Journal of Engineering and Applied Sciences 13 (Special Issue 5): 4604-4608, 2018
ISSN: 1816-949X
© Medwell Journals, 2018

Framework for Authenticating Smartphone Users Based on Touch Dynamics

Megha Prabhakar, B. Lakshmi Priya and A.G. Hari Narayanan
Department of Computer Science and IT, School of Arts and Sciences, Amrita University, Kochi, India

Abstract: Mobile phones have become an integral part of everyone's personal and professional life. Besides providing all the basic functions of a phone, most of the smart phone devices are also used to store critical information regarding a user like passwords, photos, bank account information, etc. Storing and accessing critical information and identifying legitimate users is becoming a growing challenge these days. Hence, authentication should be enforced for all mobile devices to identify legitimate users. We have based our study on behavioral biometrics specifically, touch dynamics. It and can be used to provide increased level of security when authentication is concerned and protection of mobile devices. The objective is to develop an authentication framework which uses a hybrid classification algorithm to classify legitimate users and imposters and to analyze the efficiency of the framework with existing ones. The process is performed on a publicly available touch dynamics dataset and important feature extraction is performed on the same. The accuracy of the scheme is measured on the basis of Equal Error Rate (EER).

Key words: Authentication, behavioral biometrics, touch dynamics, extraction, legitimate, available

INTRODUCTION

Smart phones are becoming easily accessible and affordable and thus a whopping increase has been seen in their usage. As protecting user data on these devices is of great importance, user authentication is increasingly necessary to prevent illegal attacks by unauthorized users. Different behavioral traits like fingerprint, swipe, etc. could be retrieved from the way a person uses a phone and these can be validated in order to properly authenticate a genuine user. Authentication can fall into any of the following three categories.

Knowledge based authentication: Validates a user based on something the user knows, e.g., a password.

Token based authentication: Validates a user based on something which the user possesses, e.g., a credit card.

Biometric based authentication: Validates a user based on something the user is physically or behaviorally, e.g., fingerprint, voice.

The first two kinds are easy to implement and provides a better level of accuracy but they are vulnerable to shoulder-surfing, smudge attack, brute force attack, theft, etc. Biometrics on the other hand is exclusive to a user and cannot be easy copied or stolen.

Biometrics can be either physical or behavioral. Examples of physical biometrics include palm, fingerprint, iris, heartbeat, gait, etc., Keystroke dynamics, mouse dynamics and touch dynamics are examples of behavioral biometrics. Behavior based biometrics can continuously authenticate users by creating normal profiles of user behavior. If the system observes differences between the profiles and current user activities it gives an alarm as illegal use. Biometric features are said to be unique because of their exclusive nature no two persons will have the same fingerprints or the same retina pattern. The sensors required to support a biometric system can be easily incorporated in an existing system thus reducing the overhead of installing additional hardware devices. They provide transparent and continuous authentication and are generally cheaper than physiological authentication. However, behavioral biometrics is dependent on the user's behavior and hence, it is inconsistent. Among the two biometric techniques, behavioral biometrics is the one that shows the finest identification performance.

Literature review: Several research works have progressed in the area of mobile phone authentication using biometrics. One of the reesarch done by Milton Ganesh *et al.* (2016) surveys various popular authentication schemes for smartphone authentication

Corresponding Author: Megha Prabhakar, Department of Computer Science and IT, School of Arts and Sciences, Amrita University, Kochi, India

with their potential advantages and disadvantages. They present a review of various research works done for each authentication scheme and compares them for providing a better understanding regarding the area. Open challenges associated with different authentication techniques including data collection, pre-processing, classification algorithms etc. are analyzed. The researcher also suggest hand waving authentication scheme as one area were few research works have progressed.

Researcher research done by Mondal and Bours (2015) uses different authentication approaches in an uninterrupted setting to validate the execution of a biometric system. They have used a publicly available touch dynamics database for their study and suggests that the same approaches could be used for biometrics other than touch dynamics. They test all fusion of techniques including threshold settings, score boosting, static vs. dynamic trust model and fusion techniques to achieve a performance which is proved to be better than the other previous research done on the same dataset. They use a combination of two classification algorithms (SVM and ANN) and plots the performance of the system in a continuous manner using DET curves and EER (Equal Error Rate) values.

Alariki *et al.* (2016) proposed an authentication framework that mainly focused on feature extraction. The framework uses feature selection scheme and random forest classification. User data regarding swipe patterns was collected by developing an android application and SFS wrapper algorithm was run on it to extract the most important features. The research claims to have achieved 91.67% accuracy in feature selection. Their research demonstrated that extracting more relevant features had a significant effect on the accuracy of the system. Further, classification is done based on these features. The performance of the proposed system was analyzed using ROC curve which plotted an EER of 0.0833. The research work also offers a comparison of various other research done in the same area.

Yet, researcher survey (Meng *et al*., 2015) provides a comparison of various biometric authentication schemes for mobile phones and focuses their study on touch dynamics. The research proposed a multimodal authentication framework (connecting either biometrics based or non-biometrics based techniques) for smartphones providing both PIN based as well as touch gesture based authentication to improve system accuracy. For this, they have collected swipe data using an android application, extracted 22 features and used a hybrid PSO-RBFN algorithm for classification. The results show that the touch dynamics based authentication achieved an average FRR of 4.76% and an average FAR of 3.82%, respectively and the multimodal framework as a whole is expected to achieve an average FRR and FAR of 0.0000431 and 0.0000223%, respectively. The study proved that a multimodal framework (PIN+touch gesture based authentication) could indeed significantly reduce error rates and improve performance of the authentication scheme.

Another framework proposed by researchers Alariki and Manaf (2014) based on touch dynamics provided a simple authentication mechanism using SVM for classification. The process is accomplished in three phases, enrollment, training and verification. The enrollment phase requires users to provide swipe features and the training phase extracts these features and stores them as a reference template. The authentication outcome is made at the third phase, based on a matching process of the newly presented features to the pre-stored reference templates. The study also suggests the use of artificial intelligence based algorithms for classification.

Antal and Szabo (2016) in their study analyzed touchscreen data collected through a psychological questionnaire and concluded that a minimum of 5 consecutive swipes are required to identify genuine users. Several one-class and two-class classifiers were evaluated and an EER value of 0.05 was achieved for single swipes. This was improved by implementing 5 consecutive swipes to achieve a better EER value around 0.002.

Meng *et al.* (2012) Particle Swarm Optimization (PSO), a Genetic algorithm is used along with a neural network classifier on a 21 feature dataset to authenticate users efficiently. PSO has been chosen to deal with the variations in user's usage patterns. The research work analyses performances of different classification algorithms including Naive bayes, back propagation neural network, radial basis function network, decision tree (s) and Kstar and found that RBFN classification algorithm has the best realization with an average error rate of 7.71%. Further, they have found that the system achieved an average error rate of 3% by analyzing the performance of the combined PSO-RBFN refined algorithm.

MATERIALS AND METHODS

Classification: Classification is a data mining functionality that categorizes collection of data into one or more classes. The main goal of classification is prediction and in order to predict a certain outcome, classification algorithms are used. A classification algorithm assorts a stack of data having a set of attributes. It identifies the relationship between the attributes of the training set and validates it using the test set.

Naive Bayes classification: Naive Bayesian is a supervised learning algorithm based on conditional probability. It uses Baye's theorem and is a statistical classifier. Baye's theorem is given below:

$$P(X \text{ given } Y) = P(Y \text{ and } X)/P(Y)$$

$$P(X|Y) = P(Y|X)\,P(X)/P(Y)$$

Where:

$P(X|Y)$ = The posterior probability of X conditioned on Y
$P(X)$ = The prior probability of X
$P(Y|X)$ = The posterior probability of Y conditioned on X
$P(Y)$ = The prior probability of Y

Naive Bayesian classifier research as follows:

- Let D be a training set containing several tuples represented as $X = (x_1, x_2, x_3, ..., x_n)$ for n attributes, respectively, $A_1, A_2, A_3, \ldots, A_n$
- Suppose that there are m classes, the algorithm tries to classify X into a class having the highest posterior probability

Radial basis function network: Radial Basis Function Network or RBFN is a neural network classification algorithm which performs classification by measuring the similarity of the input to prototypes or examples from the training set. This similarity is measured by calculating the distance between the input and prototypes from the training set and then the input is categorized into a class based on its resemblance to the class prototype.

Support vector machine: Support Vector Machine or SVM classifier categorizes inputs by defining a separating hyper plane in a multidimensional space. The objective of SVM classifier is to generate an optimal hyper plane that effectively classifies inputs into various classes. SVM classifies inputs using the inner product of any two given observations. Inner product is calculated as the sum of products of each pair of input values.

Particle swarm optimization: PSO is an algorithm inspired from bird flocking behavior. It is an optimization algorithm that can be used to revise a training set of data to improve the performance of the solution. It is similar to genetic algorithms but has evident differences from GAs and evolutionary approaches.

Every particle in the search space is associated with a pbest and an lbest value, representing the best solution any particle has achieved so far and the best value any particle has achieved among all its other neighbors in the search space. A global best solution gbest is achieved when a particle takes all other particles as its potential neighbors. The gbest value is the best solution achieved in the entire search space. After every iteration, the velocity and position of each particle gets updated. If the cost of the calculated position of a particle is less than its current pbest value, the pbest value is updated with the calculated position. And if the calculated pbest after every iteration is less than the gbest value then the gbest value is updated accordingly.

Data base: Touch dynamics refers to the patterns of touch or strokes done on a smartphone. We have chosen to base our study on a publicly available touch dynamics data set as creating a new data set is complicated and out of scope of our study. The data set contains 16 attributes and has been generated using an android application for data collection.

Proposed framework: The proposed study aims at classifying swipe gesture data taken from a publicly available data set using different classification algorithms like SVM, RBFN and Naive-Bayes. Besides this, we add a newly proposed algorithm that combines two optimization and classification (PSO-SVM) algorithms. The performance of all these four categories of classification is analyzed and compared with each other. Our proposed system of a hybrid PSO-SVM algorithm is implemented with the objective of improving the performance over other classification algorithms referred to as above and also analyzing how an evolutionary optimization algorithm like PSO will have an effect on the overall performance of the solution. The proposed framework is depicted in the flow diagram, Fig. 1 and 2 show the process of feature selection.

The objective of the proposed system is to develop an authentication framework that uses a combination of evolutionary algorithm and a conventional classification algorithm and performance evaluation of the system is done based on Equal Error Rate (EER) value. The implementation of algorithms is done on a publicly available data set which is first pre-processed. Further, an evolutionary Particle Swarm Optimization (PSO) algorithm for feature selection is run on the data and classification of users is performed using Support Vector Machine (SVM) classifier. The entire process is divided into two stages.

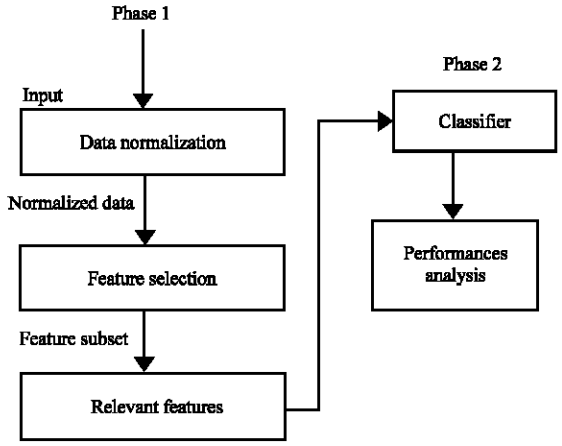

Fig. 1: Proposed framework

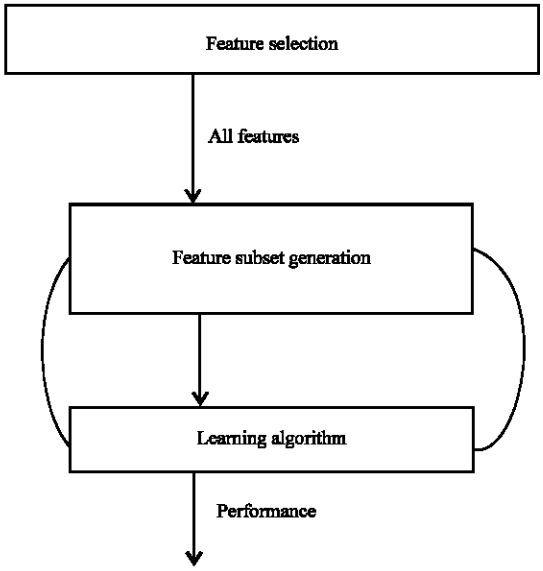

Fig. 2: Feature selection process

Stage 1; (Preprocessing) includes:

- Normalization of feature set
- Feature extraction/selection using Particle Swarm Optimization (PSO)

Stage 2; (Classification and evaluation) includes: Classification using Support Vector Machine (SVM).

RESULTS AND DISCUSSION

Feature selection or optimization technique is employed to avoid irrelevant features in the data set and produce a subset of the original data set on which further manipulation can be done conveniently and efficiently. This would also help in reducing the overall time taken for classification or training of data.

The inclusion of an optimization algorithm like PSO reduces the feature set making it relevant and efficient to research on. The PSO-SVM hybrid approach introduced here is expected to further improve the performance of the solution when compared with that of other classification algorithms. Performance analysis is done based on EER value and the lower this value, better the performance of the system. We hope the hybrid PSO-SVM classifier framework achieves an average EER value which is less than the other classifiers used in our study.

CONCLUSION

As mentioned earlier, mobile phones are becoming an integral part of our life and hence, secure ways to use it safely needs to be found out. Authentication can be done using both physical and even though physical trait authentication provides a higher level of accuracy it is often complicated and time consuming to validate using physical features. By reviewing a number of research, we learn that behavioral biometrics has an advantage over physical biometrics in that it provides uninterrupted authentication efficiently. But the demerit is that it cannot always guarantee a stable accuracy level. This leads us to choose a behavioral trait like swipe or touch pattern to be the basis of our study.

It is common to use evolutionary algorithms together with classification algorithms to form some kind of hybrid classification approach. Hence, in this research, an effort is made to develop an authentication framework that uses a combination of algorithms to differentiate among legitimate users and impostors with a higher level of accuracy and precision. The PSO-SVM hybrid algorithm used in this research is an effort made in regard to the problem of achieving a stable accuracy.

REFERENCES

Alariki, A.A. and A.A. Manaf, 2014. Touch gesture authentication framework for touch screen mobile devices. J. Theor. Appl. Inf. Technol., 62: 493-498.

Alariki, A.A., A.A. Manaf and S.M. Mousavi, 2016. Features extraction scheme for behavioural biometric authentication in touchscreen mobile devices. Intl. J Appl. Eng. Res., 11: 9331-9344.

Antal, M. and L.Z. Szabo, 2016. Biometric authentication based on touchscreen swipe patterns. Procedia Technol., 22: 862-869.

Ganesh, S. M., P. Vijayakumar and L.D. Jagatha, 2016. A comprehensive survey on gesture based authentication schemes in smartphones. Master Thesis, Department of Computer Science and Engineering, Ayyanthoppu, India.

Meng, W., D.S. Wong, S. Furnell and J. Zhou, 2015. Surveying the development of biometric user authentication on mobile phones. IEEE. Commun. Surv. Tutorials, 17: 1268-1293.

Meng, Y., D.S. Wong and R. Schlegel, 2012. Touch gestures based biometric authentication scheme for touchscreen mobile phones. Proceedings of the 8th International Conference on Information Security and Cryptology, November 28-30, 2012, Springer, Beijing, China, pp: 331-350.

Mondal, S. and P. Bours, 2015. A computational approach to the continuous authentication biometric system. Inf. Sci., 304: 28-53.