

Mobile Sensing: Spring 2015

Exercise: 3

Due on 26th March 2015 by 17:45 PM.

Instructions: All course participants are requested to submit their exercise solutions electronically to the instructors (samuli.hemminki at cs.helsinki.fi and teemu.pulkkinen at cs.helsinki.fi), as well as to the course lecturer (petteri.nurmi at cs.helsinki.fi) 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.

Ex 1. Load SensorData1 ¹. containing accelerometer measurements.

- a) Divide the data into frames of 50 samples each.
- b) Divide the data into frames of 1 second duration each.
- c) For both cases a) and b), calculate and compare the mean, median, standard deviation and interquartile range of each frame.

Ex 2. Load SensorData2 ² containing gyroscope measurements which have been collected simultaneously with the data contained in SensorData1.

- a) Perform sample synchronization for the two sensors: SensorData1 and SensorData2.
- b) Interpolate sensor values (from both sensors) with uniform 20ms intervals, starting from the timestamp given by the first accelerometer measurement.

Ex 3. Consider a scenario where 4 devices, D_1, D_2, D_3 and D_4 , sample their microphones. Table 1 summarizes similarity ρ between the audio signals, Table 2 summarized the information value α of each device and Table 3 summarizes energy costs associated with sensor sampling for each device. Additionally, we define the energy due to data transmission and processing overhead $33mW$ per device. Information gain G of adding device D_i to a set of devices $D_{j...n}$ is calculated using formula: $G(D_i, D_{j...n}) = (1 - \max(\rho(D_i, D_{j...n}))) * \alpha(D_i)$

¹In the data, columns 1 to 4 corresponds to: [timestamp, x, y, z]

²In the data, columns 1 to 4 corresponds to: [timestamp, x, y, z]

- a) Assuming an energy budget of 300mW, write a function which finds the set of devices $D_{i..j}$ yielding highest information. What is the achieved information?
- b) Assuming information requirement of 7 units, find the most energy-efficient device combination.

ρ	D_1	D_2	D_3	D_4
D_1	1	0.5	0.7	0.2
D_2	0.5	1	0.8	0.1
D_3	0.7	0.8	1	0.1
D_4	0.2	0.1	0.1	1

Table 1: Device similarity matrix

α	Information
D_1	2.8
D_2	5.5
D_3	6.2
D_4	2.4

Table 2: Information value

mW	Energy Costs
D_1	94mW
D_2	87mW
D_3	112mW
D_4	61mW

Table 3: Device energy cost

Ex 4. DTW / Distance

Let's do a little detective work!

In a user study on the movement patterns of customers within a supermarket environment, three location traces were recorded from three different users. In addition, one of these three users carried a cheaper and less accurate device (more noise, less samples). Due to a mixup in the study setup, however, the researchers forgot which user was carrying the cheaper device.

Your task is to figure out which user carried the faulty device. Specifically, implement the Dynamic Time Warping algorithm and compare the different user traces to the mystery trace. Which user would you suggest carried the device?

Traces are in CSV-format, with columns representing $[X,Y]$ -coordinates. Use the Euclidean distance as the cost function.

User 1

User 2

User 3

User Mystery