IET Biometrics

Research Article

Active detection of age groups based on touch interaction

ISSN 2047-4938
Received on 14th January 2018
Revised 26th August 2018
Accepted on 17th September 2018
E-First on 5th November 2018
doi: 10.1049/iet-bmt.2018.5003
www.ietdl.org

Alejandro Acien[1] ✉, Aythami Morales[1], Julian Fierrez[1], Ruben Vera-Rodriguez[1], Javier Hernandez-Ortega[1]

[1]BiDA Lab – Biometrics and Data Pattern Analytics Laboratory, Universidad Autonoma de Madrid, Madrid, Spain

✉ *E-mail: alejandro.acien@uam.es*

Abstract: This article studies user classification into children and adults according to their interaction with touchscreen devices. The authors analyse the performance of two sets of features derived from the sigma-lognormal theory of rapid human movements and global characterisation of touchscreen interaction. The authors propose an active detection approach aimed to continuously monitor the user patterns. The experimentation is conducted on a publicly available database with samples obtained from 89 children between 3 and 6 years old and 30 adults. The authors have used support vector machines algorithm to classify the resulting features into age groups. The sets of features are fused at the score level using data from smartphones and tablets. The results, with correct classification rates over 96%, show the discriminative ability of the proposed neuromotor-inspired features to classify age groups according to the interaction with touch devices. In the active detection set-up, the authors' method is able to identify a child using only four gestures in average.

1 Introduction

Touchscreen panels have changed the way users interact with new devices. They have increased their relevance because of their portability and ease of use. The touchscreen enables an intuitive experience of use that allows a direct interaction with what is being displayed. According to [1], the touchscreen market is growing ten times faster than other electronics markets and the revenue for these types of devices grew up from 4.3 to 23.4 billion dollars during the last decade.

Touchscreen devices provide mobile access to an unlimited number of digital content and services (e.g. more than half of YouTube visits come from mobile devices and this percentage is increasing [2]). Those digital services are used by people from everywhere, all ages, all ethnicities, and all socioeconomic status. In this context, the classification of users according to geographic and demographic attributes is crucial for the creation of accurate specific user profiles. Those profiles are used to generate digital identities of the users that can be exploited by services and platforms (e.g. recommender systems, parental control, security) [3]. Some of these attributes can be obtained from metadata associated with the device (e.g. IP address, language selection, GPS location) or can be inferred from the user behaviour (e.g. browsing history, social network contents, and keystroke dynamics) [4]. We want to highlight the spread of the use of this kind of devices by young children. The study in [5] reveals that 97% of US children under the age of four use mobile devices, regardless of family income.

In this paper, we analyse a way to classify users of touch panels according to two age groups (children and adults). Furthermore, we implement a novel active user detection (AUD) algorithm for age prediction. In this paper, we apply AUD algorithms in order to take advantage of physiological and behavioural mannerism of the user while interacting with a touchscreen device. Such mannerisms are often distinctive among different users; they are stable over a period of time and difficult to mimic [6]. Therefore, AUD systems are well protected against spoofing or hacking and recent works have shown that they report better results than single user authentication systems [7]. The main objective of the AUD method is to detect an intruder (children in our case) with the minimum possible delay from the moment he or she starts using a touch-based device. Thanks to the usage of AUD systems, in this context, it would be possible, for example, to adequate the content shown on the screen for the new user profile instantly, avoiding locking the session and asking for a password or traditional parental control systems.

The age is a human attribute which belongs to soft biometrics, like the ethnicity or the gender. According to the classification proposed in related works [8], soft biometrics traits can be divided into physical (e.g. age, ethnicity, gender), behavioural (e.g. mood, stress), and adhered human characteristics. These characteristics provide information about an individual and aim to recognise individuals, but they are not fully capable to distinguish between them due to a lack of distinctiveness and permanence.

Besides, the age is a key attribute in user profiling with direct application on several automatic systems (e.g. parental control, recommender systems, advertising). The most popular way to know the age of the user is by using online questionnaires in which the user directly reveals his age. However, this solution assumes: (i) honesty on the response of the users, and (ii) that users can read. Both assumptions cannot be guaranteed because of many practical reasons. Besides the fact that people lie, nowadays children start to use digital platforms and services before learning to read.

In the existing literature, there are many experiments exploring the use of technology by children, seeking how to improve the design of adapted interfaces and applications [9]. However, modelling and characterising mathematically how children interact with touch devices and how their conduct differs from the adult's one is a field that has not been studied deeply enough. A work related to this topic is [10] where the authors analysed different types of touching tasks like tap, rotate, or drag and drop, and they found that children have different success rates when trying to perform different tasks. Simple tasks (e.g. swiping, tapping) can be done by all children without any problem, but the more complex ones are very difficult to complete for short age children. In [11], researchers measured the touch gestures of children and compared them to gestures from adults. They discovered that children have a larger miss rate compared to adults when trying to hit small targets. The difference between adults and children is mainly caused by the different grade of maturity of their anatomy and neuromotor system. These features are less mature in children, so they have worse manual dexterity causing rougher movements [12, 13].

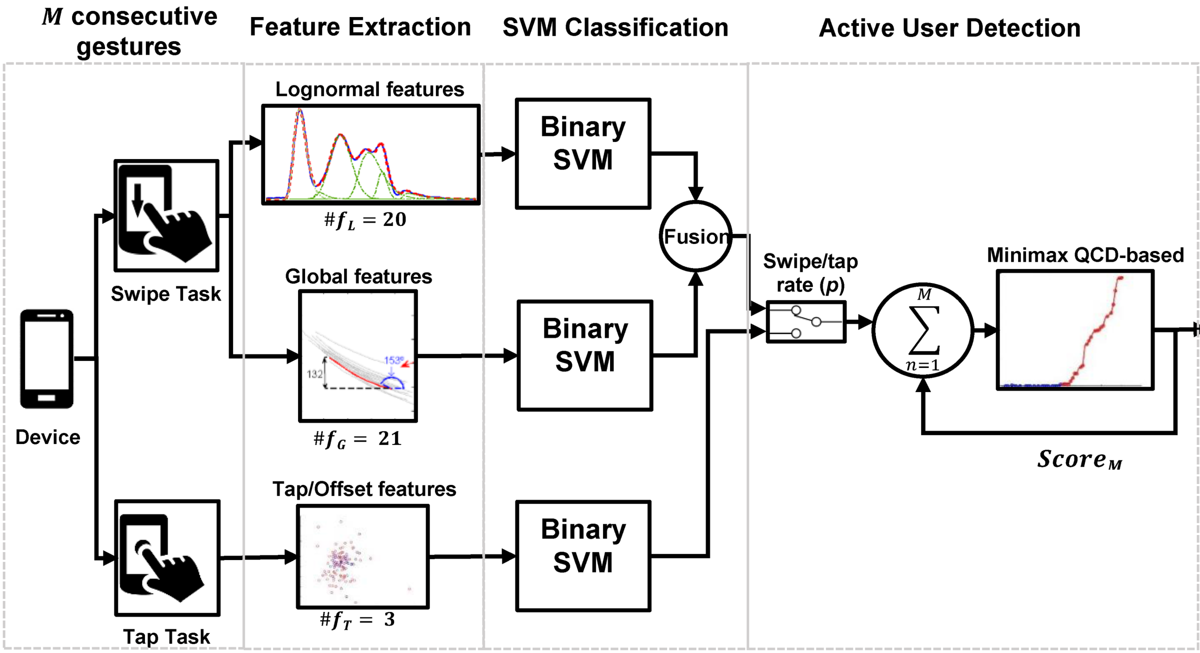

Fig. 1 *System architecture. For each m of M input consecutive touch gestures, three feature sets are generated: Lognormal (f_L), Global (f_G), and Tap/Offset (f_T)*

In a complementary case of our study [14], authors showed high classification rates between young adults (20–50 years) and older adults (70+ years) based on touch gestures, demonstrating differences in neuromotor skills during human ageing while our work studies differences between undeveloped neuromotor skills in children and total maturity in young adults.

In [15], researches show that people with a long thumb complete swipe gestures over a smartphone faster than those with shorter thumbs. This could be a key to identify children due to their shorter thumbs and therefore longer time.

In this paper, we will analyse two different types of touchscreen gestures: swipe and tap. In swipe tasks, users slide their finger over the screen, while tap tasks consist on tapping the touchscreen for a short period of time. We choose these gestures because they are the most common ones in touchscreen interaction and they are easy to be performed by children. To do this, we use information of swipe and tap patterns from a publicly available database presented in [16] comprised 119 subjects (89 children and 30 adults) using two different types of devices: a smartphone and a tablet.

Related to handwriting models, global features have shown to achieve good performance for writer identification (e.g. signatures, doodles) [17], but they have never been used to classify users regarding their age, so we will analyse whether they are suitable for this purpose. Global features refer to features extracted from the whole touchscreen patterns, such as the mean velocity or max acceleration among others.

This work extends previous research [18, 19] by: (i) studying user age group classification based on the combination (at the feature level and score level) of neuromotor characteristics with global features obtained from touch interaction; and (ii) combining swipe and tap tasks following an AUD framework for different use cases. Classification results show accuracies over 90% in several scenarios with top correct classification rates of 96%. In AUD, the algorithm shows to be able to detect children interaction in four gestures in average. The system proposed in this work is compared with the method proposed in [20]. The results show a performance improvement over 6%.

The rest of the paper is organised as follows: Section 2 depicts the systems architecture. Section 3 presents the experimental work carried out. Section 4 analyses the results obtained and Section 5 presents the conclusions and goals achieved in this work.

2 System description

The architecture proposed is divided into three consecutive stages (see Fig. 1 for details): feature extraction to compute the most suitable features for each touch-based task, SVM classification to classify between child an adult for one-time task detection, and AUD where a sequence of touch tasks performed during the interaction with the device is taken into account to decide whether it has been produced by an adult or a child.

2.1 Feature extraction

The feature extraction approach followed in this paper depends on the task performed: tap or swipe. For swipe tasks, we analyse the performance of two feature approaches, one based on the sigma-lognormal model and a different one based on global features. It is worth noting that sigma-lognormal features extract information related to neuromotor skills involved in the action performed; meanwhile, global features extract holistic information from the trajectory of the gestures. For tap tasks, we employ the feature approach based on tap/offset features proposed in [20].

2.1.1 Sigma-lognormal features: The sigma-lognormal theory of rapid human movements allows us to represent complex movements with an analytic model that describes some physical and cognitive features of human beings [21, 22]. This model has been successfully applied to handwriting tasks like handwritten signature [23, 24].

The sigma-lognormal model decomposes the complex signals that describe the speed of muscular movements into simpler ones that can be explained by few parameters. These parameters contain information about the activity itself and about the neuromotor skills of the person [25].

Studies like [25, 26] have proved that the sigma-lognormal model can be used to characterise children handwriting. They conclude there are two main groups of children that are separable by looking at their learning stage. Children's neuromotor skills become more similar to the adults' skills when they grow up, namely, when they finish their preoperational stage. At age 10, children know how to activate each little muscle properly to produce determinate fine movements [16]. As they are based on the same neuromotor skills, the principles applied to the handwriting models can be used to model touchscreen patterns.

The sigma-lognormal model [13] states that the velocity profile of human hand movements can be decomposed into strokes. Moreover, the velocity of each of these strokes, i, can be described with a speed signal $v_i(t)$ that has a lognormal shape

Table 1 Sigma-lognormal parameters description

Parameter	Description
D_i	input pulse: covered distance when executed isolated
t_{0i}	initialisation time=displacement in the time axis
μ_i	logtemporal delay
σ_i	impulse response time of the neuromotor system
θ_{si}	initial angular position of the stroke
θ_{ei}	final angular position of the stroke

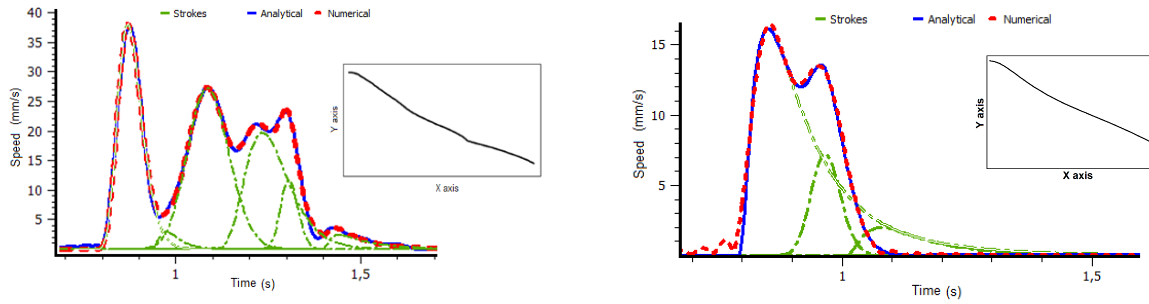

Fig. 2 *Example child (left) and adult (right) speed profiles from a touchscreen pattern (swipe). Numerical: captured velocity signal $|v(t)|$ of the touch activity (input of the model). Analytical: reconstructed sigma-lognormal velocity $v_r(t)$ profile (output of the model). Strokes: decomposition in individual strokes of the model*

$$|v_i(t)| = \frac{D_i}{\sqrt{2\pi}\sigma_i(t - t_{0i})}\exp\left(\frac{(\ln(t - t_{0i}) - \mu_i)^2}{2\pi\sigma_i^2}\right) \tag{1}$$

where each of the parameters are described in Table 1. The complete velocity profile is modelled as a sum of the different individual stroke velocity profiles as

$$v_r(t) = \sum_{i=0}^{N} v_i(t) \tag{2}$$

where N is the number of lognormals of the entire movement. A complex action, like a handwritten signature or touch task, is a summation of these lognormals, each one characterised by different values for the six parameters in Table 1.

In Fig. 2, the speed profiles of example touchscreen swipe patterns from adult and child are shown (numerical refers to the data acquired by the device while analytical refers to the velocity profile obtained by the sigma-lognormal model). This numerical signal is used as input of the sigma-lognormal model [13]. The analytical signal is calculated using the sigma-lognormal parameters extracted from the numerical signal. The parameters of the sigma-lognormal model have been adapted to calculate 18 different features that represent the neuromotor properties of the users [21]. The task time and the number of lognormals in each task have been added as features number 19 and 20. These features can be classified into three different groups (see Table 2). A visual comparison between children and adults speed profiles from Fig. 2 shows that children speed signals are usually composed of a higher number of strokes than the adult's signals. The larger maturity on the neuromotor skills of adults produces soft velocity profiles revealing a fine control of the movements:

Table 2 uses the following time variables: t_{2i}, t_{3i}, t_{4i} are zeros in the first and second derivative of (1). $t_{1i} = t_{0i} + \exp(\mu_i - 3\sigma_i)$ and $t_{5i} = t_{0i} + \exp(\mu_i + 3\sigma_i)$. They are chosen to make the $[t_{1i}, t_{5i}]$ interval to contain the 99.97% of the area under the lognormal signal.

Every task is composed of at least one lognormal. In this work, the values of the features for the different lognormals of a given task have been combined by computing the arithmetic mean of the features obtained from each $|v_i(t)|$.

Table 2 Sigma-lognormal extracted features

Space-based features
$f_1 = D_i$
$f_2 = \mu_i$
$f_3 = \sigma_i$
$f_4 = \sin(\theta_{si})$
$f_5 = \cos(\theta_{si})$
$f_6 = \sin(\theta_{ei})$
$f_7 = \cos(\theta_{ei})$

Time-based features
$f_8 = \Delta t_0 = t_{0i} - t_{0i-1}$
$f_9 = v_2 = \lvert v_i(t_{2i})\rvert$
$f_{10} = v_3 = \lvert v_i(t_{3i})\rvert$
$f_{11} = v_4 = \lvert v_i(t_{4i})\rvert$
$f_{12} = \delta t_{05} = t_{5i} - t_{0i}$
$f_{13} = \delta t_{15} = t_{5i} - t_{1i}$
$f_{14} = \delta t_{13} = t_{3i} - t_{1i}$
$f_{15} = \delta t_{35} = t_{5i} - t_{3i}$
$f_{16} = \delta t_{24} = t_{4i} - t_{2i}$
$f_{17} = \Delta t_1 = t_{1i} - t_{1i-1}$
$f_{18} = \Delta t_3 = t_{3i} - t_{3i-1}$

General features
f_{19} = task time
f_{20} = of log normals

Features are calculated for each lognormal of the decomposition of the numerical signal

2.1.2 Global features: The global features set includes features calculated from the entire touch task pattern, such as the mean velocity, max acceleration, or total duration. For this purpose, many global features have been proposed in the literature [27–29] for signature verification. We use the 28-dimensional features set applied in [30] due to good results obtained in swipe patterns, but with minor adjustments (in particular: pressure measurements are not available in the dataset used here). After removing the features related to pressure, a 21-dimensional feature vector was computed, as shown in Table 3.

2.1.3 Tap/offset features: Tap tasks are characterised by two features: (i) the distance between the target point and the point touched (offset-distance); and (ii) the time that the user touches the screen during the tap task (tap-time) [20].

2.2 Classification

In this stage, we use feature selection (based on sequential forward selection) to calculate the best subset of features for each feature set in the swipe task. The best subset for the lognormal features resulted: $f_1, f_2,\ f_9,\ f_{19}, f_{20}$. The best subset of global features is: $v',\ v_{2st}, v_{3st}, a_{2st}, a_{3st}$. As a classifier, we use support vector machines (SVM) with a radial basis function (RBF) kernel (with $C\ =\ 30$ and $\sigma\ \ =\ 10$) because of its good general performance in binary classification tasks. As shown in Fig. 1, three binary SVM classifiers were implemented, one for each feature set. We also compared SVM-RBF to decision trees, KNNs, and logistic regression; all of them resulting in worse performance.

Each SVM is trained using samples from children and adults over the training data (users sequestered for the test phase). The two approaches proposed for the swipe task achieved similar performance independently (as we will see in Section 4). We also perform the fusion of both approaches both at the feature and score level. For fusion at the feature level, both sets of features were concatenated to form the final feature vector. For fusion at the score level, the final score is obtained as the average of the previous two.

2.3 Active user detection

The AUD algorithm proposed in [6] for intrusion detection is adapted in this work for children detection. The algorithm is based on calculating a new score from the cumulative sum of previous events (swipe or tap in this paper). If the user is an adult (genuine), the cumulative sum will be almost zero. At the moment the user profile changes into a child (intruder), this score will tend to increase until reaching a certain threshold, in which the user will be detected as a child. The algorithm starts with the calculation of the log-likelihood ratio L_m

$$L_m = \log\left(\frac{f_c(x_m)}{f_a(x_m)}\right) \tag{3}$$

where f_a and f_c are the probability distributions of the SVM calculated previously on the classification stage with the training set, and x_m is the score of the event (touch gesture) number m from a test user. In Fig. 3, an example is shown with a sequence of events. It starts with an adult and changes into a child at the middle of the sequence. If $f_a(x_m)$ is higher than $f_c(x_m)$, this means that the event x_m is likely made by an adult and therefore L_m will be a negative number (see Fig. 4). By the moment the user profile changes into a child (event number 21 in Fig. 3), L_m will be positive and the cumulative sum will start to increase.

The aim is to reach a certain threshold and detect the child with the minimum number of touch gestures M. The cumulative sum calculates the score taking into account past events

$$\text{score}_M = \max(\text{score}_{M-1} + L_M, 0) \tag{4}$$

for $M \geq 1, \text{and}\, \text{score}_0 = 0$. Note that score_M is set to have a minimum value of 0, because otherwise, several adult sequences could decrease the value too much making necessary to insert a higher number of children samples to reach the decision threshold.

Fig. 3 shows an example of a sequence of events with 40 samples (20 from adult X and 20 from children Y). The events are chosen randomly between swipes and taps with p probability (tap/swipe rate). The curve shows that score_M is close to zero with adult samples and it tends to increase at the moment child samples start. The selection of the threshold to calculate when the user is detected as a child is crucial in performance terms: a high threshold could decrease the number of false detections (adults detected as a child), but also it could increase the time delay (time between the child starts to operate the device till he is detected). In order to choose the best threshold, we will present average detection delay (ADD) and probability of false detection (PFD) curves, previously used in [31]. ADD curves show the number of samples necessary to detect a child as a function of the threshold. On the other hand, PFD curves depict the percentage of false detection. Additionally, we add probability of non-detection (PND) to depict the percentage of children who are not detected by the system.

3 Experiments

3.1 Database

The experimental protocol is based on the publicly available database presented in [16]. It is a database with touchscreen activity of both children and adults performing predesigned tasks in an *ad hoc* app.

The database comprises samples from different guided activities such as tap, double tap, and drag and drop (swipe) tasks. In the present work, we have used the data from swipe and tap activities. Swipe activities consist in picking one object on the device screen and moving it to a target area. Tap activities consist in touching the screen over a target area for a moment. These tasks have been selected because they are simple and common neuromotor tasks as they consist in moving the finger on a surface and also they are widespread gestures in touchscreen device interaction. Multidevice information is available as long as the users have completed the tasks in both a smartphone and a tablet. This allows testing our solution in both scenarios.

The data set is composed of 89 children between 3 and 6 years old and 30 young adults under 25 years old. The mean age of the children is 4.6 years. The total number of samples (touch gestures) is 2912 for children and 1157 for adults (see [16] for more details). To the best of our knowledge, this is the largest database in the field of interaction with touchscreen technology with children under 6 years old.

The main issue when acquiring data from children activity is to maintain the kids' attention during a long time period. The authors of the database have adapted the activities interfaces to make the tasks more interesting to children. Thanks to this, they have managed to obtain a completion rate near 100% in tap tasks and >90% in all types of tasks.

3.2 Experimental protocol

The experiments were divided into two well-differentiated scenarios: one-time detection (OTD) and active user detection (AUD).

In OTD only one sample (touch gesture) is used to discriminate among children and adults. The users have been separated randomly in training (60%) and test (40%). It is guaranteed that users (children and adults) employed for training are different from those employed for test. The experiments were repeated 50 times and the final performance is presented in terms of average correct classification rate computed as $100 - \text{Error Rate (EER)}$. EER refers to the value where false match rate (percentage of children classified as adult) and false non-match rate (percentage of adults classified as children) are equal. Owing to the higher number of children tasks in the database compared with the adults, selecting a percentage of the total users makes the two scenarios to be unbalanced. Experiments balancing the number of both classes in training and testing processes have been made. Nevertheless, the results show small variations around 1% of accuracy (variation that can be related to the statistical variation due to the data set).

On the other hand, AUD experiments take a sequence of events during a period of time to detect a change in the user profile (from adult to child or vice versa). It is worth highlighting that active detection is based on the one-time detection system (see Fig. 1), so the performance of the first one is crucial for both scenarios.

Table 3 Global features set

Parameter	Description
T	task time
$v', v_{\sigma}, v_{1st}, v_{2st}, v_{3st}$	velocity: mean, standard deviation, first quartile, second quartile, and third quartile
$a', a_{\sigma}, a_{1st}, a_{2st}, a_{3st}$	acceleration: mean, standard deviation, first quartile, second quartile, and third quartile
d_{N-1}	distance between end points
θ	angle between line and horizontal axis
$\sum_{i=0}^{N} d_i$	summation of distance between adjacent points
d_x, d_y	distances between mean and min point
σ_{ax}, σ_{ay}	standard deviation of *x*- and *y*-axis acceleration.
H, V	horizontal and vertical span ratio
A	swipe area

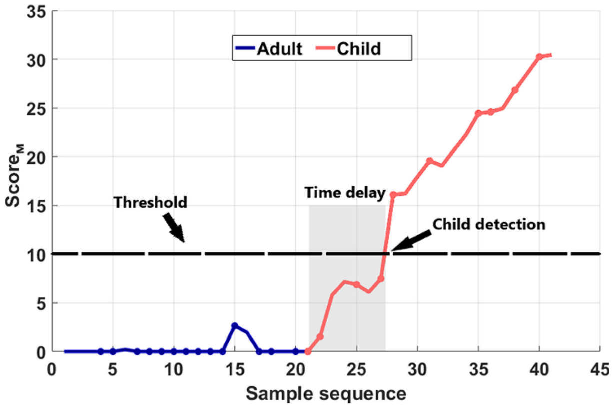

Fig. 3 *Example of QCD-based curve with a sequence of 40 samples (20 touch gestures from an adult and 20 from a children) with p = 0.5 (dots are tap samples, the rest are swipe gestures)*

3.2.1 One-time detection (OTD): A separate SVM is trained for each set of features: sigma-lognormal and global features for the swipe task, and tap/offset features for the tap task.

Swipe tasks fit into the sigma-lognormal model as they are composed of precision movements over time. The parameters of the model (see Table 1) for all tasks were extracted using the Script Studio software [13]. This software provides the $6 \times N$ parameters for each swipe task. N is the number of lognormals and it is automatically calculated by Script Studio according to the input data (coordinates x, y and their respective timestamps). In this work, we process the parameters of the model to remove the smallest lognormals. This postprocessing is intended to discard small lognormal signals that do not have the same importance as the big ones in order to differentiate between children and adults. Any lognormal with a maximum amplitude (which occurs in t_3) under a specified threshold will be removed (in this work, we use the mean value of v_3 across all the lognormals of the task as threshold).

Note that all lognormal signals including the small ones would be necessary to reconstruct the original stroke accurately but not to perform a distinction between users, as the smallest signals will have a negligible impact in the mean values of each task.

After reducing the number of lognormal signals, the features in Table 2 are computed. The features have been computed for all the lognormal signals in each swipe gesture for a total number of features equal to $18 \times N' + 2$. Where N' is the number of lognormals validated after the postprocessing ($N' \leq N$) and the two remaining features are the number of lognormals and the task time. Finally, each feature is averaged (using the N' values) to obtain a feature vector of dimension 20 for the swipe task.

The global features set (see Table 3) is also extracted for swipe tasks. Both sets (lognormal and global) are compared and combined using: fusion at the score level and fusion at the feature level [31].

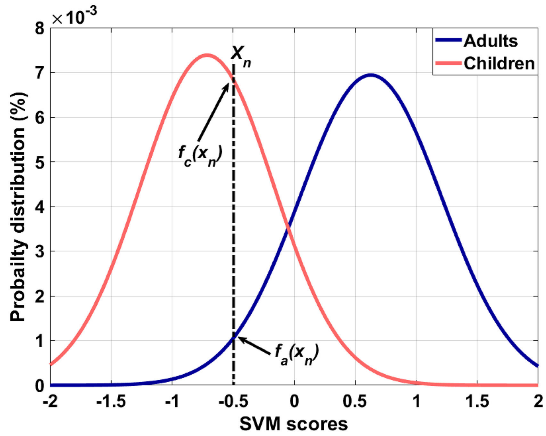

Fig. 4 *Probability distribution of adults and children for tap task. The score x_n shows that $f_c(x_n)$ is higher than $f_a(x_n)$ and the log-likelihood ratio L_n will be positive*

3.2.2 Active user detection (AUD): These experiments simulate a human device interaction with a sequence of swipe and tap gestures combined in order to detect as soon as possible if the device is being used by a child. To simulate this change in the user profile, we build test sequences with the first half of adult samples and the second half of child samples. Moreover, the rate between taps and swipes in each sequence can vary depending on the application used (remember that each sample of the sequence could be a tap or swipe); for instance, in reading applications, swipe gestures are more common than taps gestures; meanwhile, in videogames, it would be the opposite. The different combinations have a significant impact in results due to tap gestures have a worse performance in one-time user detection and it could yield a drop of performance for active detection in those sequences where tap gestures are more common. To analyse this effect, we separate the experiments into different scenarios taking into account the percentage between tap and swipe gestures in the test sequences.

Finally, we compute an error rate in order to compare among all scenarios and devices proposed in this paper.

4 Results

4.1 One-time detection

In Table 4, we summarise all experiments performed in this section. For swipe tasks, the best result was achieved by the fusion at the score level of both feature sets. The mean value of correct classification rate having into account all evaluated scenarios and both devices is 94.4%. The best results are obtained with tablets as sensors, while when using smartphone's data, slightly worse results are achieved. Note that swipe gestures have longer trajectories in tablet screens compared to smartphone screens. Larger movements imply more information available to classify users. However, tap gestures do not take advantage of the screen's size due to the target point has the same size in both devices (remember that the distance between the target point and the point touched is one of the three features in the tap/offset feature set).

Fig. 5 shows the probability distribution functions of the scores calculated in the classification process for both swipe and tap tasks in the phone device. For the swipe task, scores from children and adults are visibly separated into two different zones, making possible to get high accuracy rates (over 93%). Note there are zones where the score distributions overlap. These regions are the source of incorrect classifications. Combining scores from several samples (AUD) of the same user could reduce the overlapping areas, increasing even more the accuracy rate. For the tap task, both probability distributions show larger overlap causing a worse performance.

Table 4 also shows the classification accuracies obtained from the method proposed in [20]. The improvement in accuracy rate can be associated with the better discrimination due to a combination of lognormal and global features that describes better

Table 4 Results achieved for each one-time detection: correct classification rate (%)

Device	Our implementation of [20]	Swipe				Tap
		Sigma-lognormal	Global	Feature fusion	Score fusion	Tap/offset
phone	86.5	93.6	92.1	92.7	94.1	85.4
tablet	90.5	96.3	94.5	94.9	96.5	80.0

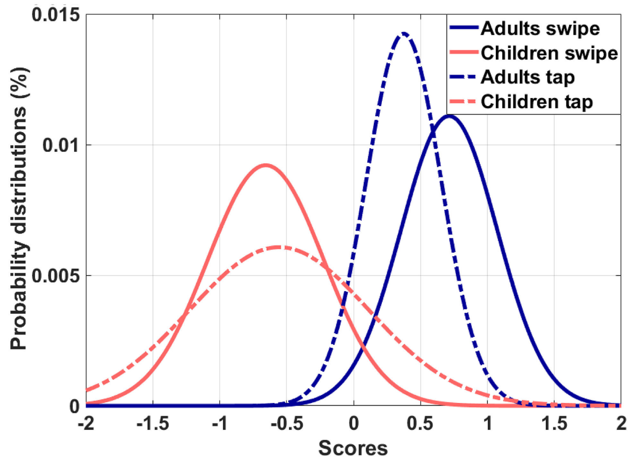

Fig. 5 *Probability distribution of adults and children for tap and swipe tasks with phone device*

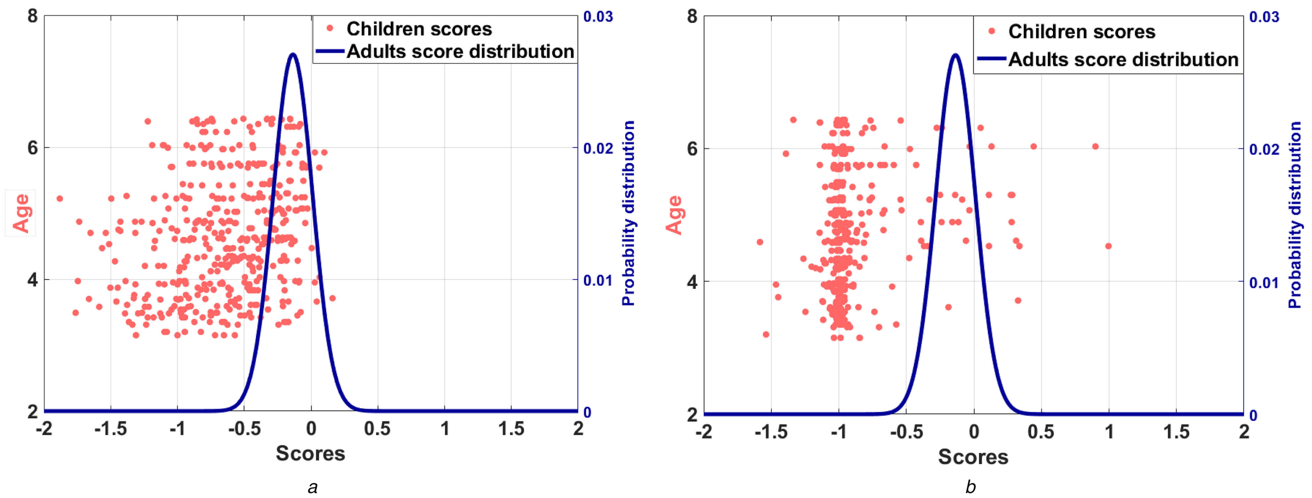

Fig. 6 *Probability distribution of adults (right y-axis) and children scores sorted by age (left y-axis) for tap task (left) and swipe task (right) with tablet device*

touch-based gestures, while in [20], their features are basically related to the precision of the gestures.

Finally, the children age is a key factor affecting the classification performance. Neuromotor skills in children become more similar to the adult ones as they grow up. At the age of ten, children have their neuromotor skills completely developed making the classification task more complicated [25]. In order to analyse the impact of the children growth, we compare children scores by age obtained in the classification task with the distributions of adults scores, expecting to observe a similarity between adults and children as children are older. Fig. 6 shows the scores obtained by the SVM of the children in tap and swipe tasks sorted by age (left *y*-axis). The distribution of adult scores is also plotted for comparison. We can observe that children scores are closer to the adult ones as children grow up due to maturity of their neuromotor skills, especially for taps. We divide children into three age groups: under 4 years old, between 4 and 5, and older than 5; the accuracies in the tablet device for taps in those groups are 85.0, 80.8, and 76.3%, respectively; and 98.2, 96.5, and 93.2%, respectively, for swipes.

4.2 Active user detection

As mentioned before, the swipe/task rate (p) could be crucial for AUD results due to worse performance in tap tasks. Real user interaction with touch screens involves swipe and tap gestures. Thus, we decided to distinguish among three scenarios taking into account different percentages of swipe and tap events:

- First scenario ($p \leq 0.25$): test sequence with a 25% or less of swipe gestures.
- Second scenario ($0.25 < p < 0.75$): test sequences with swipes and tap gestures balanced.
- Third scenario ($p \geq 0.75$): test sequences with a majority of swipe gestures.

The swipe and tap gestures from each user are randomly chosen for all three scenarios to build each test sequence. The experiments are repeated up to 100 times, so we have 100 different test sequences for each user. In order to analyse the AUD system performance, we present average detection delay (ADD), probability of false detection (PFD), and probability of non-detection (PND) curves shown in Fig. 7. All curves were calculated individually for each user, and finally averaged.

PFD curves show how many adults in percentage are identified as children (false detections). All scenarios show similar PFD performance; as it is expected, false detections decrease when thresholds increase. Furthermore, the third scenario ($p > 0.75$), where swipes gestures are more common, decrease faster due to

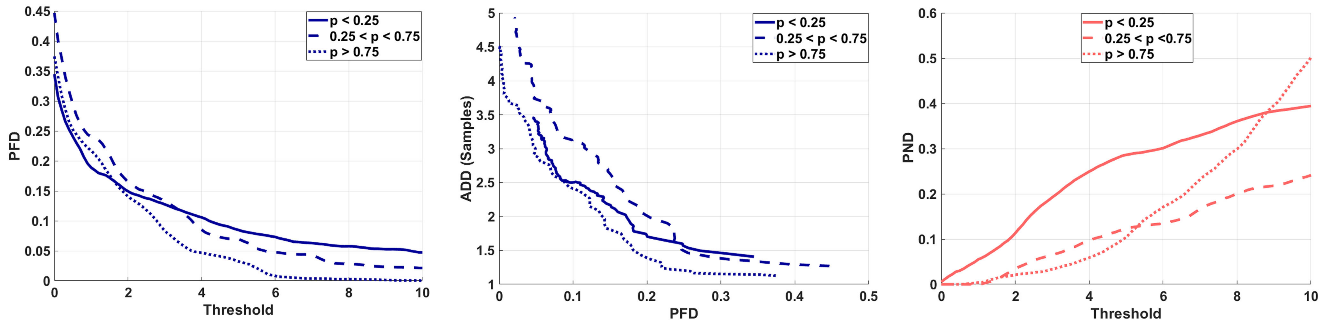

Fig. 7 *Probability of false detection (PFD, left), average detection delay (ADD) vs PFD (middle), and probability of non-detection (PND) for smartphone*

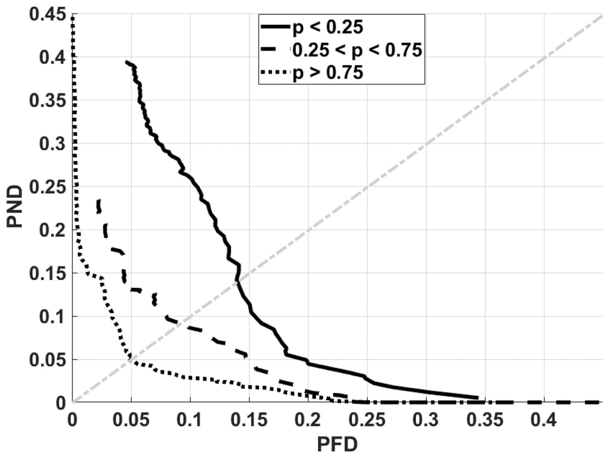

Fig. 8 *Probability of non-detection (PND) vs probability of false detection (PFD) with smartphone device. Points where curves cross the grey line are the EER values*

Table 5 Results achieved in correct classification rate (%) for both one-time detection and active user detection

Device	Active user detection					One-time detection	
	$p = 0$	$p < 0.25$	$0.25 < p < 0.75$	$p > 0.75$	$p = 1$	Swipe	Tap
phone	86.2	86.2	91.1	94.5	95.6	94.1	85.4
tablet	82.3	82.7	89.5	95.0	97	96.5	80.0

having a better performance in one-time user detection for swipe, so false detection are relatively uncommon among swipe gestures.

In addition, ADD–PFD curves denote how many samples are necessary to identify a child on average depending on false adult detections. This quantity is a significant factor to take into account when an AUD system is designed. The system tries to identify a child with the minimum amount of samples as possible in order to reduce the time delay (time between the child starts to operate the device till he is detected) but avoiding false detections as well. It can be seen that the number of samples necessary to identify a child increases when we decrease the false detection, so there is always a trade-off between both curves. The ADD curve for the third scenario ($p > 0.75$) has again the best performance.

PND curves depict the percentage of children which are not detected by the system. In this case, the first scenario ($p < 0.25$) increases faster with the threshold. Regarding the third scenario, we expected to achieve the best results, but it tends to obtain the highest PND rate with high thresholds. The main reason for this effect is that the lack of swipe samples in some children suggests that they would never be detected by the system for high thresholds. Note that the third scenario has the best results for low thresholds where few samples are enough to reach it. It can be seen that there is always a trade-off between false child detection (PFD) and non-child detection (PND), as it is possible to decrease the false adult detection at the cost of having more children who are not detected by the system and vice versa. Therefore, performance will vary depending on the system design and application.

The PND–PFD curves shown in Fig. 8 are useful performance metrics to analyse AUD systems. Note that PND–PFD in active detection are similar to FNMR–FMR in one-time detection. The main difference is that PND and PFD curves are obtained from a sequence of stacked scores; meanwhile, FNMR–FMR come from one-time detection. In order to differentiate among both cases, we decided to keep the same nomenclature as in [6]. In these curves, we can appreciate the trade-off effects: reducing the false child detection rates (PFD) makes the system more prone to non-child detection (PND) as a consequence. In this figure, we can analyse better the p rates effects over the sample sequence. Having more swipe than tap gestures yields better performance as expected.

Table 5 summarises the correct classification rates, computed as the opposite of the EER. Each correct classification rate has been calculated independently for each device and p rates. This table shows that AUD results are slightly better in smartphones, specially for small p, because the taps there perform better than swipes. In fact, the difference between smartphone and tablet devices tends to decrease when tap gestures are less common.

Additionally, the correct classification rates for one-time detection is added to Table 5 for comparison. It can be seen that AUD results are always between swipe and tap gestures in one-time detection results: results in AUD where most of samples are swipes are close to swipes results in one-time detection; meanwhile, AUD improves one-time detection when tap gestures are more common. In fact, if we only consider swipe gestures ($p = 1$) or tap gestures ($p = 0$) in AUD, the results improve one-time detection marks, so AUD could improve one-time detection at the cost of needing more time to detect a child.

5 Conclusions and future work

In this paper, we have studied user classification into children and adults according to their interaction with touchscreen devices like smartphones and tablets. Furthermore, we present an active user detection (AUD) algorithm that takes advantage from the previous classifier results to identify children during a continuous interaction session.

Firstly, we have studied feature extraction based on the parameters of the sigma-lognormal. These features illustrate the neuromotor skills of users, making possible to discern between children and adults. An evaluation of performance has been made, using a public database with touchscreen activity of both children and adults. The classification rates are over 95% combining both sigma-lognormal and global features.

Secondly, we developed an AUD system aimed to detect a child with the minimum delay as possible from the time he starts to interact with the device. We simulate this situation generating sequences of tap and swipe gestures during a period of time. Depending on the type of sequences, our methods achieve accuracies ranging between 86 and 95%. These accuracies can be obtained using features extracted from only four simple gestures made by adults or children. Although our error rates are very low, the results should be interpreted with care. A major limitation of this work comes from the database used, as it does not contain data from users with ages between 6 and 25 years old. To the best of our knowledge, there is no such database with touch screen interaction data available to the research community.

As future work, we aim at building a reliable and fast classifier of users from all ages based on their touchscreen interaction [30] based on a combination of different expert systems [31]. The main drawback of other methods like using the browsing history or social network profiles is that they need a high amount of data. Our system allows us to classify users using data from simple and short tasks (ca. 2 seconds). tasks. This makes our solution suitable for applications that require classification on the fly.

As further improvements of our developments, two aspects can be taken into account. First, the patterns used in this work are very simple: swipe and tap gestures. Better classification rates may be achieved if the information comes from more complex tasks or from continuous monitoring. Second, this study includes the analysis of touch patterns from children under 6 years. However, how to recognise users with mature neuromotor skills (from 10 years old onwards) is a challenging task and new models and methodologies should be proposed for that purpose in the future. The classification of older users using the sigma-lognormal model is a possibility since it is demonstrated that the neuromotor abilities decay with the age [22].

6 Acknowledgments

This work was funded by the project CogniMetrics (TEC2015-70627-R) and Bio-Guard (Ayudas Fundación BBVA a Equipos de Investigación Científica 2017).

References

[1] 'CDS, Crystal Display System Ltd'. Available from http://crystal-display.com/touchscreen-market-growing-10-times-faster-than-other-displays/, retrieved 10 September 2017

[2] 'Youtube, Youtube statistics'. Available from https://about.twitter.com/company, retrieved 13 September 2016

[3] 'BBA: Mobile phone apps become the UK's number one way to bank'. Available from https://www.bba.org.uk/news/pressreleases/mobile-phone-apps-become-the-uks-numberone-way-to-bank, retrieved 13 September 2016

[4] Morales, A., Fierrez, J., Tolosana, R., *et al.*: 'Keystroke biometrics ongoing competition', *IEEE Access*, 2016, **4**, pp. 7736–7746

[5] Kabali, H. K., Irigoyen, M.M., Nunez-Davis, R., *et al.*: 'Exposure and use of mobile media devices by young children', *Pediatrics*, 2015, **136**, (6), pp. 1044–1050

[6] Perera, P., Patel, V.M.: 'Efficient and low latency detection of intruders in mobile active authentication', *IEEE Trans. Inf. Forensics Sec.*, 2018, **13**, (6), pp. 1392–1405

[7] Patel, V.M., Chellappa, R., Chandra, D., *et al.*: 'Continuous user authentication on mobile devices: recent progress and remaining', *IEE Signal Process. Mag.*, 2016, **33**, (4), pp. 49–61

[8] Gonzalez-Sosa, E., Fierrez, J., Vera-Rodriguez, R., *et al.*: 'Facial soft biometrics for recognition in the wild: recent works, annotation and COTS evaluation', *IEEE Trans. Inf. Forensics Sec.*, 2018, **13**, (8), pp. 2001–2014

[9] Cassidy, B., McKnight, L.: 'Children's interaction with mobile touch-screen devices: experiences and guidelines for design', *Int. J. Mob. Hum. Comput. Interact.*, 2010, **2**, pp. 1–18

[10] Aziz, N.A.A., Batmaz, F., Stone, R., *et al.*: 'Selection of touch gestures for children's applications'. Proc. of Science and Information Conf., IEEE, 2013, pp. 721–726

[11] Anthony, L., Brown, Q., Nias, J., *et al.*: 'Interaction and recognition challenges in interpreting children's touch and gesture input on mobile devices'. Proc. of the 2012 ACM Int. Conf. Interactive tabletops and surfaces, ACM, Cambridge, MA, USA, 2012, pp. 225–234

[12] Piaget, J., Inhelder, B.: '*The psychology of the child*' (Basic Books, France, 1969), 5001)

[13] O'Reilly, C., Plamondon, R.: 'Development of a sigma–lognormal representation for on-line signatures', *Pattern Recognit.*, 2009, **42**, (12), pp. 3324–3337

[14] AL-Showarah, S., AL-Jawad, N., Sellahewa, H.: 'User-age classification using touch gestures on smartphones', *Int. J. Multidiscip. Stud.*, 2015, **2**, (1), pp. 97–109

[15] Bevan, C., Fraser, D.S.: 'Different strokes for different folks? Revealing the physical characteristics of smartphone users from their swipe gestures', *Int. J. Human-Comput. Stud.*, 2016, **88**, pp. 51–61

[16] Vatavu, R.-D., Cramariuc, G., Schipor, D.M.: 'Touch interaction for children aged 3 to 6 years: experimental findings and relationship to motor skills', *Int. J. Human-Comput. Stud.*, 2015, **74**, pp. 54–76

[17] Fierrez, J., Ortega-Garcia, J.: 'On-line signature verification', in '*Handbook of biometrics*' (Springer-Verlag, Germany Berlin, 2008), pp. 189–209

[18] Hernandez-Ortega, J., Morales, A., Fierrez, J., *et al.*: 'Detecting age groups using touch interaction based on neuromotor characteristics', *IET Electron. Lett.*, 2017, **53**, (20), pp. 1349–1350

[19] Tolosana, R., Vera-Rodriguez, R., Fierrez, J., *et al.*: 'Benchmarking desktop and mobile handwriting across COTS devices: the e-BioSign biometric database', *PLOS ONE*, 2017, **12**, (5), pp. 1–17

[20] Vatavu, R.-D., Anthony, L., Brown, Q.: 'Child or adult? Inferring smartphone users age group from touch measurements alone', in '*Human–computer interaction*' (Springer, Cham, Switzerland, 2015), pp. 1–9

[21] Fischer, A., Plamondon, A.: 'A dissimilarity measure for on-line signature verification based on the sigmalognormal model'. Proc. of 17th Biennial Conf. of the Int. Graphonomics Society, Pointe-a-Pitre, Guadeloupe, 2015

[22] Plamondon, R., O'Reilly, C., Remi, C., *et al.*: 'The lognormal handwriter: learning, performing, and declining', *Front. Psychol.*, 2013, **4**, pp. 945:1–945:14

[23] Galbally, J., Plamondon, R., Fierrez, J.: 'Synthetic on-line signature generation. Part I: methodology and algorithms', *Pattern Recognit.*, 2012, **45**, pp. 2610–2621

[24] Ferrer, M.A., Diaz-Cabrera, M., Morales, A.: 'Static signature synthesis: a neuromotor inspired approach for biometrics', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2015, **37**, (3), pp. 667–680

[25] Duval, T., Rémi, C., Plamondon, R., *et al.*: 'Combining sigma-lognormal modeling and classical features for analyzing graphomotor performances in kindergarten children', *Hum. Mov. Sci.*, 2015, **43**, pp. 183–200

[26] Meulenbroek, R., VanGalen, P.G.: 'The acquisition of skilled handwriting: discontinuous trends in kinematic variables', in Colley, A.M., Beechs, J.R. (eds.): '*Cognition and action in skilled behavior*' (North Holland, Amsterdam, 1998), pp. 273–281

[27] Serwadda, A., Phoha, V.V., Wang, Z.: 'Which verifiers work?: a benchmark evaluation of touch-based authentication algorithms'. Proc. of IEEE BTAS, Washington DC, USA, 2013, pp. 1–8

[28] Martinez-Diaz, M., Fierrez, J., Krish, R.P., *et al.*: 'Mobile signature verification: feature robustness and performance comparison', *IET Biometrics*, 2014, **3**, (4), pp. 267–277

[29] Martinez-Diaz, M., Fierrez, J., Hangai, S.: 'Signature features', in '*Encyclopedia of biometrics*' (Springer, Boston, MA, 2015), pp. 1375–1382

[30] Fierrez, J., Pozo, A., Martinez-Diaz, M., *et al.*: 'Benchmarking touchscreen biometrics for mobile authentication, *IEEE Trans. Inf. Forensics Sec.*, 2018, **13**, (11), pp. 2720–2733

[31] Fierrez, J., Morales, A., Vera-Rodriguez, R., *et al.*: 'Multiple classifiers in biometrics. Part 1: Fundamentals and review', *Information Fusion*, 2018, **44**, pp. 57–64