Three-layer Authentication in Keystroke Dynamics using Time based Tool

Namisha Bhasin\textsuperscript{1} and Sandhya Tarar\textsuperscript{2}

\textsuperscript{1}Research Scholar, School of ICT, Gautam Buddha University, Uttar Pradesh, India
\textsuperscript{2}Professor, School of ICT, Gautam Buddha University, Uttar Pradesh, India

Email: namishabhasin@gmail.com

Abstract: Having a secure and sound system is the most important need of the end-user. Confidential and authentic information about the system must be available to the genuine user when required. This paper presents a generalized Keystroke Dynamics Technique for identification of genuine users. The method works for the authentication of the user while user is entering the password to using the system. It is a three-layer approach which first check the typing pattern while entering the password then it also monitors the system while user is using the system. Different users have different typing pattern which could be used to recognize a user. For identification of the user a time-based tool is used to collect data pertaining to the typing time of each user for words of different lengths. This is a very easy and cost-effective way of collecting data for differentiating between a genuine user and impostor. At first layer elbow method is used to know unknown targets depending upon different word combinations. For the second layer principal component analysis (PCA) is used to find suitable factors where user typing pattern is making users indistinguishable. For the third step Long-Short Term Memory (LSTM) technique is used to forecast whether the user is a genuine user or an impostor.

Keywords: static text, PCA, time-based tool, elbow method, LSTM

1. Introduction

Nowadays even though systems are highly developed they still contain many flaws; which makes the system easily accessible for the attackers and that too with help of cheap tools. Till date various approaches are working on multiple theories to make the system secure. Some researchers use the static technique to find virus attacks based on specific keywords for example packet with name Trojan Horse will be considered as a malicious one. Since this method does not work for all the scenarios there was a need to identify some other failproof method. One such method is the behaviour technique in which the behaviour of the code is checked to find the attack. Many difficult problems can be solved by noticing the behaviour because that is the thing which is difficult to change.

There are various techniques used to make the system secure like face recognition, iris detection etc. But these techniques have various limitations as with age facial features change preventing face recognition from working properly or a person might be suffering from some eye ailment due to which iris detection method doesn’t work. Human beings are different. Some activities are involuntary while others are not. Every person has a different way of walking (gait activity) which is involuntary. Human nature always resists changes and has some habits, which are very difficult to change, one of them is the way in which a user presses a key. Using the same concept, it was found that users generally have a typing pattern which is very difficult to change. Identifying this pattern can be used to conclude whether the system is being used by a genuine user or an impostor.

One of the behavioural techniques is Keystroke Dynamics which is used in identifying the users depending upon their typing pattern. Users have their own way of using the keyboard. Some users are good in using(pressing) a combination of keys while others are not. Some users have habit of using keyboard with one hand while other use with both hands. Some press the keys forcefully while others use soft hands. So, by combination of all these features a very low-cost technique can be deployed where typing pattern is calculated only by knowing the difference in time which is taken for pressing the keys. This is an inexpensive method which does not need any extra hardware. As the system is
able to provide required results up to nanoseconds precision [1], having such high precision ensures that the system cannot be misled and fails any attempts to copy the pattern. This makes the system more secure by accurately recognizing the typing pattern [2].

For keystroke dynamics many factors are considered like hold time, dwell time, flight time while pressing the keys. It is found that many systems take a lot of factors into consideration to increase the accuracy of detection which in turn also leads to increase in computational time. It was found that using certain relevant factors can maintain the same accuracy level of the systems but result in a reduction in computational time [3]. Hence identification of relevant features has an important role in computation time. This research paper considers both type of data: static and dynamic. In static approach users type the same word and typing time is calculated for different users at different sessions; while in dynamic approach users are free to type words of any length. Based on their typing time they are classified as either imposter or genuine user. The data is further normalized which can be used for first layer where elbow method is applied to know the number of words if dynamic approach is followed. For static approach also elbow method is used as a word is further broken into different combinations. In second layer PCA technique is implemented to ensure only those features are used which are more relevant. This technique is applied to both dynamic and static approach to reduce the calculation time. In the third layer, deep learning technique LSTM is used to train the model to know whether the user is a genuine user or not. The outline of the paper will be as: (2) Literature Review, (3) Design and Process, (4) Experimental Setup, (5) Conclusions and future work.

2. Literature Review

In this paper for further enhancement thorough study has been carried out of all the previous work done in this field. Zhong and Deng found that combination of keystroke dynamics with other biometric techniques, like fingerprint scan or face recognition, is able to give strong, complete and more authentic solution [4]. Baynath et al. proposed the ant colony optimization as the feature subset classifier and Euclidean distance as the classifier where the recognition rate is 83 % with FAR (False Acceptance Rate) of 0.15% and FRR (False Rejection Rate) of 2% [5]. Giot and Rocha used the Siamese network, which is a neural network and used to find how one input differs from other. For the first phase which is for enrolment they used the samples provided by existing users and avoided samples from new users; and for the second phase used for authentication where the given data was compared with enrolment sample which provided Equal Error Rate of 28% in a one-shot context and 31% while using 200 enrolment samples [6]. Liu and Deng used Elbow and K-means clustering method to find the number of clusters [7]. PCA was performed to extract the features from the given feature set [8]. In this paper the gap between formulas for the LSTM network and RNN was found by considering two changes 1) the training formulas were omitted altogether; 2) unfolding an RNN was presented without justification. They derived the canonical RNN formulation from differential equations. It was analysed that storing the information over extended time intervals by recurrent backpropagation method was taking a very long time. The reason was insufficient, decaying error backflow. Then they introduced an efficient, gradient based method called long short-term memory (LSTM) and truncate the gradient so that it does not produce errors [9][10][11]. In the proposed paper author considered 1) the fuzzy ARTMAP which is an incremental supervised learning technique used for recognizing arbitrary analog or binary input vector; 2) radial basis function networks (RBFN) used as an activation function; and 3) learning vector quantization (LVQ) - a supervised learning technique based on neural network used to classify the pattern in inter-key time-based approach. Obaidat and Sadoun found that other neural network and classical pattern algorithms such as backpropagation with sigmoid activation function, sum-of-products (SOP), hybrid sum-of-products (HSOP), potential function and Bayes’ rule algorithms provide average results [12]. Bleha et al. used two types of passwords 1) phrases which are fixed and used for the identification of the system; and 2) individual names which were used as a password in the verification system and in the overall recognition system. All three systems were tested and evaluated with an error rate of 3.1% in rejecting valid users (FRR).
and 0.5% in accepting invalid users (FAR) [13]. Araujo et al. in his work shows results with the false rejection rate of 1.45% and false acceptance rate of 1.89% while considering all the features [14].

3. Design and Process
Time based calculations are used to calculate the net time. The difference between ending and starting time of the words with different combinations of sentences is calculated as shown in algorithm.

```
| Do                         |
|---------------------------|
| Do                        |
| Start time                |
| Type anything             |
| Extract time-based        |
| information for each word |
| Extract time-based        |
| information for each      |
| character                 |
| End time                  |
| Net time = (End Time -    |
| Start Time)               |
| While user is typing      |
| Use Elbow Method to       |
| Create the cluster        |
| Use Principal component   |
| analysis to perform       |
| dimension reduction       |
| Apply LSTM algorithm      |
| While other users are typing |
```

**Algorithm**

For dynamic and static text various combination of words is typed here for calculating the time. Some combinations are repeated over different time intervals in different situations. The same experiment is repeated with different users. For the first layer which is used in dynamic text “elbow” method is used to calculate the optimal number of clusters for K from K-Means. The clusters are calculated based on a range of values which is based on the time taken for typing different words. In the line chart which forms a curve like an arm and the point at which the two lines seems to connect is known as elbow point. It is used to ascertain the optimal number of clusters. where after curve model fits best at that point [7]. For second layer PCA is used to reduce the number of features which make computation fast and make the user incomparable for both static and dynamic text. Deep learning technique Long short-term memory is an artificial recurrent neural network is used for both static and dynamic text. LSTM can process entire sequence of data with the help of feedback connections. They can remember patterns for a long duration of time. A LSTM is composed of 1) input gate used to add information to the cell state, 2) output gate used to find important information from the current cell which is having given set of information, 3) forget gate is used to remove information from the cell state which is not required or less important and 4) a cell used for remembering values for a time span. LSTMs can deal with the exploding and vanishing gradient problems that can be encountered with traditional RNNs [16] [17] [18] [19].

4. Experiment Setup
In this paper analysis is performed for both static and dynamic text.

4.1. Dynamic text: For dynamic text different combination of words are typed and their time is calculated in nanoseconds as represented in Table 1.
Table 1: Different combination of words whose time is calculated in (ns).

| User | Two | Two | Three | Four | Five | Six | Seven |
|------|-----|-----|-------|------|------|-----|-------|
| U1   | 136 | 149 | 161   | 148  | 171  | 211 | 210   |
| U1   | 153 | 86  | 165   | 147  | 193  | 190 | 225   |
| U2   | 133 | 90  | 173   | 121  | 169  | 198 | 221   |
| U2   | 187 | 125 | 201   | 122  | 164  | 184 | 203   |
| U3   | 141 | 87  | 149   | 120  | 127  | 178 | 172   |
| U3   | 133 | 95  | 128   | 231  | 206  | 206 | 194   |
| U4   | 127 | 132 | 180   | 150  | 178  | 180 | 218   |

In pre-processing these values are normalized as shown in Eq 1 and represented in Table 2

\[
\text{normalized values} = \frac{(x - x_{\text{min}})}{(x_{\text{max}} - x_{\text{min}})}
\]  

Table 2: Normalization of values

| User | Two | Two | Three | Four | Five | Six | Seven |
|------|-----|-----|-------|------|------|-----|-------|
| U1   | .015| 1   | .45   | .25  | .56  | 1   | .72   |
| U1   | .043| 0   | .51   | .24  | .83  | .36 | 1     |
| U2   | 0.1 | .06 | .62   | .009 | .53  | .61 | .92   |
| U2   | 1   | .62 | 1     | .018 | .47  | .18 | .58   |
| U3   | 0.23|.016| 1.04  | 0    | 0    | 0   | 0     |
| U3   | 0.1 | .14 | 0     | 1    | 1    | .85 | .42   |
| U4   | 0   | .73 | .71   | .27  | .65  | .06 | .87   |

Depending upon the data from Table 2 “elbow” method is used to calculate the numbers of unknown targets. As in the Fig 1. shows number of targets formed are seven as no of combination of words taken are 2,2,3,4,5,6,7.

![Fig. 1: Number of clusters using elbow method.](image-url)
Depending upon the data from Fig. 1 Clusters are formed shown in Fig. 2.

**Fig 2:** Clusters formation depending upon type of words.

For second layer PCA is applied to find the features which make more impact as compared to others used. For this covariance matrix is calculated as shown in Fig 3. It is clear from Fig 4 that feature 1 and 2 have more impact as compare to other features.

**Fig 3:** Covariance matrix for dynamic text.
Fig 4: Principal Components analysis.

Features 1 and 2 are selected for further processing. For the third layer LSTM is applied and results can be seen in Fig 5.

Fig 5: LSTM result for dynamic text.

4.2 Static text: For static text Dataset is taken from www.cs.cmu.edu/~keystroke/ which is in normal form [15] and plotted as shown in Fig 6.
At first layer of authentication elbow methods is used as static word is further broken into different combinations as shown in Fig 7.

**Fig 6**: Plotting the data value for static text.

**Fig 7**: Elbow method for static text to know number of targets.
In second level of authentication PCA is applied for feature extraction. For this covaraince matrix is shown in Fig. 8.

Fig 8: Features representing the correlation to each other in static text.
PCA as shown in Fig 9 results are as follows:
Principal component: [0.26321579 0.16902472 0.05995721 0.05714044 0.05250818 0.04690555
0.0435502 0.04158057 0.03934555 0.03519659 0.03174237 0.02815475 0.02255776 0.02098883].

![Fig 9: Principal components for static text.](image)

In third layer of authenticity LSTM (Long-Short Term Memory) is applied and results are shown in Fig. 10.

![Fig 10: LSTM Result for static text.](image)

After considering 14 features there is a loss of 9.2% of data. If Training Loss>Validation Loss then results are underfit or if Training Loss<Validation Loss then results are overfit but if Training loss=Validation Loss that is a best fit.

As experiment is performed with features [H.Period, DD.period,t, DD.t.i, H.t, DD.i.e, UD.i.e, DD.five.Shift.r, DD.Shift.r.o, UD.Shift.r.o, DD.o.a, UD.o.a, DD.a.n, DD.n.l, UD.n.l,H.l] and resulting curves are approximate to each other. Hence, they are near to best fit.
5. CONCLUSION AND FUTURE WORK
This paper addresses the problem of making the system secure with the help of monitoring keystrokes at both static and dynamic level. It is a generalized approach which works at both levels. Though systems are secured by password but in case the password is cracked by the attackers they will not be able to duplicate the typing pattern of the genuine user which can be recognized by the proposed approach. In the same way the proposed approach monitors the system while user is using it and any change in the typing pattern can be recognized. Hence, works well with dynamic text too. The results are able to differentiate between different typing pattern and hence identifying users’ basis the same. Results also show that they are being followed well for different users while using LSTM with a loss of 9.2% of accuracy. In future we want to enhance the system’s functionality such that it can produce results with a decrease in the computation time and an increase in accuracy.

References
[1] Senk C and Dotzler F 2011 Biometric authentication as a service for enterprise identity management deployment: a data protection perspective. In 2011 Sixth International Conference on Availability, Reliability and Security (pp. 43-50). IEEE.
[2] Teh PS et al 2013 A survey of keystroke dynamics biometrics. The Scientific World Journal, 2013.
[3] Patil RA and Renke AL 2016 Keystroke dynamics for user authentication and identification by using typing rhythm. International Journal of Computer Applications, 144(9), 27-33.
[4] Zhong Y and Deng Y 2015 A survey on keystroke dynamics biometrics: approaches, advances, and evaluations. Recent Advances in User Authentication Using Keystroke Dynamics Biometrics, 1-22.
[5] Baynath et al 2016 Implementation of a Secure Keystroke Dynamics using Ant colony optimisation. In The International Conference on Communications Computer Science and Information Technology Vol. 2016.
[6] Giot R and Rocha A 2019 Siamese Networks for Static Keystroke Dynamics Authentication. In 2019 IEEE International Workshop on Information Forensics and Security (WIFS) (pp. 1-6). IEEE.
[7] Liu F and Deng Y 2020 Determine the number of unknown targets in Open World based on Elbow method. IEEE Transactions on Fuzzy Systems.
[8] Lever et al 2017 Points of significance: Principal component analysis.
[9] Sherstinsky A 2020 Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network. Physica D: Nonlinear Phenomena, 404, 132306.
[10] Hochreiter S and Schmidhuber J 1997 Long short-term memory. Neural computation, 9(8), 1735-1780.
[11] Du et al 2015 Hierarchical recurrent neural network for skeleton based action recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1110-1118).
[12] Obaidat M S and Sadoun B 1997 Verification of computer users using keystroke dynamics. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 27(2), 261-269.
[13] Bleha et al 1990 Computer-access security systems using keystroke dynamics. IEEE Transactions on pattern analysis and machine intelligence, 12(12), 1217-1222.
[14] Araújo et al 2005 User authentication through typing biometrics features. IEEE transactions on signal processing, 53(2), 851-855.
[15] www.cs.cmu.edu/~keystroke/
[16] Hochreiter S JA1 4 rgen Schmidhuber 1997“Long Short-Term Memory”. Neural Computation, 9(8).
[17] Graves A et al 2008 A novel connectionist system for unconstrained handwriting recognition. IEEE transactions on pattern analysis and machine intelligence, 31(5), 855-868.
[18] Sak et al 2014 Long short-term memory recurrent neural network architectures for large scale acoustic modeling.
[19] Li X and Wu X 2015 Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition. In 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 4520-4524). IEEE.