Location-Awareness: Fall 2016
Exercise set 4

Due on 30th of November 2016 by 16:15

Abstract

Instructions: All course participants are requested to submit their exercise solutions electronically to course instructors Farbod Faghihi (faghihib at cs.helsinki.fi) and Ella Peltonen (ella.peltonen at cs.helsinki.fi), and to the course lecturer (petteri.nurmi at cs.helsinki) by the due date (latest before the exercise session). In all the exercises, do not just give the answer, but also the derivation how you obtained it. Participants are encouraged to write computer programs to derive solutions to some of the given problems.

1 Concepts

a) Explain what velocity pruning is and how it can help detecting places. Download the GPS trajectory of a user from the URL[1]. The columns correspond to time in seconds, longitude, and latitude respectively. Using Velocity Pruning, how many places seem to be in the data? What velocity threshold value did you use for pruning the coordinates and why?

b) Explain the difference between a Markov chain of order one and higher orders.

c) Consider the following place visitation sequence: HWSWSHWHSHSWHWHWSWSHSWHSWSSHWHWHWS where 'H', 'W', and 'S' correspond to visits to Home, Work, and Shop respectively. Using this data create the probability transition matrices of Markov Models of order one and order two. Assuming that the user is currently located at the Shop what is the next most probable place based on each of these models? What if the user has visited Work before visiting the Shop?

2 Fingerprint-Based Place Detection

Download the file from the URL[2] containing WiFi access points. The columns of the file correspond to different access points and the rows correspond to measurements. Each entry in the matrix corresponds to the observed RSS value (or 0 if the access point was not observed).

Note: you can assume that the access points have global indexing, i.e., the i^{th} column always refers to the same access point.

a) Use the Extended-Tanimoto formula (SensLoc) given in the Lecture slides to identify places from the measurements. How many distinct places there are in the data?

b) The implementation of SensLoc (see [KKES10]) operates on **response rate** vectors instead of RSS vectors. Response rate vectors are defined as $\mathbf{r} = r_i$ where $r_i = n_i/n$. Here n_i denotes the number of scans where access point i is observed and n is the total number of scans made in a location. Construct response rates the measurements considering successive pairs of measurements (i.e., $n = 2$ with one response rate vector from measurements 1 and 2, another from measurements 3 and 4, and so forth).

[1] http://universe.hiit.fi/teaching/location-awareness/2016/velocity_pruning.csv
[2] http://universe.hiit.fi/teaching/location-awareness/data/rssVals.csv

c) Perform place detection using the response rate vectors derived in the previous subtask. How many places you can identify now (and why)?

3 Lempel-Ziv

a) Given the sequence of GSM cell values in Table 1, construct a LZ-tree from the values.

b) Determine the probabilities of the different nodes in the LZ tree. What is the likelihood of the user currently being in cell D given measurements A and N?

B	A	N	A	N	A	B	A	N	D	A	G	E	B	A	N	A	N	A

Table 1: Caption for the table.

4 Spatial Analysis: Preprocessing and Place Detection

Download GPS trajectory of a user from the URL[3]. The columns of the file represent longitude, latitude, timestamp, satellites and HDOP.

a) Perform pre-processing techniques taught in the course to remove the noise from the data (Based on the number of satellites, HDOP value, and removing outliers). Create visualizations of the processed data in KML format after each processing step. (Hint: you can perform your analysis in the Euclidean space. To get a distance measure in meters you can use 0.005 unit distance in Euclidian space = 500 meters).

b) Perform place identification on the pre-processed data and plot the cluster centers (in KML). How many places seem meaningful and why?

References

[KKES10] Donnie H. Kim, Younghun Kim, Deborah Estrin, and Mani B. Srivastava. Sensloc: Sensing everyday places and paths using less energy. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*, SenSys '10, pages 43–56, New York, NY, USA, 2010. ACM.

[3] http://universe.hiit.fi/teaching/location-awareness/ex-3/tokyo.csv