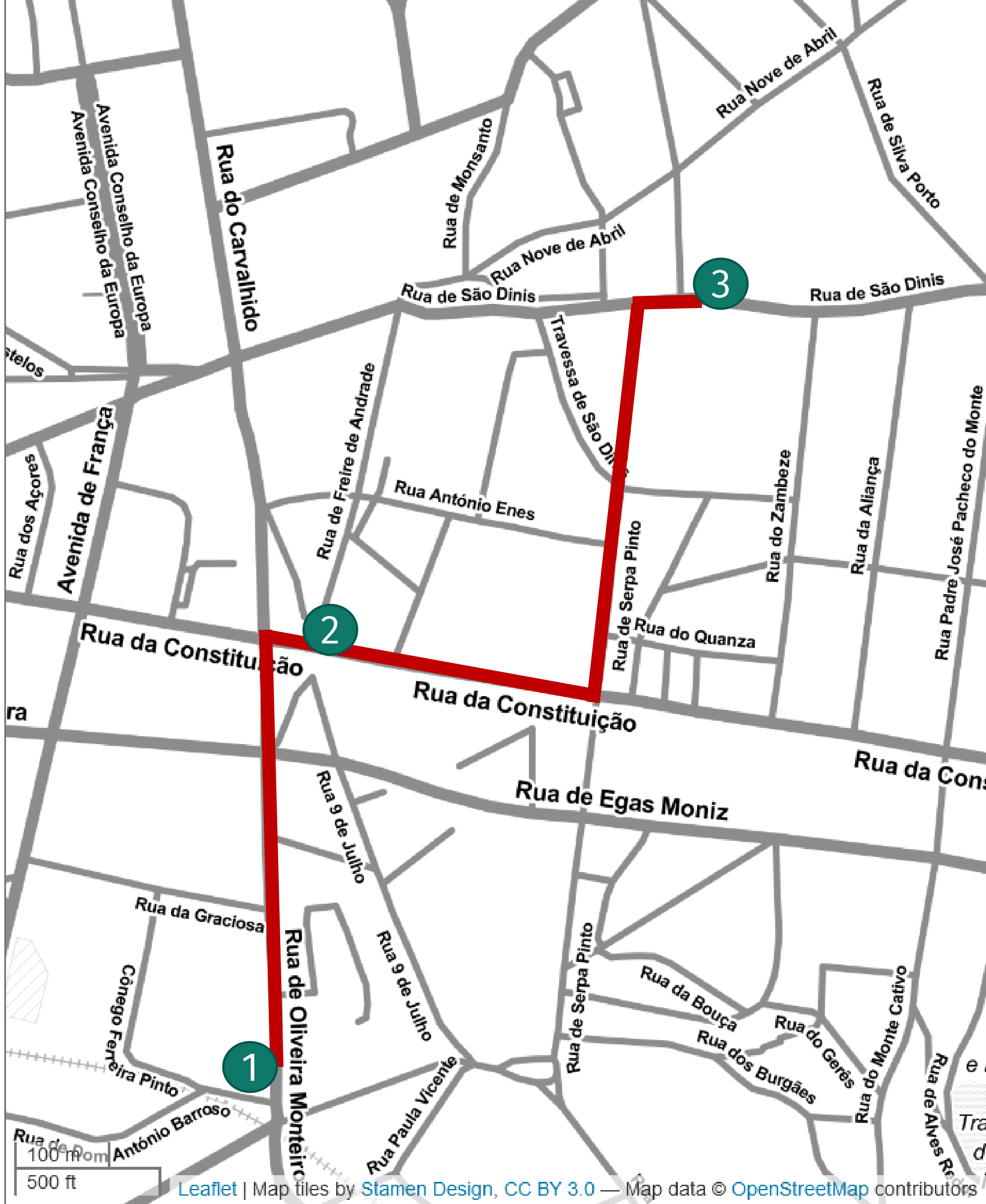
[3] Moreira-Matias, L., Gama, J., Ferreira, M., Mendes-Moreira, J., Damas, L., 2013. Predicting taxi–passenger demand using streaming data. IEEE Transactions on Intelligent Transportation Systems 14, 1393–1402.

Data Selection

- Since many VCS use cases are related to **urban** environments, we restricted our analysis to the urban area of Porto.
- We also temporally limited the investigation to three contiguous weeks, to enable future replication on different datasets that tipically span over shorter periods of time.
- After these filtering steps, we retained **~100k** trajectories
- We randomly sampled **100 taxis** and considered only the trajectories from those taxis (and repeated the case study 5 times to account for fluctuations due to this random sampling).

Map Matching

- Necessary to **align** raw GPS positions with the OSM road network.
- We did this by using the Open Source Routing Machine (OSRM)
- For each trajectory, we queried OSRM for a route passing through all the GPS points.
- OSRM returned a sequence of traversed road segments.



Measuring Road Network Coverage

- **For each road segment** in the considered map, we computed:
  - The number of **times it was traversed** by a taxi in the 3-weeks period
  - The **average timegap** between two subsequent visits.
- To gain additional insight on the traffic dynamics, we also **aggregated** these metrics by different **types** of **road segments** as defined in the OSM standard (e.g.: primary, secondary, tertiary, residential roads, etc…)

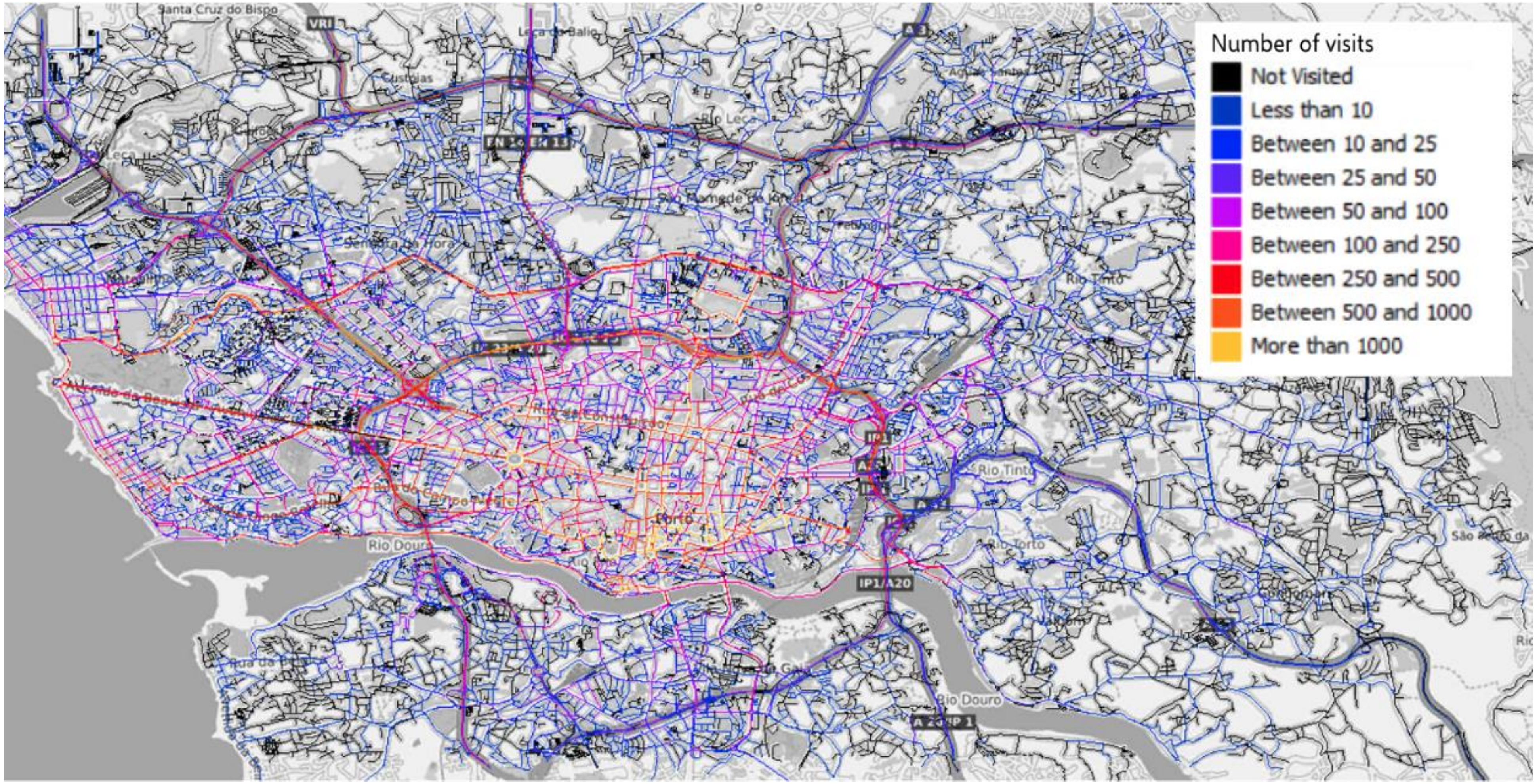
Results

Spatial Coverage Results for 100 taxis

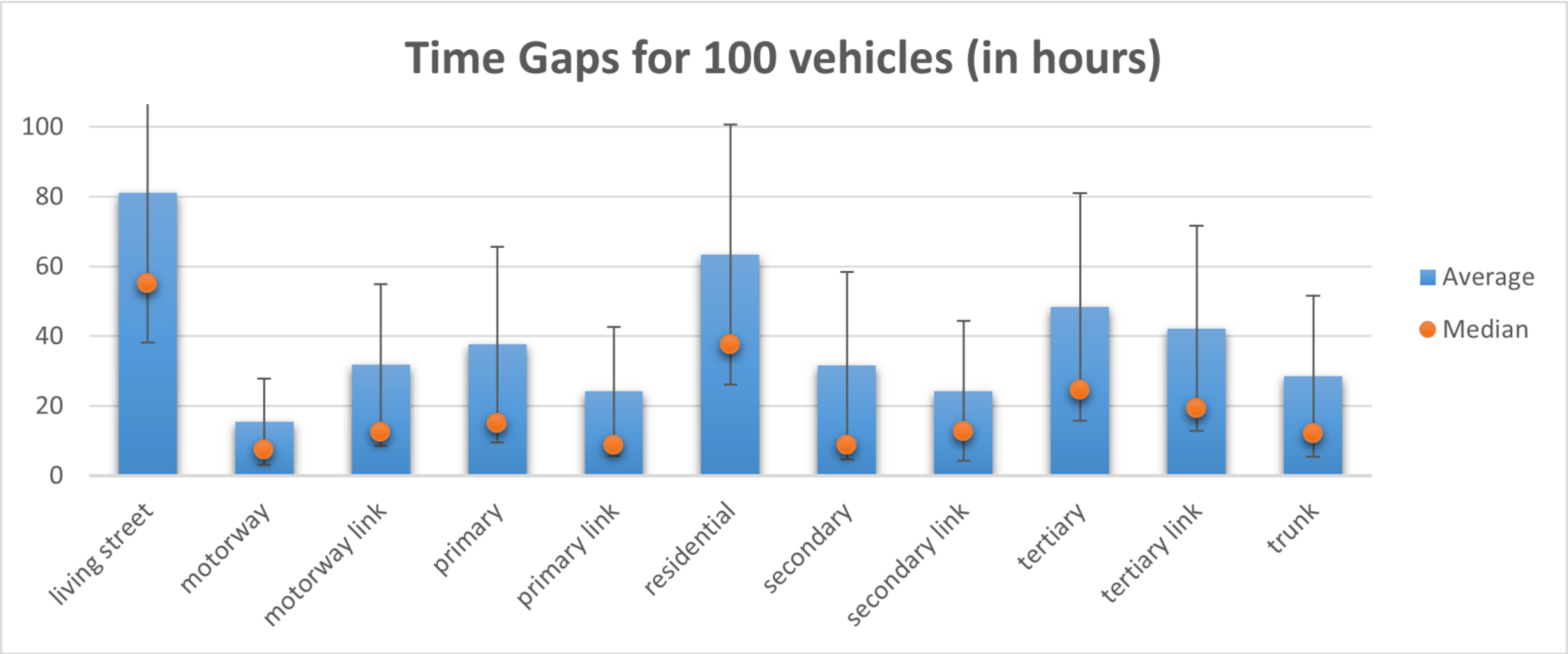
- On the right, percentage of segments visited at least once, by type
- Generally **high spatial coverage**, with differences between different road types

Segment type	**Coverage percentage**
Living street	94.4 %
Motorway	95.8 %
Motorway link	87.2 %
Primary	94.5 %
Primary link	95.6 %
Residential	60.2 %
Secondary	92.9 %
Secondary link	75.5 %
Tertiary	86.3 %
Tertiary link	87.4 %
Trunk	98.2 %

Spatial Coverage Results for 100 taxis



Temporal Coverage Results for 100 taxis



Temporal Coverage Results for 100 taxis

Conclusions

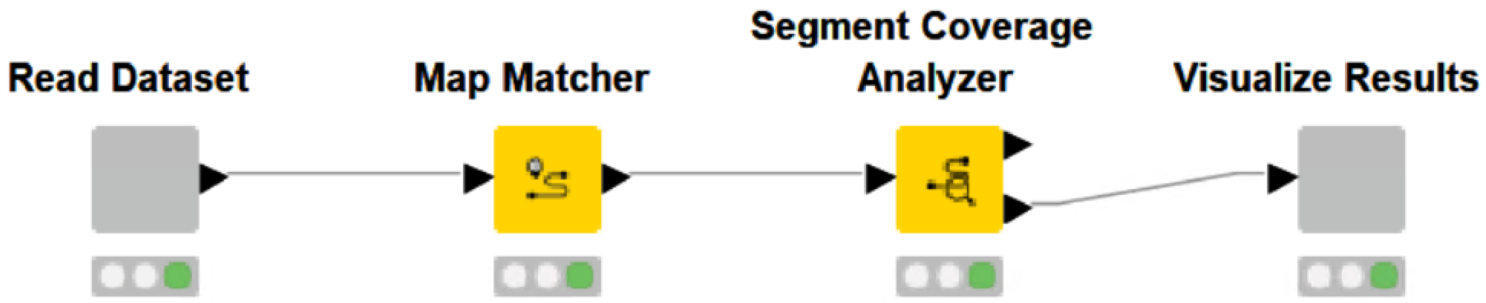
- Our case study showed that, in Porto, as few as 100 taxis can achieve a very high spatial coverage of the road network in 3 weeks.
- However, the temporal coverage might be inadequate to support VCS scenarios requiring very frequent sensings

Future works

- What if we recruit a different number (e.g.: 50, 200, 400) of taxis?
- Are there significant seasonal changes in the coverage dynamics?
- Replicate on different datasets/cities (e.g.: Rome, Beijing)

Replicability

- Replication package is available at doi.org/10.5281/zenodo.4773593
- The case study pipeline is implemented using the well-known open-source **KNIME** Analytics Platform, and an open-source extension we developed: **KNOT** (https://luistar.github.io/knot)



Vehicular Crowd-Sensing on Complex Urban Road Networks: a Case Study in the City of Porto

Vehicular Crowd-Sensing (VCS)

- Use vehicles as probes to collect contextual information on phenomena of interest:
  - Availability of on-street parking
  - Amount of pollutants in a city area
  - Presence of potholes on a street
  - And much more!
- Economically efficient, especially when exploiting high-mileage vehicles such as taxis

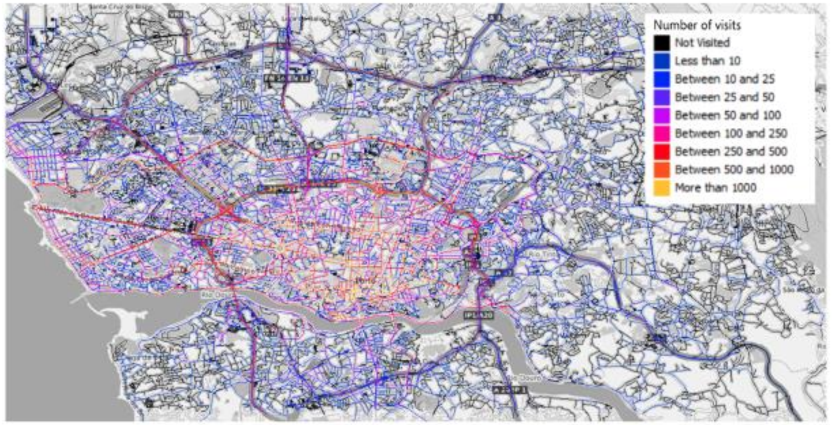
Yu TY., Zhu X., Maheswaran M. (2018) Vehicular Crowdsensing for Smart Cities. In: Handbook of Smart Cities. Springer, Cham. https://doi.org/10.1007/978-3-319-97271-8_7

The Case Study

- Goal: evaluating feasibility of using taxis as probes in VCS in cities with complex road networks.
- Leveraged a massive dataset of real taxi trajectories recorded in the City of Porto
- Measured the spatio-temporal coverage achieved by those taxis at a fine-grained level of road-segment

Dataset → Data Selection → Map Matching → Road network coverage calculation → Visualization of the results

Spatial Coverage Results



Temporal Coverage Results

Luigi Libero Lucio Starace – luigiliberolucio.starace@unina.it

Average Timegap Between Visits

- If a segment is visited by n vehicles at times $t_1, \dots, t_n$, the average timegap between subsequent visits for that segment is defined as

$$\frac{\sum_{i=1}^{n-1}(t_{i+1} - t_i)}{n-1}$$