The Influence Of Emotion On Keystroke Patterns In Android Platform: An Experimental Study Using Auditory Stimuli

A project report submitted in fulfillment of
the requirements for the degree of

Bachelor of Engineering

by

Lillita Rhea D'souza (Roll No. 7126)
Chaitrali Kiran Gandhi (Roll No. 7133)
Naman Pravin Malik (Roll No. 7146)

Under the guidance of
Prof. Prajakta Mandar Dhamanskar

DEPARTMENT OF INFORMATION TECHNOLOGY
Fr. Conceicao Rodrigues College of Engineering, Bandra (W), Mumbai - 400050
University of Mumbai

April 24, 2017

This work is dedicated to my family.

I am very thankful for their motivation and support.

Internal Approval Sheet

CERTIFICATE

This is to certify that the project entitled **"The Influence Of Emotion On Keystroke Patterns In Android Platform: An Experimental Study Using Auditory Stimuli"** is a bonafide work of **Lillita Rhea D'souza(7126), Chaitrali Kiran Gandhi(7133), Naman Pravin Malik(7146)** submitted to the University of Mumbai in fulfillment of the requirement for the award of the degree of **Bachelor of Engineering** in **Information Technology**

(Name and sign)
Supervisor/Guide

(Name and sign)
Head of Department

(Name and sign)
Principal

Approval Sheet

Project Report Approval

This project report entitled by **The Influence Of Emotion On Keystroke Patterns In Android Platform: An Experimental Study Using Auditory Stimuli** by **Lillita Rhea D'souza, Chaitrali Kiran Gandhi, Naman Pravin Malik** is approved for the degree of Bachelor of Engineering

Examiners

1.________________________

2.________________________

Date:

Place:

Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Lillita Rhea D'souza (Roll No. 7126) **(sign)**________________

Chaitrali Kiran Gandhi (Roll No. 7133) **(sign)**________________

Naman Pravin Malik (Roll No. 7146) **(sign)**________________

Date: April 24, 2017

Abstract

Keystroke Dynamics is the study of a user's typing pattern based on the various timing information obtained when a key is pressed and released. It comes under Behavioural Biometrics and has been a topic of interest for authenticating as well as identifying users based on their typing pattern. There have been numerous studies conducted on Keystroke Dynamics as a Biometrics with different data acquisition methods, user base, feature sets, classification techniques and evaluation strategies. Mobile handsets play a significant role in the modern society, providing accessibility to personal and confidential data applications anywhere, anytime. With so much importance of these devices, the focus is now on mobile security application. Majority of these applications are one factor authorization, such as verifying password and pins. Keystroke dynamics can be implemented on mobile phone which will improve the security. This type of security checks what you type and how you type. Implementation of Keystroke Dynamics on mobile phones is divided into two important phases. In the first phase, data from the user's samples is collected and stored in database. The second phase of the project is described as implementation of the algorithm and authentication of the users on the basis of data collected from the samples. Emotion is a cognitive process and is one of the important characteristics of human beings that makes them different from machines.

Traditionally, interactions between humans and machines like computers do not exhibit any emotional exchanges. If we could build any system that is intelligent enough to interact with humans that involves emotions, that is, it can detect the user emotion's and change its behaviour accordingly, then using machines could be more effective and friendly. Many approaches have been taken to detect user emotions. In recent years, a novel approach for emotion recognition has been reported, which is by keystroke dynamics. The advantages of using this approach are that the data used is rather non-intrusive and easy to obtain. However, there were only limited investigations about the phenomenon itself in previous studies. Hence, this project aims to examine the source of variance in keyboard typing patterns caused by emotions. This is a controlled experiment to collect

subjects' keystroke data in different emotional states induced by International Affective Digitized Sounds (IADS). Two-way Valence (3) x Arousal (3) ANOVAs will be used to examine the collected dataset. Our aim in this project is to prove that keystroke duration and latency are influenced by valence and arousal.

Acknowledgments

We have great pleasure in presenting the report on **"The Influence Of Emotion On Keystroke Patterns In Android platform: An Experimental Study Using Auditory Stimuli"**. We take this opportunity to express my sincere thanks towards the guide Prof. Prajakta Mandar Dhamanskar, C.R.C.E, Bandra (W), Mumbai, for providing the technical guidelines, and the suggestions regarding the line of this work. We enjoyed discussing the work progress with her during our visits to department.

We thank Prof. Jagruti Ketan Save, Head of Information Technology Dept., Principal and the management of C.R.C.E., Mumbai for encouragement and providing necessary infrastructure for pursuing the project.

We also thank all non-teaching staff for their valuable support, to complete our project.

Lillita Rhea D'souza(Roll No. 7126)
Chaitrali Kiran Gandhi(Roll No. 7133)
Naman Pravin Malik(Roll No. 7146)

Date: April 24, 2017

Contents

List of Figures

List of Tables

Glossary

ANOVA Analysis of Variance. 10

CSEA Centre for Emotion and Attention. 11

CSV Comma Seperated Values. 14

EER Equal Error Rate. 6

FAR False Acceptance Rate. 3

FRR False Rejection Rate. 3

IADS International Affective Digitized Sounds. 7

MLP Multilayer Perception. 6

PIN Personal Identification Number. 1

SAM Self-Assessment Manikin. 9

SPSS Statistical Package for the Social Sciences. 22

WEKA Waikato Environment for Knowledge Analysis. 6

Chapter 1

Introduction

The most popular access security in mobiles is either a password or personal identification number (PIN), a secret-knowledge approach that relies heavily on the user to ensure continued validity. Due to their limited length, PINs are vulnerable to surfing as well as trial-and-error attacks. Among the available techniques, biometric-based authentication is the only one that seems plausible since tokens must also be carried in token based authentication, along with the device. Biometric characteristics can be divided in two main classes:

- Physiological
  - Face
  - Fingerprint
  - Hand
  - Iris
- Behavioural
  - Keystroke
  - Signature
  - Voice

Our project deals with Keystroke Dynamics.
Keystroke dynamics or typing dynamics refers to the automated method of identifying or confirming the identity of an individual based on the manner and the rhythm of typing on a keyboard. Keystroke Dynamics is a behavioural biometric technique that uses the rhythm and manner in which an individual types characters on a keyboard. The keystroke

rhythms of a user generate a unique biometric template of the individual's typing pattern for authentication. The key measurements used to determine the unique typing rhythm of a user are as follows :-

- Keystroke Duration- The time period between key press and key release.
- Keystroke Latency- The time period between key release and next key pressed.

The recorded data is then processed through a unique algorithm which determines the user pattern for future comparison. Keystroke dynamics is mostly used for identification purposes. Every individual has a different typing style. The project will include collection of a data set using android devices with touch screen. The time based features will be studied and the identification measurements will be performed.

- **Motivation**

  Keystroke Dynamics is one of the newest and emerging field of authentication. Keystroke dynamics refers to the automated method of identifying or confirming the identity of an individual based on the manner and the rhythm of typing on a keyboard. It uses the rhythm and manner in which an individual types characters on a keyboard. The keystroke rhythms of a user generate a unique biometric template of the individual's typing pattern for authentication. The motivation for taking up this project was that keystroke dynamics is one of the latest technologies and has a wide scope of research and novelty.

- **Objective**

  Although there are several studies on PC systems, there are a few related works performed on smart phones. Obviously, the typing patterns on touchscreens will not be similar to a normal keyboard. Due to this Android is chosen as the field of study for this research. Even keystroke has proven to be effective in authentication of an individual. There is one prominent disadvantage. Keystroke patterns vary according to one's mood. If this hypothesis is true, then using keystroke dynamics as a mean of authentication will certainly not prove to be effective and will be invalid. Our study aims to find out whether this hypothesis is true. If yes, then the variance of the time duration for each emotion will be calculated.

Chapter 2

Literature Review

In general behavioral biometrics such as keystroke dynamics are less reliable than physiological biometrics. The following 7 criteria evaluate the suitability of keystroke dynamics:

Universality	This biometric solution can be used by all individuals that are able to use a keyboard.
Uniqueness	Unlike physiological biometric factors, there can be no such thing as an absolute match with behavioral biometrics. Therefore it is difficult to discuss uniqueness of a typing pattern. It must be clear that with keystroke dynamics it is not possible to have FAR and FRR as low as for the better physiological biometric factors, therefore it cannot be the sole factor to identify or authenticate a subject.
Permanence	A major problem with keystroke dynamics is that a subject's typing rhythm varies considerably in between days and even within the same day. There are numerous reasons for this: tiredness, switching computers / keyboards, mood, influence of alcohol and medications, etc.

Collectabilty	An important advantage of keystroke dynamics is that there is no special hardware needed as with other biometrics, a standard computer keyboard is sufficient. It is also possible to capture the keyboard dynamics in the background, during longer periods without causing any additional overhead for the subject. This might allow to trigger an alarm when another subject takes over the session on a logged in workstation.
Acceptability	Depending on the country or state you are in using key logging software might be a direct violation of local laws. Even if the actual typed text is not analyzed or retained, applicable legislation is sufficiently unclear to be in your disadvantage when you intend to actually use keystroke dynamics. Request legal advice before implementing or experimenting without written consent from people on the keyboard.
Circumvention	It is certainly difficult, if not impossible to mimic another person's typing rhythm. Electronically capturing using keylogging software is possible, thus implementing this biometric solution requires that data security is guaranteed from the input (keyboard) to the matching algorithm.
Performance	Behavioral biometrics have higher variations because they depend on a lot of (external) factors such as ergonomics, fatigue, mood, etc. This causes higher FAR and FRR when compared to solutions based on a physiological biometric factor such as fingerprint recognition.

Table 2.1: Table depicting criteria for evaluation of suitability of Keystroke Dynamics

The following literature survey explores these criteria in further detail:-

The work in [1] describes the concept based on using standard input devices, such as keyboard and mouse, as sources of data recognition of user's emotional states applications do not adapt to user's context. User context includes information such as their location, emotional states, or situation. This system using key strokes is more intuitive, unobtrusive and has a wider range of users. In this work, the emotional states are investigated via keystroke dynamics. The proposed method is based on to calculate the pressing time, dwell time, mean time, range and standard deviation time of keystrokes. Through the study, they have proposed a keystroke dynamics based application for recognizing emotional states of computer user. This method is inexpensive and non-intrusive to user. For keystroke dynamics method, timing features of fixed texts has been analyzed because fixed text showed better results than free text.

The authors of [2] have found about the research done in static authentication of keystroke dynamics and the future scope as well as the disadvantages of keystroke dynamics. They have listed out the various researches done on keystroke dynamics with the help of statistical and neural network classification techniques. It has been stated that it is impossible to compare both the techniques on the same standards. However, statistical approach is more popularly used pertaining to the fact that it is relatively simple, has less overhead and is easy to implement. On the other hand, neural network gives more accurate results. Through the study, it is found that keystroke dynamics has various advantages making it the best and cheapest authentication technique. But it has a major problem of being inconsistent. Typing pattern of a user can become erratic due to many reasons like being tired or injured or under medication. Emotions play an important role in keystroke dynamics and it changes the typing pattern when a person is flustered or angry. Typing pattern of an individual also changes gradually over time.

The authors of [3] say that currently people store a lot of sensitive data on their mobile devices. Moreover, touch screen allows adding features ranging from pressure of the screen or finger area to the classical time-based features used for keystroke dynamics. In this paper, they examine the effect of these additional touch screen features to the identification and verification performance by using time based features to find out the results. An Android application having its own software keyboard was developed for data collection where users had to introduce some personal data, such as gender, birth date and their experience level regarding smart phone usage in the registration phase. The majority of participants completed 2 sessions in a period of two weeks. The Measurements phase involved user identification measurements that were performed using WEKA (version 3.6.11), a popular machine learning software and used various machine learning methods such as Naive Bayes, Bayesian Network, Nearest neighbor, Decision Trees and Multilayer Perceptron (MLP). This paper demonstrates experimentally that touch screen based features improve the result and enhance the accuracy of both processes.

In [4] a user verification system on mobile phones is proposed. This system is based on behavioral biometric traits which is a keystroke dynamics derived from a touchable keyboard. A mobile application is developed for collecting those touch keystroke dynamics. The Median Vector Proximity classifier is applied on the touch keystroke data (touchable keyboard) and the performance of the system is investigated using different number of features. In this paper, they proposed to apply the Median Vector Proximity classifier on the touch keystroke data derived from mobile's touch screen. They evaluated the system on no specific text and extracted 31 and 33 touch features. The average EER were 12.9% and 12.2% respectively. They found that the average EER was reduced by about 0.7%. Therefore, the more features they used resulted in more accurate systems. As a future work, they will apply different classifiers on those features and find out which classifier would give better results.

The authors of [5] say that the automatic emotion recognition technology is an important part of building intelligent systems to prevent the computers acting inappropriately. A novel approach for recognizing emotional state by their keystroke typing patterns on a standard keyboard was developed in recent years. However, there was very limited investigation about the phenomenon itself in the previous literatures. Hence, in their study, they conduct a controlled experiment to collect subjects' keystroke data in the different emotional states induced by facial feedback. They examine the difference of the keystroke data between positive and negative emotional states. The results imply that when subjects were in different emotional states, they pressed the keyboard with different strength and demonstrated the tendency of the data on supporting the hypothesis. By conducting the controlled experiment, they validated the hypothesis about the existence of the difference on typing pattern between two opposite emotional states. The keystroke data were also applied in authentication system in previous studies to make the system more secure from hacking.

The work in [6] provides us with the results of a controlled experiment that was performed in which keystroke data of different users were collected in different emotional states induced by International Affective Digitized Sounds (IADS).The authors proposed an experiment designed to examine the effect of film-induced emotional states (PVHA, PVLA, NVHA, NVLA and NVNA (P = positive, N = negative, H = high, L = low, n = neutral, V =valence, A = arousal) in subjects, with the keystroke dynamics in regard to keystroke rate per second, average duration of keystroke (from key-down until key-up event). The results of their experiment using the fix target typing text and the 63 stimuli selected from the IADS-2 database supports the hypotheses that the keystroke duration and latency are influenced by arousal. Shorter keystroke duration was found when arousal was high compared to the keystroke duration when arousal was low. The result is in line with the findings reported by, which suggest a longer keystroke duration accompanied with negative emotional state. In addition, they found a slowest keystroke latency when arousal is medium. Their findings support the conclusion that the keystroke duration and latency are influenced by arousal and valence. This experiment was conducted in a PC using a keyboard.

The authors of [7] have simplified two main problems with current approaches for identifying emotions which are invasive and costly by using keystroke dynamics. They have conducted a field study where they collected participants' keystrokes and their emotional states via self-reports like questionnaires. Using an experience-sampling approach, users labelled the data with their level of agreement with 15 emotional states and provided additional keystrokes by typing fixed pieces of text. From that data, they extracted keystroke features and created classifiers for 15 emotional states. The results include 2-level classifiers for confidence, hesitance, nervousness, relaxation, sadness, and tiredness with accuracies ranging from 77 to 88% and show promise for anger and excitement, with accuracies of 84%. Some betterments to this approach are using a validated emotional state scale and distinguishing between different emotions rather than levels of a single emotion.

Chapter 3

Problem Statement

3.1 Drawbacks of current systems to check influence of emotions on Keystroke Dynamics

There are many papers and systems which use Keystroke Dynamics as a medium of authentication. But according to the literature survey mentioned in chapter 2, it was hypothesized that emotions play an important role in Keystroke Dynamics. Although there were limited studies exploring this perspective, the drawbacks of the current system to detect emotions was by prompting the subject at regular intervals to enter his or her emotions. This cannot be an effective way of studying the influence of emotions; a lot of context is lost in the syntax and semantics of the questions asked to the user or the user's understanding of the questions.

Also, In the previous studies it has been proved how emotions can be detected using keystroke dynamics with the help of the keyboard. But today touch screen mobiles are the new trend and people no longer use the keyboard.

3.2 Our approach

The work proposed in the project is to address the following problem statement:

To develop an Android application which will prove the Influence of Emotion on Keystroke Patterns and to conduct an Experimental Study Using Auditory Stimuli to prove the influence.

3.3 Experimental Setup

Subject is made to sit in a fully air-conditioned room without any external noise and disturbance. The subject is explained the entire process along with the significance of the SAM ratings. Two trails are done with the subject. Once the subject is familiar with the process, the actual samples are taken through the developed android application called Keydroid. The subject is made to listen to 63 sounds. For each sound the valence and arousal ratings are recorded along with the keystroke duration and keystroke latency. The user is made to type a text "748596132". The text is chosen so that the subject does not have to move much to search for the keys. The keys are placed close to each other but not in sequence so that accurate results can be obtained.

Chapter 4

Project Description

4.1 Overview

An Android application **KeyDroid** was developed to observe the role of emotions on keystroke dynamics. The advantages of using this approach are that the data used is rather non-intrusive and easy to obtain. This study aimed to examine the source of variance in keyboard typing patterns caused by emotions. A controlled experiment was performed to collect subjects' keystroke data in different emotional states induced by International Affective Digitized Sounds (IADS). We will make use of two-way Valence (3) x Arousal (3) ANOVAs to examine the collected dataset. Based on the dataset the relation between the keystroke patterns and the emotions of a particular person was observed. The project aimed at finding about how the latency and duration of a keystroke gets affected by different emotions.

4.2 Data Collection

4.2.1 Subject

Twenty-three subjects were asked to perform keyboard typing tasks right after they were presented with emotional stimuli. The subjects also reported that they have normal or corrected-to-normal vision and normal range of finger movement.

4.2.2 Procedure

A subject was made to wear earphones during the experiment and was be instructed to type-in a target typing text "748596132" once immediately after hearing each of the International Affective Digitized Sounds 2nd edition (IADS-2) sounds, for 63 trials. The experiment was conducted based on a simple dimensional view of emotion, which assumes that emotion can be defined by a coincidence of values on two different strategic dimensions that are, valence and arousal. To assess these two dimensions of the affective space, the Self-Assessment Manikin (SAM), an affective rating system will be used to acquire the affective ratings. Each trial began with an instruction ("Please type-in the target typing text after listening to the next sound") presented for 5 s. Then, the sound stimulus will be presented for 6 s. After the sound terminates, the SAM with a rating instruction ("Please rate your feeling on both the two dimensions after typing the target typing text '748596132'") was be presented. The subject was asked to first type-in the target typing text once, and then make his/her ratings of valence and arousal. A standard 15 s rating period was used, which allows ample time for the subject to make the SAM ratings. The keystroke data was recorded during the typing task. In addition to the 63 trials, 3 practice trials and a training section will be applied prior to the experiment. Three sounds (birds, female sigh, and baby cry) provided the subject with a rough range of the types of the contents that were presented. After these practice trials was the training section, in which the subject continually typed-in the target typing text using the number pad on the application. A number sequence was used as the target typing text instead of an alphabet sequence or symbols to avoid possible interference caused by linguistic context to the subject's emotional states. A comparison of keystroke typing between emotional states using different number sequences may reduce the power of statistical tests (given a same number of trials). Hence, to conduct a more conservative comparison across emotion and to enhance the generalizability of this study, we decided to use a single number sequence that is designed to be general. The target typing text was decided as "748596132" to 1) be easy to type without requiring the subjects to perform abrupt changes in their posture, 2) have the number of digits fairly distributed on a number pad, and 3) encourage all the subjects to maintain a same posture (i.e., in terms of finger usage) when typing the given sequence. The time length of the experiment is designed to be as short as possible to avoid the subjects from being tired of typing on the keyboard.

4.2.3 Stimuli and Self-Report

The stimuli used had 63 sounds selected from the IADS-2 database, which is developed and distributed by the NIMH Centre for Emotion and Attention (CSEA) at the University of Florida. The IADS-2 database contains various affective sounds proved to be capable of inducing diverse emotions in the affective space. The sounds used as the stimuli were selected from IADS-2 database complying the IADS-2 sound set selection protocol. The protocol includes the constraint about the number of sounds used in a single experiment, and the distribution of the emotions that are expected to be induced by the selected sounds. Two different stimulus orders were used to balance the position of a particular stimulus within the series across the subjects. The physical properties of these sounds were controlled to prevent clipping, and control for loudness.

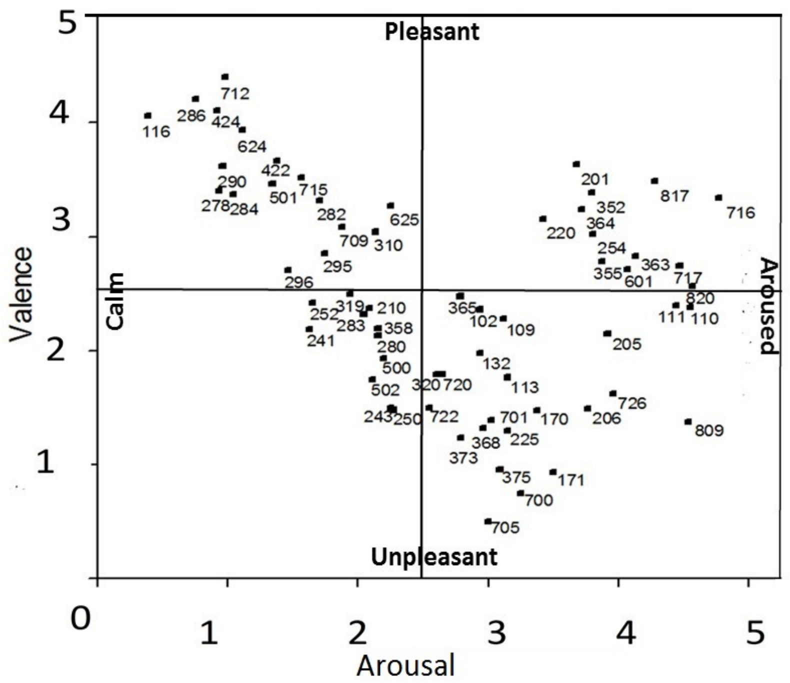

Figure 4.1: Graph of Chosen Sounds

Graph explaining how the 63 sounds that were selected change the emotions of a particular person. The graphs shows the sounds chosen to influence the emotions of a person covered a wide range of valence and arousal.

The SAM (Self-Assessment Manikin) is a non-verbal pictorial assessment designed to assess the emotional dimensions (i.e. valence and arousal) directly by means of two sets of graphical manikins. The SAM has been extensively tested in conjunction with the IADS-2 and has been used in diverse theoretical studies and applications. The SAM takes a very short time to complete (5 to 10 seconds). For using the SAM, there is little chance of confusion with terms as in verbal assessments. The SAM used is identical to the 5-point rating scale version of SAM, in which the SAM ranges from a smiling, happy figure to a frowning, unhappy figure when representing the affective valence dimension. On the other hand, for the arousal dimension, the SAM ranges from an excited, wide-eyed figure to a relaxed, sleepy figure. The SAM ratings will be rated such that 5 represents a high rating on each dimension (i.e. positive valence, high arousal), and 1 represents a low rating on each dimension (i.e. negative valence, low arousal).

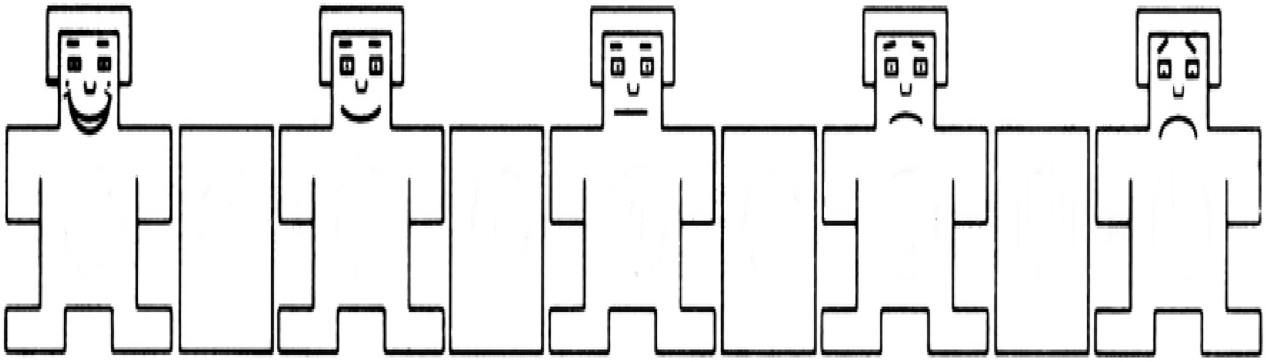

Figure 4.2: SAM Rating for Valence [8] [9]

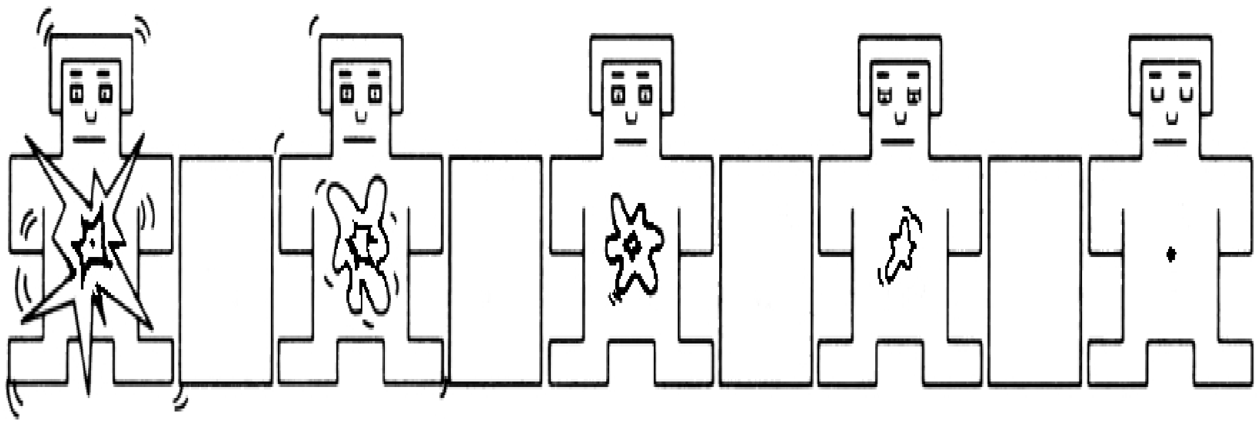

Figure 4.3: SAM Rating for Arousal [8] [9]

4.3 Preprocessing of Samples

1. The duration and latency collected in milli seconds was converted into seconds

2. The SAM rating taken as 1-5 was converted into Low, Medium and High

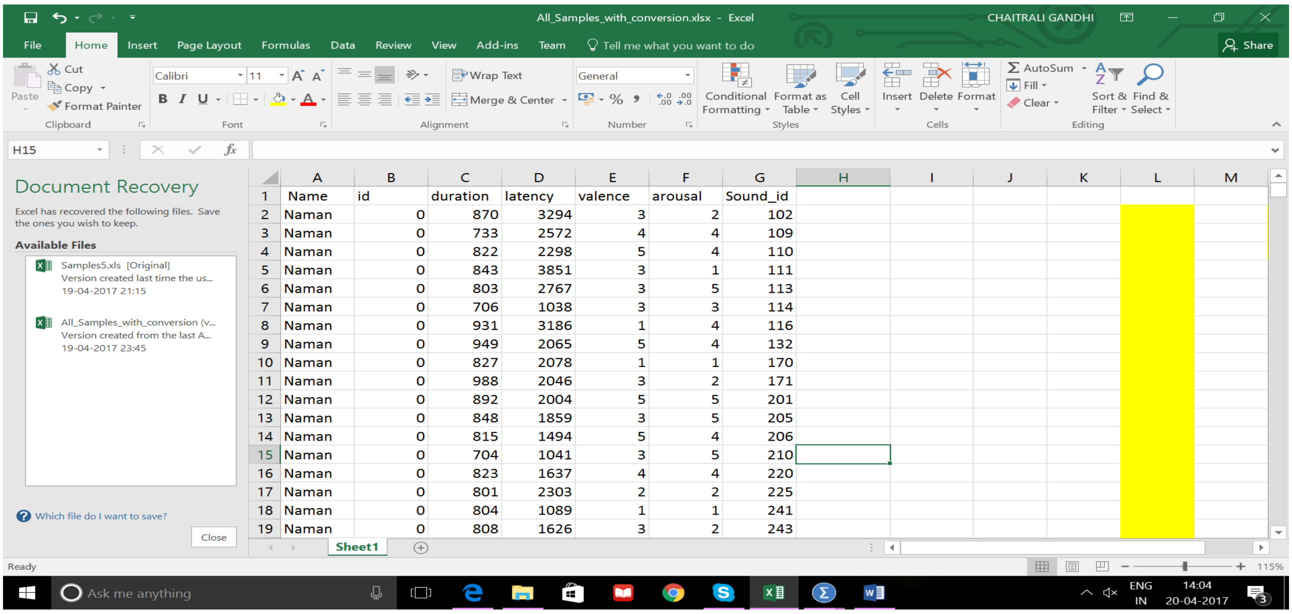

Figure 4.4: CSV file before Pre-processing

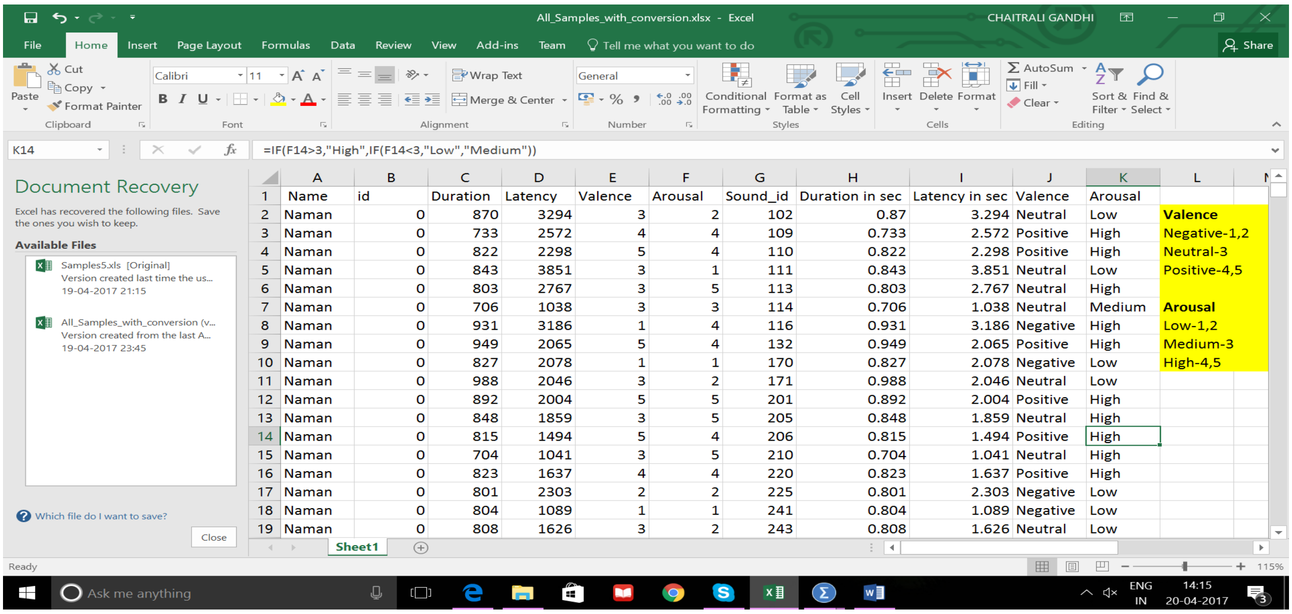

Figure 4.5: CSV file after Pre-processing

4.4 Calculations

The emotions of a subject was predicted based on relation between valence and arousal. IADS sounds elicited that the subjects' feelings of being annoyed or alarmed (i.e. reporting negative valence with medium arousal), but not being angry (i.e. reporting negative valence with high arousal) and not being tired, sad, or bored (i.e. reporting negative valence with low arousal). These values was useful to identify different emotional states. After the data collection, ANOVA was used to calculate the effect of emotions on keystroke patterns. Analysis of variance (ANOVA) is a collection of statistical models used to analyse the differences among group means and their associated procedures (such as "variation" among and between groups). Mainly the change in the keystroke duration based on different arousal and valence ratings was observed . The variance was conducted between arousal, valence, keystroke latency, duration to get the appropriate results.

4.5 Use case diagram

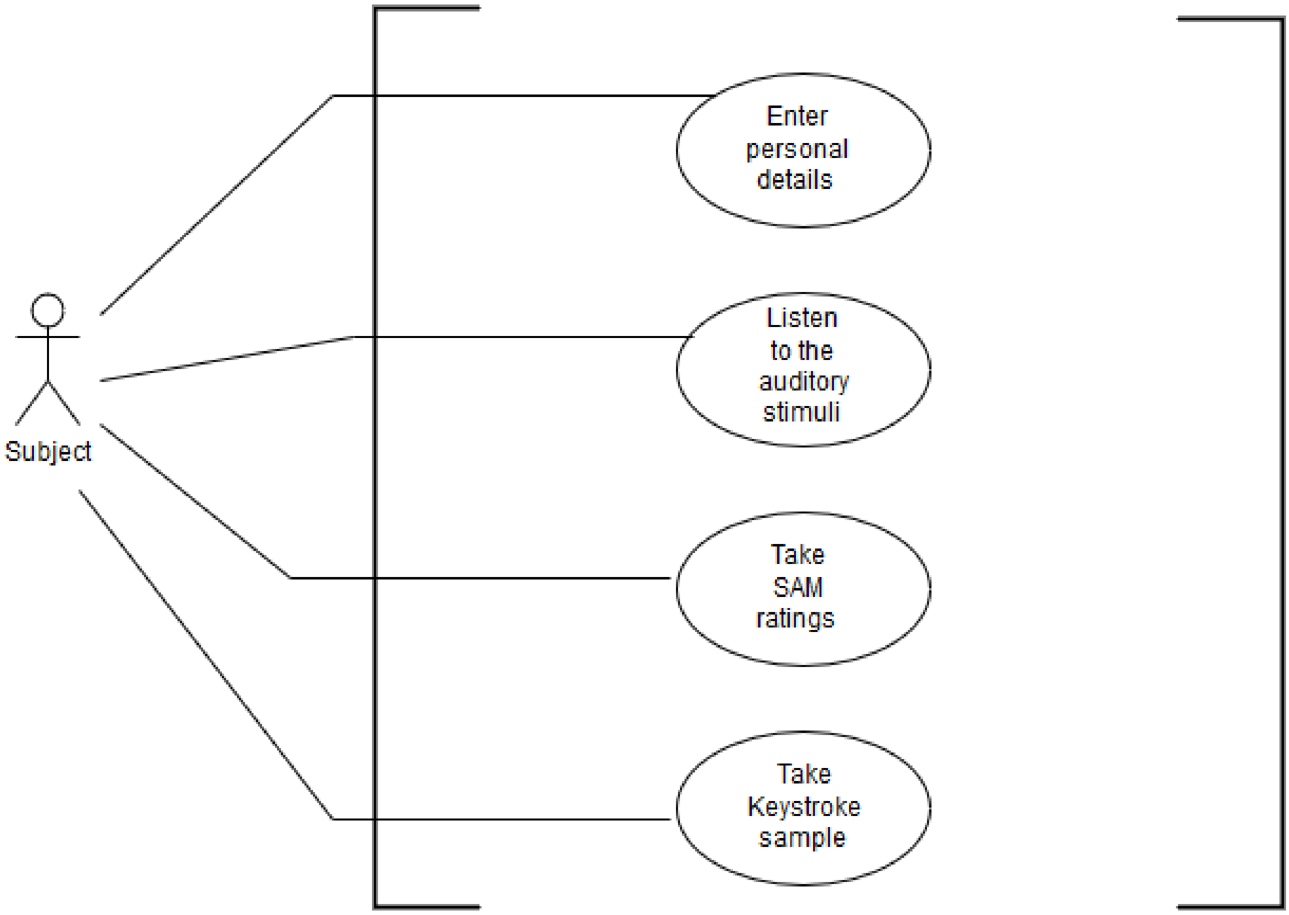

Figure 4.6: Use Case Diagram

4.6 Dataflow Diagram

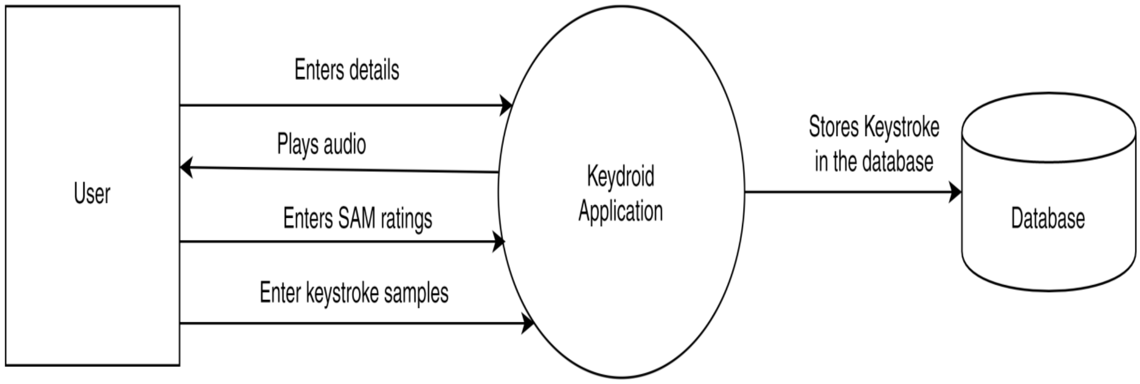

Figure 4.7: Data Flow Diagram

4.7 Flowchart

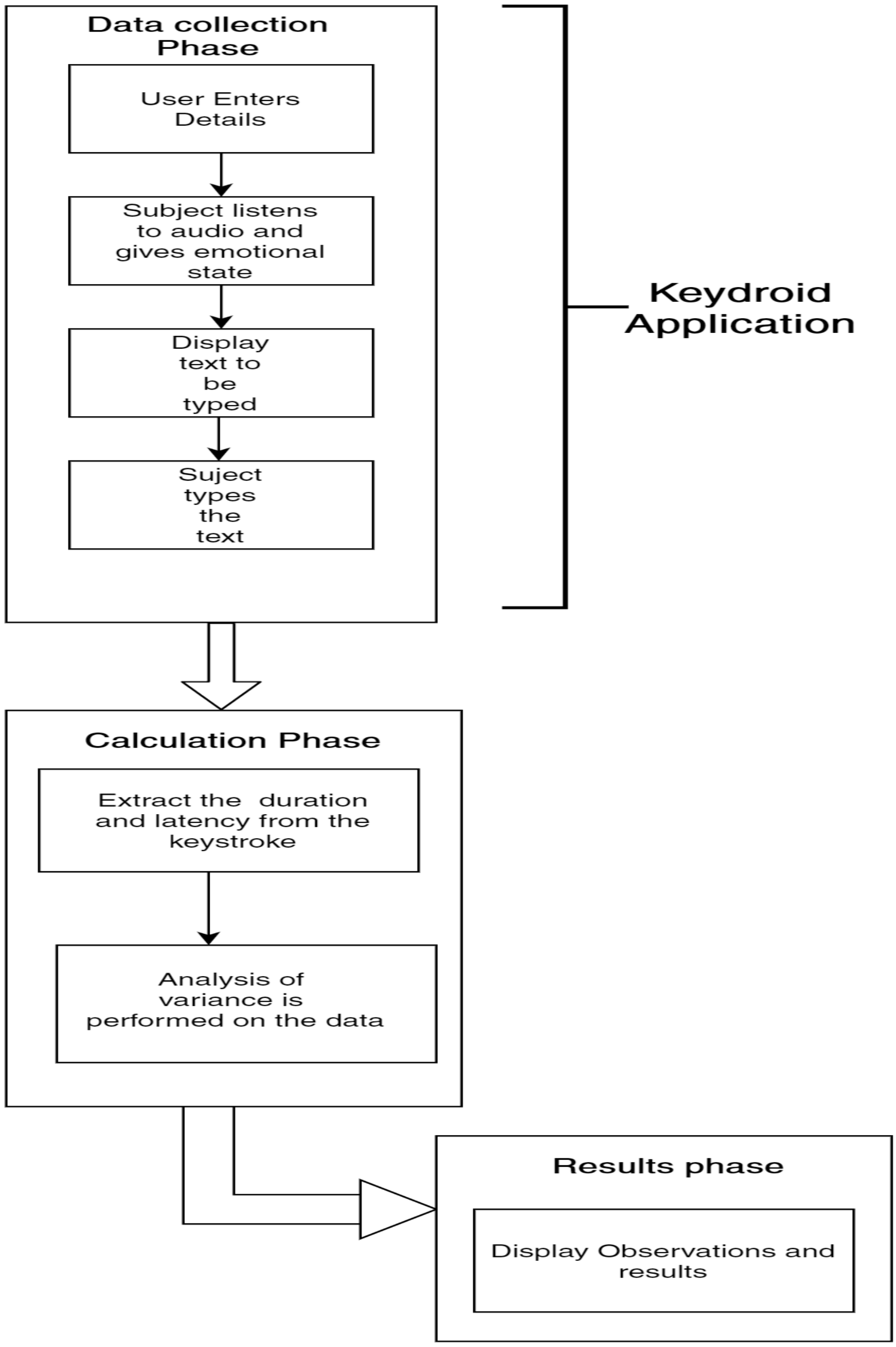

Figure 4.8: Flowchart

Chapter 5

System Testing

5.1 KeyDroid Technical Testing

To conduct the testing of KeyDroid application before taking the samples from subjects, it was tested by students who were not a part of the development phase. The following four tests were conducted:-

Test ID.	Description	Input Data	Expected Output	Actual Output	Status
1.	To check the input format of the age text box	Integer	It should not allow characters apart from Integers	The QWERTY keypad is not displayed Only Numeric keypad is given	Pass
2.	To make sure the input for the keystroke samples is collected when the screen is held in the same mode for accuracy	Text	To make sure the input for the keystroke samples is collected when the screen is held in the same mode for accuracy	If the mode is changed when the user enters the data, the text disappears and the user has to enter the text again	Pass
3.	To make sure the input is correctly entered	748596132	The subject should be asked to enter the correct text if not entered correctly	The application reloads the activity and prompts the subject to enter the correct details	Pass

4.	To make sure SAM ratings is given	The relevant valence and arousal image have to be pressed	The subject should be asked to give the SAM ratings if not given	The subject is prompted to give his/her SAM ratings	Pass

Table 5.1: Table indicating the various tests conducted to check technical competence

5.2 KeyDroid Testing by Subjects

After KeyDroid was tested for all its technical aspects, the application was handed over to the subjects for Keystroke data collections. The sample data was stored in a CSV (Comma Separated Values) file in the external storage of the experimental device - HTC Desire 828 in the following manner

Name	User_ID	Keystroke Duration	Keystroke Latency	Valence	Arousal	Sound_ID
Naman	0	870	3294	3	2	102
Naman	0	733	2572	4	4	109
Naman	0	822	2298	5	4	110
Naman	0	843	3851	3	1	111
Naman	0	803	2767	3	5	113
Naman	0	706	1038	3	3	114

Table 5.2: Table indicating how the subject's data is stored in the external storage of the Android device.

Chapter 6

Implementation Details

6.1 KeyDroid(Android Application)

KeyDroid, an android application was developed to collect Keystroke data from a large number of subjects. The application was developed in Android Studio and run on the device HTC Desire 828. The application consists of four activities. In Android an Activity is a Java code that supports a screen or UI. In other words, building block of the user interface is the activity. Activity class is a pre-defined class in Android and every application which has UI must inherit it to create window. The description of each activity is given as follows:-

- **MainActivity.java**

This activity presents the user with a form. The form is used to collect the user's name and age. There is a button submit, on the press of this button the next activity is opened and the user details is passed to this activity.

- **PlayMusic.java**

This activity displays a message asking the user to listen to the IADS sound. There is a button called "Next", on the click of this button the next activity is opened.

- **Samrating.java**

This activity displays SAM ratings i.e. multiple images depicting five different values of valence and five different values of arousal. The subject is supposed to click on one image of valence and arousal each which resonates with his/her emotion the most at that instant. These values are passed to the next activity.

- **Keyboard.java**

This is the most important activity which handles the entire application. In this activity the user is asked to enter the numeric text given. The user is presented with a numeric keypad which is not the usual Android keypad. Each numeric button is programmed to calculate the keystroke latency and duration. After the user has entered the data, a CSV (Comma Separated Values) file is created to store all the details and samples of the subject. This file is stored in the external storage (SD card) of the device.

**Due to owner permissions of HTC Desire 828, the database created cannot be accessed even after giving root permissions, also a file stored in the Internal storage of the device cannot be accessed, Due to this reason, the CSV file is stored in the external storage.*

6.1.1 Screenshots of the application

Figure 6.1: KeyDroid-Form to collect user details.

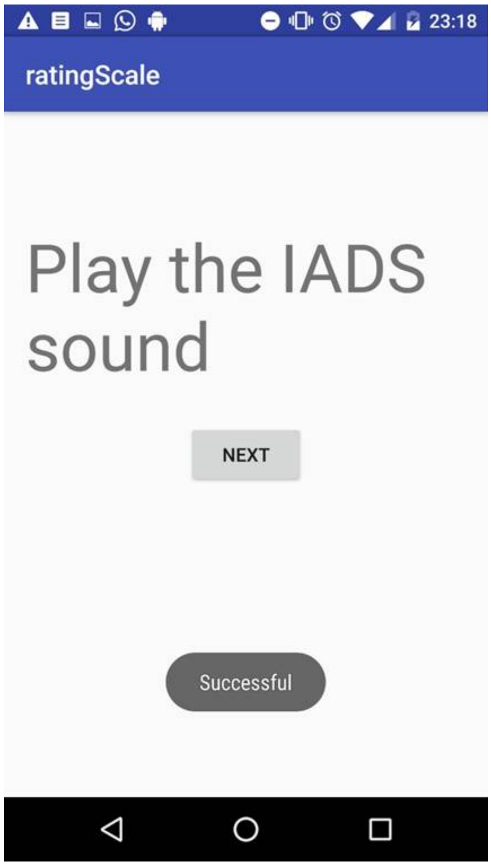

Figure 6.2: KeyDroid-Message to play IADS sound.

Figure 6.3: KeyDroid-SAM Ratings.

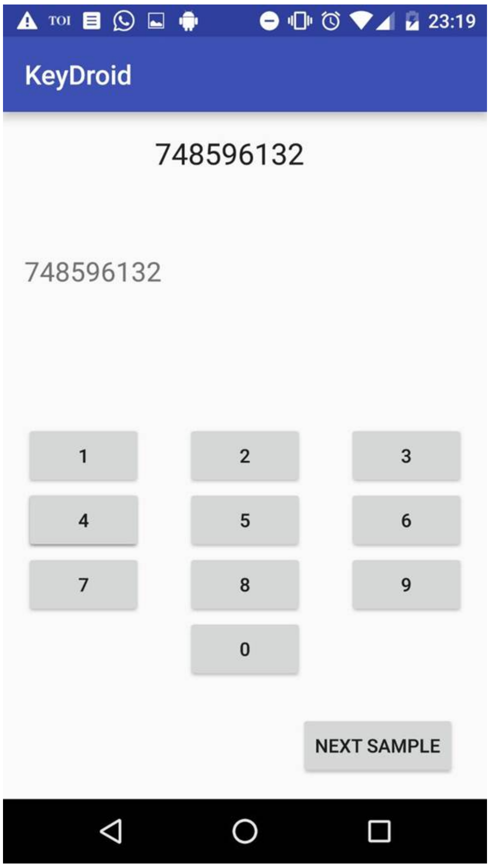

Figure 6.4: KeyDroid-Keypad for Sample Collection.

6.2 Anova (Calculation Tool)

The two-way ANOVA compares the mean differences between groups that have been split on two independent variables. The primary purpose of a two-way ANOVA is to understand if there is an interaction between the two independent variables on the dependent variable.

6.2.1 IBM SPSS Statistics

This software was used to calculate two-way ANOVA. IBM SPSS Statistics is a software package used for logical batched and non-batched statistical analysis. The software was released in its first version in 1968 as the Statistical Package for the Social Sciences (IBM SPSS) after being developed by Norman H. Nie, Dale H. Bent, and C. Hadlai Hull. Those principles incorporated as IBM SPSS Inc. in 1975. Early versions of IBM SPSS Statistics were designed for batch processing on mainframes, including IBM and ICL versions IBM SPSS is a widely used program for statistical analysis in social science. It is also used by market researchers, health researchers, survey companies, government, education researchers, marketing organizations, data miners,and others.

6.2.2 Steps to calculate two-way ANOVA using IBM SPSS Statistics

1. Go to analyze
2. Select General Linear Model
3. Select Univariate
4. Select duration or latency as dependent variable
5. Select valence and arousal as fixed variables
6. Select continue to obtain the results

Chapter 7

Experimental Results and Comparison

7.1 Influence of emotions on Keystroke Latency

Tests of Between-Subjects Effects

Dependent Variable: Latency

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	23.967[a]	8	2.996	2.196	.025
Intercept	4209.122	1	4209.122	3085.951	.000
Valence	9.013	2	4.506	3.304	.037
Arousal	12.291	2	6.145	4.506	.011
Valence * Arousal	6.673	4	1.668	1.223	.299
Error	1962.742	1439	1.364		
Total	7896.046	1448			
Corrected Total	1986.710	1447			

a. R Squared = .012 (Adjusted R Squared = .007)

Figure 7.1: ANOVA result for Keystroke Latency

2-way ANOVA results for finding the interaction between emotions and keystroke latency: The table depicts influence of valence and arousal individually on latency since value of Sigma is **less than 0.05**

Descriptive Statistics

Dependent Variable: Latency

Valence	Arousal	Mean	Std. Deviation	N
1	1	2.132930	1.0864757	273
	2	2.091282	1.3217907	78
	3	2.070640	1.2592666	89
	Total	2.112948	1.1646433	440
2	1	2.109721	1.3039234	179
	2	1.781714	.9402722	84
	3	1.688695	.8277345	82
	Total	1.929788	1.1359861	345
3	1	2.136879	1.1805752	124
	2	1.866667	1.1409045	123
	3	2.007548	1.2082435	416
	Total	2.005600	1.1919839	663
Total	1	2.126568	1.1760084	576
	2	1.903102	1.1427610	285
	3	1.972572	1.1748608	587
	Total	2.020157	1.1717446	1448

Figure 7.2: Descriptive statistics for Keystroke Latency

Descriptive Statistics depicting the mean and standard deviation for different valence and arousal combinations for Keystroke Latency.

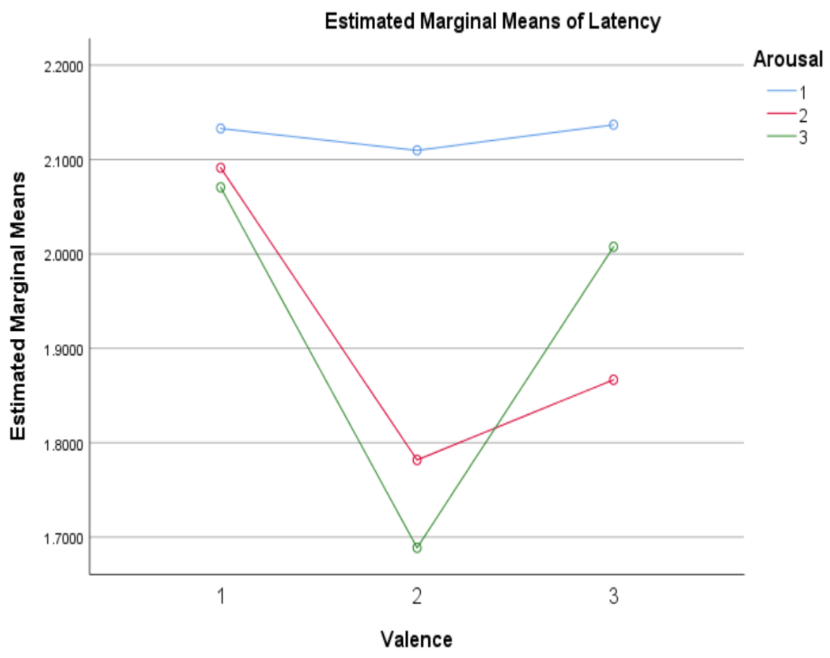

Figure 7.3: Line Graph of Keystroke Latency VS Valence and Arousal

Line Graph showing keystroke latency changes for different values of valence and arousal generated from IBM SPSS

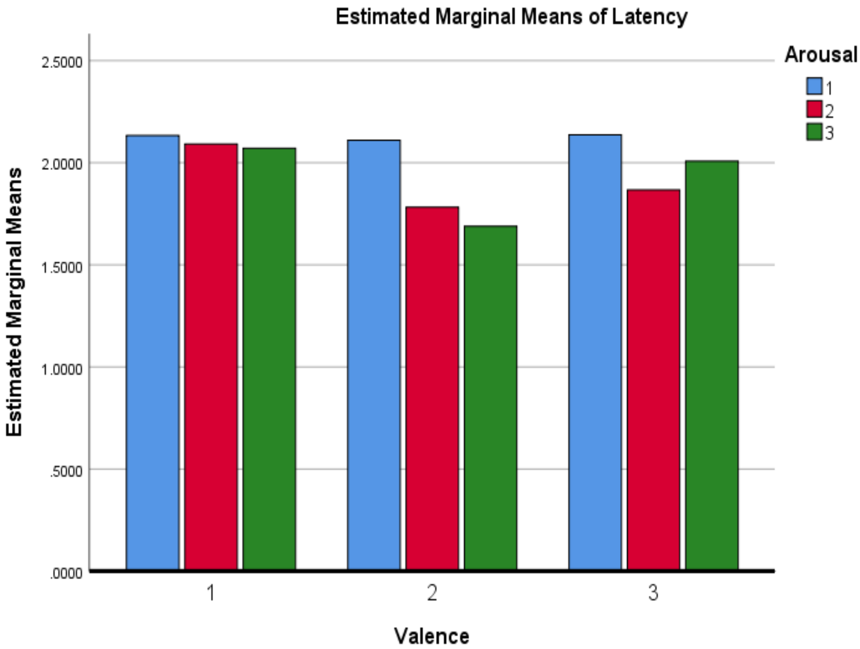

Figure 7.4: Bar Graph of Keystroke Latency VS Valence and Arousal

Bar Graph showing keystroke latency changes for different values of valence and arousal generated from IBM SPSS

7.2 Influence of emotions on Keystroke Duration

Tests of Between-Subjects Effects

Dependent Variable: Duration

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	4.284[a]	8	.536	8.930	.000
Intercept	796.385	1	796.385	13279.672	.000
Valence	2.606	2	1.303	21.726	.000
Arousal	.024	2	.012	.197	.821
Valence * Arousal	2.688	4	.672	11.205	.000
Error	86.297	1439	.060		
Total	1144.249	1448			
Corrected Total	90.581	1447			

a. R Squared = .047 (Adjusted R Squared = .042)

Figure 7.5: ANOVA result for Keystroke Duration

2-way ANOVA results for finding the interaction between emotions and keystroke duration: Table depicting influence of valence and combined vale of valence and arousal on duration since value of Sigma is **less than 0.05**

Descriptive Statistics

Dependent Variable: Duration

Valence	Arousal	Mean	Std. Deviation	N
1	1	.850304	.2300251	273
	2	.982256	.4239849	78
	3	.972955	.3690617	89
	Total	.898505	.3089205	440
2	1	.813749	.2018712	179
	2	.833071	.2166300	84
	3	.791512	.1323694	82
	Total	.813168	.1916875	345
3	1	.911589	.3472557	124
	2	.783455	.1722973	123
	3	.841130	.1904667	416
	Total	.843608	.2283850	663
Total	1	.852137	.2542882	576
	2	.852488	.2863681	285
	3	.854186	.2266730	587
	Total	.853037	.2501986	1448

Figure 7.6: Descriptive statistics for Keystroke Duration

Descriptive Statistics depicting the mean and standard deviation for different valence and arousal combinations for Keystroke duration

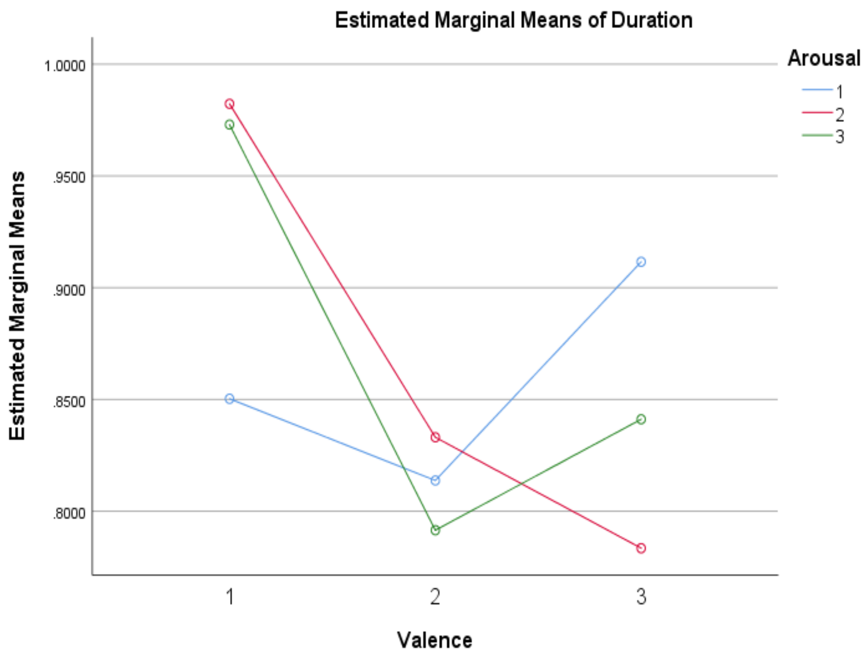

Figure 7.7: Line Graph of Keystroke Duration VS Valence and Arousal

Line Graph showing keystroke duration changes for different values of valence and arousal generated from IBM SPSS

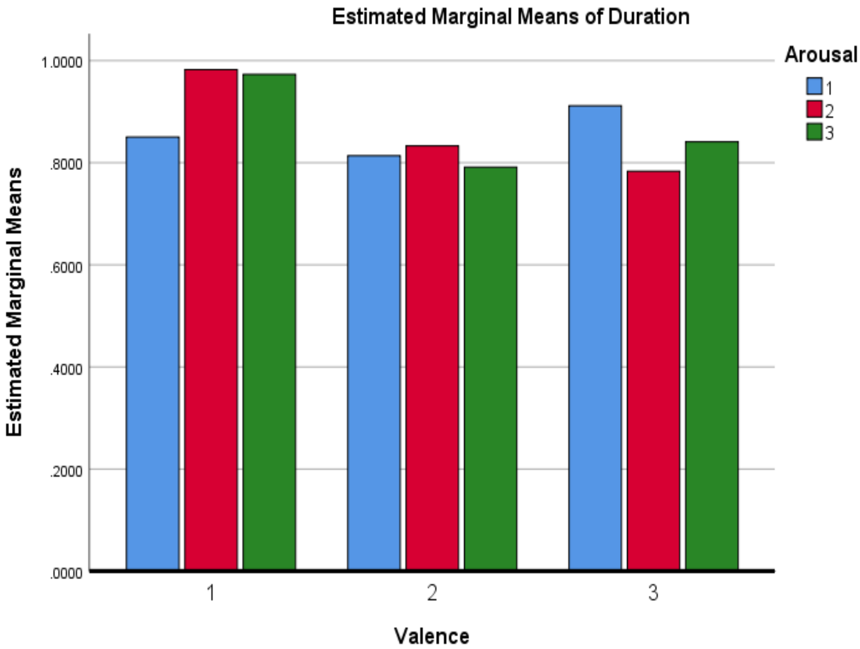

Figure 7.8: Bar Graph of Keystroke Latency VS Valence and Arousal
Bar Graph showing keystroke duration changes for different values of valence and arousal generated from IBM SPSS

7.3 Final Results

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	Duration	4.284[a]	8	.536	8.930	.000
	Latency	23.967[b]	8	2.996	2.196	.025
Intercept	Duration	796.385	1	796.385	13279.672	.000
	Latency	4209.122	1	4209.122	3085.951	.000
Valence	Duration	2.606	2	1.303	21.726	.000
	Latency	9.013	2	4.506	3.304	.037
Arousal	Duration	.024	2	.012	.197	.821
	Latency	12.291	2	6.145	4.506	.011
Valence * Arousal	Duration	2.688	4	.672	11.205	.000
	Latency	6.673	4	1.668	1.223	.299
Error	Duration	86.297	1439	.060		
	Latency	1962.742	1439	1.364		
Total	Duration	1144.249	1448			
	Latency	7896.046	1448			
Corrected Total	Duration	90.581	1447			
	Latency	1986.710	1447			

a. R Squared = .047 (Adjusted R Squared = .042)

b. R Squared = .012 (Adjusted R Squared = .007)

Figure 7.9: Table showing ANOVA result of influence of valence and arousal in both keystroke duration and keystroke latency

Chapter 8

Conclusion

Previous studies [2, 4, 5, 7] have highlighted the possibility of using keyboard typing data to detect emotions. In [2] however the method chosen for detecting emotions is the categorical approach i.e. it applies labels to emotions with some languages or words (e.g. sadness, fear, joy) However the approach that we have chosen is the dimensional approach that uses two orthogonal axes called arousal and valence to describe emotions. The method chosen in [2] repeatedly asked the user to enter his emotions in words which could not be considered as an effective measure of emotion detection. The top result for confident state was accuracy between 76% to 90% and sadness, anger state was accuracies between 78% to 85%.

In this project, auditory stimuli was used to find the influence of emotions on keystroke duration and latency. This is a non-intrusive method. During comparison of the first and last sample of each of the 23 subjects, it was observed that keystroke latency was decreased by a mean of 2.6 seconds. One of the main reason for this is that the subject became well versed with the typing pattern and typed the sequence faster than the first sample .

From ANOVA, it has been observed that there is an effect of valence and arousal in Keystroke Latency. It was also found that for low or negative valence the mean keystroke latency was the highest of 2.11 seconds. This indicates that the subject takes longer to type when he is bored or in an unpleasant mood. Thus the hypothesis that valence and arousal plays a role in Keystroke latency has been proved.

However, from ANOVA, there is an effect of valence in keystroke duration but the effect of arousal on the same could not be proven. That being said, from the descriptive statistics it is observed that for high arousal, the keystroke duration is maximum with 0.854 seconds. This could mean that when the user is angry or excited, he/she has the tendency to press the keys harder.

One was the main reasons why the effect of arousal was not proven on Keystroke duration is that there were considerably less number of samples collected for high arousal. Future work in this genre should make sure that there are more number of samples collected in high arousal values.

Thus this project was able to prove the effect of both valence and arousal in Keystroke Latency and the effect of valence in Keystroke Duration.

Chapter 9

Future Scope

This project has a wide future scope if taken further. The following are two such applications:-

1. **As an effective mood detection system in social media**
   Current social media gives a lot of importance to user data. The user's emotions can play a huge role in what recommendations should be presented to the user. Currently Facebook asks the user to enter his/her mood i.e. feeling happy, feeling angry etc. Our system will not require to enter his/her emotions on their own. Rather, judge their emotions through their keystroke. Since this is just a future scope. It will require multiple improvements in order to develop this solution further to suit the application

2. **As an effective measure for depression**
   The current system of measuring depression is tedious. The present system includes a survey with Likert's Management System. This cannot be an effective way of studying the influence of emotions; a lot of context is lost in the syntax and semantics of the questions asked to the patient or the patient's understanding of the questions. Since Auditory stimuli is effective in changing the keystroke of a person, and with keystroke dynamics we can guess the emotions of the person, The amount of depression levels in a patient can be deduced without the complexity of filling endless surveys . However this will require a few improvements in the current system.

Appendix A

Appendix

```
package com.example.android.keydroid;

import android.content.Context;
import android.content.Intent;
import android.os.Bundle;
import android.os.Environment;
import android.support.v7.app.AppCompatActivity;
import android.util.Log;
import android.view.MotionEvent;
import android.view.View;
import android.widget.Button;
import android.widget.TextView;
import android.widget.Toast;

import java.io.File;
import java.io.FileNotFoundException;
import java.io.FileOutputStream;
import java.io.IOException;
import java.io.PrintWriter;
import java.util.Vector;

public class Keyboard extends AppCompatActivity implements View.
    OnTouchListener {
    DatabaseAdapter dbhelper;
    public Button mb1, mb2, mb3, mb4, mb5, mb6, mb7, mb8, mb9, mb0;
    TextView txtView;
    String display;
    long press;
    long release;
    long keyhold;
    long duration_sum = 0, latency, latency_sum = 0;
```

```
    Vector p = new Vector<Long>();
    Vector r = new Vector<Long>();

    @Override
    protected void onCreate(Bundle savedInstanceState) {
        super.onCreate(savedInstanceState);
        setContentView(R.layout.activity_keyboard);
        dbhelper = new DatabaseAdapter(this.getApplicationContext());
        mb1 = (Button) findViewById(R.id.one);
        mb2 = (Button) findViewById(R.id.two);
        mb3 = (Button) findViewById(R.id.three);
        mb4 = (Button) findViewById(R.id.four);
        mb5 = (Button) findViewById(R.id.five);
        mb6 = (Button) findViewById(R.id.six);
        mb7 = (Button) findViewById(R.id.seven);
        mb8 = (Button) findViewById(R.id.eight);
        mb9 = (Button) findViewById(R.id.nine);
        mb0 = (Button) findViewById(R.id.zero);
        txtView = (TextView) findViewById(R.id.textbox);
        mb1.setOnTouchListener(this);
        mb2.setOnTouchListener(this);
        mb3.setOnTouchListener(this);
        mb4.setOnTouchListener(this);
        mb5.setOnTouchListener(this);
        mb6.setOnTouchListener(this);
        mb7.setOnTouchListener(this);
        mb8.setOnTouchListener(this);
        mb9.setOnTouchListener(this);
        mb0.setOnTouchListener(this);
    }

    public boolean onTouch(View v, MotionEvent event) {
        if (event.getAction() == MotionEvent.ACTION_DOWN) {
            Button b = (Button) v;
            String display = b.getText().toString();
            txtView.append(display);
            press = System.currentTimeMillis();
            if (!r.isEmpty()) {
                long prev_release = Long.valueOf(r.lastElement().toString()
                    );
                latency = press - prev_release;
            }            p.addElement(press);
```

```
        //System.out.println("The keypress time is " + press);
      } else if (event.getAction() == MotionEvent.ACTION_UP) {
        release = System.currentTimeMillis();
        keyhold = release - press;

        r.addElement(release);
        duration_sum = duration_sum + keyhold;
        latency_sum = latency_sum + latency;
      }

      System.out.println("Time duration: " + String.format("%.3f",
          duration_sum / 1000.0));
      System.out.println("Time latency: " + String.format("%.3f",
          latency_sum / 1000.0));

      //Toast.makeText(this,sum/1000.0+"",Toast.LENGTH_LONG).show();//
          store in db
      return true;
  }

   public void NextSample(View view) {
      final Context context = this;
      Bundle extras = getIntent().getExtras();
      String pleasure = extras.getString("PLEASURE");
      String arousal = extras.getString("AROUSAL");
      int user_id = extras.getInt("USER_ID");
      String name = extras.getString("user_name");
      String duration = Long.toString(duration_sum);
      String latency = Long.toString(latency_sum);
      Users u = DatabaseAdapter.insertSamples(user_id,latency, duration,
          pleasure, arousal);
      if (u == null) {
        Message.message(this, "Unsuccessful");
      } else {
        Toast.makeText(this, getFilesDir()+"", Toast.LENGTH_LONG).show
            ();

        //String filename = "myfile.txt";
        String string = name + "," + u.getId() + "," + u.getDuration()
            + "," + u.getLatency() + "," + u.getValence() + "," + u.
            getArousal();
        FileOutputStream outputStream;
```

```
            File root = android.os.Environment.getExternalStorageDirectory
                ();
            File dir = new File (root.getAbsolutePath() + "/lil");
            dir.mkdirs();
            File file = new File(dir, "Keydroid.csv");
            try {
                FileOutputStream f = new FileOutputStream(file,true);
                PrintWriter pw = new PrintWriter(f);
                pw.println(string);
                pw.flush();
                pw.close();
                f.close();
                Toast.makeText(this, file.toString(), Toast.LENGTH_LONG).
                    show();
            } catch (FileNotFoundException e) {
                e.printStackTrace();
            } catch (IOException e) {
                e.printStackTrace();
            }

            Intent intent = new Intent(context, PlayMusic.class);
            intent.putExtra("user_id",user_id);
            intent.putExtra("user_name",name);
            startActivity(intent);
        }


    }
    public boolean isExternalStorageWritable() {
        String state = Environment.getExternalStorageState();
        if (Environment.MEDIA_MOUNTED.equals(state)) {
            return true;
        }
        return false;
    }

    /* Checks if external storage is available to at least read */
    public boolean isExternalStorageReadable() {
        String state = Environment.getExternalStorageState();
        if (Environment.MEDIA_MOUNTED.equals(state) ||
                Environment.MEDIA_MOUNTED_READ_ONLY.equals(state)) {
```

```
                return true;
            }
            return false;
        }
}
```

Listing A.1: Code Snippet in JAVA where keystroke Duration and Latency is calculated and stored in the external storage

References

[1] I. L. Margit ANTAL, LászlóZsolt SZABÓ, "Keystroke dynamics on android platform," October 2014.

[2] P. V. Shivshankar Rajput, "Implementation of keystroke dynamics application for identifying emotional state," May 2015.

[3] L. A. E. Shatha J. Alghamdi, "Dynamic user verification using touch keystroke based on medians vector proximity," *Computational Intelligence, Communication Systems and Networks (CICSyN)*, June 2015.

[4] T.-C. H. Po-Ming Lee, "The effect of emotion on keystroke: An experimental study using facial feedback hypothesis," *Plos One*, June 2015.

[5] T.-C. H. Po-Ming Lee, Wei-HsuanTsui, "The influence of emotion on keyboard typing: An experimental study using auditory stimuli," *Plos One*, June 2015.

[6] M. M. Pranit Shinde, Saideep Shetty, "Survey of keystroke dynamics as a biometric for static authentication„" April 2016.

[7] R. L. M. Clayton Epp, Michael Lippold, "Identifying emotional states using keystroke dynamics," *Plos One*, June 2015.

[8] P. J. Lang, "Behavioral treatment and bio-behavioral assessment: Computer applications," *J. B. Sidowski, J. H. Johnson, T A. Williams (Eds.), Technology in mental health care delivery systems (pp. 119-137)*, 1980.

[9] . L. P. J. Bradley, M. M., "Measuring emotion: The self-assessment manikin and the semantic differential," *Journal of Behavioral Therapy and Experimental Psychiatry, 25(1), 49-59.*, 1994.