

Mobility Data Mining

Case study su “urban human mobility”

Big Data per la mobilità

- Fonti eterogenee e complementari di informazione su posizione e spostamenti:

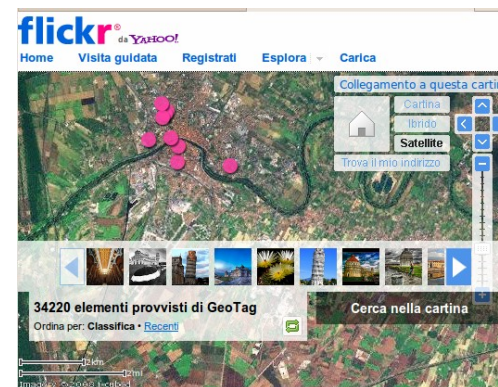
- Telefonia mobile GSM



- Sistemi tracciamento / navigazione GPS



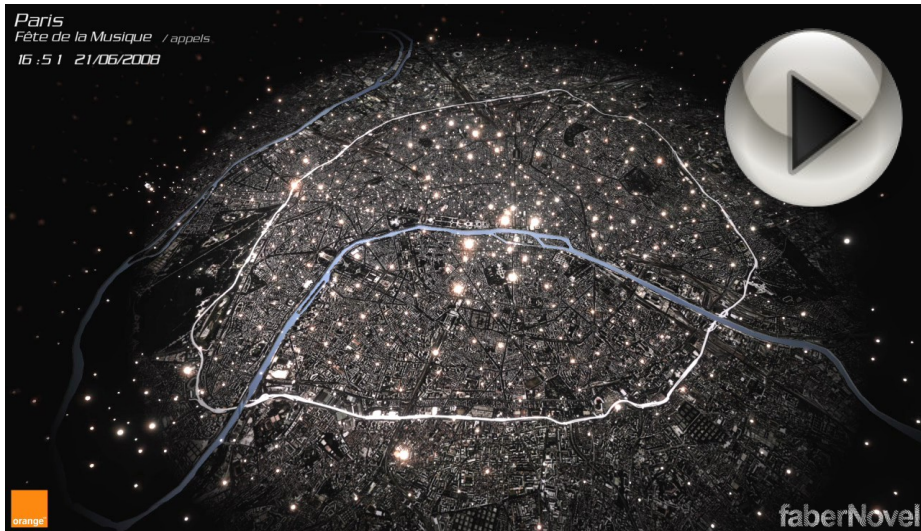
- Social networks per la condivisione dei propri dati (Flickr, FourSquare, Gowalla, ecc.)



Dati GSM

- Indicatori di presenza e movimento

Dati chiamate



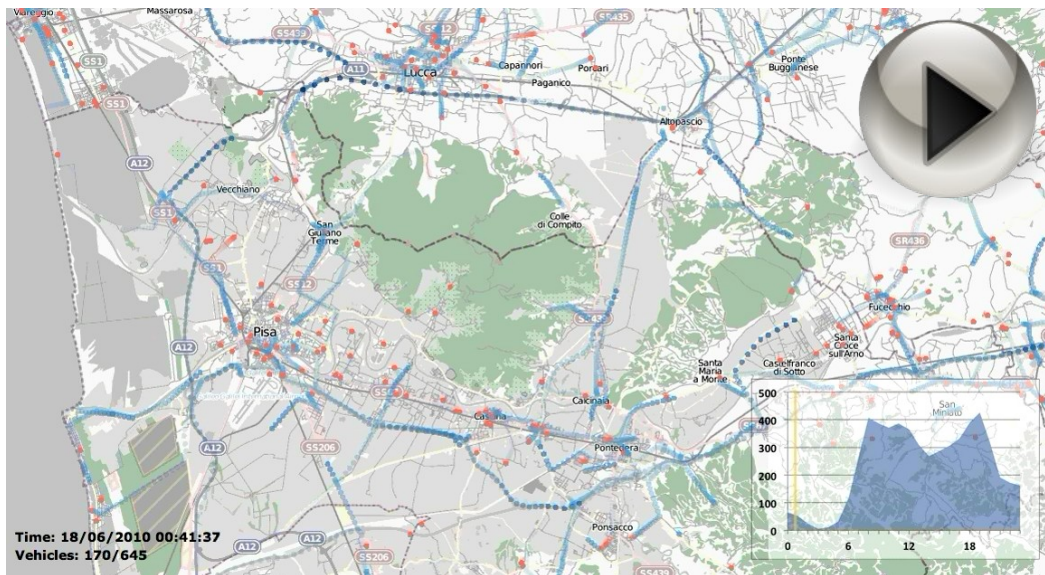
Dati hand-over



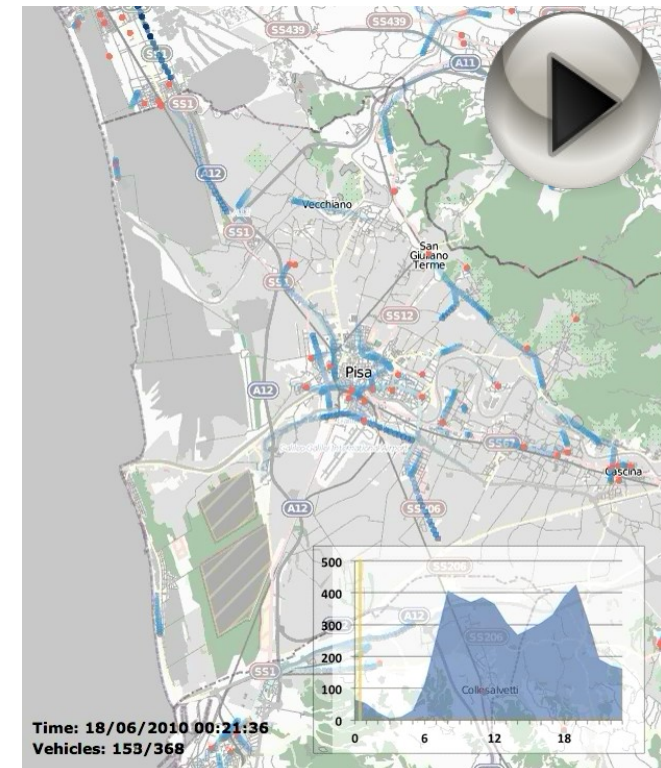
Dati GPS

- Indicatori dettagliati del movimento

Spostamenti su scala provinciale/regionale

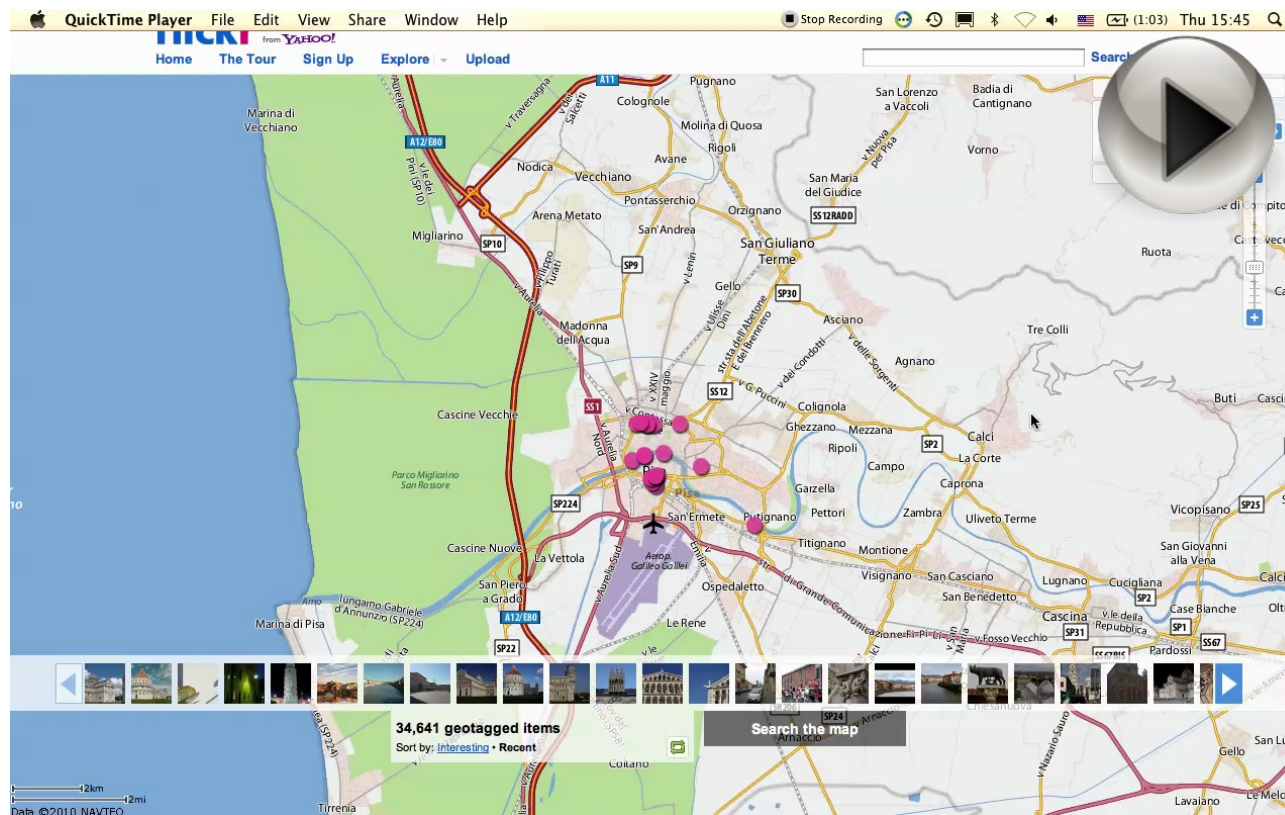


Spostamenti su scala cittadina



Dati da Social Networks

- Tendenza in crescita: condividere i propri dati (posizione, fotografie, commenti)



Sintesi

- Descrizioni complementari e a livelli di granularità molto variabile della mobilità umana
- Grandi opportunità
 - Possono nascondere una conoscenza approfondita, aggiornata e oggettiva del fenomeno mobilità
- Grande sfida
 - Dati complessi e caotici, che richiedono un trattamento sofisticato per estrarne conoscenza e risposte

Big Questions

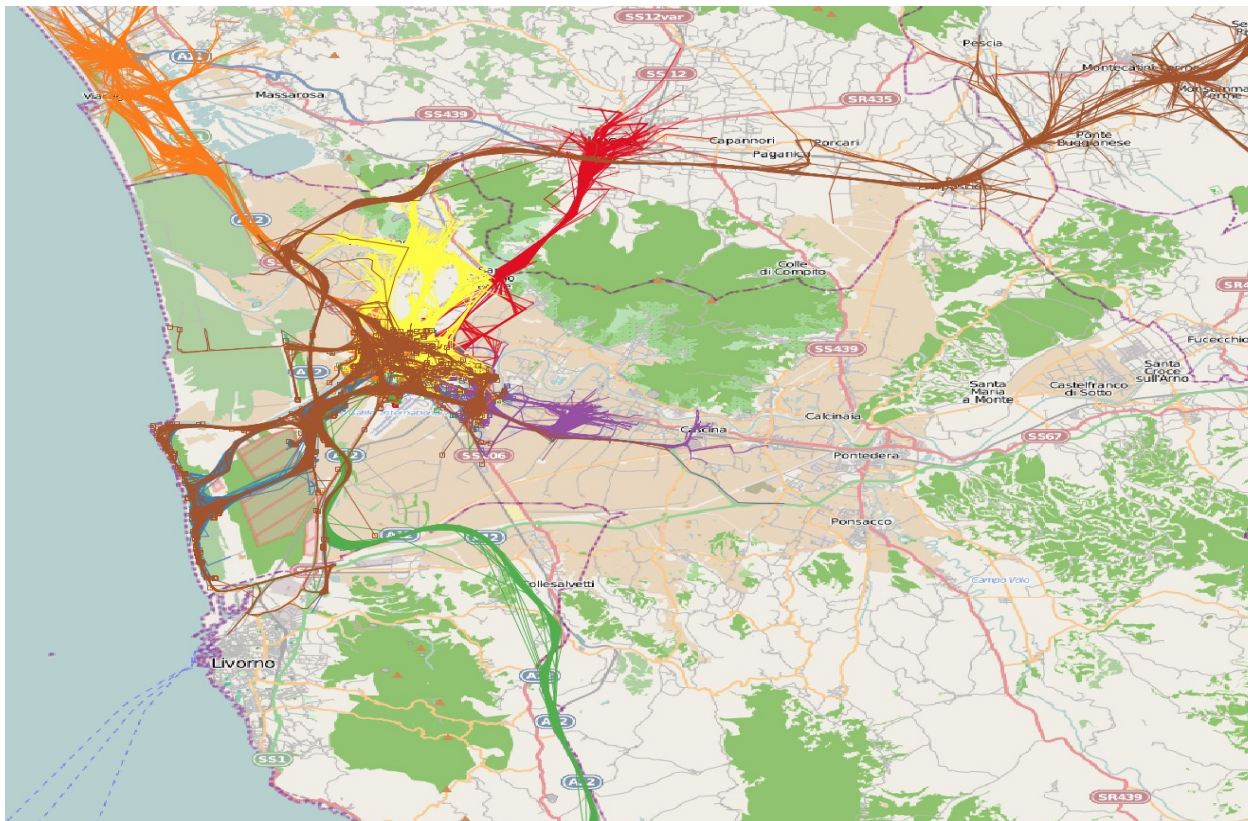
- Opportunità di affrontare un ampio spettro di quesiti legati alla mobilità
- Alcuni esempi trattati:
 - 1) Quali sono le modalità di accesso alla città?
 - 2) Quale impatto ha la mobilità sistemica?
 - 3) Quale (ri)configurazione di confini amministrativi si adatterebbe meglio alla mobilità dei cittadini?

BQ1: accesso alla città

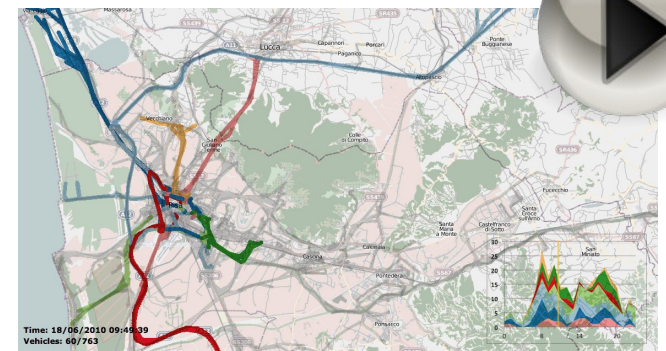
- Quali sono le modalità con cui i veicoli confluiscono nella città?
 - Percorsi di accesso
 - Area di provenienza
 - Orari
- Ricerca di una risposta analizzando la storia dei veicoli

BQ1: accesso alla città

- Emergono alcune modalità di accesso comuni

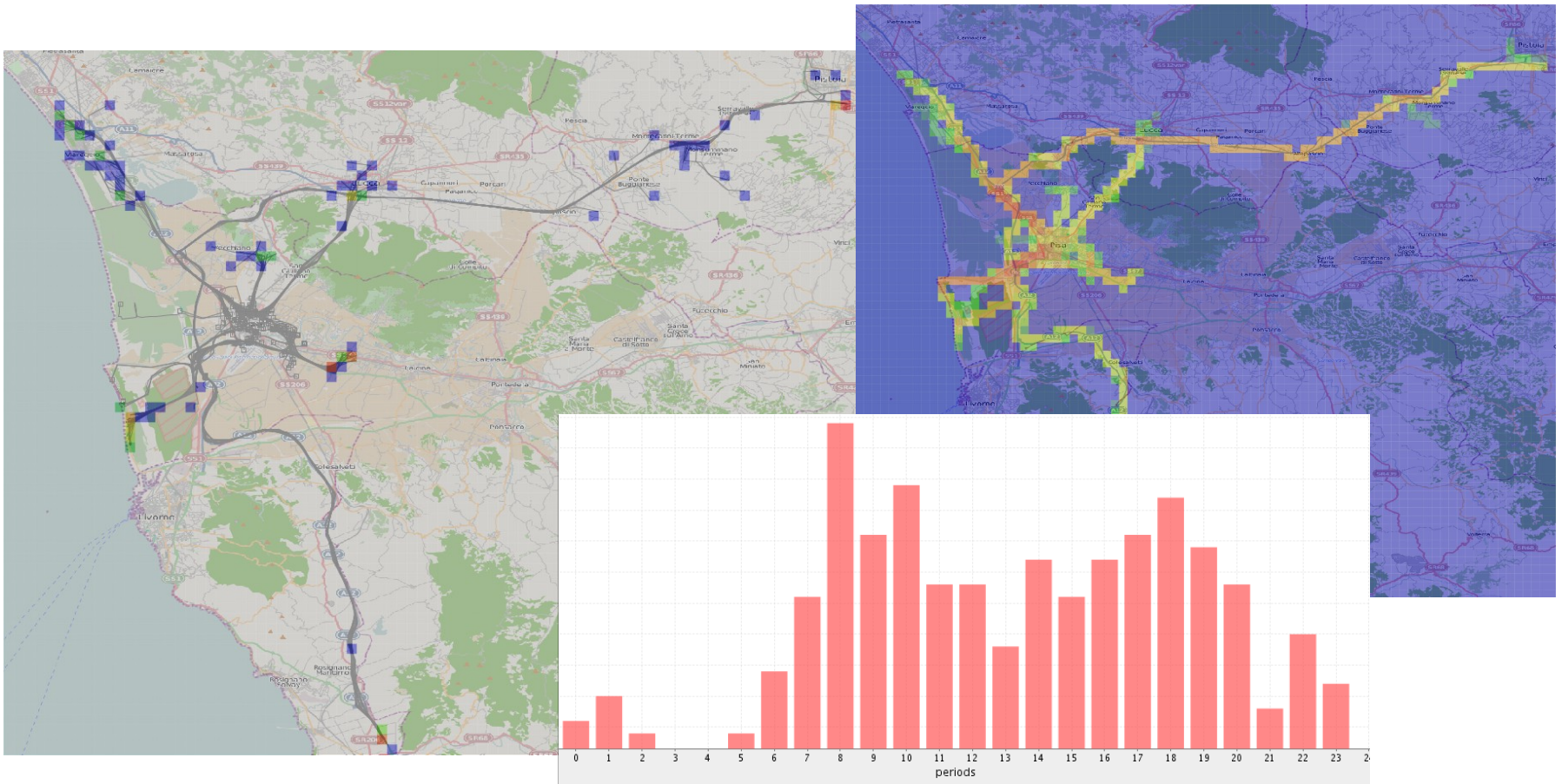


Accessi a Pisa da tutta l'area

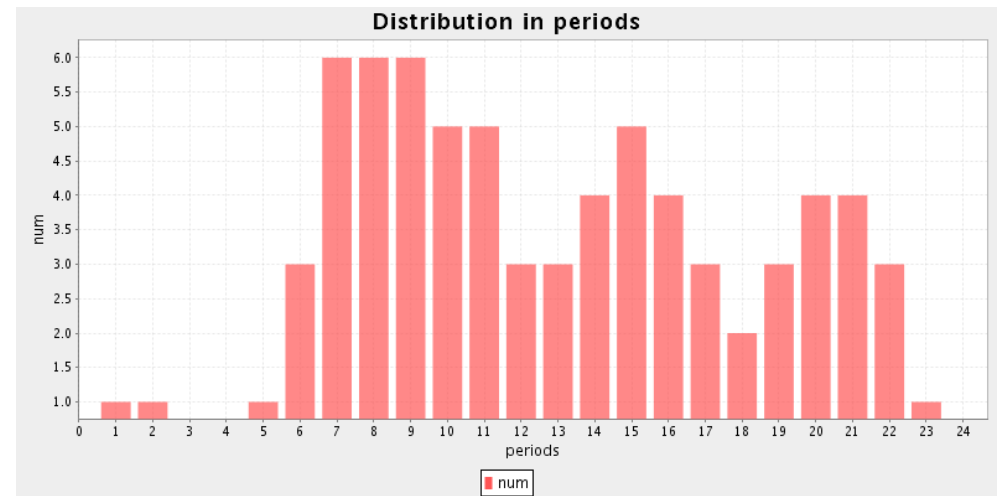
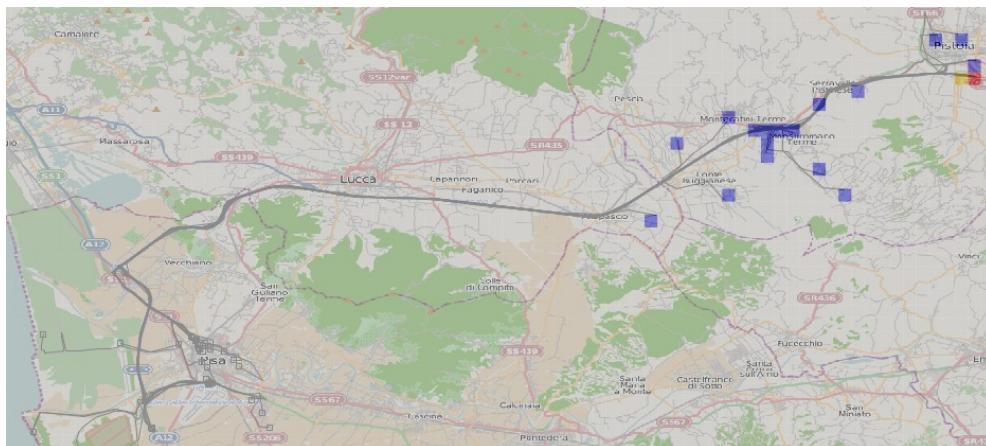
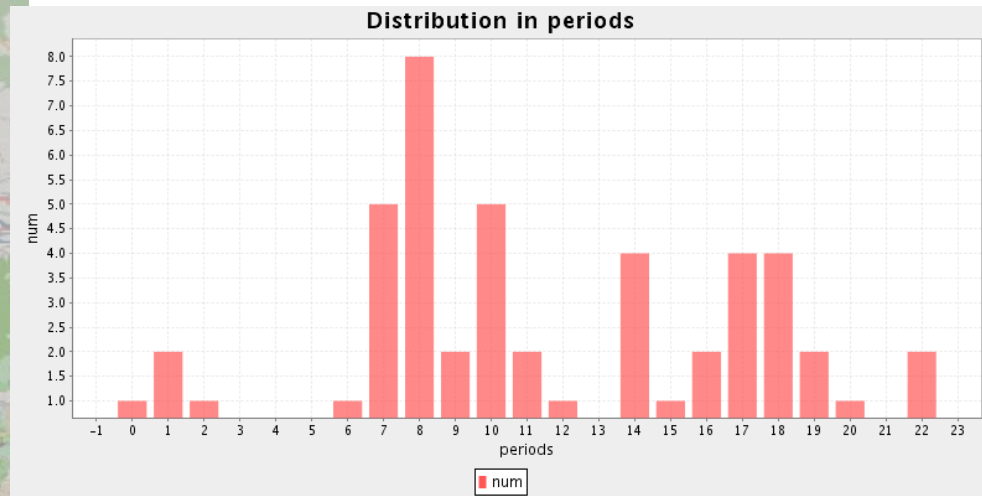
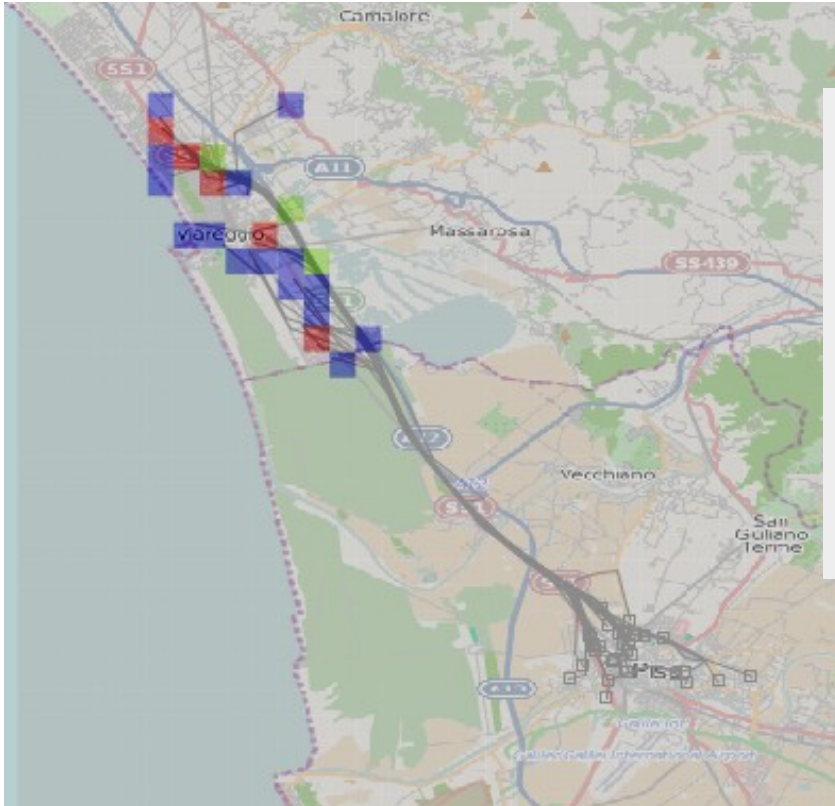


BQ1: accesso alla città

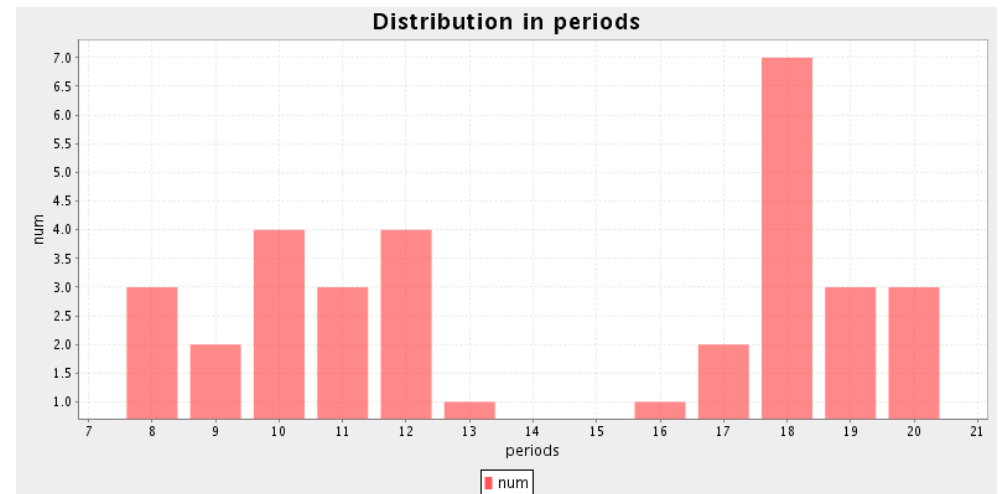
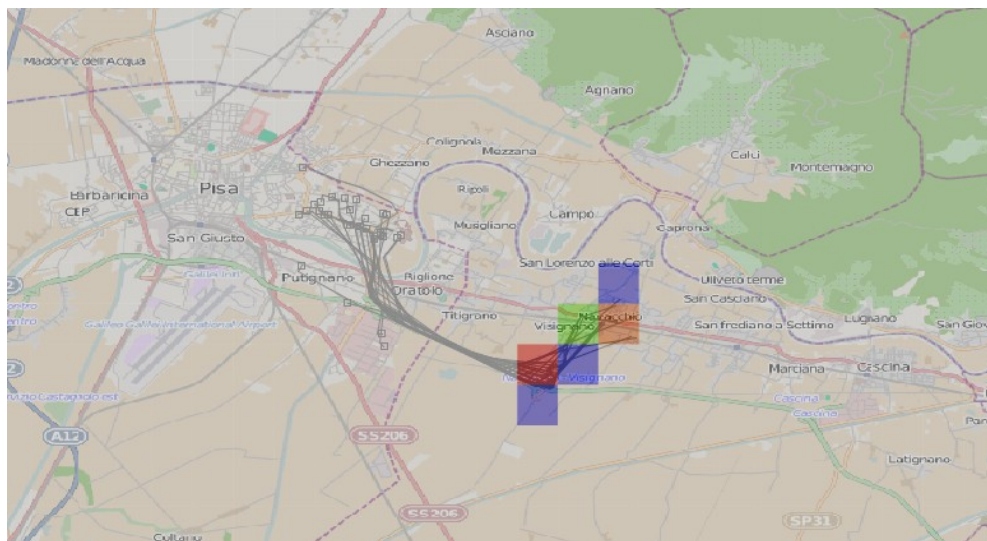
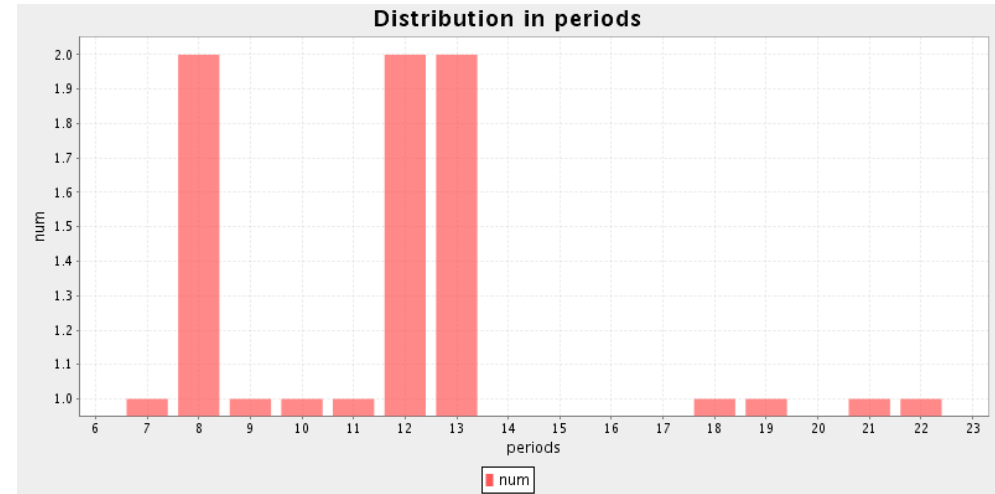
- Provenienza e distribuzione oraria dei “profili di accesso”
- Visione complessiva:



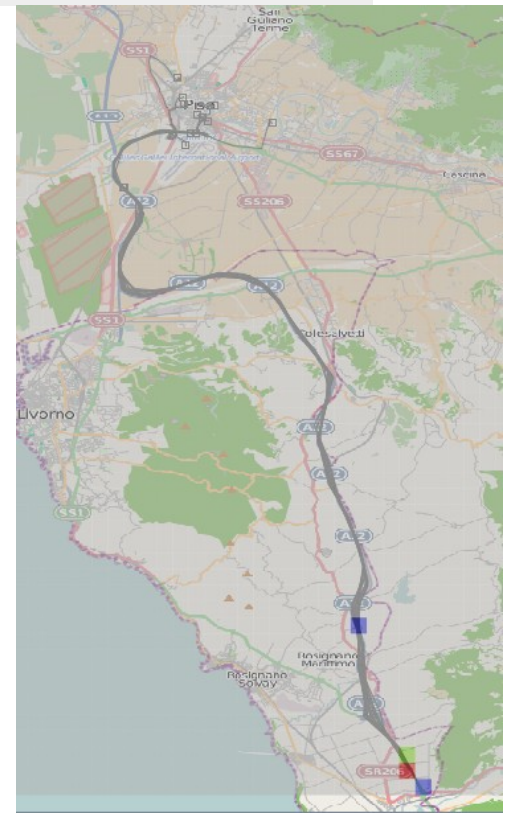
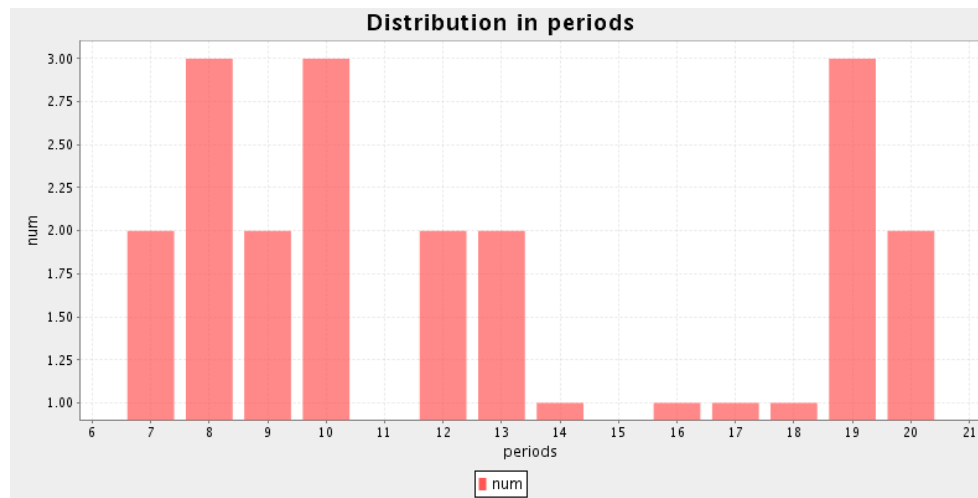
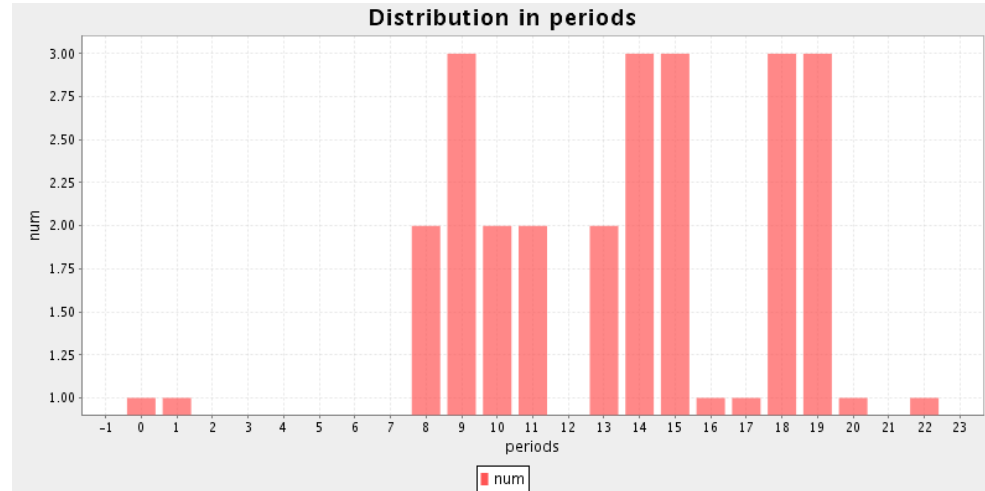
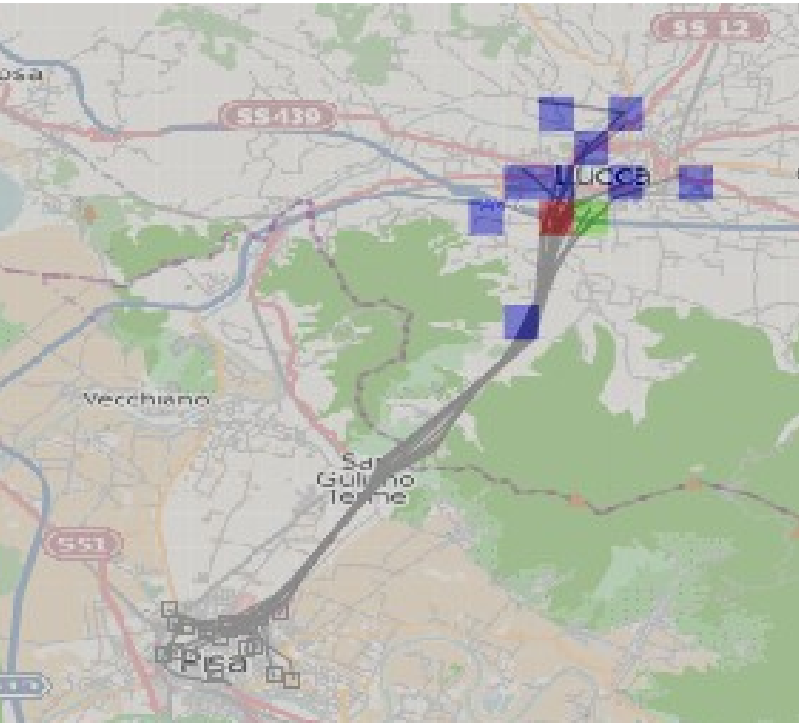
Profili 1 e 2



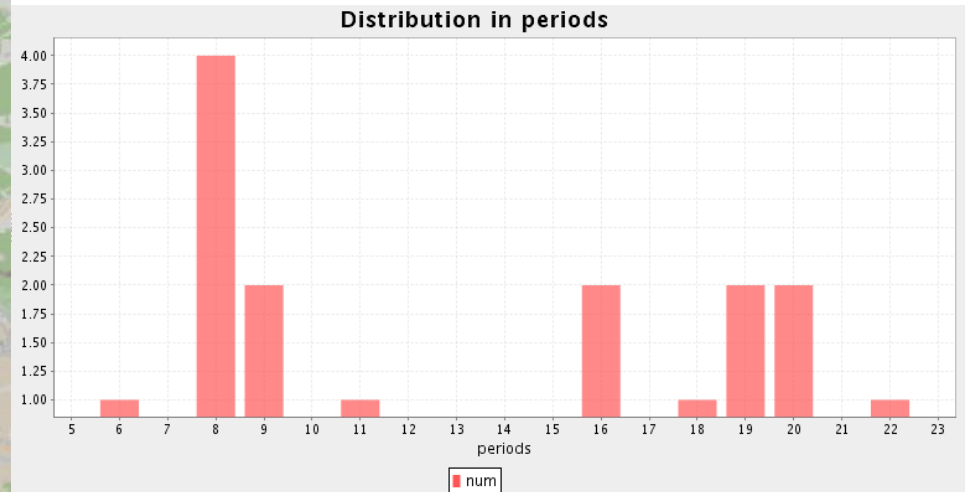
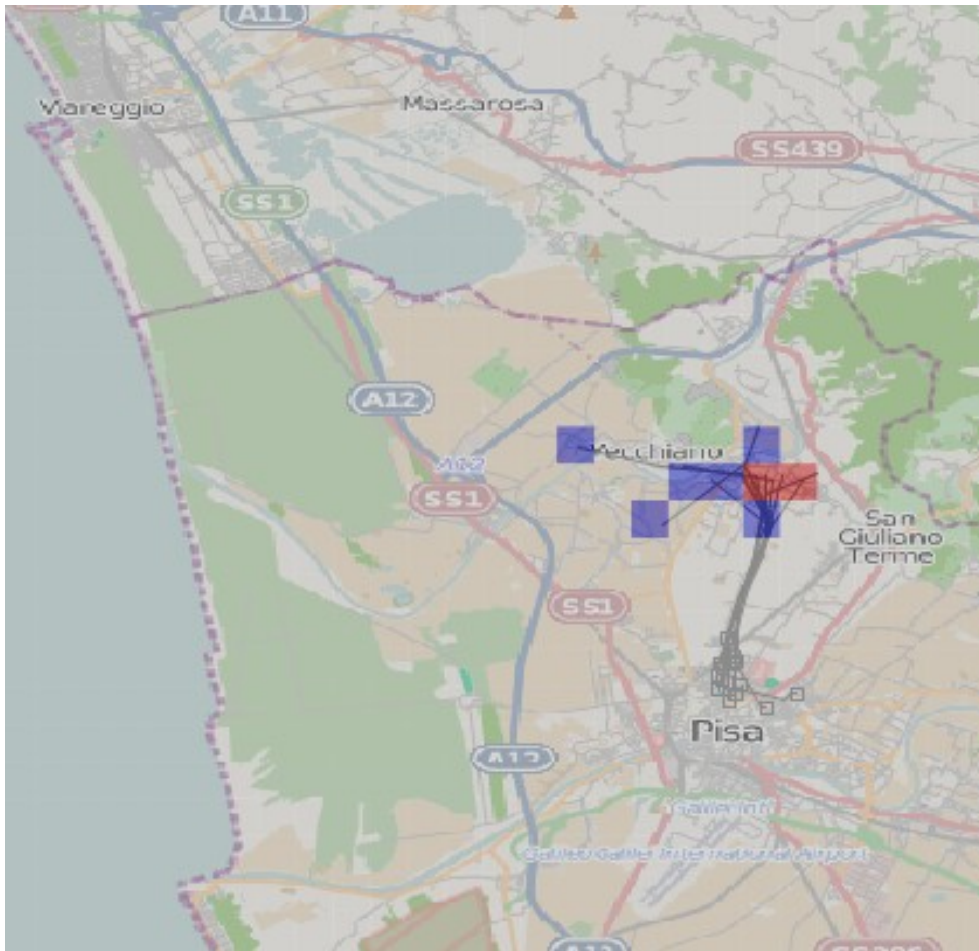
Profili 3 e 4



Profili 5 e 6

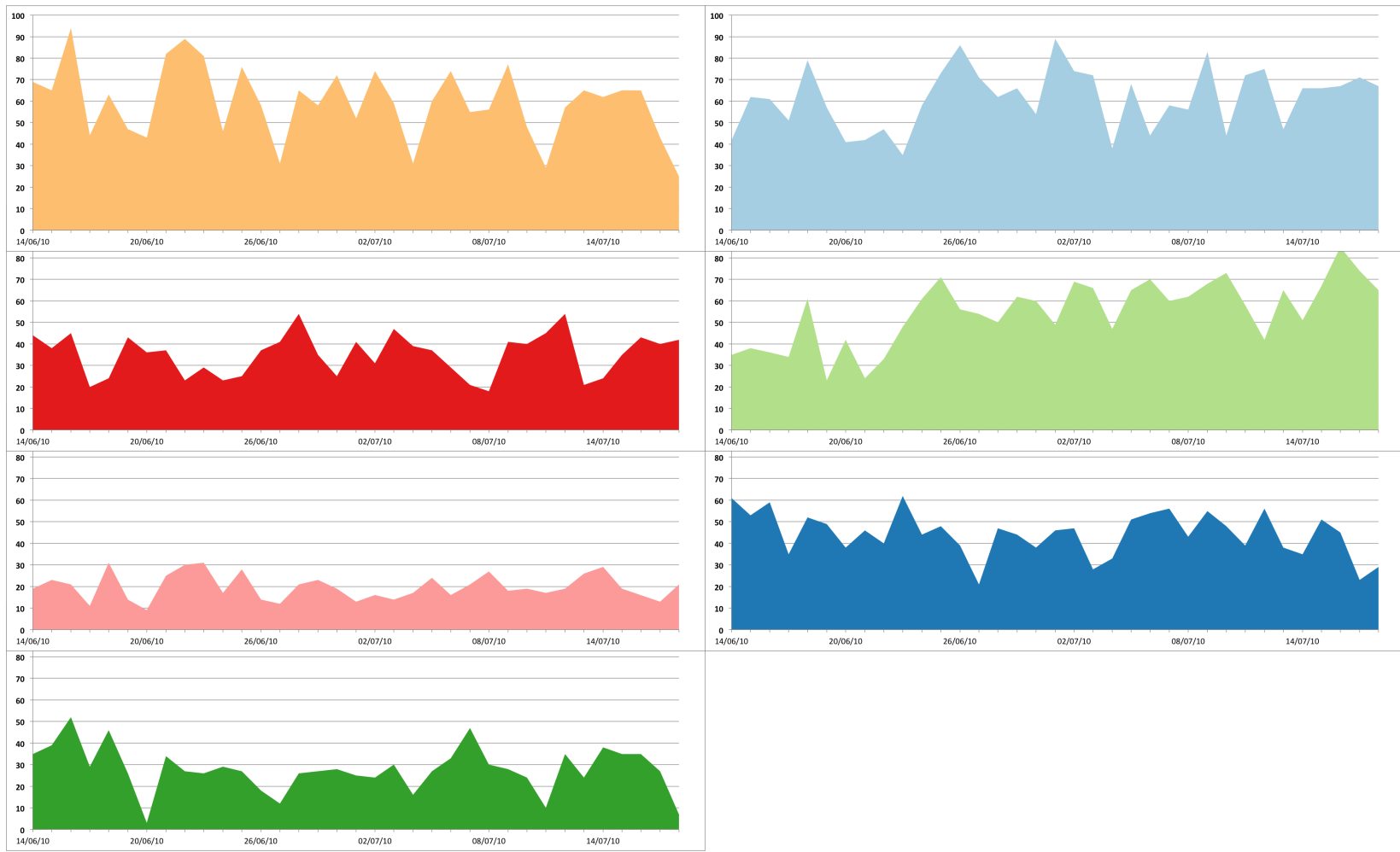


Profilo 7



BQ1: accesso alla città

- Quanto sono persistenti questi “profili di accesso” nel corso della settimana?

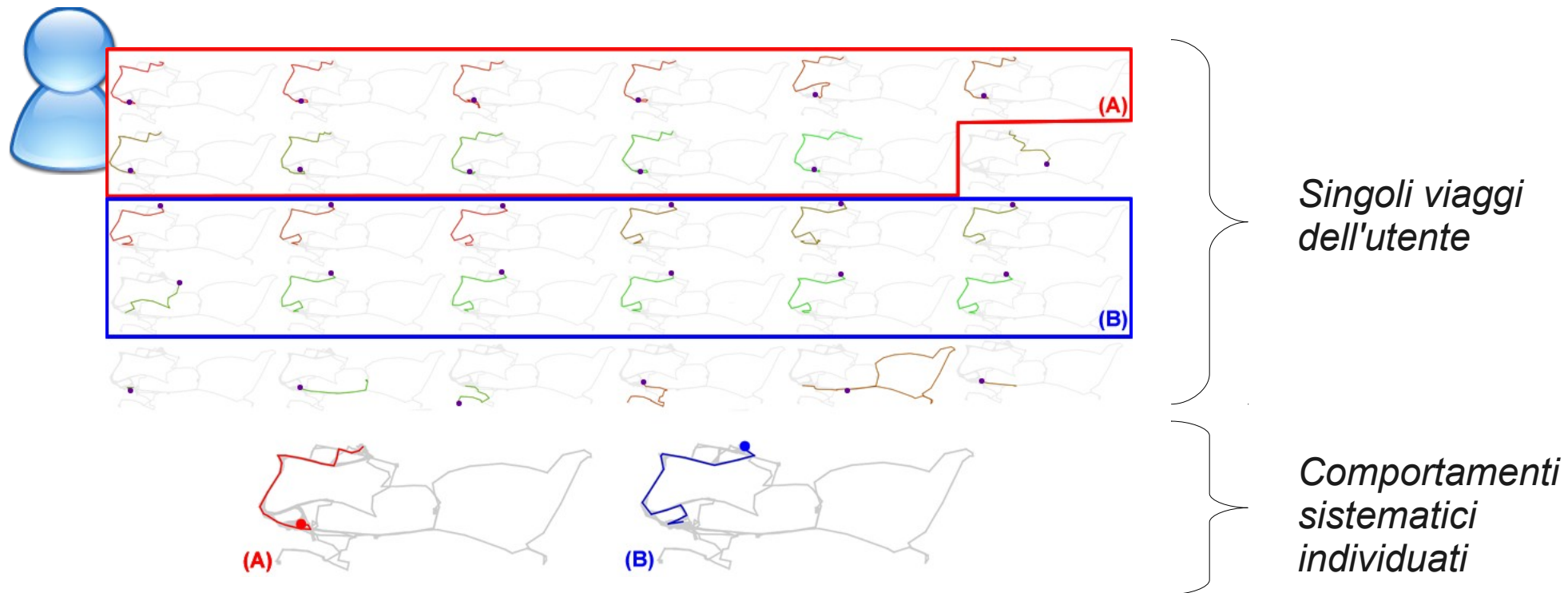


BQ2: mobilità sistematica

- Quale impatto ha la mobilità sistematica sui profili di accesso appena visti?
- Mobilità sistematica
 - Spostamenti che fanno parte della routine abituale di un utente
- La profondità storica e individuale dei dati ci consente di riconoscere la sistematicità

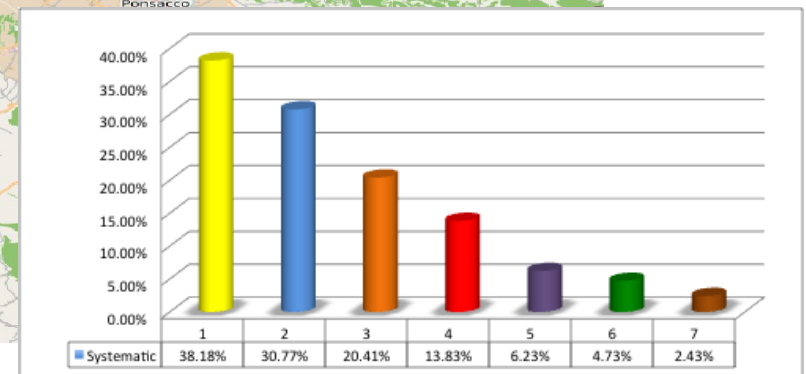
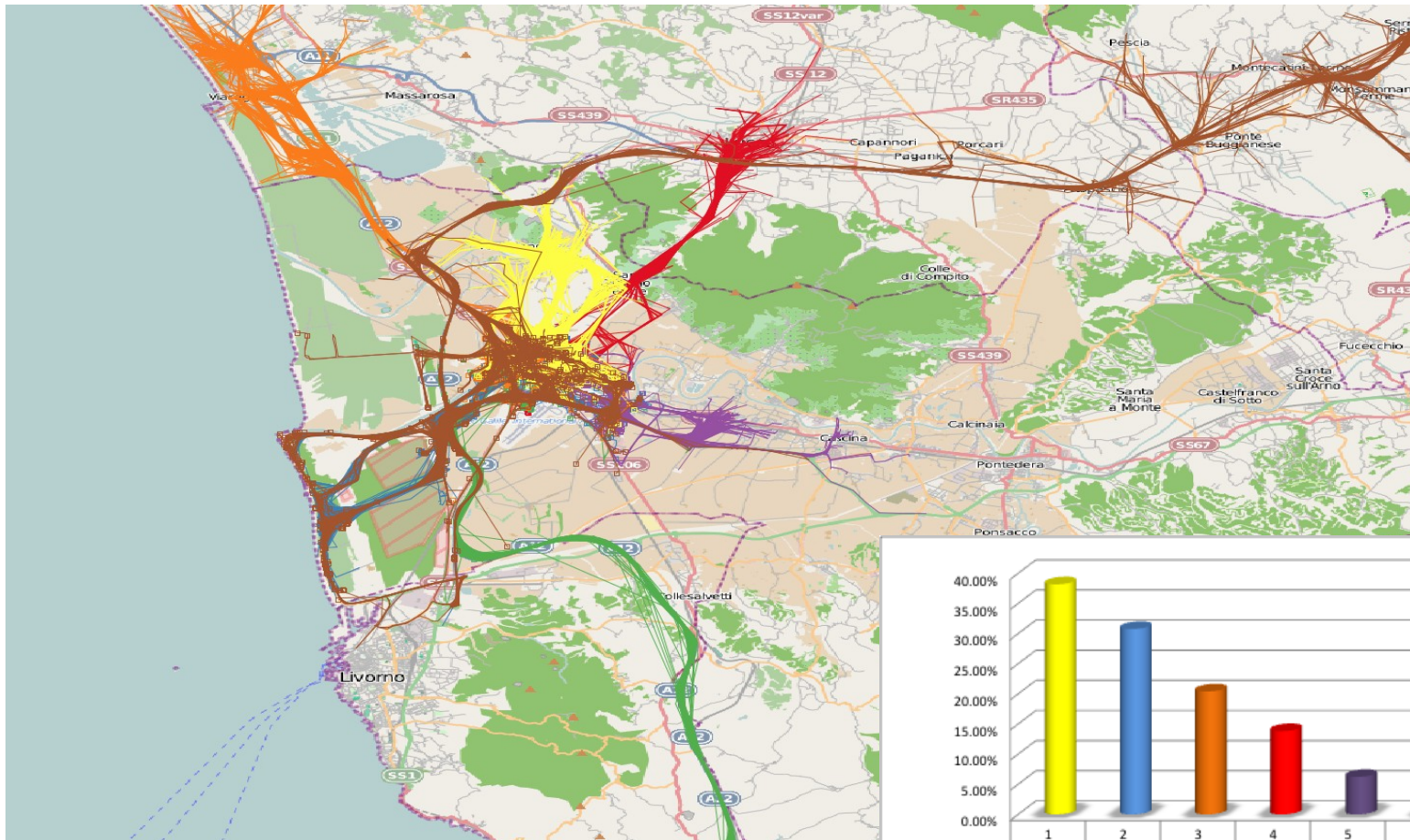
BQ2: mobilità sistemática

- Esempio di identificazione della sistematicità individuale:



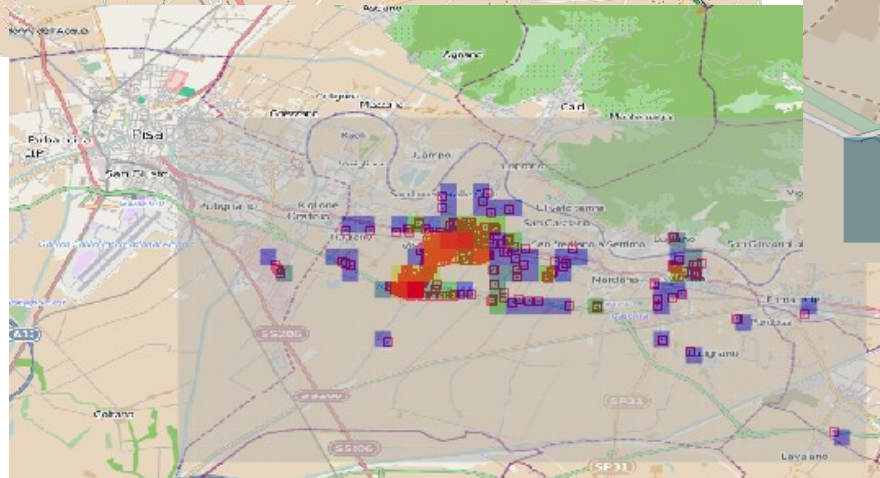
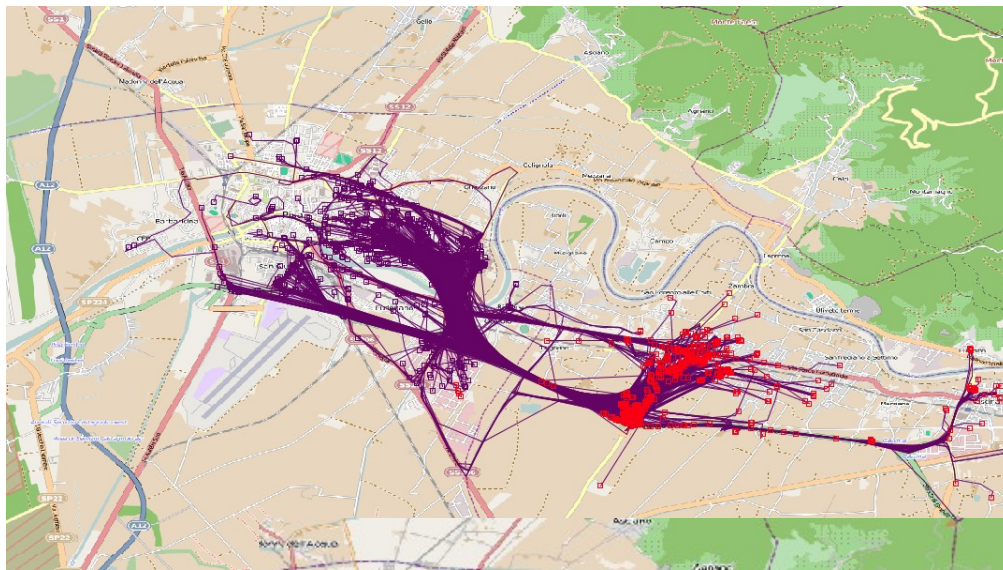
BQ2: mobilità sistemica

- Distribuzione degli spostamenti sistemici nei “profili di accesso” alla città:



BQ2: mobilità sistemica

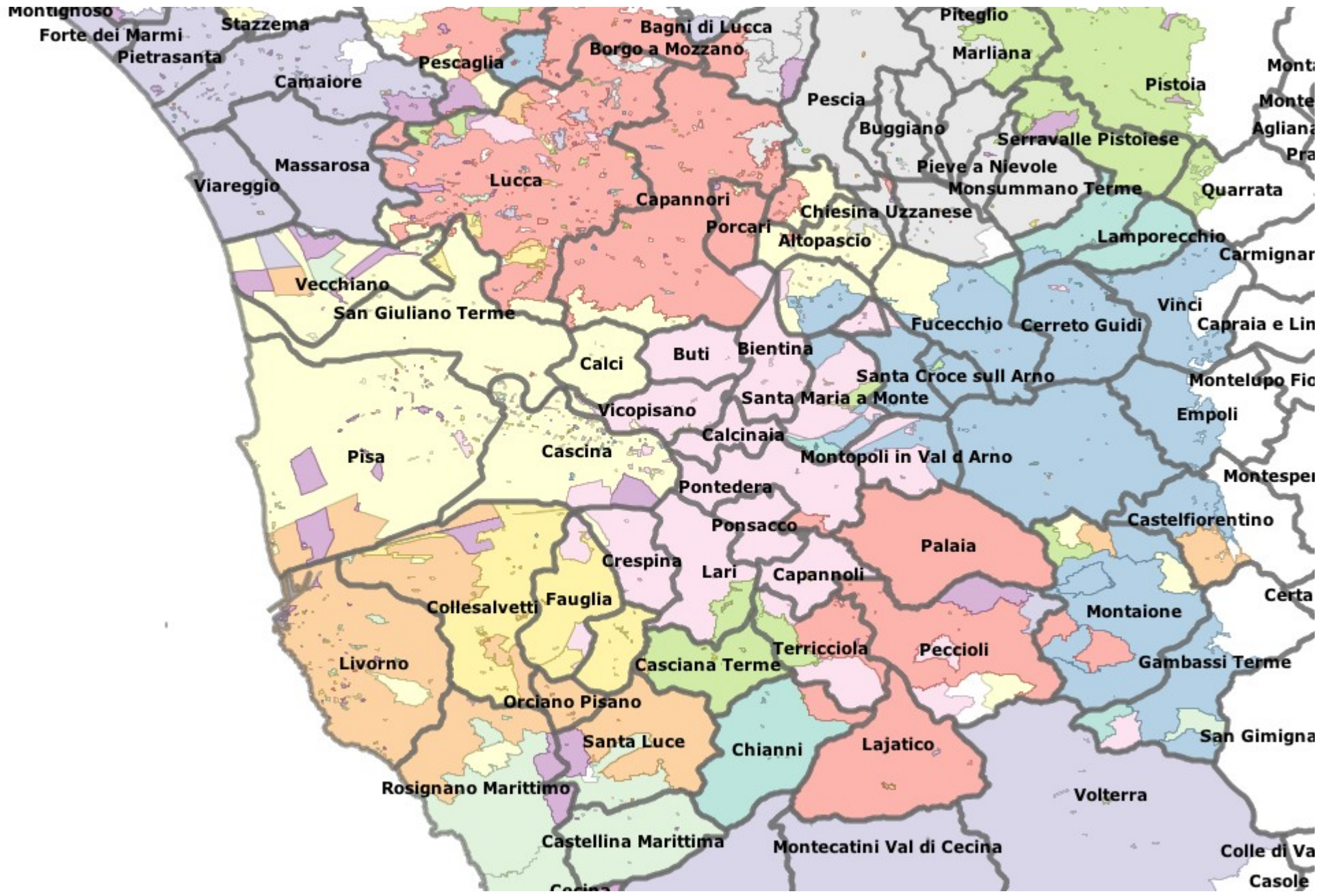
- Anomalia: scarsa sistematicità Navacchio → Pisa



BQ3: (ri)configurazione dei confini

- Quale disposizione dei confini amministrativi di un'area si adatterebbero meglio alla attuale mobilità dei cittadini?
- Principio generale:
 - aree limitrofe con grandi scambi di flusso veicolare dovrebbero appartenere ad uno stesso agglomerato

BQ3: (ri)configurazione dei confini



Dietro le quinte

- Estrarre informazioni utili e risposte a quesiti analitici richiede strumenti di analisi adeguati
- Matlas
 - Piattaforma per l'analisi di dati di mobilità, risultato di progetti europei, nazionali e regionali

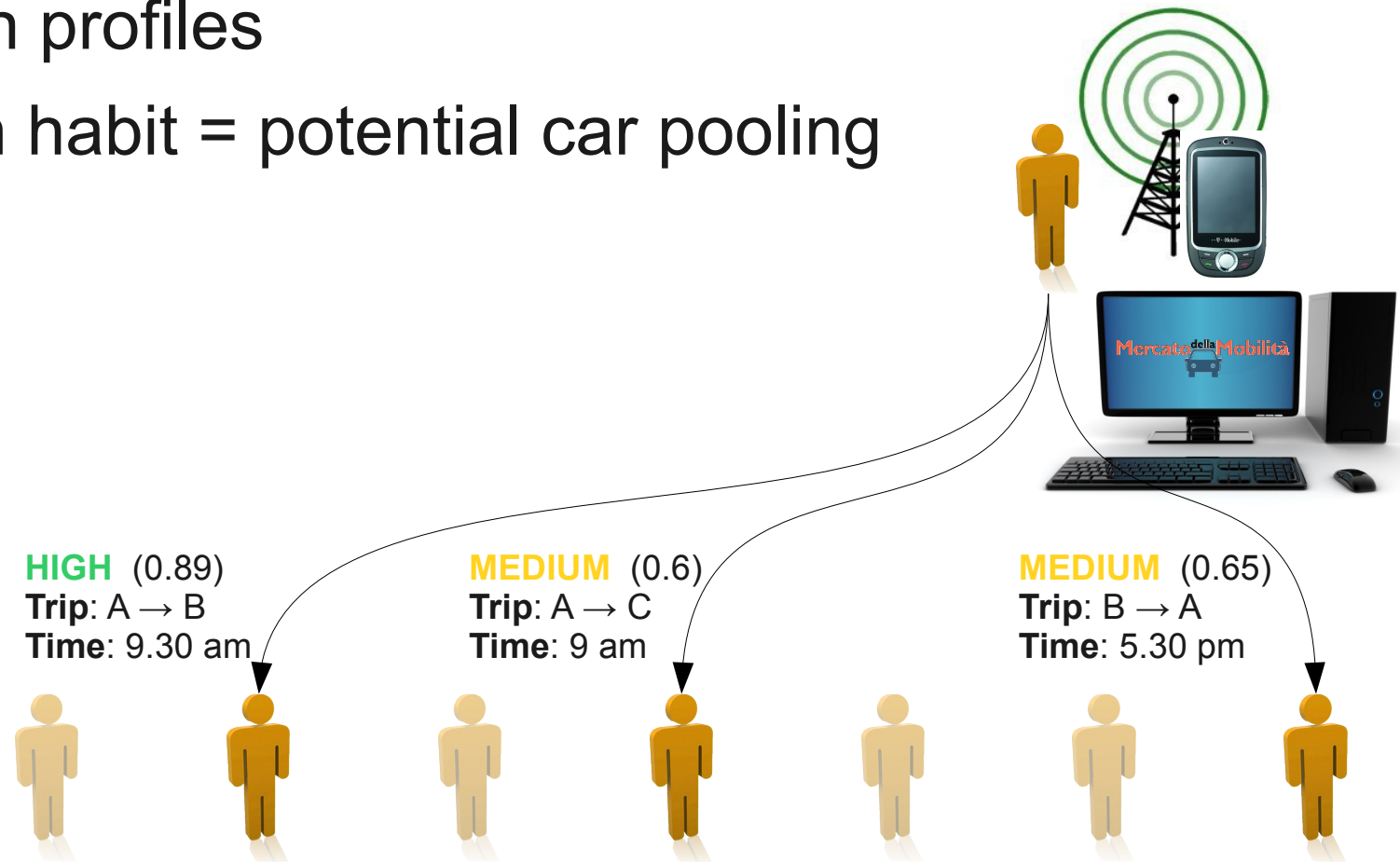


Applicazioni

- Il patrimonio di informazioni che i “big data” sulla mobilità ci forniscono possono fungere da base per applicazioni ad ampio raggio
- Qualche esempio:
 - Car pooling: una applicazione per il cittadino
 - Impatto di servizi sul territorio: una applicazione per le aziende

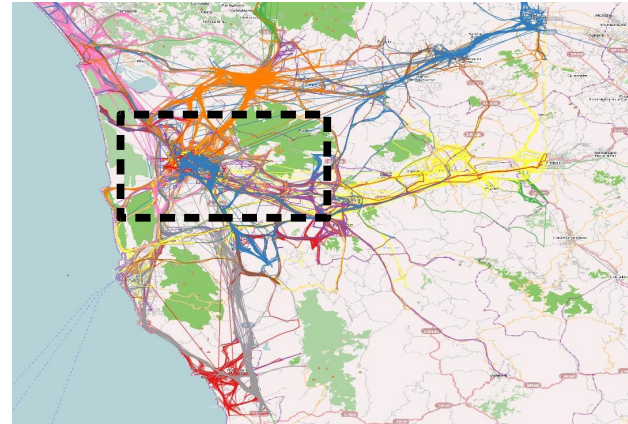
Car pooling

- Compare real trajectories to discover common habits
 - Based on profiles
 - Common habit = potential car pooling

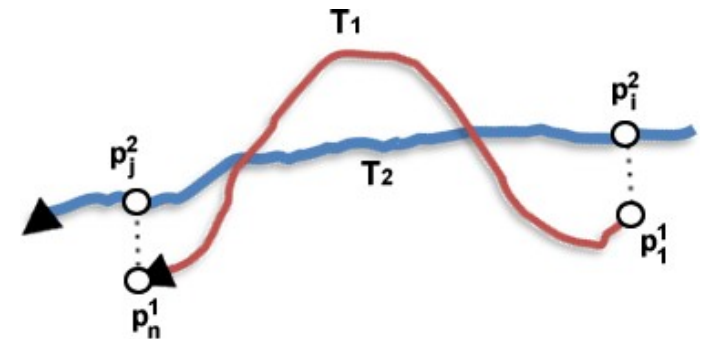


Car pooling

- Results on *domestic* traffic in Pisa:

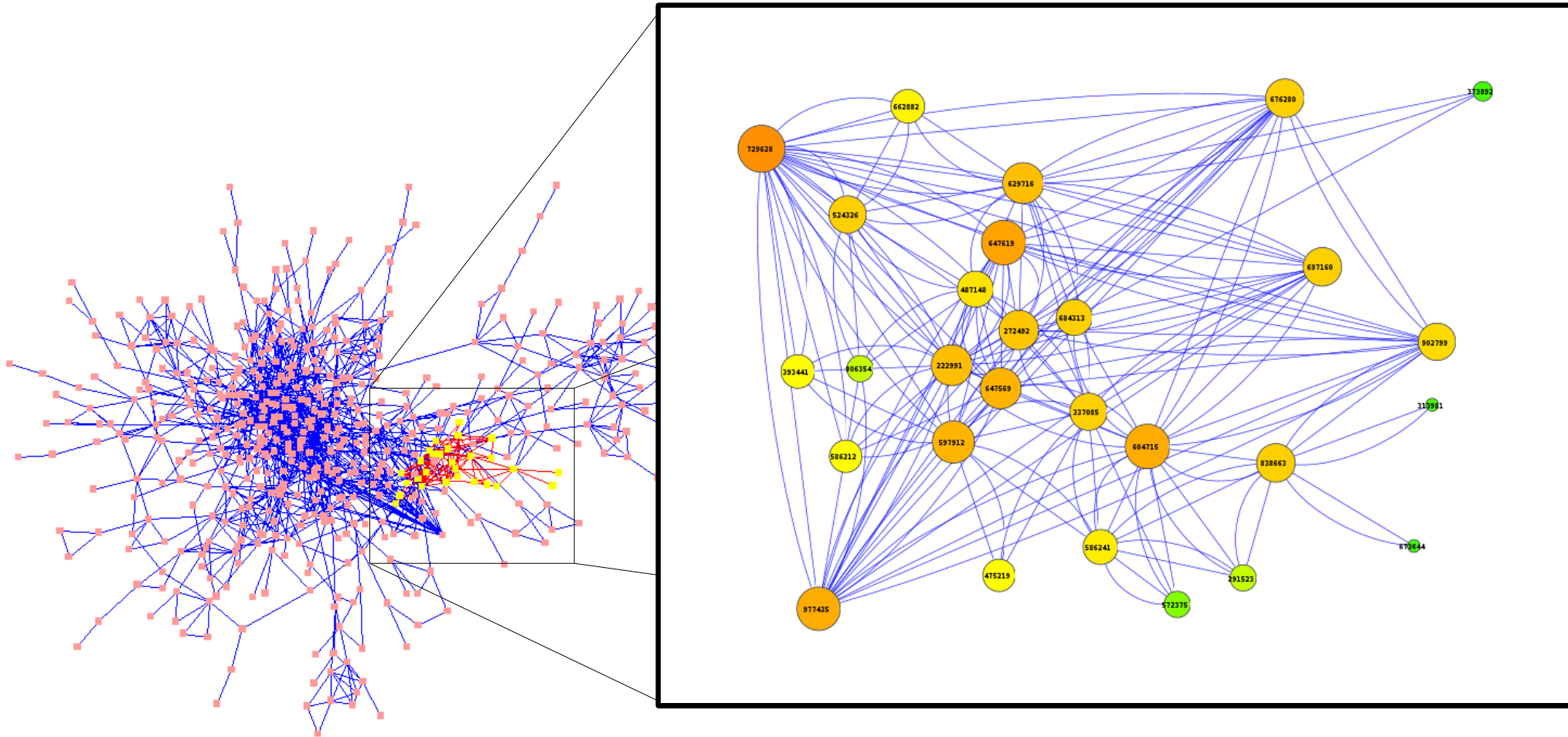


- 67.2 % of **profiles** match with one or more others
 - 32.5 % of **vehicles** share one or more profiles
- T_1 Matches T_2 if its start and end cross (or are within distance D from) T_2 with compatible timestamps

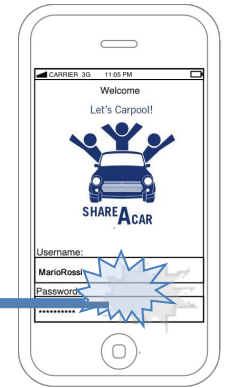
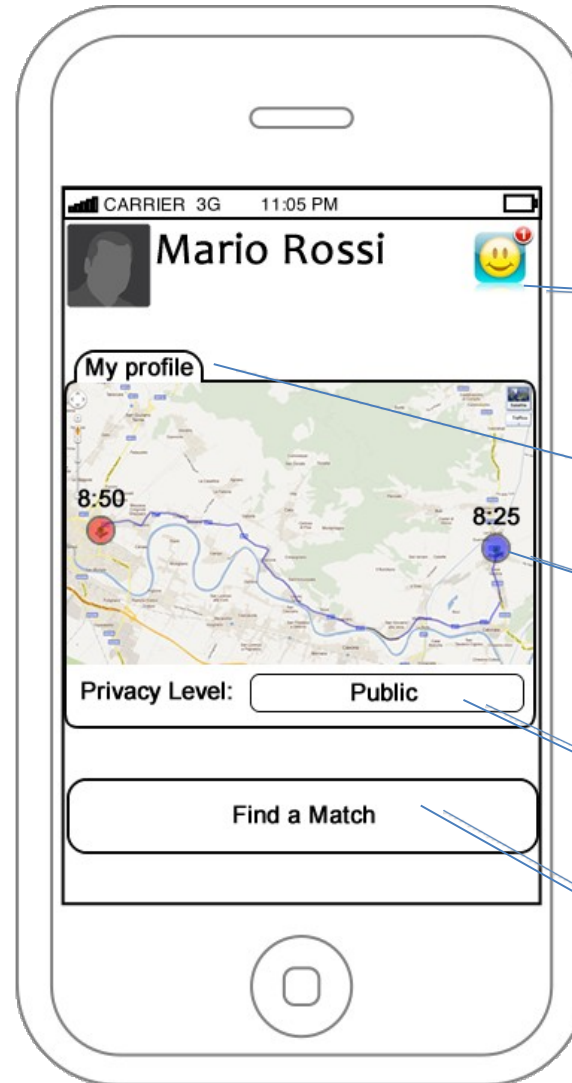
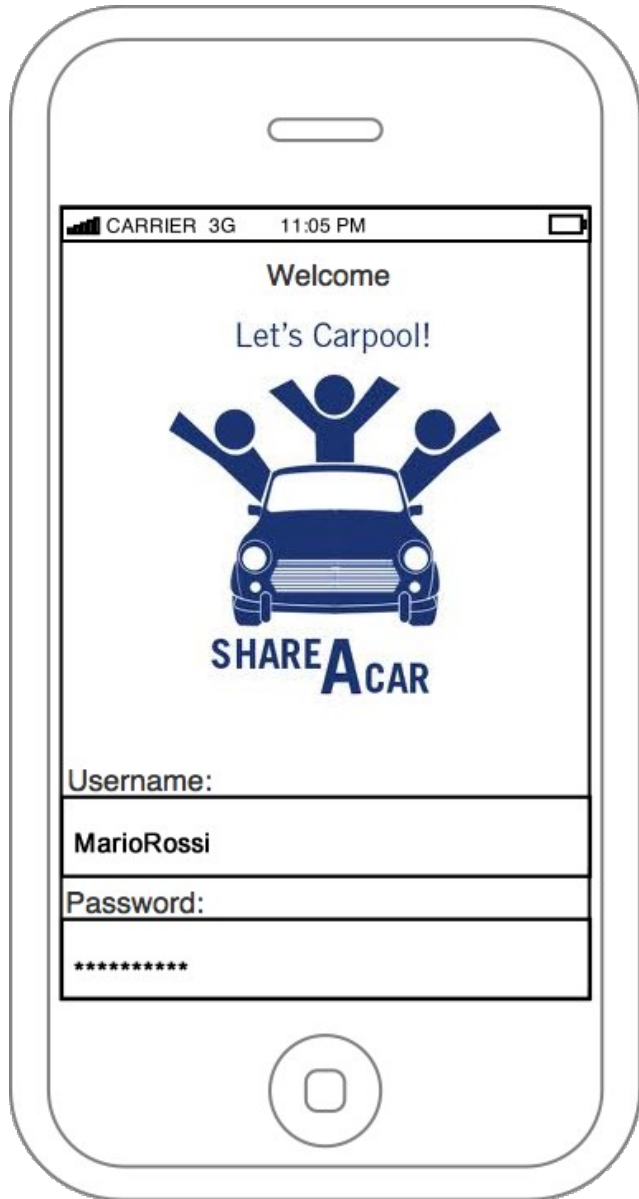


Car pooling

- There exist sub-communities with dense connections → group/public transportation?



Applicazione: car pooling



Pending requests

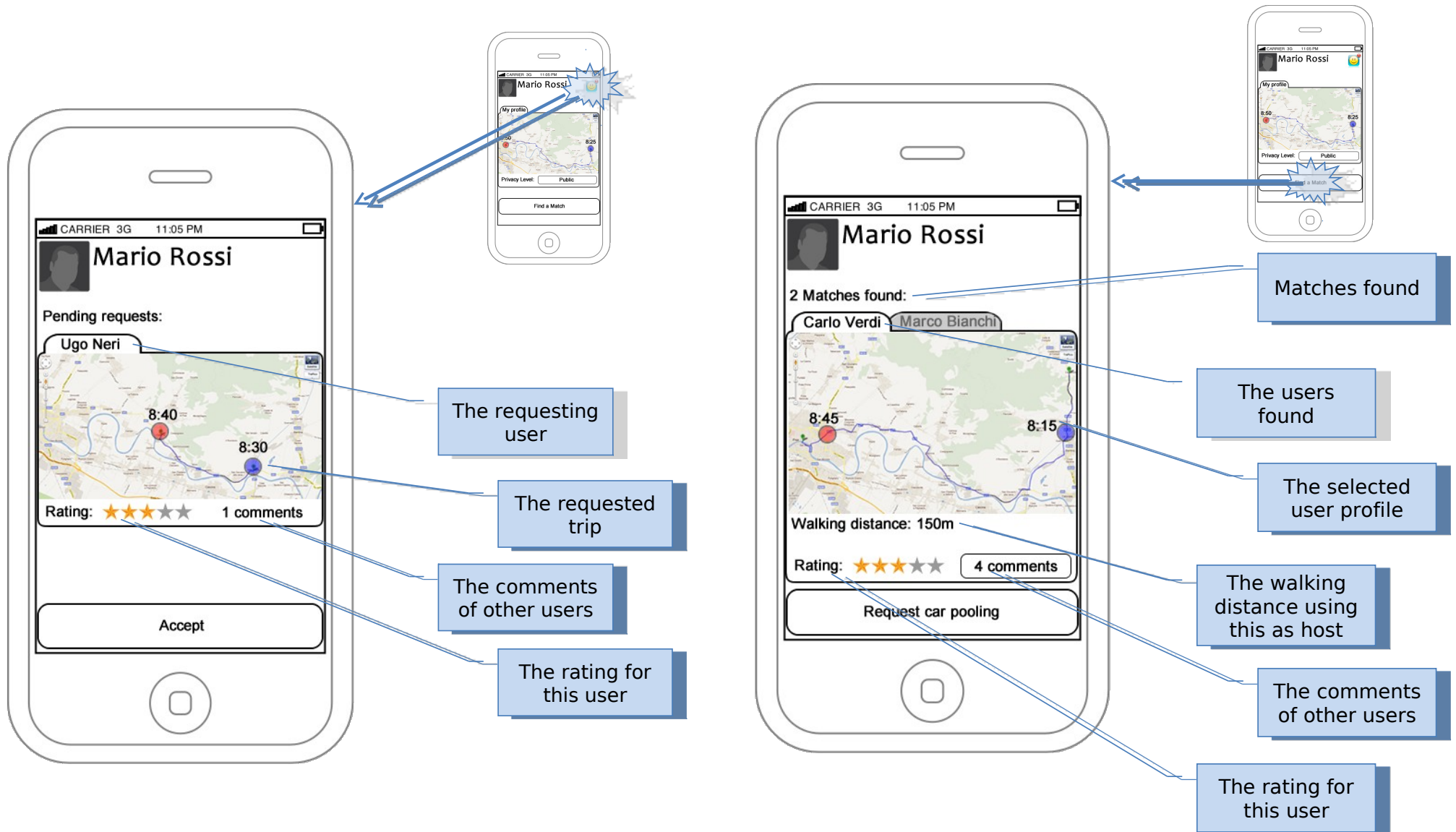
The user's profile

User's profile map

Privacy level of the profile

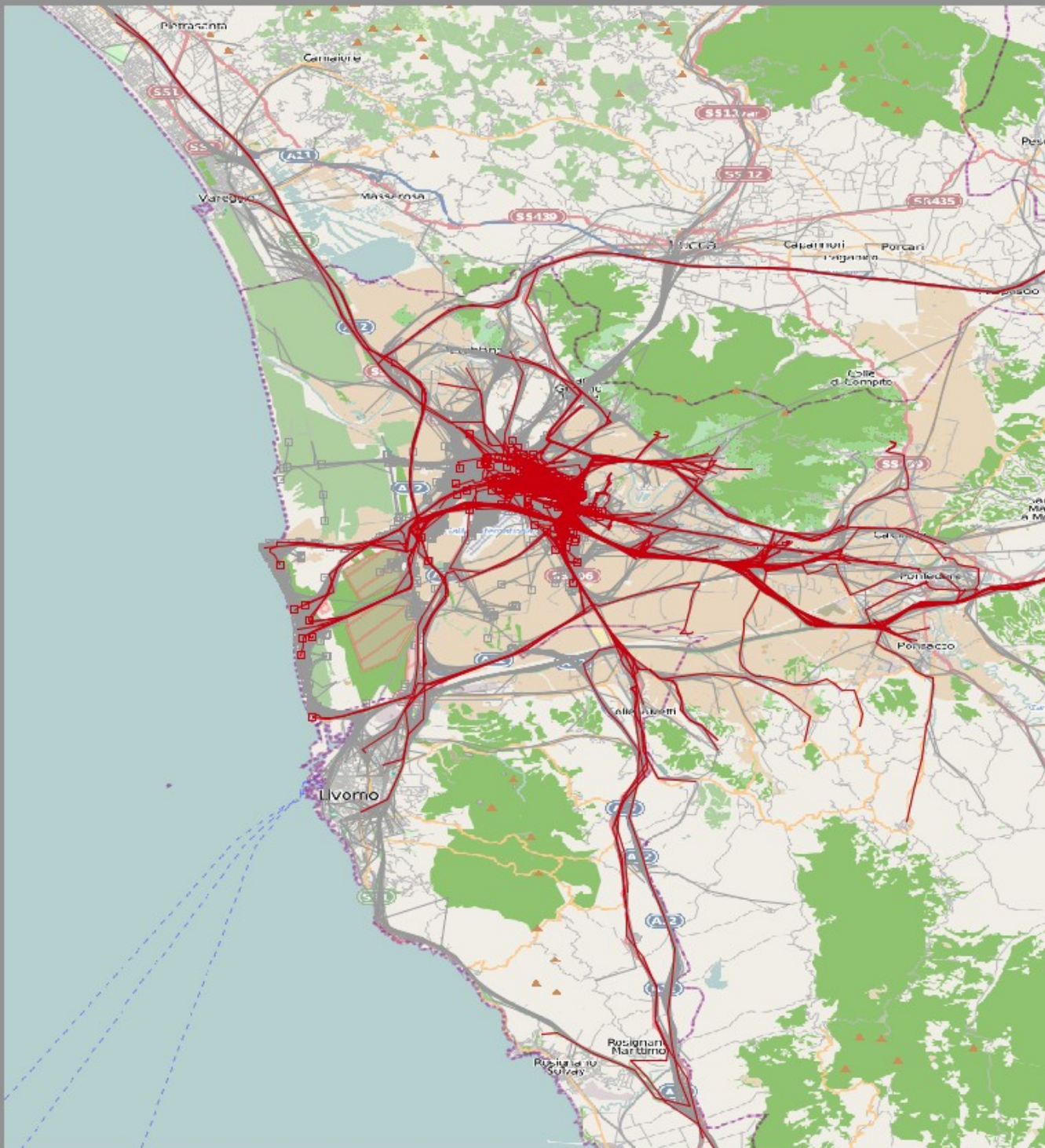
Perform a query to the central system in order to find the matches

Applicazione: car pooling/2



Applicazione: impatto servizi

- Stazioni di servizio, ecc.
 - Quale frazione di veicoli sosta in una determinata stazione di servizio?
 - Quali caratteristiche hanno tali utenti?
- Servizio di trasporto pubblico (TPL)
 - Quanta e quale parte della mobilità è coperta dal TPL, confrontato col trasporto privato?

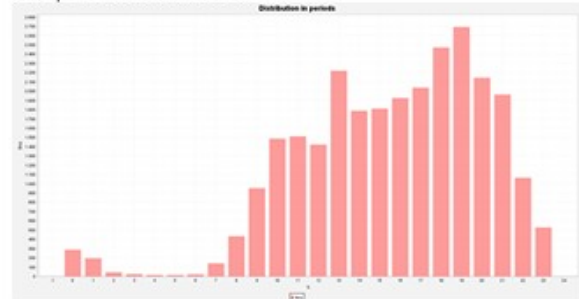


Q8 Via Cisanello

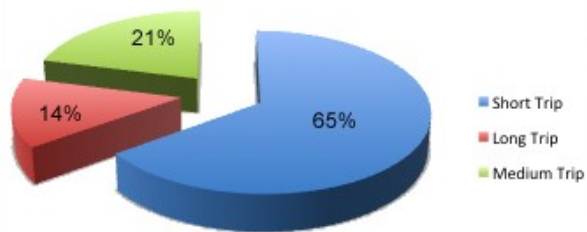
Via Cisanello, 156c
56124 Pisa



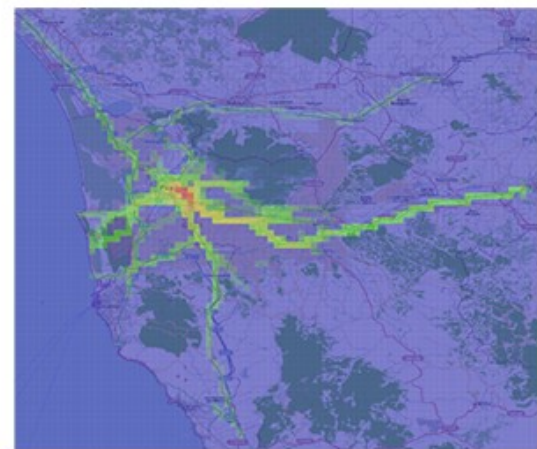
Temporal distribution of visits:



Distribution of visiting trip length:



Spatial distribution of visiting trips



Q8 Cisanello (Pisa)

Timeline:

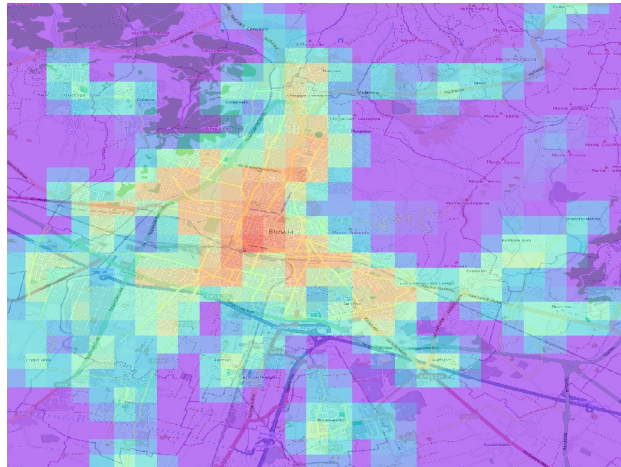
Jan 2009

Jan 2010

Jan 2011

Pub. Transp. Usage data

- Ticketing data describe where people get on
- Basic usage: distribution of origins

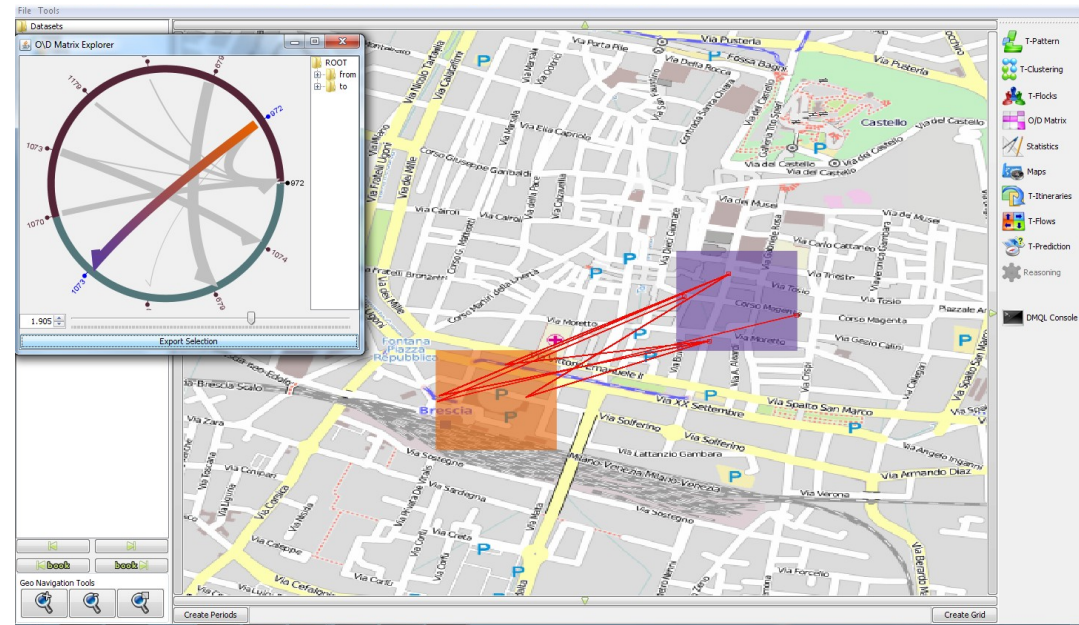


Spatial density of origins
Brescia, Italy, Jan 2009

- More advanced usage: estimate destinations
 - Possible with smart card and similar, joining information on consecutive rides
 - Multiple-buses trips can be reconstructed

Pub. Transp. Usage data

- Example: O/D matrix



- Example: build mobility profiles
 - As with GPS, yet based only on starts and ends
 - Label systematic vs. occasional mobility

GPS vs. Public transportation

- How much can actual public transport replace private transport?
- Experiments on Pisa
 - Compare GPS data with bus time tables
 - Trajectory fits \Leftrightarrow start and end points are within distance D from bus stops with suitable connections
- Results: distribution of matches

