Pattern Recognition Letters 56 (2015) 7–13

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Information revealed from scrolling interactions on mobile devices ☆

Margit Antal*, Zsolt Bokor, László Zsolt Szabó

Sapientia Hungarian University of Transylvania, Tirgu Mures, Soseaua Sighisoarei 1C, Tirgu Mures and 540485, Romania

ARTICLE INFO

Article history:
Received 5 April 2014
Available online 2 February 2015

Keywords:
Biometrics
User classification
Gestures
Touchscreen

ABSTRACT

The aim of this study is to analyze information that can be revealed from simple touch gestures such as horizontal and vertical scrolling. Touch gestures contain identity information, they can reflect the user's experience using touchscreen and they can infer the gender of the user. The statements are based on measurements on a large touch dataset collected from 71 users using 8 different mobile devices, both tablets and phones. Touch data were divided in strokes and classification measurements were investigated based on single and multiple strokes. Classification results based on single stroke are inaccurate, which can be improved by using multiple strokes. Measurements prove that identity, gender and user's touchscreen experience level can be accurately predicted from a sequence of 10 strokes. In addition to the different classification results we present statistical analysis of the collected data in order to reveal basic differences between male and female users as well as for less and more experienced touchscreen users.

1. Introduction

Each user's interaction with a touchscreen is unique. This unique behavior has been investigated by several researchers proposing authentication schemes based on user's touchscreen dynamics. However, further information may be obtained from touch dynamics. Thus in this paper we propose to study gender and touchscreen usage experience information contained in touchscreen gestures besides identity information. Advertising applications could use gender information in presenting personalized information and games may use touchscreen usage experience to adapt game difficulty for the user.

In order to study the information contained in touchscreen gestures, we designed and developed a client-server application for collecting touchscreen usage patterns from mobile devices. We were interested in the information content of horizontal and vertical strokes, therefore we designed the client application accordingly. In order to collect vertical strokes, users had to read a long text and answer some questions related to it. When collecting horizontal strokes an image gallery was used from which the users had to select their favorite one. The client application was implemented as an Android application, which always starts synchronizing with the server, downloading new texts and images. After the user solves the problems, the generated touch data are sent to the server. The server part of the application permits addition of new texts, images and provides statistics of the collected data.

Although our study is not the only one in this field, its novelty consists in collecting data regarding personal information such as gender, age and touch experience level besides touch information. Consequently, we could perform gender and touchscreen experience level related classifications. Our dataset is also unique in containing data collected from devices with different screen sizes.

The contribution of this work is twofold. We present a dataset containing data from 71 users and 8 mobile devices. Second, we show that beside user identity further pieces of information are contained in touch data, which might be used in applications by adapting the contents to the user. As for the user identity classification results, we can state that using sequences of strokes raises the accuracy of classification. Gender information classification accuracy using only one stroke is quite high, around 88%, which slowly increases using more than one stroke. We noticed that men tend to make shorter and less straight strokes than women. Regarding the users touchscreen experience level we observed that less experienced users tend to make longer strokes with higher velocity.

In Section 2 we present a comprehensive review of the state-of-the-art in touch-based biometrics. Section 3 presents the research aims of our study. In Section 4 our dataset is presented beginning from data acquisition to feature extraction. Section 5 is dedicated to experimental evaluation. User identity, gender and touch experience level classifications are presented using multiple strokes. Beside classification results we present statistical analysis of the collected data in order to reveal main differences between male and female users. The two final sections contain discussions and conclusions.

☆ This paper has been recommended for acceptance by S. Sarkar.

* Corresponding author. Tel.: +40 265 250 620; fax: +40 265 206 211.
E-mail address: manyi@ms.sapientia.ro, margitantal68@gmail.com (M. Antal).

http://dx.doi.org/10.1016/j.patrec.2015.01.011
0167-8655/

2. Related work

Most desktop computers and mobile devices offer only entry-point based authentication schemes, which usually require a username/password combination. In many institutions people share their authentication data in order to transfer work to other employees. If software was able to continuously track the identity of the user, this abnormal usage pattern would be impossible. Keystroke dynamics [1] is a proposal to solve such usage anomalies. On the one hand it may be used in the username/password typing phase, allowing verification of the user's identity not only by a proper password, but also by typing rhythm, whereas on the other hand continuous authentication would be possible in applications requiring users' typing. Killourhy and Maxion [2] compared several anomaly detectors for keystroke dynamics and identified the best ones. Bours [3] went one step further and proposed continuous, dynamic authentication.

Due to mobile device usage patterns, typing rhythm based continuous authentication schemes are not suitable for smartphones. However, compared to desktop computers, mobile devices are more vulnerable concerning their security. Hence recent studies propose touch-based biometrics instead of keystroke dynamics.

2.1. Touch usage based user profiling

Seo et al. [4] conducted two studies for biometric identification based on touchscreen data. In their first experiment only simple touch data were used, such as touch duration, pressure level and the area covered by finger. In their second experiment scroll wheel touch data (strokes) were used in addition to basic touch data. From each stroke start and stop position, speed and stroke length were used. Neither input pattern classification features nor collecting data methods are very clear. However, the paper reports 100% identification accuracy using 12 features and a Back Propagation Neural network (BPN).

A large scale study to observe the users' touchscreen behavior on standard UI elements was conducted by [5]. Hold time characteristics and pressure dynamics of 14000 users through a dedicated game were studied.

Meng et al. [6] present another exploration of touch dynamics for user authentication. The implemented system generates authentication signatures for each session containing 21 features. Users had to complete 6 sessions within 3 days, each lasting for 10 min. The paper reports best error rates using particle swarm optimization (PSO) combined with radial basis function (RBF) network: 2.5% false acceptance rate (FAR), 3.34% false rejection rate (FRR). However, the 10 min session length is a serious limitation of this authentication scheme, as a phone should detect an intruder in a much shorter time.

Damopoulos et al. [7] implemented a touchlogger for iOS based smartphones to demonstrate the ability of touch data for user identification. This paper presents touch data collected through natural usage of smartphones. Participants used their own devices for 24 h, while the touchlogger application running in the background was constantly logging their touch data, which was retrieved from the devices at the end of the experiment. The feature vector contained the x and y coordinates and the timestamp corresponding to the touch event. After feature extraction different supervised machine learning algorithms were tested using Weka. The best error rate was obtained by Random forests: 0.205% equal error rate (EER).

2.2. Touch gestures for unlocking devices

De Luca et al. [8] present two experiments using touch strokes for unlocking mobile devices. The aim of the study was to investigate whether it was possible to recognize users based on the way they perform unlocks. The users had to unlock the device using 4 unlock types: horizontal, vertical, diagonal and two-finger vertical strokes. Both experiments were conducted using Android devices and touch data obtained through standard Android API were collected during two sessions. Strokes were compared using dynamic time warping (DTW) algorithm. The best reported accuracy was 57% for diagonal strokes. They concluded that longer time series lead to better results, therefore a second study was designed in order to allow the collection of longer time series. The participants had to unlock their phones once per day for 21 days using password patterns having different levels of difficulties. The best result was 77% accuracy using feature vectors containing pressure, finger area and speed.

Angulo and Wastlund [9] also investigated graphical lock patterns biometrics. Features similar to keystroke dynamics were employed, measuring drawing pattern speed. Neither pressure nor finger area size was employed. A number of 32 users participated in this experiment using an Android application for data collection. A total of 3 graphical patterns were used, each consisting of 6 dots. Fifty trials of each pattern were collected from each user, resulting in 150 trials per user. The paper reports 10.39% EER using Random forests [10] classifier.

Authentication using a set of gestures involving multiple fingers was studied by [11]. A 90% accuracy was reported on a dataset containing data from 34 users by using a single gesture, which can be improved by taking the authentication decision based on multiple gestures. These results are considered to be very promising however, the applicability is still reduced to tablets with large screens due to the special five-finger gestures.

2.3. Touch based continuous user authentication

Feng et al. [12] examined gesture based continuous user authentication for the first time. Horizontal, vertical and zooming gestures were involved, reporting classification accuracies separately for each type of gesture. Random forests, Bayes Net and C4.5 (Weka J48) were applied as classifiers. They used 7 consecutive strokes in the authentication measurements, accepting a sequence as valid only if 3 or more strokes are recognized as inputs from the authorized user. Using this setting FAR = 4.66% and FRR = 0.13% were achieved.

Li et al. [13] presented a study related to continuous authentication in which tap, sliding left, right, up and down gestures were collected from 75 users. Data collection was performed in an unsupervised manner, participants were allowed to use the phone without restrictions. The lowest FAR (≈3%) and FRR (≈3%) error rates were obtained when feature vector was computed from a sequence of 14 gestures.

One of the deepest touch data analysis for continuous user authentication is presented by [14]. This thorough paper presents an experiment in which touch data are collected during text comprehension and image comparison tasks. In order to complete these tasks, users had to make several vertical and horizontal strokes using the touchscreen. Forty one users participated in this study in which data were collected in multiple sessions. The authors propose 30 touch features which were evaluated using the k-NN and the SVM (with Gaussian RBF kernel) classifiers. EER less than 4% was achieved using 11 strokes in the long-term authentication scenario.

Micro-movement characteristics of the device were added to the users' touch behaviors by [15] in their SilentSense system. Touchscreen usage actions were classified as tap, scroll (vertical scroll) and fling (horizontal scroll). After 10 user actions, the system identified users with FAR and FRR below 1% using SVM classifier on a dataset containing data from 100 users.

A novel graphic touch gesture feature (GTGF) to extract the identity features from touch strokes was proposed by [16]. A 2.62% EER was achieved on a dataset from 30 subjects using 6 gestures.

Serwadda et al. [17] collected horizontal and vertical strokes from 190 users in 2 sessions. Testing was performed using a sliding window, in which 10 strokes were used to compute a single feature vector. Measurement results are reported based on eight machine learning algorithms and two others using distance metrics. They reported that

Table 1
Touch features based studies. ACC – accuracy, EER – equal error rate, FAR – false acceptance rate, FRR – false rejection rate, NA – not applicable.

Study	#Users	#Strokes	Algorithms	Best result(s) (%)
User profiling				
[4]	50	NA	BPN	ACC: 100
[5]	14000	NA	Naive Bayes	ACC: 80
[6]	20	NA	PSO-RBF	FAR: 2.5 FRR: 3.34
[7]	18	NA	Random forests	EER: 0.205
Unlocking devices				
[8]	48	NA	DTW	ACC: 77
[9]	32	NA	Random forests	EER: 10.39
[11]	34	NA	DTW	ACC: 90
Continuous user authentication				
[12]	40	7	Random forests, C4.5, Bayes net	FAR: 4.66 FRR: 0.13
[13]	75	20	SVM (LibSVM)	FAR: ≈3 FRR: ≈3
[14]	41	11	k-NN, SVM	EER: <4
[15]	100	10	SVM	FAR: ≈1 FRR: ≈1
[16]	30	6	Image processing: score between two images	EER: 2.62
[17]	190	10	Logistic regression, SVM, Random forests	EER: 10.5
[18]	50	3	SVDE	EER: 0.5
[19]	41	11	HMM	EER: <3

authentication based on a block of strokes gave better performance than authentication based on a single stroke.

GEAT, a gesture based user authentication scheme for touchscreen devices was proposed by [18]. A 0.5% EER using 3 gestures on a dataset containing gestures from 50 users was reported using support vector distribution estimation (SVDE) with RBF kernel.

Hidden Markov models (HMM) were tested by [19]. Separate models were created from users' horizontal and vertical scrolling data. Evaluation measurements were conducted on the Touchalytics dataset [14]. A 1.75% EER for 11 horizontal strokes and 2.80% EER for the same number of vertical strokes were reported, which can be considered an improvement of the performance reported by the dataset creators.

Implicit authentication (IA) schemes were presented by [20] and [21]. As both papers focus on the framework presentation, the accuracy of the framework for touch gestures are not evaluated at all. Based on the above literature review we can conclude that there have been two major data collection methods so far: during specific tasks in order to produce the desired touch data, and during usual everyday activities. Table 1 lists the details of all touch biometrics studies.

3. Research aims

Our primary aim is to extend the existing studies for user identification and authentication by using devices with different screen sizes. The existing studies are almost limited to mobile phones, and our study includes tablets as well. Furthermore, in our study the basic usage pattern is the stroke, since for data collection we designed tasks constraining the user to make horizontal and vertical strokes. The first research question related to this study is: how many strokes are necessary to accurately identify the user? The second group of questions is related to the gender of the user: is it possible to find out the gender of the user from a sequence of strokes? How many strokes are necessary to accurately identify the gender of the user? What are the main differences between male and female touchscreen users? The third group of questions is related to the touchscreen experience level of the user. Since we have collected data related to users' experience level, we tried to find answers for the following questions: Is it possible to detect the touchscreen experience level of the user? How many strokes are necessary for an accurate detection? What are the differences between beginner and advanced touchscreen users?

4. Dataset

4.1. Data acquisition

A client server application was developed for data acquisition. The Android client presents to users two types of tasks requiring scrolling interactions. One type of task is reading, whereas the other one is an image gallery. While the reading task requires vertical scrolling, the image gallery task requires horizontal scrolling. During the first task users have to read a text and answer questions regarding its comprehension. After user login Android clients always synchronize with the server and fetch new texts uploaded by an administrator user. For the second task the images were uploaded to the server application side and grouped into albums. Android clients synchronize with the server and download only the newest album. In this case the task was to select the favorite picture. Data were continuously sent to the server during the two tasks. For each touch point the data collector application recorded: the timestamp of the touch, the x and y coordinates, the pressure exerted on the screen, and the area occluded between finger and screen. Data collection was performed during 4 weeks; a new text comprehension task and a new album were added for each week. Seventy one users, 56 men and 15 women participated in this study. The average age of the participants was 29.8 with a standard deviation of 9.08. The youngest was 19 whereas the oldest one was 47 years old. 8 different Android devices were used with resolutions ranging from 320×480 to 1080×1205 pixels. Each participant provided the data in multiple sessions, sometimes from multiple devices. In the registration phase several extra pieces of information were required from the participants such as gender, birth date and touchscreen user experience level. Device information such as Android version and resolution were also stored. The user experience level was quantized into 4 classes: inexperienced (level 0), moderately experienced (level1), experienced (level 2), very experienced (level 3).

4.2. Data cleansing

Data acquisition was followed by data cleansing. In this process sequences of pixels were divided into strokes, a procedure, which provoked several problems to solve. First of all, several users did not rise their finger between consecutive strokes. This resulted in compound strokes that had to be divided into individual strokes using pressure and time information. Second, one device turned out to be inappropriate in data acquisition due to its incapacity to return the accurate pressure value. This device always returned the value 1, consequently all the strokes collected through this device were excluded from the dataset. Finally, short strokes were detected, containing no more than three pixels. These strokes were also excluded from the final dataset. After cleaning the data 14,316 strokes remained (11,584 horizontal and 2,732 vertical), which means that on average we obtained 200 strokes/user.

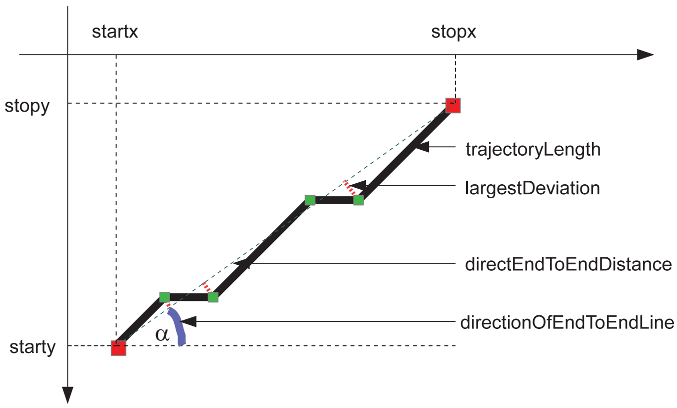

Fig. 1. Features extracted from touch stroke.

4.3. Feature extraction

The collected raw data were divided up into strokes. Each stroke consists of a sequence of touchscreen points between a starting and stopping point. The kth stroke is defined as $v^k = (x_i^k, y_i^k, t_i^k, p_i^k, A_i^k), i \in 1, 2, \ldots, N^k$, where x_i^k, y_i^k are the coordinates of the touching position, t_i^k is the timestamp, p_i^k is the pressure, A_i^k is the area covered by the finger and N^k denotes the number of points belonging to the stroke. One feature vector was extracted from each stroke. We used a similar terminology as in [14], but with fewer features. Our feature vector contains the following elements (several are shown in Fig. 1):

1. strokeDuration: the time needed for a stroke expressed in milliseconds;
2. startX: the x coordinate of the stroke starting point;
3. startY: the y coordinate of the stroke starting point;
4. stopX: the x coordinate of the stroke ending point;
5. stopY: the y coordinate of the stroke ending point;
6. directEndToEndDistance: the length of the segment defined by the two endpoints;
7. meanResultantLength: a feature characterizing the straightness of the stroke [14];
8. upDownLeftRightFlag: orientation of the stroke; a stroke is classified horizontal (left, right) if its horizontal displacement exceeds its vertical displacement;
9. directionOfEndToEndLine: the slope of the segment defined by the two endpoints;
10. largestDeviationFromEndToEndLine: the maximum of the distances between points belonging to the stroke and the segment defined by the two endpoints;
11. averageDirection: the average slope of the segments belonging to the stroke trajectory;
12. lengthOfTrajectory: the length of the stroke;
13. averageVelocity: the average velocity of the stroke;
14. midStrokePressure: the pressure at the midpoint of the stroke;
15. midStrokeArea: the area covered by finger at the midpoint of the stroke.

Additional information were available for each stroke such as information regarding the device, the task in which the stroke was produced and information regarding the user who produced the stroke. Some of these additional information were used only as attributes for labeling the class in different classifications.

5. Experimental evaluation

5.1. Classification algorithms

Evaluation was performed using a Java program with Weka library (version 3.6.11)[22] for classifier algorithms. This was necessary because Weka provides only instance classification through the interface, which can be used only for single stroke classification, while we performed stroke sequence classifications.

For stroke classification we used k-NN (IBk in Weka), Random forests [10] and Support Vector Machine (SVM) algorithms. k-NN is an instance based classification algorithm, where a new instance label is decided by the k closest neighbors. For each dataset we tested the algorithm for several odd values for parameter k and reported the best accuracy. We always mention the value of parameter k which produce the best accuracy.

Random forests classifier builds up a number of decision trees following specific rules for tree growing, tree combination, self-testing and post-processing. Among the benefits of Random forests classifier we mention the following: it can be used for both regression and classification, there is no need for prior feature selection and it trains rapidly.

Support vector machines build a linear discriminant function separating the instances of classes. If no linear separation is possible, a kernel maps the instances into a high-dimensional feature space. We used the LibSVM implementation through Weka with RBF kernel. The C and γ kernel parameters were optimized by a grid search algorithm distinctly for each dataset and all input features were normalized (0–1).

All measurements were carried out using 3-fold cross-validation, except classifier learning curve calculation, where we used 5-fold stratified cross-validation.

5.2. Stroke sequence classification

Let us denote N the number of classes and X the sequence of strokes to be classified:

$$X = \{x_1, x_2, \ldots, x_T\}, \qquad x_i \in R^D, \tag{1}$$

where T is the number of strokes and D is the dimension of the feature vector. We compute for each stroke the prediction distribution (Eq. (2)).

$$P_i = \{p_i^1, p_i^2, \ldots, p_i^N\}, p_i^k \in [0, 1], k = 1 \ldots N, i = 1 \ldots T \tag{2}$$

where p_i^k is the probability that x_i belongs to class k.

This is followed by computing the average probability for each class and choosing the maximum one (Eq. (3)).

$$\text{Class}(X) = \arg\max_{k=1}^{N} \left\{ \frac{\sum_{i=1}^{T} p_i^k}{T} \right\} \tag{3}$$

Consequently, a sequence of strokes is classified belonging to the kth class if the average probability for this class is the maximum one.

Stroke sequences considered for testing always contained the strokes from a user in chronological order.

5.3. Datasets

Table 2 shows 4 different datasets for user identity, stroke and gender classifications (The data used in this work can be accessed at http://www.ms.sapientia.ro/~manyi/bioident.html). For each type of classification we created a subset of the original dataset which contains the same number of samples from each user. Each dataset contains 15 attributes (class attribute is not counted).

All datasets were used without any transformation or feature selection (excepting normalization). No boosting or other tuning methods were used, except the case of random forests classifier, which is based on internal randomization.

5.4. User identity classification

User identity classification was performed both on dataset1 and dataset2. The second one contains fewer users and it is more homogeneous, containing only horizontal strokes. Table 3 shows

Table 2
Datasets and their characteristics used in measurements.

Name	#Users	#Strokes/user	Gender		Experience				Type
			#Male	#Female	#Level0	#Level1	#Level2	#Level3	
Dataset1	71	200 (On average)	56	15	15	9	31	16	Horizontal, vertical
Dataset2	51	100	42	9	12	6	21	12	Horizontal
Dataset3	18	100	9	9	4	2	8	4	Horizontal
Dataset4	24	100	20	4	6	6	6	6	Horizontal

the best parameters for the classification algorithms using the two datasets.

Fig. 2 shows the classification results obtained for the two datasets, using a varying number of strokes for classification. It can be seen that the classification rate is quite low for a single stroke, but it increases as the number of strokes grows. The Random forests classification performance is slightly better than the k-NN performance, especially for a number of reduced strokes. Thus we can state that it is possible to identify the user from strokes, although a single stroke does not contain enough user specific information. The results also show that an accurate user identification may be done using at least 10 strokes.

Learning curves (see Fig. 3) were calculated from datasets created with sampling without replacement of 5, 10 … 95% of dataset2 with Weka's Resample filter. The mean value of correctly classified instances is presented. The used parameters were $k = 1$ for k-NN, 100 trees for the Random forests algorithm and $C = 8, \gamma = 8$ for SVM with RBF kernel. Since the balanced dataset contains 100 strokes for each user, and stratified cross-validation was used, one could estimate the necessary number of strokes for a learning rate in percent, reported to our best result.

5.5. Gender classification

Gender classification is a two-class classification problem. In order to get a more accurate result, for this measurement we createda subset of users from the original dataset containing 9 male and 9 female users (dataset3). Fig. 4 shows that each classification algorithm predicts with 88% accuracy the gender of the user based on the stroke. Using a 10 stroke sequence the accuracy is over 99% for each classification algorithm. The best parameters for the classification algorithms were: $k = 3$ (k-NN), 100 trees (Random forests) and $C = 2, \quad \gamma = 8$ (SVM, RBF kernel).

In order to find the differences between male and female users Wilcoxon signed-rank test (0.05 significance level) was performed for all the 15 attributes involved in classification. No differences were found for the following attributes: startX, upDownLeftRight, largestDeviationFromEndToEndLine, averageVelocity and midStrokeArea. The most significant differences were found at the following attributes (in decreasing order): startY, stopY, directEndToEndDistance, lengthOfTrajectory, meanResultantLength, strokeDuration, directionOfEndToEndLine, stopX, midStrokePressure, averageDirection. As shown in Fig. 5 men tend to make shorter and less straight strokes than women.

5.6. Touch experience level classification

We have set up four classes of users {inexperienced (0), moderately experienced (1), experienced (2), very experienced (3)}. That being so, user experience level labels the classes. For this measurement we created a subset of users from the original dataset containing the same number of users and strokes from each class (dataset4). The best parameters for the classification algorithms were: $k = 1$ (k-NN), 100 trees (Random forests) and $C = 2, \quad \gamma = 8$ (SVM, RBF kernel).

As shown in Fig. 6 the classification accuracies for the best performer (Random forests) varies from 81% for single stroke to 100% for 20 strokes.

Table 3
User identity classification: algorithm parameters.

Dataset	k-NN (k)	Random forests (numTrees)	SVM, RBF kernel(C, γ)
Dataset1	3	10	2, 8
Dataset2	3	100	8, 8

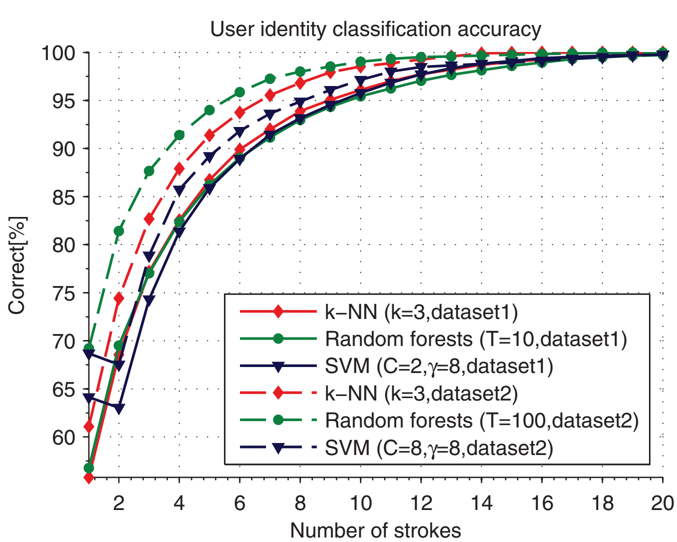

Fig. 2. User identity classification for dataset1 and dataset2. The x axis shows the length of stroke sequence used for classification, while the y axis represents the accuracy of this classification.

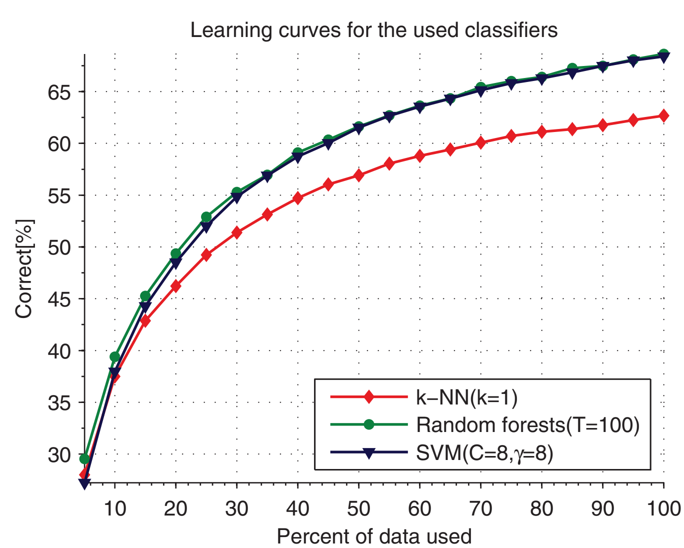

Fig. 3. User identity classification learning curves. The x axis represents percents of data obtained from dataset2 by sampling without replacement using the Resample filter from Weka.

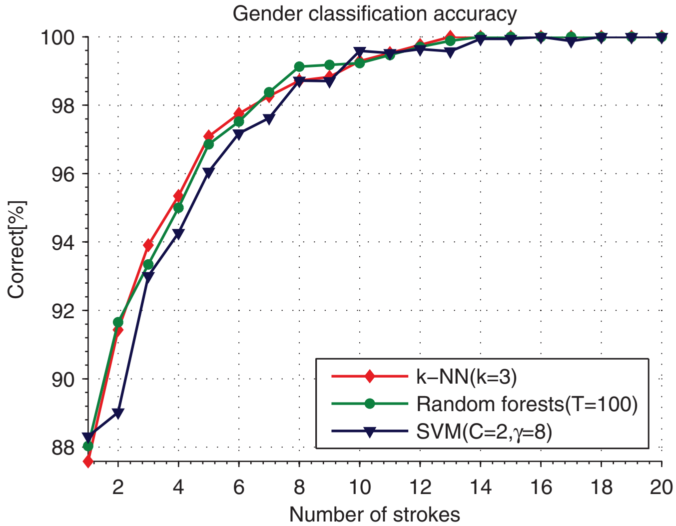

Fig. 4. Gender classification for dataset3. The *x* axis shows the length of stroke sequence used for classification, while the *y* axis represents the accuracy of this classification.

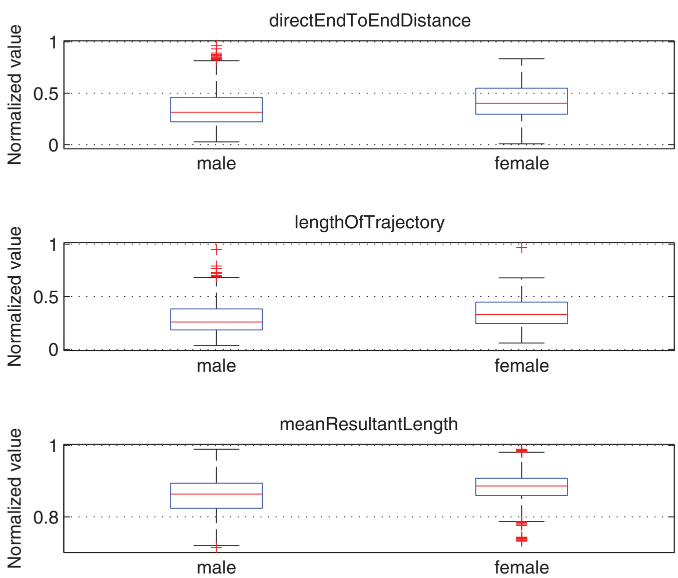

Fig. 5. Feature distributions for male and female users.

Fig. 7 exemplifies the basic differences between users belonging to inexperienced (level 0) and very experienced (level 3) categories (the same Wilcoxon signed-rank test was used). Less experienced users tend to make longer strokes with higher velocity.

6. Discussion

This study proves that – besides user identity information – there are other types of information in users' touch data, such as gender and touchscreen experience level. The more data we use for user identification, the better identification accuracy we obtain, regardless of the classification algorithm. This is true for gender and experience level classifications too. Using stroke sequences of length 10 resulted in classification accuracies over 98% for each type of classification. Developers of mobile applications may benefit from the results of our study by adapting applications according to the users' gender and touchscreen usage experience. Applications may use gender information in more adapted information/offers to users, whereas touchscreen usage experience may be used for better adapted game difficulty. Our findings correlate with other type of biometric identi-

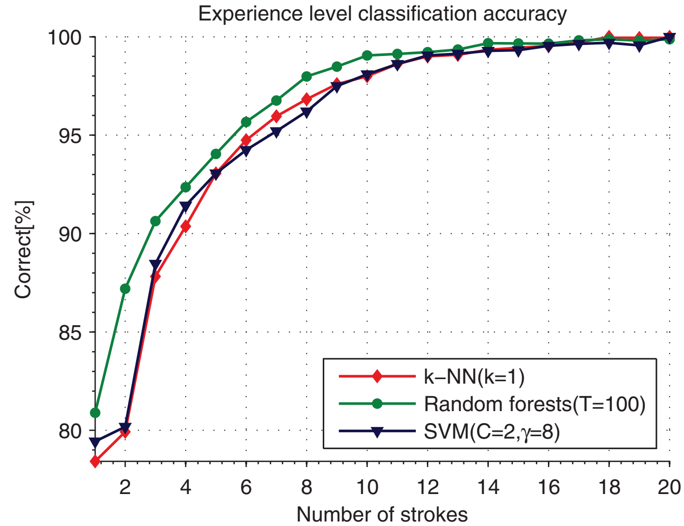

Fig. 6. Experience level classification for dataset4. The *x* axis shows the length of stroke sequence used for classification, while the *y* axis represents the accuracy of this classification.

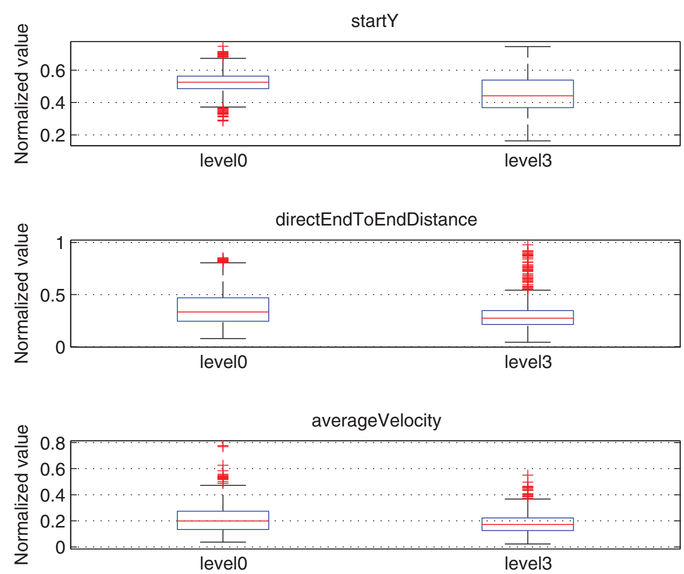

Fig. 7. Feature distributions for inexperienced (level0) and very experienced (level3) users.

fication results. Rising the amount of test data leads to better identification accuracy. The basic limitation of this study is the selection of the participants, who are predominantly students and professors from our university. Furthermore, the whole dataset is unbalanced in favor of male participants. However, we performed user identity, gender and touchscreen experience level classifications on class balanced subsets. Despite the fact that touchscreen related experience level is a subjective data, we tried to identify a few basic characteristics of less and more experienced users.

7. Conclusions

Revealing gender or touch experience level from touch data brings new perspective for the mobile application developer.

An experiment was designed in which touch data were collected from 71 users. Data were collected through 8 different devices, both tablets and smartphones. Three different type of classifications were performed: user identity, gender and touchscreen experience level of the users. For all type of classifications k-NN, Random forests and

SVM algorithms were used. All the measurements were performed by a Java program using the Weka library or the Weka built-in tools. Besides single stroke classification we performed multiple strokes classifications trying to determine the optimal number of strokes for each type of classification.

User identity classification starts at about 65% classification accuracy using a single stroke and reaches 100% accuracy for 20 strokes. Gender classification were performed on a subset containing equal number of male and female users. Classification accuracies in this case are between 88% and 100% for all classification algorithms used in this study. For experience level classification accuracies between 80% and 100% were obtained by varying the number of strokes from 1 to 20. All the classifications were performed using the same 15 attributes for a stroke.

Besides the classification results, we presented the basic differences between male and female users, as well as between less and more experienced touchscreen users.

In conclusion, user identity, gender and touchscreen experience recognition can be performed with high accuracy (over 98%) for blocks of 10 strokes.

Acknowledgments

The research has been supported by Sapientia Foundation – Institute for Scientific Research and partly 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.

References

[1] F. Monrose, A. Rubin, Authentication via keystroke dynamics, in: Proceedings of the 4th ACM Conference on Computer and Communications Security, 1997, pp. 48–56.
[2] K. Killourhy, R. Maxion, Comparing anomaly-detection algorithms for keystroke dynamics, in: IEEE/IFIP International Conference on Dependable Systems Networks, DSN '09. 2009, pp. 125–134.
[3] P. Bours, Continuous keystroke dynamics: a different perspective towards biometric evaluation, Inf. Secur. Tech. Rep. 17(1–2) (2012) 36–43.
[4] H. Seo, E. Kim, H.K. Kim, A novel biometric identification based on a user's input pattern analysis for intelligent mobile devices, Int. J. Adv. Robot Syst. 1 (1) (2012) 1–10.
[5] S.M. Kolly, R. Wattenhofer, S. Welten, A personal touch: recognizing users based on touch screen behavior, in: Proceedings of the Third International Workshop on Sensing Applications on Mobile Phones, ACM, New York, NY, USA, 2012, pp. 1–5.
[6] Y. Meng, D. Wong, R. Schlegel, L.f. Kwok, Touch gestures based biometric authentication scheme for touchscreen mobile phones, in: Information Security and Cryptology, vol. 7763, pp. 331–350, 2013.
[7] D. Damopoulos, G. Kambourakis, S. Gritzalis, From keyloggers to touchloggers: take the rough with the smooth, Comput. Secur. 32 (2013) 102–114.
[8] A. De Luca, A. Hang, F. Brudy, C. Lindner, H. Hussmann, 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, pp. 987–996.
[9] J. Angulo, E. Wastlund, Exploring touch-screen biometrics for user identification on smart phones, in: Privacy and Identity Management for Life, vol. 375 IFIP Advances in Information and Communication Technology, 2012, pp. 130–143.
[10] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.
[11] N. Sae-Bae, K. Ahmed, K. Isbister, N. Memon, Biometric-rich gestures: a novel approach to authentication on multi-touch devices, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, New York, NY, USA, 2012, pp. 977–986.
[12] T. Feng, Z. Liu, K.A. Kwon, W. Shi, Continuous mobile authentication using touchscreen gestures, in: IEEE Conference on Technologies for Homeland Secutity (HST), 2012, pp. 451–456.
[13] L. Li, X. Zhao, G. Xue, Unobservable re-authentication for smartphones., in: NDSS, The Internet Society, 2013.
[14] M. Frank, R. Biedert, E. Ma, I. Martinovic, D. Song, Touchalytics: on the applicability of touchscreen input as a behavioral biometric for continuous authentication, IEEE Trans. Inf. Forensics Secur. 8 (1) (2013) 136–148.
[15] C. Bo, L. Zhang, X.Y. Li, Q. Huang, Y. Wang, Silentsense: silent user identification via touch and movement behavioral biometrics, in: Proceedings of the 19th Annual International Conference on Mobile Computing; Networking, ACM, 2013, pp. 187–190.
[16] X. Zhao, T. Feng, W. Shi, Continuous mobile authentication using a novel graphic touch gesture feature, in: IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2013, pp. 1–6.
[17] A. Serwadda, V. Phoha, Z. Wang, Which verifiers work?: a benchmark evaluation of touch-based authentication algorithms, in: IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2013, pp. 1–8.
[18] M. Shahzad, A.X. Liu, A. Samuel, 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, ACM, New York, NY, USA, 2013, pp. 39–50.
[19] A. Roy, T. Halevi, N. Memon, An hmm-based behavior modeling approach for continuous mobile authentication, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014, pp. 3789–3793.
[20] H. Crawford, K. Renaud, T. Storer, A framework for continuous, transparent mobile device authentication, Comput. Secur. 39 (2013) 127–136.
[21] H. Khan, A. Atwater, U. Hengartner, Itus: an implicit authentication framework for android, in: Proceedings of the 20th Annual International Conference on Mobile Computing and Networking, ACM, New York, NY, USA, 2014, pp. 507–518.
[22] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I.H. Witten, The weka data mining software: an update, SIGKDD Explor. Newsl. 11 (1) (2009) 10–18.