MPRI Web Data Management Project

Aligning personal data : ccTV tracker using geolocalisation

Andreea Beica, Robin Champenois

February 24, 2014

1 Overview

The purpose of the project is to locate surveillance cameras on whose footage a person could appear during their daily trajectories, using geolocalisation information gathered by generally available API's, such as Google Maps Location, or Open Street Map.

2 Organisation

The project is organised as follows:

- geolocalisation information collection: we use the Google Location History feature of Google Maps (accessible via a Google account) in order to track our trajectory and then retrieve and store all the visited locations in a suitable format (i.e., latitute + longitude)
- surveillance camera database construction: using data collected via data.gouv.fr, we construct a database of all the positions of surveillance cameras; that data comes from the Open Street Maps API and from public information by the Police Department of Paris. We envisaged using a multiscale database, with several tables - one for every region we would decide upon-, in order anticipate large-scale storage or more efficient querying, but ultimately found that doing so did not improve things in a significant manner
- external information collection: we then proceed to locate other potential surveillance cameras that might have caught footage of us but aren't contained in the information provided by our sources; for this, we use the geolocalisation information gathered earlier on and the Google Places API to find places of interest we might have passed by; we then filter these places, keeping only the ones that are usually equipped with ccTV cameras (e.g., ATM's, banks, public institutions, public transport stations)

- locating potential footage: the final step is combining the personal data collected with the ccTV data in order to locate ccTV cameras that might have filmed us during our trajectories

3 Usage

Usage is :

```
python3 main.py fichier.kml
```

It outputs a report.html file containing the list of the cameras that may have recorded you, the time (with html metadata using schema.org conventions)

This program uses Google APIs. You may need to specify your own API key in the file commons.py.