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

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

Bachelor of Engineering

by

Lillita Rhea D'souza (Roll No. 7126)
Chaitrali Gandhi (Roll No. 7133)
Naman 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
November 4, 2016

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 Gandhi(7133), Naman Malik(7146)** submitted to the University of Mumbai in partial 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 Gandhi, Naman 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 Gandhi (Roll No. 7133) **(sign)**________________

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

Date: November 4, 2016

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. It has been observed that keystrokes vary according to a person's emotions. If any system could be built that is intelligent enough to interact with humans that involves emotions, that is, it can detect the user emotions and change its behaviour accordingly, then using machines could be more effective and friendly. Many approaches have been taken to detect user emotions. This project will be based on the use of auditory stimuli to detect the emotions of the individual. 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 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"**. I 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 Gandhi(Roll No. 7133)

Naman Malik(Roll No. 7146)

Date: November 4, 2016

Contents

List of Figures

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.

Chapter 2

Problem Statement

Emotions play a very important part in today's life. In 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 key board. Therefore this research based project will be developed 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.

Chapter 3

Literature Review

The authors of [1] 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. They made use of time based features to perform the experiment and find out the results. An Android application having its own software keyboard was developed for data collection. Users had to introduce some personal data, such as gender, birth date and their experience level regarding smart phone usage in the registration phase. Because typing pattern can be influenced by several factors, data should be collected in several sessions. 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. The authors selected some well-known algorithms implemented in WEKA, covering various machine learning methods such as Naive Bayes, Bayesian Network, Nearest neighbor , Decision Trees and Multilayer Perceptrons (MLP). In this paper they demonstrated experimentally that touch screen based features improve keystroke dynamics based identification and verification. And the results show that these additional features enhance the accuracy of both processes.

[2] 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 the present paper, 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. In this paper 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.

In [3] 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 [4] 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. The results also 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.

[5] 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, which implied that button presses may have been carried out with less strength when arousal was low. This result indicates an increased keystroke duration when the subjects experienced tired, sad, or bored. 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. This experiment was conducted in a PC using a keyboard.

Chapter 4

Project Description

4.1 Introduction

In this Project, an Android application will be 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 aims to examine the source of variance in keyboard typing patterns caused by emotions. A controlled experiment to collect subjects' keystroke data in different emotional states induced by International Affective Digitized Sounds (IADS) will be performed and a two-way Valence (3) x Arousal (3) ANOVAs will be used to examine the collected dataset. Based on the dataset we are going to observe the relation between the keystroke patterns and the emotions of a particular person. The work in this project mainly aims at finding about how the latency and duration of a keystroke gets affected by different emotions.

4.2 Data Collection

4.2.1 Subject

Fifty-two subjects ranging in age between 20 and 26 perform keyboard typing tasks right after presented with emotional stimuli. All the subjects will be self-reporting that they are non- smokers, healthy, with no history of brain injury and cardiovascular problems. The subjects will also report that they have normal or corrected-to-normal vision and normal range of finger movement. They will be right-handed.

4.2.2 Procedure

A subject will wear earphones during the experiment and will 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 will be 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 will begin 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 6s. 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'") will be presented. The subject will first type-in the target typing text once, and then make his/her ratings of valence and arousal. A standard 15 s rating period will be used, which allows ample time for the subject to make the SAM ratings. The keystroke data will be 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) provide 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 will be 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, a single number sequence has been designed to be general. The target typing text "748596132" has been designed in such a way that it will

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 we will use has 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 which will be 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 will be used to balance the position of a particular stimulus within the series across the subjects. The physical properties of these sounds will also be controlled to prevent clipping, and to control for loudness.

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 that will be used is identical to the 9-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 9 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).

4.3 Calculations

The emotions of a subject will be predicted based on relation between valence and arousal. IADS sounds have 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 will help us to identify different emotional states.

After the data collection, ANOVA will be used to calculate the effect of emotions on keystroke patterns. Analysis of variance (ANOVA) is a collection of statistical models used to analyze 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 will be observed. The variance will be conducted between arousal, valence, keystroke latency and accuracy to get the appropriate results.

4.4 Project Design

4.4.1 Block Diagram

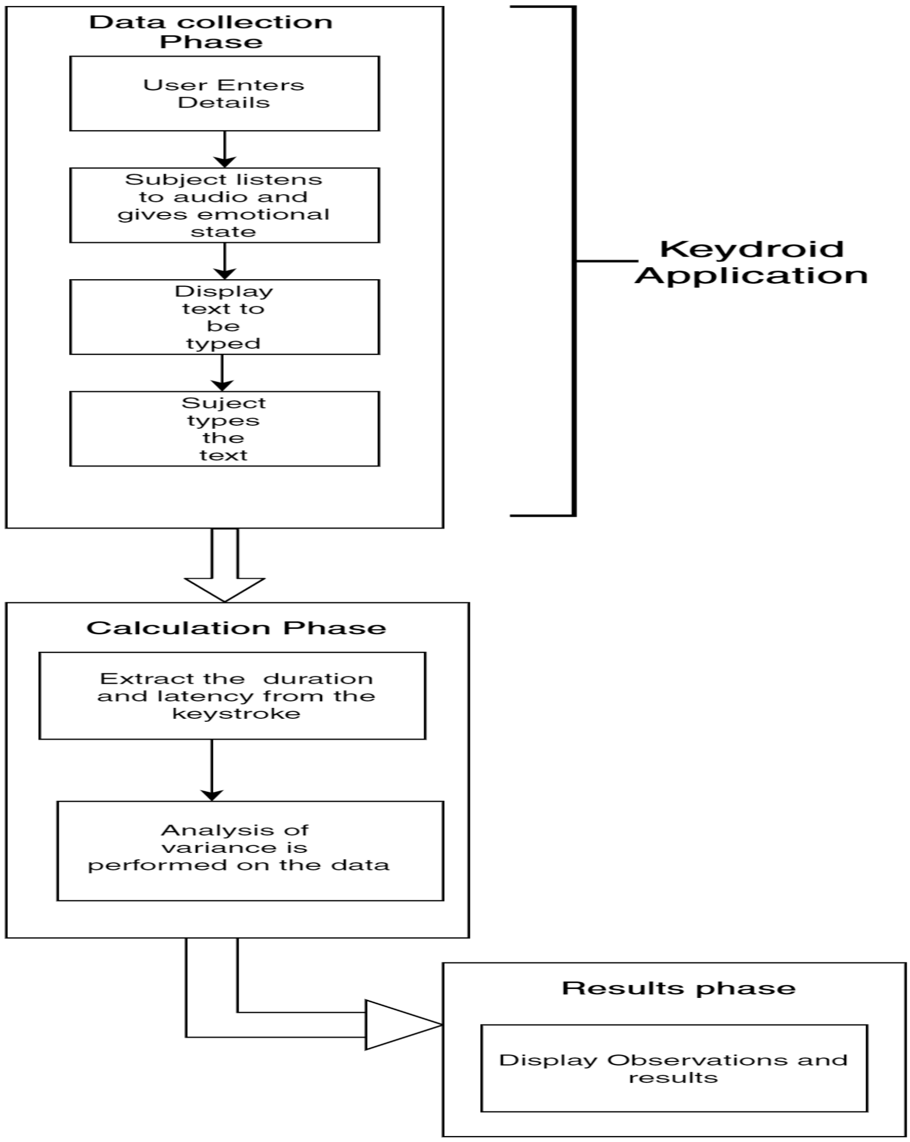

Figure 4.1: Block Diagram

.

4.4.2 Use Case Diagram

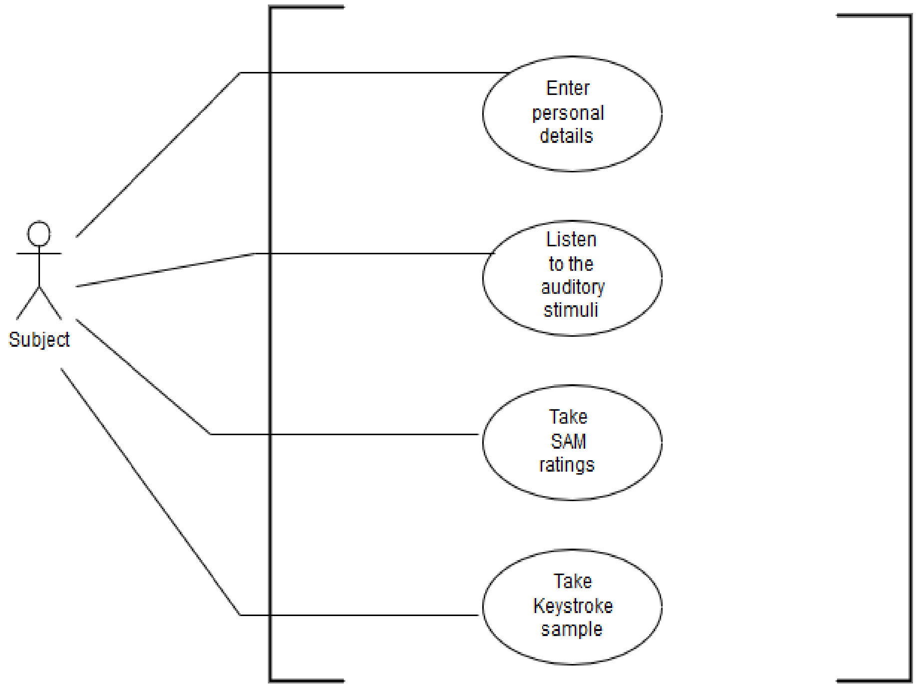

Figure 4.2: Use Case Diagram

.

4.4.3 Data Flow Diagram

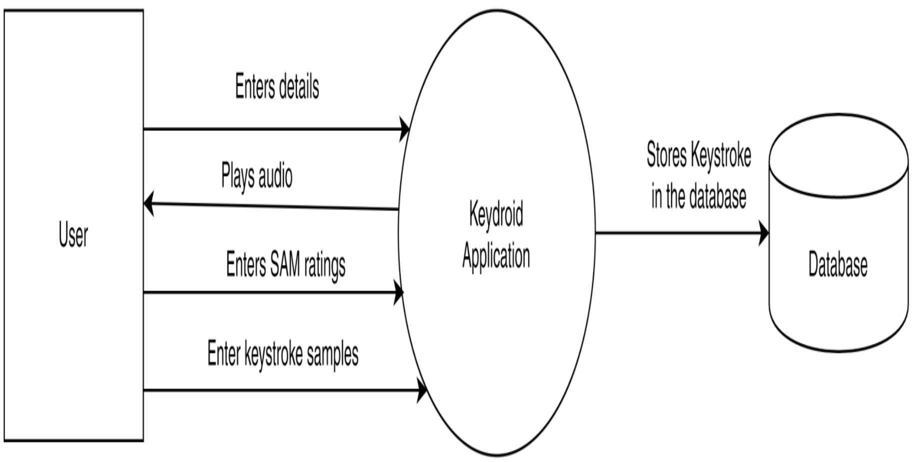

Figure 4.3: Data Flow Diagram

Chapter 5

Project Schedule

	Task Name	Duration	Start	Finish	Predecessors
	Learning Curve	**56 days**	**Mon 8/8/16**	**Tue 10/25/16**	
	Literature Survey	21 days	Mon 8/8/16	Mon 9/5/16	
	Learning Android	30 days	Mon 9/5/16	Fri 10/14/16	
	Learning ANOVA	20 days	Mon 9/5/16	Fri 9/30/16	
	Milestone: Learning curve is com	0 days	Tue 10/25/16	Tue 10/25/16	
	Implementation and Coding	**51 days?**	**Wed 10/26/16**	**Wed 1/4/17**	**1**
	User Interface Designing	7 days	Wed 10/26/16	Thu 11/3/16	
	Backend Development	15 days	Fri 11/4/16	Thu 11/24/16	7
	Database Connectivity	10 days?	Thu 12/22/16	Wed 1/4/17	
	Milestone: Implementation is Com	0 days	Wed 1/4/17	Wed 1/4/17	
	Research	**39 days?**	**Mon 1/16/17**	**Thu 3/9/17**	**1,6**
	Sample collection from subjects	30 days?	Mon 1/16/17	Fri 2/24/17	
	Applying ANOVA	10 days?	Fri 2/24/17	Thu 3/9/17	
	Milestone: Research is complete	0 days	Thu 3/9/17	Thu 3/9/17	
	Findings	**10 days?**	**Fri 3/10/17**	**Thu 3/23/17**	**1,6,11**
	Report Generation	10 days?	Fri 3/10/17	Thu 3/23/17	
	Milestone: Report has been gene	0 days	Wed 3/22/17	Wed 3/22/17	

Figure 5.1: Project Schedule developed in Microsoft Project

Chapter 6

Summary

Previous studies have highlighted the possibility of using keyboard typing data to detect emotions. Specifically, keystroke duration, keystroke latency, and accuracy rate of keyboard typing were used as input features for model building. These results have led to three hypothesized relationships. That are, the relationship between keystroke duration and emotion, the relationship between keystroke latency and emotion, and the relationship between accuracy rate of keyboard typing and emotion. Hence, the current study will test the three hypothesized relationships. Once implemented, the relationship between emotions and keystrokes will be found . Also, how much does every factor i.e. valence and arousal affect the key latency and duration will be known and the variance between them will be observed. The Influence of Emotion on Keystroke Patterns will be proved in this project. How much does every factor i.e. valence and arousal affect the key latency, duration and variance between them will be found.

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.