UNOBTRUSIVE AUTHENTICATION OF SMARTPHONE USERS BY CONTEXT-AWARE BEHAVIOURAL BIOMETRICS

by

SIVAA SHAANTH SUMAN R S 2015103048
AISHWARYA V 2015103582
SHOPHINE S 2015103609

A project report submitted to the

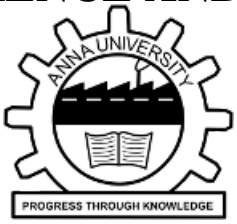
FACULTY OF INFORMATION AND COMMUNICATION ENGINEERING

in partial fulfillment of the requirements for the award of the degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

ANNA UNIVERSITY, CHENNAI – 25

APRIL 2019

BONAFIDE CERTIFICATE

Certified that this project report titled **UNOBTRUSIVE AUTHENTICATION OF SMARTPHONE USERS BY CONTEXT-AWARE BEHAVIOURAL BIOMETRICS** is the *bonafide* work of **SIVAA SHAANTH SUMAN R S (2015103048)**, **AISHWARYA V (2015103582)** and **SHOPHINE S (2015103609)** who carried out the project work under my supervision, for the fulfillment of the requirements for the award of the degree of Bachelor of Engineering in Computer Science and Engineering. Certified further that to the best of my knowledge, the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or an award was conferred on an earlier occasion on these or any other candidates.

Place: Chennai
Date:

Dr Sudha S
Assistant Professor (Sl. Gr.)
Department of Computer Science and Engineering
Anna University, Chennai – 25

COUNTERSIGNED

Head of the Department,
Department of Computer Science and Engineering,
Anna University Chennai,
Chennai – 600025

ABSTRACT

The widespread use of mobile phone in conjunction with the advent of smart things enable seamless and frictionless user experience. The traditional authentication techniques make it inadequate for users to fully realize the benefits of seamless and frictionless authentication.

Existing authentication techniques, face a number of weaknesses that includes degraded user experiences (requiring users to authenticate multiple times when the device is used), lack of user-specific service access rights (e.g., only user has access to application A), poor security practices (e.g., automatic sign-in, disabling passwords, or setting longer timeouts), insufficient security provided (e.g., PINs and passwords are susceptible to brute force, smudge, and social engineering attacks), and lack of continuous authentication (e.g., knowledge-based and biometric solutions that authenticates user once and leaves device in trusted state till timeout).

We aim at developing a solution using continuous and multifaceted behavioural biometrics authentication that passively authenticates the user in the background, and a context-aware risk assessment and access control that provides access to applications based on the perceived threat level.

ABSTRACT

ACKNOWLEDGEMENT

We express our atmost gratitude to our guide **Dr.S.Sudha** for guiding us with patience throughout the project. We thank her for leading us in the right direction through useful discussions and ideas. She also encouraged us during every phase of the project to come out with different perspectives to solve the problems. We value and appreciate her knowledge, principles and ability to bring out the best in us. Without her constant support and appreciation, we could have never done it to this extent.

We are very thankful to **Dr.S.Valli**, Head of the Department of Computer Science and Engineering, Anna University, for providing us with the facilities of the Department and for the constant support.

We are grateful to the panel of reviewers **Dr.S.Chitrakala,** Professor, Department of Computer Science and Engineering, Anna University, **Dr.Angelin Gladston,** Associate Professor, Department of Computer Science and Engineering, Anna University, **Dr.S.Renugadevi,** Assistant Professor(SI.Gr.), Department of Computer Science and Engineering, Anna University, for their insightful suggestions and critical reviews throughout the course of our project.

Sivaa Shaanth Suman R S **Aishwarya V** **Shophine S**

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

PCA Principal Component Analysis

SVM Support Vector Machine

HMM Hidden Markov Model

SM Sequence Matching

EER Equal Error Rate

GPS Global Positioning System

CHAPTER 1

INTRODUCTION

1.1 PROBLEM DOMAIN

Smartphones and tablets have become ubiquitous in our daily lives. Smartphones, in particular, have become more than personal assistants. These devices have provided new avenues for consumers to play, work, and socialize whenever and wherever they want. However, mobile devices are also susceptible to various problems[4].

Traditional knowledge-based authentication techniques such as PINs, passwords and pattern locks are inadequate to fully realize the current needs of the users due to low user-friendliness and insufficient security. On an average user check their device 150 times per day which in turn implies that users need to authenticate at most 150 times whenever the device is accessed. This inconvenience forces users to make less security conscious decisions such as leaving device unprotected or setting long timeouts before the device locks [1].

One of the greatest concerns is the possibility of a breach of security and privacy if the device is seized by an outside party. It is possible that threats can come from friends as well as strangers. Due to the size of smart devices, they can be easily lost and may expose details of users' private lives. In addition, this might enable pervasive observation or imitation of one's movements and activities, such as sending messages

to contacts, accessing private communication, shopping with a credit card, and relaying information about where one has been. This also increases concerns over the resilience of existing mobile authentication methods and their ability to safeguard the growing amount of sensitive information stored and processed on these devices[5]. This project presents context-aware security technology for responsive and adaptive protection (CASTRA), an always-on context-aware authentication and access control framework that seamlessly and unobtrusively authenticate users to mobile applications of varying sensitivity levels.

1.2 PROBLEM DESCRIPTION

The widespread use of mobile phone in conjunction with the advent of smart things enable seamless and frictionless user experience. The traditional authentication techniques make it inadequate for users to fully realize the benefits of seamless and frictionless authentication. Existing authentication techniques, face a number of weaknesses that include: degraded user experiences (requiring users to authenticate multiple times when the device is used), lack of user-specific service access rights (e.g., only user i has access to application A), poor security practices (e.g., automatic sign-in, disabling passwords, or setting longer timeouts), insufficient security provided (e.g., PINs and passwords are susceptible to brute force, smudge, and social engineering attacks), and lack of continuous authentication (e.g., knowledge-based and biometric solutions that authenticates user once and leaves device in trusted state till timeout). We aim at developing a solution using continuous and multifaceted behavioural biometrics authentication that passively authenticates the user in the background, and a context-aware risk assessment and access control that provides access to applications based on the per-

ceived threat level.

1.3 SCOPE

The number of mobile applications is also exploding these days and these various applications might require different levels of security. Indeed, a trade-off between usability and security needs to be taken into consideration. For instance, retrieving the users daily schedule does not require the same level of security as making an Internet banking transaction. Applying the same verification scheme to all applications requiring different levels of security would be somewhat cumbersome. Thus, it is necessary to provide miscellaneous authentication mechanisms on the mobile adapting to different security level requirements to optimize the user-device interaction. Accordingly, an implicit authentication technique needs to be investigated, which aims to enhance the user experience and ameliorate mobile security[2].

1.4 CONTRIBUTION

A novel gait recognition scheme which can be used for user verification or identification on mobile device that can adapt to the actual usage in reality is proposed. We pay attention to the context that the mobile is placed in the front pocket, which is the most appropriate location for the device in daily use.The study mainly focuses on addressing the instability problem of sensors orientation that frequently arises when the device is flexibly attached with its owner in practice. Furthermore, gait is likely to be considered as a behavioral biometric which is not as robust as other physiological traits since it is affected by many physical and environmental conditions, such as the clothing, footwear, ground material, mood, health, age, weight, etc. Therefore,

applying pattern matching to deal with all these circumstances could be inefficient. What is more, since the mobile is generally carried and accessed by its owner, gait signals can be captured frequently and continuously. We prefer to leverage machine learning techniques to adapt to the variation of the gait characteristics over time.

The proposed hidden Markov models to capture human path traces for authentication. One issue that stems from the fact of using location merely as an authentication factor is its inability to detect shared or insider attacks as both adversary and owner share the same location. To tackle this, we propose a solution using continuous and multifaceted behavioural biometrics authentication that passively authenticates the user in the background, and a context-aware risk assessment and access control that provides access to applications based on the perceived threat level.

1.5 SWOT ANALYSIS

1.5.1 Strength

The Support Vector machine is used to train the system which provides higher accuracy when compared to other machine learning models. We have used Principal component analysis to give better results. This helps in dimensionality reduction which leads to increase of accuracy.

1.5.2 Weakness

The model works under certain constraints only. The real-time implementation of the project will require vigorous data collection for a extended period of time. The gait model development for relative position of the phone is still under research.

1.5.3 Opportunity

The Biometric authentication system will act as an additional layer of security which will be impossible to break when implemented accurately.

1.5.4 Threat

The project can only work with fixed positions. The system is only implemented with a limited number of subjects. The smartphones have to be rooted to implement this application and allowed to run in the background to monitor all applications. Doing so may sometime lead to negative results.

1.6 ORGANISATION OF THESIS

Chapter 2 includes the literature survey which was performed on the project relevant topics. It includes a concept wise survey on the various ideologies that were put forth in the project.

Chapter 3 consists of the requirement analysis which includes functional and non-functional requirements where non-functional requirements brief us upon the user interface, hardware and software. This chapter also gives a touch up on the dataset description, usecase and sequence diagrams.

Chapter 4 consists of the system architecture which includes the block diagram and its detailed description. This chapter also includes UI design and module wise description of the block diagram.

Chapter 5 consists of the system development which includes the prototype across the modules and the various pseudocodes and algorithm which were used for the project.

Chapter 6 consists of the results of the various modules that were discussed in the previous chapters. This includes screenshots of the vari-

ous stages in the project. In the end of the chapter the evaluation metrics are also discussed. Chapter 7 consists of the conclusion of the project and the future works.

CHAPTER 2

RELATED WORK

The section shows us the similar works carried out in the biometric authentication. The unobtrusive authentication of smartphones is an additional layer of security. Combining gait,location and proximity and building a trustscore gives better results than the existing methods.

2.1 CASTRA: SEAMLESS AND UNOBTRUSIVE AUTHENTICATION OF USERS TO DIVERSE MOBILE SERVICES

D. M. Shila and K. Srivastava [4] presents context-aware security technology for responsive and adaptive protection (CASTRA), an always-on context-aware authentication and access control framework that seamlessly and unobtrusively authenticate users to mobile applications of varying sensitivity levels. It uses a continuous and multifaceted behavioural biometrics authentication that passively authenticates the user in the background while the device is being in contact with the user, and a context-aware risk assessment and access control that provides access to applications based on the perceived threat level around the device.

The performance of CASTRA was evaluated under natural settings, on 15 subjects, using different variants of the Samsung devices. Multiple realistic attack scenarios targeting mobile devices were designed to prove the security and user-friendliness of the proposed scheme[3].

They also present techniques to reduce energy and bandwidth consumption and ways to unobtrusively acquire data.

2.2 ON THE INSTABILITY OF SENSOR ORIENTATION IN GAIT VERIFICATION ON MOBILE PHONE

Authentication schemes using tokens or biometric modalities have been proposed to ameliorate the security strength on mobile devices. While the gait signal captured by inertial sensors is understood to be a reliable profile for effective implicit authentication, recent studies have been conducted in ideal conditions and might therefore be inapplicable in the real mobile context. The acquiring sensor is always fixed to a specific position and orientation.

Thang Hoang , Deokjai Choi and Thuc Nguyen[5], focus on addressing the instability of the sensors orientation which mostly happens in reality. A flexible solution taking advantages of available sensors on mobile devices which can help to handle this problem is presented. Moreover, a novel gait recognition method utilizes statistical analysis and supervised learning to adapt itself to the instability of the biometric gait under various circumstances is also proposed. By adopting PCA+SVM to construct the gait model, the proposed method outperformed other state-of-the-art studies, with an equal error rate of 2.45 percent and accuracy rate of 99.14 percent in terms of the verification and identification aspects being achieved, respectively. They addressed the sensor disorientation problem in gait verification or identification systems which can frequently arise in reality, especially in the mobile context. A simple but effective solution taking advantages of available sensors in a mobile device was proposed.

2.3 A MULTI-FACETED APPROACH TO USER AUTHENTICATION FOR MOBILE DEVICES - USING HUMAN MOVEMENT, USAGE AND LOCATION PATTERNS

Physiological biometrics such as fingerprint, voice or facial recognition can enable user- friendly and strong security, but they only provide single-shot authentication and lack ability to continuously authenticate the user. In this paper, they take a different approach by designing a multi-faceted authentication scheme termed as mAuth that continuously and unobtrusively authenticates the user while the device is being in contact with the user[1]. By leveraging a combination of supervised and unsupervised learning techniques on the raw low-level sensor data from the mobile device, multiple inferences about the user (or higher-level contexts) such as the frequently visited locations, physical proximity with the device (carrying in the pocket or placed on the table), and gait patterns are extracted.

These multiple high-level contexts regarding the user are further fused to generate a dynamic trust score that determines the degree to which the user is trustworthy to access the applications. Experiments demonstrate the performance of the individual learning algorithms as well as the overall method in identifying users under natural settings. Various attack scenarios targeting mobile devices are designed to prove the security of the proposed approach. They also explore ways to unobtrusively acquire data for supervised learning algorithms without explicit user annotation.

2.4 PATH: PERSON AUTHENTICATION USING TRACE HISTORIES

Recognizing human behavior and understanding user mobility from sensor data is an interesting and challenging problem in ubiquitous computing.

Upal Mahbub and Rama Chellappa [2], propose a solution to the problem of Active Authentication using trace histories is addressed. Specifically, the task is to perform user verification on mobile devices using historical location traces of the user as a function of time. Considering the movement of a human as a Markovian motion, a modified Hidden Markov Model (HMM)-based solution is proposed.

The proposed method, namely the Marginally Smoothed HMM (MSHMM), utilizes the marginal probabilities of location and timing information of the observations to smooth- out the emission probabilities while training. Hence, it can efficiently handle unforeseen observations during the test phase. The verification performance of this method is compared to a sequence matching (SM) method, a Markov Chain-based method (MC) and an HMM with basic Laplace Smoothing (HMM- lap). Experimental results using the location information of the UMD Active Authentication Dataset-02 (UMDAA02) and the GeoLife dataset are presented. The proposed MSHMM method outperforms the compared methods in terms of equal error rate (EER). Additionally, the effects of different parameters on the proposed method are discussed.

CHAPTER 3

REQUIREMENTS ANALYSIS

3.1 FUNCTIONAL REQUIREMENTS

The application output detects an intruder for a given set of dataset. The output of the detection should adhere to the following requirements:

- The output should be able to detect the level of accessibility that should be given to an user.
- The application must be able to block the required applications automatically in case of an intruder detection.
- For any user, the access to applications should be according to the trust score generated.

3.2 NON-FUNCTIONAL REQUIREMENTS

3.2.1 User Interface

There must be a simple and easy to use application where the user should be authenticated automatically based on his gait and location data. Machine learning is used to build the model using the selected features as input.

3.2.2 Hardware

The hardware requirements for this android app for end user side includes :

- 2 GB ram in the smart phone.
- Android version greater than 4.

- Permission setup in the mobile.
- Good internet connection.
- Mobile sensors

3.2.3 Software

- 3 Gb RAM minimum (8 Gb is recommended).
- Java Development Kit (JDK 7).

  For the Node server , it will be implemented in laptop or PC with some requirements which includes:
- 8 GB ram is recommended
- NodeJs Setup
- MATLAB R2018 with 'libsvm'

3.3 DATASET DESCRIPTION

The first component of the system is the dataset. The data for the experiments can be gathered using a sensor tracking application built for Android. The basic architecture of the application focuses on a launcher class which receives updates from sensors and launches an intent service thread in the background at a user specified interval (our experiments used 50Hz). The application makes use of Google ActivityRecognition API in addition to all of the sensor data available (GPS, Accelerometer, Magnetometer).

The first sensor needs to be activated to capture gait signal is the mobile accelerometer. The accelerometer senses forces acting on the mobile in the three orthogonal axes of X,Y,Z (Figure 1a). A sequence of acceleration samples output by the accelerometer during walking is recognized as the gait signal. Each sample is a 3-dimensional vec- tor, wherein each component is a combination of the forces of gravity and user motion acting on each dimension.

a = (a(X),a(Y),a(Z))

Due to the characteristics of the accelerometer, the raw acceleration samples always comprise gravitational acceleration components. In order to obtain samples which only involve pure gait signals of individuals, we eliminate the impact of gravity by additionally activating a virtual sensor of gravity to determine the gravitational acceleration components on the 3 axes of the mobile during the gait capture process. The output of the gravity sensor is a 3-component vector.

g = (g(X),g(Y),g(Z))

The gait data is collected and preprocessed using dimensionality reduction.The dataset set collected is used for training and testing the system.

3.3.1 Constraints And Assumptions

Constraints

The model is built using the dataset that was already available. To collect a proper gait dataset the required time is more than a year. The training and testing was done using the existing dataset.

Assumptions

The user places the phone in the pocket with the gps always on (i.e) the gait segment and session data can be obtained only if the user keeps the phone in the pocket and the GPS should be always on to validate the user's location.

3.4 SYSTEM MODELS

3.4.1 Use Case Diagram

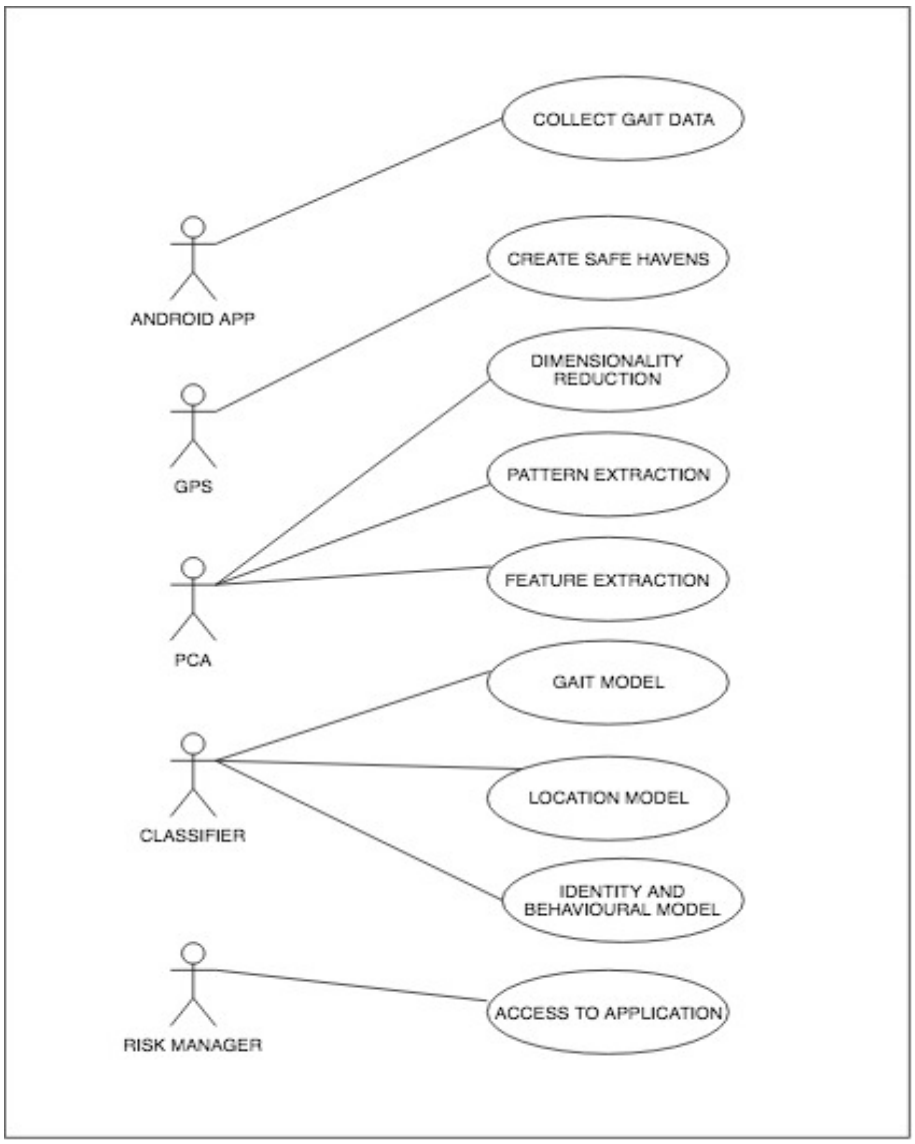


Figure 3.1 Use Case Diagram

Figure 3.1 shows the Use Case Diagram of the project. It represents the overall outline of the project modules.

3.4.2 Sequence Diagram

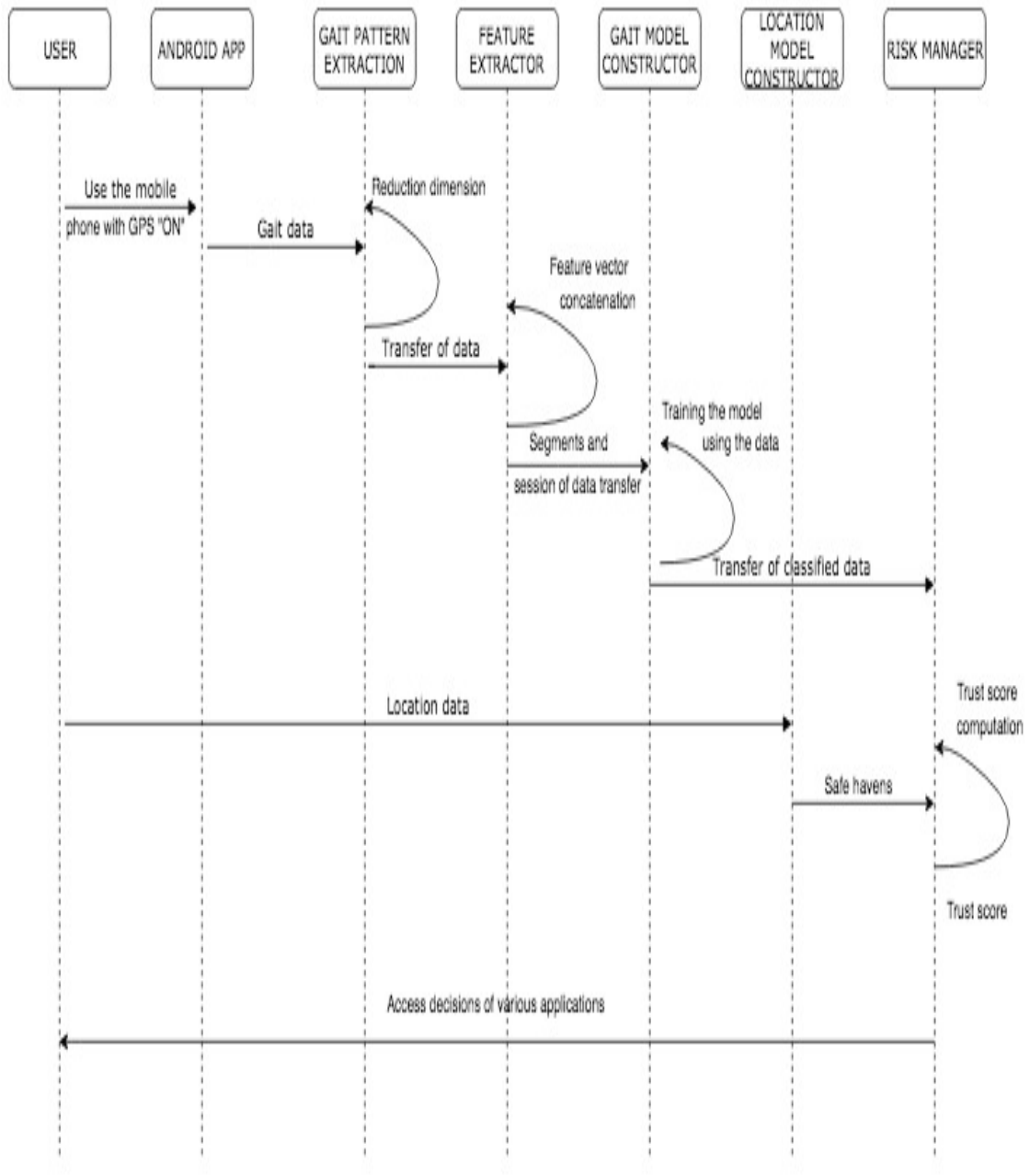


Figure 3.2 Sequence Diagram

Figure 3.2 shows the Sequence Diagram of the project. It helps visualize the flow of the project.

CHAPTER 4

SYSTEM DESIGN

4.1 ARCHITECTURE DIAGRAM

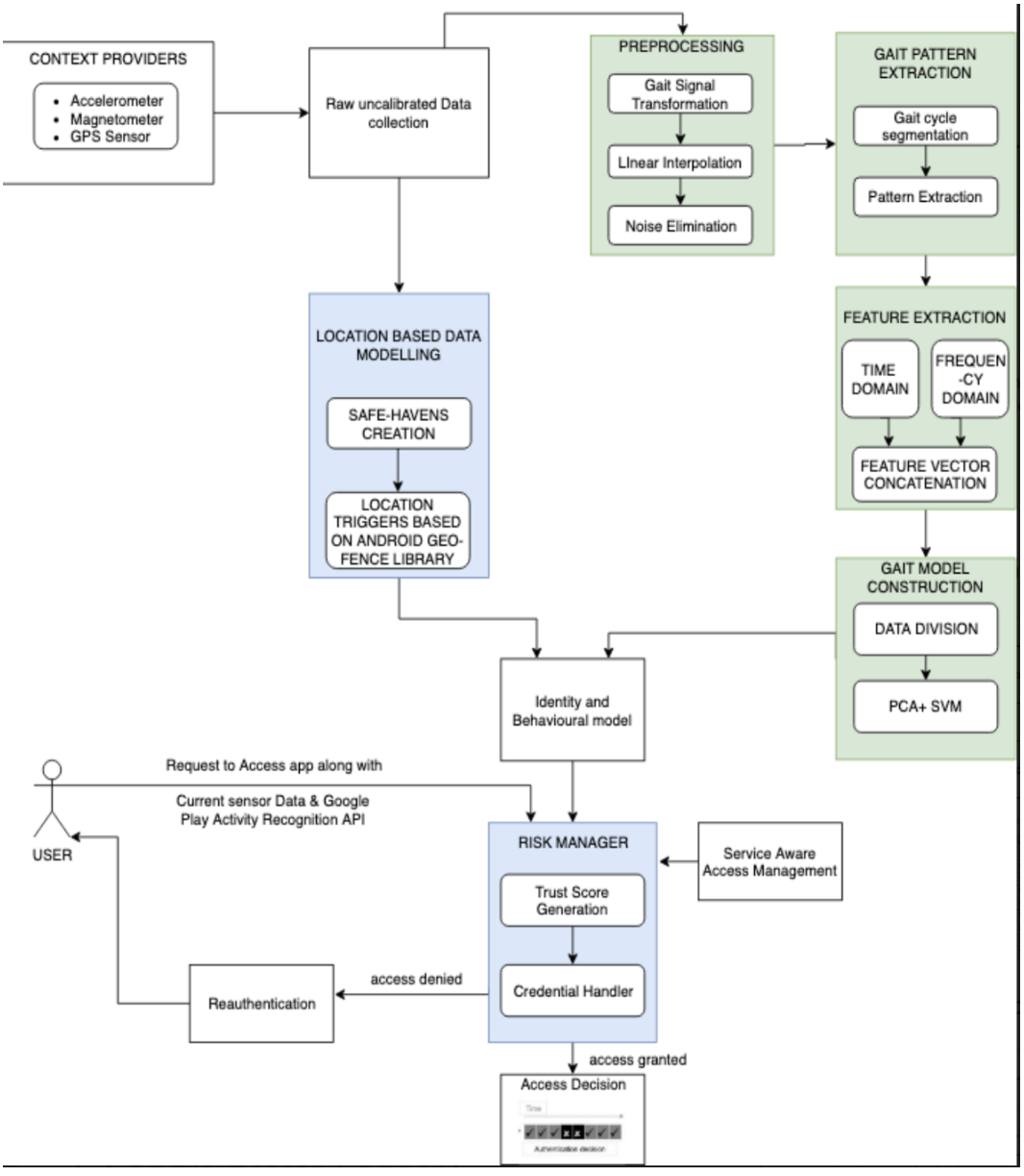


Figure 4.1 Architecture Diagram

Figure 4.1 shows the complete architecture diagram of the project. An android application is created to collect senor data from the users.Location model and Gait model are constructed. The two models are combined to build the Identity and Behavioural Model. The data from the model is used to compute the trust score.

4.2 MODULES

Table 4.1 Modules

Module	Description
Context Logger	An application is created and installed to collect the raw data.
Gait Processing	It takes the raw data log and transforms into earth coordinates to account for disorientation.
Gait Pattern Extraction	The module is used to segment separate gait signals from the preprocessed data.
Gait Model Construction	The module focuses on feature extraction from the extracted gait patterns.
Location Based Data Model	It builds a location based data model.
Risk Manager	This module decides whether an identified user/app has authorization rights to access a specific resource under the given context.

Table 4.1 shows the list of modules and their descriptions.

4.3 DETAILED MODULE DESIGN

4.3.1 Context Logger

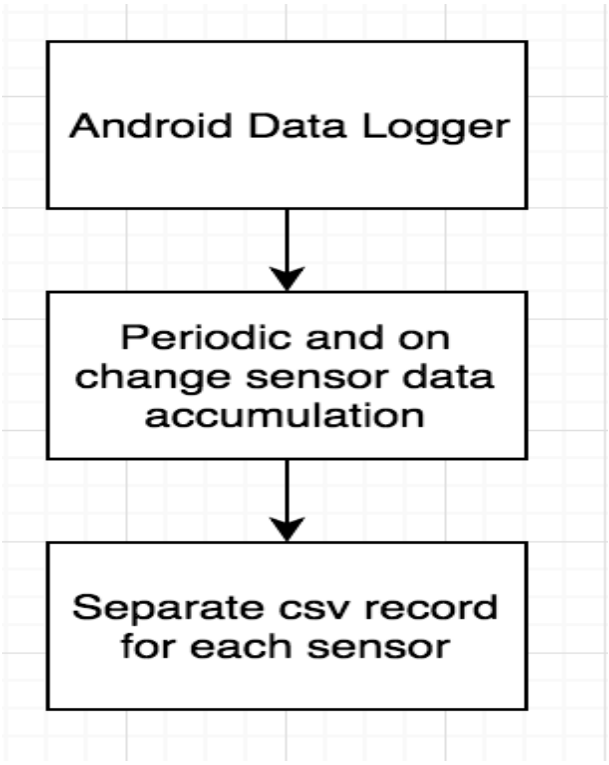


Figure 4.2 Context Logger

Figure 4.2 shows that the data for the experiments can be gathered using a sensor tracking application built for Android. An application will be installed to the user's phone which will periodically collect the data such as the location of the user, frequently visited locations of the user, accelerometer and magnetometer sensor data and log them into separate csv files.

4.3.2 Gait Processing

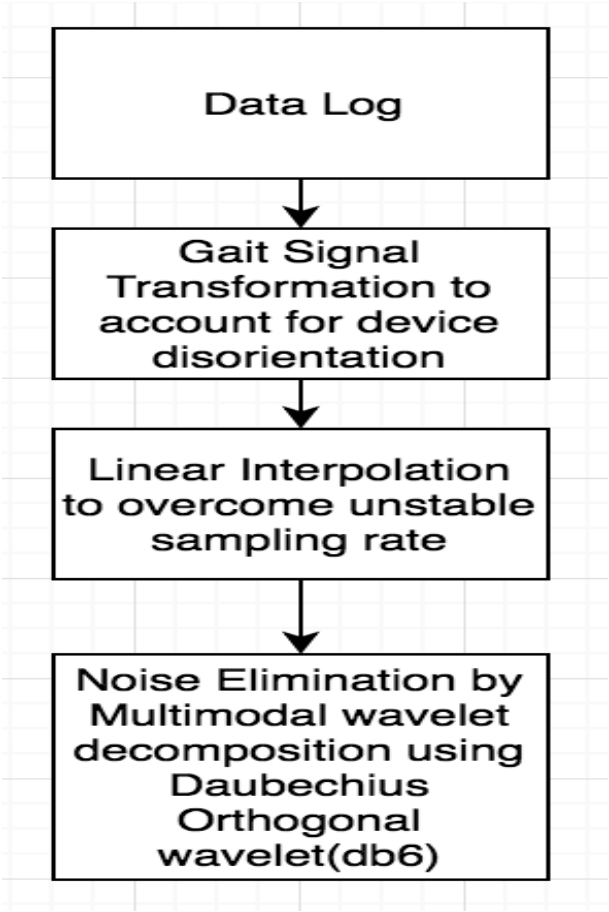


Figure 4.3 Gait Processor

Figure 4.3 shows the Gait processing module. It takes the raw data log from the accelerometer and further transforms them from mobile co-ordiantes to Earth coordinates to account for device disorientation. Since the sampling rate of mobile phones are unstable, linear interpolation is carried out to make it stable. Further the noise from the data is eliminated by using multi-model wavelet decomposition using Daubechies Orthogonal Wavelet (db6).

4.3.3 Gait Pattern Extraction

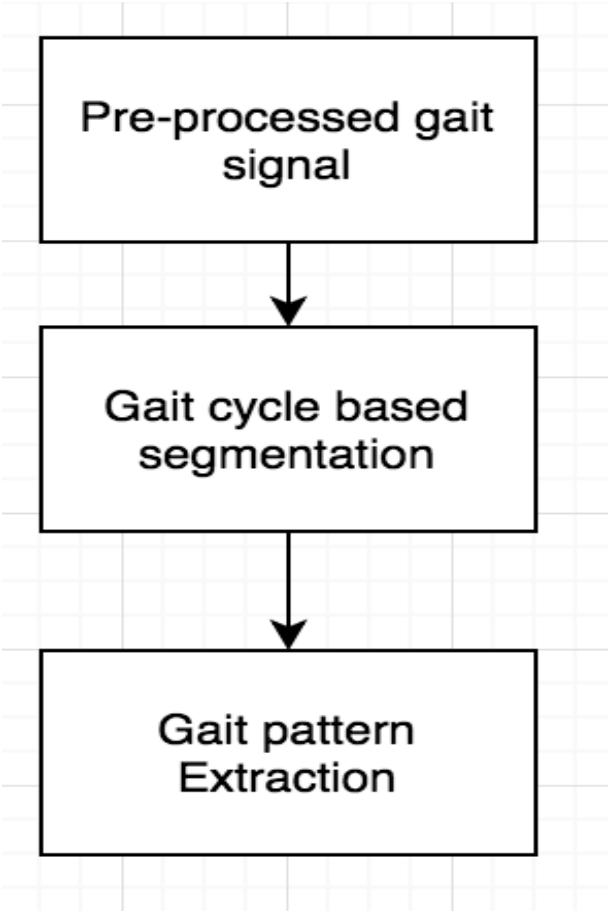
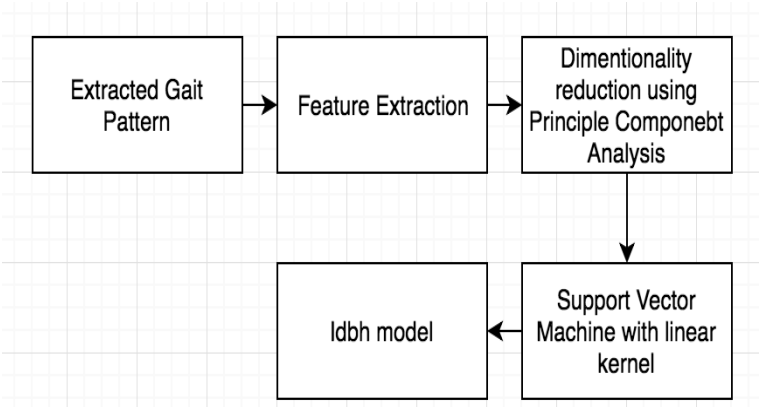


Figure 4.4 Gait Pattern Extraction

Figure 4.4 shows how to segment separate gait signals from the preprocessed data. First we filter the insignificant peaks that fall below a certain threshold. Several Gait segments are combined together to produce gait patterns. These patterns will be unique for every individual. These patterns will be used further in the next module.

4.3.4 Gait Model Construction



EXTRACTED FEATURES FROM SENSOR STREAMS

Sensor	Feature
Accelerometer (Tri-axial) Gyroscope (Tri-axial)	Mean value
	Standard Deviation
	Sum of absolute deviation from mean
	Binned histograms (ten bins)
	Pairwise correlation between axes
	Mean Log Energy
	RMS value
Magnetometer (Tri-axial)	Mean value
	Standard Deviation
	Sum of absolute deviation from mean
Proximity Sensor	Binned histograms (two bins)

Figure 4.5 Gait Model

Figure 4.5 focuses on feature extraction from the extracted gait patterns (Both in Frequency and Time domain) . Since the computation is primarily going to be handled in mobile phones, dimensionality reduction is carried out to reduce the workload by using PCA. Further SVM with a linear kernel is used to build the Gait model which in turn contributes to the Identity and Behavioral model.

4.3.5 Location Based Data Model

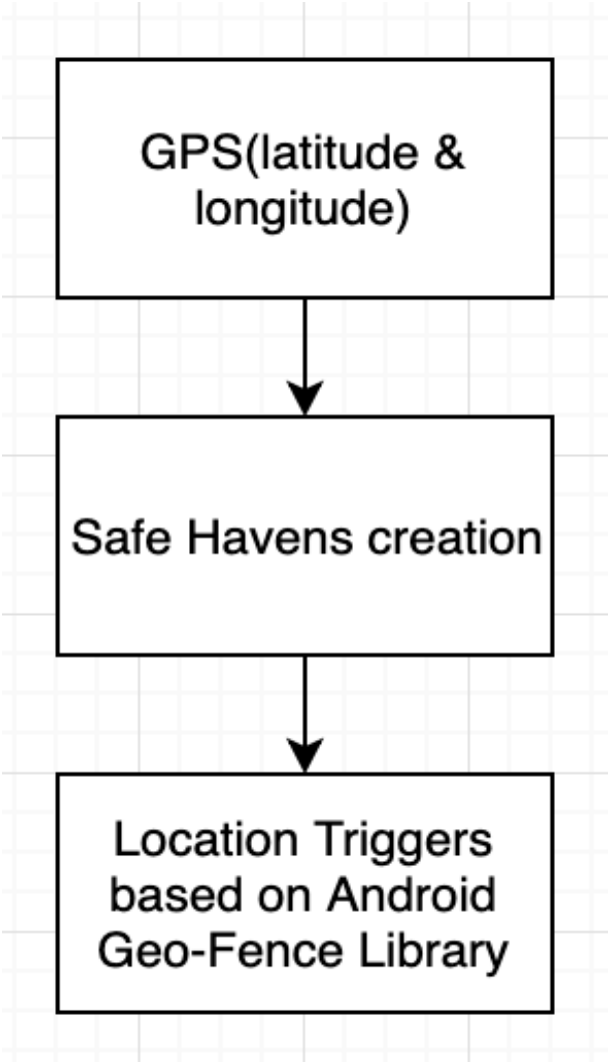


Figure 4.6 Location Based Data Model

Figure 4.6 shows how to construct a Location based data model that also contributes to the Identity and Behaviour model with some weightage. Android The frequently visited locations are collected and safe havens is created. Then separate weightage is given based on the location of the user.

4.3.6 Risk Manager

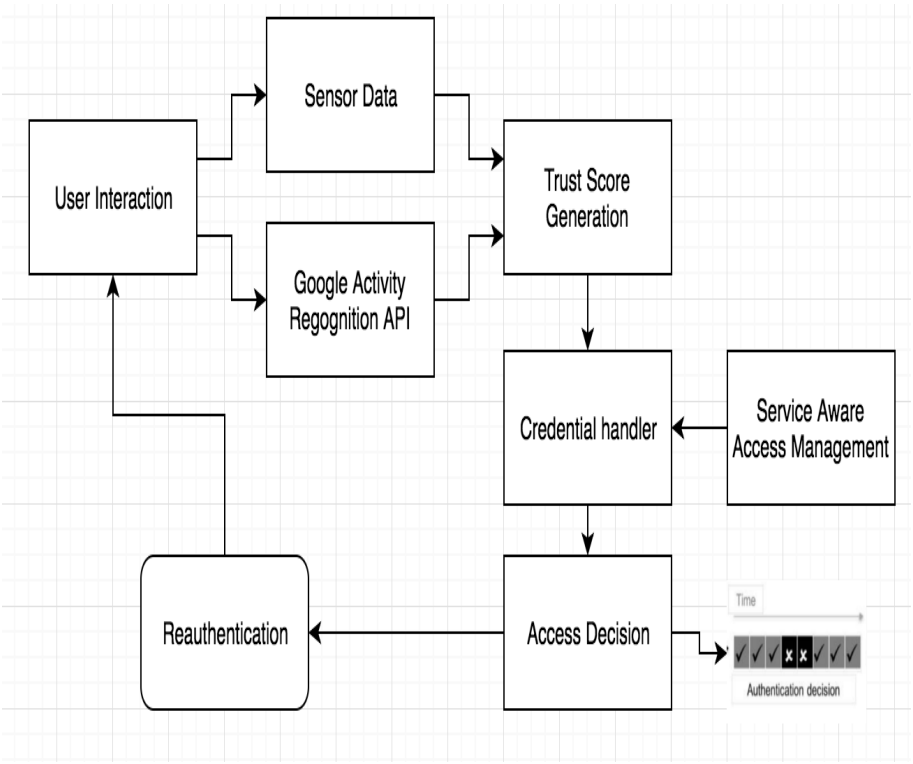


Figure 4.7 Risk Manager

Figure 4.7 shows how the risk manager tries to decide whether an identified user/app has authorization rights to access a specific application based on the overall trustscore. Based on the trust score generated the user will be given access to high level secured, medium level secured and low level secured applications. For instance, assume high and low trustscores for financial and entertainment apps, respectively. If the trustscore is very low then the user will not be granted access to a financial app unless a high trust score is observed around the device.

CHAPTER 5

SYSTEM DEVELOPMENT

The input and output to each module of the system is described in this section.

Figure 5.1 shows the list of modules used to develop the android application.

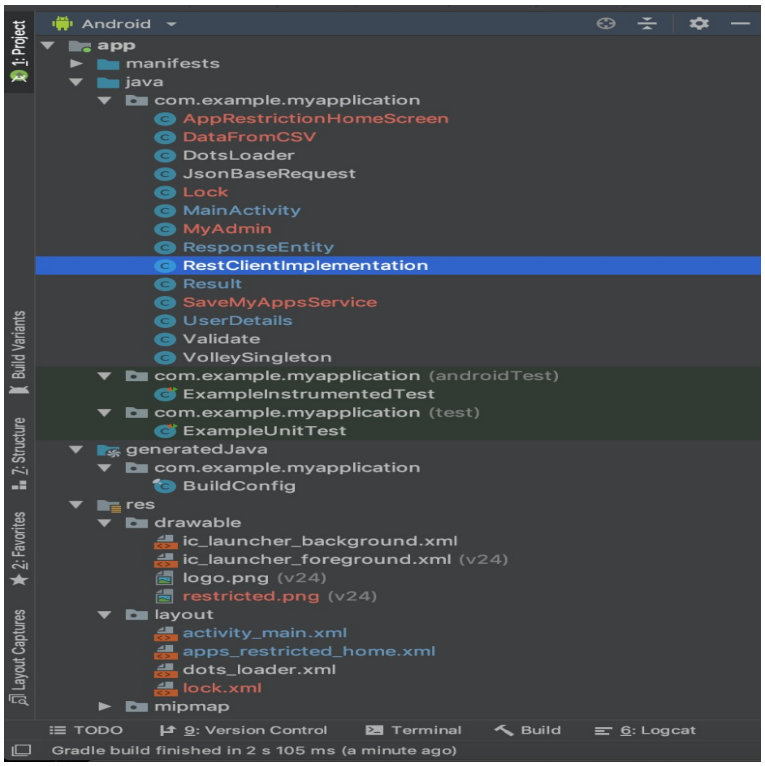


Figure 5.1 Code Overview Of Android Application

5.1 MODULE IMPLEMENTATION DETAILS

The input and output to each module of the system along with the algorithms used are described in this section.

5.1.1 Gait Processing

The input to this module would be the sensor data collected. The raw accelerometer data log is preprocessed to eliminate the noise in the data. Linear Interpolation is applied to stabilize the sampling rate. After that the db6 wavelet is used to eliminate noise in the stabilized data.

Algorithm 1: Algorithm for Gait Processing

Data: Data Log

Pseudocode:

1. preprocess()
2. signal-transf()
3. transform sample to earth coordinates from mobile coordinates;

$$\mathbf{A} = \left[\mathbf{a}^{(Z)}\ \mathbf{a}^{(XY)}\ \mathbf{a}^{(M)}\right] = \begin{bmatrix} a_1^{(Z)} & a_1^{(XY)} & a_1^{(M)} \\ \vdots & \vdots & \vdots \\ a_i^{(Z)} & a_i^{(XY)} & a_i^{(M)} \\ \vdots & \vdots & \vdots \\ a_n^{(Z)} & a_n^{(XY)} & a_n^{(M)} \end{bmatrix}$$

4. linear-interpolation()
5. Stabilize the sampling rate;
6. noise-elimination()
7. multi-level-wavelet-decomposition(Daubechisius Orthogonal wavelet, level 2);

Result: Pre-processed Gait Data

5.1.2 Gait Pattern Extraction

The Gait segment is separated from the obtained processed data. Then the peak positions are obtained. The insignificant peak positions are filtered. From the filtered segments patterns are extracted.

Algorithm 2: Algorithm for Gait Patternn Extraction

Data: Pre-processed Gait Signal

Pseudocode:

1. GaitPatternExtract()
2. GaitCycleSegment()
3. getAllPeaksPositions();
4. filterInsignificantPeaks();
5. Calculate autocorrelation();

$$c_t = \frac{N}{N-t} * \frac{\sum_{i=1}^{N-t} a_i^{(Z)} a_{i+t}^{(Z)}}{\sum_{i=1}^{N} (a_i^{(Z)})^2} \tag{5.1}$$

6. FindGaitLengthCycle();
7. ExtractGaitCyclemarks();
8. StoreGaitStartingPoints();
9. GaitPatternExtraction()
10. Define n;
11. Concatenate n segments into 1 pattern each;
12. Plot the patterns;

Result: Extracted Gait Pattern

5.1.3 Gait Model Construction

Dimensionality reduction is done and the model is constructed.

Algorithm 3: Algorithm for Gait Model Construction

Data: Extracted Gait Pattern

Pseudocode:

1. GaitModelConstruction()
2. forEachGaitPattern :
3. Extract TimeDomainFeatures();
4. Extract FrequencyDomainFeatures();
5. featureVectorDimRedPCA()
6. Feature Component,

$$\mathbf{F}^{\top} = \begin{bmatrix} \mathbf{v}_1^{(1)} \\ \vdots \\ \mathbf{v}_j^{(i)} \\ \vdots \\ \mathbf{v}_M^{(N)} \end{bmatrix} = \begin{bmatrix} f_{1,1}^{(1)} & f_{1,k}^{(1)} & f_{1,n_f}^{(1)} \\ \vdots & \vdots & \vdots \\ f_{j,1}^{(i)} & f_{j,k}^{(i)} & f_{j,n_f}^{(i)} \\ \vdots & \vdots & \vdots \\ f_{M,1}^{(N)} & f_{M,k}^{(N)} & f_{M,n_f}^{(N)} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_t \\ \vdots \\ \mathbf{v}_{MN} \end{bmatrix} \in \mathbb{R}^{MN \times n_f}.$$

7. Covariance of F =

$$\sum = \frac{1}{MN} \sum_{i=1}^{MN} (v_i - \overline{v})(v_i - \overline{v})^T \in R^{n_f * n_f} \tag{5.2}$$

8. Sort Eigen Values in descending order;
9. Construct eigen vector;

$$U = [u_1 ... u_i ... u_k] \in R n_f * k \tag{5.3}$$

10. Dimension reduced matrix =¿

$$\overline{F}^T = F^T U \tag{5.4}$$

Result: Gait Model is added to Idbh model

5.1.4 Location Data Model

The smartphone GPS is used to collect the Location Data. The frequent sites visited by the user is collected and safe havens is created. Based on the current location of the user a small weightage is given. Also if the user is inside the safe haven a notification is sent to the user.

Algorithm 4: Algorithm for Location Data Model

Data: Raw Data log

Pseudocode:

```
1. get-API-key();
2.     getSystemPermission(Location,Notification);
3.     getCurrentGPSvalue();
4.     setReminder();
5.         createGeoFence(circumference — LatLong);
6.     onLocationChange();
7.         if(LocationWithinFence)
8.             sendNotification(user,location);
9.     buildTrustScore();
```

Result: Location model added to Idbh model

5.2 RISK MANAGER

The idbh model is connected to the Risk manager. Based on the gait score and location score a combined trust score is generated. Based on the trustscore user access to various appliations is granted.

Algorithm 5: Algorithm for Risk Manager

Data: User interactions Idbh Model

Pseudocode:

```
1. Predicted(userWalkingSegment)
2.     [Predicted-label, label-probability] =
   svmpredict(model,UserWalkingSegment)
3.     Refurn [Predicted-label, label-probability]
4.     gait-score= label-probability ;
5. GpsScoreCalc()
6.     Case inside safe haven : gps-score =1 ;
7.     Case Outside safe haven : gps-score =0;
8.     trust score = 0.9*gait-score +0.1*gps-score ;
9. CompareTrustScore()
10.    AcquireAccessLevel();
11.    AuthenticateAccess() ;
12.    case true :
13.        allow user to access;
14.    case false :
15. reauthenticate;
```

Result: Authentication and Access level Authorisation

CHAPTER 6

RESULTS AND DISCUSSION

6.1 SENSOR DATA COLLECTION

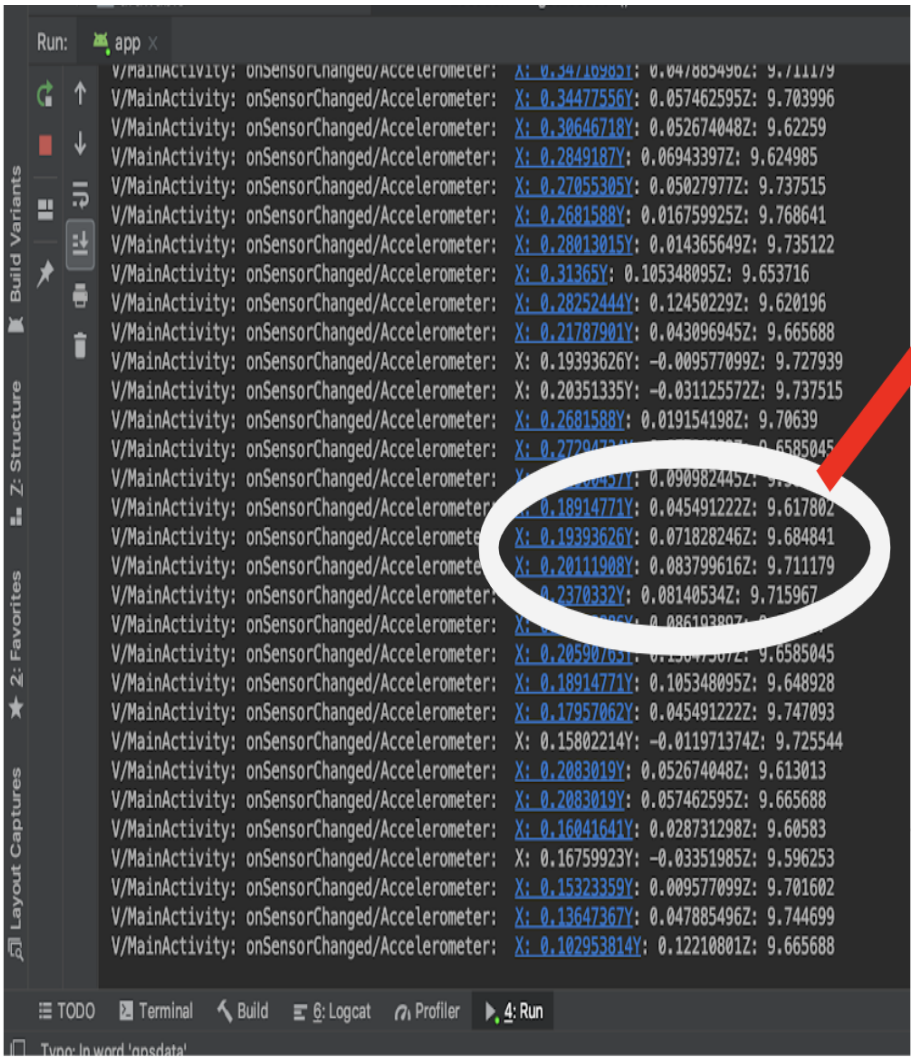


Figure 6.1 Accelerometer Data Collection

Figure 6.1 shows accelerometer data being collected from the mobile phone sensors.

gyrodata

X	Y	Z	Timestamp
4.2638395E-04	5.5466563E-04	-0.0028252518	1548657628
-0.0026279422	5.5466[illegible]-04	-0.0016035212	1548657628
0.0010372492	[illegible]5309	-0.003436117	1548657628
-0.0014062119	[illegible]655309	-0.0016035212	1548657628
-1.8448131E	[illegible]1655309	-0.0028252518	1548657628
-0.00[illegible]077	0.0011655309	-0.0016035212	1548657629
-0[illegible]6279422	0.0011655309	-0.0016035212	1548657629
[illegible].0014062119	0.0011655309	-0.0028252518	1548657629
-0.002017077	0.0011655309	-0.0022143864	1548657629
4.2638395E-04	0.0011655309	-0.0022143864	1548657629
-1.8448131E-04	0.0011655309	-0.0022143864	1548657629
-1.8448131E-04	5.5466563E-04	-0.0022143864	1548657630
-7.9534657E-04	0.0011655309	-0.0022143864	1548657630
0.0010372492	0.0072741834	-0.0022143864	1548657630
-0.002017077	5.5466563E-04	-0.0022143864	1548657630
-0.0014062119	0.0011655309	-0.0022143864	1548657630
-7.9534657E-04	5.5466563E-04	-0.0022143864	1548657631
-0.002017077	5.5466563E-04	-0.0022143864	1548657631
-1.8448131E-04	5.5466563E-04	-0.0022143864	1548657631
-0.0014062119	5.5466563E-04	-0.0028252518	1548657631
-1.8448131E-04	0.0011655309	-0.0022143864	1548657631
-0.0026279422	0.0017763962	-0.0016035212	1548657631
-7.9534657E-04	5.5466563E-04	-0.0016035212	1548657632
-1.8448131E-04	0.0011655309	-0.0022143864	1548657632
-1.8448131E-04	0.0011655309	-0.0016035212	1548657632

magnodata

X	Y	Z	Timestamp
-15	133.6	118.4	1548657628
-12.2	134.5	117.5	1548657628
-11.6	134.7	116.3	1548657628
-12.400001	134.7	117.9	1548657628
-13.5	135.8	117.3	1548657628
-12.6	134.6	119.200005	1548657628
-14.1	134.90001	116.1	1548657629
-12.900001	135.40001	117.4	1548657629
-11.1	135.1	118	1548657629
-11	135.1	118	1548657629
-12.8	135.6	117	1548657629
-13.1	136	117.700005	1548657629
-12.8	135.6	117.700005	1548657630
-13.1	135.5	118.700005	1548657630
-11.3	135.6	118	1548657630
-14	134.90001	117.4	1548657630
-13	136.3	118.4	1548657630
-11.900001	134.8	118.3	1548657631
-13.7	135.2	118	1548657631
-14.5	134.6	118.700005	1548657631
-13.1	135.2	118.4	1548657631
-12.400001	135	119.3	1548657631
-13.3	134.8	119	1548657631
-12.7	135.2	119.6	1548657632
-13.900001	135.2	119	1548657632

Figure 6.2 Collected Sensor Data

Figure 6.2 shows the collected sensor data.

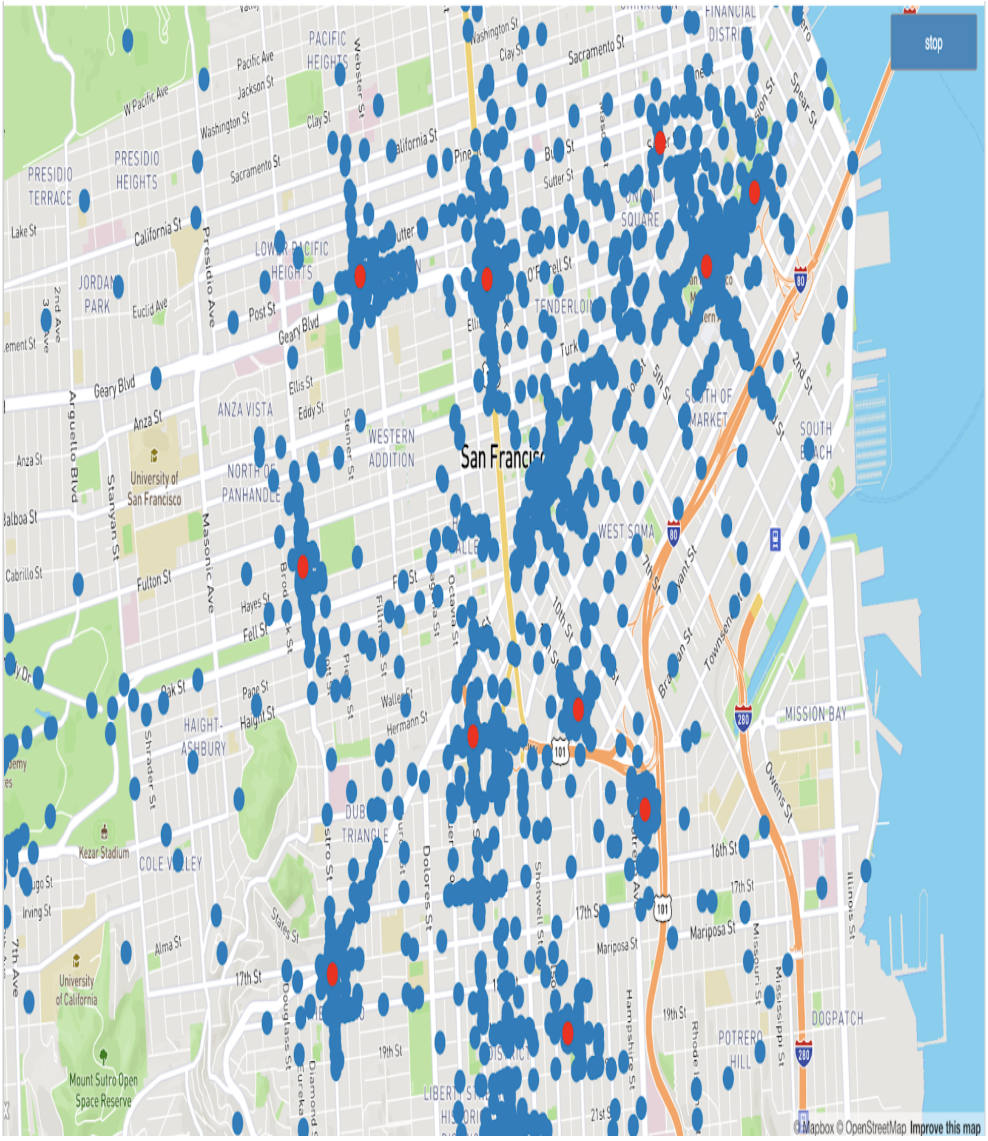
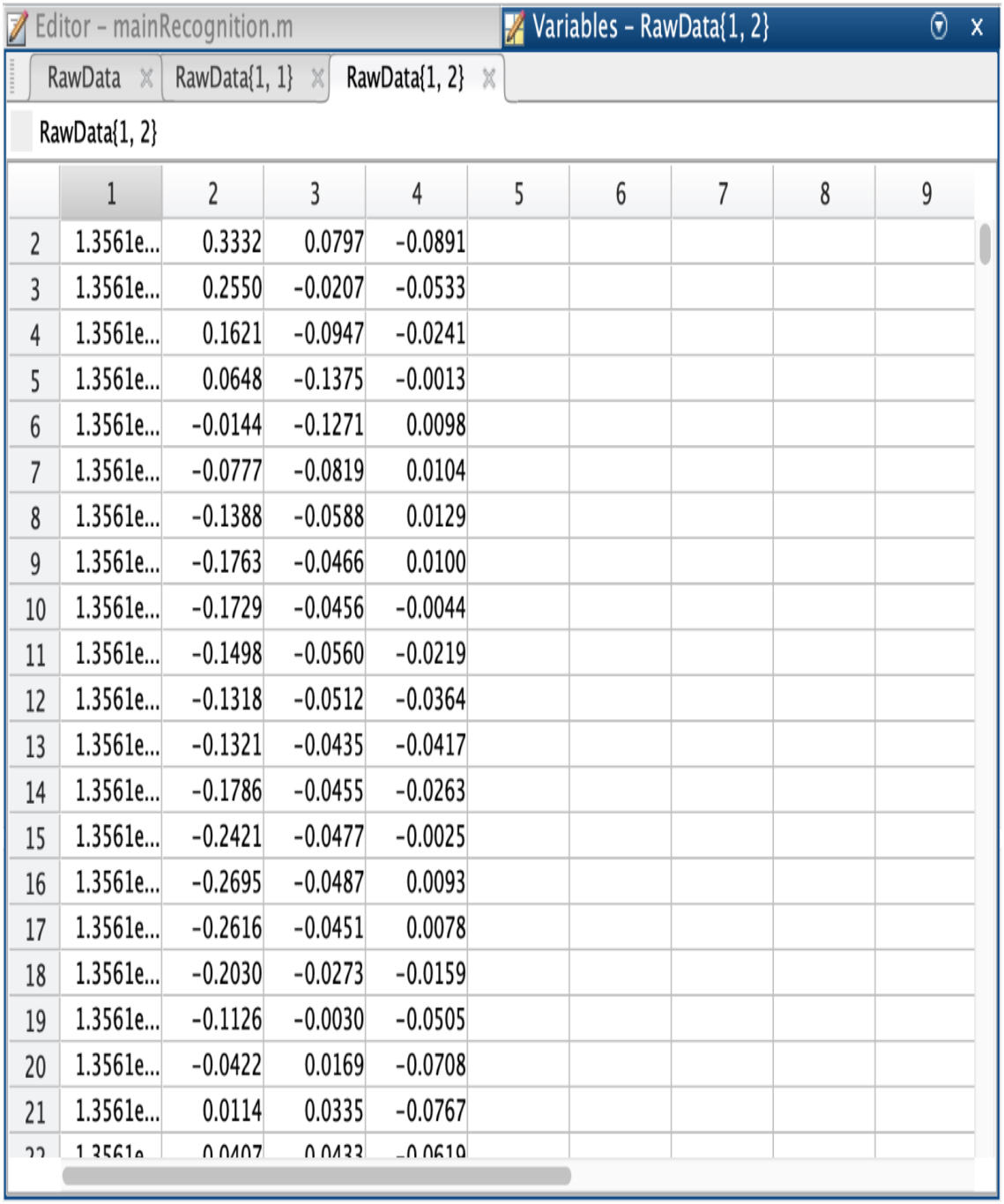


Figure 6.3 Location Cluster Visualization

Figure 6.3 shows clustering of user location data to plot his frequently visited places.



	1	2	3	4	5	6	7	8	9
2	1.3561e...	0.3332	0.0797	-0.0891					
3	1.3561e...	0.2550	-0.0207	-0.0533					
4	1.3561e...	0.1621	-0.0947	-0.0241					
5	1.3561e...	0.0648	-0.1375	-0.0013					
6	1.3561e...	-0.0144	-0.1271	0.0098					
7	1.3561e...	-0.0777	-0.0819	0.0104					
8	1.3561e...	-0.1388	-0.0588	0.0129					
9	1.3561e...	-0.1763	-0.0466	0.0100					
10	1.3561e...	-0.1729	-0.0456	-0.0044					
11	1.3561e...	-0.1498	-0.0560	-0.0219					
12	1.3561e...	-0.1318	-0.0512	-0.0364					
13	1.3561e...	-0.1321	-0.0435	-0.0417					
14	1.3561e...	-0.1786	-0.0455	-0.0263					
15	1.3561e...	-0.2421	-0.0477	-0.0025					
16	1.3561e...	-0.2695	-0.0487	0.0093					
17	1.3561e...	-0.2616	-0.0451	0.0078					
18	1.3561e...	-0.2030	-0.0273	-0.0159					
19	1.3561e...	-0.1126	-0.0030	-0.0505					
20	1.3561e...	-0.0422	0.0169	-0.0708					
21	1.3561e...	0.0114	0.0335	-0.0767					
22	1.3561e	0.0407	0.0433	-0.0619					

Figure 6.4 Calibrated And Interpolated Gait Data

Figure 6.4 shows gait data that has been calibrated to address dis-orientation issues and interpolated data.

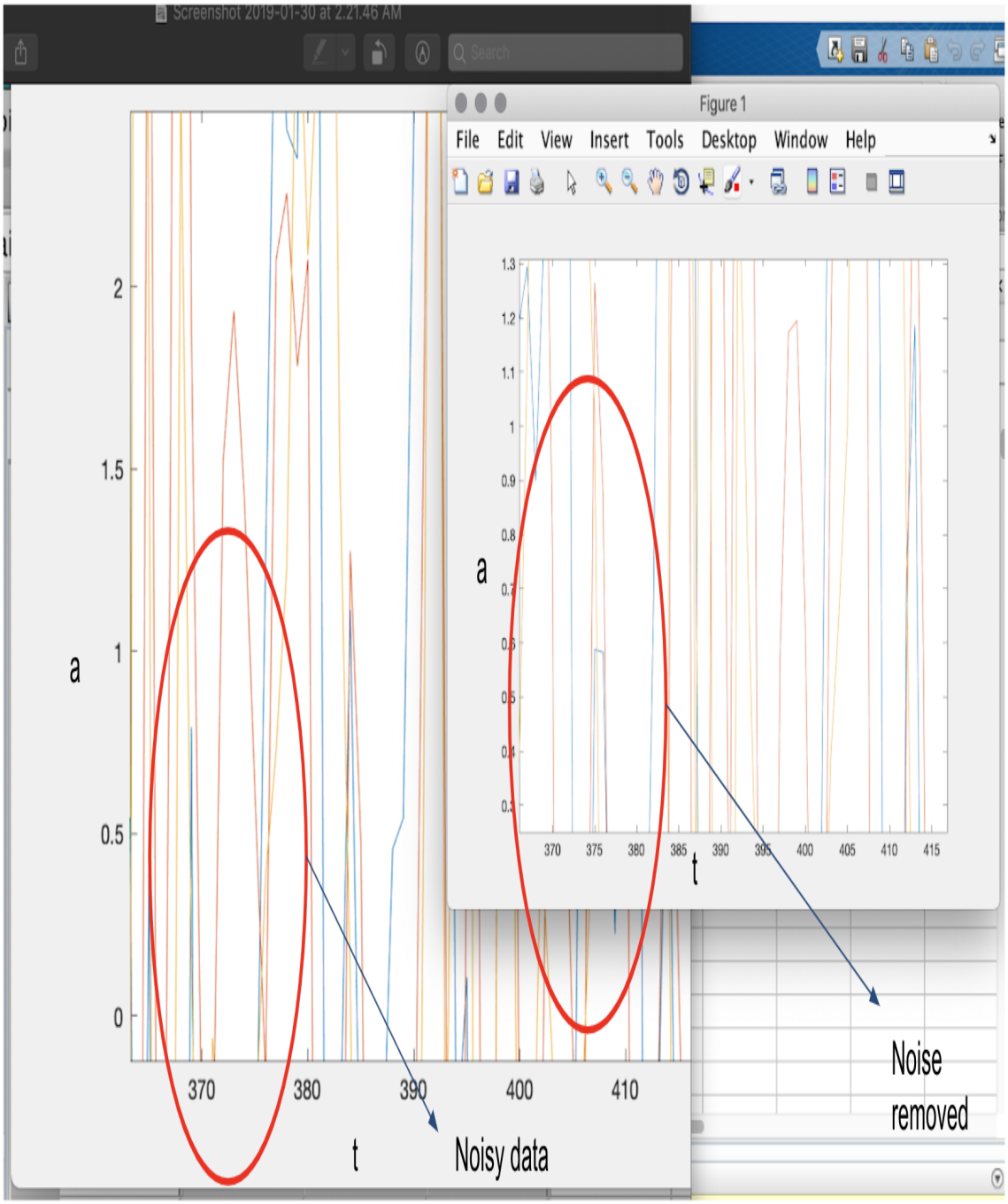


Figure 6.5 Noise Elimination

Figure 6.5 shows Noise elimination done by multi level wavelet decomposition.

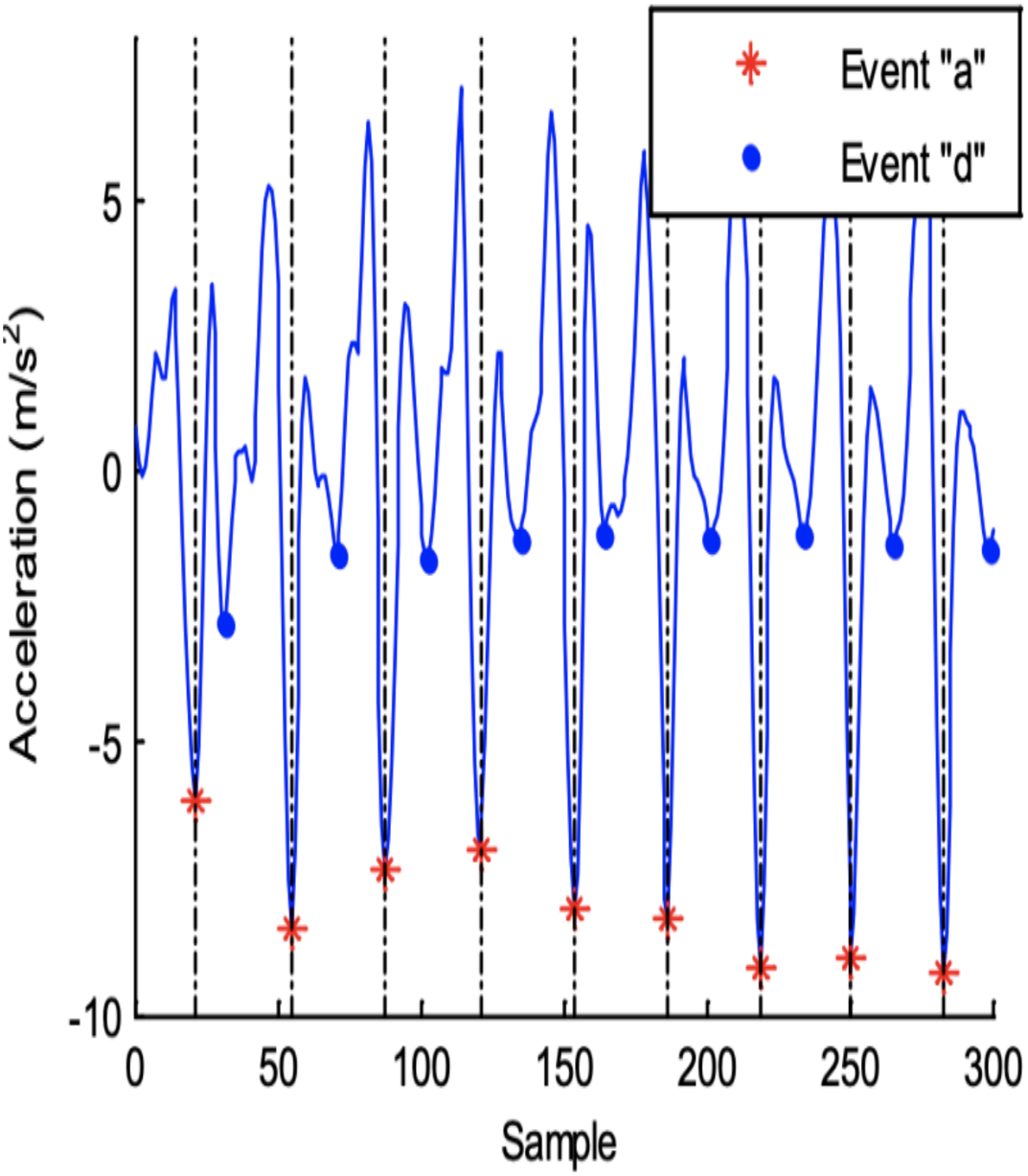
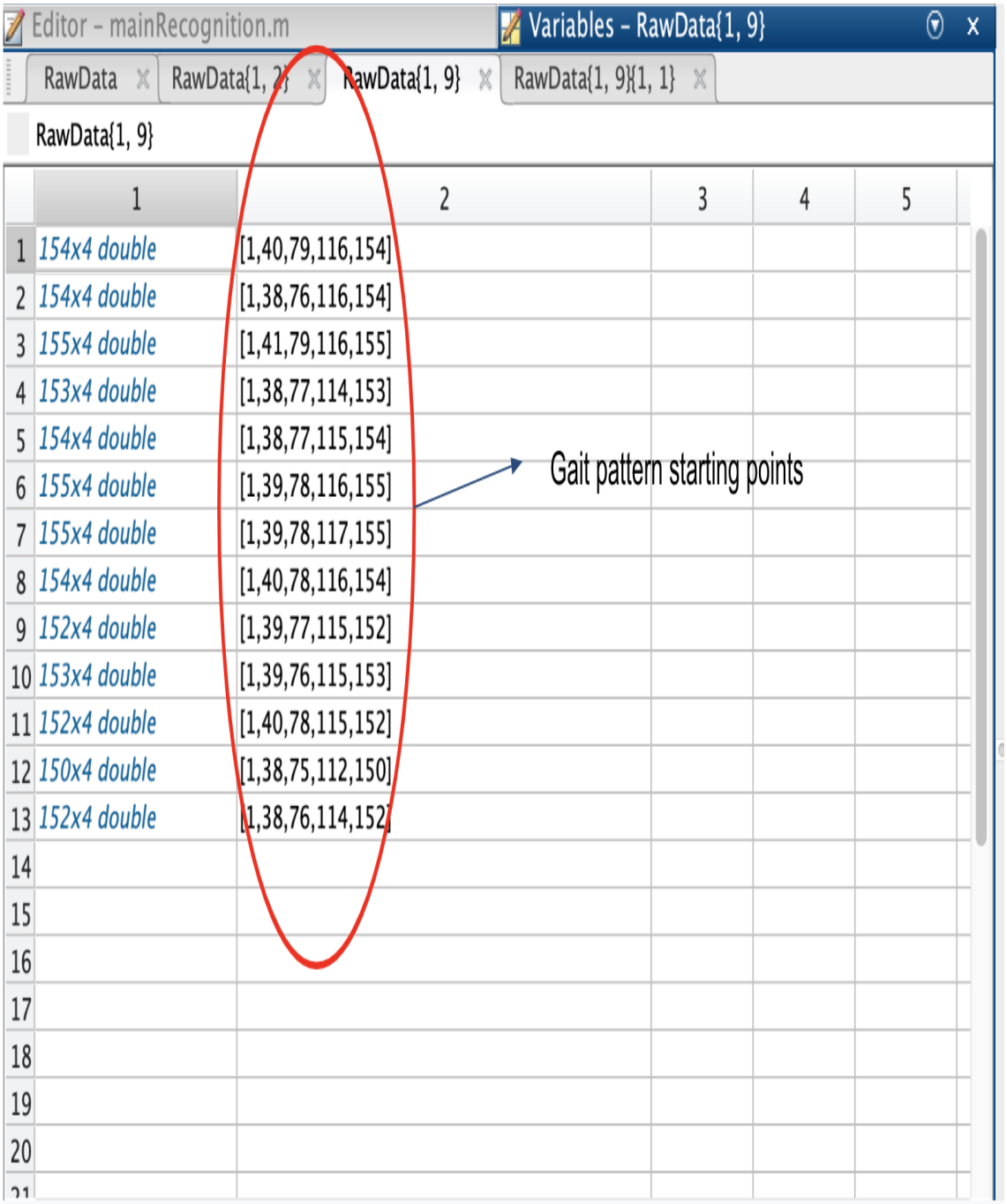


Figure 6.6 Segmentation Of Gait Cycle

Figure 6.6 shows Segmented Gait Cycles each denoting the events ’a’ and events ’d’.



	1	2	3	4	5
1	154x4 double	[1,40,79,116,154]			
2	154x4 double	[1,38,76,116,154]			
3	155x4 double	[1,41,79,116,155]			
4	153x4 double	[1,38,77,114,153]			
5	154x4 double	[1,38,77,115,154]			
6	155x4 double	[1,39,78,116,155]			
7	155x4 double	[1,39,78,117,155]			
8	154x4 double	[1,40,78,116,154]			
9	152x4 double	[1,39,77,115,152]			
10	153x4 double	[1,39,76,115,153]			
11	152x4 double	[1,40,78,115,152]			
12	150x4 double	[1,38,75,112,150]			
13	152x4 double	[1,38,76,114,152]			
14					
15					
16					
17					
18					
19					
20					

Figure 6.7 Gait Pattern Extraction

Figure 6.7 shows extracted gait patterns that denote combinations of several gait segments.

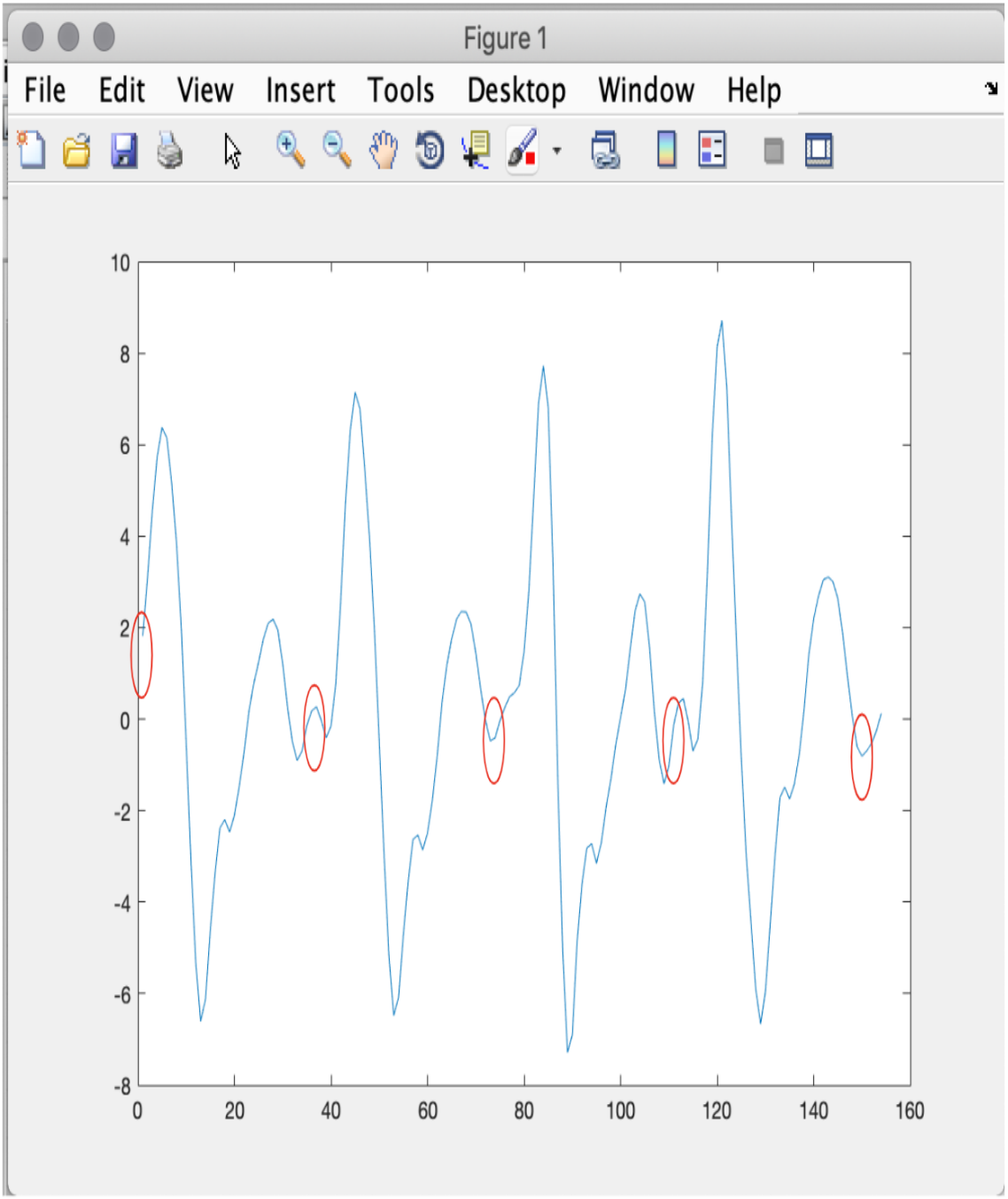


Figure 6.8 Visualization Of Gait Patterns

Figure 6.8 shows visualization of starting point of each segment.

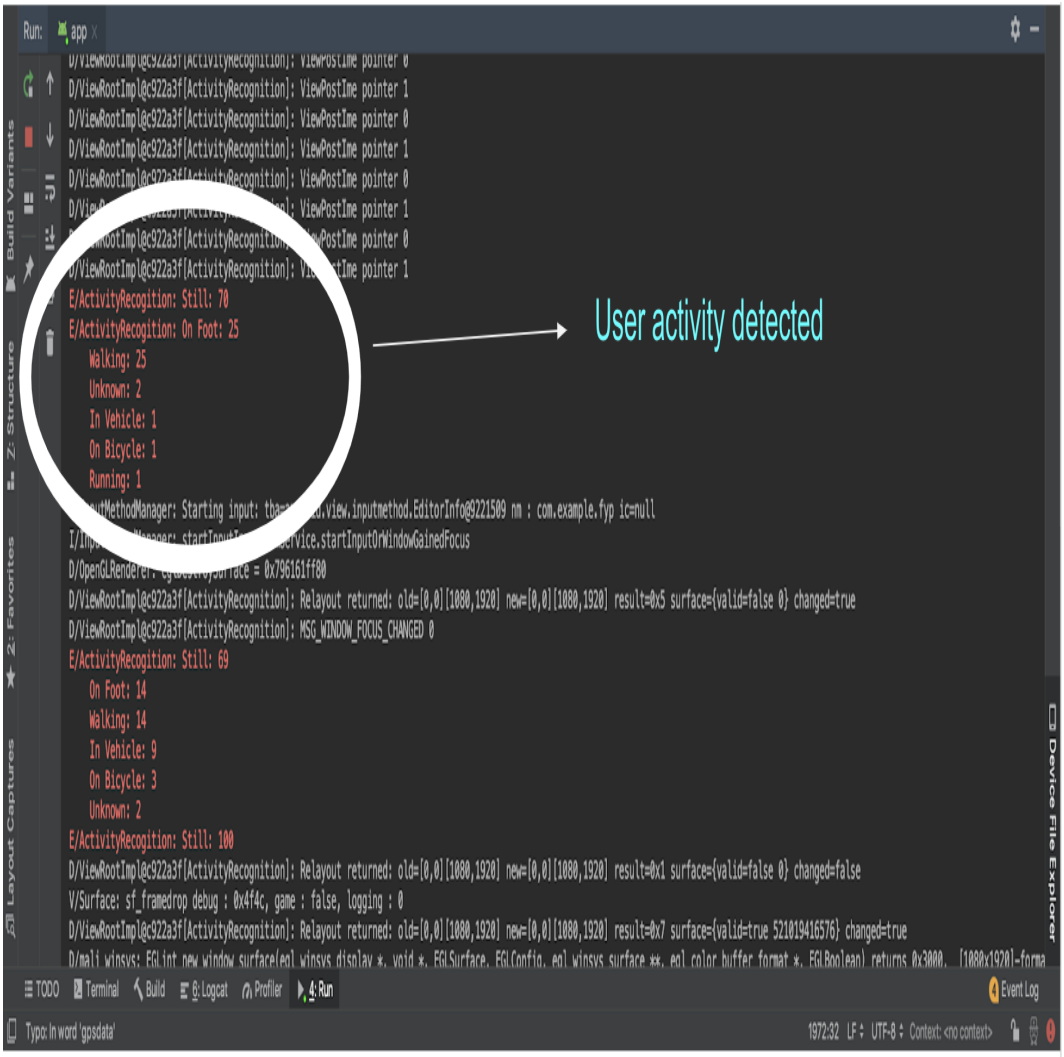


Figure 6.9 User Activity Recognition

Figure 6.9 shows the detection of user activity using Google Play Activity Recognition API.

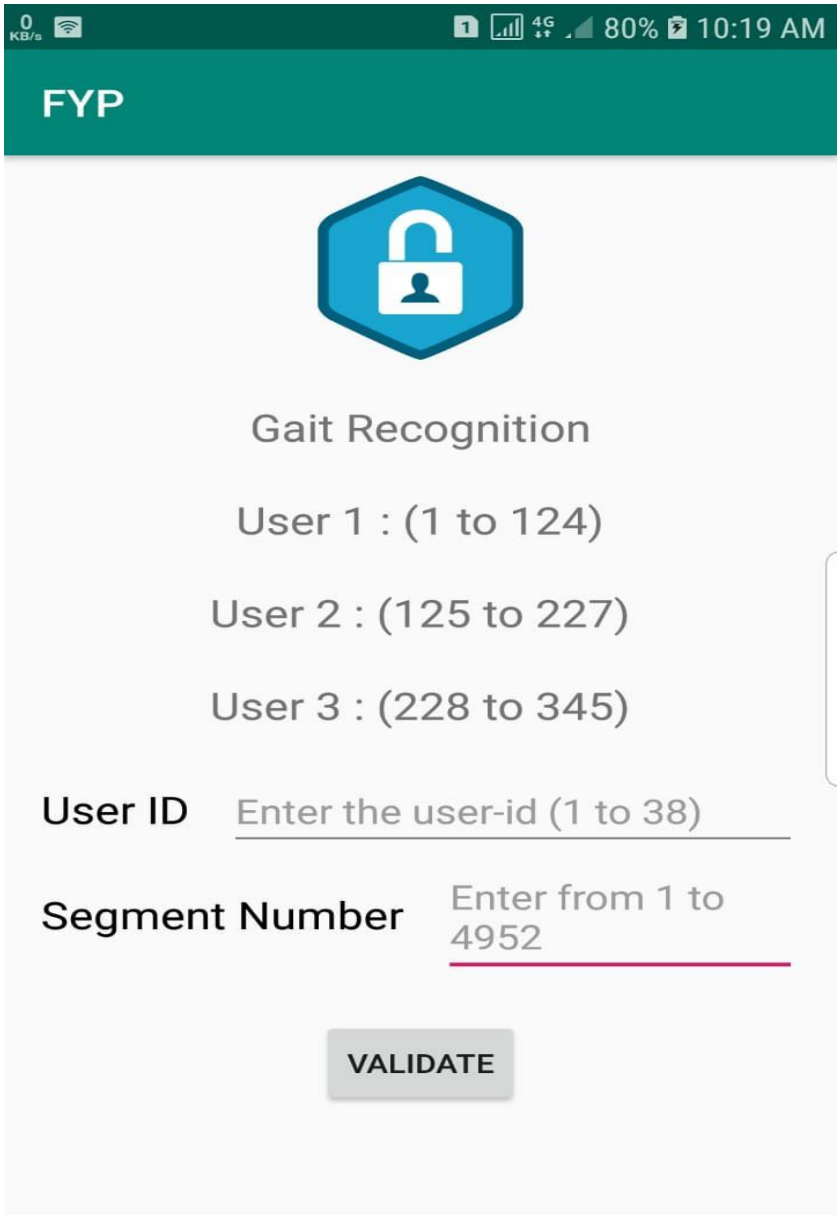


Figure 6.10 User Authentication UI

Figure 6.10 shows the User Interface to test the authentication of a user by supplying mock data.

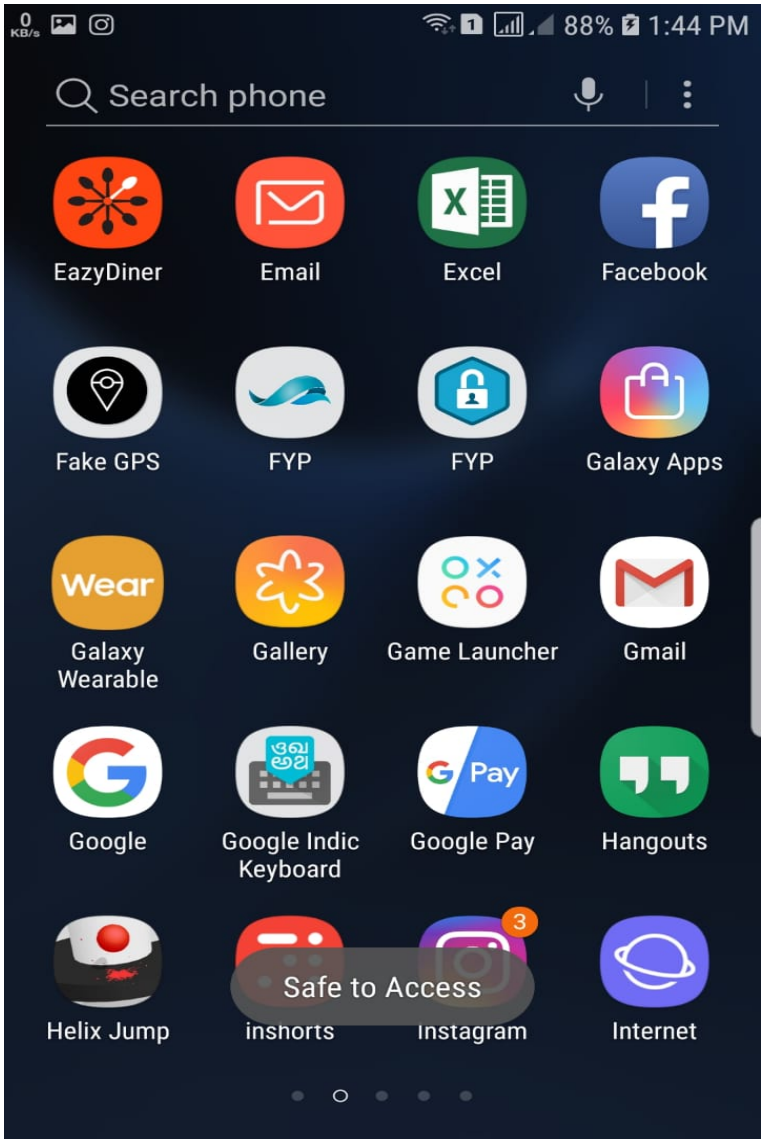


Figure 6.11 Safe Access

Figure 6.11 shows the feedback from our application when the user authentication score falls above 0.85 .

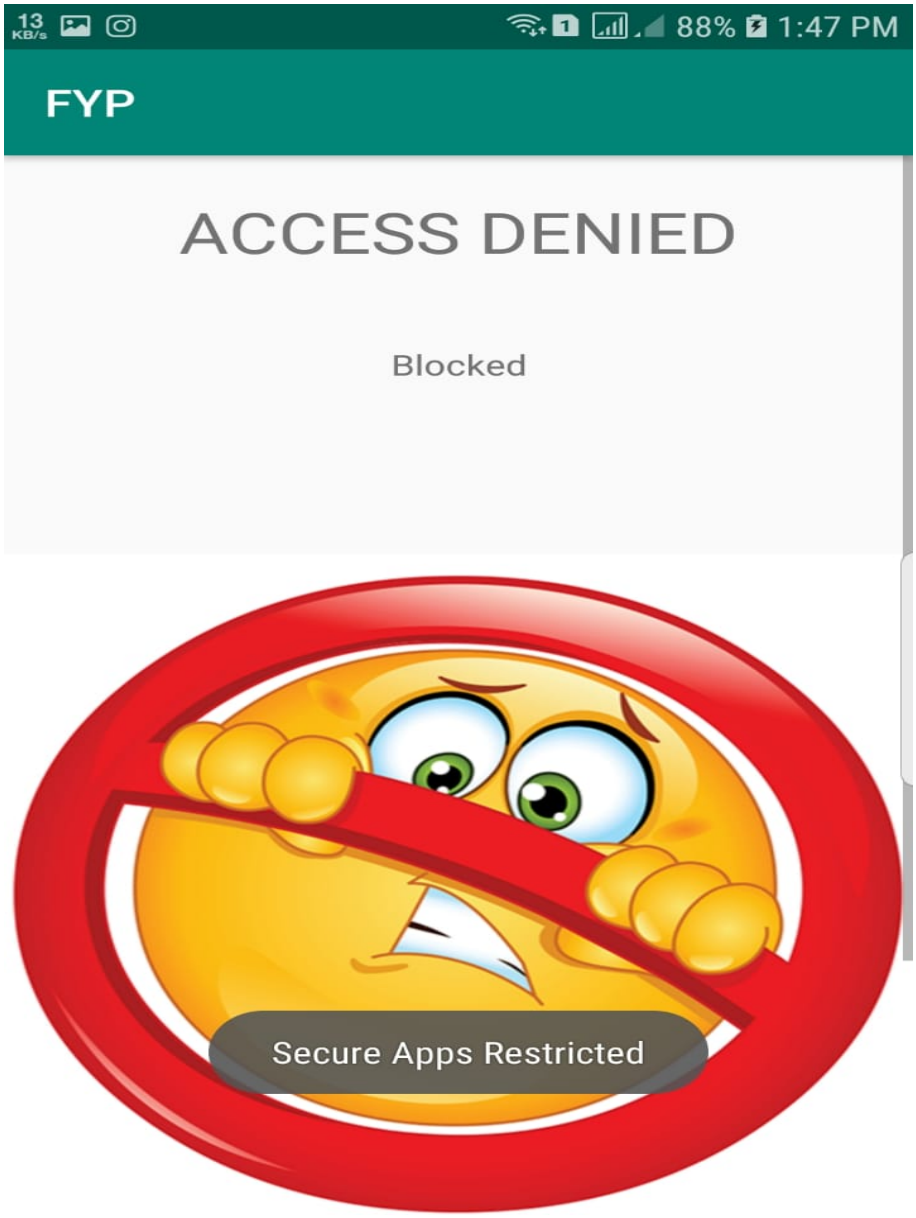


Figure 6.12 Secure Apps Restricted

Figure 6.12 shows the restriction of some highly secure apps when the user score falls within 0.85 and 0.5 .

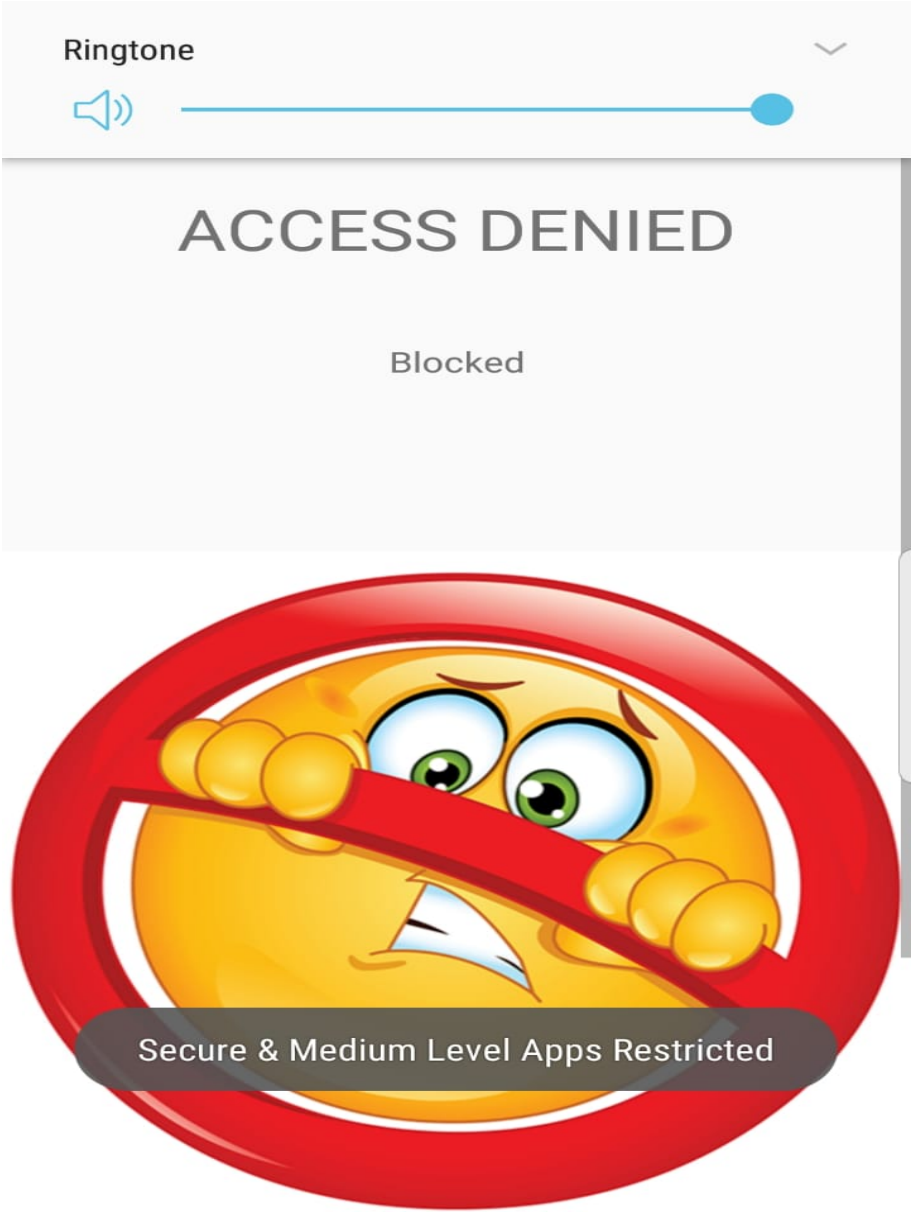


Figure 6.13 Secure and Medium Apps Restricted

Figure 6.13 shows the restriction of some highly secure and medium level apps when the user score falls within 0.5 and 0.1 .

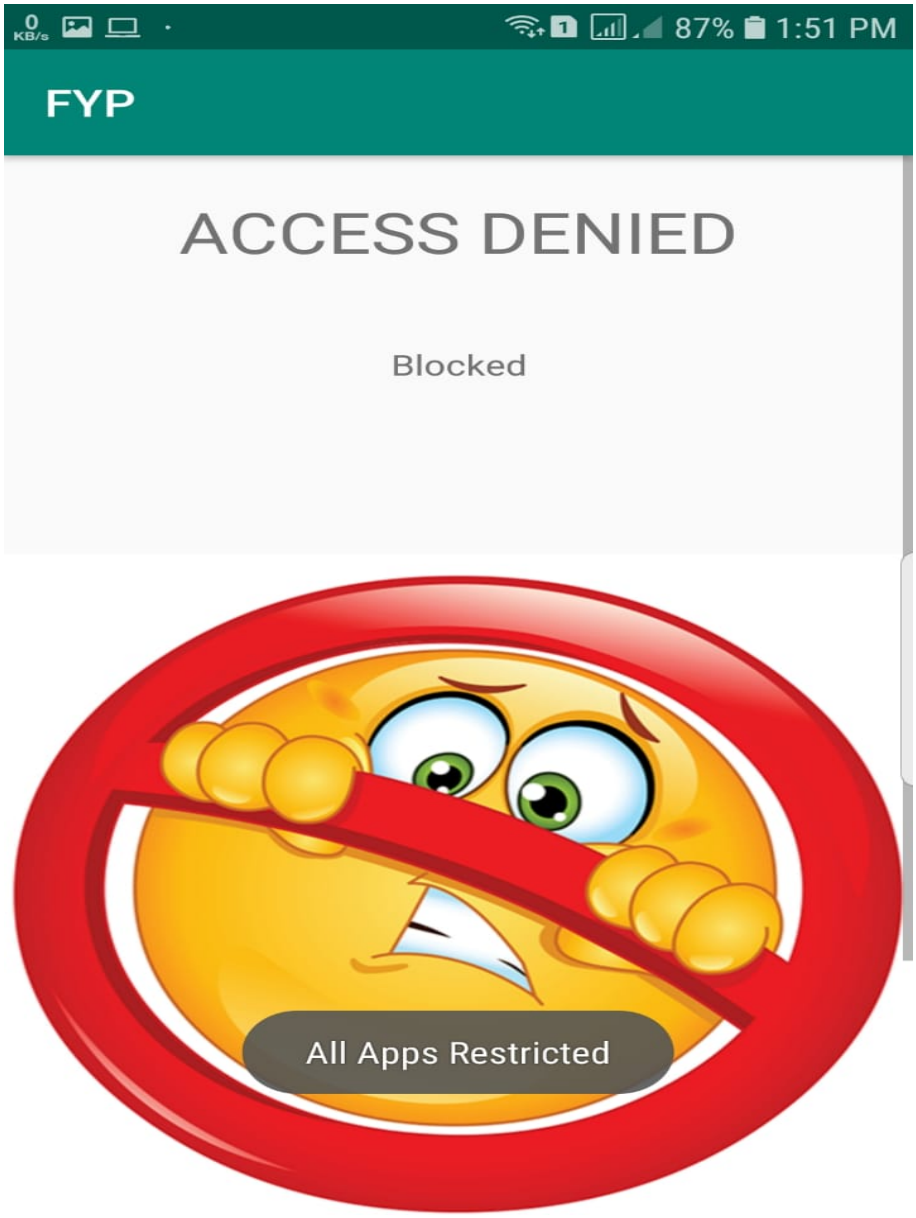


Figure 6.14 All Apps Restricted

Figure 6.14 shows the restriction of some highly secure apps when the user score falls less than 0.1 .

6.2 TESTCASES

The system takes set of applications as input and gives predicted output.

Table 6.1 Testcases

TC-ID	MODULE NO.	INPUT	OUTPUT	Expected Result
01	5	Location Data	Safe Havens will be created.	Yes
02	5	User moves inside safe haven	gps-score=1	Yes
03	5	User moves outside safe haven	gps-score=0	Yes
04	3	Gait pattern by SVM and walking for less than a second (segments)	Predicted class label and gait-score= probability	Yes with accuracy of 95.01
05	3	Gait session by SVM and walking for more than 3 seconds (patterns)	Predicted class label and gait-score= probability	Yes with accuracy of 99.1
06	3	Gait pattern by KNN and walking for less than a second (segments)	Predicted class label and gait-score= probability	Yes with accuracy of 85.4
07	3	Gait pattern by KNN and walking for more than 3 seconds (patterns)	Predicted class label and gait-score= probability	Yes with accuracy of 96.5

Table 6.1 shows the various test cases considered and their expected results.

TC-ID	MODULE NO.	INPUT	OUTPUT	Expected Result
08	6	trust-score= 0.9* gait-score + 0.1* gps-score> 0.85	User can access all apps	Yes
09	6	trust-score= 0.9* gait-score + 0.1* gps-score 0.5< x < 0.85	User cant access highly secure apps	Yes
10	6	trust-score= 0.9 * gait-score+ 0.1 * gps-score 0.1 < x < 0.5	User cant access highly secure apps and medium-level apps	Yes
11	6	trust-score= 0.9 * gait-score+ 0.1 * gps-score< 0.1	User cant access any apps	Yes

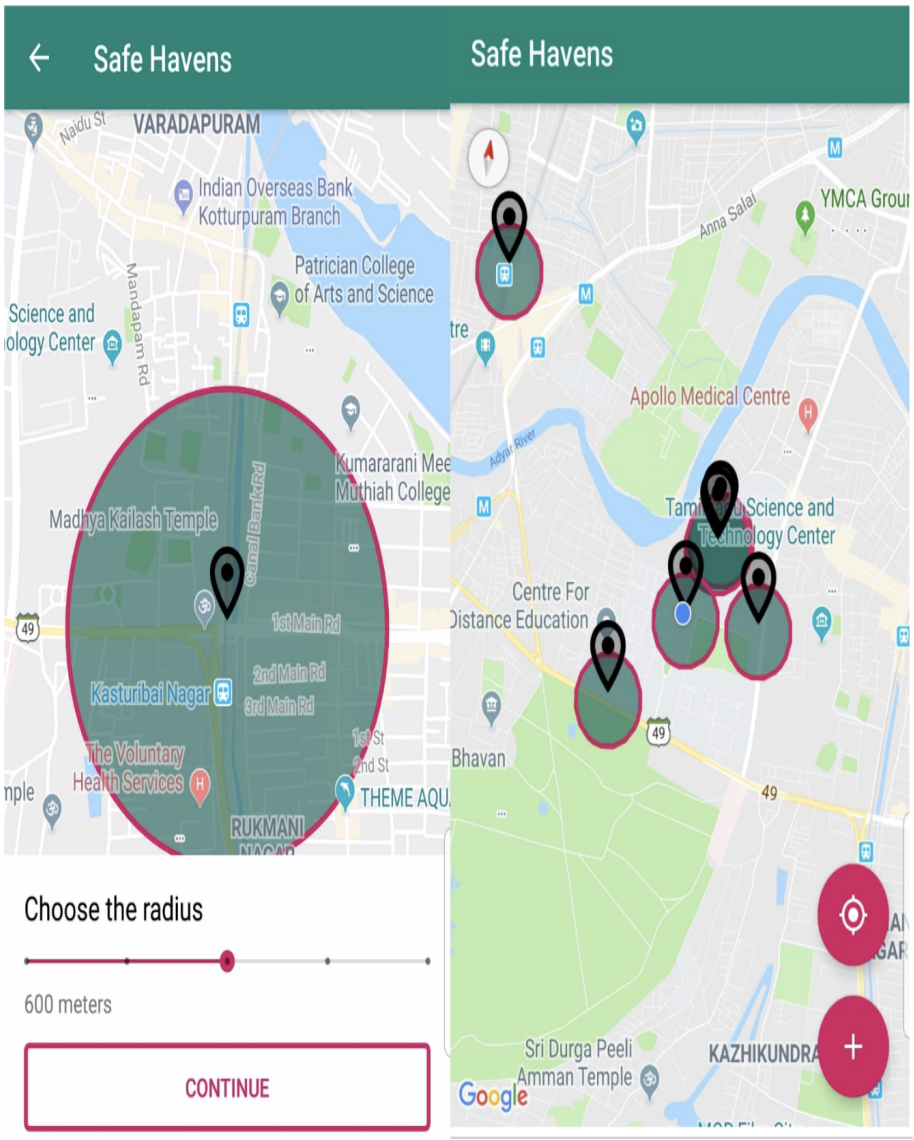


Figure 6.15 Safe Havens Creation

Figure 6.15 shows the creation of safe havens based on the places frequently visited by the user.

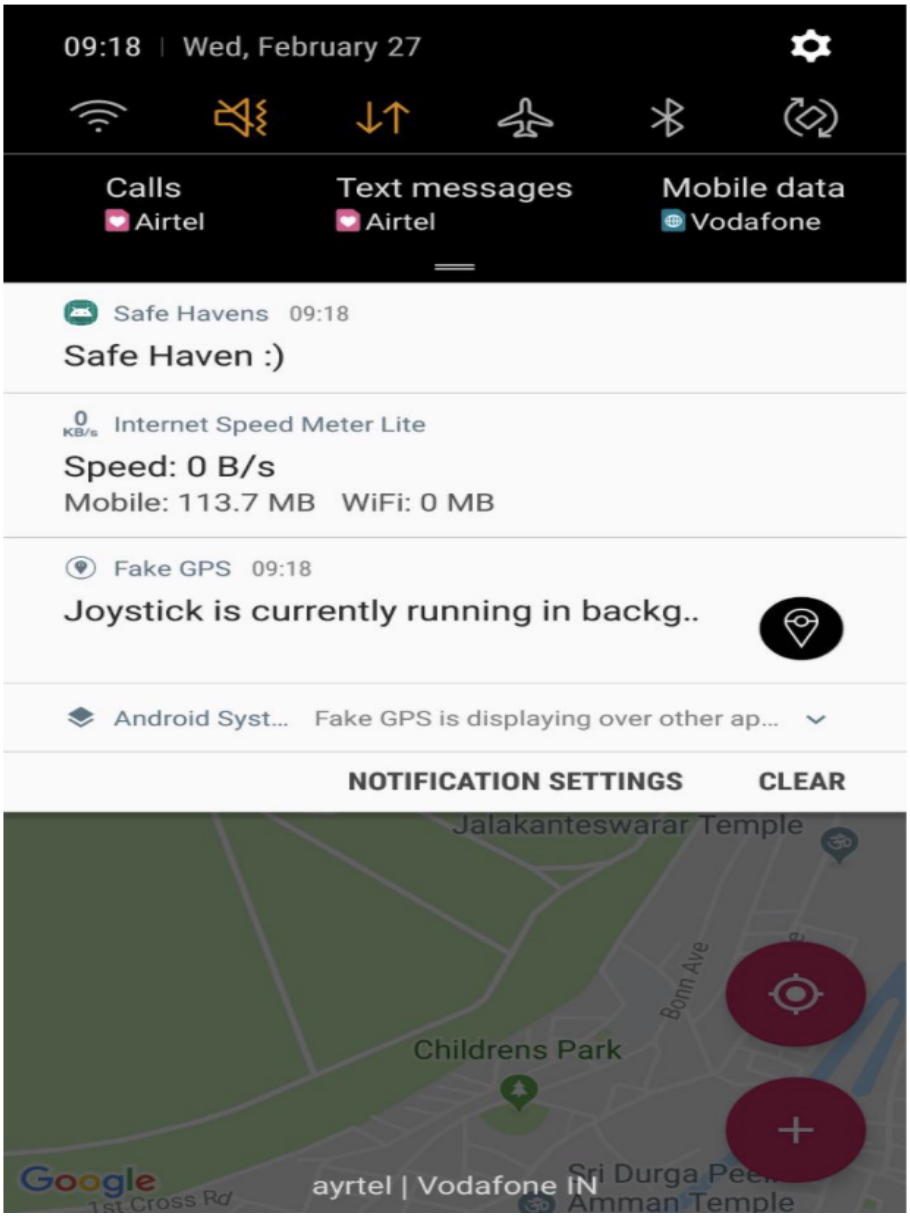


Figure 6.16 Safe Havens Notification

Figure 6.16 shows the notification sent to the user when he enters the safe haven.

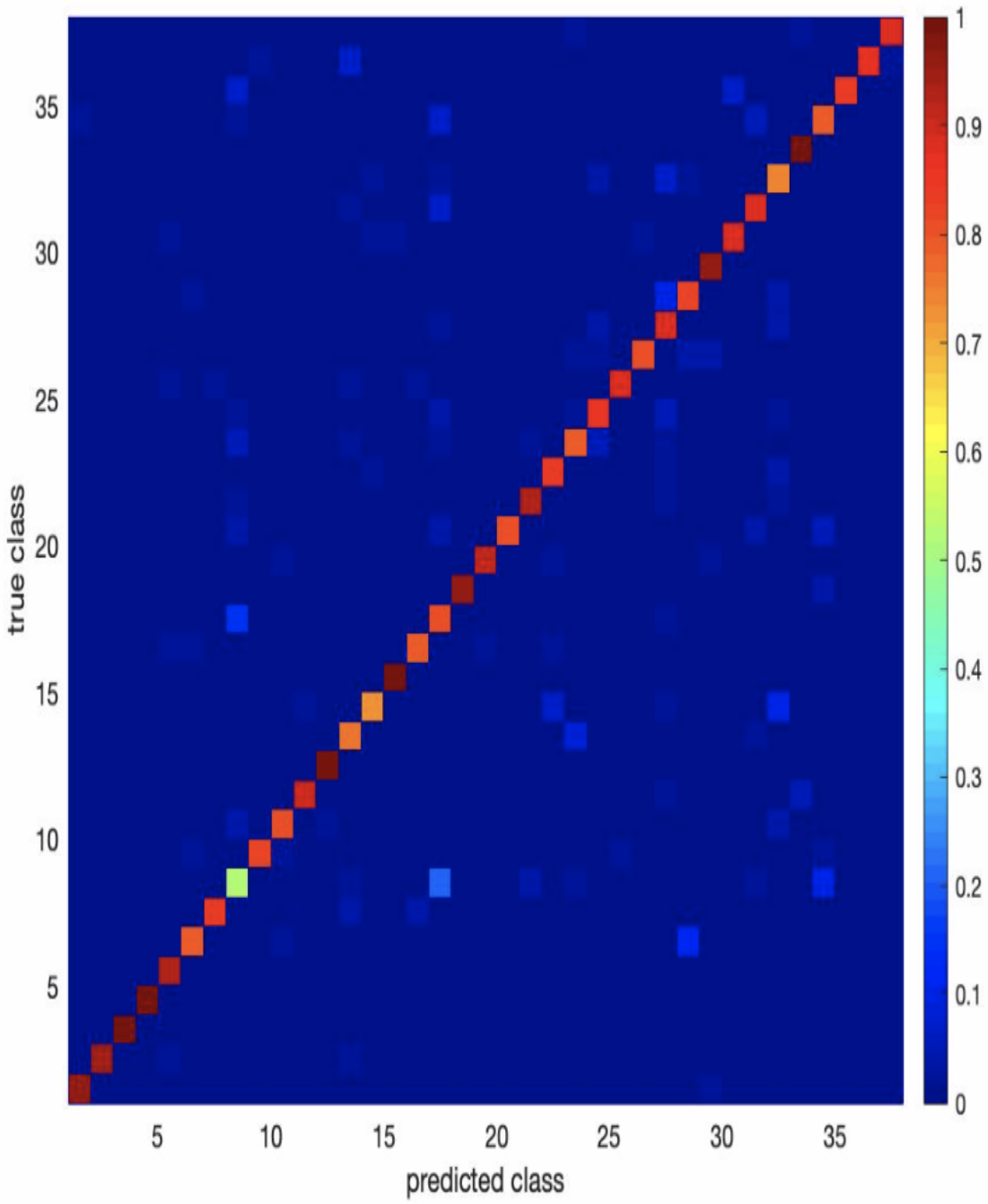


Figure 6.17 Each Segment as a Testing Sample

Figure 6.17 shows the confusion matrix for each segment using SVM.

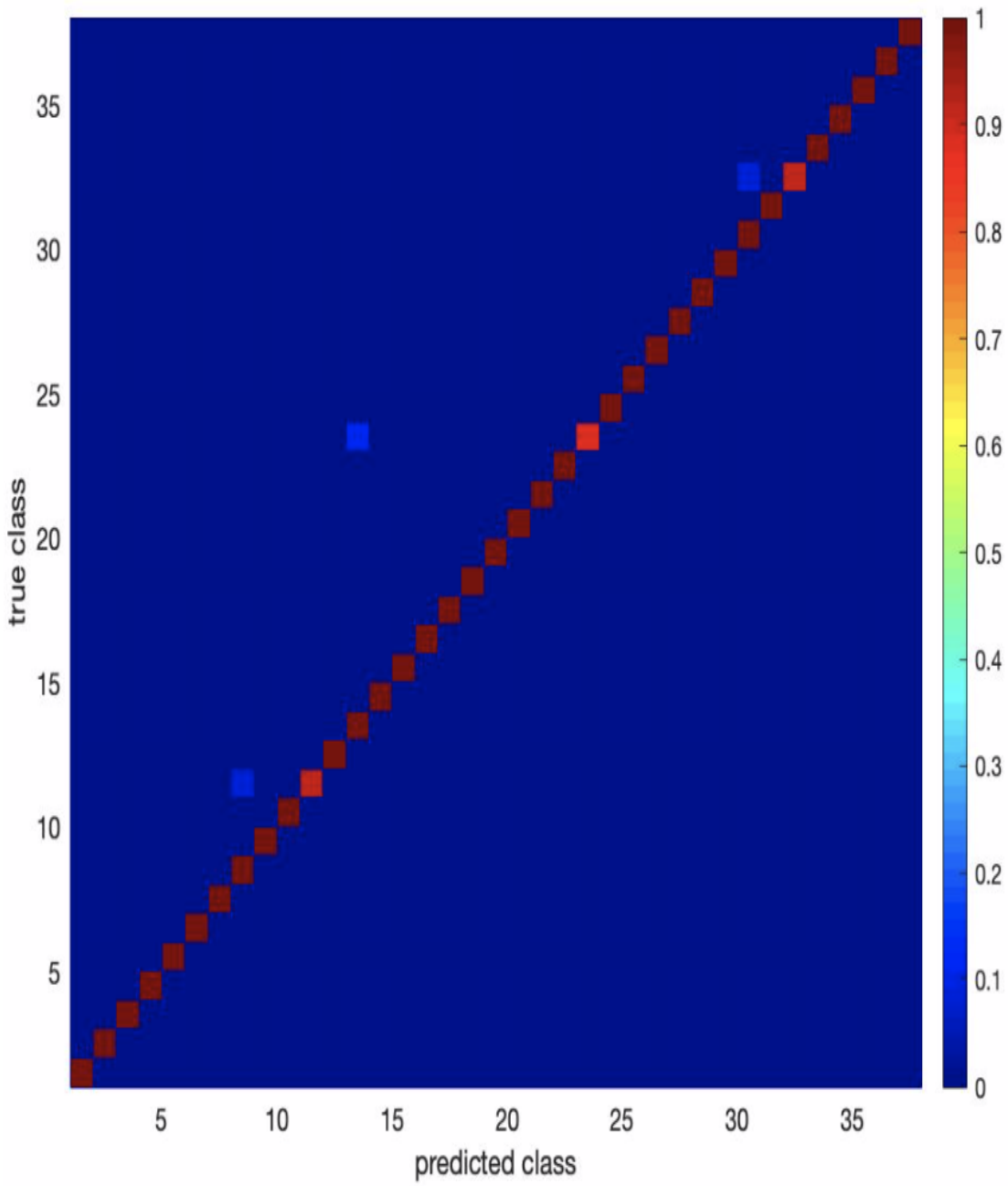


Figure 6.18 Each Session as a Testing Sample

Figure 6.18 shows the confusion matrix for each session using SVM.

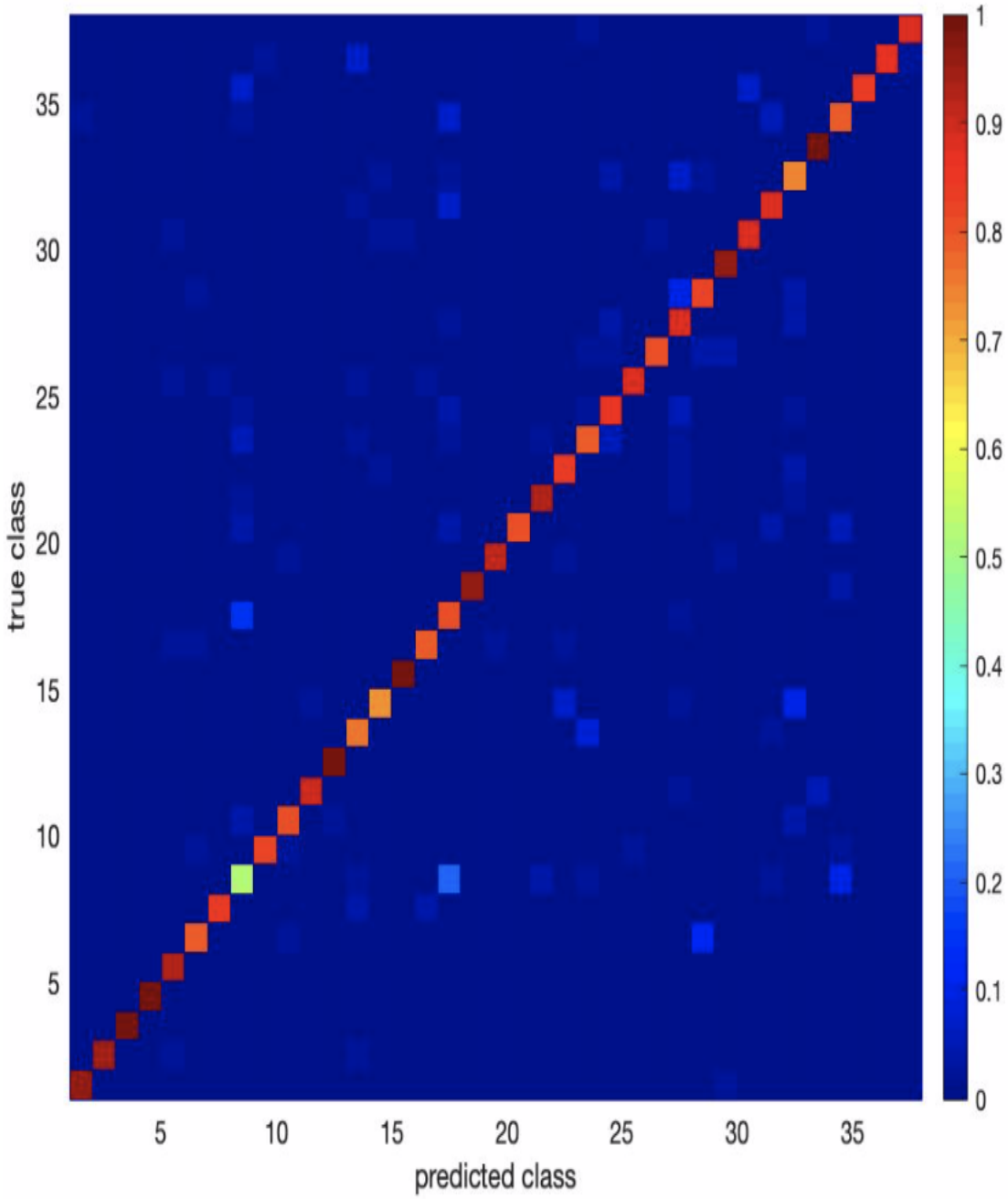


Figure 6.19 Each Segment as a Testing Sample

Figure 6.19 shows the confusion matrix for each segment using K-Nearest Neighbour.

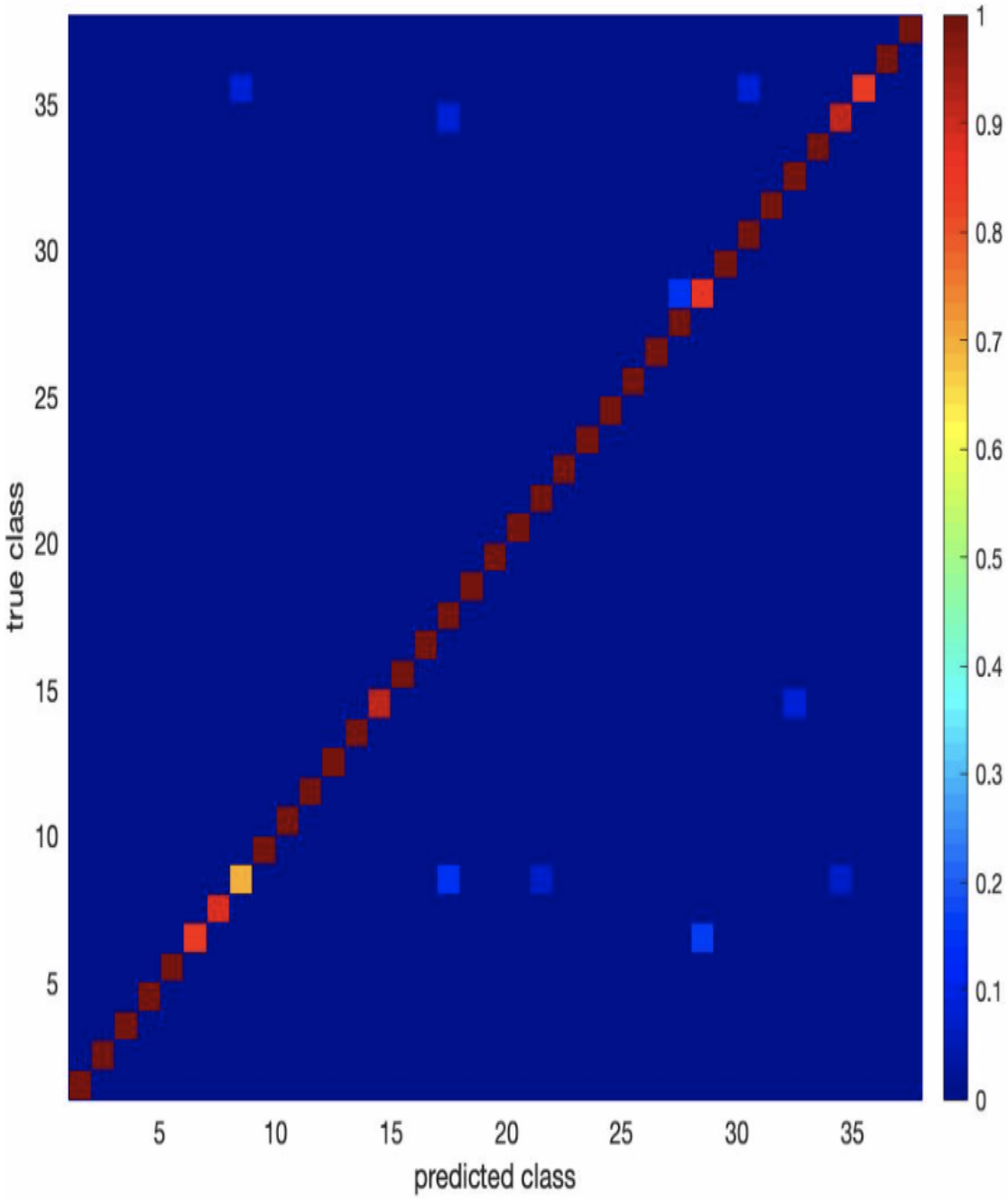


Figure 6.20 Each Session as a Testing Sample

Figure 6.20 shows the confusion matrix for each session using K-Nearest Neighbour.

6.3 EVALUATION PARAMETERS

False Acceptance Rate (FAR) is the measure of the likelihood that the biometric security system will incorrectly accept an access attempt by an unauthorized user.

$$FAR = FP/(FP+TN) \tag{6.1}$$

False Rejection Rate (FRR) is the measure of the likelihood that the biometric security system will incorrectly reject an access attempt by an authorized user.

$$FRR = FN/(TP+FN). \tag{6.2}$$

Equal Error Rate (EER) is a biometric security system algorithm used to predetermine the threshold values for its false acceptance rate and its false rejection rate. When the rates are equal, the common value is referred to as the equal error rate.

Intersection of FAR and FRR in ROC.

SEGMENT VS THEIR PREDICTION PROBABILITY

Figure 6.21 Shows prediction probability rates of each sample segments. Many lower prediction probabilities mean that the system is still novice and can't replace existing authentication systems.

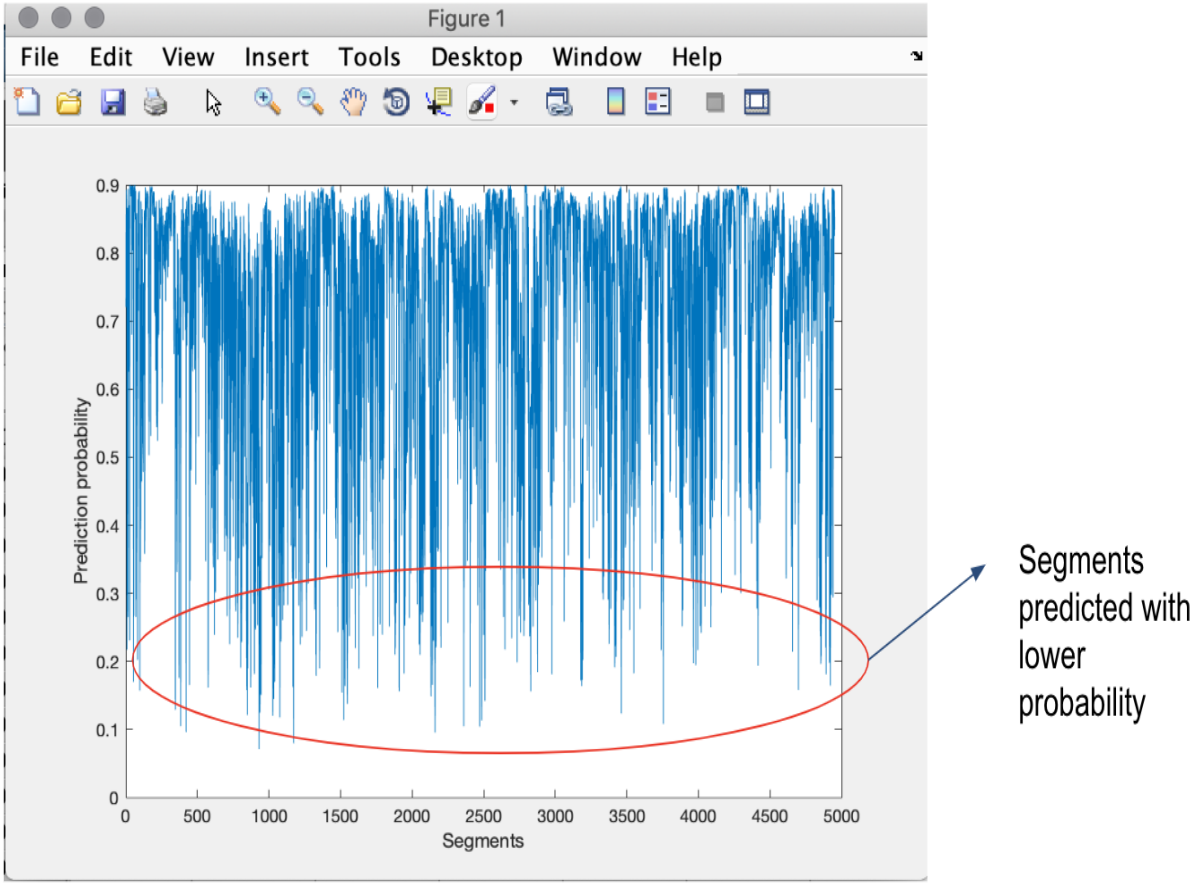


Figure 6.21 Prediction Probability

ROC CURVE-AUTHENTICATION USING SVM

Figure 6.22 Shows equal error rate when each segment is considered as a testing sample.

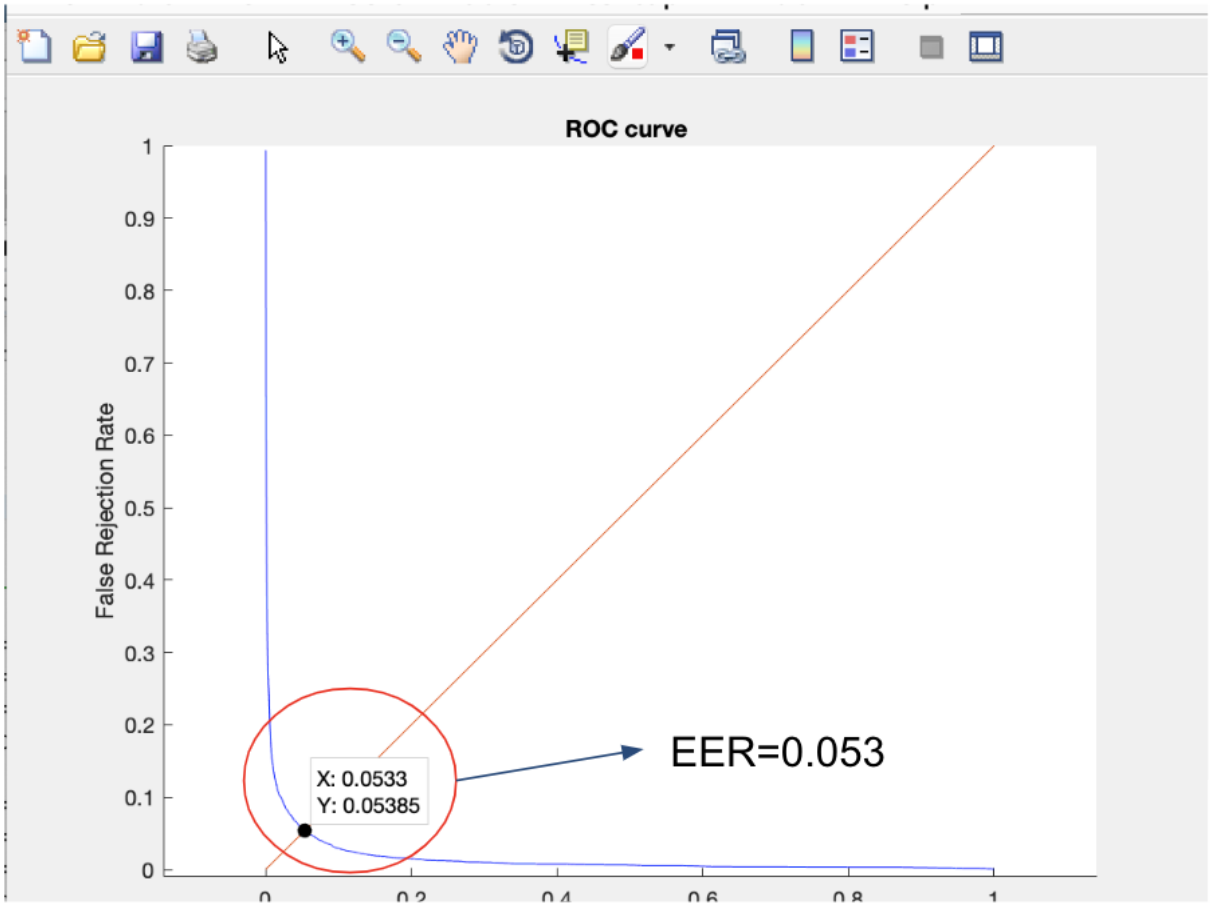


Figure 6.22 Each Segment as a Testing Sample

ROC CURVE-AUTHENTICATION USING SVM

Figure 6.23 Shows equal error rate when each session is considered as a testing sample.

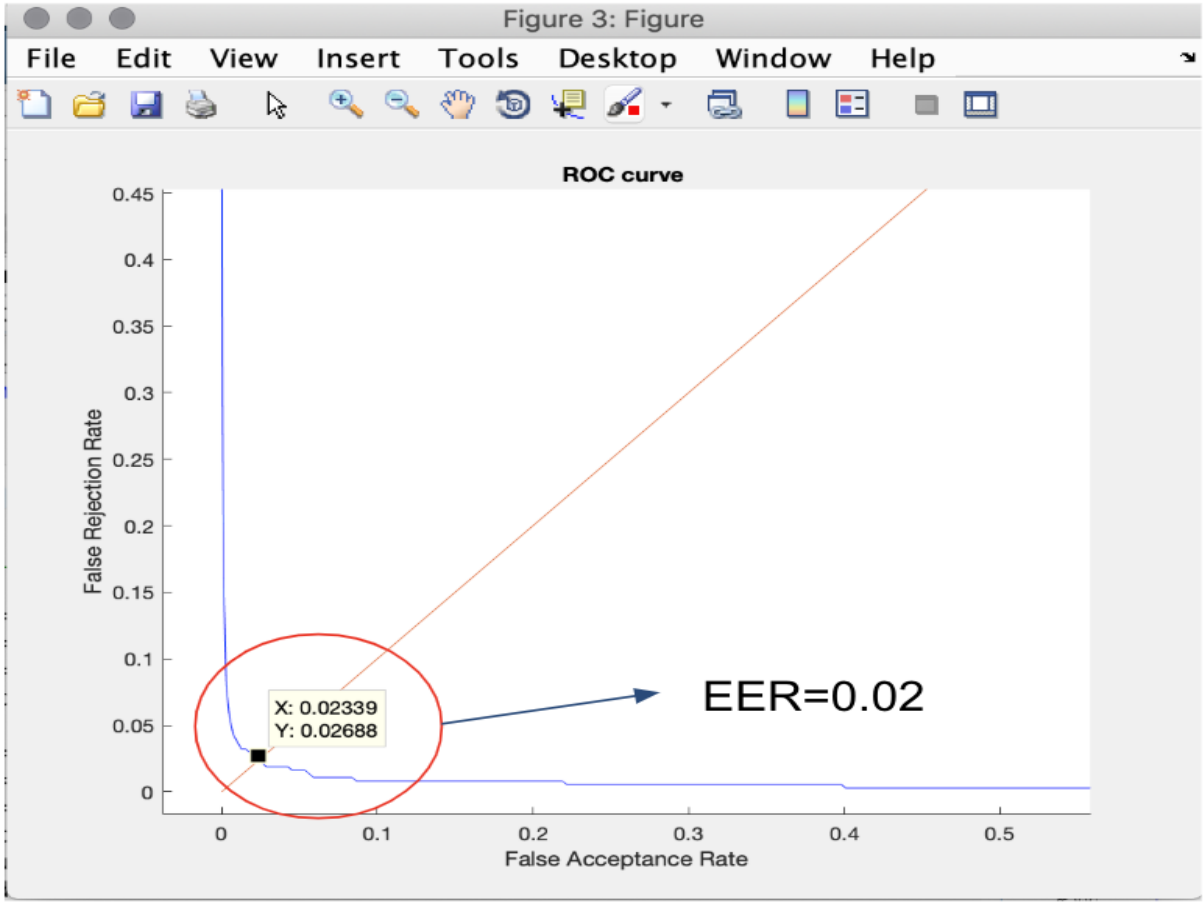


Figure 6.23 Each Session as a Testing Sample

CHAPTER 7

CONCLUSION AND FUTURE WORK

7.1 CONCLUSION

In this project, we addressed the sensor disorientation problem in gait verification or identification systems which can frequently arise in reality, especially in the mobile context. A simple but effective solution taking advantages of available sensors in mobile device was proposed. A gait recognition model leveraging statistical analysis and supervised machine learning in addition with GeoFencing Location Model which could be used to verify or identify mobile user was presented. The results achieved are highly promising, especially with regard to identification. They reflect the good potential of deploying a gait-based and location based authentication to ameliorate the security on portable devices.

Our proposed method does not aim to completely replace the existing explicit authentication schemes on mobiles, since at this moment it is infeasible to achieve a perfect security level (e.g., the zero-FAR is always achieved) of any behavioral biometric-based verification systems. However, the proposed method can be used as an additional authentication scheme, to enhance the usability of the device.

7.2 FUTURE WORK

In the future work, we would like to investigate on developing a unique gait recognition model working effectively regardless of the relative position of the mobile to its owner. A protection scheme with real-time tracking and real-time authentication will be the further step of development. We are planning to expand the analysis to incorporate various supervised learning techniques for profiling. The effectiveness of different model parameters, and of different sensors (perhaps in different situations or for different users), that might lead to a weighted comfort computation will be analysed.

REFERENCES

[1] Paul ONeill Kishore Reddy Devu Manikantan Shila, Kunal Srivastava, *A multi-faceted approach to user authentication for mobile devices - using human movement, usage, and location patterns*, In the Proceedings of IEEE Symposium on Technologies for Homeland Security, pp. 1-6, 2016.

[2] Upal Mahbub and Rama Chellappa, *PATH: Person Authentication using Trace Histories*, In IEEE Transactions on Information Systems and Security, pp. 1-8, 2016.

[3] Oriana Riva, Chuan Qin, Karin Strauss, and Dimitrios Lymberopoulos, *Progressive Authentication: Deciding When to Authenticate on Mobile Phones*, USENIX, Bellevue, WA, 2012.

[4] D. M. Shila and K. Srivastava, *CASTRA: Seamless and Unobtrusive Authentication of Users to Diverse Mobile Services*, In IEEE Internet of Things Journal, vol. 5, no. 5, pp. 4042-4057, 2018.

[5] Deokjai Choi Thang Hoang and Thuc Nguyen3, *On the Instability of Sensor Orientation in Gait Verification on Mobile Phone*, In the Proceedings of IEEE 12th International conference on e-business and telecommunications, pp. 148-159, 2016.