

NYIT
Computer Science

**CSCI 860: Biometrics and its
Applications (Spring 2018)**

Programming Assignment 1

**Submitted To
Professor Kiran Balagani**

**Prepared By
Rohanraje Sawant
Student id:1113947**

Problem Statement:-

You are required to implement Manhattan verifier and report false accept (impostor pass) and false reject rates on a publicly available keystroke biometric dataset. You may use any programming language, as long as it can be compiled.

Dataset:-

The data consist of keystroke-timing information from 51 subjects (typists), each typing a password (.tie5Roanl) 400 times.

Programming Approach:-

I have written the program in python 2.7 and compiled in a Jupyter environment for the above problem statement. Used the scientific library Scipy and Pandas package to compute the statistical operation. Steps performed for the analysis of data:

1. Imported the csv file of the keystroke data as kdata.csv in the python environment .
2. kdata.csv is read in a data frame which is loaded with the help of Panda and contains the entire contents of the kdata.csv file where subject column is declared as unique which is the labels of all the 51 subjects.
3. Depending on the value of $N(100,200,300)$ the training(model) sample set of data is created from the 400 entries of the each user . And, the rest is considered as the testing sample for the genuine user data for particular user.
4. Extracted training sample set of data is considered for the mean vector template for each user which will be parameter for generating genuine and impostor score.
5. The mean vector template and genuine testing sample is computed in manhattan distance formula for each user which generates the genuine scores as the further parameter for the False Reject Rate.
6. Then impostor score is calculated in Manhattan distance which has the parameter mean vector template of particular user and impostor testing sample which is 400-N from the all other subjects or user.

7. The threshold value is set for the verification. Then program scans through all the genuine score to generate the False reject Rate and also scans through impostor score to generate the False Accept rate.

Outputs:-

1. Well documented, compilable software codes and executables performing template calculation; genuine and impostor score computation with Manhattan distance; and calculation of false accept and false reject rates at a given threshold T.
=> Threshold set as thr=15 and N=200. Output in output folder by file name FARFRR15N200.pdf.
2. A well-written report containing false accept and false reject rates for N = 200 and various threshold values (choose five threshold values that give you the best tradeoff between the false accept and false reject rates).
=> Five threshold values are:
 - a. Threshold set as thr=0.05 and N=200. Output in output folder by file name FARFRR0.05N200.pdf.
 - b. Threshold set as thr=0.5 and N=200. Output in output folder by file name FARFRR0.5N200.pdf.
 - c. Threshold set as thr=1 and N=200. Output in output folder by file name FARFRR1N200.pdf.
 - d. Threshold set as thr=5 and N=200. Output in output folder by file name FARFRR5N200.pdf.
 - e. Threshold set as thr=10 and N=200. Output in output folder by file name FARFRR10N200.pdf.
3. Report the false accept rate at 0 false reject rate, when N = 100, 200, and 300.
=> 0 False Reject Rate at:
 - a. Threshold set as thr=28 and N=100. Output in output folder by file name FAR0FRR100N.pdf.
 - b. Threshold set as thr=15.5 and N=200. Output in output folder by file name FAR0FRR200N.pdf.
 - c. Threshold set as thr=15 and N=300. Output in output folder by file name FAR0FRR300N..pdf.