Programming Assignment 1  
CSCI-860: Biometrics in a Networked Society  
  
Presented To  
Kiran Balagani  
Associate Professor of Computer Science  
New York Institute of Technology  
  
Prepared By:  
Pranav S. Krishnamurthy  
Major – Computer Science  
Student ID: 0704111  
  
Due Date – 2nd December 2014

# Problem Statement

This assignment requires the implementation of the Manhattan verifier to report *false accept (impostor pass)* and *false reject rates* on a publicly available keystroke biometric dataset. Any programming language may be used, as long as the program can be compiled on computers in Harry Schure Hall Room 212. In addition, a demonstration and explanation of our written code will be required.

# My Approach

For this programming assignment, I have used the C#.net programming language, the Visual Studio 2013 Ultimate IDE for Windows 8.1 along with the PasswordData.csv file which I manipulated into 51 separate csv files, and when running the program, I have the code written such that depending upon the user number and the value of N, which represents the sample size: the training and testing vectors are extracted from the original data, the template vectors, genuine scores, impostor scores, and the various rates are calculated appropriately. The approach that I have taken for this assignment is detailed below:

1. Taking the CSV file from the link provided in the PDF file given, I took that file and broke that into 51 separate CSV files which represents each user.
2. Depending upon the user number that the end user enters (the user number will act as the “tag” for the subject ID), the program will then find the file in the respective directory, and convert the entire CSV file into a 2D double array.
3. When the end user enters a value for N, which can have the value of either 100, 200, or 300; the program will then parse through using the 2D double array, and the value of N and extract the testing samples which will be used to generate the mean vector.
4. The mean vector, the value of N, and the remaining samples are then parameters for generating the genuine scores. Here, I have assumed that the value of n in the formula for the Manhattan Distance is equal to the number of samples.
5. Then to generate the impostor scores, I will then be taking all the other user data which contain 400-N samples, and the mean vector of the user as parameters.
6. *Finally, will be taking a look at the various genuine scores and using a threshold T and using the same threshold T, and looking at the various impostor scores, will calculate the Impostor Pass Rate and False Reject Rates.*  – This step I have tried, but due to time constraints I was not able to accomplish this step.