ENHANCING KEYSTROKE DYNAMIC BASED BIOMETRIC AUTHENTICATION FOR MOBILE DEVICES USING SENSOR DATA

Dissertation submitted in fulfilment of the requirements for the Degree of

MASTER OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

By

POTHINA YOGI SAI PRASANNA

11307087

Supervisor

BALJIT SINGH SAINI

School of Computer Science and Engineering

Lovely Professional University

Phagwara, Punjab (India)

November 2017

PAC Form

TOPIC APPROVAL PERFORMA

School of Computer Science and Engineering

Program : L252D::B. Tech (Hons.)-M-Tech(Dual Degree) - CSE

COURSE CODE : CSE548 **REGULAR/BACKLOG :** Regular **GROUP NUMBER :** CSERGD0031

Supervisor Name : Baljit Singh Saini **UID :** 22078 **Designation :** Assistant Professor

Qualification : _____ **Research Experience :** _____

SR.NO.	NAME OF STUDENT	REGISTRATION NO	BATCH	SECTION	CONTACT NUMBER
1	Pothina Yogi Sai Prasanna	11307087	2013	K1309	8194871536

SPECIALIZATION AREA : System Programming **Supervisor Signature:** _____

PROPOSED TOPIC : User Authentication using keystroke dynamics

Qualitative Assessment of Proposed Topic by PAC		
Sr.No.	Parameter	Rating (out of 10)
1	Project Novelty: Potential of the project to create new knowledge	7.00
2	Project Feasibility: Project can be timely carried out in-house with low-cost and available resources in the University by the students.	8.00
3	Project Academic Inputs: Project topic is relevant and makes extensive use of academic inputs in UG program and serves as a culminating effort for core study area of the degree program.	7.50
4	Project Supervision: Project supervisor's is technically competent to guide students, resolve any issues, and impart necessary skills.	8.00
5	Social Applicability: Project work intends to solve a practical problem.	7.75
6	Future Scope: Project has potential to become basis of future research work, publication or patent.	8.00

PAC Committee Members		
PAC Member 1 Name: Prateek Agrawal	UID: 13714	Recommended (Y/N): Yes
PAC Member 2 Name: Deepak Prashar	UID: 13897	Recommended (Y/N): Yes
PAC Member 3 Name: Raj Karan Singh	UID: 14307	Recommended (Y/N): NA
PAC Member 4 Name: Pushpendra Kumar Pateriya	UID: 14623	Recommended (Y/N): Yes
PAC Member 5 Name: Sawal Tandon	UID: 14770	Recommended (Y/N): NA
PAC Member 6 Name: Aditya Khamparia	UID: 17862	Recommended (Y/N): NA
PAC Member 7 Name: Anupinder Singh	UID: 19385	Recommended (Y/N): Yes
DAA Nominee Name: Kuldeep Kumar Kushwaha	UID: 17118	Recommended (Y/N): NA

Final Topic Approved by PAC: User Authentication using keystroke dynamics

Overall Remarks: Approved

PAC CHAIRPERSON Name: 10034::Amandeep Nagpal **Approval Date:** 04 Nov 2017

11/20/2017 4:01:22 PM

DECLARATION STATEMENT

I hereby declare that the research work reported in the dissertation/dissertation proposal entitled "ENHANCING KEYSTROKE DYNAMIC BASED BIOMETRIC AUTHENTICATION FOR MOBILE DEVICES USING SENSOR DATA" in partial fulfilment of the requirement for the award of Degree for Master of Technology in Computer Science and Engineering at Lovely Professional University, Phagwara, Punjab is an authentic work carried out under supervision of my research supervisor Mr./Mrs. BALJIT SINGH SAINI. I have not submitted this work elsewhere for any degree or diploma.

I understand that the work presented herewith is in direct compliance with Lovely Professional University's Policy on plagiarism, intellectual property rights, and highest standards of moral and ethical conduct. Therefore, to the best of my knowledge, the content of this dissertation represents authentic and honest research effort conducted, in its entirety, by me. I am fully responsible for the contents of my dissertation work.

Signature of Candidate

Pothina Yogi Sai Prasanna

R. No. 11307087

ABSTRACT

With the advent of artificial intelligence, the technology has transformed into smart and sophisticated technology. Since few decades, the simple password authentication has either replaced or compounded with biometrics (such as Facial Recognition, Fingerprint Scan etc.) to provide better security. Now, these onetime authentication systems are shifting towards a smart (i.e., continuous) authentication system. Keystroke Dynamics is behavioral biometrics that can perform continuous authentication to detect intruders. In this paper, the origin and various works carried out in Keystroke Dynamics is discussed along with a proposal to do new research.

Keywords: artificial intelligence, authentication, biometrics, intruder, keystroke, typing rhythms

SUPERVISOR'S CERTIFICATE

This is to certify that the work reported in the M.Tech Dissertation/dissertation proposal entitled "**ENHANCING KEYSTROKE DYNAMIC BASED BIOMETRIC AUTHENTICATION FOR MOBILE DEVICES USING SENSOR DATA"**, submitted by **Pothina Yogi Sai Prasanna** at **Lovely Professional University, Phagwara, India** is a bonafide record of his original work carried out under my supervision. This work has not been submitted elsewhere for any other degree.

Signature of Supervisor
(Baljit Singh Saini)
Date:

Counter Signed by:

1) **Concerned HOD:**
   HoD's Signature: _____
   HoD Name: _____
   Date:_____

2) **Neutral Examiners:**

External Examiner

Signature: _____

Name: _____

Affiliation: _____

Date: _____

Internal Examiner

Signature: _____

Name: _____

Date: _____

ACKNOWLEDGEMENT

I would like to convey my most heartfelt and sincere gratitude to my mentor Mr. Baljit Singh Saini of Lovely Professional University, for his valuable guidance and advice. His willingness to motivate me contributed tremendously to achieve the goal successfully. I would also like to my parents and friends who have always inspired and encouraged me to achieve my goal successfully.

Pothina Yogi Sai Prasanna

TABLE OF CONTENTS

CONTENTS **PAGE NO.**

TABLE OF CONTENTS

CONTENTS **PAGE NO.**

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Artificial Intelligence(AI) has changed the way how modern machines work. It is currently being used in all the major fields of work (such as Medicine, Business analytics, Digital Security etc.). AI facilitates a machine to learn on its own and update by itself. These cognitive abilities of the machine are now considered as key to many of the real-world problems. Data Security is considered one of the major problems of digital world. For example, On 3rd October 2017, Yahoo revealed that the August 2013, data breach of Yahoo has affected over 3 billion user accounts [1]. To prevent these data breeches, the digital resources (data, computing, network etc.) are protected from the intruders by means of two processes: authentication and authorization. Authentication is the process of recognizing and verifying the identity of claimed user [2]. Authorization is the process of granting access to a resource. Authentication is broadly classified into three types:

1. Knowledge based (e.g. password, 4-digit pin)
2. Ownership based (e.g. access card, token)
3. Biometric based
    a) Physiological (e.g. finger print, Iris)
    b) Behavioral (e.g. hand writing, typing rhythms, voice)

The above two or more methods are combined to achieve a strong authentication. Amongst the three, biometric based authentication is an excellent way to verify user's identity. Unlike, passwords, tokens or smart cards, biometrics couldn't be lost, stolen, or shoulder surfed [3]. Amid the Physiological and Behavioral, Physiological is static in nature, thereby, it could be exploited using the existing technology. Thus, Behavioral biometrics has a better potential to be used for user authentication. This type of authentication is achieved by using machine learning algorithms such as neural networks. This paper provides brief insight of a typing rhythm based behavioral biometrics, Keystroke Dynamics.

1.1 KEYSTROKE DYNAMICS

Keystroke Dynamics is behavioral biometrics that is based on user's typing patterns. Every user has unique typing pattern on computer keyboard like that of user's signature [4]. The origin of keystroke dynamics dates to World War II where the sender of the telegraph is identified by a

methodology known as the 'Fist of the Sender' [5]. It uses the rhythm, syncopation and pace of the telegraph keys. The following graph is plotted using the similar approach followed by P. H. Pisani and A. C. Lorena [6]. The total number of published papers that are titled with the words "Keystroke Dynamics" in the following reputed publications: ACM Digital Library, IEEE Xplore, Science Direct and Springer (are searched using their respectve website's advanced search option that facilitates to filter the research papers based on title and year) is ploted chronologically from the year 1990 to 2016. Figure 1.1 shows that its popularity has increased a lot over the span of 26 years. There is a greater increase in published research papers from the year 2014 to 2015.

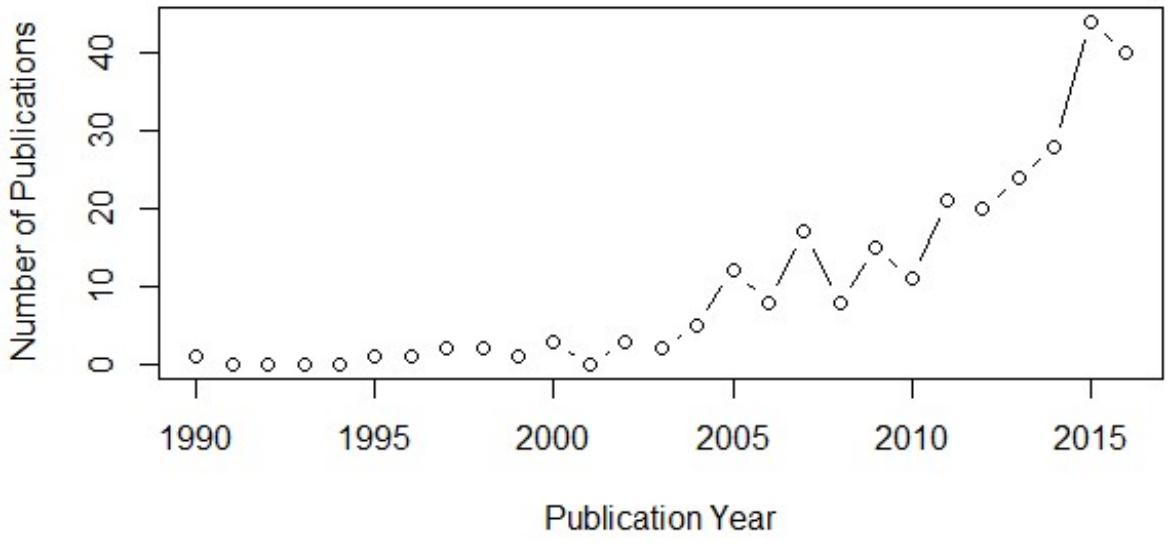

Figure 1.1 Number of published research papers from 1990-2016.

The researchers' effort has made enhanced commercial software possible. E.g. Biopassword™ [7], This uses a technique called *password hardening* where the user's typing rhythm is checked along with the password to verify the user.

Keystroke dynamics has the following advantages over the other authentication systems:

- It doesn't need an extra hardware. (i.e., cost effective)
- It is highly non-intrusive.
- The user doesn't need to put any extra effort.
- Continuous authentication can be done throughout the session.
- It can be used in an intrusion detection system

A typical keystroke dynamics system consists of two phases with four modules:

I. Enrollment Phase: In this phase, the user's data is recorded to extract features that are unique to the user and then store them in the database.

II. Authentication Phase: Here, the user's input is validated with the saved data in the database.

Figure 1.2 shows the biometric system. The four modules are explained in the later sections.

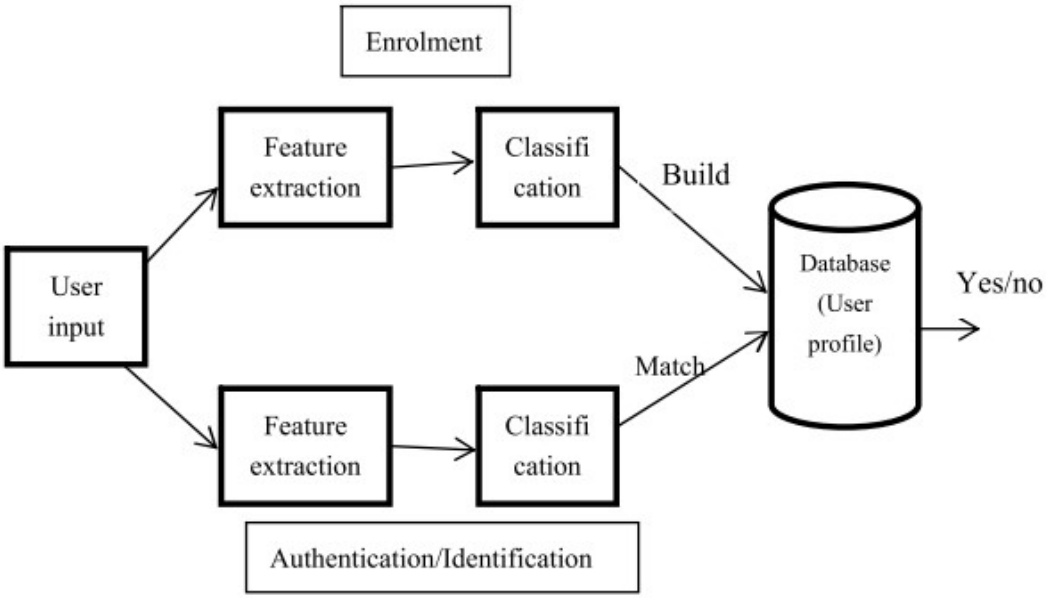

Figure 1.2 Biometric System [8]

1.2 USER INPUT

The input text length, clock precision and training data influence the performance of keystroke dynamics system [4]. The user input can be either static or dynamic. Static input consists a fixed structured text (transcription) that should be entered by user. Whereas in dynamic input the user can either have a fixed text or enter random text (free composition) in every input, i.e. the user may enter different text in both the phases. The former type of input is used for entry level authentication purpose i.e. at the login time only [2]. The dynamic can be used for both, entry level authentication as well as continuous authentication (free composition). According to an earlier study [6], there is no significant difference between the two methods, transcription and free composition. Spillane [8] was first to suggest the use of Computer Keyboards for capturing user's

keystroke dynamics. The raw data of the input consists of the key and its press and release timestamps. If mistakes are not allowed during acquisition of user password, Giot et al. [9] found that a huge number of failures occur. Figure 1.3 shows the graph of correct and erroneous acquisitions.

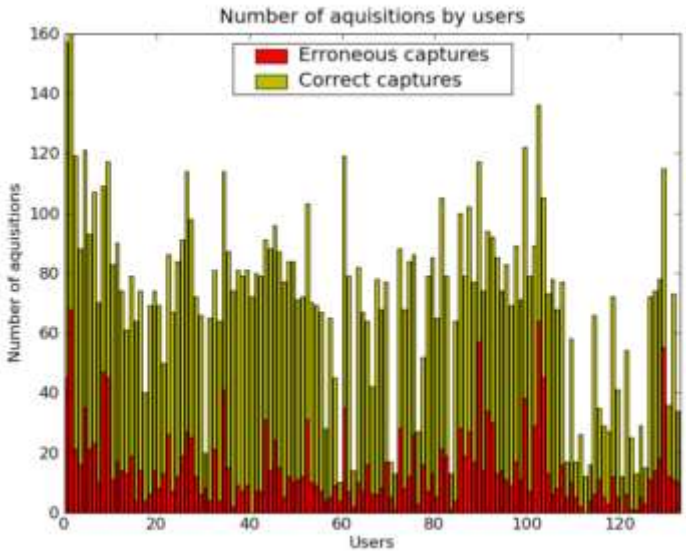

Figure 1.3 Number of correct and erroneous acquisitions for each user [10]

These erroneous captures are due to several reasons:

- Long length of the password (17 characters)
- Difference in user keyboard. i.e. the user is not accustomed with the keyboard used for getting the input.
- User's state of mind. i.e. forgetting password, disturbed by the environment etc.

When the keystroke latencies were paired with key pressure, the error rate has decreased from 12.2 to 6.9 in [10]. Thereby pressure can also be taken during the user input for better authentication. To measure pressure, special keyboards which are generally expensive than the standard keyboard are required. In general, desktop computers, laptops and smartphones are not pressure sensitive. Therefore, majority of the researchers limit their work to standard keyboard only. Notably, the environment plays a vital role in influencing an individual's typing behavior. Depending on the computer's operating system, resolution of data captured, keyboard dimensions, device, posture, lighting condition, the user's typing rhythm may be affected. For example, In MS Windows, the key press and key release keyboard events could not be

distinguished if their latency is below 15.625ms [6]. An environment where all the parameters that influence an individual's typing rhythm are relatively neutral to all the users is known as controlled environment. Otherwise it is known as uncontrolled environment. Uncontrolled environment provides a realistic scenario to test and deploy the system, the first experiment was performed by Monrose and Rubin [3]. They mentioned that the separate data recorded at different time is more reliable. Controlled environment can be used for gaining knowledge by comparing various approaches and algorithms used in making the system. To facilitate peer review, publishing the collected dataset and specifying the environment conditions (such as OS and applications used, computer's configurations, its work load etc.) would yield comparable results.

Application of keystroke dynamics in smart phones has some advantages over hard keyboards used in Personal Computers [11].

- Comparatively, less number of fingers are used with soft keyboard of smartphones.
- The position has no considerable impact on the user's typing rhythm as the smartphone would be handled in the same manner in different postures.

Due to the high availability of movement sensors (such as gyroscope, air pressure, accelerometer etc.) that had been proven effective for authentication [12], sensor enhanced keystroke dynamics came into existence. In this, the sensor data must be collected along with traditional keystroke rhythms.

1.3 FEATURE EXTRACTION

Ever since the debut of the term, Keystroke Dynamics, latency has been the most commonly used feature by early researchers [13]. The basic types of latencies are:

1) Press to Press (PP) or Digraph latency [13], [14]: It is the time elapsed between two successive key presses.
2) Press to Release (PR) or dwell/hold time [15]: It is the time delay between a key press and its release.
3) Release to Press (RP) or Flight Time [13], [14]: It is the latency between release of a key and press of next key.
4) Release to Release (RR): It is time duration between two consecutive key releases.

The first feasibility study to use the above features was conducted by Gaines et al. [16] at Rand Corporation.

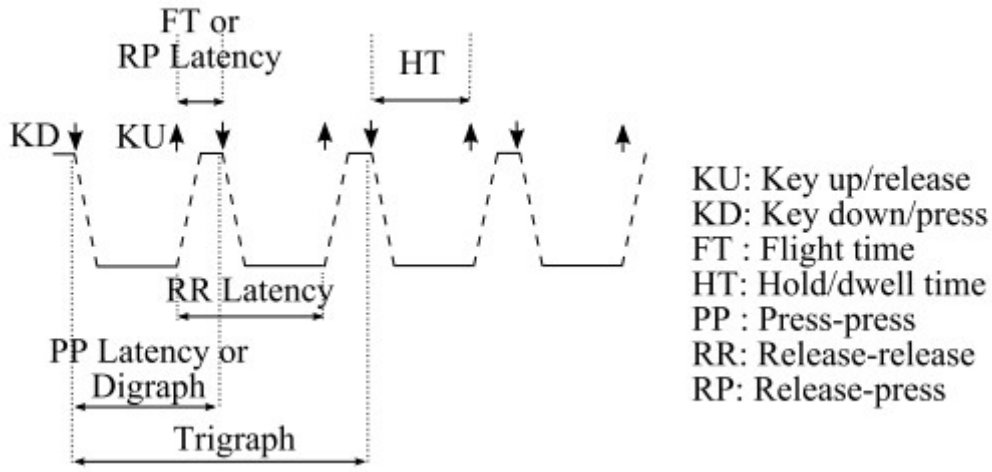

Figure 1.4Most common features extracted in keystroke dynamics [13]

The following latencies are considered when more than two keys are used for feature extraction:

a) Tri-Graph latency: It is like digraph latency, however here we use three successive keystrokes to derive the metrics.
b) N-Graph latency: It is computed by using more than three successive keystrokes. It is popularly known as elapse time between a key and n^{th} key of a string.

Using these features either individually or in combination, sub-features (like typing speed, key pressure etc.) can be derived [17]. It is observed that trigraph latency gives better classification results than digraphs and higher ordered n-graphs [13]. Robinson [18], found that the hold times are very important than inter key timings.

1.4 PERFORMANCE EVALUATION

The performance of keystroke dynamics system is measured by using the following three basic types of errors:

- False Acceptance Rate (FAR): It is the percentage of erroneously accepted imposters (invalid users) as legitimate user.

$$\text{FAR} = \frac{\text{Number of incorrect acceptances}}{\text{Total number of attempts made by intruder}} * 100$$

- False Rejection Rate (FRR): It is the percentage of legitimate users who are incorrectly categorized as imposters.

$$\text{FRR} = \frac{\text{Number of wrong rejections}}{\text{Total number of attempts made by legitimate user}} * 100$$

- Equal Error Rate (EER) or Cross Error Rate (CER): It is the value at which FAR and FRR are equal. Figure 1.5 shows the relation between FAR, FRR and ERR.

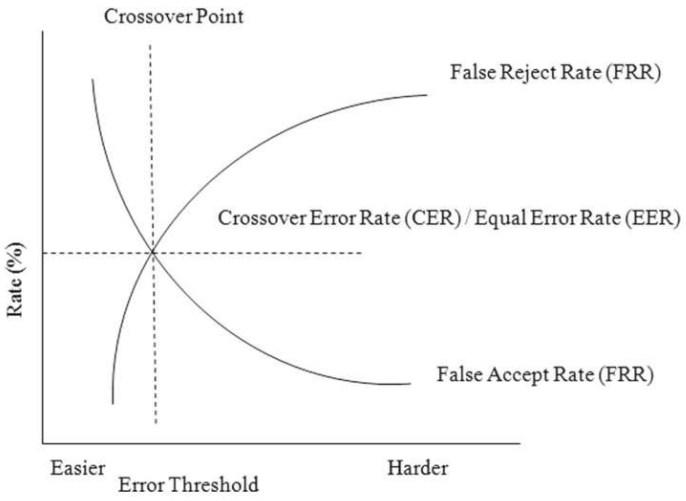

Figure 1.5 Relationship between FAR, FRR and ERR [19]

The lesser these values are, the better is the performance of the system. For e.g. Sally et al. [19] proved that for PIN's whose length falls between the four to twelve-digit range, the FRR was 0 and FAR was below 20%.

1.5 CLASSIFICATION

Classification aims at finding the best category/group which is closest to a classified pattern. The extracted features of the user in the enrollment phase are used to create template of the user behavior (profile). During the authentication phase, the extracted features are compared with the existing user template to find the best match to determine if the user is legitimate user or imposter. There are four major classification methods:

1.5.1 STATISTICAL METHODS

In the earlier times, these were the most popularly used method by majority of the researchers. To build templates, simple statistical methods like mean and standard deviation were used. For comparison of hypothesis, t-test or distance measures like Euclidean or Manhattan distance were used. Table 1.1 shows various statistical methods used by various researchers along with their results.

Table 1.1: Statistical Methods with FAR and FRR values

Author	Year	Algorithm	Features	E	Text Type	No. of subjects	Samples	FAR	FRR
Gaines el. al [16]	1980	T-test	1	C	S	7	36	0	4
Umphress el. al [20]	1985	Mean and Standard Deviation latency	1	C	S	17	34	6	12
Leggett el. al [21]	1988	keystroke Digraph Latencies	1	C	S, D	17	72	5	5.5
Bleha S. et al [22]	1990	Bayes Classifier	1	C	S	32	171	2.8	8.1
Joyce & Gupta [23]	1990	absolute distances	1	C	S	33	975	0.25	16.36
D. Gunetti el. al [24]	2005	Distance Measure	[1,2]	U	D	31	-	≈0	≈ 2.0
S. Giroux el. al [25]	2009	Mean	[1,3]	U	S	11	880	0	-

E- Environment, 1 - Latency, 2 - Trigraph/N-graph, 3 - Key hold time, S - Static, D - Dynamic, C - Controlled, U - Uncontrolled.

1.5.2 ARTIFICIAL NEURAL NETWORKS

Artificial Neural Networks (also referred as Neural Networks) are adaptive non-linear statistical learning algorithms that are inspired from biological neural networks. The type of leaning can be either supervised or unsupervised. The backpropagation algorithm is most popular method in supervised learning. Table 1.2 shows various Neural Network algorithms used by authors along with their performance measure:

Table 1.2: Neural Network approaches

Author	Algorithm	E	Text Type	No. of subjects	Samples	FAR	FRR
Brown and Rogers [26]	Geometric Distance Approach	C	S	46	1867	14.9	0
	Kohenen Neural Network					17.4	0
	back propagation Neural Network					12	0
Bleha & Obaidat [27]	Perceptron Algorithm	C	S	24	1400	8	9
M. S. Obaidat and B. Sadoun [28]	Potential Function	C	S	30	6750	2.2	4.7
						0	0
John A. Robinson et al [18]	Minimum Intra Class Distance Classifier	-	-	137	-	23	24
	Non-Linear Classifier			138		31	31
	Inductive Learning Classifier			139		10	62
D.-T. Lin [29]	Backpropagation	C	S	151	-	1.11	0
J. Bechtel et al [30]	ART-2	U	S	50	-	-	~10

E- Environment, S - Static, D - Dynamic, C - Controlled, U - Uncontrolled.

1.5.3 PATTERN RECOGNITION

Pattern recognition aims to classify patterns into different classes. Several pattern recognition techniques had been proposed and used by various researchers. It contains algorithms such as the Fishers linear discriminant (FLD), Bayes classifier and support vector machine (SVM). Table 1.3 displays various pattern recognition algorithms that are being used by researchers.

Table 1.3: Pattern Recognition Algorithms and their results

Author	F	Algorithm	E	Text Type	#	S	FAR	FRR
Robinson et al. [18]	1,2	Minimum intra-class distance, non-linear, Inductive learning	D	U	20	-	9	10
Haider et al. [31]	1	Fuzzy logic, Neural Network & statistical techniques	S	C	-	-	6	2
Gutirrez et al. [32]	1	Naive Bayesian, Statistical Decision Trees, Instance based	S	C	-	-	1.6	14.3
Kacholia & Pandit [33]	1	Random dist. function	S	-	20	440	3.4	2.9
Arau´jo et al. [34]	1,2	Fuzzy Logic	S	C	10	200	3.4	2.9
Eltahir et al. [35]	1,3	Autoregressive	S	C	22	22	5	-
Sang et al. [36]	1	Support Vector Machine	S	-	10	-	0.02	0.1
Joshi & Phoha [37]	1	Competition between Self Organizing Maps for Authentication	S	C	43	873	0.88	3.55
Chang [38]	1,2	Hidden Markov Models	S	C	20	600	0	-
Sheng et al. [39]	1,2	Parallel decision trees	S	C	43	387	0.88	9.62
Bartlow & Cukic [40]	1,2	Random Forest	S	U	41	8775	3.2	5.5
Hosseinzadeh et al. [41]	1,2	Gaussian Mixture Models	S	U	8	80	2.1	2.4

E- Environment, F- Features, #- No. of subjects, S- Samples, 1 - Latency, 2 - Key hold time, 3 - Key Pressure, S - Static, D - Dynamic, C - Controlled, U – Uncontrolled.

1.5.4 OTHER TECHNIQUES

The optimizing techniques like Evolutionary algorithms (e.g. genetic algorithm) and hybrid algorithms fall into this category. Ant colony optimization and Extreme machine learning achieved a feature reduction of 46.51% compared to Particle swarm optimization and Genetic Algorithm

[42]. To analyze hold time and latency, the concept of artificial rhythm with cues was used by Hwang et al. [43]and obtained EER of 4%. In free text(continuous) authentication, unstable (noisy/ gibberish) keystrokes are generated from user activities like playing game etc. It was found that this noisy data increases FRR. Filtering this type of data using spelling checker yielded better FRR instead of using regular expressions [44]. With Ant Colony Propagation, Marcus Karnan et al. [45] achieved classification error of 0.059 and Accuracy of 92.8. Though the hybrid techniques yield good results, yet they consume more system resources for computation. If the complexity of the combined algorithm is exponential or factorial in nature then a standalone application for system with limited resources (e.g. mobiles, laptops etc.) wouldn't possible.

1.6 BENCHMARK DATASETS

Keystroke data collection is quite time consuming. To ease the researchers across the world to focus on performance optimization (i.e. on algorithmic point of view) by saving their time from being spent on data collection, benchmark datasets were introduced. Table 1.4 shows few such datasets.

Table 1.4 Table of Benchmark Datasets

Dataset	No. of Users	Description	URL
CMU [46]	51	Input: ".tie5Roanl"	[47]
CMU-2 [48]	20	Contains both free text input and fixed text input	[49]
GREYC [9]	100	Input: greyc laboratory Duration: 2 months	[50]
Web-GREYC [51]	118	Input: user's own password Duration: 1 year	[52]
Queen Mary University [53]	100	Input: try4-mbs Additional Data: Pressure exerted	
Pressure sensitive [54]	104	3 Input strings: "pr7q1z","jeffrey allen" and "drizzle" Additional Data: Pressure exerted	[55]
Si6 Labs [56]	63	Language: Spanish	

		Input: 20 long sentences	
Beihang University [57]	117	This is a fixed Input(transcription) dataset	* [58] [59]
University of Torino [60]	165	Free text Input: 700-900 characters Language: Italian Duration: 6 months	*
Clarkson University [61]	39	Mixed Input (free text and fixed text) Additional Data: Facial expression videos and hand movements	
TapDynamics [62]	55	Device: mobile Additional Data: accelerometer readings	[63]
BioChaves [64]	47	4 Input strings: “chocolate”, “zebra”, “banana” and “taxi”	[65]
Graphical Password [10]	100	Device: 2 touch screen mobiles Additional Data: pressure	[66]
GREYC-NISLAB [67]	110	Additional Data: Soft Biometrics Data	[68]
Kevin Killourhy and Roy Maxion [69]	51	Input: user’s own password	[70]

* request to author needed

1.7 PATENTS

- J. Garcia. [71] Personal identification apparatus.
- J.R. Young and R.W. Hammon. [72]. Method and apparatus for verifying an individual's identity.
- Brown ME, Rogers SJ. [73] Method and apparatus for verification of a computer user’s identification based on keystroke characteristics.
- Blonder, G.E. [74]. Graphical Passwords.
- Cho SZ, Han DH. [75] Apparatus for authenticating an individual based on a typing pattern by using neural network system.
- Gordon Ross [76]Intellectual property protection and verification utilizing keystroke dynamics.

- Steven S. Bender, Howard J. Postley [77] Key sequence rhythm recognition system and method.
- S. Blender and H. Postley. [78]. Key sequence rhythm recognition system and method
- P. Nordström, J. Johansson. [79] Security system and method for detecting intrusion in a computerized system.
- Awad and I. Traore [80] System and method for determining a computer user profile from a motion-based input device.
- Vir V. Phoha, Shashi Phoha, Asok Ray, Shrijit Sudhakar Joshi, Sampath Kumar Vuyyuru [81]Hidden markov model ("HMM")-based user authentication using keystroke dynamics.
- DA SILVA FERREIRA João Pedro SOARES [82] System and method for intrusion detection through keystroke dynamics.

CHAPTER 2: LITERATURE REVIEW

In this section, the referred research papers during the dissertation have been listed along with brief description about them:

"KEYSTROKE DYNAMICS AS A BIOMETRIC FOR AUTHENTICATION" (Monrose and Rubin 2000) [3]

This paper presents the working of various algorithms used in keystroke dynamics. The authors built a C++ toolkit by using *xview* library for analyzing the data. This toolkit facilitates quick way to create rough properties on the data set by dividing the users in distinct groups.

They have explained the formulae used in the following algorithms: Euclidean distance measure, Weighted probability measure, Non-weighted probability, Bayesian-like classifier.They achieved best identification rate 87.18% (approximately) using the weighted probabilistic classifier on a dataset of 63 users. The identification rated obtained by using Euclidean distance and the non-weighted scoring approach are 83.22% and 85.63% respectively.

They favor the using structured text (transcription) instead of arbitrary text (i.e., free composition).

"KEYSTROKE DYNAMICS ON A MOBILE HANDSET: A FEASIBILITY STUDY" (Clarke et al. 2003) [83]

This paper investigates the ability of neural networks to successfully authenticate users, according to their interactions with a mobile phone's keypad. Using PIN code, they got 18.1% FAR, 12.5% FRR and 15% ERR. They suggested that Keystroke dynamics can become a part of hybrid authentication system.

"ARE DIGRAPHS GOOD FOR FREE-TEXT KEYSTROKE DYNAMICS?"(Sim and Janakiraman 2007) [84]:

The researchers found that word specific digraphs are good for free text keystroke dynamics. They have also proved that typing behavior of the user is dependent on context (E.g. the user's typing of the word "IN" as a whole word is different from "IN" in "BEGIN"). Of the top ten most frequently used words, they obtained perfect classification with the word "IN". Compared to non-word specific n- graphs, word specific n-graphs perform better. The authors defined a finger set to identify trained users and non-trained users.

"COMPARING ANOMALY DETECTORS FOR KEYSTROKE DYNAMICS" (Killourhy and Maxion 2009) [46]:

In this paper, 14 detectors(algorithms) were compared on a dataset that was collected from 51 users with 400 entries of the passphrase ".tie5Roanl" in 8 sessions(50 entries per session). They used Windows XP operating system with an external reference clock that is accurate upto ±200 micro seconds. They achieved best performance (least EER) with Manhattan distance.

"KEYSTROKE DYNAMICS-BASED AUTHENTICATION FOR MOBILE DEVICES" (Hwang, Cho, and Park 2009) [43]

To overcome the problems of shorter PIN length (e.g. 3 characters long), they have used artificial rhythms and tempo cues, by using these they obtained better performance with reduction of error(EER) from 13% to 4%. They suggested to apply the analysis to more diverse group of users. They mentioned that in future one amongst the FAR and FRR would be important and the issue could be addressed by selecting proper threshold.

They have also shown following accepted hypothesis:

a) Artificial Rhythms with cues are useful for users using both the hands for typing.
b) The average error rate(EER) involving Artificial Rhythms with cues is lower than Natural Rhythm without cues.

"GREYC KEYSTROKE: A BENCHMARK FOR KEYSTROKE DYNAMICS BIOMETRIC SYSTEMS" (Giot, El-Abed and Rosenberger 2009) [9]

This paper focuses on benchmarking datasets to help researchers directly work on their proposed algorithm instead of spending time for data collection. They developed a software "GREYC-Keystroke" for data collection. It can be download from [50]. The software can be used to share the collected data with the world. They observed that there were huge number of failures during static data acquisitions. The reasons were already cited earlier, in the report. Their database is available for free and can be used to compare the efficiency of the methods used for identification in keystroke dynamics.

"STUDY ON THE BEIHANG KEYSTROKE DYNAMICS DATABASE" (Li, Zhang and Cao 2011) [57]

The BeiHang database is like GREYC Keystroke database, which is open for public. But to use the existing database, a request should be sent to the corresponding author. It is available through the two Chinese websites: [58] and [59].

They used Gaussian, Neural network classifier and variants of SVM algorithm to compare their performance against 2 variants of databases. They obtained best results with SVM expansion reduction method with an ERR of 11.8327 and 26.6014 for database A and B respectively.

"BIOMETRIC PERSONAL AUTHENTICATION USING KEYSTROKE DYNAMICS: A REVIEW" (Karnan, Akila and Krishnaraj 2011) [17]

This paper summarizes the well-known approaches (statistical, neural network etc.) used in keystroke dynamics. They mentioned that user acceptance is one of the major advantage of keystroke dynamics, as users are already accustomed with typing. They found that combined use of the features key duration, latency or digraph give low FAR and IPR (Imposter pass rate) value. The author states that further research needs to be carried out to reduce the false alarms (FAR and IPR).

"KEYSTROKE DYNAMICS FOR BIOMETRIC AUTHENTICATION - A SURVEY" (Bhatt and Santhanam 2013) [2]:

This paper explains the 2013's trends of keystroke dynamics. They explained components and working of biometric system. This paper cites the research done by various researchers. This paper mentions that keystroke dynamics can be used for finding, if person's influence of alcohol or not. They cited J.R.Young et al. for using the term keystroke dynamics for the first time. They mention that appropriate length of password needs to be determined, to authenticate the user easily. Also, the template stored to authenticate the user needs to be able to adopt themselves with change in behavioral nature of the user.

"FUZZY MODEL IN HUMAN EMOTIONS RECOGNITION" (Bakhtiyari and Husain 2014) [85]

This paper deals with recognition of multi-level human emotions using a fuzzy model. They recognize the human emotions in five levels. Unlike other researchers they consider the region of the users also. The author shares the tabular data that is used as reference for determine the human emotions. The system was trained and tested by using SVM (Support Vector Machine). They obtained false positive rate of 16.7%. This area of research can be further extended to obtained higher accuracy.

“KEYSTROKE DYNAMICS FOR MOBILE PHONES: A SURVEY” (Saini, Kaur and Bhatia 2016) [11]

The researchers presented a comprehensive survey of research done in keystroke dynamics by representing various approaches (statistical, neural network etc.) in tabular format. Unlike most researchers, they have cited the work done related to both devices, mobile as well as personal computers. The advantage mobile device offers in this field of research are already cited earlier in this paper. They mentioned that the research can be done to find the ideal: password length, password type and minimum samples required for keystroke dynamics.

“AGE DETECTION THROUGH KEYSTROKE DYNAMICS FROM USER AUTHENTICATION FAILURES” (Tsimperidis, Rostami and Katos 2017) [86]:

The researchers used cross validation to get independence data and classifier parameters. The data was split into K folds of equal size. These folds are used for testing the system. The user’s age is categorized into the following 4 classes: 18-25, 26-35, 36-45, 46+. The researchers obtained a success rate of 56.7%, 68.6%, 60.2% and 72.7% respectively. The researchers obtained over all best result of 73.3% success rate when they have considered only two classes (18-35 and 36+). This research can be further explored to obtain better performance.

“EFFECTS OF TEXT FILTERING ON AUTHENTICATION PERFORMANCE OF KEYSTROKE BIOMETRICS” (Huang et al. 2017) [44]:

This research is based on free text keystroke dynamics. The authors introduced a new term called “gibberish” (noisy or unstable) text. This text is due to user activities such as playing games. In comparison to normal text keystrokes, often the keystrokes of gibberish text are typed fast. In their dataset, 23.3% of the keystrokes are of gibberish text. They investigated the impact of this gibberish text on authentication performance and found that it has no significant impact on FAR but worsens FRR by altering the range and mean of the distribution of the digraph latencies.

The removal of this noisy input has reduced the FRR (from 10.8 to 5.3). They filtered the gibberish text by using regular expressions and spell checker. Of these two methods, spell checker based filtering has given better performance but combining both methods gave best performance.

CHAPTER 3: PRESENT WORK

3.1 PROBLEM FORMULATION

Researchers across the world has done a lot of research in keystroke dynamics since 1980s but the technology is yet to achieve FAR of 0.001% and FRR of 1% to meet the European standard for commercial biometric [87]. Keystroke dynamics is based on the fact that every individual has unique typing behavior [3]. To authenticate a user successfully, the keystroke data should be consistent and discriminable. Artificial cues (such as Pauses and musical rhythms) had proven to be successful in achieving high discriminability [43]. With the hypothesis that user type their own password more consistently and uniquely than any regular text, I would like to investigate (research) if user's own password or artificially rhythmed text gives better authentication results.

The performance of the keystroke dynamic system was found to be enhanced when combined with sensory input such as pressure exerted on key [10]. Smart phones are already equipped with sensors such as accelerometer. To provide cost effective and better authentication performance, I would like to utilize the sensory input from the accelerometer and check its impact on both types of data (i.e., user's own password and artificially rhythmed text input).

Every computer is constrained with its resources. To achieve the best with limited resources is a desired and challenging goal. To obtain the best performance, I would like to apply an optimization technique and compare the optimization results of both datasets.

3.2 SCOPE OF STUDY

The current study is limited to static keystroke dynamics based authentication, but the results of the study could be used in dynamic keystroke dynamics to perform continuous authentication of the user more efficiently (for e.g. if user's own password gives better results then, in continuous evaluation, the user need not be evaluated for every text s/he types. Instead user's consistent texts (such as preposition "IN" etc.) are evaluated to reduce system's work load and increase overall evaluation speed/performance of the system. But necessary precautions such as evaluation of input after fixed number of keystrokes should be done to prevent infinite wait or starvation for consistent text input).

This study helps us answer whether a user need to remember s/he password or not. Also, this study would help us determine whether keystroke dynamics be used as a password recovery tool or not. Similarly, it would let us answer, whether keystroke dynamics can be used as an alternative for OTP or not.

Based on the obtained results tailored virtual keyboards can be designed to increase the input data quality, thereby increasing the performance of authentication.

3.3 OBJECTIVES OF THE STUDY

i. To compare the performance of Keystroke Dynamics system by using User's choice of password and fixed artificially rhythmed string of same length, and determine which amongst them is best suited for static authentication.
ii. To evaluate the individual and combined performance of authenticating user with accelerometer data and keystroke data for the above two types of datasets.
iii. To apply an optimization technique and enhance performance of the system.

3.4 RESEARCH METHADOLOGY

Step 1: Data collection

The sample data should be collected from at least 20 individuals (preferably equal number of male and female members of similar age) using an android application (that only accepts input without typographical errors) at random intervals with 25 entries for each type of data (i.e., 25 entries for user's own password and 25 entries for artificial string) per session. Only one session is taken per week. A minimum of 4 sessions must be taken from each user to obtain significant amount of data.

Step 2: Feature extraction

From the collected data, the hold time and flight time features would be extracted. Hold time is computed by calculating the difference between the time stamps of press and release of the key. The flight time is calculated as the difference between release timestamp of a key and immediate press timestamp of next key. From the sensor data, the change in magnitude along x, y and z axis are extracted.

Step 3: Classification

The extracted features from individual datasets are used for classifying users into separate groups. The feature data is used for analyzing in various aspects such as measuring discriminability by graphically plotting the overlapping rhythm of different users etc. A neural network based implementation would be used for training the system and then use it for user identification.

Step 4: Result and Analysis

From the obtained results, inferences would be made to generalize the results. (by answering the objectives)

3.5 EXPECTED OUTCOMES

Through the proposed research, the answers to the following questions are expected:

- Amongst user's own password and artificially rhythmed string based keystroke dynamics authentication, which one gives better performance?
- Impact of sensory data on the above two types of datasets (user's own password and artificially rhythmed string)?
- Which type of dataset is better optimized using an optimization technique?

The obtained results can be used to infer whether keystroke dynamics can be used as password-free authentication system or not.

CHAPTER 4: CONCLUSION

This report attempts to present an overview of research on keystroke dynamics over the past few decades along with the research proposal. Keystroke Dynamics is a behavioral biometric authentication technique that has great potential for becoming the standard authentication model in near future. It offers continuous authentication unlike the traditional login time authentication system. The boundaries for application of keystroke dynamics are not yet well defined. The following are a few applications of keystroke dynamics:

a) Detecting the age of a person. [86]
b) Detecting the mood of the person. [85]
c) Both, One time and continuous Authentication of User.
d) Determine if a person is under the influence of alcohol or not. [2]

Though there is significant increase in this field of research since 2002, yet the field has many unexplored parameters. The FAR, FRR and ERR of keystroke dynamics is usually high that it do not meet standards in some standard access control systems, such as the EN-50133-1 [46]. There are many parameters such as keyboard dimensions, operating system used, light conditions, human mental health/condition, the device being used etc., that influence the keystroke biometric system. The research work considering more parameters would yield better results. This is a promising field to carry out the research work. For e.g. the following areas can be explored:

I. The statistical influence of operating system on keystroke biometrics.
II. Estimate the percentage of alcohol consumed by a user using keystroke dynamics
III. Determine the human personality using key patterns of the user.
IV. Security standards to govern the use of keystroke dynamics as a web based service.
V. Determine the optimum length of the string to be used for keystroke based user identification system.

My study(research) focuses on static keystroke biometric authentication on mobile devices using accelerometer sensor. This study aims to find whether any artificially rhythmed string or user's own password be used for authentication users with keystroke dynamics. Also, the impact

on overall performance of keystroke dynamic system by using sensory data in combination with the keystroke data would be investigated for both the cases. The results of this research would help us questions such as whether keystroke dynamics be used as a password reset tool or not. This study would help the developers of keystroke dynamics based authentication system, in their designing phase to take proper decisions related to training the system (by using user's choice of password or system generated text).

REFERENCES

[1] A. D. Rushe, "Yahoo says all of its 3bn accounts were affected by 2013 hacking," 03 Oct 2017.

[2] S. Bhatt and T. Santhanam, "Keystroke dynamics for biometric authentication - A survey," in *Proceedings of the 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME)*, February 21-22, 2013.

[3] F. Monrose and A. D. Rubin, "Keystroke dynamics as a biometric for authentication," *Future Generation Computer Systems,* vol. 16, no. 4, p. 351–359, 2000.

[4] I. Bhardwaj, N. Londhe and S. Kopparapu, "Study of Imposter Attacks on Novel Fingerprint Dynamics Based Verification System," *IEEE Access,* p. 595–606, 2016.

[5] P. S. Rinky Solanki, "Estimation of the User's Emotional State by Keystroke Dynamics," *International Journal of Computer Applications,* vol. 94, no. 13, p. 21–23, 2014.

[6] P. H. Pisani and A. C. Lorena, "A systematic review on keystroke dynamics," *Journal of the Brazilian Computer Society,* vol. 19, no. 4, pp. 573-587, 2013.

[7] [Online]. Available: http://www.biopassword.com.

[8] S. R, "Keyboard Apparatus for Personal Identification.," *Technical Disclosure Bulletin 17, 3346, IBM,* 1975.

[9] R. Giot, M. El-Abed and C. Rosenberger, "GREYC keystroke: A benchmark for keystroke dynamics biometric systems," in *IEEE 3rd International Conference on Biometrics: Theory, Applications and Systems, BTAS 2009*, 2009.

[10] T.-Y. Chang, C.-J. Tsai and J.-H. Lin, "A graphical-based password keystroke dynamic authentication system for touch screen handheld mobile devices," *Journal of Systems and Software,* vol. 85, no. 5, pp. 1157-1165, 2012.

[11] B. S. Saini, N. Kaur and K. S. Bhatia, "Keystroke dynamics for mobile phones: A survey," *Indian Journal of Science and Technology,* vol. 9, no. 6, pp. 1-8, 2016.

[12] C. Giuffrida, K. Majdanik, M. Conti and H. Bos, "I sensed it was you: authenticating mobile users with sensor-enhanced keystroke dynamics," *In Proceedings of DIMVA,* 2014.

[13] S. P. Banerjee and D. L. Woodard, "Biometric Authentication and Identification Using Keystroke Dynamics: A Survey," *Journal of Pattern Recognition Research,* vol. 7, no. 1, pp. 116-139, 2012.

[14] R. Saifan, A. Salem, D. Zaidan and A. Swidan, "A Survey of behavioral authentication using keystroke dynamics: Touch screens and mobile devices," *Journal of Social Sciences,* pp. 29-41, 2016.

[15] P. Shinde, S. Shetty and M. Mehra, "Survey of Keystroke Dynamics as a Biometric for Static Authentication," *International Journal of Computer Science and Information Security,* pp. 203-207, 2016.

[16] R. S. Gaines, W. Lisowski, S. J. Press and N. Shapiro, "Authentication by Keystroke Timing: Some Preliminary Results," *Technical Report R-2526-NSF, Rand Corporation,* May 1980.

[17] M. Karnana, M. Akila and N. Krishnaraj, "Biometric personal authentication using keystroke dynamics: A review," *Applied Soft Computing Journal,* vol. 11, no. 2, pp. 1565-1573, 2011.

[18] J. A. Robinson, V. M. Liang, J. A. M. Chambers and C. L. MacKenzie, "Computer User Verification Using Login String Keystroke Dynamics," *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans,* vol. 28, no. 2, p. 236–241, Mar. 1998.

[19] S. D. Abualgasim and I. Osman, "An Application of the Keystroke Dynamics Biometric for Securing PINS and Passwords," *World of Computer Science and Information Technology Journal, vol. 1, no. 9,* pp. 398-404, 2011.

[20] G. Williams and D. Umphress, "Identity verification through keyboard characteristics," *International Journal of man Machine and Systems,vol. 23, no. 3,* pp. 263-273, 1985.

[21] L. J. and W. G, "Verifying Identity via Keystroke Characteristics," *International Journal of Man- Machine Studies, vol. 28, no. 1,* pp. 67-76, 1998.

[22] S. Bleha, C. Slivinsky and B. Hussien, "Computer Access Security Systems Using Keystroke Dynamics," *IEEE Transcations on Pattern Analysis and Machine Intelligence,* pp. 1217-1222, 1990.

[23] Gupta, R. Joyce and Gopal, "Identity Authorization Based on Keystroke Latencies," *Communications of the ACM,* pp. 168-176, 1990.

[24] D. Gunetti, C. Picardi and G. Ruffo, "Keystroke analysis of different languages: A case study," *Advances in Intelligent Data Analysis VI,* pp. 133-144, 2005.

[25] S. Giroux, R. Wachowiak-Smolikova and M. P. Wachowiak, "Keystroke-based authentica- tion by key press intervals as a complementary behavioral biometric," *IEEE International Conference on Systems, Man and Cybernetics,* p. 80–85, Oct. 2009.

[26] Rogers, M. Brown and S. Joe, "User Identification Via Keystroke Characteristics of Typed Names Using Neural Networks," *International Journal of Man-Machine Studies,* pp. 999-1014, 1993.

[27] B. SA and O. MS, "Computer Users verification using the Perceptron Algorithm," *IEEE Transactions on Systems, Man and Cyberetics,* pp. vol. 23, no. 3, 1993.

[28] M. S. Obaidat and B. Sadoun, "Verification of Computer Users Using Keystroke Dynamics," *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics,* p. 27(2):261–269, Apr. 1997.

[29] L. D.T., "Computer-access authentication with neural network based keystroke identity verification," *International Conference on Neural Networks,* pp. volume 1, pages 174 –178, June 1997.

[30] J. Bechtel, G. Serpen and M. Brown, "Passphrase Authentication Based on Typing Style Through an Art 2 Neural Network," *International Journal of Computational Intelligence and Applications,* p. 131–152, 2002.

[31] S. Haider, A. Abbas and A. K. Zaidi, "A multi-technique approach for user identification through keystroke dynamics," *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics 2,* p. 1336–1341, 2000.

[32] F. Gutirrez, M. Lerma-Rascn, L. Salgado-Garza and F. Cant, "Biometrics and Data Mining: Comparison of Data Mining-Based Keystroke Dynamics Methods for Identity Verificatio," in *Advances in Artificial Intelligence*, MICAI, 2002, p. 221–245.

[33] V. Kacholia and S. Pandit, "Biometric Authentication using Random distributions (BioART)," *Canadian IT Security Symposium,* 2003.

[34] L. F. Ara´ ujo, M. G. Liz'arraga, L. R. Sucupira, J. T. Yabu-uti and L. L. LEE, "Typing Biometrics User Authentication based on Fuzzy Logic," *IEEE Latin America Transactions,* vol. 2, no. 1, p. 69–74, march 2004.

[35] W. Eltahir, M. Salami, A. Ismail and W. Lai, "Dynamic keystroke analysis using AR model," *IEEE International Conference on Industrial Technology,* pp. volume 3, pages 1555 – 1560, Dec. 2004.

[36] Y. Sang, H. Shen and P. Fan, "Novel Impostors Detection in Keystroke Dynamics by Support Vector Machine," in *Parallel and Distributed Computing: Applications and Technologies, volume 3320*, 2005, p. 37–38.

[37] S. S. Joshi and V. V. Phoha, "Investigating hidden markov models capabilities in anomaly detection," *Proceedings of the 43rd annual Southeast regional conference - Volume 1,* p. 98–103, 2005.

[38] C. W, "Improving hidden Markov models with a similarity histogram for typing pattern biometrics," *Proceedings of the IEEE International Conference on Information Reuse and Integration (IRI '05),* p. 487–493, August 2005.

[39] Y. Sheng, V. Phoha and S. Rovnyak, "A parallel decision tree-based method for user authen- tication based on keystroke patterns," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 35(4),* p. 826 –833, Aug. 2005.

[40] N. Bartlow and B. Cukic, "Evaluating the Reliability of Credential Hardening through Keystroke Dynamics," *17th International Symposium on Software Reliability Engineering,* p. 117 –126, Nov. 2006.

[41] D. Hosseinzadeh, S. Krishnan and A. Khademi, "Keystroke Identification Based on Gaussian MixtureModels," *IEEE International Conference on Acoustics, Speech and Signal Processing Proceedings, volume 3,* p. III–1144 – III–1146, May 2006.

[42] D. Shanmugapriya and G. Padmavathi, "A survey of biometric keystroke dynamics: approaches, security and challenges," *International Journal of computer Science and Information Security, Vol. 5, No. 1,* pp. 115-119, 2009.

[43] S.-s. Hwang, S. Cho and S. P. Seoul, "Keystroke dynamics-based authentication for mobile devices," *Computers and Security,* vol. 28, no. 1-2, pp. 85-93, 2009.

[44] J. Huang, D. Hou, S. Schuckers and S. Upadhyaya, "Effects of text filtering on authentication performance of keystroke biometrics," in *8th IEEE International Workshop on Information Forensics and Security, WIFS 2016*, 2017.

[45] M. Karnan and M. Akila, "Personal Authentication Based on Keystroke Dynamics Using Ant Colony Optimization and Back Propagation Neural Network," *IJCNS, vol. 1, no. 2,* 2009.

[46] K. S. Killourhy and R. A. Maxion, "Comparing Anomaly-Detection Algorithms for Keystroke Dynamics," in *39th Annual International Conference on Dependable Systems and Networks (DSN-2009)*, Los Alamitos, 2009.

[47] [Online]. Available: http://www.cs.cmu.edu/keystroke/.

[48] K. KS and M. RA, "Free vs. transcribed text for keystroke-dynamics evaluations," *Proceedings of the 2012work- shop on learning from authoritative security experiment results, LASER '12, ACM, New York,* p. 1–8, 2012.

[49] [Online]. Available: http://www.cs.cmu.edu/ keystroke/laser-2012.

[50] [Online]. Available: http://www. ecole.ensicaen.fr/~rosenber/keystroke.html.

[51] R. Giot, M. El-Abed and C. Rosenberger, "Web-based benchmark for keystroke dynamics biometric systems: a statistical analysis," *Intelligent information hiding andmultimedia signal processing (IIH-MSP),* p. 11–15, 2012.

[52] [Online]. Available: http://www. epaymentbiometrics.ensicaen.fr/index.php/app/resources/ 84..

[53] C. Loy, C. Lim and W. Lai, "Pressure-based typing biometrics user authen- tication using the fuzzy ARTMAP neural network," *International Conference on Neural Information Processing (ICONIP),* 2005.

[54] J. Allen, "Ananalysis of pressure-based keystroke dynamics algorithms," Master's thesis,Southern Methodist University, Dallas.

[55] [Online]. Available: http://jdadesign.net/2010/04/pressure-sensitive-keystroke-dynamics-dataset/.

[56] L. Bello, M. Bertacchini, C. Benitez, J. Pizzoni and M. Cipriano, "Collection and publication of a fixed text keystroke dynamics dataset," in *In XVI Congreso Argentino de Ciencias de la Computación (CACIC)*, 2010.

[57] Y. Li, B. Zhang, Y. Cao, S. Zhao and Y. Gao, "Study on the BeiHang Keystroke Dynamics Database," 2011.

[58] [Online]. Available: http://www.microdone.cn/ballet_login. php.

[59] [Online]. Available: http://www.u1ge.com/help/passdancer.

[60] D. Gunetti and C. Picardi, "Keystroke analysis of free text," *ACM Transactions on Information and System Security,* vol. 8, no. 3, p. 312–347, 2005.

[61] E. Vural, J. Huang, D. Hou and S. Schuckers, "Shared research dataset to sup- port development of keystroke authentication," in *International Joint Conference on Biometrics (IJCB)*, 2014.

[62] G. Ho, "Tapdynamics: strengthening user authentication on mobile phones with keystroke dynamics," Technical report, Stanford University, 2014.

[63] [Online]. Available: https://github.com/grantho/TapDynamics/.

[64] M. J, A. C and F. E, "Equalization of key- stroke timing histograms for improved identification performance," *Telecommunications symposium, 2006 International,* p. 560– 565, 2006.

[65] [Online]. Available: http://www.biochaves.com/en/download. htm.

[66] [Online]. Available: http://ty.ncue.edu.tw/N27/data.html..

[67] S. Idrus, E. Cherrier, C. Rosenberger and P. Bours, "Soft biometrics database: a benchmark for keystroke dynamics biometric systems," 2013.

[68] [Online]. Available: http://www.epaymentbio metrics.ensicaen.fr/images/pdf/greyc-nislab%20keystroke%20benchmark%2 0dataset.xls..

[69] K. S. Killourhy and R. A. Maxion, "Comparing Anomaly Detectors for Keystroke Dynamics," in *Proceedings of the 39th Annual International Conference on Dependable Systems and Networks*, California, June 29-July 2, 2009.

[70] [Online]. Available: http://www.citefa.gov.ar/si6/kprofiler/dataset. .

[71] J. Garcia.U.S. Patent and Trademark Office Patent 4,621,334, November, 1986.

[72] J. Young and R. Hammon, "Method and apparatus for verifying an individual's identity". U. S. Patent and Trademark Office, Washington D.C. Patent 4,805,222, February 14, 1989.

[73] B. ME and R. SJ.U.S. Patent 5557686, Seprember 1996.

[74] B. G.E.United States Patent 5,559,961, 1996.

[75] C. SZ and H. DH.U.S. Patent 6151593, November 2000.

[76] G. Ross.U.S. Patent 20040034788, 2004.

[77] S. S. Bender and H. J. Postley.U.S. Patent 7206938, 2007.

[78] S. Blender and H. Postley.U.S. Patent and Trademark Office Patent 7,206,938, 2007.

[79] P. Nordström and J. Johansson.European Patent Office Patent 2,069,993, 2009.

[80] A. Awad and I. Traore.U.S. Patent and Trademark Office Patent 8,230,232, 2012.

[81] V. V. Phoha, S. Phoha, A. Ray, S. S. Joshi and S. K. Vuyyuru.U.S. Patent 8,136,154, 2012.

[82] D. S. F. J. P. SOARES.Patent WO2013006071 A1, Jan 10, 2013.

[83] N. L. Clarke, S. M. Furnell, B. M. Lines and P. L. Reynolds, "Keystroke dynamics on a mobile handset: a feasibility study," *Information Management & Computer Security,* vol. 11, no. 4, p. 161–166, 2003.

[84] T. Sim and R. Janakiraman, "Are digraphs good for free-text keystroke dynamics?," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2007.

[85] B. Kaveh and H. hafizah, "Fuzzy Model in Human Emotions Recognition," 2014.

[86] I. Tsimperidis, S. Rostami and V. Katos, "Age Detection Through Keystroke Dynamics from User Authentication Failures," *International Journal of Digital Crime and Forensics,* vol. 9, no. 1, pp. 1-16, 2017.

[87] A. R. Sonawane and H. V. Kumbhar, "Graphical-Based Password Keystroke Dynamic Authentication System for Android Phone," *International Journal of Advanced Research in Computer and Communication Engineering,* vol. 5, no. 12, pp. 377-381, 2016.

[88] K. K and M. R, "Why did my detector do that?! predicting keystroke-dynamics error rates," *Recent advances in intrusion detection, lecture notes in computer science, vol 6307. Springer, Berlin/Heidelberg,* p. 256–276, 2010.