Probable Path Inference for GPS Traces in Cities

Elena Agapie, Jason Ryder, Jeff Burke, Deborah Estrin

Urban Sensing, CENS - http://urban.cens.ucla.edu/

Introduction: Using GPS to generate a travel diary

Motivation – PEIR (Personal Environmental Impact Report)

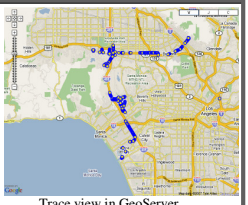
- Automatically generate travel diary from GPS measurements during daily activities to provide feedback and infer location statistics and information related to impact on and exposure to pollution.

Method

- Data is collected from a person using the phone and the GPS unit daily
- Data is processed using GIS tools to infer: **path** taken during GPS gaps, **traveled distances, average speed, percentage of time** spent in different zones of the city (residential, commercial, transportation, etc)

Feedback possibilities in the context of PEIR

- **Visualization** of recorded data
- **Statistical information** about travelling period - in what type of zones is the person spending time; time spent near polluted areas
- **Exposure** to different types of pollution (traffic exposure, ozone exposure)
- **Impact** of the user on the environment - distance and time travelled, means of transportation used, congestion impacts, local-air quality impacts, density of the activities in time and space, energy consumption



Trace view in GeoServer

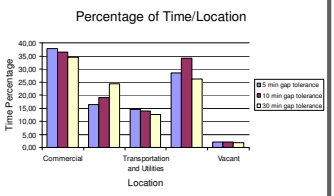
Problem Description: Inferring the path during GPS gaps

Dealing with gaps in the GPS data

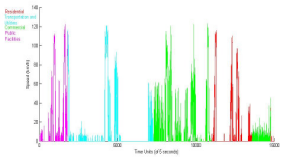
- Gaps in GPS data occur when the device looses reception or when it is turned off.
- Finding information about these periods is important in order to use the GPS data with more precision in subsequent analyses (building statistics on time spent in different zones or close to highways, etc)
- Speed, distance, time spent in different zones are analyzed for a better prediction of the path

Identifying Challenges

- Using information from GPS data and from a road map to infer a path
- Determining average speed
  - how to deal with recordings with speed zero
- Using information of speed and time spent in different zones
- Dealing with information near the boundary of the zones (in which zone is time spent)
- Dealing with large and small gaps (as distance or time)
- Percentage of time in different zones is computed based on how it takes gaps into account
  - A tolerance for gaps of 5, 10 and 30 minutes is set
  - Maximum speed is evaluated through different zones



Variation of Speed Through Zones



Proposed Solution: Finding the best path using speed constraints

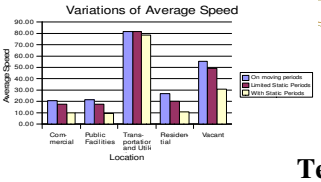
Outline

- Percentage of time spent in different zones of the city is analyzed
  - GPS data is passed through the processing pipeline
- A* algorithm is used for finding the probable path during GPS gaps
  - A heuristic is built
  - Speed information is extracted from GPS data

Processing pipeline

Speed analysis

- Speed computed under three conditions:
  - Periods of movement, only
  - With tolerance of 1-minute static periods
  - Considering all static periods
- Variation of speed values gives informati on the extent to which a person is static o moving in different zones



A* - best path detection algorithm

- Complete and optimal for a good heuristic
- Uses a cost function to find the best path

$f(x) = g(x) + h(x)$

f is the cost at position x, g is the cost of the path so far, h is the **heuristic** based on **speed**

Road map coding

http://www.esip.lsu.edu/projects/cms/images/

Heuristic

- **Extremities speed** – estimated at the extremities of the gap on two-minute segments, before and after
- **Straight-line speed** – estimated speed from straight line distance between extremities and time length of the gap
- **Zone speed** – average speed in zones
- **Road speed** – estimated speed on different types of roads small roads (25miles/h), large roads (45miles/h), highways (65miles/h)
- **Expected speed** – average speed of the extremities and of the zone speed

$$h(x) = |road\ speed - expected\ speed|$$

Estimated path in Roadwav

- **Challenges**
  - Computing expected speed using recordings that cross zones
  - Considering road limit speed
  - Computing the variation of user average speed to the limit speed of the road

Tests and Behavior

- Tested a small set of points on which zone speed and average speed have been varied
- Highways are avoided for small speed values and preferred with high values
- If extremity points have small speeds over large distances, highways are avoided, even if they were used in the original trace