

Kent Academic Repository

Full text document (pdf)

Citation for published version

Miguel-Hurtado, Oscar and Stevenage, Sarah V. and Bevan, Chris and Guest, Richard (2016) Predicting sex as a soft-biometrics from device interaction swipe gestures. *Pattern Recognition Letters*, 79 . pp. 44-51. ISSN 0167-8655.

DOI

<https://doi.org/10.1016/j.patrec.2016.04.024>

Link to record in KAR

<http://kar.kent.ac.uk/55702/>

Document Version

Publisher pdf

Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research

The version in the Kent Academic Repository may differ from the final published version.

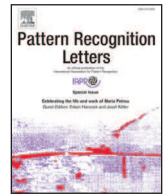
Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

Enquiries

For any further enquiries regarding the licence status of this document, please contact:

researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>



Predicting sex as a soft-biometrics from device interaction swipe gestures[☆]



Oscar Miguel-Hurtado^{a,*}, Sarah V. Stevenage^b, Chris Bevan^c, Richard Guest^a

^aSchool of Engineering and Digital Arts, University of Kent, Canterbury, UK

^bDepartment of Psychology, University of Southampton, Southampton, UK

^cCREATE Laboratory, University of Bath, Bath, UK

ARTICLE INFO

Article history:

Received 15 October 2015

Available online 17 May 2016

Keywords:

Soft-biometrics

Sex prediction

Swipe gestures

Feature selection

Classifiers

ABSTRACT

Touch and multi-touch gestures are becoming the most common way to interact with technology such as smart phones, tablets and other mobile devices. The latest touch-screen input capacities have tremendously increased the quantity and quality of available gesture data, which has led to the exploration of its use in multiple disciplines from psychology to biometrics. Following research studies undertaken in similar modalities such as keystroke and mouse usage biometrics, the present work proposes the use of swipe gesture data for the prediction of soft-biometrics, specifically the user's sex. This paper details the software and protocol used for the data collection, the feature set extracted and subsequent machine learning analysis. Within this analysis, the BestFirst feature selection technique and classification algorithms (naïve Bayes, logistic regression, support vector machine and decision tree) have been tested. The results of this exploratory analysis have confirmed the possibility of sex prediction from the swipe gesture data, obtaining an encouraging 78% accuracy rate using swipe gesture data from two different directions. These results will hopefully encourage further research in this area, where the prediction of soft-biometrics traits from swipe gesture data can play an important role in enhancing the authentication processes based on touch-screen devices.

© 2016 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Soft-biometrics traits are defined as “anatomical or behavioural characteristics that provides some information about the identity of a person, but does not provide sufficient evidence to precisely determine the identity” [1]. They include characteristics such as age, ethnicity, sex¹, height, weight, scars and tattoos. These traits have been used within biometrics deployment in combination with hard-biometrics modalities such as fingerprint [2], iris [3] and face [4]. Studies have shown that the use of a soft-biometrics can enhance biometrics system performance and can greatly decrease search time in large databases [1]. Whilst soft-biometrics are high level cues and as such are largely incapable of

differentiating one individual from another, their use has been proposed for deployment within continuous authentication scenarios [5]. In these cases, hard-biometrics techniques are used for initial authentication, while a combination of soft-biometrics traits are used to continuously authenticate the subject.

In parallel, there is also a growing interest in using soft-biometrics in non-biometrics scenarios. The Human-Computer Interaction (HCI) community are looking to the prediction of traits (e.g. sex, age, handedness, emotional states) as a way to enhance the interaction between computer-based systems and users [6]. For example, the use of dynamic keystroke and mouse movement information has been suggested as a way to predict the level of stress of computer users [7]. There has also been attempts to predict emotional states such as happiness from dynamic keystroke data [8]. These kind of predictions may able to provide enhanced interfaces tailored to specific demographic groups and/or different emotional states.

Human-computer interaction based on touch-screen technology can be dated back to 1965 when Johnson published his work [9] using wires on a CRT device. Recently, the growth of mobile technologies such as smartphone has resulted in the ubiquity of touch as an interface methodology. Moreover, the

[☆] This paper has been recommended for acceptance by Maria De Marsico.

* Corresponding author. Tel.: +44 1227 764000; fax: +44 1227 764000.

E-mail addresses: O.Miguel-Hurtado-98@kent.ac.uk, omiguelh@gmail.com (O. Miguel-Hurtado), S.V.Stevenage@soton.ac.uk (S.V. Stevenage), c.r.bevan@bath.ac.uk (C. Bevan), R.M.Guest@kent.ac.uk (R. Guest).

¹ Sex prediction is commonly named “gender prediction” within the biometrics community. Sex refers to the biological characteristic of a person, while gender refers to the sociocultural roles. Due to this, the authors prefer the use of “sex prediction”.

latest generation of touch-screen devices have enhanced input capabilities tremendously, which raises the possibility of using touch input data in order to predict information about the user (so called ‘soft-biometrics’). These capture capabilities have been used to develop more secured means of authentication through swipe gestures [10–13]. However, as far as the authors’ knowledge, there has not been any previous related work in soft-biometrics prediction based on touch gestures data.

In an analogous way to how keystroke and mouse data has been proposed to enhance the human computer interaction [7,8], we believe that it is possible to extract soft-biometrics traits from swipe gesture information. The predicted soft-biometrics traits can allow touchscreen computers-based systems to tailor their interaction to better suit the user’s characteristics. In addition, this information could also improve the performance of continuous authentication biometrics systems deployed in touchscreen devices.

Taking into account the aforementioned potential uses of soft-biometrics predictions from touch gesture data within both biometrics and HCI fields, the work presented in this paper analyses the possibility of predicting user’s sex from swipe gesture data captured via a smart phone. In particular we assess how differences in swipe direction affect prediction performance, indicating relevant swipe gesture features for user’s sex prediction and which classification algorithm best suits this data. Moreover, sex prediction fusion schemes based on individual swipe direction are analysed.

2. Literature review

Swipe gesture-related techniques have been proposed in several studies as an authentication method in order to enhance the access of touch-screen devices. In one of the first studies in this field, Jermyn et al [10] proposed the use of the graphical password to replace the text-based password. This work was motivated by the graphical input capabilities of the first personal digital assistant devices (PDAs).

More recently, in [11] the authors combine the idea of a use of Android lock pattern authentication with the dynamics of drawing such a pattern. The Android lock pattern authentication allows users to unlock their devices by drawing with the finger a pattern on a 3×3 grid of 9 circles displayed on the device touch-screen. The authors proved that security can be significantly enhanced by using the dynamic information of the finger movement while drawing this pattern. The use of touch graphical passwords for tablets with multi-touch-screens has been also analysed in [12], where the dynamic information of multi-touch input sequence were analysed for user authentication. In this work, the authors report an equal error rate of 10% using a single multi-touch gesture. This work has also been used to obtain a US Patent [14].

In [13] the authors explore the possibilities of swipe gesture data (combined with accelerometer data) to perform continuous user identification for mobile devices. In this work, the interaction with the mobile device was analysed according to three gestures: Tap, Scroll and Fling. The touch-based features include touch coordinates on the screen, touch pressure and duration across three different apps (Message, Album and Twitter) with 100 users interacting with the smartphones. The results show 80% accuracy with 10 interaction for identifying a non-owner using the smartphone, and nearly 100% identification accuracy for the owner of the smartphone within six interactions.

Another example of continuous authentication using swipe gesture can be found at [15], where the authors presented the FAST (Finger-gestures Authentication System using Touchscreen) framework. In this work the six most frequently used swipe gestures were assessed: down-to-up swipe, up-to-down swipe, left-to-right swipe, right-to-left swipe, zoom in, and zoom out. Across multiple

interactions, the proposed system achieved a False Acceptance rate (FAR) of 4.6% and a False Reject rate (FRR) of 0.13% in a dataset composed of 40 users. In addition, using just a single interaction, this work showed different FAR and FRR performances for the six different swipe gestures analysed: down-to-up, up-to-down, left-to-right, right-to-left, zoom in and zoom out.

In 1997, Wayman [16] proposed the use of soft-biometrics in order to optimise search in large surveillance databases. Since then, soft-biometrics have continuously received the attention of the biometrics research community [1]. Studies have shown the benefits of soft-biometrics traits in combination with biometrics modalities to improve both identification/verification algorithm performances and computational search times. Soft-biometrics data can help improve biometrics systems [17] and also can be used to tailored applications and the information displayed based on user’s characteristics. Thus, the prediction of soft-biometrics, specifically a subject’s sex, have been thoroughly analysed in several biometrics.

Face images have been comprehensively analysed for sex prediction [18], reaching accuracy rates of 99% using frontal face images and a support vector machine (SVM) classifier [19]. In [20], the authors analysed sex prediction from unconstrained face images in which there was a high level of variation in viewpoint, pose, articulation and occlusion from the images, simulating a personal photo album. The proposed method achieved 82% accuracy for sex prediction using Poselet-level features and SVM classifiers. Gait-based sex prediction has also obtained a high accuracy rate: 98%, based on video sequences of a person walking from 11 different camera views in the same scene [21].

In a study by Amayeh et al. [22] the authors examined hand features to obtain sex prediction as a soft-biometric. Hands were analysed using image processing techniques and three different machine learning classifiers: minimum distance, k-nearest neighbours and linear discriminant analysis. 98% accuracy rate was obtained by score-level fusion using MPEG-7 Fourier descriptors as the hand features and linear discriminant analysis as the classifier.

Biometric modalities such as iris or fingerprint, not naturally linked with everyday sex classification, have obtained reliable accuracy rates. From iris images, it has been possible to obtain an accuracy rate of 91% analysing iris texture using a local binary pattern and SVM classifier [23]. In [24] a method used on fingerprint images based on discrete wavelets transform and singular value decomposition is proposed. Using a K-nearest neighbour as a classifier, 88% accuracy is reported.

More aligned to the present work are studies undertaken on mouse and keystroke interaction. These are used as the main interaction with many computer-based systems. Idrus et al. [25] conducted a thorough analysis for profiling users from keystroke data. In this work, the authors analysed the possibility of predicting whether the user: (a) used one or two hands to type; (b) belongs to a particular age category; (c) is male or female; and (d) is right- or left-handed. Using four features based on digraphs (two consecutive keystrokes) and SVM classifiers, the authors achieved recognition rates for free text of $> 90\%$ for whether the user typed with one or two hands, between 79% and 84% for user’s sex, 72–75% for age categories and 83–88% for handedness. Furthermore, recent studies have focused on predicting emotional states such as stress [7,8] and happiness [8] obtaining promising results.

To the best of our knowledge, there has not been previous work analysing the possibilities of soft-biometrics prediction from swipe gesture data.

3. Methodology

This section describes the steps undertaken within our experimental methodology from swipe data acquisition to subject



Fig. 1. Software for swipe gesture data acquisition.

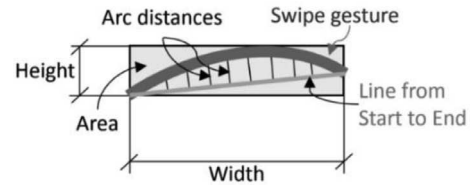


Fig. 2. Swipe feature details.

Table 1
Swipe feature set.

#	Description	#	Description
1	Total length (px)	8	Maxima speed (px/ms)
2	Total time (ms)	9	Average speed (px/ms)
3	Width (px)	10	Maxima acceleration (px/ms ²)
4	Height (px)	11	Average acceleration (px/ms ²)
5	Area (px ²)	12	Average arc distance (px)
6	Average thickness (px)	13	Max arc distance (px)
7	Average pressure	14	Angle start to end (degrees)

sex prediction. It will detail how the data were collected, which features were extracted from the raw data. It will also detail how those features were used in combination with feature selection techniques and machine learning algorithms in order to perform the sex prediction of the mobile users.

3.1. Swipe gesture data collection

The swipe gesture data used in this study were taken from the SSD [26] dataset. The SSD is a comprehensive multi-modal biometrics dataset (face, iris, swipe, keystroke, signature, gait, hand, voice and fingerprint) along with demographics (such as sex, height, weight, handedness) created for the SuperIdentity project [27]. The SSD contains 116 participants, with an even sex distribution (57 males and 59 females) and ages ranging from 18 to 35 years.

The swipe gestures data were captured using a Samsung GT-I9100 ‘Galaxy S2’ smartphone. The GT-I9100 has a 4.3” capacitive touchscreen display (W480 x H800 pixels, 219dpi). Participants were instructed to operate the smartphone one-handed (according to their preference) in portrait orientation, using the thumb of the same hand to interact with the screen (Fig. 1).

The swipe gestures data were captured using a Samsung GT-I9100 ‘Galaxy S2’ smartphone. The GT-I9100 has a 4.3” capacitive touchscreen display (W480 x H800 pixels, 219dpi). No screen protectors or external cases were used. Participants were instructed to operate the smartphone one-handed (according to their preference) in portrait orientation, using only the thumb of the same hand to interact with the screen (Fig. 1). The swipe gestures were captured using an Android OS application that was custom built for this purpose. The application detected and automatically recorded all swipe gestures made in four directions (left-to-right, right-to-left, up-to-down and down-to-up). To elicit swipe gestures, the capture application used a simple reading task. Participants were presented with a series of short jokes (in random order) presented as slides. To complete the task (to read each joke and its punchline), participants were required to perform a swipe gesture in the direction indicated on the screen. Each participant submitted 120 swipe gestures in total, divided evenly across the four directions.

Raw data were recorded about each swipe and were stored locally in a text file. This included x and y positional information, pressure and thickness time series.

3.2. Feature extraction

After the collection of the dataset, a pre-processing step was performed in order to remove swipes which weren’t properly acquired (due to software or user input errors) or were too short

for features extraction (fewer than 4 sample points). From the raw data, instantaneous swipe speed and acceleration were derived as first and second time derivatives. The arc distance between the swipe gesture and an imaginary line joining the start and end point was also calculated. Fig. 2 depicts these and other extracted geometry features such as swipe height, width and area.

From the time series (x and y position, speed, acceleration, pressure, thickness and arc distance) 14 features were extracted, which are detailed in Table 1 along with their units. Screen pixels (px) is used as screen movements unit and milliseconds (ps) for time differences. Android system provides the thickness, as the approximate size of the touch contact area, also in pixels, whilst pressure is provided without any unit scale:

These 14 features were extracted from each swipe, and each feature averaged across all samples from each subject for each swipe direction (left-to-right, right-to-left, up-to-down and down-to-up), resulting in 4×14 features for each subject.

3.3. Feature differences between male and female groups

A Wilcoxon rank-sum test analysis was performed in order to determine whether mean feature values from male and female groups were significantly different. The Wilcoxon rank-sum test was used as it was not possible to assume that feature values are normally distributed. This assumption was made after performing a Lilliefors tests to the distributions.

The Wilcoxon rank-sum test revealed differences between male and female populations in several swipe features, specifically in the down-to-up direction: Width, Area and Angle Start to End; and in the left-to-right: Total Time, Average speed, Average Arc Distance and Max arc Distance. For the two remaining directions, the up-to-down direction only showed significance differences in the Width feature, whereas the right-to-left direction failed to show any significance difference.

These results are comparable to those found in [28] using a similar swipe gesture dataset. In [28], the authors concluded that swipe features are linked to specific physical characteristics of the hand. Users with longer thumbs performed swipe gestures with higher speed and acceleration and, therefore, shorter completion times. This work also highlighted that, on average, male subjects have longer thumbs than females. Based on these differences between male and female populations, this work analyses the use of prediction tools such as feature selection techniques combined with machine learning classifiers in order to find the best combination of features for the sex prediction based on swipe gestures.

3.4. Machine learning approach for sex prediction

The 14 averaged swipe gesture features for a particular swipe direction were used as inputs to machine learning classifiers in order to predict the subject's sex. The WEKA v3.7 suite was used as the machine learning tool [29]. The WEKA machine learning suite has been successfully used in a broad range of scientific fields from speech segmentation [30] to medical domains [31].

Our method employed an initial feature selection step to identify and select the most promising features for sex prediction. After this selection, the predictive power of these feature sets was analysed through 10-fold cross validation using, separately four machine learning classifiers. These steps are explained in the following subsection, indicating the algorithm setting values used to ensure reproducibility.

3.4.1. Feature selection

Feature selection has been carried out in WEKA v 3.7 using the Classifier subset evaluator ("ClassifierSubsetEval v1.0.4") in combination with the BestFirst [32] attribute selection implementation. Both, the classifier subset evaluator and the BestFirst feature selection algorithm (included within WEKA), have been used with their default setting values.

The BestFirst feature selection algorithm searches the attribute space greedily in one of three possible directions: forward, backward or bidirectional. Our experimentation used all three directions independently in order to identify an optimal selection.

As an exploratory analysis, and due to the WEKA implementation of the feature selection algorithms, all the samples from the dataset were used at the feature selection step.

The feature selection step will identify the most promising feature subsets for each classifier and for each feature selection search direction. These feature subsets will be used to create the sex prediction models to evaluate their success ratios using 10 folds cross-validation.

3.4.2. Machine learning classifiers

The classifiers used in this study have been used successfully in different fields as Machine Learning engines [30,31,33]. Four different machine learning classifiers have been tested as possible candidates to predict the sex of the subject. The chosen classifiers cover a range of popular modes of classification: decision trees (J48), probabilistic (naïve Bayes), support vector machines (SVM) and logistic regression, and have been selected for complementarity in assessment.

Decision tree (J48): Decision tree learning is one of the most commonly used algorithms for automatic learning. The decision trees are composed of nodes (which test the value of an attribute), branches (path to follow based on the attribute value) and leaves (which provide the classification of the instance). The decision tree employed in this work is the C4.5 implementation developed by Quinlan [34] implemented in WEKA as the J48 algorithm.

The settings values used for the decision tree classifier (J48) are detailed in Table 2:

Table 2
Decision tree setting values.

Option	Value	Option	Value
BinarySplits	False	SaveInstanceData	False
CollapseTree	True	Seed	1
ConfidenceFactor	0.25	SubtreeRaising	True
MinNumObj	2	Unpruned	False
NumFolds	3	UseLaplace	False
ReduceErrorPruning	False	UseMDLcorrection	True

Table 3
Support vector machine setting values.

Option	Value	Option	Value
SVMType	C-SVC	KernelType	linear u*v
Cachesize	40.0	Loss	0.1
Coef0	0.0	Normalize	True
Cost	1.0	Nu	0.5
Degree	3	ProbabilityEstimates	False
Eps	0.001	Seed	1
Gamma	0.0	Shrinking	True

Table 4
Multilinear logistic regression setting values.

Option	Value	Option	Value
MaxIts	-1	Ridge	1.0E-8

Table 5
Naïve Bayes setting values.

Option	Value	Option	Value
UseKernelEstimator	False	useSupervisedDiscretization	False

Support vector machine (SVM): Support vector machines were introduced by Cortes and Vapnik [35] in 1995 and have been successfully used in a wide range of different areas. SVM algorithms are based on finding the optimal separating hyperplane that maximizes the margin, in other words, the hyperplane that gives the largest minimum distance to the training examples. The implementation used in this investigation is the LibSVM v1.0.6 [36] as an add-on to the WEKA system. The options selected for SVM classifier were, (Table 3):

Multilinear logistic regression: Multilinear logistic regression [37] is one of the most commonly used tools for discrete data analysis. Multinomial logistic regression is used to predict the probabilities of the different classes analysed given a set of independent variables. It represents a particular solution to the classification problem that assumes that a linear combination of the observed features can be used to determine the probability of each particular outcome of the dependent variable.

Table 4 details the setting values used for the multilinear logistic regression classifier:

Naïve Bayes: The naïve Bayes classifier [38] is based on the probabilistic Bayes' rule and is particularly suited when the dimensionality of the inputs is high. In order to reduce the complexity of the high dimensionality, the naïve Bayes classifier assumes that the effect of the value of a particular feature on a given class is independent of the values of the other predictors. Despite its oversimplified and generally unrealistic assumptions, the naïve Bayes classifier has been shown to perform remarkably well in a wide range of applications such as text classification [38] and internet traffic identification [39].

This classifier can be found implemented in WEKA and was analysed with the following settings, Table 5:

3.4.3. Sex prediction evaluation

The machine learning models (following the feature selection step) were created for the four different classifiers and the four swipe directions. The evaluation of the models was undertaken by means of 10-fold cross validation. This method is a model validation technique to estimate the performance of a statistical prediction model. The technique randomly partitions the original sample data into 10 equal sized subsamples. Nine of the subsamples are used for training the model and the remaining subsamples is used

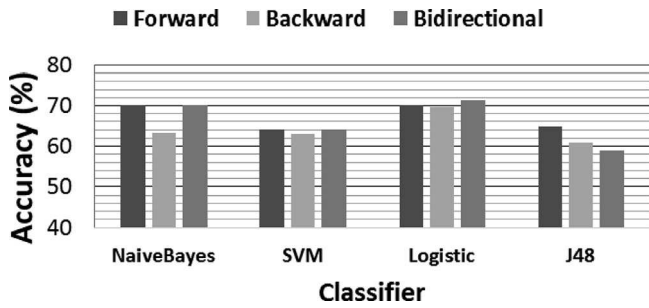


Fig. 3. Down-to-up sex prediction score distribution.

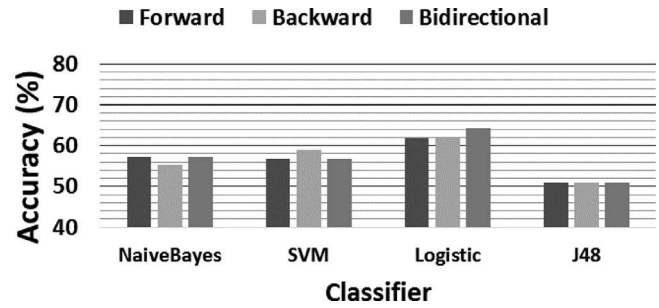


Fig. 6. Up-to-down sex prediction score distribution.

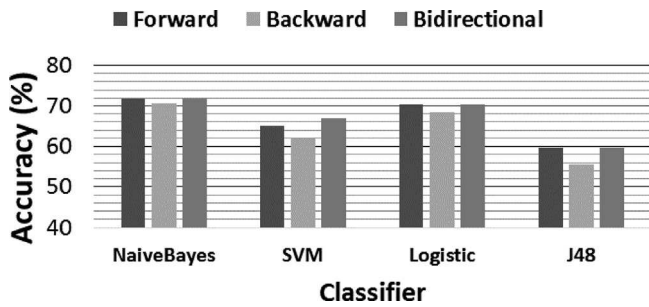


Fig. 4. Left-to-right sex prediction score distribution.

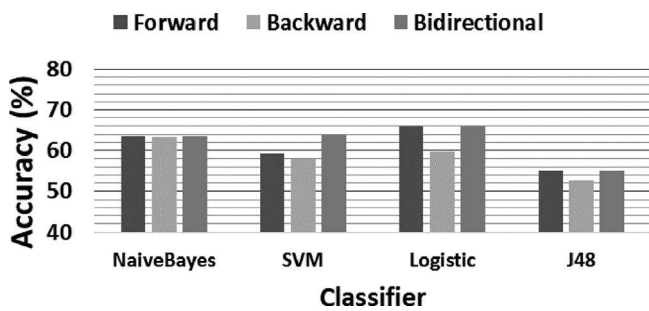


Fig. 5. Right-to-left sex prediction score distribution.

as the validation data for testing the model. To reduce variability, this process is repeated 10 times using different subsamples data for validation.

The 10-fold cross validation was carried out 25 times using different random seed numbers (values from 1 to 25, which ensure that fold observations are different on each evaluation) in order to obtain a statistically significant average of the model performance, with confidence intervals ($\alpha=0.05$) for the average accuracy around 1%.

4. Results

The results obtained from the evaluation of the machine learning models created are presented in this section. In the first subsection the sex prediction performance for each classifier and each swipe direction is disclosed. Following this, the score distributions for the best classifier of each direction are presented as an introduction to the last subsection, which attempts to enhance sex prediction performance by means of score and decision fusion techniques.

4.1. Individual directions sex prediction performance

Figs. 3–6, one figure for each swipe direction, show the performance of the four classifiers analysed using the feature set

selected by the BestFirst algorithm for the three different search directions.

As it can be seen in this figures, there are clear differences in performance between the four swipe gesture directions. Down-to-up and left-to-right direction both show the highest accuracy rates of around 71%, being significantly higher than the accuracies from right-to-left and up-to-down, which average around 65%. Down-to-up and left-to-right swipe gestures involve the extension of the thumb whilst up-to-down and right-to-left involve flexion. Subjects might be more stable in extension movements and, therefore, the classification of this movements obtain a higher classification performance. These results are consistent with the Wilcoxon rank-sum test results. More significant differences were found for these two directions (down-to-up and left-to-right), which could be explained by a greater difference in swipe gesture performance between males and females groups for these two directions.

Regarding the most suitable classifier algorithms for sex prediction based on swipe gesture data, the multilinear logistic regression classifier shows good accuracy rates across all four swipe directions. Only the naive Bayes algorithm slightly bettered the logistic classifier for left-to-right swipe direction. Table 6 summarise the best accuracy rates (and their average accuracy confidence interval, CI) for each swipe gesture direction, giving further details such the feature set used and the feature selection search direction.

The best classifier (71.8% accuracy), for the left-to-right swipe gesture features, is based on *average thickness* (feature ID 6), *maxima speed* (feature ID 8) and the *average arc distance* (feature ID 12). For this swipe direction, women presented a higher average arc distance (6.57px for females compared with 5.21px for males), slightly lower thickness (40.6px for females compared with 41.6px for males) and maxima speed (2.54px/ms for females compared with 2.77px/ms for males). *Average thickness* swipe gesture feature was also selected within the feature set for the four swipe directions. It is also worth highlighting that the *average pressure* (feature ID 7) was selected within three of the directions). These two features, thickness and pressure, are closely related. Thickness is the approximate size of the touch contact area. The pressure is not directly measured from the Android smartphone used in our experimentation as it is estimated from the size of the touch point on the assumption that more pressure means your finger flattens out. *Height* (feature ID 4), *area* (feature ID 5) and *maxima acceleration* (feature ID 10) were included in both down-to-up and up-to-down swipe gesture directions models whilst *Total time* (feature ID 2), *Maxima speed* (feature ID 8) and *average acceleration* (feature ID 11) were included in two swipe direction models.

In order to analyse the influence of each feature on the prediction models, Table 7 details the coefficient values for the multilinear logistic regression models (naive Bayes model has not been included due to lack of detailed model information at WEKA machine learning suite):

Table 6
Best accuracy rates for each swipe gesture direction.

Swipe direction	Classifier	Feature selection direction	Feature IDs	Accuracy and CI
Down-to-up	Logistic	Bi-dir.	2,4,5,6,7,9,10,11,14	71.3 ± 0.8 %
Left-to-right	Naïve Bayes	Forward/Bi-dir	6,8,12	71.8 ± 0.7 %
Right-to-left	Logistic	Forward/Bi-dir	2,6,7,11	66.0 ± 0.8 %
Up-to-down	Logistic	Bi-dir.	1,4,5,6,7,8,10,13	64.3 ± 1.1 %

Table 7
Multilinear logistic regression coefficients for the sex prediction models.

Feature ID	DU	RL	UD
1 Total length	-	-	46.13
2 Total time	10.8	-1.29	-
3 Width	-	-	-
4 Height	-11.55	-	-48.39
5 Area	2.64	-	-0.18
6 Average thickness	15.09	20.27	12.13
7 Average pressure	-14.39	-20.15	-13.20
8 Maxima speed	-	-	3.03
9 Average speed	11.09	-	-
10 Maxima acceleration	4.04	-	-3.55
11 Average acceleration	-9.10	-1.68	0.90
12 Average arc distance	-	-	-
13 Max arc distance	-	-	-
14 Angle start to end	-1.0	-	-

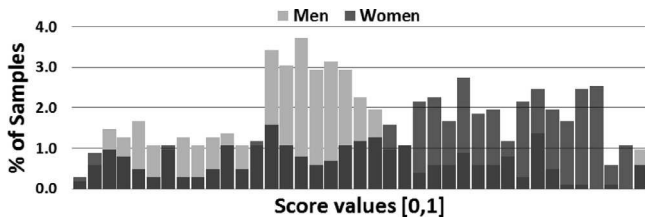


Fig. 7. Down-to-up sex prediction score distribution.

To calculate this coefficients, all the features have been normalised (mean 0 and standard deviation 0) to be able to compare the coefficients between each other. At Table 7, the higher the value of the coefficients, the higher the “log odds” increment of being female. It is worth to highlight the high weight values of “Total length” and “Height” for the up-to-down swipe sex prediction model, being remarkable higher than the rest of coefficients. It is also interesting the positive and negative sign of the “Average Thickness” and “Average Pressure” weights for these three swipe directions. As mentioned before, the pressure values are estimated by the android operating system from the thickness values. The different signs and the similar weight values will mitigate the overall importance of these two features. An increase of the “Average Thickness” values will imply an increase of the “log odds” of being female, however, it will also imply an increase of the “Average Pressure” which will lead to a decrease of the “log odds” to be female due to the negative sign of its weight. Due to the contradictory sign of these two values, the overall importance of the thickness and pressure can be consider low.

4.2. Direction score distributions

In Figs. 7–10, the sex prediction score distributions on each swipe gesture direction using the best classifier identified in Table 6 are depicted. The sample scores obtained from each swipe direction classifier represent the probability of the input sample to belong to a female user, in a scale from 0 to 1. If the score is close to 0, it means that the probability of the input sample to belong

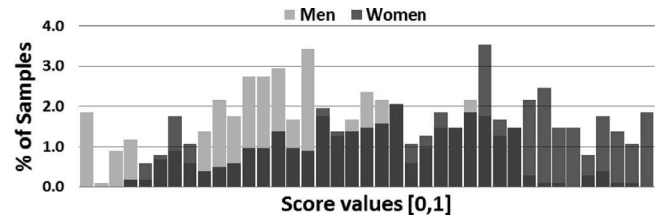


Fig. 8. Left-to-right sex prediction score distribution.

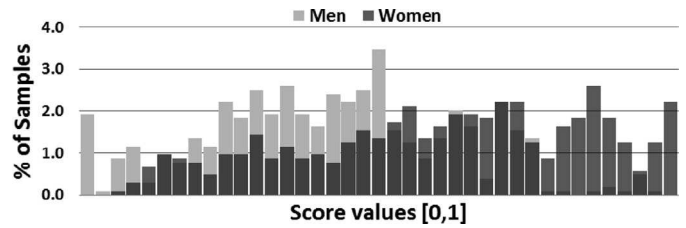


Fig. 9. Right-to-left sex prediction score distribution.

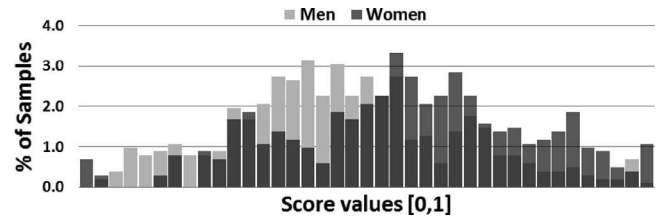


Fig. 10. Up-to-down sex prediction score distribution.

to a male user is very high. On the other hand, if the score is close to 1, it will likely belong to a female user. These figures show male score distribution in light grey, female score distribution in dark grey, and how they overlap in mid-grey. It can be seen how the distribution from each direction follows different patterns. Specifically, it can be observed how down-to-up score graph has a lower overlap between men and women distribution. This overlap was the reason that this direction obtained the highest accuracy performance (see Table 6). These differences will be used to analyse several fusion techniques in order to improve the accuracy rates obtained from individual swipe gesture direction.

4.3. Fusion scheme sex predictions results

Fusion techniques are commonly used in multi-biometrics systems [40]. These systems have been proven to obtain enhanced performance over individual modality biometrics system. Fusion can be performed at different levels:

- (i) **Feature level:** where the features from different biometrics modalities are combined. The combined feature set will be used as an input to the matching algorithm.
- (ii) **Matching score level:** where the matching score levels from each biometrics modality are combined to obtain a single score level.

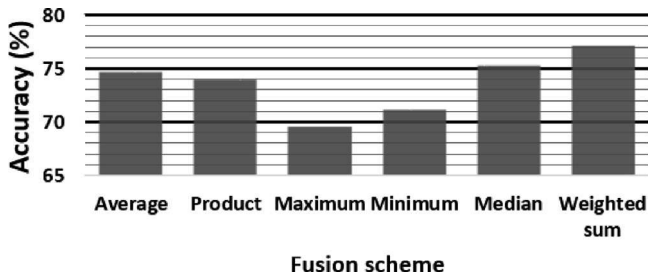


Fig. 11. Up-to-down sex prediction score distribution.

- (iii) **Decision level:** where the binary decision of the classification algorithms are combined to reach a single decision.

In this work, matching score level and decision level techniques will be analysed. These fusion techniques have been selected due to their popularity and success in previous studies [40], along with their simplicity in implementation.

For matching score level fusion, the following matching score level rules have been analysed and applied to the prediction of sex based on swipe gesture scores:

- (i) Weighted sum of scores: $S_{fusion} = \sum_i w_i \cdot s_i$
- (ii) Product of scores: $S_{fusion} = \prod_i s_i$
- (iii) Maximum score: $S_{fusion} = \max\{s_i\}$
- (iv) Minimum score: $S_{fusion} = \min\{s_i\}$
- (v) Median score: $S_{fusion} = \text{median}\{s_i\}$

where w_i is the weight of the i th swipe direction with $\sum_i w_i = 1$ and s_i is the score obtained from the i th swipe direction classifier.

The decision level approach was implemented by using the individual classifier swipe gesture direction decisions. Three different voting thresholds have been analysed: (i) at least one classifier across the four swipe directions classified the subject as female, (ii) at least two and (iii) at least three. Otherwise, if the number of classifiers labelling the subject as a female is lower than the threshold, the sample is considered from a male subject.

Fig. 11 shows the accuracy obtained when the score level fusion techniques were applied to the swipe gesture prediction scores.

From all the possible weights combinations, the higher accuracy was obtained by the weighted sum of scores fusion technique, which achieves 77.1% of accuracy rate. The best combination of weights found was:

$$S_{fusion} = 0.3 \cdot S_{UD} + 0.7 \cdot S_{LR}$$

It can be seen how the two best swipe directions in terms of accuracy: left-to-right (71.8% accuracy) and up-to-down (70%) account for the most of the score fusion, with a higher weight for the best swipe direction.

This combination of swipe direction scores means an improvement of 5% compared with the best individual swipe gesture direction sex prediction rate, 71.8%.

The accuracy rates obtained when using the decision level fusion techniques is presented in Fig. 12.

This approach obtain a 78.2% of accuracy when two or more swipe gesture direction agree on the sex of the users. This rate means an improvement of 6% from the best individual accuracy rate.

5. Conclusion and future work

The increasing adoption of touch-screen devices and their continuous data capture enrichment will bring the possibility of collecting high quality swipe gesture data from users interactions and the opportunity of using these data to predict soft-biometrics such

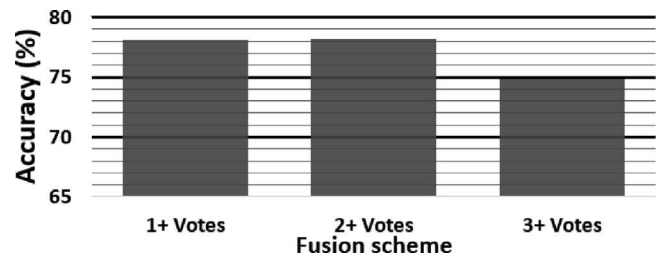


Fig. 12. Up-to-down sex prediction score distribution.

sex, age category, single or-two handed usage, handedness or even emotion prediction.

This soft-biometric information can be used to improve authentication systems (i.e. continuous authentication based on mobile devices use), to tailor applications interfaces to specific user groups, or to enhance the interaction between computer-based systems and users.

Following these ideas, this paper has analysed the possibility of sex prediction using swipe gesture data collected from the user interaction with a touch-screen device. This interaction involves the placement of a finger on the screen following with a fast movement in one specific directions. These gestures are frequently used with touch-screen devices while navigating websites, list menus or picture galleries.

The results of this exploratory analysis have confirmed the possibility of sex prediction from the swipe gesture data, obtaining an encouraging 71% accuracy from an individual swipe gesture direction. Furthermore, the results have shown a significant difference in the sex prediction power based on swipe directions. The swipe directions involving finger extensions (down-to-up and left-to-right) obtained around 71% accuracy, while the swipe directions involving finger flexion (up-to-down and right-to-left) obtained around 65% accuracy.

Regarding the most suitable machine learning classifier for this task, the multilinear logistic regression has shown a good performance across all swipe gesture direction, only slightly bettered by the naïve Bayes classifier for left-to-right swipe direction.

We have also analysed sex prediction accuracy when combining the data from the four directions. Several score level and decision level fusion techniques have been implemented. The results of this analysis showed that the combination of direction swipe data enhanced the sex prediction accuracy by 6%, achieving a 78% accuracy rate using a decision voting scheme.

It is important to acknowledge the limitation of this research to drawn general conclusions from small sample populations. Yet, it is valuable to acknowledge this kind of information can be extracted from swipe gesture data pointing to further research in this area. This will help the research community to find ways to use this information to enhance the user interaction with technology and also to improve continuous authentication on mobile devices. Moreover, the acknowledgement of the potential prediction of this kind of information can be essential to prevent this leak of information when it could imply a privacy risk.

These results will hopefully encourage further research on this area. Predicting soft-biometrics information from swipe gesture is a new field in biometrics and Human-Computer interaction field and, therefore, there is scope for improvements in all data analysis steps. Specifically for the work presented in this paper, the analysis of sex prediction based on swipe gesture data, the following areas should be investigated:

- Feature extraction: the swipe gesture is a multi-dimensional sequence of numerical data points. This characteristic allows the creation of new swipe features that can be analysed to im-

prove sex prediction accuracy rates. New swipe features could be found through a deeper analysis of the biological mechanism between swipe gestures, swipe direction and user's sex.

- Feature selection: more advanced techniques could be investigated for the selection of the best combination of swipe features.
- Classifiers: specific parameterisation of classifiers could also bring improved accuracies. Furthermore, the use of ensembles of classifiers could be another option for the improvement of the performance.
- Swipe gesture directions: thorough analysis of the difference between swipe features from different swipe gesture directions and how they impact on the sex prediction performance can enable better fusion strategies of swipe features from different directions.

Moreover, the prediction of other soft-biometrics traits such as age categories, handedness, stress detection and emotion recognition are other areas where swipe gesture data could lead to results with a practical applications.

Acknowledgments

This work was supported by UK EPSRC Grant [EP/J004995/1](#) (SID: An Exploration of SuperIdentity). The authors would like to thank colleagues on this grant for helpful comments on this work.

References

- [1] K. Nandakumar, A.K. Jain, Soft biometrics, in: S. Li, A. Jain (Eds.), *Encyclopedia of Biometrics*, Springer US, Boston, MA, 2009, pp. 1235–1239.
- [2] H. Ailisto, et al., Soft biometrics—combining body weight and fat measurements with fingerprint biometrics, *Pattern Recognition Lett.* 27 (5) (2006) 325–334.
- [3] R. Zewail, et al., Soft and hard biometrics fusion for improved identity verification, in: *Proceedings of the 2004 47th Midwest Symposium on Circuits and Systems*, 1, IEEE, 2004.
- [4] U. Park, A.K. Jain, Face matching and retrieval using soft biometrics, *IEEE Trans. Inf. Forensics Secur.* 5 (3) (2010) 406–415.
- [5] K. Niinuma, U. Park, A.K. Jain, Soft biometric traits for continuous user authentication, *IEEE Trans. Inf. Forensics Secur.* 5 (4) (2010) 771–780.
- [6] A. Kolakowska, A review of emotion recognition methods based on keystroke dynamics and mouse movements, in: *Proceedings of the 6th International Conference on Human System Interaction (HSI)*, IEEE, 2013.
- [7] J. Hernandez, et al., Under pressure: sensing stress of computer users, in: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2014.
- [8] M. Fairhurst, L. Cheng, E. Meryem, Exploiting biometric measurements for prediction of emotional state: A preliminary study for healthcare applications using keystroke analysis, in: *Biometric Measurements and Systems for Security and Medical Applications Proceedings, Workshop on*. IEEE, 2014.
- [9] E.A. Johnson, Touch display—a novel input/output device for computers, *Electron. Lett.* 1 (8) (1965) 219–220.
- [10] W. Hu, X. Wu, G. Wei, The security analysis of graphical passwords, in: *Proceedings of Communications and Intelligence Information Security (ICCIS)*, International Conference on, IEEE, 2010.
- [11] J. Angulo, E. Wästlund, Exploring touch-screen biometrics for user identification on smart phones, *Privacy and Identity Management for Life*, Springer, Berlin Heidelberg, 2011, pp. 130–143.
- [12] N. Sae-Bae, N. Memon, K. Isbister, Investigating multi-touch gestures as a novel biometric modality, in: *Proceedings of IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems*, 2012.
- [13] C. Bo, et al., Silentsense: silent user identification via touch and movement behavioral biometrics, in: *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking*, ACM, 2013.
- [14] K. Isbister, N. Memon, S.B. Napa, Biometric-rich gestures for authentication on multi-touch devices, U.S. Patent No. 9,147,059 (29 Sep. 2015).
- [15] T. Feng, et al., Continuous mobile authentication using touchscreen gestures, in: *Proceedings of 2012 IEEE Conference on Technologies for Homeland Security (HST)*, 2012.
- [16] J.L. Wayman, Large-scale civilian biometric systems—issues and feasibility, in: *Proceedings of Card Tech/Secur Tech ID*, 732, 1997.
- [17] A.K. Jain, S.C. Dass, K. Nandakumar, *Soft biometric traits for personal recognition systems*, Biometric Authentication, Springer, Berlin Heidelberg, 2004, pp. 731–738.
- [18] C.B. Ng, Y.H. Tay, B.-M. Goi, Recognizing human gender in computer vision: a survey, *PRICAI 2012: Trends in Artificial Intelligence*, Springer, Berlin Heidelberg, 2012, pp. 335–346.
- [19] J. Zheng, B.L. Lu, A support vector machine classifier with automatic confidence and its application to gender classification, *Neurocomputing* 74 (11) (2011) 1926–1935.
- [20] L. Bourdev, S. Maji, J. Malik, Describing people: A poselet-based approach to attribute classification, in: *Proceedings of International Conference on Computer Vision (ICCV)*, 2011.
- [21] F.B. Osuikie, K. Faez, Gender classification using a novel gait template: radon transform of mean gait energy image, *Image Analysis and Recognition*, Springer, Berlin Heidelberg, 2011, pp. 161–169.
- [22] G. Amayeh, G. Bebis, M. Nicolescu, Gender classification from hand shape, in: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2008.
- [23] J.E. Tapia, C.A. Perez, K.W. Bowyer, Gender Classification from Iris Images Using Fusion of Uniform Local Binary Patterns, in: *Proceedings of Computer Vision-ECCV 2014 Workshops*, Springer International Publishing, 2014.
- [24] P. Gnanasivam, S. Muttan, Fingerprint gender classification using wavelet transform and singular value decomposition, *Int. J. Comput. Sci. Iss.* 9 (3) (2012). <http://arxiv.org/abs/1205.6745>.
- [25] S. Idrus, S. Zulkarnain, et al., Soft biometrics for keystroke dynamics: Profiling individuals while typing passwords, *Comput. Secur.* 45 (2014) 147–155.
- [26] R. Guest, et al., Biometrics within the SuperIdentity project: a new approach to spanning multiple identity domains, in: *Proceedings of 2014 International Carnahan Conference on Security Technology (ICCST)*, IEEE, 2014.
- [27] S.M. Black, et al., "Superidentity: Fusion of identity across real and cyber domains." in: *ID360: Global Identity* (2012).
- [28] C. Bevan, D.S. Fraser, Different strokes for different folks? Revealing the physical characteristics of smartphone users from their swipe gestures, *Int. J. Human-Comput. Stud.* 88 (2016) 51–61.
- [29] M. Hall, et al., The WEKA data mining software: an update, *ACM SIGKDD Explor. Newsl.* 11 (1) (2009) 10–18.
- [30] B. Ziólko, et al., Logitboost weka classifier speech segmentation, in: *Proceedings of 2008 IEEE International Conference on Multimedia and Expo, IEEE*, 2008.
- [31] B.M. Patil, D. Toshniwal, R.C. Joshi, Predicting burn patient survivability using decision tree in weka environment, in: *Proceedings of IEEE International Advance Computing Conference*, 2009.
- [32] M. Dash, L. Huan, Feature selection for classification, *Intell. Data Anal.* 1 (3) (1997) 131–156.
- [33] T. Garg, S.S. Khurana, Comparison of classification techniques for intrusion detection dataset using WEKA, in: *Recent Advances and Innovations in Engineering*, IEEE, 2014.
- [34] J.R. Quinlan, *C4.5: programs for machine learning*, Elsevier, 2014.
- [35] C. Cortes, V. Vladimir, Support-vector networks, *Mach. Learn.* 20 (3) (1995) 273–297.
- [36] C.-C. Chang, C.J. Lin, LIBSVM: a library for support vector machines, *ACM Trans. Intell. Syst. Technol.* 2 (3) (2011) 27.
- [37] D.W. Hosmer Jr., S. Lemeshow, *Applied Logistic Regression*, John Wiley & Sons, 2004.
- [38] D.D. Lewis, Naive (Bayes) at forty: the independence assumption in information retrieval, *Machine Learning: ECML-98*, Springer, Berlin Heidelberg, 1998, pp. 4–15.
- [39] J. Erman, A. Mahanti, M. Arlitt, Qrp05-4: Internet traffic identification using machine learning, in: *Proceedings of Global Telecommunications Conference, IEEE*, 2006.
- [40] A.A. Ross, K. Nandakumar, A. Jain, *Handbook of Multibiometrics*, 6, Springer Science & Business Media, 2006.