Journal of Social Sciences (COES&RJ-JSS)
ISSN (E): 2305-9249 ISSN (P): 2305-9494
Publisher: Centre of Excellence for Scientific & Research Journalism, COES&RJ LLC
Online Publication Date: 1st January 2016
Online Issue: Volume 5, Number 1, January 2016
http://www.centreofexcellence.net/J/JSS/JSS Mainpage.htm

A Survey of behavioral authentication using keystroke dynamics: Touch screens and mobile devices

Ramzi Saifan, Asma Salem, Dema Zaidan, Andraws Swidan

Abstract:
Nowadays most systems became computerized and use internet for remote access, including systems which have critical and sensitive data such as banks and governmental institutions. This led to the huge need for a reliable and efficient authentication system to secure data. User authentication is mostly done using passwords. But it is not a sufficient way to use just a password since it has many drawbacks, like guessing them, brute force attacks, key-loggers and social engineering. Additional authentication procedure is needed to enhance password security. Keystroke dynamics is one of the famous behavioral measurements that rely on utilizing the typing rhythm of each individual. It is used to strengthen password authentication in an efficient and cheap way since no hardware will be added. This paper presents a comprehensive survey on research in the last two decades on keystroke dynamics authentication. The objective is to discuss, summarize and provide insightful comparison about the well-known approaches used in keystroke dynamics such as statistical and neural network approaches, as well as offering suggestions and possible future research directions, especially for touch-screen and mobile devices. Keystroke dynamics could provide a second authentication factor for touch screen devices, as they are rapidly increasing in their use and are replacing the classical keyboards in the markets.

Keywords:
Authentication; security; keystroke dynamics; user authentication; biometric; typing

Citation:
Ramzi, Saifan; Salem, Asma; Zaidan, Dema; Swidan, Andraws (2016); A Survey of behavioral authentication using keystroke dynamics: Touch screens and mobile devices; Journal of Social Sciences (COES&RJ-JSS), Vol.5, No.1, pp: 29-41.

Nowadays, computers have become a main part in the modern society and they are used in all aspects of our lives. In the last few years, a huge rapid increase in technology has connected businesses via internet. Many websites are accessed by too many users daily. These websites have sensitive data that must be secured [2]. Authentication is the most commonly used procedure for access control. It allows legal users to use the resources and services in an authorized manner and denying illegal ones.

Many recent computerized systems use the simple scheme of username and password for authentication. But this scheme suffers from many security limitations [23]. Moreover, users attempt to ease the burden by making multiple authentication services to use the same password, writing them down (may be theft or stolen), or using simple ones (easy to be guessed) [5] [8] [23]. This forms a real threat against authentication systems [3]. Therefore, a new scheme is needed to solve these issues.

Use of human characteristics is one of the solutions. Biometrics such as face, fingerprints and iris are possible biometric characteristic [1] [4] [5]. Each biometric characteristic has its own advantages and disadvantages. The way to choose which biometric characteristic should be used for a given authentication application, depends on the application requirements. The main disadvantage of biometric authentication is that it requires additional tools which leads to cost increase [6] [8].

The second solution that utilizes the human for authentication is human behavior. It makes use of behaving pattern of an individual [4]. Behavioral authentication can be represented in multiple ways. In this survey we focus on typing behavioral with more focus on touch screens and mobile devices. Keystroke dynamics depends on typing behavior, which can be extracted using existing hardware such as the standard keyboard. This makes it inexpensive and extremely attractive technique [6]. Typing behavior is considered to be unique for each user. Therefore, including the typing behavior in authentication will strengthen authentication. The hacker is required to know the user typing rythm in addition to the password itself in order to break it [1][3].

The remaining of this paper is organized as follow; Section two talks about authentication types and some examples on each type. Section three explains the two factor authentication and what are the main advantages and disadvantages of it. Later in Section four, we explain keystroke dynamics basic features. Section five explains the performance metrics. Section six talks about behavioral based authentication. Future work directions and conclusion are discussed in sections seven and eight respectively.

II. AUTHENTICATION

Authentication is the process of determining whether someone is the one who it is declared to be or not. Authentication is categorized into four main types, as shown in Fig. 1:

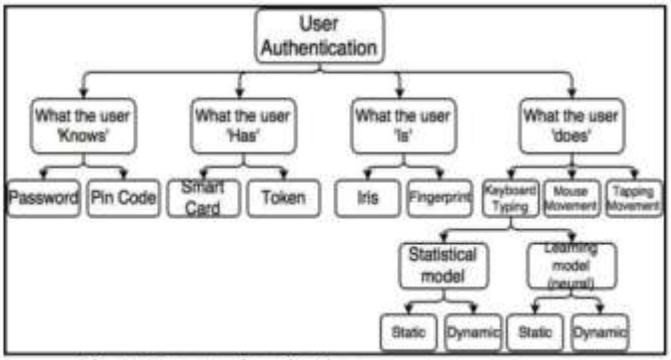


Fig.1: User Authentication.

A.What the user knows (Knowledge based)

This type depends on something that the user knows, such as passwords, PIN codes and security question. It is the most commonly used method. The simple logic here, if you know the secret password for an account, then you must be the owner of that account [5] [20] [21].

The advantages of this type are: there is no additional hardware and all the needed is a database which saves the passwords or the hash value of the password. The problems associated with this type of authentication are: the password can be forgotten, stolen, or guessed [7] [20] [23].

B.What the user has (smart card or credit card)

This type depends on something the user own, such as smart cards, mobile phone and mini devices. It is used in automated teller machine (ATM). The logic here is if you have the smart card with you, you must be the owner of the account. The problems of this type: the card may be lost, stolen, or duplicated by someone else [20] [21].

C.What the user is (Biometric based)

This type depends on the human body characteristics, such as finger print, iris, retina and hand geometry [5] [6] [20]. It is used in some banks to identify the users. This requires that the user possesses some human attribute that can be scanned and digitally documented [20].

The advantage of this type of system is that biometrics could not be stolen neither lost. The problem is that this system needs special tools to do the scanning of the biometrics [10] [20] [21]. Also, biometric authentication is not suitable in all situations, especially for remote authentication. Also, we cannot provide the machine which reads the biometric data in all situations that need authentication. Moreover, not all biometric authentication techniques are acceptable by all people. In addition to that, biometric authentication techniques require specific conditions which are not available all the time. For example, iris authentication require good light, fingerprint authentication requires the fingers to be clean and not sweaty.

D.What the user does (Behavioral based)

This type depends on the user behavior, such as typing rythm, mouse movement pattern and tapping pattern on touch screens. Typing rythm, or as it is called keystroke dynamics, sometimes considered as the best behavioral based authentication system. It is the cheapest way for almost free cost, because no hardware is being required, just the keyboard or keypad [5] [11] [20].

III. TWO FACTOR AUTHENTICATION

Two factor authentication is now taking the lead in this field. Password is not sufficient to authenticate users remotely, especially when connecting to private data, banking systems or even social networks.

Password authentication suffers from many drawbacks and flaws due to its nature. Since many people tend to use very simple and easy to remember passwords, they are usually related to themselves and can be guessed easily [23]. Moreover, people use the same passwords for many applications and websites. Some of these websites may transfer the password in an unencrypted manner which leads to be easily sniffed and restored. Another problem with passwords is typing the password very close to other person (Shoulder Surfing), some passwords characters are revealed. In tablets, this issue is more dangerous, where the typed character appears until the next character is typed where the typed character is converted to “*” [4] [20] [23].

Another major problem of password happens when a user signs up for a website, the process requires a security question and being answered for future use. Users either use answers which are dummies or they choose correct answers. The later suffers from social engineering threats. And the former suffers from being forgotten [23] [25].

Therefore, there is a need for another technique to be added to password authentication. “Something you have” and “Something you are” are not applicable over the internet. Then moving to “Something you do” factor to be added to the password will strengthen the password authentication mechanism.

Typing behavior is considered to be unique for each user; different people tend to have different typing behaviors. Therefore, including the typing behavior in authentication will strengthen password authentication. The hacker is required to know the user typing behavior in addition to the password itself [2] [23].

IV. KEYSTROKE DYNAMICS BASIC FEATURES

This survey paper emphasizes on the behavioral based authentication and focuses on keystroke authentication. In order to collect data, most of the studies use one of two ways, static-based or dynamics-based. Static-based mean fixed text, so the user has to enter the same text several times. Dynamic-based or free-text mode, the data can be collected from any typed text, where the user may be required to type long text, and features get collected [4].

There are several features which can be detected when the users press keys on keyboards or touching keys on smart phones or tablets. Basic keystroke features are:-

•Duration or Dwell time or Hold (H): It is the time interval between a key is pressed and released (by considering individual keys) [7] [8] [9] [14] [18]. For example, the time between pressing the letter”a” and releasing it.

•Latency or Flight time or Up-Down time (UD): It is the time interval between releasing a key and pressing the next key (by considering 2 keys) [4] [7] [8] [9] [14] [17] [18].

•Digraph latency or Down-Down (DD): It is the latencies between two successive key down presses (time duration between pressing 1st key and 2nd key). It is considered as the major feature data represented in keystroke dynamics domain [7] [14] [17] [18]. For example, the time from pressing “a” until the time of pressing “b”, which includes the pressing time of “a”.

•Trigraph latency: It is the latency between every three consecutive key down presses (time duration between pressing 1st key and pressing 3rd key) [14] [18]. for example in typing “car”, the time from pressing “c” until the time of pressing “r”.

•Up-Up (UU): It is the time between key-up of the first key and key-up of the second key; it is equal to UD + H of the second key [7] [17] [18].

•Down-Up (DU): It is time between key-down of the first key and key-up of the second key; it is equal to DD + H of the second key [7] [8] [17] [18].

•Pressing time: It is the time while the key is held down [4] [7].

•Releasing time: It is the time while the key is released [18].

•Overall speed: Variations of speed moving between specific keys [8].

•Frequency of errors: How much some specific error repeats [9].

•Pressure: Used when hitting keys while typing (used only for keyboard in touch screens) [9] [17].

•Finger Placement: Where the finger is placed or even the angel of the finger when pressing the key [8] [9].

•Finger choice: Which finger is used on the key of the keyboard [8] [9].

•Capital letters and special characters. This happens when the user prints any character followed by capital letter or special character which requires pressing on ‘shift’; these in many cases require longer time than two consecutive small letters [20].

•Distance between letters. For example the distance between ‘a’ and‘s’ is larger than the distance between ‘a’ and ‘p’ on the keyboard which may require more printing time [2].

V. PERFORMANCE MEASURES

Performance of such keystroke analysis is typically measured in terms of various error rates, namely False Accept Rate (FAR), False Reject Rate (FRR) and Equal Error Rate (EER).

A.False Reject Rate (FRR)

It measures the percent of valid users who are rejected as impostors. In statistics this type of errors is referred to as a Type I error.

B.False Accept Rate (FAR)

It is the probability of an impostor being able to successfully gain access to a secured system, which is in statistics is referred to as a Type II error.

Both error rates should ideally be 0%. From a security point of view, type II errors should be minimized, that is no chance for an unauthorized user to login. However, type I errors should also be infrequent because valid users get annoyed if the system rejects them incorrectly.

Many performance results were reported for typing behavior authentication. For example, in 2014 the authors of [11] have used statistical model and static text on mobile devices and got (0.92%) FAR, and (1%) FRR with 315 users. While authors in [17], in 2013 have used dynamic text with statistical model using touch screens and got results (9.78%) FAR and (10%) FRR for 100 users.

On the other hand, for artificial neural networks (ANN) models, authors in [16] reported (0.0277) FAR and (0.0862%) FRR with experiments done on 140 users. While authors in [15] got (9.04 %) FAR and (6.66%) FRR with hybrid model and using dynamic text.

C.Equal Error Rate (EER)

One of the most common measures of biometric systems is the rate at which both accept and reject errors are equal. It is also known as the Cross-Over Error Rate (CER). The value indicates that the proportion of false acceptances is equal to the proportion of false rejections. The lower the equal error rate value, the higher the accuracy of the biometric systems.

Author in [3] reported (2%) EER with 58 users for their work, using dynamic text and with statistical model, while authors in [2] reported (12.82%) EER with 63 users, using static text with ANN model.

VI. TYPING BEHAVIOR AUTHENTICATION

The typing behavior authentication; as most biometric techniques has two phases: 1) the enrolment, where reference features are stored to compare with them at the actual authentication phase. 2) Actual authentication where a user wants to access a system. Both phases are necessary for typing behavior authentication [1] [15]. See fig. 2.

In order to analyze the collected data, there are two main ways, statistical model and learning model. Statistical model compare reference typing characteristic with user typing characteristics. However, learning model uses learning algorithms such as ANN and Genetic algorithms [15] to learn and modify the user identity each time. A third approach is the hybrid which uses both statistical and learning models.

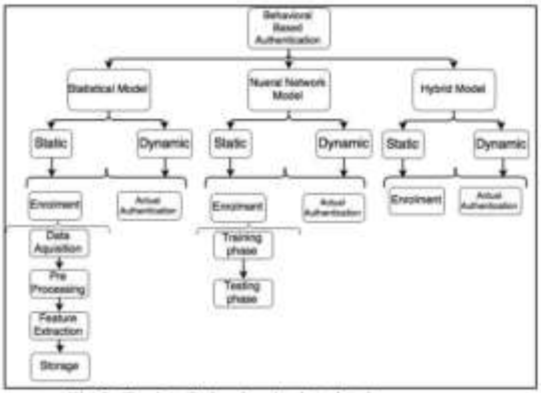


Fig.2: Typing Behavior Authentication

First typing behavior authentication in desktop field have evolved between 1985 and 1990 [1]. Where studies showed a good fault resistance and they were improved over years. With the usage of mobile devices many researchers conducted many experiments on personal mobile digital devices [15] [17], personal computer with their classical keyboards [19], laptops with special keyboards embedded in, and handheld devices with their special keyboards screens [15] [11].

The basic idea of typing behavior authentication approach is to compare a reference set of typing characteristics of a certain user with a test set of typing characteristics of the same user or a test set of a hacker. The distance between these two sets (reference and test) should be below a certain threshold or else the user is recognized as a hacker. This method has two families: static [7] [11] [12] [19] [20] [21] and dynamic [1] [3] [15] [17].

A.Static family:

It is the family in which a user is asked to type several times the same sequence of characters in order to build the model. During the authentication phase, the user is supposed to provide the same string captured at enrollment. A methodology is really suitable to authenticate an individual by asking him to type his own password, before login to a computer session, and verifying if the way of typing matches the model. Changing the password implies to enroll the identity again, because the methods are not able to work with a different password [2] [4] [5].

In static method verification, the keystrokes are analyzed only at specific times e.g., during login. Static approaches provide more robust user verification than simple passwords but it doesn't provide continuous security. Therefore they cannot detect substitution of the user after the initial verification [2] [11] [19].

In statistical approaches for example, authors in [1] in early 1990, proposed a statistical model for identity authentication based on keystrokes latencies. They described a method of verifying the identity of a user based on keystrokes latencies. A static text in early 1990's was sufficient to describe the statistical model; they conducted the experiment on classical keyboards with 33 users in which they have reached FAR (0.25 %) and FRR (16.36%).

Authors in [19] for example, proposed a statistical model that built on top of Manhattan Distance, Manhattan with Standard Deviation Distance and Euclidean Distance. For a fixed password (static which was "kolkata123") they reached near 13.3% EER. They got (0.40) and (0.53) EER for the passwords "password" and "123456" respectively.

Authors in [20] proposed a statistical model based on new distance metric. It was evaluated on the CMU typing behavior authentication benchmark dataset. With 15 users, they got performance ratio of 8.7% EER. CMU [13] consists of the Dwell time for each key and the suggested latencies between two successive keys for suggested static password strings. These informative data helped the researchers to build their model based on unbiased dataset, because sometimes the collected data would be biased and unreal, especially when number of users in the experiment is not enough to do the experiment.

Authors in [21] proposed a statistical model dedicated for cell phone devices. They targeted mobile cloud computing users. With 150 users, they reported FAR of (2.67%).

On the other hand, authors in [8] proposed a statistical simple model mainly based on fixed text for a password with a length of 4 characters.

B.Dynamic family:

Usually, it asks the users to type long text to create their model, or the model is created indirectly by monitoring the users typing rhythm for a certain period. It is called continuous authentication [3] [12] [15].

Authors in [15] for example, did the experiments on different mobile devices such as Samsung S2, HTC desire, and Galaxy Nexus. They focused on several scenarios covering numerical and alphabetic dynamic input data collected from 12-Key layout and old QWERTZ-layout.

While authors in [11] proposed a unique research on continuous authentication on mobile devices and touch screens. By analyzing motion behavior, they proposed a suggestion for using a linear Support Vector Machine SVM which indeed was used to identify the authorized user. They used timely based feature extraction over Android OS devices only. They achieved FAR of (92%) and FRR of (1%).

Authors in [17] proposed a statistical classifier model based on mean absolute deviation for each feature being selected. They did their work on handheld mobile devices with QWERTY keyboard layout. They conducted their experiments with different combinations of features such as (time), (time & size), and (time, size, and pressure). With a 100 volunteered users, they got 9.78% FAR and 10% FRR.

Authors in [3] proposed a statistical model using hidden Markov chain method. This statistical learning model was built for web based authentication and targeting the fixed-text typing behavior authentication analysis. In addition, Fuzzy Logic was used for characterizing the typing behavior. With a dynamic text instead of static one and with 58 users they gained a (2%) ERR with a minimum target password length of 9 characters.

i.Statistical models

1)Enrolment.

The enrolment process has four different steps which are presented in Fig. 2.

a)Data acquisition

Typically typing behavior generates information from the keyboard (hardware or virtual) at the authentication device [11] [20]. In [11], for example, they use a software keyboard for Android mobiles that enabled them to collect the typing behavior of 300 users in a field study, where in [20] they proposed a use of CMU typing behavior authentication benchmark data set. Researchers in [7] [13] are developing an intelligent way for data acquisition based on raw data generated from a benchmark tool such as CMU.

The data are recorded via the operating system and can be stored as a stream of events. After this, the raw data is stored as enrolment samples for later evaluations [1] [15].

b) Pre-processing

Since behavioral features cannot be extracted in the same quality every time, depending on human interaction and his/her behavior, pre-processing has to be done. For example, in [2] researchers preprocessed collected data in the form of vectors in order to prepare for the next step to extract features. Some other researchers for example [1] [13] count this step as an embedded step in the first one and named it as data representation. In which they represent their collected data in weighted and un-weighted features in vectors, matrixes and other forms [13]. Others are dividing collected data into different groups that are classified into two types of typing behavior authentication dynamic text and static text [11].

c) Feature extraction

typing behavior authentication features could be extracted form user behavior and can be classified according to many different mechanisms and models. The features are similar to the features explained in Section IV [8][9] .

Extraction of features is one of the most important steps of behavior authentication. The error rates depend on the selection of the right features. The extraction of characteristic features from the input data is the net result of a well preprocessed and represented collection of features in a pattern vectors as in [2]. This step indeed is very critical; because the number of degrees of freedom of variation in the chosen features across the input data may affect the efficiency of the encoding and the reliability of the identification pattern. However, extraction of more extra features, increases computational complexity of the problem [1] [11].

d) Storage

Authentication attempts are stored in the database [14]. With the storage of these data, changes can be recognized. For example, typing can change over time. This can be recorded to modify the reference data for authentication. The main disadvantage of behavioral features is that they cannot be done hundred percent in the same way (e. g. movements, and speed). This is a big advantage concerning replay attacks, because changes can be recognized with storage of these data, these types of attacks are relying on a known and fixed profile. Typing behavior authentication is a behavioral mechanism which could be changed with time where the system should be adapted for that [1] [2] [15].

Storage includes: classification the data collected at the enrollment stage. This step speeds up the comparison at the actual authentication stage [19].

2)Actual authentication.

In this phase, real users get authenticated and imposters are denied. Keep in mind that behavior authentication has the disadvantage that some people are falsely rejected (FRR) and others are falsely accepted (FAR). [1].

Authors in [3] proposed two modules to authenticate the user: Profile Building module which is the enrollment in our classification, and Authentication Module which is the same step in our classification to features classification and user typing rhythm identification.

Authors in [19] proposed intensive and more accurate steps. They captured, extracted and did the comparison. Then, they have another phase for actual authentication which splits the authentication into two steps: first, password generation step, which made the password at the end unpredictable. They stored an encrypted combination of a password and timing template along with the password in this step. Second step is password verification, which did the comparison between the input one and the stored one.

While authors in [12] in 2014, worked also on statistical approaches. They proposed a new summarizing phases for their model: registration and authentication phases. They use maximum likelihood estimation for the algorithm with FAR of (4%) and (8%) FRR with a password of four characters length only.

Authors in [10] proposed a statistical model that demonstrated a novel feedback and training interface named Mimesis. It provided both positive and negative feedback on the differences between input pattern and a reference one. For a group of 84 participants playing a role of different attackers and 2 eight–characters passwords of different difficulty, the FAR was nearly (0.99) for both passwords

ii.Artificial Neural Network models (ANN)

The second main type of key stroke dynamics modeling is the ANN [16]. This is defined as an adaptive non-linear statistical data modeling tool which have been inspired by biological interconnection of neurons.

In ANN model, researchers first build a prediction model from historical data, and then they use this model to predict the outcome of a new trial or to classify a new observation (i.e. authentication). Although studies tend to vary in what typing behavior information they utilize to the pattern classification techniques they employ, all have attempted to solve the problem of providing a robust and inexpensive authentication mechanism [2][15].

Authentication in ANN is also consisting of two main parts: enrollment and actual authentication. But most researchers divide enrollment in ANN into two main phases: ANN training phase and ANN testing phase. Also the four steps of enrollment still exist. These are: data acquisition, preprocessing, feature extraction, and storage [16].

1)Enrollment in ANN

For ANN models, authors in [22] in early 1993, presented a way to classify inter-character times using ANN. During the investigation phase, three different ANN architectures were used: tested-back propagation, sum-of-products and hybrid sum-of-products. Based on experiments, hybrid sum-of-products was found to perform better than other architectures and achieved FAR equals to (2.2%).

Author in [16], proposed a model for dynamic typing behavior authentication normality statistics, using ANN. Features were extracted and verified by measuring two main features: digraph time and digraph latency. Results were 0.0277% FAR and 0.0862% FRR.

While authors in [15] proposed ANN model with dynamic text, which consists of enrollment process, data acquisition, pre-processing, feature extraction and storage phase. They measured the features consisting of digraphs and trigraph with numerical inputs and alphabetic inputs. They reported two experiments: 12-key keyboard layout with FAR of (9.04%) and FRR (6.66%), and with QWERTZ keyboard layout with FAR (12.13%) and FRR (8.75%).

Authors in [2] for example, depend on using an ANN and they are working on K-nearest neighbors for feature extractions and matching. The hierarchal clustering of feature patterns vector is the main algorithm. They used different major classifiers summed in (Euclidean distance measure, non-weighted probabilities, and weighted probabilities measures). They presented combined measures for classifying and identifying identities of users.

On the other hand, author in [6] suggested use of weightless ANN for classifying users. They scaled the data before discretizing it into linear and non-linear intervals. They also observed that the non-linear intervals gave better results than linear intervals.

2)Actual authentication in ANN

ANN has an advantage that they can handle many parameters. However, they are slow not only during the training but also in the authentication phase. In ANN it is difficult to decide which features are important for classification due to its “black box” mode of operation. As in ANN we are noticing the function results rather that the entire function steps itself. This could be a problem for continuous typing behavior authentication where results typically are required in real time [2] [4].

ANN approach requires a training phase to be added to the classification in statistical, where it is a mandatory phase that should be added due to the need for training at each enrollment step done in ANN model. For this purpose, authors are now working on creating a large user database for comparing and identifying users. Many directions can be taken starting from these points; especially they are moving towards learning models and ANN models.

The author in [7] proposed a neural learning model which has built into two main phases: training phase, in which a user profile is created through repeated entry of password. And testing phase, in which the password of typing rhythm of the user is compared with the stored one in the typing profile. This learning model was used for anomaly detection of typing behavior based on median vector. In which the proximity distance threshold for a feature is chosen to be the value of standard deviation of that feature. The researcher used CMU benchmark to evaluate the dataset. It served the experiments with an independent comparison for the reference data regarding the evaluation of the classifiers and anomaly detection algorithms. This tool is usually used in comparative studies of detection performance.

iii.Hybrid models

The two main major analysis approaches in typing behavior authentication are statistical and ANN techniques. In hybrid, statistical and ANN methods are combined together to provide more security level. Table (2) provides the improvements have been done in this field.

Authors in [16] proposed a hybrid model with statistical and ANN. They proposed a model for static text and got results of FAR (0.0277%) and FRR (0.0862%).

The authors in [18] proposed hybrid techniques with statistical and ANN model targeting cell phone devices only in their experiment. With 15 users, they reported a ratio of (0.806 %) for EER.

VII. FUTURE WORK DIRECTIONS

The latest typing behavior authentication research is moving towards different main paths which can be summarized as follow:

1-Features collection and extractions

Most of the research on feature extraction was on regular keyboards and on specific features such as the H time, UD time, and DD time. The features can be much more as we have seen in Section VI. More features can be added. For example the pressing force especially with touch-screen devices where the force can be measured. In addition, the finger used for typing can be read in touch screen devices and also the angle of pressing the character. Also, the physical distance between characters is important. Moreover, distinguishing between capital letters, special characters, and small letters have some effects on typing behavior.

Another research direction in this area is the adaptability of feature selection. Since we have many features for typing behavior authentication, there is a need for the model to be adaptive where it selects the best set of features that fit specific user. Not all features are used for all users.

2-Typing rhythm types

Most research papers were focusing on static text typing behavior authentication. These days, with the widespread use of cell phones and the problems of losing them, there is more need to go towards dynamic authentication. In dynamic authentication, if somebody logged in to an online service on the cell phone, and then he/she lost the cell phone, dynamic authentication can detect that the current user (the one who found it) is not the owner of the cell phone because there is a difference in typing behavior between the two users. This behavior can only be discovered in dynamic authentication.

3-Modeling algorithms

There is a need for more hybrid modeling algorithms which may merge multiple learning algorithms like ANN, genetic, and instance based learning methods. Also, the models are required to be adaptive where they keep learning after each correct and incorrect logins.

4-Benchmarking

There is a need for benchmarking to compare the different ideas in this area. The CMU benchmark exists. But, the touch screen requires different types of benchmarks.

VIII. CONCLUSION

Password authentication is the most commonly used authentication method for local access, network access, and internet access. However, password authentication suffers from many drawbacks due to password nature. Therefore, some techniques are required to strengthen password. This survey focused on typing behavior strengthening techniques (also called keystroke dynamics).

Majority of the latest dynamic typing behavior authentication researches have been summed up and analyzed in this paper. From the last three decades, typing behavior authentication were being invested starting from classical keyboards and end up with touch screens in digital devices and cell phones, which have replaced the classical keyboards. Our survey is focusing on research results on these devices mainly. We discussed also important features being collected with Statistical and Neural learning models as well as the hybrid ones.

IX. REFERENCES

[1]Joyce, Rick, and Gopal Gupta. "Identity authentication based on keystroke latencies." Communications of the ACM 33.2 (1990): 168-176.

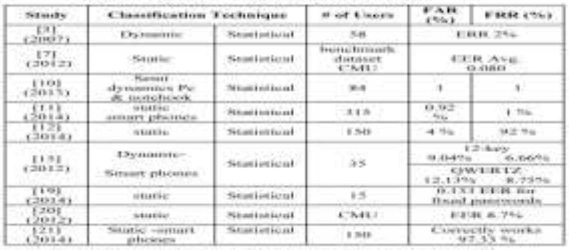
[2]Monrose, Fabian, and Aviel D. Rubin. "Keystroke dynamics as a biometric for authentication." Future Generation computer systems 16.4 (2000): 351-359.

[3]Jiang, Cheng-Huang, Shiuhpyng Shieh, and Jen-Chien Liu. "Keystroke statistical learning model for web authentication." Proceedings of the 2nd ACM symposium on Information, computer and communications security. ACM, 2007.

[4]Karnan, M., M. Akila, and N. Krishnaraj. "Biometric personal authentication using keystroke dynamics: A review." Applied Soft Computing 11.2 (2011): 1565-1573.

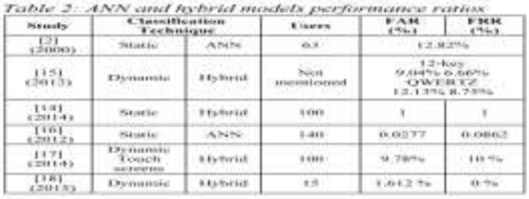
[5]Shanmugapriya, D., and Ganapathi Padmavathi. "A survey of biometric keystroke dynamics: Approaches, security and challenges." arXiv preprint arXiv: 0910.0817 (2009).

Table 1 : Statistical performance ratios



Study	Classification Technique		# of Users	FAR (%)	FRR (%)
[3] (2007)	Dynamic	Statistical	58	ERR 2%	
[7] (2012)	Static	Statistical	benchmark dataset CMU	EER Avg. 0.080	
[10] (2013)	Semi dynamics Pc & notebook	Statistical	84	1	1
[11] (2014)	static smart phones	Statistical	315	0.92 %	1 %
[12] (2014)	static	Statistical	150	4 %	92 %
[13] (2012)	Dynamic- Smart phones	Statistical	35	12-key 9.04% 6.66%	
				QWERTZ 12.13% 8.75%	
[19] (2014)	static	Statistical	15	0.133 EER for fixed passwords	
[20] (2012)	static	Statistical	CMU	EER 8.7%	
[21] (2014)	Static -smart phones	Statistical	150	Correctly works 97.33 %	

Table 2: ANN and hybrid models performance ratios



Study	Classification Technique		Users	FAR (%)	FRR (%)
[2] (2000)	Static	ANN	63	12.82%	
[15] (2013)	Dynamic	Hybrid	Not mentioned	12-key 9.04% 6.66% QWERTZ 12.13% 8.75%	
[14] (2014)	Static	Hybrid	100	1	1
[16] (2012)	Static	ANN	140	0.0277	0.0862
[17] (2014)	Dynamic Touch screens	Hybrid	100	9.78%	10 %
[18] (2013)	Dynamic	Hybrid	15	1.612 %	0 %

[6]Peacock, Alen, Xian Ke, and Matthew Wilkerson. "Typing patterns: A key to user identification." IEEE Security & Privacy 2.5 (2004): 40-47.

[7]Mudhafar, M. "Al-Jarrah, an Anomaly Detector for Keystroke Dynamics Based on Medians Vector Proximity." Journal of Emerging Trends in Computing and Information Sciences VOL3 6 (2012).

[8]Bajaj, Swarna, and Sumeet Kaur. "Typing Speed Analysis of Human for Password Protection (Based On Keystrokes Dynamics)." no 2 (2013): 88-91.

[9]Schlenker, Anna, and Milan Sarek. "Behavioural Biometrics for Multi-Factor Authentication in Biomedicine." EJBI 8.5 (2012): 19-24.

[10]Tey, Chee Meng, Payas Gupta, and Debin Gao. "I can be you: Questioning the use of keystroke dynamics as biometrics." The 20th Annual Network & Distributed System Security Symposium (NDSS 2013), 2013.

[11]Gascon, Hugo, et al. "Continuous Authentication on Mobile Devices by Analysis of Typing Motion Behavior." Sicherheit. 2014.

[12]Alves, D. D., G. Cruz, and C. Vinhal. "Authentication system using behavioral biometrics through keystroke dynamics." Computational Intelligence in Biometrics and Identity Management (CIBIM), 2014 IEEE Symposium on. IEEE, 2014.

[13]Bhatt, Shanthi, and T. Santhanam. "Keystroke dynamics for biometric authentication—A survey." Pattern Recognition, Informatics and Mobile Engineering (PRIME), 2013 International Conference on. IEEE, 2013.

[14]Chandrasekar, V., V. Shanmugavalli, and P. Krishna Sankar. "Multimodel Biometric Authentication Based on Finger Print and Keystroke Dynamics Using Fuzzy Set." Australian Journal of Basic & Applied Sciences 8.7 (2014).

[15]Trojahn, Matthias, and Frank Ortmeier. "Biometric authentication through a virtual keyboard for smartphones." International Journal of Computer Science & Information Technology (IJCSIT) 4.5 (2012).

[16]Abernethy, Mark, and Shri Rai. "Applying Feature Selection to Reduce Variability in Keystroke Dynamics Data for Authentication Systems." (2012).

[17]Tasia, Cheng-Jung, et al. "Two novel biometric features in keystroke dynamics authentication systems for touch screen devices." Security and Communication Networks 7.4 (2014): 750-758.

[18]Dhage, Sudhir, et al. "Mobile authentication using keystroke dynamics." Communication, Information & Computing Technology (ICCICT), 2015 International Conference on. IEEE, 2015.

[19]Roy, Soumen, Utpal Roy, and D. D. Sinha. "Enhanced knowledge-based user authentication technique via keystroke dynamics." Int. J. Eng. Sci. Invention (IJESI) 3.9 (2014): 41-48.

[20]Zhong, Yu, Yunbin Deng, and Anil K. Jain. "Keystroke dynamics for user authentication." Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on. IEEE, 2012.

[21]Babaeizadeh, Mahnoush, Majid Bakhtiari, and Mohd Aizaini Maarof. "Authentication Method through Keystrokes Measurement of Mobile users in Cloud Environment." Int. J. Advance Soft Compu. Appl 6.3 (2014).

[22] M. Obaidat and D. Macchiarolo. “An online neural network system for computer access security”. IEEE Transactions on Industrial Electronics, 40(2):235 –242, Apr. 1993.

[23]Raza, Mudassar, et al. "A survey of password attacks and comparative analysis on methods for secure authentication." World Applied Sciences Journal 19.4 (2012): 439-444.

[24]Halakou, F. Feature Selection in Keystroke Dynamics Authentication Systems.

[25]Banerjee, Salil P., and Damon L. Woodard. "Biometric authentication and identification using keystroke dynamics: A survey." Journal of Pattern Recognition Research 7.1 (2012): 116-139.
[26]Maren, A. J., Harston, C. T., & Pap, R. M. (2014). Handbook of neural computing applications. Academic Press.
[27]Li, C. T., Weng, C. Y., & Fan, C. I. (2012). Two-factor user authentication in multi-server networks. International Journal of Security and Its Applications, 6(2), 261-267.

------------------------- xxxxxxxxxxxxxxx -------------------------