ANT 2019, EDI40 2019 and Workshops
Special Session Track & Know

Track & Know

Big Data for Mobility Tracking Knowledge Extraction in Urban Areas

Pilot 1: Car Telematics for Insurance & Innovative Mobility Services

Leonardo Longhi

Sistematica

Mirco Nanni

Summary of contents

- Context
- Business cases of the Pilot
- Technical challenges & Initial results
- Conclusions

Context

Car telematics core business:

- **collect** data from telematics devices
- **develop** advanced solutions and algorithms for sophisticated data analysis

Typical applications:

- help insurance companies **assessing correct policy prices**
- provide services for the **management of accidents & facilitate communication** between companies and customers
- provide **services to car manufacturers** (identify driving styles and habits, create custom warranty programs, offer personalized services)

Demonstrator of the EU project Track & Know: mainly targets applications for

- car insurance
- electric vehicle mobility
- shared-mobility

Context

The demonstrator data

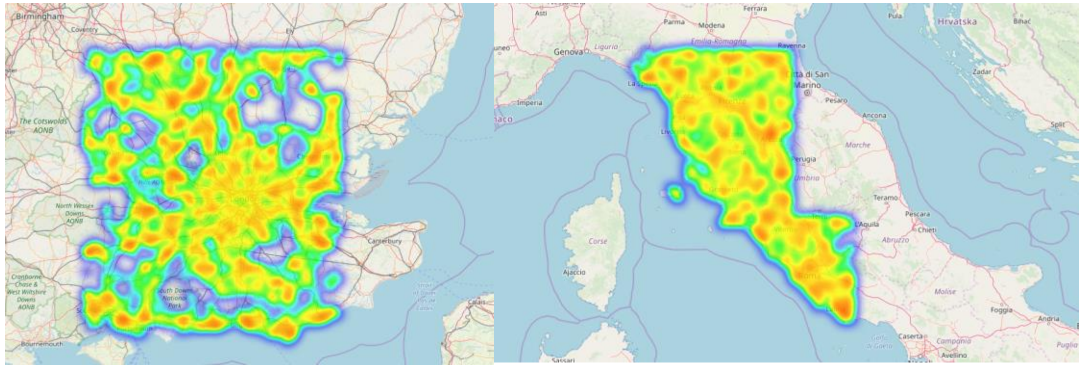
What

Traces of ~500k vehicles over 1 year, consisting of:

- **Positions**: WGS84 GPS position (latitude and longitude) related to anonymized vehicles (via anonymous IDs)
- **Events**: position data (as above) enriched with threshold-base labels describing motion events occurring in a given time stamp
    - harsh accelerations, brakings, cornerings, etc.
- **Crashes**: position data (as above) related to crash events
- **Car models**: info about car age, brand and model

Where

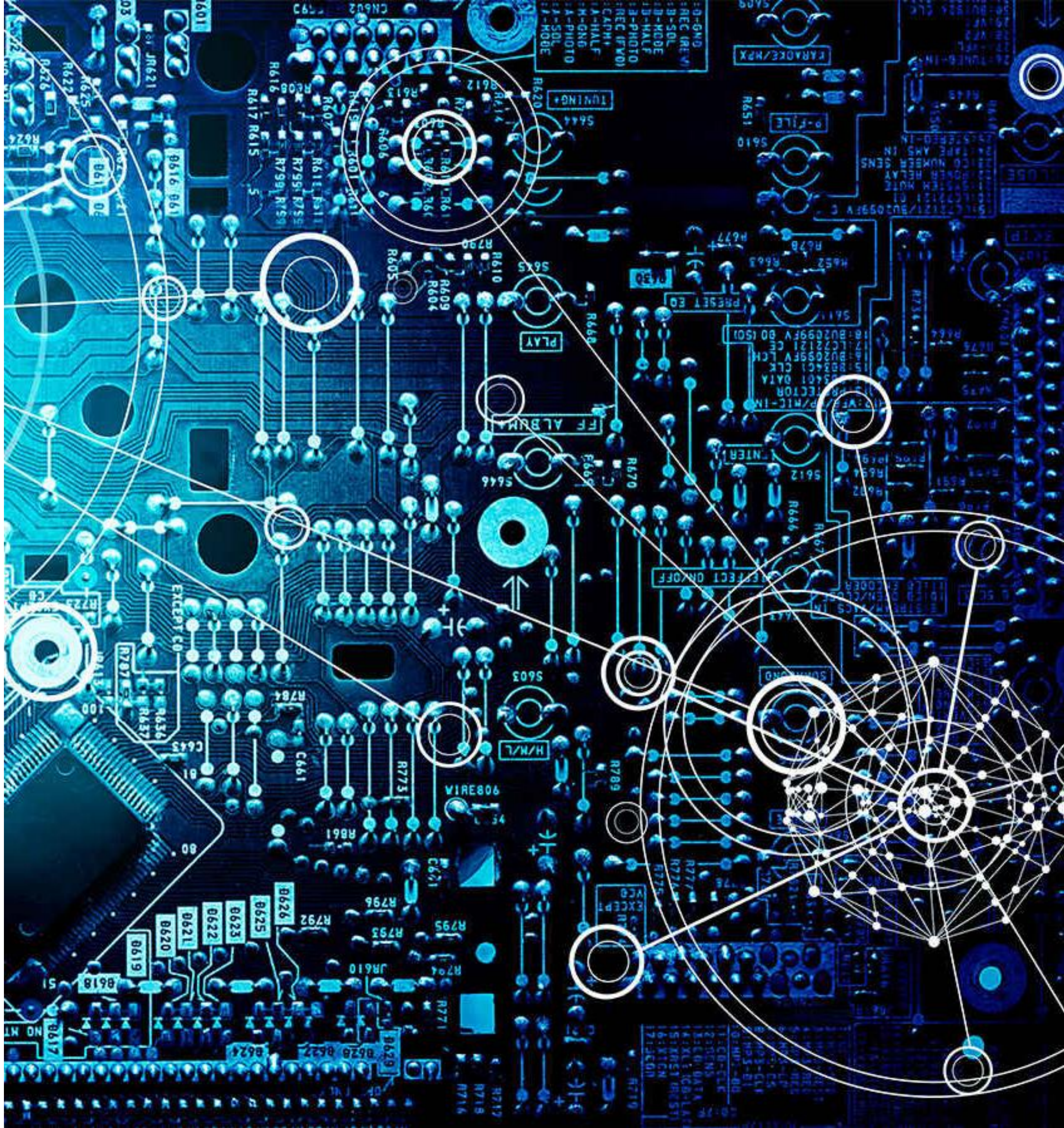
- Three different context types:
    - A very large city (London, UK, >8M inhabitants)
    - A moderately large city (Rome, Italy, 2.8M inhabitants)
    - A whole region, composed of variable-size cities (Tuscany, Italy, 3.7M inhabitants)



Track & Know

Business Cases

- Car insurance
- Electric car
- Car pooling



Car Insurance Business Case

Car insurance

One of the most important application fields of car telematics

Movement data is typically used to provide several services to end users:

- pay-as-you drive contracts
- anti-theft control
- prompt emergency rescue in case of accidents

Fundamental task of the demonstrator: find the most appropriate policy pricing for a customer

- consists in a trade-off between profit and competitiveness
- usually based on estimates of customer’s risk of having accidents in the near future

Car insurance

Main objectives of the demonstrator:

- **Predicting the Customer’s risk score**
  - given a car insurance customer, provide a risk score relative to the near future, e.g. the next year or the next three months
  - expected to be dependent on how the customer drives and the conditions of the surrounding environment (traffic, etc.)
- **Inferring risk mitigation strategies**
  - given a car insurance customer and her risk score, identify the characteristics of her driving that mostly determine her risk score
  - prescriptive viewpoint: provide the customer indications of how to improve her risk score
    - benefits for her → safety and insurance costs
    - benefits for the insurance company → costs for accidents

Electric Car Business Case

Electric cars

- EVs industry and their adoption is expanding in most EU countries
- Need to understand effects of switch from fossil fuel to Evs
  - pros and cons
  - habit changes required to each user

Overall target of the business case:

- provide objective, data-driven means to measure such aspects
- let companies and individual users to evaluate the ease of conversion to EV mobility

Electric cars

The main focus of this demonstrator: understanding the impact of EV switching on the individual

- **Estimate Costs/Benefits of EVs for the individual**:
  - given an individual customer with her mobility history, evaluate her costs or savings in terms of money and time in case of switching towards an EV.
  - That should take into consideration daily mobility needs, and therefore usual paths, as well as charging point availability (with corresponding detours from the fastest trip) and charging times.

Key tasks:

- understand the mobility needs of the users both at the individual and at the collective level
- derive the consequences of the limited autonomy and longer recharge times of EVs compared to traditional vehicles
  - these two factors might make some fossil fuel-based trip impossible for an EV
  - finding EV-compliant alternative routes for the travels of a user and measuring their efficiency is a basic need

Car Pooling Business Case

Car pooling

Basic measure to reduce traffic congestions, the derived environment footprint, and the personal costs (fuel consumption & maintenance)

- Simple, ecologically sustainable and alternative paradigm of mobility
- Growing market: expected to grow at a compound annual growth rate of 8% from ~22 M users in 2017 to ~47 M in 2025 (Frost and Sullivan)

Car telematics companies are venturing into providing technological solutions for Shared Mobility, including:

- advanced fleet management
- insurance telematics for operators in the Mobility sector
- rental companies.

Several common points with EVs mobility:

- both would greatly benefit from a clearer understanding of the pros, cons and habit changes that each user is going to experience when she joins it

Car pooling

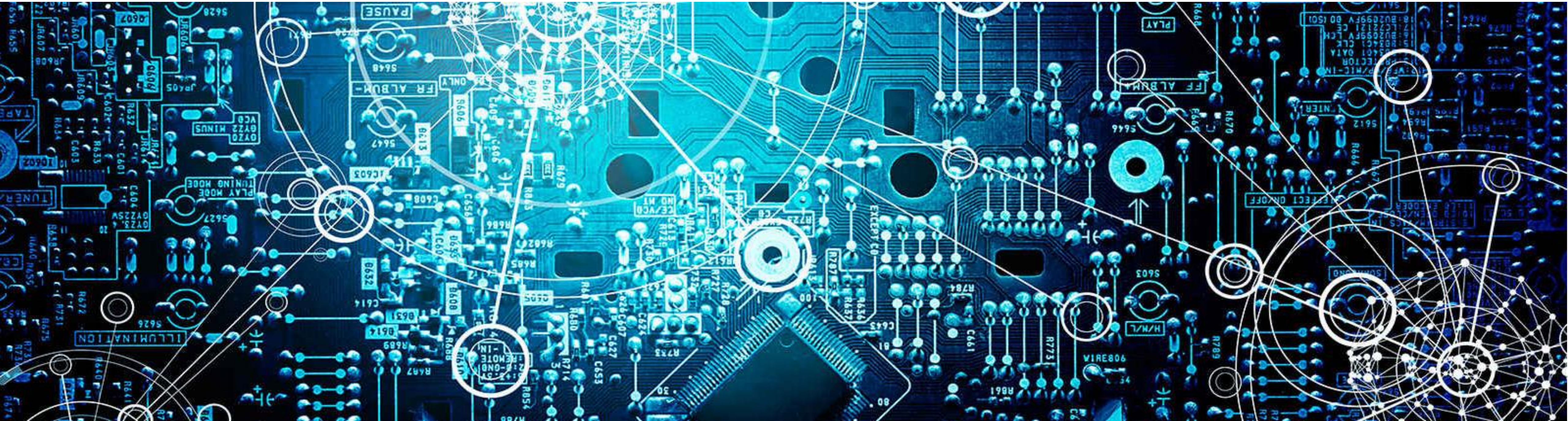
Main focus of the demonstrator:

- **Estimate Costs/Benefits of car/ride-sharing for the individual**
  - given an individual customer with her mobility history, evaluate her costs or savings in terms of money and time in case of adoption of car/ride-sharing
  - That should take into consideration daily mobility needs, and therefore the importance that each trip has in the overall mobility demand of the individual

- Synchronizing with other users usually affects the efficiency of the travel in terms of time delays and slight changes in the itinerary.
- Finding car/ride-sharing alternative routes for the travels of a user and measuring their efficiency is a basic need

Technical Challenges

(...and some initial results)



Challenges

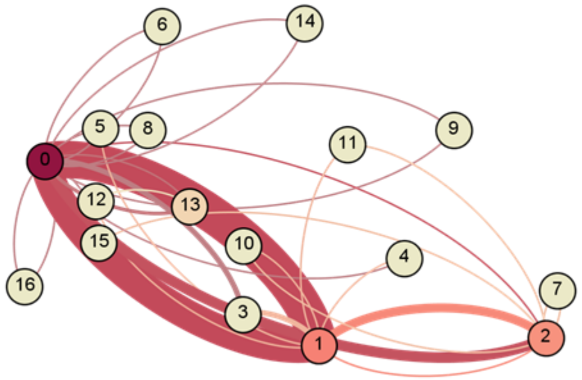
- ***Preprocessing, or Dealing with data quality issues***
  - Inevitably quality issues due to a number of causes
  - Ad hoc mitigation countermeasures are needed
  - E.g. reconstruction of trips out of raw GPS points recorded for each device:
    - determining the start and end of each trip in a precise way requires specific heuristics;
    - reconstructing the detailed path (which roads were traversed) might require map matching and similar solutions;
    - singling out noise and errors of moderate size in the data can be very challenging.

- ***Semantic enrichment***

Adding semantics to the raw data and its aggregations (e.g. trips) can be realized

- joining **external information**, e.g. attaching to each GPS location the weather conditions, local traffic and points-of-interest around it
- inferring it **from the data itself**, e.g. identification of recurrent trips or the spatial aggregation of driving events aimed to identify areas were some specific behaviours are more frequent

Challenges



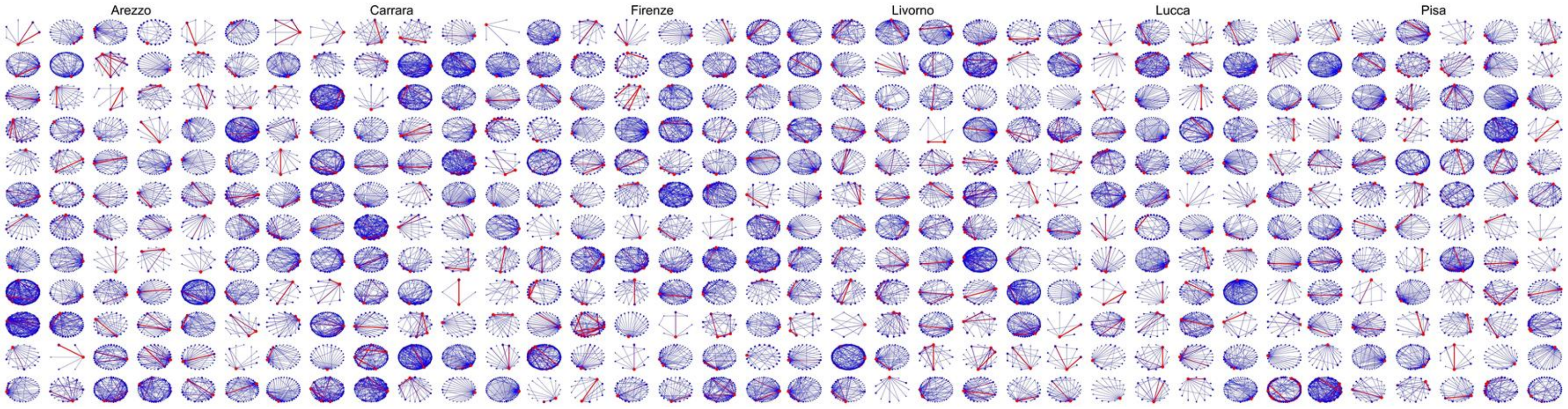
- ***Individual-centered mobility modeling***

Semantics related to the meaning that the different parts of the mobility have for the individual:

- recurrent vs. systematic trips, frequent locations vs. single visit ones, transit locations vs. long stays, etc.

Individual Mobility Networks: model the mobility of the individual as a whole, creating a single, complete picture of it

Challenge: integrating as much information as possible in this formalism and inferring from it mobility indicators useful for the predictive/prescriptive purposes of the demonstrator



Sample IMNs for 6 different cities in Tuscany, with O/D areas only. Apparently, no clear visual feature characterize different cities.

Challenges

- ***Prediction of (crash) risk probability***

Difficult task:

- Accidents are (in statistical terms) rare events
- Lack of a clear set of predictive indicators to adopt

Approach under development: consider the driving behaviour of the user + the types of environment she usually traverses (road categories, weather during driving time, etc.)

clf	features	f1-score	precision	recall	test_accuracy	train_accuracy
RF	all	0.659024	0.709972	0.655093	0.712644	0.815271
RF	evnt	0.641626	0.734188	0.663095	0.672414	0.757389
RF	traj	0.636443	0.723219	0.655688	0.669540	0.748768
RF	trev	0.603573	0.651994	0.609672	0.658046	0.795567

Crash prediction performances of Random Forest models on various sets of basic features: trips, events, car model. All feature categories bring some improvement

- ***From prediction to prescription***

Achieving a good prediction accuracy often conflicts with the understandability of the predictive model.
Objective: extraction of risk mitigation guidelines for the user (the driver), i.e. understanding which factors made a driver a risky one
The project will explore “explainable AI” methodologies community, aimed to extract from a black-box model an explanation for each prediction

Challenges

- ***Models Transferability***

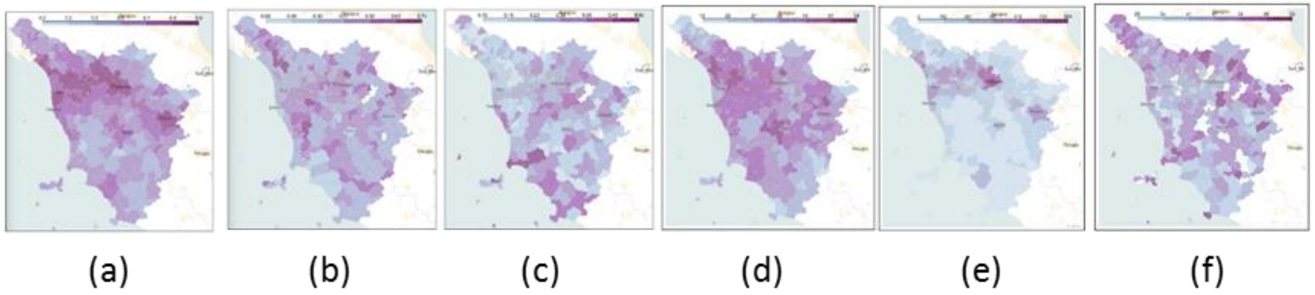
The mobility models are expected to be highly dependent on the specific geographical area under study

At the same time, not all areas of interest for the demonstrator are equally well covered by data

This calls for methodologies that make it possible to adapt models built in data-rich areas to less rich ones

- geographical instance of the general transfer learning problem

Ongoing work starts from recognizing relations between model transferability and features of the territory



Spatial distribution of sample city indicators over the Tuscany region:
(a) Population entropy, (b) Modularity,
(c) Fit to gravitation model, (d) N. of nodes in IMNs,
(e) Roads concentration, (f) Traffic concentration.

Track & Know

Challenges

- ***Defining proper notions of electrificability and shareability of individual mobility***

How much is the mobility of an individual compatible with alternative transport modalities?

- in our case, EVs and car sharing/pooling, both with their own constraints
- realistic evaluation of the effort required on behalf of the user to adapt her whole mobility is needed:
    - Evs: changing times and routes to intercept charging stations when needed
    - Car sharing/pooling: change times of travels, or even reschedule part of them.

Challenges:

- providing such compatibility measure definitions
- building the tools for computing their values

Conclusions

- Several **opportunities** to transform the raw mobility data collected by telematics companies into insights and valuable services
- Technical **challenges** require improvements of current research state-of-art
- Preliminary **results** show promising signals of meaningful solutions for the identified problems

- The Track & Know platform will provide the development and testing environment for the solutions envisioned

- Stay tuned!

THANK YOU!

Mirco Nanni
mirco.nanni@isti.cnr.it
ISTI-CNR, Italy