Keystroke Dynamics for User Verification

Ashwini Sridhar and H. R. Mamatha

Abstract With the evolution of internet, the dependency of humans on them has increased. This has led to an increase in attacks, forgery, impersonation and so on, which require that a user and his privacy be maintained. Thus the need to protect a user has increased intensifying protection, authentication and verification methods of a user. There are many methods of authenticating a user, which include traditional methods of authentication such as passwords, personal identification numbers and so on, However, these methods have their drawbacks and hence biometrics have replaced these methods in some cases and in some cases biometrics has turned out be an additional layer of security, therefore providing better security. In this paper we propose one of the behavioral methods of biometric authentication called keystroke dynamics which uses a user's typing rhythm to verify a user. One of the most common examples of this method is the verification of user using CAPTCHA, where the user is asked to type the letters to be verified as a genuine user and thus the user's typing rhythm is captured based on which a match is generated and the user is verified. This method is most commonly used in applications such as online banking, email verifications and other such areas. This method acts as an additional layer of security to an existing system and helps protect the sensitive information of the user.

Keywords Biometrics · User verification · Keystroke dynamics · Authentication · Security · Typing rhythm

A. Sridhar (✉) · H. R. Mamatha
PES University, Bengaluru, India
e-mail: asashwini38@gmail.com

H. R. Mamatha
e-mail: mamathahr@pes.edu

V. Suma et al. (eds.), *Evolutionary Computing and Mobile Sustainable Networks*,
Lecture Notes on Data Engineering and Communications Technologies 53,
https://doi.org/10.1007/978-981-15-5258-8_14

1 Introduction

For over many decades, the combination of username/password has been used for protecting electronic information systems and services. Although there are variations to this, like usage of email address or user ID instead of username, the fundamental concept has remained the same.

The combination of username/password for securing information systems is nearly 50 years old. This method was at first developed in 1961 at MIT (Massachusetts Institute of Technology) and has been in use thereon for securing most of the online services that comprise email service, banking systems and so on. Figure 1 presents a traditional authentication system.

However, due to availability of modern commodity hardware systems with better processing and storage capacity, it is becoming easier for hackers to crack the password. Hence the research community in the security domain has been working on novel type of authentication and authorization system for securing the systems.

Biometric authentication has replaced the traditional authentication method. There are two types of biometrics: physical and behavioral. This work focuses on a behavioral-based biometric called keystroke dynamics. Keystroke dynamics is the analysis of a user's typing pattern based on which a user can be verified as genuine or not. The basic features usually collected are keydown-keydown time, hold time and keyup-keydown time. Figure 2 presents these features.

This work adopts the method of keystroke dynamics as a means to verify the genuity of a user. This method thus provides better protection and an additional layer of security when combined with the traditional methods. It is also proved to be a strong method of authentication when used alone. The rest of the paper is divided as follows: Related works are presented in Sects. 2; Sect. 3 outlines the methodology and implementation used; the results and discussions are depicted in Sect. 4; Sect. 5 and Sect. 6 represents the conclusion and future work respectively.

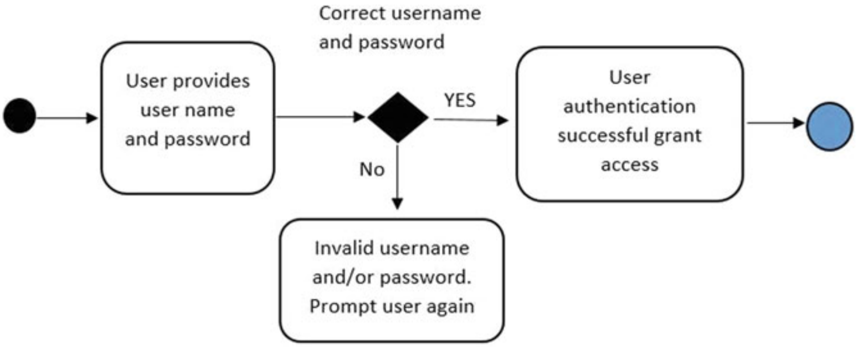

Correct username and password
User provides user name and password
YES
No
User authentication successful grant access
Invalid username and/or password. Prompt user again

Fig. 1 Traditional authentication system

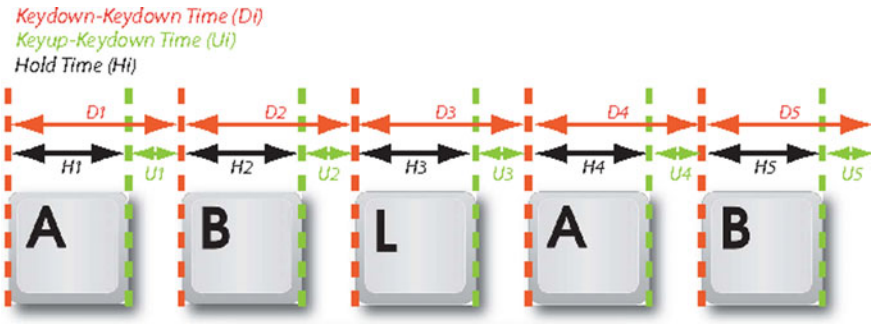


Fig. 2 Basic features of keystroke [6]

2 Related Works

Keystroke dynamics analysis has been done by different people in different ways, and each of them have arrived at their own results. This section describes in brief the work done by different people and the algorithms used to analyze the keystroke patterns of a user.

According to the work done by Killourhy and Maxion, comparing anomaly detection algorithms [1] states the best performing algorithms based on the equal error rate which was calculated on the dataset collected consisting of the user's typing patterns. The dataset comprised 51 users, which were then evaluated for a total of 14 different classifiers.

Keystroke dynamics has proved to be a wide field of research and a lot of studies have been conducted recently. There are many parameters that are taken into consideration while considering a user's typing pattern. In [2] the author talks about such parameters. This work also focuses on increasing the reliability of authentication of a user and hence makes use of keystroke dynamics as a biometric method.

The work done in the field of keystroke dynamics consists of multiple features and methods of evaluation. The work by Abdullah et al. [2] talks about an algorithm called dynamic time warping (DTW) which makes use of waveforms in order to arrive at a suitable estimation of performance.

The traditional methods of authentication make use of passwords, PINs and so on as a method of authentication. However, with the evolution of technology, it was observed that these methods of authentication alone do not provide enough security for the user data. Hence to improve the security, keystroke dynamics is used as an additional layer of security. The author in [3] includes keystroke dynamics as an additional layer of authentication to the traditional password-based authentication. The anomaly scores are calculated by using various distance-metric algorithms such as Manhattan distance and Mahalanobis.

With the increase in risk to security everyone requires a safe, quick and trustable source of communication. This requires protection of data by means of authentication. The work done by Maheshwary et al. [4] describes the method of safe, quick and trustable source of communication. The work makes use of keystroke dynamics as a method of authentication, and this is done by using the nearest neighbor algorithm.

3 Methodology

In this work of verifying a user based on their keystroke dynamics, we have studied the performance of different algorithms. The general methods followed are: data loading, data selection, training, testing and calculation of equal error rate (EER).

In the data loading phase, the data are loaded from a text file into the system. These data are then split into training and testing sets, where the first 15 vectors are for training and the rest for testing. The data are then trained where the current user and his data are taken as a genuine user data for training and the rest of the data are treated as imposters.

For the test phase, user and imposter scores are calculated. If the score is high, it is proved that the user is not genuine. EER is calculated as the total number of incorrect predictions divided by the total number of values in the dataset. 0.0 and 1.0 are considered to be the worst and the best error rates, respectively. It can be represented as follows:

$$\mathrm{EER} = \frac{\mathrm{FP} + \mathrm{FN}}{\mathrm{P} + \mathrm{N}}$$

where FP: false positive, FN: false negative, P: positive, N: negative

Accuracy (ACC) can be calculated as the number of correct predictions to the total number of values in the dataset. 0.0 and 1.0 are considered to be the worst and the best accuracies, respectively. It is calculated as

$$\mathrm{ACC} = 1 - \mathrm{EER}$$

Therefore, it can be concluded that better the EER, better the accuracy.

We have used various classification systems in order to measure the performance of the algorithms on the system and verify a user. The algorithms used are Manhattan scaled distance [1], nearest neighbor Mahalanobis [1], outlier count [1], K-nearest neighbor (KNN) [5], recurrent neural network (RNN) [6], dynamic time warping (DTW) [2], convolutional neural network (CNN) [6] and decision tree.

3.1 Implementation

The accuracy and effectiveness of the authentication system depend on the input dataset used. The dataset should comprise large data in order to successfully verify a user's identity. For the current project we have collected a dataset from 78 users,

Table 1 Data collected

Subject	H.period	DD.period.t	UD.period.t	H.t	DD.t.i	UD.t.i
Ashwini	0.078	0.531	0.453	0.078	0.281	0.203
Ashwini	0.063	0.391	0.328	0.078	0.266	0.188
Ashwini	0.079	0.547	0.468	0.078	0.188	0.11
Ashwini	0.062	0.546	0.484	0.079	0.282	0.203
Ashwini	0.063	0.5	0.437	0.094	0.157	0.063
Ashwini	0.078	0.437	0.359	0.094	0.125	0.031

each of them typing the password used in [1], “tie5roanl”, 30 times. Table 1 presents the dataset collected, which represents the timing data of each key press. The basic features include keyup-keydown time, which is the time between release of one key and the press of next; keydown-keydown is the time between continuous key presses; and the hold time which is the time between the press and release of each key. These features are collected for each letter of the password. In order to increase the efficiency of the algorithm, attributes like age, gender, trigram and bigram time are also added.

This dataset is evaluated using different detection algorithms like KNN, RNN, CNN, and the top performing algorithms used in [1] that are Mahalanobis and Manhattan scaled.

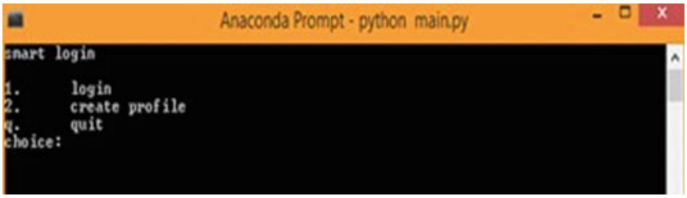
In Table 1 the first column represents the subject, that is, the user; the second column represents the hold period duration for the password typed by the user where each row represents the password typed by the user once. The third column represents the keydown-keydown period, and the fourth column represents the keyup-keydown period. The rest of the columns represents the hold time (H time), keydown-keydown time(DD time) and keyup-keydown time (UD) for each letter in the password typed by the user.

The dataset was collected with the help of a console-based application that was developed. Figure 3 presents this application. There were two options provided in the application:

1. Login: In this option the user is asked to login with his username and password which is used to authenticate the user.
2. Create profile: This option is used whenever a new user profile has to be created in order to collect the features. Once this option is selected the user is asked to type his username and the password which are stored in the text file.

This file is then used as the basis for authentication of a user.

In order to provide accurate results, it is important that we have the right functional, data and system requirements. Since it is based on machine learning algorithms used, it is important that we have enough data. Hence we collected a total of 2500 keystrokes. The general requirement includes a Windows or Linux OS with 4 or 8 GB RAM with suitable python environment and packages. We used Python 2.7 environment along with the scipy, numpy packages.



```
smart login

1.      login
2.      create profile
q.      quit
choice:
```

Fig. 3 Dataset collection application

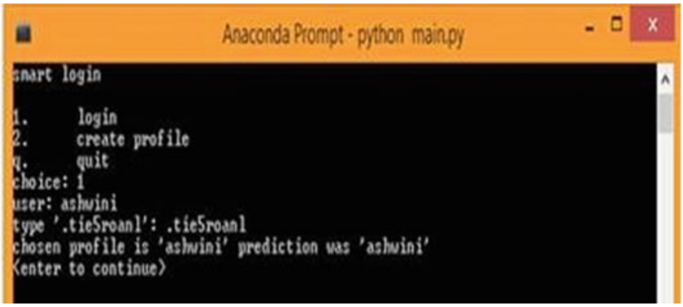
4 Results and Discussions

The data of approximately 80 users was collected. Figure 4 presents the authentication of a user.

The initial accuracy of the data seemed good. Table 2 presents the initial accuracy. As the number of users increased for the data when tested, it seemed to decrease the accuracy.

Hence the number of users was reduced with each turn and the accuracy was tested. Table 3 presents the variance in the accuracy for the dataset as the users are reduced.

In spite of reducing the users, it was seen that the maximum accuracy obtained was 50% for 40 users. Hence there was a need to re-evaluate the same data with additional features and algorithms to achieve better accuracy.



```
smart login

1.      login
2.      create profile
q.      quit
choice: 1
user: ashwini
type '.tie5roanl': .tie5roanl
chosen profile is 'ashwini' prediction was 'ashwini'
<enter to continue>
```

Fig. 4 Authentication of a user

Table 2 Initial accuracy

User Id	No. of times password entered	Correct	Wrong	Accuracy (%)
1	10	6	4	60
2	10	8	2	80
3	10	9	1	90
4	10	4	6	40

Therefore we added attributes such as age, gender and trigram time. The system performance was also measured with other algorithms, such as K-nearest neighbor (KNN), recurrent neural network (RNN), convolutional neural network (CNN), dynamic time warping (DTW) and decision tree classifiers.

After the addition of new features, we re-evaluated the algorithms. Table 4 presents the evaluation done based on the equal error rate (EER).

The first column in Table 4 represents the top performing algorithms based on the work done by Killourhy and Maxion [1]. The second column represents the EER results obtained in the benchmark dataset, that is, the evaluation done in [1]. The last column represents the evaluation based on EER for the dataset collected. Based on this EER, the top performing algorithms were established.

Similarly, the other algorithms such as K-nearest neighbors (KNN), recurrent neural networks (RNN), convolutional neural networks (CNN) and decision tree were evaluated. Table 5 summarizes the algorithms used with their accuracy.

Table 3 Variance in the accuracy of data

Total number of users in the dataset	No. of times password typed by a single user	No. of correct predictions	Accuracy
78	10	3	30
56	10	3	30
43	10	4	40
40	10	5	50

Table 4 Top performing algorithms based on EER

Algorithms	EER of benchmark dataset	EER of dataset collected
Manhattan scaled	0.17681169012847736, 0.10416561236462442	0.0674319757690701, 0.04717726419913844
Nearest neighbor (Mahalanobis)	0.3063765640274061, 0.10974436873298003	0.10893319797198889, 0.18134704000510013
Outlier count	0.13730802044159202, 0.09555389158452095	0.06919125305634136, 0.06404148467115772

Table 5 Accuracy of algorithms evaluated

S. No.	Algorithm	Accuracy (%)
1	KNN	90
2	RNN	85
3	CNN	2
4	Decision Tree	7

Dynamic time warping presents the comparison between two waveforms of a user. Figure 5 presents the peak comparison of a single user. From Fig. 5 it is observed that the peaks of a single user vary each time the user inputs the password. This is because of the key press and typing rhythm of the user which also varies with each input. Figure 6 presents the peak comparison of different users. From Fig. 6 it can be observed that peaks of each user vary due to the difference in the typing rhythm as well as the key press durations.

In order to arrive at the best performing algorithm, it is important that the factors like false positive and true positive be considered. This helps in determining the accuracy of a system. Thus it leads to an appropriate conclusion. Table 6 presents the false positive and true positive for all the algorithms used in the evaluation.

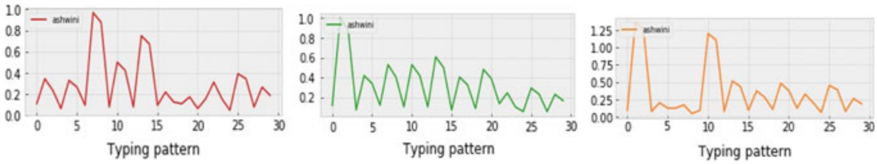


Fig. 5 Peak comparison of one user

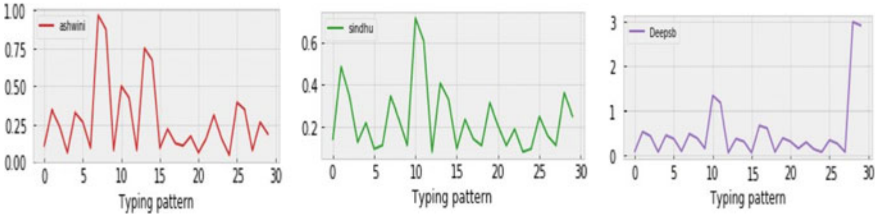


Fig. 6 Waveform Comparison of different users

Table 6 False positive and true positive for algorithms

Algorithm	Number of samples	False positive	True positive
Manhattan scaled	77	9	68
Nearest neighbor (Mahalanobis)	77	43	34
Outlier count (Z-score)	77	8	69
CNN	474	463	11
RNN	15	5	10
Decision tree	712	473	39

On the basis of our analysis, it was found that the best performing algorithm based on the equal error rate (EER) from Table 4 when compared with the benchmark dataset is Manhattan scaled algorithm. However, the outlier count and nearest neighbor (Mahalanobis) were found to be the second and third best when compared to the benchmark dataset. This may be due to slight variations in the data collected.

Of the algorithms in Table 5, KNN was found to be the most accurate algorithm, while RNN was slightly less accurate in comparison to KNN. The algorithms CNN and decision tree were found to be the least accurate algorithms with accuracy of below 10%.

Hence from Table 5 it can be concluded that CNN and decision tree algorithms are not suitable for time series data because CNN requires a large amount of multidimensional data collected over a long period of time for each individual in order for it to be thoroughly trained and tested. The data we have collected here are not enough. Therefore they do not produce accurate results and cannot be used.

It is therefore clear that the performance of the verification systems depends on the data collected and the features used. The performance of the algorithm, as well as the accuracy also, depends on the data and the features. Thus it can be concluded that the dataset and the features play an important role.

The proposed method of user verification using keystroke dynamics when compared to the existing techniques provide better security in terms of user privacy, verification, imposter user and other such things. In the techniques that are usually used, such as authentication through passwords, it becomes easy for an imposter to impersonate the password and the user's passwords and PINS can be hacked easily.

Keystroke dynamics acts as an additional layer of security protecting the user's privacy and user information as the typing and key press rhythm of each user is different. The difference in the typing and key press of each user makes this method better when compared to the traditional methods of security, and thus it is impossible for an imposter to impersonate the user. Hence keystroke dynamics proves to be one of the most preferred methods of user verification.

5 Conclusion

As witnessed in the design and implementation of verification of a user using keystroke dynamics, it can be concluded that the experiment is successful in achieving the targeted application feature.

The goal of authenticating a user based on the user's keystroke dynamics by building a security application has been successfully achieved. It was observed that as the number of users increased, the accuracy decreased. Hence it was necessary that different algorithms be applied and additional features be added in order to improve the efficiency of the system.

Based on Sect. 4 from Tables 4 and 5, the best performing algorithms were found to be Manhattan scaled and KNN with an accuracy of 88.3 and 90%, respectively, while the least performing algorithms were found to be decision trees and CNN with an accuracy of 7 and 2%, respectively.

The user verification method proposed in this work can be used as an additional layer of security for many applications, such as banking, various transactions and other such areas, therefore improving user authenticity, genuity and thus help preserve user security.

6 Future Work

This work is only limited to desktop applications and makes use of basic features, such as keyup-keydown time, keydown-keydown time, hold time and trigram time. It can be further extended to other computing devices such as smart phones and tablets with the addition of features suh as right handed or left handed etc.

References

1. Killourhy KS, Maxion RA (2009) Comparing anomaly-detection algorithms for keystroke dynamics. In: 2009 IEEE/IFIP International conference on dependable systems & networks. https://doi.org/10.1109/dsn.2009.5270346
2. Sulavko AE, Eremenko AV, Fedotov AA (2017) Users' identification through keystroke dynamics based on vibration parameters and keyboard pressure. In: 2017 IEEE dynamics of systems, mechanisms and machines (dynamics) (Omsk, Russia) 14 Nov–16. https://doi.org/10.1109/dynamics.2017.8239514
3. Abdullah A, Frans C, Danushka B (2016) Towards keystroke continuous authentication using time series analytics, Springer International Publishing AG 2016 M. Bramer and M. Petridis (eds.), Research and Development in Intelligent Systems XXXIII, https://doi.org/10.1007/978-3-319-47175-4_24
4. SoumenRoy,Utpal Roy, D. D. Sinha, September 2014. Enhanced Knowledge- Based User Authentication Technique via Keystroke Dynamics. International Journal of Engineering Science Invention ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726 www.ijesi.org Volume 3 Issue 9 || September 2014 || PP.41–48.33
5. Lu X, Zhang S, Yi S (2018) Continuous authentication by free-text keystroke based on CNN plus RNN. In: 2018 International Conference on Identification, Information and Knowledge in the Internet of Things, IIKI 2018, 147, pp 314–318, https://doi.org/10.1016/j.procs.2019.01.270
6. Venugopalan S, Juefei-Xu F, Cowley B, Savvides M (2015) Electromyograph and keystroke dynamics for spoof-resistant biometric authentication. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). https://doi.org/10.1109/cvprw.2015.7301326
7. Saket M, Vikram P (2016) Mining keystroke timing pattern for user authentication, Springer International Publishing AG 2017 A. Appice et al. (Eds.): NFMCP 2016, LNAI 10312, pp 213–227. https://doi.org/10.1007/978-3-319-61461-814
8. Obaidat MS, Macchairolo DT (1994) A multilayer neural network system for computer access security. IEEE Trans Syst Man Cybernet 24(5):806–813. https://doi.org/10.1109/21.293498