Available online at www.sciencedirect.com
ScienceDirect
Procedia Technology 22 (2016) 862 – 869
Procedia Technology
ELSEVIER

9th International Conference Interdisciplinarity in Engineering, INTER-ENG 2015, 8-9 October 2015, Tirgu-Mures, Romania

Biometric authentication based on touchscreen swipe patterns

Margit Antal[a], László Zsolt Szabó[a]

[a]*Sapientia University, Sighisoarei 1C, Tirgu Mures 540485, Romania*

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 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.

Keywords: 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.

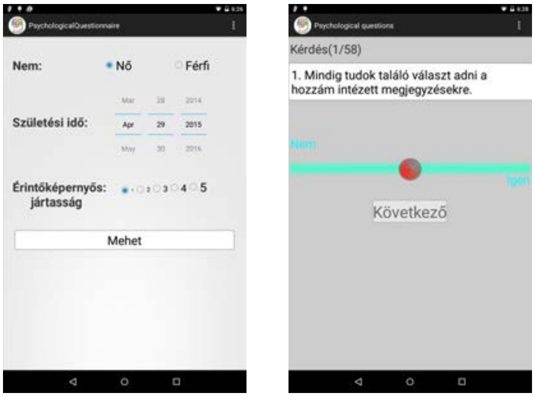
Mobile devices simplify the collection of some behavioral biometric data due to their powerful sensors. The way people use their touchscreen devices yields new types of behavioral biometrics. Several research studies have analyzed touch behavior in terms of new biometrics and the applicability of touch 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.

In this paper we analyze touchscreen and motion data collected through a psychological questionnaire in order to determine the user identity. The main research questions of this study are as follows:

- Is it possible to authenticate the user through constrained swipes on touchscreen-based mobile devices?

[*] Corresponding author. Tel.: +40-265-250-620 ; fax: +40-265-206-211.
E-mail address: manyi@ms.sapientia.ro

doi:10.1016/j.protcy.2016.01.061



Nem: ● Nő ○ Férfi

Születési idő: Apr 29 2015

Érintőképernyős: jártasság ● 1 ○ 2 ○ 3 ○ 4 ○ 5

Kérdés(1/58)

1. Mindig tudok találó választ adni a hozzám intézett megjegyzésekre.

(a) Personal data (b) Questionnaire

Fig. 1: Android application: Psychological questionnaire.

- How many swipes are necessary to accurately authenticate a user?
- Which set of features is the most discriminative?

The rest of this paper is organized as follows. Section 2 discusses related work. Section 3 presents data collection, feature extraction and the authentication framework used in this study. Section 4 describes the datasets used in experiments and the obtained results. Section 5 concludes the study and presents future research directions.

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 [1], [2], [3]. Other researchers constructed phone unlock systems based on touch sequences [4], [5], but none of these two studies used touchscreen specific features, such as pressure or finger area.

Some researchers have investigated characteristics of scrolling interactions [6], [7], [8], [9], [10], [11]. These studies reported the necessity of using more than one swipe for accurate user authentication. Authentication using multitouch gestures were also studied [12], [13]. Phone usage behavior was captured by motion sensors and added to touchscreen usage behavior by Bo et al. in their SilentSense system [14]. They analyzed authentication in both static and dynamic scenarios (the user using mobile phone while in motion).

The novelty 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. Utilizing a slider, all of the swipes are constrained straight swipes, in contrast to those used in previous studies.

3. Materials 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. In order to answer the questions, users had to use a slider. Therefore for each answer they made a horizontal swipe on the touchscreen (see Fig. 1 (b)). 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.

Table 1: Details of data acquisition.

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)

We define a swipe as a sequence of touch points from a touch down until a touch release event. In each touch point the following raw data were captured: touch action, x coordinate, y coordinate, x gravity, y gravity, z gravity, the pressure exerted, and the area occluded between finger and screen (finger area). All this information was obtained from the standard Android API, where touch action has three distinct values: ACTION_UP, ACTION_MOVE and ACTION_DOWN. Based on these action values the raw data were divided into swipes, and then features were extracted. Sometimes more than one swipe resulted during an answer. The following 11 features were extracted from each swipe:

- duration: the time between touch down and touch release;
- length_of_trajectory: the length of the segment defined by the two endpoints. It is computed as the sum of sub-segment length. A sub-segment is a segment between two consecutive touch points;
- average_velocity: it is computed as a fraction of the length of trajectory and duration;
- accelerationatstart: it is computed as the average acceleration at the first 4 touch points;
- midstrokpressure: the pressure at the middle point of the swipe;
- midstrokefingerarea: the finger area at the middle point of the swipe;
- meanpressure: the average of pressures in touch points;
- meanfingerarea: the average of finger areas in touch points;
- meangx: the average of x gravities in touch points;
- meangy: the average of y gravities in touch points;
- meangz: the average of z gravities in touch points;

Details regarding data acquisition are presented in Table 1. Data was collected from 40 subjects, at least 58 samples/subject (the questionnaire contains 58 questions).

3.2. Authentication methods

An authentication system always has to decide whether a sample or a sequence of samples belongs to the genuine user or 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).

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. 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.

The above procedure was applied also in the evaluation of one-class classifiers, except that only positive samples were used in the training phase.

Authentication results are presented using Equal Error Rate (EER) with confidence bounds and Detection Error Trade-off (DET) curves [15]. For DET curves one has to compute scores for positive and for negative samples. We used classifiers which, in addition to the classification decision, yield a score (likelihood value), which indicates the measure of being part of the positive class. DET curves are plotted by varying the decision threshold of the classifier across the ordered classifier output scores. Each presented DET error curve is an averaged error curve, based on 4000 evaluation (10 times x 10 folds x 40 users), and was calculated by using the *perfcurve* function from MATLAB (Statistics Toolbox, MATLAB, The Mathworks, Inc., Natick, MA).

Evaluation was performed several times, each time using a different length swipe sequence. Let us denote by C a trained classifier and $X = \{x_1^+, x_2^+, \ldots, x_{N_1}^+, x_1^-, x_2^-, \ldots, x_{N_2}^-\}, x_i \in R^D$ the testing set containing N_1 positive and N_2 negative samples, where D is the number of features. We compute the scores for each swipe using the trained classifier C. Let us denote $P = \{p_1^+, p_2^+, \ldots p_{N_1}^+, p_1^-, p_2^-, \ldots p_{N_2}^-\}$ 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 $N_1 - k + 1$ sequences containing positive samples: $S_i = \{x_i^+, x_{i+1}^+, \ldots x_{i+k-1}^+\}$, $i = \overline{1, N_1 - k + 1}$. The scores for these sequences were computed using formula $p(S_i) = \frac{\sum_{j=i}^{i+k-1} p_j^+}{k}$. Sequences of negative samples were treated similarly.

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

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 [16] 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 [17] 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).

All of the experiments were performed by cross-validation methods (described in section 3) for both two- and one-class 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 Fig. 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.

As for the one-class classification algorithms (see Table 3 and Fig. 3), they usually performed better for small feature sets if these were discriminative. 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. 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.

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

	Num			NumSwipes		
Classifier	Features	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**

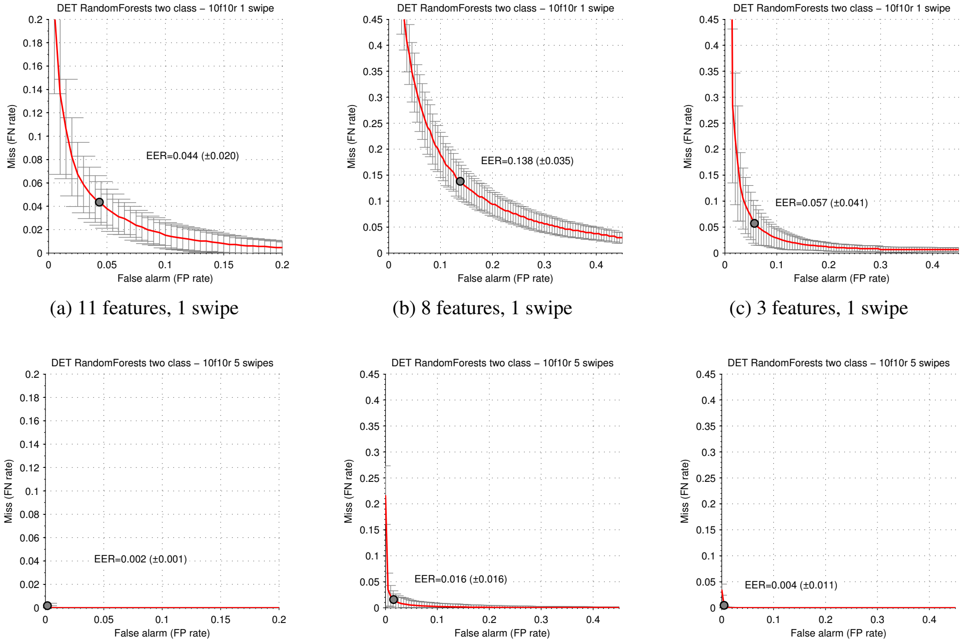
Table 3: One-class classification. Authentication EERs with confidence bounds. Swipe sequences of length: 1, 2, 3, 4, 5.

	Num			NumSwipes		
Classifier	Features	1	2	3	4	5
knndd	3	0.065 ± 0.054	0.045 ± 0.045	0.030 ± 0.023	0.025 ± 0.018	0.024 ± 0.020
mogdd	3	0.106 ± 0.037	0.120 ± 0.039	0.133 ± 0.039	0.137 ± 0.039	0.136 ± 0.040
parzendd	3	0.072 ± 0.047	0.070 ± 0.044	0.072 ± 0.045	0.071 ± 0.041	0.073 ± 0.037
incsvdd	3	0.084 ± 0.049	0.056 ± 0.037	0.038 ± 0.021	0.027 ± 0.019	**0.023 ± 0.019**
knndd	8	0.213 ± 0.039	0.180 ± 0.037	0.158 ± 0.041	0.149 ± 0.039	0.139 ± 0.039
mogdd	8	0.234 ± 0.054	0.216 ± 0.065	0.225 ± 0.076	0.234 ± 0.083	0.239 ± 0.096
parzendd	8	0.215 ± 0.039	0.208 ± 0.041	0.211 ± 0.049	0.214 ± 0.053	0.213 ± 0.055
incsvdd	8	0.217 ± 0.041	0.175 ± 0.041	0.142 ± 0.041	0.119 ± 0.038	**0.104 ± 0.036**
knndd	11	0.090 ± 0.027	0.069 ± 0.021	0.049 ± 0.018	0.046 ± 0.018	0.045 ± 0.020
mogdd	11	0.114 ± 0.031	0.107 ± 0.092	0.118 ± 0.113	0.130 ± 0.113	0.143 ± 0.117
parzendd	11	0.092 ± 0.030	0.073 ± 0.035	0.074 ± 0.038	0.077 ± 0.034	0.078 ± 0.033
incsvdd	11	0.093 ± 0.027	0.059 ± 0.023	0.041 ± 0.019	0.029 ± 0.018	**0.025 ± 0.016**

5. Conclusions

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



(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

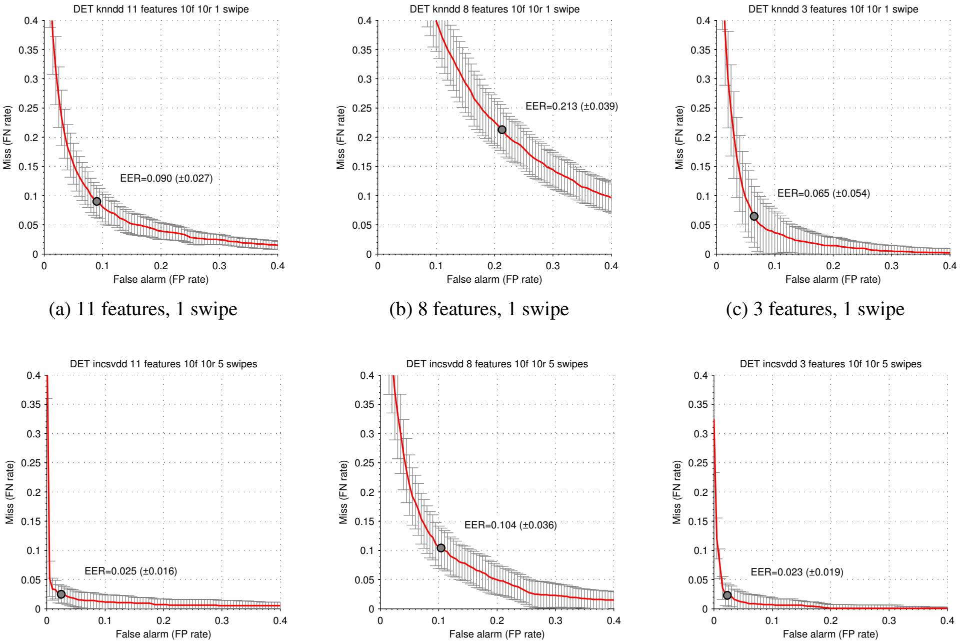
Fig. 2: DET curves, two-class authentication, Random Forests classifier.

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 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.



(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, one-class authentication, knndd (1 swipe) and incsvdd (5 swipes) classifiers.

Acknowledgements

The research has been supported by the European Union and Hungary and co-financed by the European Social Fund through the project TAMOP–4.2.2.C–11/1/KONV–2012–0004–National Research Center for Development and Market Introduction of Advanced Information and Communication Technologies.

The publication has been supported by Sapientia Foundation - Institute for Scientific Research.

We would also like to thank Lóránt Krausz who developed the Android application and all the people who have filled in the questionnaire.

References

[1] Seo, H., Kim, E., Kim, H.K.. A novel biometric identification based on a users input pattern analysis for intelligent mobile devices. Int J Adv Robot Syst 2012;1(1).

[2] Kolly, S.M., Wattenhofer, R., Welten, S.. A personal touch: Recognizing users based on touch screen behavior. In: Proceedings of the Third International Workshop on Sensing Applications on Mobile Phones. PhoneSense '12; New York, NY, USA: ACM. ISBN 978-1-4503-1778-8; 2012, p. 1–5.

[3] Meng, Y., Wong, D., Schlegel, R., Kwok, L.f.. Touch gestures based biometric authentication scheme for touchscreen mobile phones. In: Information Security and Cryptology; vol. 7763. 2013, p. 331–350.

[4] Angulo, J., Wastlund, E.. Exploring touch-screen biometrics for user identification on smart phones. In: Privacy and Identity Management for Life; vol. 375 of *IFIP Advances in Information and Communication Technology*. ISBN 978-3-642-31667-8; 2012, p. 130–143.

[5] De Luca, A., Hang, A., Brudy, F., Lindner, C., Hussmann, H.. Touch me once and i know it's you!: Implicit authentication based on touch screen patterns. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012, p. 987–996.

[6] Frank, M., Biedert, R., Ma, E., Martinovic, I., Song, D.. Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication. Information Forensics and Security, IEEE Transactions on 2013;8(1):136–148.

[7] Zhao, X., Feng, T., Shi, W.. Continuous mobile authentication using a novel graphic touch gesture feature. In: Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on. 2013, p. 1–6.

[8] Li, L., Zhao, X., Xue, G.. Unobservable re-authentication for smartphones. In: NDSS. The Internet Society; 2013,.

[9] Serwadda, A., Phoha, V., Wang, Z.. Which verifiers work?: A benchmark evaluation of touch-based authentication algorithms. In: Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on. 2013, p. 1–8.

[10] Antal, M., Bokor, Z., Szabó, L.Z.. Information revealed from scrolling interactions on mobile devices. Pattern Recognition Letters 2015;56:7–13.

[11] Roy, A., Halevi, T., Memon, N.. An hmm-based behavior modeling approach for continuous mobile authentication. In: Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on. 2014, p. 3789–3793.

[12] Sae-Bae, N., Ahmed, K., Isbister, K., Memon, N.. Biometric-rich gestures: A novel approach to authentication on multi-touch devices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '12; New York, NY, USA: ACM. ISBN 978-1-4503-1015-4; 2012, p. 977–986.

[13] Shahzad, M., Liu, A.X., Samuel, A.. Secure unlocking of mobile touch screen devices by simple gestures: You can see it but you can not do it. In: Proceedings of the 19th Annual International Conference on Mobile Computing; Networking. MobiCom '13; New York, NY, USA: ACM. ISBN 978-1-4503-1999-7; 2013, p. 39–50.

[14] Bo, C., Zhang, L., Li, X.Y., Huang, Q., Wang, Y.. Silentsense: Silent user identification via touch and movement behavioral biometrics. In: Proceedings of the 19th Annual International Conference on Mobile Computing; Networking. MobiCom '13; ACM. ISBN 978-1-4503-1999-7; 2013, p. 187–190.

[15] Martin, A., Doddington, G., Kamm, T., Ordowski, M., Przybocki, M.. The det curve in assessment of detection task performance. Tech. Rep.; DTIC Document; 1997.

[16] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.. The weka data mining software: An update. SIGKDD Explor Newsl 2009;11(1):10–18.

[17] Tax, D.. Ddtools, the data description toolbox for matlab. 2014. Version 2.1.1.