Key-Stroke Pattern-based Behavioral Bio-metric Authentication system

Mudit Rathore

Indian Institute of Information Technology Allahabad, ICM2015502

***Abstract*— This is the era of information, where every byte of data has its own vitality and importance. Access to information is the symbol of power and the pathway to success. It has become an integral part of our lives. But this new Information age has opened gates for new threats. Thus in order to protect the information, we must understand the key concepts and methods of digital information security. One such security process is behavioral user authentication, the advancements in technology to recognize behavioral characteristics will bring about a new dimension to security in general as the defining traits of an individual, his/her behavior can neither be inherited or stolen. In this paper, we examine the emerging behavioral bio-metric authentication technique based on the key-stoke pattern.**

I. INTRODUCTION

Data is the fuel that drives almost everything. Computers control communication, businesses, finances, and every other field. We are generating over 2.5 Quintilian bytes of data every single day, that is over 1.7 MB of data every second. Being extensively dependent on the internet, we have every sort of information over the internet, including sensitive information relating to national security, individual finances, healthcare, personal life, etc. However, this has unveiled new threats to the computer system and information security. The cases of fraud are increasing day by day. It has become challenging to innovate new solutions for identification and authentication of users.

We have many user authentication techniques that are knowledge-based, object-based or bio-metric based. The biometric authentication system includes physiological and behavioral characters. Using keystroke dynamics seems to be a natural choice because it requires no end user training, no additional hardware, it is easy to deploy and the typing pattern or rhythms becomes almost intrinsic to users behavior with prolonged use of a computer.

Keystroke dynamics is the process of analyzing the way a user types at a terminal by monitoring the keyboard inputs thousands of times per second and attempts to identify them based on habitual rhythm patterns in the way they type. In this is paper we present our techniques to implement Key-stroke based behavioral authentication system.

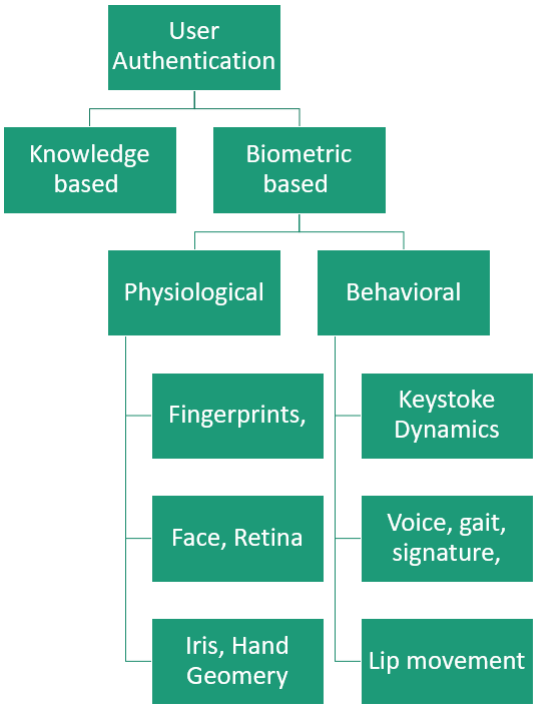

Fig. 1. User Authentication techniques

A. Biometrics

Biometrics, the physical traits and behavioral characteristics that make each of us unique, are a natural choice for identity verification. Physical characteristics like fingerprints, tongue prints, retina, iris, are good candidates for verification because they are unique traits of the individuals. Behavioral characteristics, on

the other hand, have some physiological basis, but also reflect a person's psychological state. Biometric technologies are defined as "automated methods of verifying or recognizing the identity of a living person based on a physiological or behavioral characteristic". Biometrics provides an extra level of security as these cannot be lost, stolen or imitated and these can also be used with the traditional method for verification. Currently, the available measures to overcome the authentication problems are:-

1) Something a person knows.: Passphrases or passwords fall under this category, in this method of authentication, a person needs identifying himself through a user id and authenticate himself via passwords.

2) Something a person physically possesses.: Identification cards fall under these categories, in this method, a physical entity that identifies a person using the details helps in authenticating the person.

3) Characteristics of a person: Physical traits and Behavioural characteristics of an individual are used for identification in this category. The first two measures can easily be compromised by stealing or copying the necessary attribute. But the physical traits of a person can neither be inherited nor be copied. The behavior of a person depends upon his/her neurophysiological characteristics which are impossible to imitate, thus making the false authentication very difficult. Thus, hereon our primary focus must be to integrate the conventional ways of authentication to the behavioral aspects of identity. With the evolution and improvement in technology, behavioral authentication will eliminate the need for physical identification and will provide a higher level of security. In this paper, we will analyze one of the behavioral authentication techniques which are based on Key-stroke dynamics.

II. Key-Stroke Dynamics

Key-stroke dynamics are a part of the behavior of a person. Because of the widespread use of a keyboard as an input medium, no external hardware is needed for implementation.

This revolves around the pattern recognition of the time-series of an individual user, the features for identification might include the following

- words-per-minute count
- Error rate
- Key-hold time
- Key seek time
- Method of error correction
- Frequency error

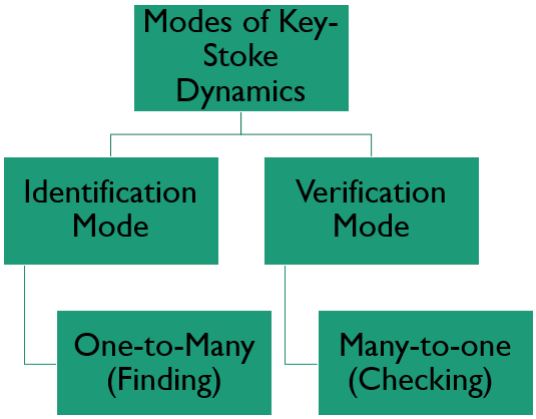

Fig. 2. User Authentication techniques

For automated pattern recognition system involves

- Representation
- Extraction
- Classification

1) Representation: The representation of the input data measures the characteristics of the pattern to be recognized.

2) Extraction: Here, we have to preprocess the data so that it yields the features and reduces the dimensionality of the data-set. The complexity of the algorithm should be kept in mind because too many features can increase the computational cost without much gain in the extraction of the features. Thus, there is a trade-off between the computational-power and the desired accuracy.

3) Classification: Classification involves the determination of the optimum decision procedures. In this phase, we try to classify and label the data for further training of the model.

III. Methodology

Keystroke dynamics is the process of analyzing the way user types by monitoring the pattern or rhythm to identify users based on his behavior linked with neurophysiological characteristics. Typing rhythm is a good authentication technique as it adds an increased level of security and is also cost-efficient.

A. Data Selection and representation

The database is created using a paragraph that is input by the user. This paragraph consists of all the English alphabets in upper as well as lower case. Each user has to input 7 paragraphs in order to get the typing rhythms and patterns for the recognition algorithm to run. The general form of the data entry starts with

running a python script at the back end. The user has to input his/her name. The prompt screen will then ask the user whether he is ready to enter the paragraph and the paragraph will pop up for the user to type. After the paragraph is entered, a scatter plot will be generated and the data will be exported in CSV format to a file named "Nametime" the name will be consisting of the first name and the time will be last 5 digits of the system timer. Each of the users can run the script at their convenience and enter the paragraph. The graph that will be obtained by running the script on one of the paragraphs is shown.

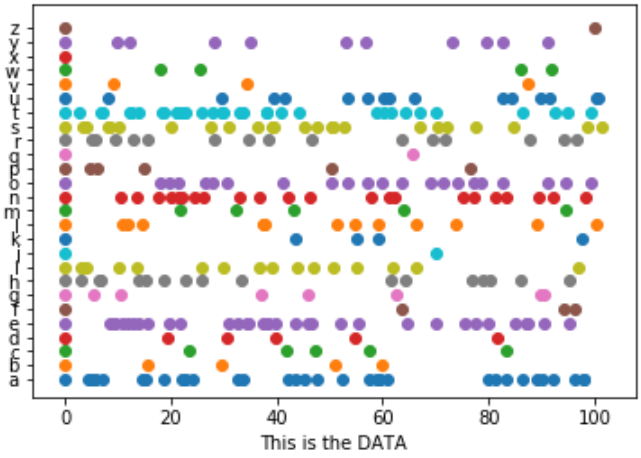

Fig. 3. User Authentication techniques

B. Keystroke Verification Techniques

The Keystroke verification techniques use one of the two modes to verify, they are

1) Static Verification Techniques: This is the verification technique in which we use the fixed text mode. Here, we are prompted to enter a predefined text by the system. For example, this type of verification can be based on the password typing rhythm, or authentication at the login time using static paragraphs.

2) Dynamic Verification Techniques: This is the verification technique in which we use the dynamic text mode. Here, we are not given any text, we generate the pattern taking into consideration whatever the user is typing. It can be continuous or periodic monitoring of the user during the sessions as well. This type of verification technique does not need a predetermined text string, which may be complex for the user to remember.

C. Data Extraction

We have used Neural Networks for the extraction of the patterns and information about the data from user input. Because Artificial Neural Network(ANN) uses the processing style of the brain as a basis to develop learning models, it can be used to solve complex patterns and prediction problems. Also, ANNs have the ability to learn and model non-linear and complex relationships. ANNs can generalize — i.e. after learning from the initial inputs and their relationships, it can infer unseen relationships on unseen data as well, thus making the model perform well even on unseen data. ANNs do not impose any restriction over the input variables. It assigns weights to features as per their importance in the prediction process. This model also learns from error on every data point to improve overall accuracy. This is achieved by "Backpropagation".

1) Advantages of Keystroke Dynamics: The advantages of using the keystroke dynamics are

- Based on neuro-physiological characteristics, thus cannot be stolen, lost or imitated.
- Only Software Requirement, No costly application specific hardware required.
- Simple to Deploy
- Cost Effective
- No end user training
- Simple and natural way to increase computer security

2) Disadvantages of Keystroke Dynamics: The disadvantages of using the keystroke dynamics are

- User susceptibility fatigue.
- Dynamic change in the typing pattern.
- Injuries might affect the typing patterns.
- Typing patterns might change with the hardware i.e. keyboard
- Requires adaptive learning
- Access is granted if typing pattern of a user matches within a particular threshold of the other user but falsely claimed.

IV. Summary

In this paper, we presented the importance of using biometric authentication using one of the behavioral characteristics i.e Key-stoke dynamics. Keystroke dynamics is the process of analyzing the way users type by monitoring keyboard inputs and authenticating them based on habitual patterns in their typing rhythm. we asked the user to create a data set for himself and then used the Artificial neural network for the classification. We have considered that the behavioral traits

of a person are their neurophysiological characteristics. thus, these are very difficult to imitate, these cannot be lost or stolen and hence are more secure. We also presented that these provide a more robust model of authentication of a user, that the conventional methods used. These techniques have quite a few advantages and some major drawbacks as well, Cost-effective, robustness, less prone to attack are few advantages and adaptive learning, frequently changing typing patterns or rhythms and mood of the user can add to false rejection of the user.

REFERENCES

[1] "On Neural Networks for Biometric Authentication Based on Keystroke Dynamics," Sensors and Materials, p. 385, 2018.

[2] F. Monrose and A. D. Rubin, "Keystroke dynamics as a biometric for authentication," Future Generation Computer Systems, pp. 351–359, 2000.

[3] B. A. Forouzan, Introduction to cryptography and network security. Boston: McGraw-Hill Higher Education, 2008.

[4] "Keystroke dynamics," Wikipedia, 19-Mar-2019. [Online]. Available: https://en.wikipedia.org/wiki/Keystrokedynamics.

[5] R. Giot, M. El-Abed, and C. Rosenberger, "Keystroke Dynamics Overview," Biometrics, 2011.

[6] Towards Data Science. (2019). Introduction to Neural Networks, Advantages and Applications. [online] Available at: https://towardsdatascience.com/introduction-to-neural-networks-advantages-and-applications-96851bd1a207 [Accessed 14 Apr. 2019].

[7] "Neural Network Classification," solver, 28-Dec-2015. [Online]. Available: https://www.solver.com/xlminer/help/neural-networks-classification-intro. [Accessed: 14-Apr-2019].