BIOMETRICS ON KEY STROKE DYNAMICS LOGGING SYSTEM WITH VERIFICATION

Submitted in partial fulfillment of the requirements for the degree of

Bachelor of Technology

in

Computer Science and Engineering Spl. in Information Security

by

RITIK RAMAIYA 18BCI0264

Under the guidance of

Prof. / Dr.

Boominathan P

VIT, Vellore.



June,2022

DECLARATION

We hereby declare that the thesis entitled "biometrics on keystroke dynamics

logging system with verification" submitted by me, for the award of the degree of

Bachelor of Technology in Computer Science and Engineering to VIT is a record of

bonafide work carried out by me under the supervision of Dr. Boominathan P.

I further declare that the work reported in this thesis has not been submitted and

will not be submitted, either in part or in full, for the award of any other degree or

diploma in this institute or any other institute or university.

Place: Vellore

Date: 29/05/2022

Ritik Ramaiya

Signature of Candidate

2

CERTIFICATE

This is to certify that the thesis entitled "Biometrics on keystroke dynamics

logging system with verification" submitted by Ritik Ramaiya, 18BCI0264 of

SCOPE, VIT, for the award of the degree of *Bachelor of Technology in Computer*

Science and Engineering with specialization in information security, is a record of

bonafide work carried out by them under my supervision during the period, 15.01.2022

to 03.06.2022, as per the VIT code of academic and research ethics.

The contents of this report have not been submitted and will not be submitted

either in part or in full, for the award of any other degree or diploma in this institute or

any other institute or university. The thesis fulfills the requirements and regulations of

the University and in my opinion meets the necessary standards for submission.

Place: Vellore

Date: 29/06/2022

Signature of the guide

Internal Examiner

External Examiner

DR SATHYARAJ R

HEAD OF DEPARTMENT INFORMATION SECURITY

3

ACKNOWLEDGEMENTS

I would like to thank the school of computer science and engineering (SCOPE) for providing me the opportunity to be capable to carry out the project and learn a lot of new things while working on it besides my MS in computer science in USA. I would also like to thank my guide in VIT, Dr Boominathan P for supporting me since the day I started working on this project and remained calm throuhout my small mistakes I might have done and supporting me to the end.

Overall the journey of my undergraduate studies has been an overwhelming journey and I have learnt a lot of new things which will help me brighten up my future. This thesis has been a great opportunity to showcase my work and appreciate the people who have supported me since I started this thesis.

Ritik Ramaiya

Executive Summary

Everybody needs to authenticate himself on his computer before using it, or even before using different applications. in most cases usually it is always that we authenticate a user by taking in his login details like the username and password and sometimes the use of captcha is used to avoid robotic users from spamming the system from brute force attack. To be efficient and secure, the client should take on a severe administration of its certifications. As these conditions are quite strict and difficult to be applied for most users, they do not not respect them. We have seen that keystroke elements permits to get the verification interaction by confirming the approach to composing the accreditations. It can likewise be utilized to get the meeting after its opening by identifying the changing of composing conduct in the meeting. It is able to recognize if another individual uses the keyboard, because the way of interacting with it is different. Moreover, keystroke dynamics can also prevent the steal of data or non authorized computer use by attackers. In this thesis we will be discussing on how to eradicate the problem and try to minimize it as much as we can.

CONTENTS

		Page No	
Acknowl	i		
Executiv	re Summary	ii	
Table of	contents	iii	
List of F	igures	iv	
List of T	of Tables v		
Abbrevia	ations	i ii iii iii iv v vi ON 11 11 11 11 11 11 11 11 11 11 11 11 11	
1. INTR	ODUCTION	11	
1.1.	Objective	11	
1.2.	Motivation	11	
1.3.	Background	7	
2. PROJ	JECT DESCRIPTION AND GOALS	13	
3. TECI	HNICAL SPECIFICATIONS	15	
4. DESI	GN APPROACH AND DETAILS	21	
4.1	design approach / materials and Methods	21	
4.2	codes and standards	23	
4.3	Constraints, Alternatives and tradeoffs	24	
5. SCHI	EDULE, TASKS AND MILESTONES	29	
6. PROJ	JECT DEMONSTRATION	31	
7. COS	7. COST ANALYSIS / RESULT & DISCUSSION 35		
8. SUM	39		

List of Figures

Figure No.	Figure No. Title		
1	Three authentication factors	10	
1.1	Classification of user authentication approaches	12	
2	Traits of biometric system	14	
4	Architectural diagram of system	23	
4.1	Crossoverrate graph	26	
6.1	user interface for creating new user	31	
6.2	Registration details for new user	31	
6.3	Architecture of ResNet	32	
6.4	Capturing key release times	32	
6.5	Existing user login	33	
6.6	Verification of existing user	33	
6.7	Pop up message generation	34	
6.8	6.8 Pressing and Releasing time generated		
6.9	Signature profile for different Users	35	
7.1	large population EER graph plot		
7.2	Small population EER graph plot	37	
	List of Tables		
Table No	Title	Page No.	
Table 5.1	Schedule and Task	Schedule and Task 29	
Table 5.2 Gantt Chart		30	

List of abbreviatons

SD- standard deviation

LM – Latency mean

LMA – latency mean average

KPT/KRT - Key-press/key-release timestamps

KHT (KDT) - Key-hold (key-dwell) time, lapse passing from the press and release of a key

PRT - Key press-release time, lapse passing from the press to the release of different keys

PPT - Key press-press time, lapse passing between the press events of different keys

RPT (**KFT**) - Key release-press (key-flight) time, lapse passing from the release to the press of different keys

RRT- Key release-release time, lapse passing between the release events of different keys

WPM - Word per minute, a measure of typing speed

CPS- Characters per second, typically excluding backspaces

AdjWPM - Adjusted word per minute, modified version of WPM considering errors

KPS - Keystrokes per second, similar to CPS (including backspace)

1. INTRODUCTION

1.1. OBJECTIVE

In this project we have designed a biometric system which takes biometric as the trait it takes each person types on a keyboard in a characteristic way. In this project biometric is not expected to be unique to each individual but it may be expected to offer sufficient discriminatory information to permit identity verification. The keystrokes of an individual could be observed subtly as that individual is entering in data. This biometric grants "persistent check" of a singular's character over a meeting after the individual logs in utilizing a more grounded biometric like unique mark or iris so go about as dynamic biometric framework. so we have used digraph to catch the words pressing and releasing time of each word is captured this process is repeated for n trials which provides us the train data and then when a different user tries to enter the system he would not be allowed to enter in the system as he would be having a different keystroke pattern. The variable is average lantency, standard. deviation with the trained data. We have also made an interface for the user to interact with the system there we are going to store the signature profile in the database where the data trained we have latency average average and much more information suing which we are going to calculate the standard. deviation which serves as important factor in threshold calculation.

Authentication Factor	Description	Examples
Knowledge Factors	Something a user knows	username &password,
		passphrase, PIN
Possession Factors	Something a user physically	hardware keys, ID card, smart
	possesses	phone software
Biological Factors	Something inherent to a user	iris, fingerprint, and behavioral
	something	characteristics: gait, typing

Fig 1. Three authentication factors

1.2. Motivation

We have seen that keystroke dynamics permits to get the authentication cycle by checking the approach to composing the qualifications. we can likewise use to make the framework more hearty by looking at the progressions in the squeezing and composing variety from the client. For this situation, we discuss constant authentication, the PC knows how the client connects with its console. It can perceive if another singular purposes the console, on the grounds that the approach to cooperating with it is unique. Furthermore, keystroke dynamics can in like manner thwart the take of data or non supported PC use by aggressors.

1.3. Background

The conduct biometric of Keystroke Dynamics involves the way and mood wherein a singular kinds characters on a console or keypad. The keystroke rhythms of a client are estimated to foster a novel biometric layout of the client's composing design for future authentication. Keystrokes are isolated into static and dynamic composing, which are utilized to assist with recognizing approved and unapproved clients. Vibration data might be utilized to make an example for later use in both ID and verification undertakings.

Biometrics advancements are acquiring fame because of the explanation that when utilized related to conventional strategies for confirmation they give an additional a degree of safety. Biometrics includes something an individual is or does.Information expected to break down keystroke elements is acquired by keystroke logging. Regularly, all that is held while logging a composing meeting is the grouping of characters comparing to the request in which keys were squeezed. It is disposed of to Time data. For instance, while understanding email, the collector can't tell from perusing the expression "I saw 3 zebras!" whether:

- (a)it was composed quickly or gradually,
- (b)the source utilized the left shift key, the right shift key, or the covers lock key to make the "I" transform into an uppercase letter "I,"
- (c) the letters were totally composed at a similar speed, or on the other hand on the off chance that there was a long respite before any characters while searching for that key.

(d) the source composed any letters wrong at first and afterward returned and adjusted them, or on the other hand assuming they got them right the initial time.

Huge investigation into the practicality of keystroke dynamics as a potential biometric verification technique has occurred since the appearance of PCs. The examinations have advanced from looking at composing designs on work area keyboards utilizing factual example classifiers to portable keyboards utilizing brain networks as an example classifier. The investigations don't have a bringing together technique for contrasting outcomes, which limits correlation between the strategies introduced. Without the capacity to look at concentrates inside research regions, future review is restricted in its capacity to give significant alterations to the work being referred to. This paper surveys a delegate subset of the flow research in keystroke dynamics, and gives proposals on the likely heading of future work around here. This will give a bunch of rules that can be trailed by scientists proposing to accomplish further work in the space of keystroke dynamics.

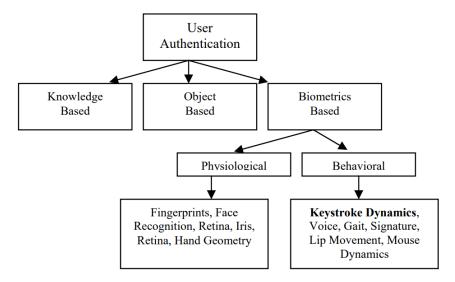


Fig 1.1. Classification of User authentication Approaches

2. PROJECT DESCRIPTION AND GOALS

In this project i have designed a biometric system which takes biometric as the trait it takes each person types on a keyboard in a characteristic way. In this project biometrics are not aimed to be unique for all users but instead it is to put an extra layer of authentiction and verification to the existing users who are trying to access the system by preventing them from the intruders. The keystrokes of an individual could be checked discreetly as that individual is entering in data. I have used digraph to find the words pressing and conveying time of each word is gotten this cycle is reiterated for n starters which gives us the train data and a while later when an other client endeavors to enter the structure he wouldn't be allowed to enter there of psyche as he would have an other keystroke. I have also made an interface for the user to interact with the system there we are going to store the signature profile in the database where the data trained we have latency average average and much more information suing which we are going to calculate the Stdeviation which serves as important factor in threshold calculation.

Keystroke dynamics is the most common way of breaking down the manner in which a client types on a console and recognize him in light of his routine composing cadence. Keystroke dynamics is not what you type, but it is the manner in which you type. A significant part of the past exploration in keystroke dynamics has been for confirmation and check purposes. It contains social event and denoting clients' keystroke and mouse data. This cycle runs behind the scenes, gathering keystrokes paying little mind to anything more the application that is eventually in center. The main noticeable sign that the application is running has a symbol in the work area framework plate. The information blend server is permitted us to protect part dimness what's more, kept up with distant clients. Our important goal is to get the Users way to deal with creating in his login nuances with an overall articulation, in light of how he types down his login certificates our structure is stooped to get the forming configuration by getting the times for conveyance and pressing key on the control center so as our model can set up the data and make a compass for which when the client endeavors to again login, he/she is verified according to how they type.

Requirement	Description	
Acceptability	Biometric characteristic is accepted by users.	
Circumvention	Biometric characteristic is impostor proof.	
Collectability	Biometric characteristic can be measured quantitatively via detectors.	
Distinctive	stinctive Biometric characteristic is unique for each individual.	
Performance	Biometric feature collection can be achieved with a desired recognition	
	accuracy and speed.	
Permanence	rmanence Biometric characteristic is immutable.	
Universality	All individuals possess the biometric characteristic.	

Fig 2 Traits of a biometric system

Keystroke occasion can be compared milliseconds accuracy by programming. Along these lines, it is unfeasible to repeat one's keystroke design at such high goal without huge measures of exertion. Rather than customary physiological biometric frameworks, for example, palm print, iris, and finger impression acknowledgment that depend on committed gadget and equipment foundation, keystroke dynamics acknowledgment is altogether programming implementable. The advantage of low reliance on particular equipment not exclusively can altogether lessen sending cost yet additionally makes an optimal situation for execution in distant validation climate. Keystroke dynamics confirmation can be arranged as verification and identification. Verification alludes to the most common way of sealing a legitimacy of guaranteed character. As such, "is this individual truly who the person in question pronounces to be." This is a nice evaluation procedure that significant immaterial vertical and is what's going on in our general populace's security access control climate. In fact, distinguishing proof connotes "is this person in our informational collection, if to be sure, to whom this presented character has a spot with." Identification is overall extra dreary, all the more delayed in responsiveness, and require higher taking care of breaking point. Coincidentally, distinguishing proof mode has its own excellent usage like legitimate assessment and interference revelation.

3. TECHNICAL SPECIFICATION

Terminologies to know

a) False reject rate

A false rejection happens when an approved subject is rejected by the biometric framework as unapproved. False rejections cause disappointment of the supported clients, decline in work due to awful access conditions, and utilization of resources for revalidate endorsed clients.

b) False accept rate

A false acceptance is when an intruder is authorized in the system and is treated as an trusted user. On the off chance that an association's biometric control is delivering a ton of false rejections, the general control could need to bring down the exactness of the framework by decreasing how much information it gathers while verifying subjects. At the point when the information focuses are brought down, the association gambles with an expansion in the false acceptance rate. The association takes a chance with an unapproved client obtaining entrance.

c) Crossover Error Rate

The Crossover Error Rate (CER) portrays where the False Reject Rate (FRR) and False Accept Rate (FAR) are equivalent. CER is otherwise called the Equal Error Rate (EER). The Crossover Error Rate depicts the general precision of a biometric framework.

i) Error rates

The fundamental idea driving involving keystroke dynamics as an authenticator is to identify the remarkable examples that exist when a client types on a keyboard. These examples can be perceived in a wide range of ways, including factual classifiers and brain organizations. The aftereffects of the order (and accordingly the presentation of the classifier) can be estimated utilizing two error rates: the False Accept Rate (FAR) and False Reject Rate (FRR). FAR communicates the probability that an unapproved client (i.e., a sham) will be allowed admittance to the safeguarded asset, and FRR addresses the probability that an approved client will be denied admittance to the safeguarded asset. Thusly, the FRR addresses an irritation to the client since being falsely dismissed basically implies that the client should make one more endeavor to validate. There is a particular compromise among FAR and FRR, a low FAR in mix with a high FRR addresses a high-security climate that is not prone to be acknowledged by clients since they are probably going to experience countless reconfirmation demands. The opposite environment has a high FAR and a low FRR what's

more, addresses a low-security environment since it has the potential for enduring various unapproved clients, yet on occasion requires a supported client to re-affirm.

ii) Static vs dynamic text inputs

The ongoing keystroke dynamics studies pick either static or on the other hand unique text passage as a reason for looking at a composing example to an example captured during enlistment. Static text entry requires the client to type a pre-portrayed text string, for example, their username and secret expression and contemplates the keystroke guides to those gathered at selection using something fundamentally the same as pre-described string. Dynamic text passage permits the client to type any text they wish, with confirmation occurring by contrasting example similitudes of generally composed letters.

iii) Metrics

The standard measurements that are utilized in keystroke dynamics research are between key idleness and key hold-time. The previous is a proportion of how much time between when a key is delivered and the resulting key is squeezed. The last option is a proportion of how much time between when a key is squeezed and when a similar key is delivered.

a) Simple code for capturing the data from the new user, which is the Name, password, confirm password.

```
public class User {
   private int id ;
   private String userName;
   private String password;
   private int successfulLoginNo;
   private int userNameLoginNo;
   private int passwordLoginNo;
   private int phraseLoginNo;
   private UserSignatureProfile profile;
   public User(){}
   private int loginNo;
   public int getLoginNo() {
       return loginNo;
   public void setLoginNo(int loginNo) {
       this.loginNo = loginNo;
   public int getSuccessfulLoginNo() {
       return successfulLoginNo;
   public void setSuccessfulLoginNo(int successfulLoginNo) {
       this.successfulLoginNo = successfulLoginNo;
    public int getUserNameLoginNo() {
        return userNameLoginNo;
    public void setUserNameLoginNo(int userNameLoginNo) {
        this.userNameLoginNo = userNameLoginNo;
    public int getPasswordLoginNo() {
        return passwordLoginNo;
    public void setPasswordLoginNo(int passwordLoginNo) {
        this.passwordLoginNo = passwordLoginNo;
    public int getPhraseLoginNo() {
        return phraseLoginNo;
    public void setPhraseLoginNo(int phraseLoginNo) {
        this.phraseLoginNo = phraseLoginNo;
    public User(String userName, String password) {
        super();
        this.userName = userName;
        this.password = password;
    }
```

b) The below code verifies if all the text fields required to register or login as a new user are satisfied or not:

```
private boolean verifyFormTextFields() {
     String userName = userNameField.getText();
     String password = passField.getText();
     String stPhrase = userPhraseField.getText();
     boolean flag = false;
     if(userName != null && userName.trim().length()>0){
         if(password != null && password.trim().length()>0){
             if(stPhrase != null && stPhrase.trim().length()>0){
                 flag = true;
             } else {
                 retryAgain("Missing Standard Phrase, please retry again.");
                 flag = false;
         } else {
             retryAgain("Missing Password, please retry again.");
             flag = false;
     } else{
         retryAgain("Missing Username, please retry again.");
         flag = false;
     return flag ;
 }
```

d) code to verify the original user typing pattern against the person who is claiming to be the user, it checks whether you are who you claim to be or not.

```
public boolean verify(User user, StringProfile profile, ArrayList<Digraph> digraphs, String type){
    VerificationStatistics stats = generateVerificationStatistics(user, profile, digraphs, type);
    int match = 0;
    double totalWeight = 0;
    for(int i = 0; i < digraphs.size(); i++){</pre>
        double meanLetancy = profile.getStringProfileEntries().get(i).getMeanLatency();
        double digraphLetancy = digraphs.get(i).getLatency();
        double stdDev = stats.getEntry(i).getStdDev();
        if((digraphLetancy > (meanLetancy - stdDev))) && (digraphLetancy < (meanLetancy + stdDev))){
             totalWeight += stats.getEntry(i).getFinalWeight();
        }
    }
    System.out.println("Total Weight " + totalWeight);
System.out.println("match " + match+ " from "+digraphs.size());
    if (totalWeight > 0.5 || ((match / (digraphs.size())) > 0.75))
        return true;
    return false;
}
```

E) Calculating the average of the typing pattern and generating the standrd deviaton

```
public static StringProfile constructStringProfile(ArrayList<ArrayList<Digraph>> trainingData){
    StringProfile profile = new StringProfile();
    for(int i = 0; i < trainingData.get(0).size(); i++){</pre>
        StringProfileEntry entry = new StringProfileEntry();
        entry.setDigraph(trainingData.get(0).get(i));
        double meanLatency = 0;
        double sumOfX = 0;
        double sumOfXSquared = 0;
        for(int j = 0; j < trainingData.size(); j++){</pre>
            sumOfX += trainingData.get(j).get(i).getLatency();
            sumOfXSquared += Math.pow(trainingData.get(j).get(i).getLatency(), 2);
        entry.setSumOfX(sumOfX);
        meanLatency = sumOfX / Constants.NUMBER_OF_TRAINING_TRIALS;
        entry.setMeanLatency(meanLatency);
        entry.setSumOfX(sumOfX);
        entry.setSumOfSquaredX(sumOfXSquared);
        profile.addProfileEntry(entry);
    }
    return profile;
```

Keystroke dynamics has gotten a fundamental measure of thought in endorsement research circles since it gets the chance of giving an immediate, great methodology for endorsement that can executed use existing equipment. specialists have kept on advancing by performing focuses on extra control place sorts, separating new model getting sorted out calculations, and perceiving new keystroke qualities. Moreover, experts have embraced commonly used connection techniques, for instance, False Accept Rate (FAR), Bogus Reject Rate (FRR), and Equal Error Rate (EER) to show upgrades in keystroke dynamics strategy quantitatively. From these evaluations we have observed that it is far-gotten that keystroke dynamics alone will be sufficiently great to inconceivably see clients, yet it shows marvelous confirmation as a piece of a more noteworthy multimodal biometric certification technique.

All in all, keystroke dynamics make a benchmark for clients' composing and afterward utilize that gauge to look for irregularities. In case an abnormality is recognized, it can demand an other confirmation component to check the client or end the gathering immediately.

Biometrics factors offer a tempting possibility to IT leaders: a verification factor programmers can't take or reproduce. Danger entertainers can every now and again break passwords, or even supposition them from online entertainment data. SMS informing can wind up blocked by cunning programmers, and some could send counterfeit SMS messages in a novel phishing assault. With time and persistence, danger entertainers can repeat the mark of hard tokens, or they could take the low-tech course of taking them.

Obviously, for this reason ventures should send multifaceted verification as opposed to single-factor confirmation or two-factor approval; the more factors you convey, the more developers deterred or hindered from access. Be that as it may, no advanced edge made through personality remains totally impenetrable. In the long run, with the right apparatuses and abilities, programmers can get entrance.

Biometric factors offer an elective that programmers can't take or phony; this is generally speaking considering the way that as yet, no software engineer has really taken biometric data and a short time later elaborate it in an attack. Regardless, this doesn't mean developers can never do accordingly; it essentially suggests that software engineers have not done as such as of now.

On the other hand, keystroke dynamics fall under the class of "behavioral biometrics;" these utilization the ways of behaving of the clients as a verification factor. Thusly, programmers can't "take them" since they are necessary to the character of the clients; they likewise can't duplicate them for similar reasons.

4. DESIGN APPROACH AND DETAILS

4.1 Design Approach

The system starts with a simple user interface where by you have to prompt whether you are a new user or an existing user. If you are a new user then it asks for your name and password for the first time in order to register you in the system.

Once it has taken the new user details from you, it will prompt you to enter the details 10 times so as it can train the data against your typing patterns which are the pressing and

releasing times of the keys you type and captures the time and calculates a average which we call latency mean and calculate the standard devation which will help it determine the existing user is genuine or trying to replicate the original user.

On the other hand if you choose Existing user instead of new user, it asks for your name and password and a general phrase so that it can match ur typing characteristics against the database in which your timings for pressing and relasing the keys are stored. The application uses Java as a language to run through functionalities and MYSQL database to store the user data.

The application upholds different clients attempting to login and distinguishing regardless of whether they are approved to login by contrasting their composing qualities and the idleness normal and standard deviation taken by the quantity of preliminaries which for our situation is 10, so in view of those values the entrance is allowed or denied. Besides I have likewise executed a check framework by which in the event that a field is missing eg. Secret key ought to be between 8 to 12 characters, in the event that it isn't the framework prompts you to restart the cycle once more.

Most of the techniques exploiting KD for biometric affirmation purposes rely upon the timing information to isolate among clients. Cells regularly offer the availability of timestamp information, recording the minutes when different keys are pressed and conveyed. From such central information, a couple of idleness measures can be then figured.

- Key-press timestamp (*KPT*), which is the system timestamp value, at the moment of key press, typically recorded in milliseconds;
- Key-release timestamp (*KRT*), which is also the system timestamp value, at the moment of key release, in milliseconds.
- Key press-release time (*PRT*), corresponding to the lapse from the press of a key to the release of the other one. This is often referred to as generalized key dwell time;
- Key press-press time (*PPT*), corresponding to the lapse from the press of a key to the press of the other one;
- Key release-press time (*RPT*), corresponding to the lapse from the release of a key to the press of the other one. This is often referred to as key flight time (*KFT*);
- key release-release time (*RRT*), corresponding to the lapse from the release of a key to the release of the other one.

From the collected KD data, the user typing speed can be determined by means of several measures, such as the words per minute (WPM), defined in as the average number of words that could be typed by the considered user in a minute. PINs have been always used in the access control protocols of mobile devices, mainly to unlock SIMs. They are also often required to access online services such as bank accounts. Not by chance, most of the research on KD-based biometric recognition for mobile devices has focused on PINs as input.

Performance Factors

- False acceptance rate (FAR), providing the percentage of recognition attempts made by impostors that are falsely accepted in systems working in verification modality.
 Low values of FAR correspond to secure systems;
- False rejection rate (FRR), providing the percentage of recognition attempts made by legitimate users that are falsely rejected by the system, in systems working in verification modality. Low values of FRR correspond to usable systems;
- Equal error rate (EER), the operating point where FRR and FAR are equal. For a verification system to be accurate, the EER should be as low as possible;
- Correct identification rate, giving the probability of correctly determining the identity of the presented subject among a set of possible users, for systems working in identification modality.

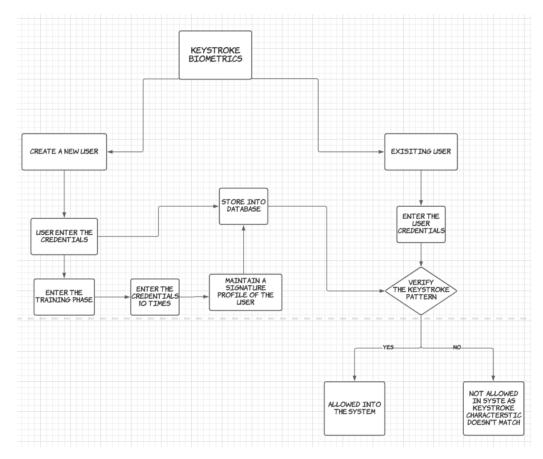


Fig 4. An Architectural diagram of the system implemented

4.2 Codes and standards

These are the standards that ensure reliability of the product or the organization based on various parameters like social, environmental, health etc. Different standard organizations have different codes which are used to certify the products or the organization business. The codes that maybe applicable to this product are as follows:

- **ISO/IEC TR 20004:2015** This code is standardized for Information Technology and security techniques. It aims at refining the software vulnerability of any product.
- ISO/IEC 27000:2018 This code is standardized for Information Technology and security techniques along with management systems. It aims at performing the overview and the vocabulary check.

- **ISO 14005:2019** This code is standardized for Environmental management systems. It aims at providing a flexible approach for a phased implementation keeping the environmental hazards in check.
- ISO/IEC/IEEE 15288 This code is standardized for Software Technology. It
 defines the cycle of a software product from the time of no-software till the time
 software was implemented.
- **ISO/PAS 45005:2020** This code is standardized for Occupational Health and Safety Management. It aims at providing a guideline for safe and secure working during Covid-19.
- **ISO 9004:2018** This code is standardized for Quality Management. It aims at guiding the path to achieving sustained success of an organization or of a product.
- IEEE/EIA 12207— This code is standardized for Information Technology. This
 talks about the Lifecycle framework, from the inception till the deployment of the
 product.

4.3 Constraints, Alternatives and tradeoffs

Keystroke dynamics biometrics are second rate as far as confirmation exactness because of the varieties in composing mood that brought about by outside variables like injury, weariness, or interruption. Also most conduct biometrics by and large experience lower permanency contrasted with physiological biometrics. Composing example of a human may steadily change following the accustomization towards a secret word, developing composing capability, variation to enter gadgets, and other natural elements. At the point when a subject is chosen, the model or test data used in the assessment will by and by be coordinated into, and become a piece of the readiness set data. Each subsequent test is as of now appeared differently in relation to the sum of the previous data from got keystrokes. The furthest down the line data can be even more enthusiastically weighted to give more emphasis on the most recent keystrokes that the subject made, thinking about a tireless update of the selection data. Each resulting test is presently contrasted with the entirety of the former information from caught keystrokes. The latest information can be all the more vigorously weighted to give more accentuation on the latest keystrokes that the subject made, taking into consideration a persistent update of the enlistment information. On the off chance that we are capable to

foresee the pace of rot in the precision of the keystroke biometric, then, at that point, the impacts can be invalidated by integrating them into the biometric. We recommend that this will not just limit the corruption of exactness over time, yet in addition wipe out the requirement for occasional reenrollment.

Bayesian order has been utilized in a large number areas of examination including highlight order of information related with keystroke dynamics. In any case, we propose that Bayesian examination can be a helpful device in defeating the versatility limits and expanding the ability and capability of keystroke biometric frameworks as a device in character verification in a multicomponent framework. Bayesian factual examination depends on extra data applied to unique probabilities or suspicions described as measured probabilities. By adding the results of keystroke biometric data examination (long-text online test results, for example) to what we most certainly have some knowledge of the student's character, we could extend our precision in phrasing of Identity.

FAR shows the capacity of a technique to accurately separate certified and faker. Execution markers utilized by the explores are summed up as follow. False Rejection Rate (FRR) implies the rate extent between falsely denied affirmed clients against irrefutably the quantity of authentic clients getting to the system. At times known as False Nonmatch Rate (FNMR) . A lower FRR infers less dismissal and simpler access by certified client.

False Acceptance Rate (FAR) is depicted as the rate degree between falsely perceived unapproved clients against the full scale number of scams getting to the framework. A more modest FAR demonstrates less faker acknowledged. Equal Error Rate (EER) is utilized to decide the general exactness as well as a similar estimation against different frameworks. It very well might be in some cases referred to as Crossover Error Rate (CER).

The crossover error rate is the region where FAR = FRR. CER is generally called the Equal error rate (EER). The crossover error rate portrays the overall precision of a biometric system. As the responsiveness of a biometric framework builds, FRRs will rise and FARs will drop. Alternately, as the responsiveness is brought down, FRRs will drop and FARs will rise.

Assuming the framework is utilizing biometric gadgets, the False Acceptance Rates (FAR), the False Reject Rates (FRR), and the Cross-over Error Rates (CER) ought to be tried. A biometric gadget is more accurate and dependable as the CER goes down and you will need

to lay out acceptable edges in your test plan. Different measurements to think about for biometrics incorporate the Failure to Enroll (FTE) rate and the Failure to Acquire (FTA) rate. FTE means how much individuals who can't utilize the framework because of some kind of inconsistency and FTA signifies the quantity of clients who can't deliver an acceptable enlistment picture to utilize the gadget.

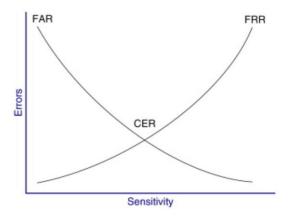


Fig 4.1 Crossover Error rate graph

Confidentiality tests determine if unauthorized disclosure is possible. When you perform confidentiality tests, you are trying to determine if data is disclosed to people that it is not intended for. You are also trying to determine that data is readable and executable by the people it is intended for.

Before you can set up tests to ensure confidentiality, you have to understand a bit about confidentiality risks and vulnerabilities. Data traveling in plaintext over communications lines is vulnerable to sniffing. Weak passwords can be compromised using password crackers. Confidentiality tests look to ensure that authentication and encryption mechanisms work according to the security requirements. It's also important to ensure that the authentication and encryption mechanisms have not just been implemented, but that they have safeguards built around them to protect them from being sabotaged.

i) Normal Hardware

One of the great advantages of keystroke dynamics biometrics is low reliance on devoted equipment framework. Consequently, it is plain as day why most analysts go for promptly accessible equipment for study. The most notable choice is the by and large available QWERTY console, followed by worked in PC console.

Some exploration works, not at all like others, just utilized explicit part of an equipment. The examination confined client to utilize num-cushion of a keyboard with only one finger to reproduce a devastated trial condition. That is the thing they confided if extraordinary result was achieved in such silly game plan, execution in a less restrictive environment could likely accomplish better execution.

Then again, used Synaptic Touchpad joined to a note pad to quantify finger tension and position. Their goal was to execute keystroke dynamics biometrics on touch screen cell phones, however because of the innovation bottleneck by then of time, it is figured out why a less expensive option had been picked. Albeit the gadget responsiveness probably won't be any place equivalent to a genuine touch screen innovation, the thought was helpful for scientists when the innovation opens up.

ii) Customized hardware

Customary information gadgets, for example, typical PC keyboards are just fit for delivering keystroke timing vector as component information for investigation. An optional element information that might be demonstrated more unmistakable is the strain grouping while at the same time collaborating with the info gadgets. Accordingly, various specialists have attempted to change the current gadgets to incorporate strain delicate recipients.

One more change was made to an elastic layer keypad that looks like an ATM machine input arrangement, with the goal of further developing security on a numeric PIN framework. The first mounted printed circuit board under the keypad was supplanted by exclusively manufactured force delicate resistors. In any case, the real execution to the monetary region is fairly dubious in view of the cost of replacement to the entire gear structure.

iii) Mobile Devices

While typographical contribution from PC keyboard has been the principal center at the outset phase of keystroke dynamics research, mathematical base contribution from compact communicational gadgets has slowly acquired consideration since the wide spread utilization of cell all around the world in the twentieth century.

Research works, for example, performed probes regular mathematical key cushion mobile phone in endeavor to confirm client by means of short info message. The drive was engaging anyway the issue of cross-stage comparability across arranged model of contraptions remains an open request

Alongside the fast advancement of innovation, cell phones have likewise acquired more noteworthy handling capacity. Java empowered Symbian telephone was chosen by as the stage for their review. They endeavored to involve a few computational costly brain network calculations for acknowledgment and have yielded a few empowering results. Sadly, a significant difficulty was the debasement of reaction time to the cell phone that could influence client acceptance.

iv) Development Platform

Since the most well-known client communication including text and mathematical information is through a PC, scientists who were working on keystroke dynamics are practically completely founded on neighborhood PC stage. Before the 21st 100 years, keystroke dynamics try prototype was created on working framework (OS) stage utilizing third-age programming language (3GL) like FORTRAN and Turbo Pascal. Some other time when Microsoft items overwhelm most working framework, an exploratory prototype was based on top of MS DOS and windows climate by utilizing dialects like C++ and Visual Basic .

Attributable to the speed of web advancement somewhat recently, exploratory stage has been moved to the online climate with web programming apparatuses like JavaScript , Java Applet , and Flash . It is just over the most recent few years; a few works have been created in light of cell phone climate. Getting going with portable emulator , Symbian working framework , and most as of late Android stage.

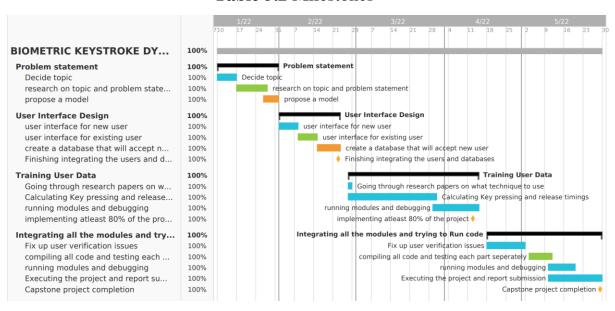
5. Schedule, Tasks and Milestones

Table 5.1 schedule and Tasks

	TASKS & MILESTONES	DEADLINE	STATUS
1.	Decide on a relevant thesis topic	15/01/2022	COMPLETE
2.	Research on the topic, in order to evaluate previous work and identify gaps.	20/01/2022	COMPLETE
3.	Come up with an appropriate problem statement and the aim of the thesis, based on the research undertaken	01/02/2022	COMPLETE
4.	Decide on a proposed model to bring out the solution for the said problem statement	10//02/2022	COMPLETE
5.	Write up code for user interface	25/02/2022	COMPLETE
6.	Write up code for capturing times	08/03/2022	COMPLETE
7.	Implementation of code and debugging	20/03/2021	COMPLETE
8.	Train and test the data, n trials So to check if the output is valid	17/04/2022	COMPLETE
9.	Running the whole project smoothly And making a full report	03/05/2021	COMPLETE

5.2 Gantt Chart

Table 5.2 Milestones



6. PROJECT DEMONSTRATION

In the initial phase the system starts with a simple user interface where by you have to prompt whether you are a new user or an existing user. If you are a new user then it asks for your name and password for the first time in order to register you in the system.

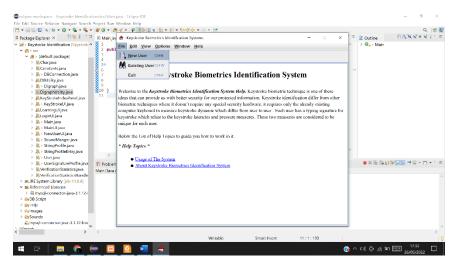


Figure 6.1 user interface for creating new user



Figure 6.2 registration details for new user

Once it has taken the new user details from you, it will prompt you to enter the details 10 times so as it can train the data against your typing patterns which are the pressing and releasing times of the keys you type and captures the time and calculates a average which we call latency mean and calculate the standard devation which will help it determine the existing user is genuine or trying to replicate the original user.

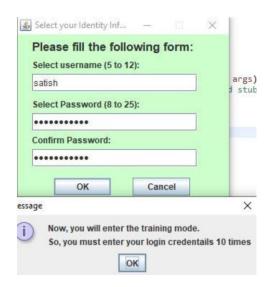


Figure 6.3 Training mode where by the user typing patterns are captured



Figure 6.4 The data has been captured 10 times and training of data has been successfully completed.

On the other hand if you choose Existing user instead of new user, it asks for your name and password and a general phrase so that it can match ur typing characteristics against the database in which your timings for pressing and relasing the keys are stored.

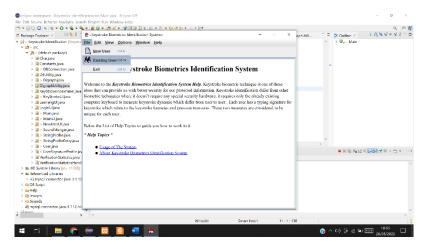


Figure 6.5 an instance whereby the existing user is now trying to login.

The claimed user will then have to pass the verification test which is get the typing pattern correctly and within the time stored in the system which is the latency mean time and the standard deviation.



Figure 6.6 User now goes through verification phase.

If the user passes the verification, they are granted appropriate access else they are rejected by the system, in this case either one of the messages pop up.



Figure 6.7 one of the either messages pops up

The pressing and release times are captured when the user shifts from one key to the other.

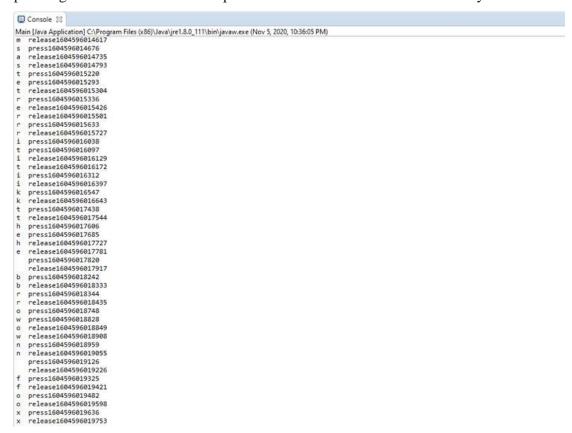


Fig 6.8 pressing and release time generated

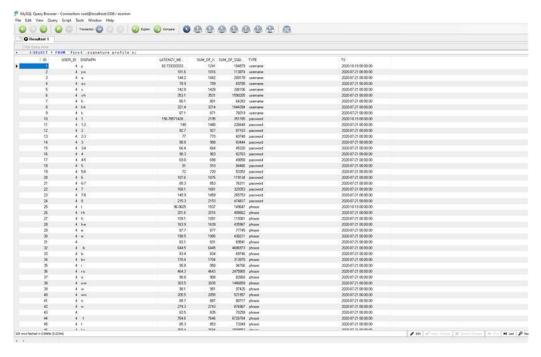


Fig 6.9 Signature profiles for different users

7. COST ANALYSIS/ RESULTS AND DISCUSSIONS

Keystroke dynamics biometrics are conduct qualities, which suggests that it is difficult to secure a precise composing example of even from a similar person. This is helpful for validation, by which the uniqueness can be utilized to separate one's keystroke dynamics from others. Then again, it might likewise create some issues due to intraclass changeability.

Data input plays a very huge role as if the characters are widely spaced on the keyboard the results might be vague sometimes as the user might try to find the keys if they purchase a new machine and try to login. This causes some abiguities as biometric systems are a behavioural trait. The writing concentrate on recommended that keystroke dynamics biometrics are probably not going to supplant existing information based verification completely and it is likewise not powerful enough to be a sole biometric authenticator. In any case, the benefit of keystroke dynamics is unquestionable, for example, the capacity to operate in covertness mode, low execution cost, high client acceptance, and simplicity of coordination to existing security frameworks.

These make the premise of a possibly viable approach to improving in general security rating by having a huge job in impact of a bigger multifaceted validation component. With regards to execution improvement strategy, a ton of examination works have been zeroing in on further developing characterization calculations. In any case, suggested that quality extent of keystroke plans is an essentially more determinant rules than classifier used. Nature of client design directly influences the introduction of a check system, in this way, arranging a fair and discriminative keystroke incorporate profile is a fundamental cycle that should not be subverted. Information getting is the starter and key period of keystroke elements research. Because of the lower development contrasted and other laid out biometrics, openly accessible benchmark data sets are restricted. Yet a couple of researchers have moved forward and share their hand made enlightening file, due to the different improvement plans and factors, many

have chosen to create in-house educational assortment. It is aggregately agreed that attempts that consolidates colossal number of subjects better suggest the versatility of study. Deplorably the greater part of the investigations performed include just modest number of subjects. This is justifiable because of different issues and challenges experienced in information assortment process.

Information gathered will ultimately be utilized for execution assessment. The most notable technique for execution assessment is the degree of accuracy of a structure's ability to perceive genuine and farce. Fraud tests are normally acquired by either a similar person who adds to the age of certifiable examples in data set or through one more gathering of people going after or reenacting the veritable examples put away in the data set. The previous forces members to give more data sources and dedicate additional time in the trial. The extensive interaction might prevent volunteer cooperation.

Generally speaking, as per the performed research, it is sensible that biometric recognition on cell phones can be acted, in actuality, applications taking advantage of the particular attributes of every client's composing patters. Having the option to accomplish such objective would be exceptionally advantageous for incorporating KD-based recognition frameworks into the entrance control components previously utilized while utilizing cell phones. Secure recognition techniques would hence permit gaining admittance to the actual gadgets, as well as to administrations, for example, on-line installments, web based business, email, or financial balances, without requiring a particular activity from the client. In any case, a few endeavors should be made prior to making the previously mentioned situations relevant, in

actuality, since numerous perspectives influencing the recognition capacities connected with the composing patterns ought to in any case be appropriately investigated, like the impacts of the embraced pose, the possibility of cross-stage recognition, or the assessment of the viability of the picked inputs.

COMPARATIVE ANALYSIS AGAINST OTHER ALGORITHMS

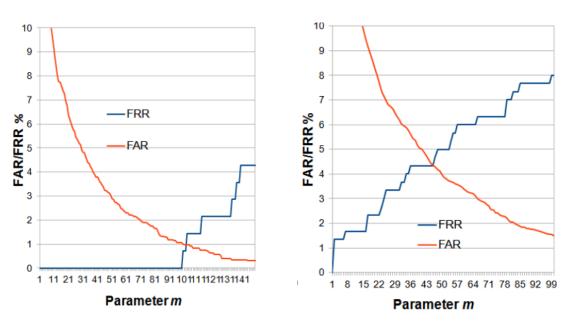


Fig 7.1 Large population = Low EER

Fig 7.2 Small population = Higher EER

The main contribution of this study was the evaluation of the text-input performance as a function after training on the longest keystroke samples available. As the number of keystrokes per test sample was increased, the EER decreased roughly logarithmically, and the EER increased with the increase in population size.

As we can see the Figure on the Left side has an EER of 1% with a large number of keystrokes, on the left side we can observe that the EER is roughly 4% with a smller population which indicates that large population provides a much better authentication mechanism.

Using a random forest classifier, they achieved an equal error rate (EER) of 8.6%. Bakelman's results were not as robust as the results obtained by the team at Carnegie Mellon when used in conjunction with the features used in the Carnegie Mellon study, obtaining an EER of 10.5%. However, when Bakelman and team ran their experiments using their own feature data, they recorded an EER of 6.1%. However we have managed to get much better results of about 4% with small population and much better results i.e almost 1% with our proposed model.

Firstly the model implemented has the aspect of training data 10 times which makes the aunthentication more secure as the number of times the data has been fed into the system by the user is more which increases the security as the time is spread over a range which will avoid the system from denying access to the original user.

Secondly the current algorithms which have already been implemented by other scholars have the gap of whereby they accept passwords/usernames which are incorrect and allow them to use the backspace key or delete key which allows brute force attacks much more faster and easier, While in our case the model which is implemented does not aaccept any other charaters apart from letters, numbers and symbols. In the case where by any user tries to hit backspace the system will automatically redirect it to the start page again.

Also the system which are currently implemented have only username and password as an input which in most cases allows the fake user to login. In the system implemented by me, I have added a a random sentence in addition to username and password which allows the user to type keys and record his pattern so that even if the intruder who get the user name and password correct will still take to input the sentence and will get rejected by the system. Most biometric approaches attempt to only capture pressing time for the keyboard keys as opposed to our system which adds an extra layer of authentication by capturing the time taken for the user to release key as well as how long the whole process takes from opening the page to click the final login button.

8. SUMMARY

The above results demonstrate to us the properties which are required for the system in order to authenticate the system aggresively without any stopping. This biometric grants "nonstop confirmation" of a singular's character over a meeting after the individual logs in utilizing a more grounded biometric like finger impression or iris so go about as dynamic biometric framework, so we have used digraph to catch the words pressing and releasing time of each word is captured this process is repeated for n trials which provides us the train data and then when a different user tries to enter the system he would not be allowed to enter in the system as he would be having a different keystroke pattern. We have also made an interface for the user to interact with the system there we are going to store the signature profile in the database where the data trained we have latency average average and much more information suing which we are going to calculate the standard, deviation which serves as important factor in threshold calculation. This strategy can assist with improving a great deal of utilizations in the space of network protection with no additional equipment cost and non rudeness. The user can expect to have a accuracy of about 90 % with the implemented system, it is also very cost effective and resources are minimized in this situation.

Getting to basic data, for example, individual information, financial balances, individual wellbeing data and oversee basic framework frameworks might be open by means of cell phones. Getting such basic frameworks against digital dangers is crucial for the public wellbeing and postures many difficulties as there are numerous related parts in any figuring framework. Client verification assumes a significant part in giving admittance to asset and it is a fundamental security layer for all frameworks. The more verification factors a framework has, the more troublesome it is to by-pass the confirmation or mimic a genuine client.

9. REFERENCES

- Revett, K., Gorunescu, F., Gorunescu, M., Ene, M., Magalhaes, S., & Santos, H. (2007). A machine learning approach to keystroke dynamics based user authentication. *International Journal of Electronic Security and Digital Forensics*, 1(1), 55-70
- 2. P. Magalhaes, K. Revett, and H. Santos, —Keystroke dynamics: stepping forward in authentication, 2006.
- 3. Killourhy, K. S., & Maxion, R. A. (2009, June). Comparing anomaly-detection algorithms for keystroke dynamics. In 2009 IEEE/IFIP International Conference on Dependable Systems & Networks (pp. 125-134). IEEE.
- 4. Bartlow, N., & Cukic, B. (2006, November). Evaluating the reliability of credential hardening through keystroke dynamics. In *2006 17th International Symposium on Software Reliability Engineering* (pp. 117-126). IEEE.
- 5. F.Monrose, M.K.Reiter, and S. Wetzel, "Password hardening based on keystroke dynamics," International Journal of Information security, pp. 69–83, 2002.
- Anil K.Jain, Arun Ross, Salil Prabhakar, "An Introduction to Biometric Recognition," in IEEE Transactions On Circuits and Systems for Video Technology, vol 14, no. 1, 2004.
- 7. len Peacock, Xian Ke and Mathew Wilkerson, "Typing Patterns: A Key to User Identification," in IEEE Security & Privacy, pp. 40–47, 2004.
- 8. NSA. Commercial National Security Algorithm Suite and Quantum Computing FAQ. [Online]. Available: https://cryptome.org/2016/01/CNSA-Suite-and-QuantumComputing-FAQ.pdf
- 9. S. Maheshwary, S. Ganguly, and V. Pudi, "Deep secure: A fast and simple neural network based approach for user authentication and identification via keystroke dynamics," in IWAISe
- 10. M. E. Karim, K. S. Balagani, A. Elliott, D. Irakiza, M. O'Neal, and V. Phoha, "Active authentication of keyboard users: Performance evaluation on 736 subjects," arXiv preprint arXiv:1804.08180, 2018.
- 11. R. S. Milad Mohammadi, Rohit Mundra. CS 224D: Deep learning for NLP, lecture notes part IV, spring 2015.

- 12. J. Kim and P. Kang, "Recurrent neural network-based user authentication for freely typed keystroke data," arXiv preprint arXiv:1806.06190, 2018.
- 13. MITRE. Cwe-327:use of a broken or risky cryptographic algorithm. [Online]. Available: https://cwe.mitre.org/data/definitions/327.html
- 14. F. T. Sheldon, J. M. Weber, S.-M. Yoo, and W. D. Pan, "The insecurity of wireless networks," IEEE Security & Privacy, vol. 10, no. 4, pp. 54–61, 2012.
- 15. T. Koponen, P. Eronen, M. Särelä et al., "Resilient connections for ssh and tls." in USENIX Annual Technical Conference, General Track, 2006, pp. 329–340.
- 16. P. Kobojek and K. Saeed, "Application of recurrent neural networks for user verification based on keystroke dynamics," Journal of telecommunications and information technology, no. 3, pp. 80–90, 2016.
- 17. Sheng, Y., Phoha, V., and Rovnyak, S. A paralleldecision treebased method for user authentication based on keystroke patterns. IEEE Transactions on Systems, Man, and Cybernetics 35, 4 (2005), 826-833.
- 18. L. M. Vizer, L. Zhou, and A. Sears. Automated stress detection using keystroke and linguistic features: An exploratory study. International Journal of Human-Computer Studies, 67(10), 2009.