Fig. 1: Personal Data.

Biometric Authentication Based on Touchscreen Swipe Patterns

Andrey Preobrazhenskiy[1] and Dmitriy Chekulaev[2]

[1] Samara University, Samara 443086, Russia; *ssau@ssau.ru*

[2] Samara University, Samara 443086, Russia; *ssau@ssau.ru*

Abstract In this work we investigated user authentication on mobile devices using touch behavior and micro movements of the device. The novelty of our work lies in the collection of user behavior data during the filling in of a psychological questionnaire (implemented

Fig. 2: Questionnaire.

as an Android application). In order to answer the questions, users were required to use a slider. Therefore users were constrained to using only straight horizontal swipes. Extensive evaluations were conducted on the resulting dataset using one- and two-class classification algorithms. Although authentication EER based on single swipe is around 4%, this was improved by using sequences of 5 swipes (0.2% EER). Features related to micro movement of the device proved to be the most discriminating ones.

Key words: Security; Authentication; Behavioral Biometrics; Touchscreen; Swipe; Eysenck Personality Questionnaire

1 INTRODUCTION

There has been a huge increase in the use of touchscreen based mobile devices in people's everyday life. People store personal information on these devices, hence protecting user data is of paramount importance. Mobile devices simplify the collection of some behavioral biometric data due to their powerful sensors. The way people use their touchscreen devices yields new types of behavioral biometrics. Several research

Information	Description
Number of subjects	40
Number of samples	2729 (at least 58 samples/subject)
Device	Nexus 7 tablet
Number of swipes/question	unlimited
Controlled acquisition	Yes
Age range	20-49 (average: 25.92)
Gender	22 male, 18 female
Touchscreen experience	5 (level 1), 3 (level 2), 9 (level 3), 11 (level 4), 12 (level 5)

Table 1: Details of data acquisition.

studies have analyzed touch behavior in terms of new biometrics and the applicability of touch behavior for continuous authentication, indicating the huge potential in this type of biometric. None of the previous studies have exploited the full range of user specific features extractable from data collected from mobile devices.

2 RELATED WORK

In the last few years there has been an explosion in the study of touch biometrics. The importance of these type of behavioral biometrics is supported by the rapidly growing spread of touchscreen based smartphones. Some early research constructed user profiles based on general touch dynamics/behavior. Other researchers constructed phone unlock systems based on touch sequences, but none of these two studies used touchscreen specific features, such as pressure or finger area. Some researchers have investigated characteristics of scrolling interactions. These studies reported the necessity of using more than one swipe for accurate user authentication. Authentication using multitouch gestures were also studied. Phone usage behavior was captured by motion sensors and added to touchscreen usage behavior by Bo et al. in their SilentSense system. They analyzed authentication in both static and dynamic scenarios (the user using mobile phone while in motion). Thenovelty of our data collection method consists of collecting spatial, touch and motion data through a personality questionnaire in which users are required to use a slider in order to answer the questions (Table (1)). Utilizing a slider, all of the swipes are constrained straight swipes, in contrast to those used in previous studies.

3 MATERIAL AND METHODS

3.1 Data collection and feature extraction

In this paper we analyzed touch data collected through the Hungarian 58-question Eysenck Personality Questionnaire. An Android application was implemented in order to collect users' behavioral data during the filling in of the questionnaire (Figure (1)). In order to answer the questions, users had to use a slider. Therefore for each answer they made a horizontal swipe on the touchscreen. Besides touchscreen data - such as touch position, pressure and finger area - accelerometer data were also collected and user specific features were extracted from the raw data.

3.2 Authentication methods

An authentication system always has to decide whether a sample or a sequence of samples belongs to the genuine useror not. In pattern recognition or machine learning context, a two-class classifier can be employed for authentication systems if negative samples (impostor data) are available, otherwise one-class classifiers should be used. Several one-class classifiers were evaluated, such as the Parzen density estimator, the nearest-neighbor, Gaussian mixtures method and Support Vector Data Description method. Regarding two-class classification we chose to evaluate Random Forests, Bayes Net and k-nearest neighbors (k-NN) (Figure (3)). In order to construct two-class classifiers for authentication purposes, one has to select both positive and negative samples. Positive samples are always from the legitimate user and negative ones from impostors(Formula (2)). In order to use a class-balanced set for evaluation purpose, we selected all samples from a given user as positive samples, and two random samples from every other user as negative samples. For evaluation we used 10-fold cross-validation, namely 90% of the data was used for training and 10% of the data for testing and this was repeated for each fold combination. To get a more accurate result, we repeated the above evaluations 10 times (10 runs), using a different seed for randomization in each run.

Evaluation was performed several times, each time using a different length swipe sequence. Let us denote by C a trained classifier and he testing set containing N1 positive and N2 negative samples, where D is the number of features. We compute the scores for each swipe using the trained classifier C. Let us denote the set of scores obtained. The prediction score for a swipe sequence was computed by averaging the scores for each swipe. Let us denote by k the length of the swipe sequence used for authentication purpose.

Then, we can form (N1-k+1) sequences containing positive samples (Formula (3)):

The scores for these sequences were computed using formula (4):

$$X = \{x_1^+, x_2^+, ..., x_{N_1}^+, x_1^-, x_2^-, ..., x_{N_2}^-\}, x_i \in R^D \tag{1}$$

$$P = \{p_1^+, p_2^+, \ldots, p_{N_1}^+, p_1^-, p_2^-, \ldots, p_{N_2}^-\}, \tag{2}$$

$$S_i = \{x_i^+, x_{i+1}^+, ..., x_{i+k-1}^+\}, i = 1, N_1 - k + 1. \tag{3}$$

$$p(S_i) = \frac{\sum_{j=i}^{i+k-1} p_j^+}{k} \tag{4}$$

Sequences of negative samples were treated similarly. (Equation (1))

4 RESULTS

4.1 Datasets

All the measurements were performed on the following three datasets:

- dataset_11f, all 11 features

- dataset_8f, 8 touch features: duration, length of trajectory, average velocity, accelerationatstart, midstrokpressure, midstrokefingerarea, meanpressure, meanfingerarea

- dataset_3f, 3 gravity features: meangx, meangy, meangz

Classifier	Num Features	NumSwipes				
-	-	1	2	3	4	5
Bayes Net	3	0.067 ± 0.062	0.032 ± 0.026	0.015 ± 0.019	0.0009 ± 0.011	0.005 ± 0.017
k-NN	3	0.054 ± 0.144	0.071 ± 0.029	0.015 ± 0.038	0.0170 ± 0.012	0.005 ± 0.054
Random forests	3	0.057 ± 0.041	0.025 ± 0.014	0.010 ± 0.016	0.0050 ± 0.019	0.004 ± 0.001
Bayes Net	8	0.185 ± 0.048	0.113 ± 0.044	0.087 ± 0.038	0.0059 ± 0.036	0.046 ± 0.033
k-NN	8	0.193 ± 0.143	0.281 ± 0.058	0.102 ± 0.077	0.1450 ± 0.041	0.060 ± 0.054
Random forests	8	0.138 ± 0.035	0.076 ± 0.027	0.043 ± 0.025	0.0250 ± 0.022	0.016 ± 0.016
Bayes Net	11	0.056 ± 0.029	0.018 ± 0.024	0.011 ± 0.019	0.0005 ± 0.068	0.005 ± 0.013
k-NN	11	0.107 ± 0.143	0.115 ± 0.028	0.037 ± 0.038	0.0280 ± 0.037	0.016 ± 0.035
Random forests	11	0.044 ± 0.020	0.010 ± 0.011	0.005 ± 0.084	0.0050 ± 0.021	0.002 ± 0.000

Table 2: Two-class classification. Authentication EERs with confidence bounds. Swipe sequences of length: 1, 2, 3, 4, 5.

Datasets were normalized (range 0-1) in order to be used with classifiers sensible to features belonging to various numerical ranges.

4.2 Authentication

Two-class classifiers were used from the WEKA Data Mining Software package with the following parameters: 100 trees in the case of Random Forests, default settings for the Bayes Net classifier and k = 1 for the nearest neighbors (k-NN) classifier. With respect to one-class classifiers, the Dd tools Toolbox was used, with classifier parameters selected after some preliminary optimization tests as follows: default values for the Parzen density estimator (parzendd), two mixtures for the positive class in case of Mixture of Gaussians (mogdd), k = 3 for the nearest-neighbor data description and exponential kernel (P = 0.1) for the incremental Support Vector Data Description (incsvdd). Allof the experiments were performed by cross-validation methods (described in section 3) for both two- and oneclass classifiers and as a main measure for these tasks we used the mean EER value with confidence bounds across users and test folds. Results of authentication experiments formulated for two-class classifiers are presented in Table (2) and Figure (2). For two-class classifiers the average of gravity features (dataset 3f) already assured a reasonable authentication for one swipe (0.054±0.144 for the k-NN classifier), which was gradually improved for all classifiers by using the average scores of 2-5 consecutive swipes, achieving finally 0.004±0.001 EER value for the Random Forests classifier. The 8 touch feature group (dataset 8f) produced the poorest performance among our feature sets, but even this one achieved an EER value of 0.016±0.016 for five swipes. The 11-feature set gave the best results for the two-class classifier group, achieving EER values between 0.044±0.020 – 0.107±0.143 for one swipe, and 0.002±0.000 – 0.016±0.035 for 5 swipes. The best results were obtained for the 3 gravity feature set for all four classifiers (considering a single swipe). The best EER value was obtained for the knndd classifier (0.065±0.054). Large confidence bounds show that variance among users is high for the one-class classifiers (Figure (4)). Consecutive swipes reduced the EER value only for the distance based classifiers (knndd and incsvdd), but yielded no better results for density methods (parzendd and mogdd), except a slight improvement in the case of 11 features (parzendd). The best EER value for 5 swipes (0.023±0.019) was achieved for the incsvdd classifier.
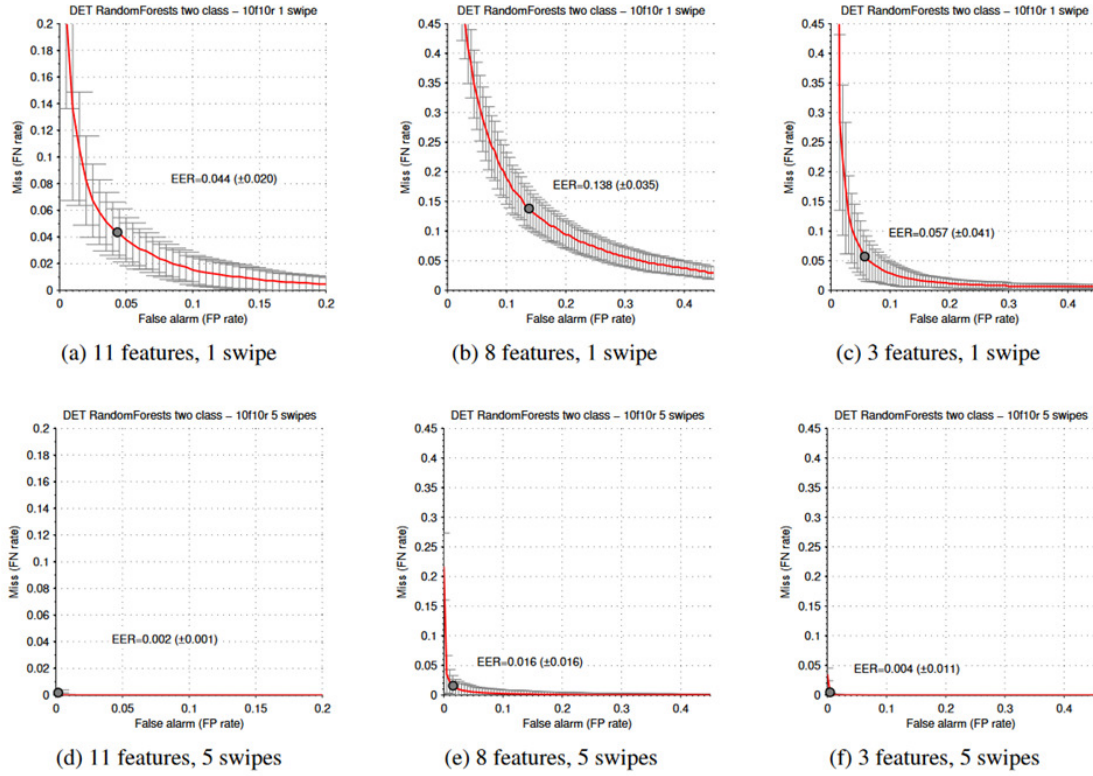
(a) 11 features, 1 swipe   (b) 8 features, 1 swipe   (c) 3 features, 1 swipe

(d) 11 features, 5 swipes   (e) 8 features, 5 swipes   (f) 3 features, 5 swipes

Fig. 3: DET curves, two-class authentication, Random Forests classifier.

5 CONCLUSION

Though it may be inadvisable to implement an authentication procedure using only a slider and some control questions, these results suggest, that constrained horizontal swipe movements have the potential to yield good authentication results for both one- and two-class classifiers. On the one hand, the best EER values achieved for single swipes (around 0.05) will not permit implementation of an authentication procedure. On the other hand, classifier performance was improved (except for one-class density based classifiers) by using the average scores of consecutive swipes.

Consequently, using 5 consecutive swipes in the case of our datasets produced a good EER value (0.002±0.000 for the 11-feature set and 0.004±0.001 for the 3-feature set) by using two class-classification (Random Forests classifier), so authentication is possible if impostor samples are available. Secondly, in case of missing impostor samples, the distance based one-class classifiers achieved good EER values (knndd: 0.024±0.020, parzendd: 0.023±0.019). These results could be further improved, if users with bad authentication results are excluded during the enrollment phase of an authentication system (with the proposal of using other biometric methods). Surprisingly, the most discriminative features were the features calculated as the average of x, y and z gravities. Though two-class classifiers generally perform poorly on smaller feature sets, they achieved good results even for a single swipe on these 3 features, close to the 11-feature set results (the k-NN classifier performed even better than for the 11-feature set). We conclude that device movement and holding position are the most user specific information (all feature selection algorithms confirmed this, the best touch features, like accelerationatstart and duration, generally appeared behind the gravity features - data not shown in this study). Feature sets which contained no information
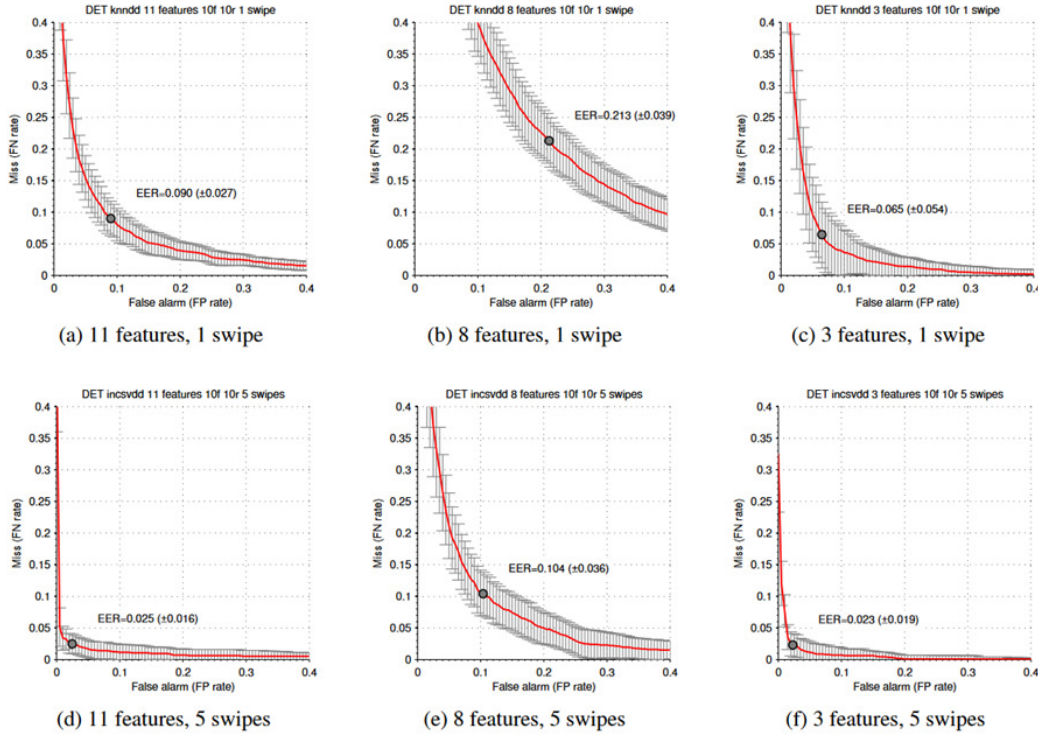
(a) 11 features, 1 swipe   (b) 8 features, 1 swipe   (c) 3 features, 1 swipe

(d) 11 features, 5 swipes   (e) 8 features, 5 swipes   (f) 3 features, 5 swipes

Fig. 4: DET curves, one-class authentication, knndd (1 swipe) and incsvdd (5 swipes) classifiers.

recorded from the accelerometer performed generally poorer, this is visible also from the results of the 8-feature set. In the case of distance based one-class algorithms, the best results were obtained by using the 3-feature set, in addition, the mean EER could not be improved by adding touch features (except for improvement in the confidence interval, which became narrower). The worst results were obtained by using only touch features in the case of one-class classifiers. Two-class classification methods also yielded a high accuracy concerning error rates both on the positive and negative class, meanwhile all one-class methods showed better performance on classifying the negative class than the positive class, however the scores were usable for expressing good EER results. In the near future we plan to continue data collection and correlate user specific EER with user's personality type.