Project Assignment

Comparing anomaly detection using Manhattan Detector for Keystroke Dynamics

Submitted to,

Professor Kiran Balagani

New York Institute of Technology

By,

Salman Siddiqui

Student ID: 1204013

CSCI-860, Spring 2019

New York Institute of Technology

Contents

Problem Statement

The scope of the project is to analyse and implement a Manhattan verifier on a publicly available keystroke biometric dataset. To verify the accuracy of the verifier, we compute the false reject rate (false positive rate) and the impostor pass rate (false accept rate) on the population. Furthermore, an optional step was taken to calculate the equal error rate. For a good verifier, it is always good to have a low equal error rate. In addition, a ROC curve was generated and the area under the curve was calculated to understand the accuracy of the verifier.

Dataset

The related data was collected at Carnegie Melon University. The number of participants was 51. Each subject attended 8 sessions and typed a password 50 times per session. The password was ".tie5Roanl", which is of length 10. The features of the password are the key hold times, key interval and trigraph. We can calculate the trigraph knowing the hold times and interval. However, all three features were used for the computation. All in all, there were 31 unique features of 400 repetitions per subject.

Approach

Python programming language was used due to its versatility when handling a large dataset. Furthermore, as stated in the verification task, we take in two inputs from the user, i.e. the sample size (N) and the threshold (t). First the dataset is loaded in data frames in python. We split the data according to the user entered sample size, N. In other words, N is also the number of repetitions we take per user. For example, if N is 200, we take the first 200 repetitions of each user the training set and the remaining 200 as the testing set. Using the training set, we calculate the template of each user across the features.

We treat the dataset as population and not as sample. Which implies that using the testing sample, we calculate the genuine score of every user. For N of 200, this creates a dataset of 51x200 genuine scores. Next, we calculate the impostor score for each user on the testing sample. For N of 200, we have the 50x200 impostor score for each user. For a user group of 51 subjects, we have 51x50x200 impostor scores. After splitting and calculating the scores, we should have 10,200 genuine scores and 510,000 impostor scores.

Results

Using the threshold input from the user, the false positive rate (false reject rate) and impostor pass rate (false accept rate) is calculated. For example, for a sample size of 200 and threshold of 2, the false positive rate was 0.269 and the impostor pass rate was 0.349.

```
Enter the size of the sample: 200
Enter a threshold value: 2

--------------------------------------------------------
False positive rate (false reject rate) for threshold of 2.00  (N=200) is 0.269
Impostor pass rate (false accept rate) for threshold of 2.00  (N=200) is 0.347

-------------------------------------------------------
```

Figure 1: A snippet of the output on the IDE

For this test run, there is a reason why we selected 2 as our threshold which we will explore later in the document. Another requirement of the project was select 5 threshold values that give the best trade-off between the false accept and reject rates. To estimate the optimum threshold values that gives the best trade-off, we use the roc_curve method from sklearn library. Using this method we plot a graph of detection error trade off for a sample size of 200.

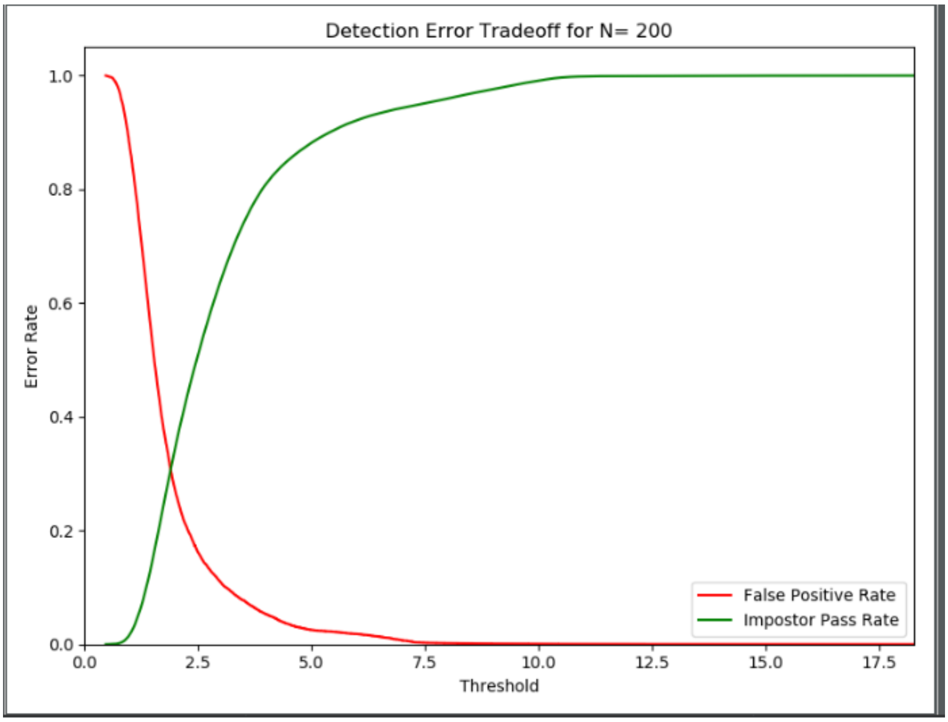

Figure 2: Detection error trade off for N=200

Using this graph, we can estimate 5 optimal thresholds. We select 0, 1, 2, 3, 8 as our threshold. It was important to include 1 and 2 as we can observe in figure 2 that our equal error rate occurs at a threshold between 1.5 and 2. The false positive rate (false reject rate) and impostor pass rate (false accept rate) for the above thresholds are printed below.

```
At selected thresholds [0, 1, 2, 3, 8]
False positive rate (false reject rate) for threshold of 0.00 (N=200) is 1.0
Impostor pass rate (false accept rate) for threshold of 0.00 (N=200) is 0.0

False positive rate (false reject rate) for threshold of 1.00 (N=200) is 0.872
Impostor pass rate (false accept rate) for threshold of 1.00 (N=200) is 0.0177

False positive rate (false reject rate) for threshold of 2.00 (N=200) is 0.269
Impostor pass rate (false accept rate) for threshold of 2.00 (N=200) is 0.347

False positive rate (false reject rate) for threshold of 3.00 (N=200) is 0.109
Impostor pass rate (false accept rate) for threshold of 3.00 (N=200) is 0.64

False positive rate (false reject rate) for threshold of 8.00 (N=200) is 0.00196
Impostor pass rate (false accept rate) for threshold of 8.00 (N=200) is 0.96
```

Figure 3: A snippet of the result of selected threshold values

We can also verify our selection of the thresholds by plotting a graph which has the impostor pass rate and false positive rate at the selected thresholds. In figure 4, the curve looks similar to the curve in figure 3.

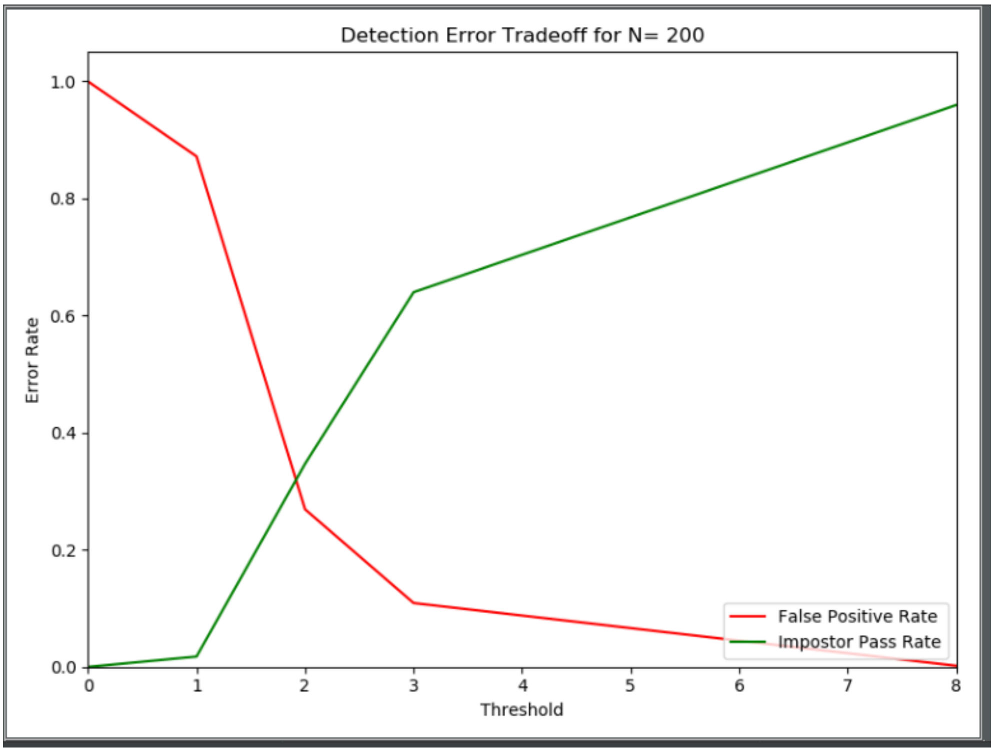

Figure 4: DET curve for selected threshold

Another requirement was to display the impostor pass rate (false accept rate) at 0 false positive rate (false reject rate) and the false positive rate (false reject rate) at 0 impostor pass rate (false accept rate) at 3 different sample sizes. To obtain these values, by intuition we know that the false positive rate (false reject rate) will be 0 when all genuine scores are classified as genuine. To find the threshold at which all genuine users are classified as genuine, we find the maximum score in the genuine scores. Using the maximum score in the genuine score, we calculate the impostor pass rate (false accept rate). To compute the value for false positive rate (false reject rate) at 0 impostor pass rate, we know that in order to obtain a 0-impostor pass rate, the threshold must be less than the minimum impostor score in the dataset. We find the minimum score in the impostor data set and use a value less than that since we say that anything less than or equal to threshold is genuine and anything else is positive. Using the less than minimum value in the impostor score as our threshold we calculate the false reject rate. We repeat the same process for the different sample size

```
-------------------------------------------------------
Impostor pass rate (false accept rate) at zero false reject rate is (N=100) 0.99994 at threshold (max score of genuine) 27.8
At the same threshold, false reject rate is (verification): 0.0

False positive rate (false reject rate) at zero impostor pass rate is (N=100)1.00000 at threshold (minimum score impostor) 0.430
At the same threshold, impostor pass rate is (verification): 0.0

-------------------------------------------------------
```

Figure 5: A snippet from the output in the IDE for N = 100

```
------------------------------------------------------
Impostor pass rate (false accept rate) at zero false reject rate is (N=200) 0.99986 at threshold (max score of genuine) 15.1
At the same threshold, false reject rate is (verification): 0.0

False positive rate (false reject rate) at zero impostor pass rate is (N=200)1.00000 at threshold (minimum score impostor) 0.463
At the same threshold, impostor pass rate is (verification): 0.0

------------------------------------------------------
```

Figure 6: A snippet from the output for N=200

```
------------------------------------------------------
Impostor pass rate (false accept rate) at zero false reject rate is (N=300) 0.99964 at threshold (max score of genuine) 14.5
At the same threshold, false reject rate is (verification): 0.0

False positive rate (false reject rate) at zero impostor pass rate is (N=300)1.00000 at threshold (minimum score impostor) 0.438
At the same threshold, impostor pass rate is (verification): 0.0

------------------------------------------------------
```

Figure 7: A snippet from the output for N=300

We also calculate the false reject rate at the maximum score of genuine to verify our claim. Similarly, we calculate the impostor pass rate at minimum score of impostors to verify the claim. We notice that the false reject rate at all the 3 different sample size is a little less than 1. We can deduce that there are certain impostor scores that are higher than the maximum genuine scores.

Further analysis which were beyond the scope of the assignment was also done to verify and to facilitate the understanding of the concept and the idea of keystroke dynamic biometric systems. We calculate the equal error rate for 3 different sample sizes.

```
------------------------------------------------------
Equal error rate for selected threshold (N=100) is 0.347
Threshold at which equal error rate occurs (N= 100) is 2.142
```

Figure 8: Equal error for N=100

```
------------------------------------------------------
Equal error rate for selected threshold (N=200) is 0.307
Threshold at which equal error rate occurs (N= 200) is 1.897
```

Figure 9: Equal error rate for N=200

```
------------------------------------------------------
Equal error rate for selected threshold (N=300) is 0.283
Threshold at which equal error rate occurs (N= 300) is 1.761
```

Figure 10: Equal error rate for N=300

We can observe that as we increase the size of our training data, the equal error rate improves. Which is true since we know if a verifier is trained with more data, the better score it can produce.

Furthermore, we calculate the area under the curve and the relative operator characteristics for different sample size. We observe that as we increase the sample size, the area under the curve increases justifying our claim that the model can perform better if we have more training data.

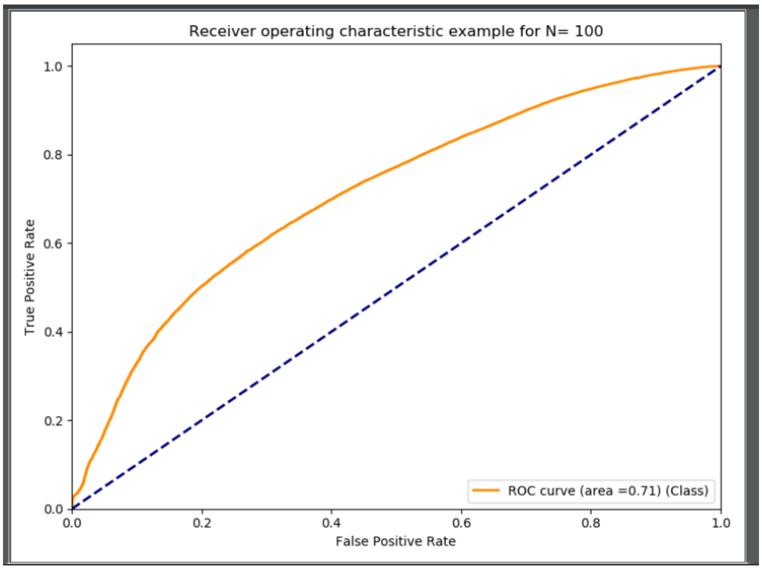

Figure 11: ROC for N=100

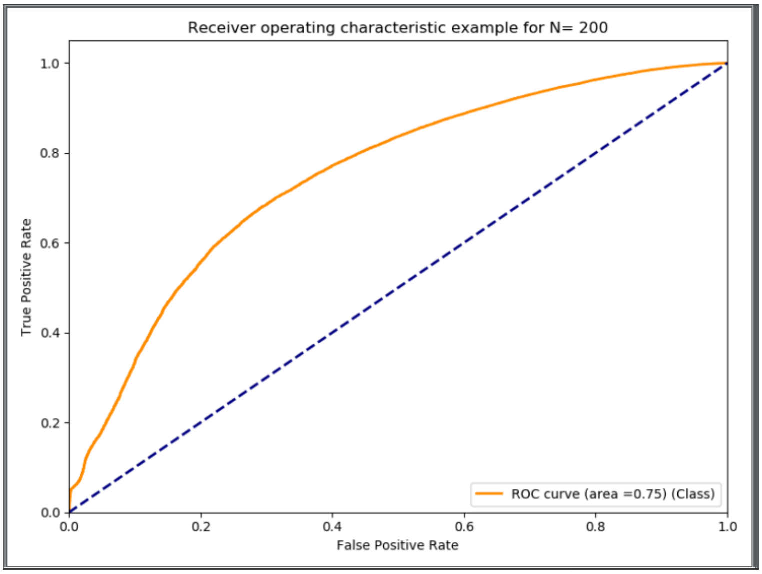

Figure 12: ROC for N=200

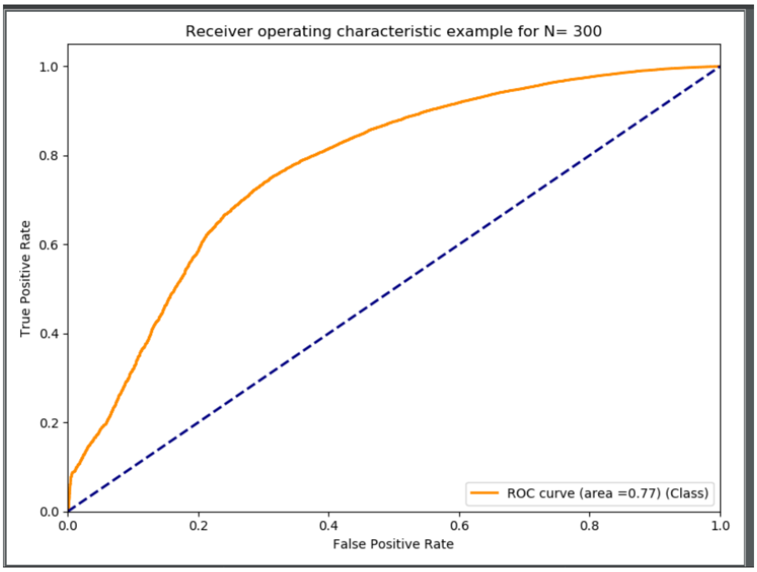

Figure 13: ROC for N=300

Conclusion

We can observe that our verifier leaves room to be improved. In the future, we can have a larger dataset and can change the password length. We can also change the feature size using only the key intervals and key hold times. Furthermore, we can use different verifiers to improve our equal error rate and the area under the curve. To recap, completing the scope or objective of this project, we have done the following

1. Calculate the false accept rate and false reject rates at a given threshold
2. Present the false accept rate and false rejects for a sample size of 200 at give different threshold values
3. Display the false accept rate at 0 false reject rate and false reject rate at 0 false accept rate at sample sizes of 100,200 and 300
4. All of the above was calculated using user inputs of sample size and threshold

Appendix

```
"C:\Program Files\Python37\python.exe" "C:/Users/Salman Siddiqui/PycharmProjects/Biometrics/Driver.py"
Enter the size of the sample: 300
Enter a threshold value: 8

---------------------------------------------------------
False positive rate (false reject rate) for threshold of 8.00  (N=300) is 0.00196
Impostor pass rate (false accept rate) for threshold of 8.00  (N=300) is 0.976

--------------------------------------------------------


At selected thresholds [0, 1, 2, 3, 8]
False positive rate (false reject rate) for threshold of 0.00 (N=300) is 1.0
Impostor pass rate (false accept rate) for threshold of 0.00 (N=300) is 0.0

False positive rate (false reject rate) for threshold of 1.00 (N=300) is 0.802
Impostor pass rate (false accept rate) for threshold of 1.00 (N=300) is 0.0235

False positive rate (false reject rate) for threshold of 2.00 (N=300) is 0.215
Impostor pass rate (false accept rate) for threshold of 2.00 (N=300) is 0.377

False positive rate (false reject rate) for threshold of 3.00 (N=300) is 0.102
Impostor pass rate (false accept rate) for threshold of 3.00 (N=300) is 0.677

False positive rate (false reject rate) for threshold of 8.00 (N=300) is 0.00196
Impostor pass rate (false accept rate) for threshold of 8.00 (N=300) is 0.976



---------------------------------------------------------
Impostor pass rate (false accept rate) at zero false reject rate is (N=300) 0.99964 at threshold (max score of genuine) 14.5
At the same threshold, false reject rate is (verification): 0.0

False positive rate (false reject rate) at zero impostor pass rate is (N=300)1.00000 at threshold (minimum score impostor) 0.438
At the same threshold, impostor pass rate is (verification): 0.0

---------------------------------------------------------
Equal error rate for selected threshold (N=300) is 0.283
Threshold at which equal error rate occurs (N= 300) is 1.761
```

Figure 14: Output on the idea for N=300