

An Study of User Validation using Keystroke Dynamics

K. Vinodhini, Dr. Rama moorthy

¹Research Scholar, ²Dean

^{1 2} Department of Computer Applications, Dr. M.G.R. Educational and Research Institute, Chennai

Abstract - Nowadays while to identify the user authentication to access the validation system by using key stroke patterns. This is supported to biometric based data authentication. In this examine is to find the efficiency of key stroke dynamic to ascertain the user identity. This study is reorganization users while user entering a value through keyboard using key stroke pattern. This paper is used novelty is authorizing by using keystroke data. In this study use the novelty of machine learning methods and to the performance of individual methods as well as the novelty performance both concerning accuracy and time complexity. Similarly, the problem is used to apply equally to other pattern identification methods and to optimize a set of investigated feature for user identification via keystroke dynamics.

Keywords: User Coupling, Keystroke Dynamics, Biometrics, User Identification

1. Introduction

In recent years to developing highly to protected authentication techniques in system security is a most important task. Herculean attempts have been taken to advance security of computer systems. The user preventing unauthorized access is user Authentication. Among them the process of authentication is differentiate between the genuine user and fraud user. User authentication is the process of verifying claimed identity. It is a challenging area of security in research and practice. Many methods were discussed till date but none of them is completely secure. The authentication is accomplished by matching some short-form indicator of identity, such as a shared secret that has been pre-arranged during enrollment or registration for authorized users. Usually, the client verification and validation is classified into three sub classes, there are Token base or Object based, Knowledge based and Biometric based.

Authentication is the process to verify, that a person is what he claims to be. A user can be verified in one of the following ways Password: Something a user knows like a password or a PIN and Token: Something a user has like smart card Biometric attribute: Something you are. Biometric attribute can be of two types-physical (retina scan, finger scan etc) and behavioural (signature, keystroke dynamics etc)

2. Review of Literature

Most of the papers used statistical methods and neural networks for keystroke based authentication.[3] proposed a Monte Carlo approach for data collection and parallel decision tree (DT) for identifying the genuine user. Data collection included six basic parameters by comparing key press and key release of successive keys. A vector formed on the basis of raw

data. Wavelet analysis was performed on four 16-element sub vectors by splitting the keystroke feature vectors and eight DT classifiers were trained for every user.

Mrs. D. Shanmugapriya and Dr. G. Padmavathi discussed about the study in Keystroke Dynamics is one the biometric methods, authors tried and implemented to identify the genuineness of a user while the user is key in a keyboard. The verification process is done by observing the change in the typing pattern of the user. A complete survey of the various approaches of existing keystroke dynamics techniques its metrics. This research also discusses about the various security issues and challenges faced by keystroke dynamics.

Abdullah Osamah AL-Rahmani Enhanced the Classifier of Authentication in Keystroke Dynamics Using Experimental Data in the problem of cyber-attacks on information systems and networks, for various illegal purposes, has become a major threat to society and individuals. Computer hackers are using all possible means to get access to private data, or to destroy such data. It has become necessary to improve computer security through more advanced access control mechanisms. Recently the use of biometrics has been employed to strengthen access control through user authentication that is based on users' measureable features.

Baljit Singh Saini and etal explained a survey of Keystroke Dynamics for Mobile Phones. The Biometric is the science of authenticating a user based on his physical or behavioral attributes. Keystroke dynamics is behavioral study which analyses the typing rhythm of the user. We adopted a systematic procedure for studying the state of the art in keystroke dynamics in mobile phones. We analyzed the features extracted, the classification techniques, the input text, length of the input text, number of users, hardware used and the results.

3. Methodology

Keystroke dynamics characteristically consists of the examination of characteristics such as duration of a key press or group of keys and the latency between consecutive keys i.e. time elapsed from one key to a subsequent key. Figure 1 show that in generally all keystroke-dynamics evaluations involve

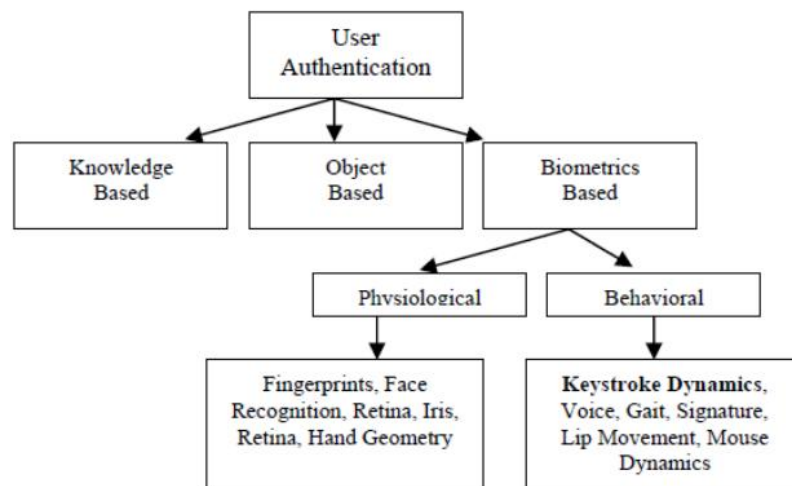


Figure-1: User Authentication / Identification

4. Keystroke Dynamics

Keystroke dynamics is a behavioral biometric which identifies users based on their typing rhythm. The rhythm of each and every user is identified by extracting the features like key hold time, latency and etc. Identifying a user based on key entry goes back to the days of second World War where the operators or users identified each other by “fist of the sender”. Keystroke dynamics has obvious advantages of No extra hardware is required, No extra effort. Typing is what a user will obviously and Continuous authentication can be done for the entire session till the user is typing.

Use of keystroke dynamics in mobile phones offers some advantages over hard keyboards used in PCs

1. Use of two thumbs or one index finger only. While typing on hard keyboards we use all our fingers but while using mobile we normally use one index finger (right / left hand) or both our thumbs.
2. The variation in the typing pattern is less in different. Positions that we might take during typing e.g. typing message on mobile while sitting or while lying on bed will alter the pattern less as we can handle the mobile in almost the same manner in both the situations. But this is not the case with hard keyboards

5. Evaluation

The study discussed and implemented with existing keystroke dynamics methods. The first suggested use of keystroke characteristics for identification appeared in 1975, but observations

about the uniqueness of an individual typing characteristics stretch as far back as the end of the 19th century. Telegraph operators at the time could often identify each other by listening to the rhythm of their Morse code keying patterns. Let's look at some of the pertinent and interesting ways in which existing keystroke dynamics algorithms implemented and applied by Killourhy and Maxion indicated that the top performers are the classifier using scaled Manhattan distance, with an equal error rate of 0.087, and the nearest neighbor classifier using the Mahalanobis distance with an equal error rate of 0.15 on their keystroke dynamics benchmark dataset. In the following section we discuss the advantages of both Manhattan distance and Mahalanobis distance to understand why they succeed in matching keystroke dynamics patterns. We also point out their limitations. A new distance metric is then proposed to combine the benefits of these two distance metrics while overcoming their limitations.

Mahalanobis Distance

Euclidean distance has been the default distance metric for its simplicity and geometrical intuitiveness. However, it has two main limits:

- ❖ It is very sensitive to scale variations in the feature variables
- ❖ It has no means to deal with the correlation between feature variables

Mahalanobis distance, on the other hand, takes into account the covariance of data variables to correct for the heterogeneity and non-isotropy observed in most real data. The squared Mahalanobis distance between two feature vectors x and y is defined as

$$\|x - y\|^2 = (x - y)^T S^{-1} (x - y)$$

where S is the covariance matrix of the data. It not only weights the distance calculation according to the statistical variation of each feature component, but also decouples the interactions between features based on their covariance matrix to provide a useful distance metric for feature comparisons in pattern analysis.

Manhattan Distance

The Manhattan distance metric, also called L1 distance or city block distance, is defined as follows:

$$\|x - y\|_1 = \sum_i |x_i - y_i|$$

The Manhattan distance has the advantages of simplicity in computation and easy decomposition into contributions made by each variable. Most importantly, it is more robust to the influence of outliers compared to higher order distance metrics including Euclidean distance and Mahalanobis distance.

A New Distance Metric

With the above two method discussion, it is easy to understand why keystroke biometrics using Mahalanobis distance and Manhattan distance outperformed other algorithms including some of the more advanced machine learning techniques. The keystroke dynamics features consist of both dwell and latency timings to have large variations in individual components. The feature variables tend to interact with each other as well. These evident scale variations and feature correlations are handled well using the Mahalanobis distance metric.

Accuracy

After acquiring the features, a profile is built for every user during enrolment phase. Then during authentication, the features are again extracted and biometric is built. This biometric is then compared with the existing profile using a matching algorithm. Typically, there are few metrics describe in biometric classifier to performance with regard for accuracy

FRR: false rejection rate the percentage of valid (genuine) user attempts identified as

$$\text{imposters } FRR = \frac{\text{Number of wrong rejections}}{\text{Total number of genuine attempts made}} * 100$$

FAR: false acceptance rate the percentage of imposter access attempts identified as a

$$\text{valid users } FAR = \frac{\text{Number of wrong accepces}}{\text{Total number of intruder attempts made}} * 100$$

ERR: equal error rate the crossover point at which FRR equals FAR. For identification systems accuracy is the measure of performance

In much of the literature regarding keystroke dynamics, the imposter pass rate (IPR) and FAR are used interchangeably with FAR and FRR as defined here, respectively. The existence of two terms whose meanings are opposite but both denoted by FAR can cause some confusion. We've adopted the former terminology throughout this article.

Although ERR is a desirable metric in terms of its ability to condense FAR and FRR into one value, few researchers report ERR in their published results. For that reason, we present an alternative approach to combining FAR and FRR: averaging the two values. We call this value the average false rate (AFR). Empirically, AFR closely approximates ERR for those few papers that did report ERR.

6. Keystroke Dynamics Classifier

Keystroke dynamics based authentication as a one-class classification problem which learns a technique for a client, discards difference to the studied model as imposters, and accept inliers as the genuine user. Although the use of negative examples in training could significantly increase

the truthfulness of the classifier, it is impractical to assume prior knowledge about the keystroke features from imposters, let alone the availability of their training data.

Table 1 : Results of Classifier Algorithms

Algorithm	Equal-error rate	Zero-miss false-alarm rate
Manhattan	0.087 (0.060)	0.423 (0.269)
Mahalanobis	0.096 (0.069)	0.482 (0.273)
New distance metric (Nearest Neighbor)	0.100 (0.064)	0.482 (0.273)

keystroke biometric algorithms outperform existing detectors reported in Shown in bold in the table are the average equal error rate (with the standard deviation shown in brackets) and the zero-miss false-alarm rate of our two keystroke dynamics algorithms: Nearest neighbor classifier with the proposed new distance metric for keystroke dynamics features, and NN classifier using the new distance metric with additional outlier removal in training phase. We also show the performances of the top two performers for either of the error categories reported. The proposed new distance metric is shown to be advantageous in handling the challenges intrinsic to the keystroke dynamics data by reducing both errors

7. Conclusions

We studied the characteristics of keystroke dynamics for computer user authentication and proposed a new distance metric which decouples correlated data, normalizes feature variations, and suppresses outliers. As outliers and data correlations are typical in keystroke dynamics data, it is not surprising that classifiers using the new distance metric outperform existing top performing keystroke dynamics classifiers which use traditional distance metrics. Although we applied the new distance metric to the problem of matching keystroke dynamics features, it is a general distance metric that can be applied to any distance computation in feature vector spaces where the traditional Mahalanobis distance is applicable, with the additional advantage of robustness to outliers.

References

- [1] K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition", IEEE Trans. on Circuits and Systems for Video Technology, vol. 14, pp. 4–20, Jan 2004
- [2] Y. Li, B. Zhang, Y. Cao, S. Zhao, Y. Gao and J. Liu, "Study on the Beihang Keystroke Dynamics Database", Int'l Joint Conf. on Biometrics (IJCB), pp. 1-5, 2011.
- [3] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, Handbook of Fingerprint Recognition, Springer, NY, 2003
- [4] S. Cho, C. Han, D. H. Han, and H. Kim. "Web-based keystroke dynamics identity verification using neural network", Journal of Organizational Computing and Electronic Commerce, 10(4):295–307, 2000.

- [5] S. Prabhakar, S. Pankanti, and A. K. Jain, “Biometric Recognition: Security and Privacy Concerns”, IEEE Security and Privacy Magazine, Vol. 1, No. 2, pp. 33-42, 2003.
- [6] F. Bergadano, D. Gunetti, and C. Picardi, “User Authentication through Keystroke Dynamics”, ACM Trans. Information and System Security, 5(4), pp. 367–397, 2002.
- [7] Romain Giot, Baptiste Hemery, and Christophe Rosenberger. Low cost and usable multimodal biometric system based on keystroke dynamics and 2d face recognition. In ICPR, 2010
- [8] Yu Zhong, Yunbin Deng, and Anil K. Jain. Keystroke dynamics for user authentication. In CVPR, 2012.
- [9] Kevin S. Killourhy and Roy A. Maxion. Comparing anomaly-detection algorithms for keystroke dynamics. In DSN, 2009.

