BigData@Mobility
Understanding Human Mobility

Mirco Nanni
Knowledge Discovery and Data Mining Lab
(ISTI-CNR & Univ. Pisa)
kdd.isti.cnr.it
BigData @ Mobility: Objectives

• Infer human mobility from mobility proxies
  – GPS traces
  – Mobile phone records
  – Social media networks – Twitter, Flickr, …

• extract mobility models and patterns
  – Regularities – routines, traffic distribution, …
  – Predictive models

• apply them to specific services
  – Towards mobility planners / managers
  – Towards citizens, etc.
(A basic) Urban Mobility Atlas

- Vehicle mobility from GPS traces
Analysis Methods

- Main general line: infer semantics of places and movements from raw mobility data
- Methods vary with data sources
  - GPS trajectories
    - Mobility profiles
    - Individual Mobility Network
  - GSM call records
    - Routines discovery
    - Sociometer
- Others: e.g. PT ticket validation data
Methods.GPS.Mobility Profiles

- Describe an abstraction in space and time of the systematic movements of a user.
- The exceptional movements are completely ignored.

Routines = trips that most likely will take place also in the future
→ Applications in prediction and carpooling
Methods.GPS.IMN

*Individual Mobility Networks*

- Extract relevant locations and their trips
Methods.GPS.IMN

- Target: guess activities for locations and trips
- Method: learn a sequential classifier from labelled IMNs over network features
(A better) Urban Mobility Atlas

- Show collectively frequent paths
- Systematic vs. occasional mobility
Application 1
GPS.Proactive Carpooling

Routine containment: driver r2 can give a lift to driver r1

Carpooling Network: who might share a ride with whom

Carpooling Assignment: optimal matching to maximize impact
Application 2
GPS. Trajectory Prediction

- Combination of individual habits (profiles) and collective common behaviours
Methods.GSM.Routines

• A single trace of an individual can be poorly informative about his/her movements
Methods.GSM.Routines

- Yet, several daily traces of the same individual might allow to identify regular places and trips
Methods.GSM.Routines

- The whole individual mobility is then summarized by its systematic movements

- They will be used as typical daily schedule of the individual
For each user, extract significant L1 → L2

Aggregate (individual) systematic movements into (collective) systematic flows

Examples:

Outgoing traffic

Incoming traffic
Simulated traffic in Abidjan, Ivory Coast
Methods.GSM.Sociometer

- Derive presence distribution for each < user, municipality >

- $t1 = [00:00-08:00)$
- $t2 = [08:00-19:00)$
- $t3 = [19:00-24:00)$
Methods.GSM.Sociometer

- Recognize user's role through his “fingerprints”
  - Based on learning a “fingerprints” classifier
Methods.GSM.Sociometer

- Sample application: measuring exceptional events
Other methods
GSM.Correlation Patterns

• GSM data for event detection in space & time

• Extract correlated groups/sequences of events

{(Cell27, +35%)}
→ {(Cell7, +15%), (Cell5, +10%)}
→ {(Cell13, +5%)}
The future: multidimensional data

• Data Integration at the **individual level**
  – Integrate GPS, GSM, Flickr, Twitter
  – Personal Data Store view
    • Controlled by the user
    • Releases only essential information for services

• Data Integration at the **collective level**
  – Comparative analysis of phenomena from different data sources
    • E.g. event discovery & understanding w/ tweets & GSM
Individual Mobility Networks

- Individual Mobility Data Store
- Semantic Amplifier for Big Data
- Individual vs Collective modeling

Heterogeneity Model

\[ r_g = \sqrt{\frac{1}{N} \sum_{i \in L} n_i (\vec{r}_i - \vec{r}_{cm})^2}, \]

the characteristic distance traveled by individuals

\[ r_g^{(k)} = \sqrt{\frac{1}{N_k} \sum_{i=1}^{k} n_i (\vec{r}_i - \vec{r}_{cm}^{(k)})^2} \]

the radius computed on the k most visited locations

Well-being Indicators

Mobility Diversity

Social Diversity

![Graph showing Mobility Diversity](image1)

![Graph showing Social Diversity](image2)