2022-02

Real-World Smartphone-based Gait Recognition

Alobaidi, H

http://hdl.handle.net/10026.1/18438

Journal Pre-proof

Real-World Smartphone-based Gait Recognition

Hind Alobaidi , Nathan Clarke , Fudong Li , AbdulrahmanAlruban

Please cite this article as: Hind Alobaidi , Nathan Clarke , Fudong Li , AbdulrahmanAlruban , Real-World Smartphone-based Gait Recognition, *Computers & Security* (2021), doi: https://doi.org/10.1016/j.cose.2021.102557

Real-World Smartphone-based Gait Recognition

Hind Alobaidi[1], Nathan Clarke[1, 2], Fudong Li[3,] Abdulrahman Alruban[4]

[1]Centre for Security, Communications and Network Research, University of Plymouth, Plymouth, UK
[2]Security Research Institute, Edith Cowan University, Perth, Western Australia
[3]School of Computing, University of Portsmouth, Portsmouth, UK
[4]Department of Information Technology, Computer Sciences and Information Technology College, Majmaah University, Al-Majmaah, 11952, Saudi Arabia
{hind.al-Obaid, n.clarke}@plymouth.ac.uk; fudong.li@port.ac.uk; a.alruban@mu.edu.sa

Abstract

As the smartphone and the services it provides are becoming targets of cybercrime, it is critical to secure smartphones. However, it is important security controls are designed to provide continuous and user-friendly security. Amongst the most important of these is user authentication, where users have experienced a significant rise in the need to authenticate to the device and individually to the numerous apps that it contains. Gait authentication has gained attention as a mean of non-intrusive or transparent authentication on mobile devices, capturing the information required to verify the authenticity of the user whilst the person is walking. Whilst prior research in this field has shown promise with good levels of recognition performance, the results are constrained by the gait datasets utilised being based upon highly controlled laboratory-based experiments which lack the variability of real-life environments. This paper introduces an advanced real-world smartphone-based gait recognition system that recognises the subject within real-world unconstrained environments. The proposed model is applied to the uncontrolled gait dataset, which consists of 44 users over a 7-10 day capture – where users were merely asked to go about their daily activities. No conditions, controls or expectations of particular activities were placed upon the participants. The experiment has modelled four types of motion normal walking, fast walking and down and upstairs for each of the users. The evaluation of the proposed model has achieved an equal error rate of 11.38%, 11.32%, 24.52%, 27.33% and 15.08% for the normal, fast, down and upstairs and all activities respectively. The results illustrate, within an appropriate framework, that gait recognition is a viable technique for real-world use.

Keywords: Smartphone authentication, Transparent authentication, Continuous authentication, Gait recognition, Biometrics.

1 Introduction

During the last decade, smartphones have become a ubiquitous technology, with more than 6.3 billion users currently around the world (Statista, 2021). Currently, smartphones provide a wide range of services and features (e.g. personal communications, entertainment, and business) and are used to access and store sensitive and confidential information such as financial data and health-based records. Indeed, it is highly likely that the stored data is far more valuable than the device itself (Saevanee et al., 2015). As a result, smartphones should be kept secure against any illegitimate access. Current authentication approaches (e.g. password or fingerprint) that are deployed upon smartphones are typically intrusive, insecure, and fail to take into account user satisfaction and convenience (Clarke, 2011). Therefore, transparent and continuous biometric authentication systems have been proposed to provide more convenient, secure protections for mobile devices (Muaaz, 2013).

Transparent and continuous authentication schemes (TAS), also referred to as Active Authentication or Implicit Authentication seek to verify the identity of a user through capturing biometric-based information during a user's normal interaction with a mobile device (Clarke, 2011). For example, capturing a facial image for use in facial recognition whilst the user is reading a web page. Rather than providing authentication at the point of entry (or request) these systems keep a constant or continuous confidence in the identity of a user and use this confidence when a user requests access to a service or information. Should they have sufficient confidence, a user can be granted access without having to explicitly provide a sample, and thereby reducing user inconvenience. In practice the TASs could be the gatekeeper to a password manager or keychain that subsequently releases the username/password pair to a service. Prior research has also explored the use of TAS across devices in an approach referred to as an Authentication Aura (Hocking et al., 2013). However, to enable this to take place seamlessly, it is important that the TAS can draw upon a range of biometric modalities, as users will be undertaking a range of activities where one or other biometric modalities would not be available. In contrary to traditional biometrics, where the conditions for capture can be highly controlled (e.g. facial recognition at border control, where light, height and distance from camera can be managed), transparent biometric systems require modalities that can adapt to the varying external environments they find themselves in. As such, the current focus of research is upon the identification and development of biometric techniques that would operate in this manner.

One of those is Gait Recognition, in which people are verified by the way they walk. Many studies in the fields of psychology, medicine, and biometrics suggest that every person's gait is unique (Derawi, 2012; Sprager and Juric, 2015) and it can be deployed as a transparent technique for user identification and verification purposes (Gafurov, 2008). Currently, the majority of smartphones have built-in sensors (e.g. accelerometer and gyroscope) that can be used to record the user's gait information (e.g. non-gravitational accelerations and rotational paces) (Rana, 2015) By using gait recognition for the purpose of user authentication, the user does not need an explicit action to authenticate as related data is continuously recorded while the person walks. Therefore, gait-based authentication can be a valuable approach, amongst other modalities (Al Abdulwahid, 2017), for providing a multimodal transparent and continuous protection for smartphones devices (Clarke, 2011). It is, however, the responsibility of the wider TAS to *intelligently* determine what modalities to use and when.

The feasibility of using a mobile device for gait recognition has been explored by a number of researchers (Sprager and Juric, 2015). However, all the previous studies were applied within a highly controlled environment (i.e. a fixed set of activities for each participate to undertake, such as walking on a flat floor at their usual pace)(Phan et al., 2015; Watanabe, 2015). While this approach is arguably suitable when initially evaluating whether a biometric modality has merit (i.e. whether sufficient discriminate information exists), it does not reflect the type of use one might expect in practice where a large number of variables can impact the reliability. Real-world means completely unconstrained, with no experimental controls on the nature of the activity or where and when it would take place.

This paper aims to investigate the extent to which gait recognition is feasible within uncontrolled, real-world use. The data was collected over 7-10 days using real-world smartphones and real-world use. A series of investigations were undertaken to assess the suitability and efficiency of utilising such gait data. As it was anticipated that real-world data might well be nosier/more variable than controlled data, the study sought to evaluate whether a multi-algorithmic based approach, based upon differing activities, would perform better than a single classification approach. Moving beyond classification, the study also sought to

evaluate the role a majority-voting based decision process could have upon overall system performance.

The rest of the paper is organised as follows: Section 2 highlights the background and related work in the area of using smartphone motion sensors for authentication. Section 3 explains the data collection settings and experimental methodology. Section 4 presents the experimental results of the different tests undertaken to evaluate the proposed approach. Section 5 discusses the findings and the implications for future research directions. The paper concludes in Section 6.

2   Background and Related Work

Gait recognition is a wide research field, and initially focussed upon the use of video observation of a person's gait. Such approaches are useful if the identify of a user is verified from a distance. More recent research has explored the use of personal devices, both wearable sensors and smartphones, to enable or facilitate local user authentication. In comparison with wearable sensors, a key advantage of Smartphones is that the sensors required to capture motion are embedded at no additional cost within the device itself. Indeed, more recent sensor additions such as the gyroscope provide the potential for additional motion-based data (Capela et al., 2016; Shoaib et al., 2016).

Methodologically, studies that utilised smartphone-embedded sensors, the device is either placed in a pouch or inside the trouser pocket (Antos et al., 2014; Ganti et al., 2010). Typically, gait data is collected at a rate ranging from 20 to 50 samples per second, with an average of around 35 samples per second. The recording of user's gait information is typically gathered either on the same day (SD scenario) or across two different days (CD scenario). The dataset is an essential part of the authentication process, and an algorithm could give different results depending upon the tested data (Gadaleta and Rossi, 2010). One of the most extensive gait datasets, which are publicly accessible is from Osaka University (Ngo et al., 2014). It is based on three internal sensors placed on the subjects' belt, with a triaxle accelerometer and gyroscope. However, a smartphone was worn in the centre back waist, and only measured triaxle accelerometer data. Although the collected data consists of 744 subjects, it was collected in a controlled environment, and for each participant, there are only two data sequences available (each session lasting about 1 minute). With such limited signals for each individual, it is challenging to validate the approach extensively. Moreover, the gyroscope data was not included. As illustrated in Table 1 a number of other datasets are available, but for a much smaller number of participants.

The data generated by embedded motion sensors are raw signals which are typically pre-processed to enable feature extraction. The two principal approaches for the feature extraction process are cycle and segment-based techniques. In the cycle-based approach, the captured activity data are supposed to be a periodic signal in which each cycle begins once a foot touches the ground and finishes when the same foot touches the ground for the second time (i.e. two steps for a human) (Derawi and Bours, 2013). In the segment-based method, signals are divided into fixed time-length windows (e.g., 10 seconds) in which each segmented window is processed independently. Some gait activities signals are periodic, as each time segment is reasonably assumed to contain similar signal features. While other activity streams, such as standing and sitting, do not necessarily generate cycle-like patterns. The segmenting of the signal based on a time sequence requires less computational overhead than the cycle-based method. Numerous features from the time (TD) and frequency (FD) domains can be extracted from processing the raw data.

Table 1 presents an analysis of prior research in smartphone-based gait authentication. It can be seen that those studies that were carried out under the SD scenario achieved better performance than those that were under the CD scenario. This is understandable as for the SD scenario both the enrolment and testing data were collected on the same day/session, and the change or variability in user's activities pattern would arguably be smaller than those collected on different days.

Table 1: Prior studies on gait authentication systems using mobile sensors.

Study	Device	Approach	Feature Domain	Classification methods	Users	Performance %	Data duration
Sprager and Zazula, 2009	Nokia N95	C	TD	SVM	6	CCR 93.30	CD
Derawi et al., 2010	Google G1	C	TD	DTW	51	EER 20	CD
Frank et al., 2010a	HTC G1	S	TD	SVM	6	CCR 85.48	CD
Frank et al., 2010b	-	S	TD	Nearest neighbours	40	CCR 100	SD
Kwapisz et al., 2010	Nexus One, HTC Hero, Motorola	S	TD	J48 decision trees &Neural network	36	CCR 90	SD
Kwapisz et al., 2011	Nexus One, HTC Hero, Motorola	S	TD	J48 decision trees &Neural network	5	CCR 100	SD
Nickel et al., 2011a	Google G1	S	FD	SVM	48	EER 6.1	CD
Nickel et al., 2011b	Motorola milestone	S	TD&FD	SVM, HMM	36	EER 10 & 2.36	CD
Nickel et al., 2011d	Motorola	C	TD	Manhattan & DTW	48	EER 21.7	CD
Nickel et al., 2011e	Google G1	S	TD	HMM	48	EER 6.15	CD
Hestbek et al., 2012	Motorola Milestone	S	TD&FD	SVM	36	EER 10.1	CD
Nickel et al., 2012	Motorola Milestone	S	TD&FD	K-NN	36	EER 8.24	CD
Hoang et al., 2013	HTC Nexus One	S	TD&FD	SVM	38	EER 1.95	SD
Muaaz and Nickel 2012	WS &Google G1	C	TD	DTW	48	EER 29.39	CD
Hoang et al., 2013	Google Nexus	C	TD	SVM	32	CCR 100	SD
Derawi and Bours, 2013	Samsung Nexus	C	TD	Euclidean distance and DTW	5	CCR 89.3	SD
Muaaz and Mayrhofer 2013	Google G1	C	TD	DTW	51	EER 33.3	CD
Hoang et al., 2013	HTC Nexus One & LG Optimus G	C	TD&FD	SVM& RBF	14	CCR 91.33	SD
Nickel and Busch 2013)	Google G1	S	FD	HMM	48	EER 6.15	CD
Watanabe, 2014	IOS	S	TD	Neural Network	5	EER 1.82	SD
(Hoang et al., 2012)	HTC Google one	C	TD	Hamming distance	34	EER 8.09	SD
(Watanabe, 2015)	iOS iPhone 5	S	TD	Neural Network	8	CCR 97.9	SD

Legend: **C**: Cycle-based; **S**: Segment-based; **TD**: Time Domain; **FD**: Frequency Domain; **DTW**: Dynamic Time Warping; **HMM**: Hidden Markov Model; **SVM**: Support Vector Machine; **K-NN**: k-nearest Neighbour; **EER**: Equal Error Rate; **CCR**: Correct Classification Rate; **SD**: Same-Day; **CD**: Cross-Day; **-**: not defined.

Studies to date have used data recorded under laboratory conditions and have sought to profile a range of activities to explore the recognition performance. For example, carrying extra weight, climbing stairs, jogging, and running (Kwapisz et al., 2010; Kwapisz et al., 2011; Nickel et al., 2011c). The results from the same-day and cross-day experiments demonstrate a high degree of variance within the feature vector, which raises the concern whether successful classification could be achieved in practice over time and across a range of differing activities. The composition of the feature vector itself, time and frequency-based features will also play a significant role in recognition performance; however, far too little attention has been paid to explore this across accelerometer and gyroscope sensors in both time and frequency domains.

Nakano and Chakraborty (2017) studied the impact of the dynamic features on the activity recognition system performance. Their analysis revealed that the performance of the efficiency of dynamic features is better than static features in classifying different types of activities.

Notably, none of the previous studies explored the viability of a multi-algorithmic approach (separation of the classifiers depending upon activities) compared with a single classifier approach. Furthermore, none of the prior art attempted to evaluate their approach using data captured over a prolonged period of time under real-life conditions (i.e. days rather than minutes). This includes completely unrestricted movement that could include changes to clothes and shoes, being in a rush, carrying luggage, running and exercising, variations in human mood, time of day effects and ground/surface changes, to name but a few.

3 Experimental Methodology

The study sought to investigate the following:

- Whether real-world use of gait recognition would be a feasible authentication approach
- To investigate the impact upon performance by using a multi-algorithmic approach to classification over a single classifier
- To identify where feature vector variability exists to appreciate under what circumstances gait recognition might or might not be successfully achieved
- To investigate the impact of a decision-based majority voting scheme would have on overall performance.

In each case, where available, the results are compared against the prior art as a baseline to understand the relative merits of the approach.

This study builds upon the outcomes of two prior papers published by the authors:

- In (Al-Obaidi et al., 2018), the authors undertook an investigation into gait recognition using a controlled data capturing environment. This study enabled the authors to propose and evaluate a feature vector algorithm, explore both gyroscope and accelerator sensors feeds and optimise the classification performance. The feature vector algorithm and configurations established in this paper form the baseline starting point in this study.
- In (Alruban et al., 2018), the authors propose and evaluate an activity recognition system which, from core motion data provided by the gyroscope and accelerator, is able to identify what activity an individual is undertaking at any point (e.g. walking, running, going up or down stairs). This system was used as a basis for sampling the real-world dataset into differing sets of activity.

Figure 1 presents the biometric system and flow of data from initial capture through to final decision. The capture, processing, segmentation and feature generation processes have been taken from the authors prior study (Al-Obaidi et al., 2018) (which itself is taken from the prior art). The Activity Identification Model is taken from the second study by the authors (Alruban et al., 2018). This paper seeks to focus upon the classification and decision phases of the biometric system using a unique real-world dataset.
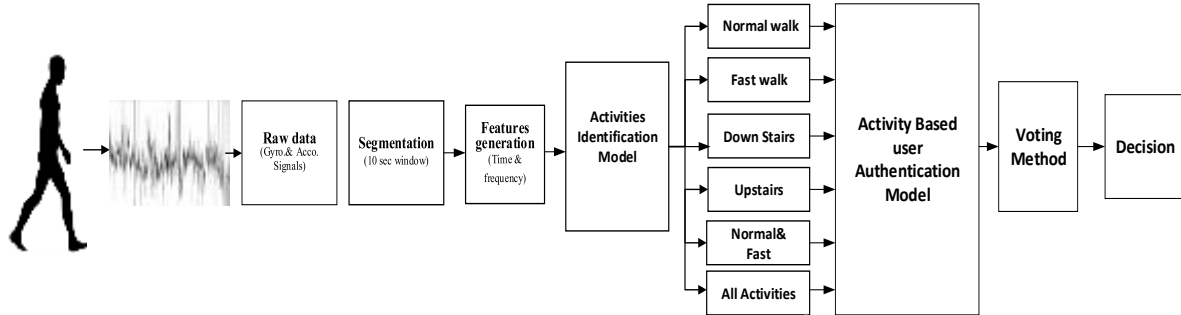

Figure 1: Gait-Based Biometric Process

3.1 Data Collection

Motion data from both the gyroscope and accelerometer were captured during the study. A gyroscope is used to maintain a reference direction in the motion by sensing the degree of orientation in the x, y, and z directions of the smartphone. The axis signal is affected by the direction of the devices orientation. The accelerometer sensor measures the acceleration in metres per second squared (m/s^2) in the x, y, and z directions of the smartphone. An Android application called 'AndroSensor' was used to record the sensor data as it supports most of the sensors that the Android device can offer (AndroSensor, 2018). A Samsung Galaxy S6 smartphone was carried by each to record the sensor data generated by different human physical activities. Each user was asked to place the smartphone in a belt pouch, as presented in Figure 2. Whilst this does not necessarily reflect normal placement for the participants, it was felt an important variable to control and to ensure it remained on the participant (and not placed in a bag). All participants utilised identical handsets and pouches purchased by the authors for the project. The generated data was collected continuously at a rate of 30-32 Hz for the x, y, and z-axes across both the accelerometer and gyroscope sensors. They were asked to start recording using the 'AnroSensor' application each day and stop recording at the end of the day (in order to reduce the size of the resulting dataset capture).

Figure 2: Phone is placed in the right or left belt pouch

The study recruited 44 participants, with a 23/21 male/female split and all were aged between 18 and 56 years old. The collection exercise last for between 7-10 days for each user. This was to help ensure a minimum of 7 days of movement data was collected, as some users would not walk anywhere on a particular day or possibly forget to take the device. Overall, an average of 8 days of data was captured per participant. Figure 3 illustrates the distribution of daily gait activities in minutes for all users per day in which the median of the dataset is 80 minutes a day.
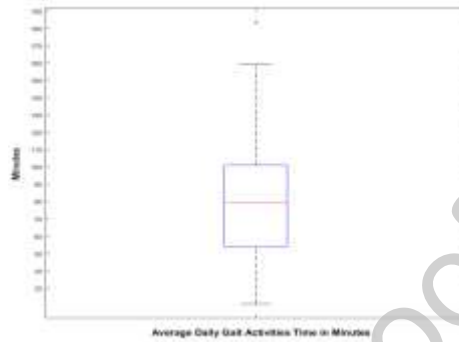

Figure 3: The Average of Daily Gait Activities Time in Minutes for All Users

Having processed the raw data, Table 2 presents the total number of samples of each gait activity as classified by the *Activity Identification Model*. A significant proportion of the samples were identified as normal walking, as would have been expected. The *Activity Identification Model* was able to identify a total of 174,396 samples, an average of 3,963 samples per participant. Notably the other samples category comprised of 576,439 samples. Whilst it would have been useful to be able to classify these samples, either within the existing samples or through the creation of more activities, this does not impact the research being presented in this study. The concept behind a viable continuous and transparent authentication model is to have a sufficient number of samples to be able to perform authentication, not that every sample that has been identified is actually used. At an average of 3,963 samples across an average 8-day capture results in 495 samples per day – this is arguably a more than sufficient volume to achieve the stated goal in this stage of the research. Future research will focus on opportunities to improve the *Activity Identification Model* recognition.

Table 2: Percentage of Identified Real life Activities Samples

Activity Type	No. of Samples	Samples Utilised (%)
Normal	139,907	80%
Fast	12,315	7%
Downstairs	5,175	3%
Up Stairs	16,999	10%
Other samples	576,439	-

3.2 Feature Extraction & Selection

The raw signal data generated by the gyroscope and accelerometer were processed by computing the time and frequency domain features using features identified from the prior art. Table 3 presents a full list of the features. The time-domain features were calculated

directly from the raw data samples, while a Fourier transform was applied to the raw signals across the three sensor axes before computing the frequency domain-based feature set. This process generated 304 unique features from the two domains.

Table 3: Generated Features

Features	Domain	Description
Mean (3)	TD, FD	The mean values in the segment.
Standard Deviation (3)	TD, FD	The standard deviation of the data in the segment.
Median (3)	TD, FD	The median values of the data points in the segment.
Variance (3)	TD, FD	A measure of how far each value in the segment points is from the mean.
Covariance (3)	TD, FD	A measure of how much two variables change together.
Zero crossing rate Minimum	TD, FD	The rate value of sign changes in the segment.
Interquartile range	TD, FD	The range amidst the data. It is the distinction between the upper and lower quartiles in the segment.
Average Absolute Difference (3)	TD, FD	Average absolute difference between the value of each of the segment points from the mean value over the segment values (for each axis).
Root mean square (3)	TD, FD	Square root of the mean of the squares of the acceleration values of the segment.
Skewness (3)	TD, FD	A measure of the symmetry of distributions around the mean value of the segment.
Kurtosis (3)	TD, FD	A measure of the shape of the curve for the segment point's values.
Percentile 25 (3)	TD, FD	The percentile rank is measured by the following formula: $R= (P/100)*(N+1)$. Where R is the rank order of values, P percentile rank, N total number of the data points in the segment.
Percentile 50 (3)	TD, FD	Similar to the Percentile 25 feature; but with the setting of $P=50$.
Percentile 75 (3)	TD, FD	Similar to the percentile 25 feature but with the setting of $P=75$.
Maximum (3)	TD, FD	The largest four values of the segment are calculated and averaged.
Minimum (3)	TD, FD	The largest four values of the segment are calculated and averaged.
Correlation coefficients (3)	TD, FD	The relationship between two axes is calculated. The correlation coefficient is measured between X and Y axes, X and Z axes and Y and Z axes.
Average resultant acceleration (1)	TD, FD	Average of the square roots of the sum of the values of each x, y and z axes in the segment squared.
Difference (3)	TD	Difference of maximal and minimal value of the segment (each axes).
Maximum value (4)	TD	The largest four values of the segment are calculated and averaged.
Minimum value (4)	TD	The smallest four values of the segment are calculated and averaged.
Binned distribution (3)	TD	Relative histogram distribution in linear spaced bins be- tween the minimum and the maximum acceleration in the segment. Ten bins are used for each segment.
Maximum peaks (3)	TD	The average of the largest 4 peaks in the segment.
Minimum peaks (3)	TD	The average of the smallest 4 peaks in the segment.
Peak Occurrence (3)	TD	Calculate how many peaks are in the segment.
Time between peaks (3)	TD	Time in milliseconds between peaks in the sinusoidal waves associated with most activities is calculated and averaged (for each axis).
Interquartile range (3)	TD	Calculating the median of the lower and upper half of the data.
Entropy (3)	FD	The average amount of information produced by a probabilistic stochastic source of data
Energy (3)	FD	The signal energy is equal to the summation across all frequency components of the signal's spectral energy density.

Features were normalised prior to use within the classification stage into a 0-1 range using the maximum value of the feature in the dataset as a base of 1. To validate the effectiveness of the generated feature vectors (comprising of a possible 304 unique features), the dataset was divided to form the reference/training and testing datasets). Due to the large number of samples, undertaking training on 60-70% of the data (which is typical in many biometric-based studies) would have resulted in a very large training dataset (comprising of tens of thousands of samples), which would impact the time taken to create an appropriate classifier in a practical implementation. It also had an impact experimentally, as training a single classifier across the population was taking 3-4 days to complete. However, given the variability in the data particularly over time, it was felt merely taking samples from day 1 and then testing against the remaining days might not be the most appropriate strategy. As shown in the prior art, same and cross-day results illustrated the impact of using data from a single day. As such, the methodology deployed in this study would utilise data from the first x days but would sample only 10% of the data for training. It was hoped this would provide a better capture of a participant's walking style. Each experiment was repeated 10 times in order to establish a baseline performance and remove any bias introduced in one iteration of the sampling.

Recognising a large number of features would place a burden on the classification (particularly on processing/battery limited mobile devices), a dynamic feature selection approach was incorporated within the experiment to explore the impact upon performance of different lengths of feature vector. The algorithm used to create the per user dynamic feature vector was from the authors' prior work (Al-Obaidi et al., 2018). It envisaged that the effectiveness of each feature towards the classification for each user could vary; with some features having a more significant impact for some users over others. The dynamic feature selection mechanism selects features based upon a calculation of the standard deviation of a user's features with the smaller standard deviation chosen to prioritise features

3.3 Classification

The prior art and indeed the authors prior work had identified a number of candidate classifiers. The decision was to focus upon the Feed-Forward Multilayered Perceptron (FFMLP). Whilst other classifiers, such as SVM had demonstrated good performance in the prior work (highlighted in Table 1), SVM suffers when there are large training sets and also lacks the ability to optimise and refine the generalisation and subsequently classification performance (Karatzouni 2014; Saevanee et al. 2015). For each identified activity, eleven different FF MLP neural network configurations were evaluation (hidden layer size varying from 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, and 60).

For each classification, one user was selected as the authorised user and all remaining users were used as unauthorised. Each user in the population would have the chance of acting as the authorised user, with performance rates averaged across the population. Table 4 illustrates a break down of the experimental configurations:

Table 4: Experimental Configuration

Variable	No. of Configurations
Network size	11
Feature size	8
Activity	6
Participants	44
Training (repeated)	10

Total		**232,320**

4 Experimental Results

The following two sections will provide an extraction of the key results resulting from all the experimental tests performed, with a focus upon addressing each of the identified research questions. Results will be presented using the Equal Error Rate (EER) as is typical for biometric-based studies to enable comparison between the results themselves and the prior art.

4.1 Classification Results

The first set of results, sought to explore the feature vector and the impact upon performance. As illustrated in Table 5, the general performance achieved increased as the number of features increased (to the maximum of 304). This is true against all three activities of normal, fast and the combined normal and fast walking activities. In the authors' prior work (Al-Obaidi et al., 2018) utilising the controlled data, the results demonstrated that a dynamic feature vector of between 10-160 features provided the best classification performance and presents an interesting insight into the realisation of a gait recognition scheme, moving from controlled data to actual real-life data. The results would suggest, a longer feature vector is beneficial in such circumstances, with the possible effect of helping to mitigate against the impacts of larger signal variations that exist in real-life data.

Table 5: Population Average EER (%) of Normal, Fast and Normal and Fast Walking Activity Utilising Different Feature Subsets (with a fixed 40 hidden node neural network)

No. of Features	Normal Walking (EER (%))	Fast Walking (EER (%))	Normal & Fast Walking (EER (%))
10	28.69	26.8	27.74
50	17.50	19.47	17.56
100	16.39	17.56	15.20
110	14.53	15.20	14.45
160	15.90	14.60	14.56
200	14.50	13.84	14.17
250	14.04	13.38	13.69
304	11.38	11.32	12.49

Drawing upon the best configuration from Table 5 and the full feature vector, Table 6 presents a comparison of the EERs obtained against the activities listed. For typical walking (normal and fast), the use of an individual classifier rather than a single classifier for all walking data, provides a clear advantage, supporting the use of an activity recognition and multi-algorithmic approach to classification. The error rates themselves are also in a similar range to many of the prior studies that utilised highly controlled datasets, which is promising. It is clear however, the classifiers struggle with walking activities involving stairs. In one sense, the rhythmic gait cycle which this biometric utilises is arguably disrupted, which a single classifier would subsequently struggle with, given quite differing signals all linked to the authorised user. However, notably, the results using the single activity classifier performed very poorly, suggesting even though the cycle is different to walking, the classifier is struggling to identify a suitable discriminate signal from the data.

Table 6: EER (%) for Individual per activity for all features

Activity Type	EER (%)
Normal	11.38
Fast	11.32
Down Stairs	24.52
Upstairs	27.33
Normal & fast	12.49
All Activity	15.08

Figures 4 and 5 present a breakdown of individual EERs across the normal and all activity classifiers. An analysis of the general pattern of results against both classifiers illustrates a strong correlation. This is largely a result of the normal walking data representing the largest proportion of the dataset. The multi-algorithmic approach using single activity based recognition has a clear improvement in performance across a significant proportion of the population. The largest impact was *User 30* that experienced +10% worsening in the performance. Further analysis of this user showed a relatively higher proportion of fast walking samples over normal walking. Figure 4 exhibits that a significant proportion of user's performance was under 10% EER (a value typically linked to positive viability of the approach given its continuous evaluation). *User 29* achieved the best performance of 1.94% EER. In comparison *User 30* achieved the worse EER of 46.80%. It was found that 12 participants (*Users 5, 10, 15, 16, 20, 23, 25, 26, 28, 29, 39 & 42*) achieved an EER of less than 5% each whilst another 6 (*Users 26, 22, 30, 33, 37 & 38*) accomplished an EER of more than 20% each. As with most behavioural-based biometric modalities, the results suggest that for a proportion of the population, the technique proves viable; however, for a smaller proportion it is not.
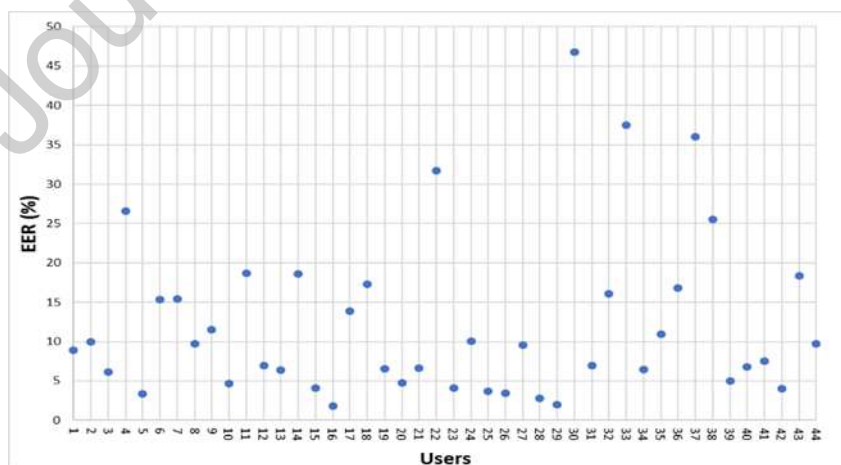

Figure 4: The EER (%) of Individual Performance for Normal Walking Activity
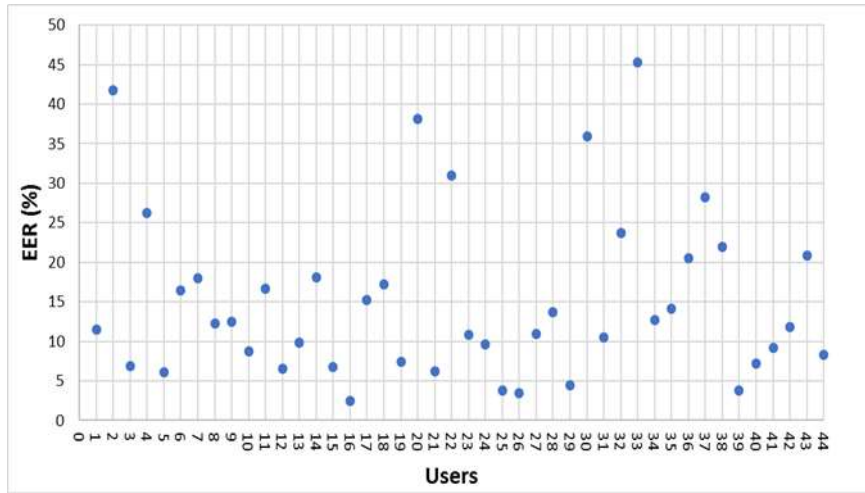
Figure 5: The EER (%) of Individual Performance for All Activities

From the analysis (illustrated in Table 7) , it is clear that the majority of the participants score less than 10% EER with a multi-algorithmic approach. In contrast, with a single classifier, the majority of participants scores were greater than 10% EER, possibly pushing it outside of the area of viability.

Table 7: Summary of Individual Performance for Each Activity

Activity Type	# Users				
	ERR (%) <=5	ERR (%) >5-10	ERR (%) >10-15	ERR (%) >15-20	ERR (%) >20
Normal	12	16	4	6	6
Fast	13	13	7	5	6
Normal and Fast	8	18	7	5	6
All Activates	5	12	10	7	10

4.2 Decision Results – Applying Majority Voting

In a continuous authentication scheme, it becomes less critical to make a final authentication decision on a single sample (as would be the case in most point-of-entry approaches). With a sufficient frequency of samples, it is possible to make a final identity decision based upon a range of samples. The prior art have primarily focussed upon two approaches – majority or quorum voting, with the former having achieved stronger performance and therefore was utilised in this study (Nickel et al. 2011a).

Table 8 presents the results of applying the majority voting scheme. The results present both the mean and median averages – the former providing a good illustration of overall average performance and the latter enabling a greater insight into the impact of outliers on the mean average performance. The results produce significant enhancements on the system performance. In comparison with a single- sample evaluation, normal, fast, down and upstairs walking activities are improved by an average rate of 53%, 46%, 53% and 25% respectively. Moreover, analysing the performances for the merged normal and fast and the four combined activities also demonstrates a significant improvement with an average rate of 47% and 49% accordingly.

Focusing on the median, which is less sensitive to outliers, it can be seen that EER median-based is quite a bit better than the EER mean values. The EER dropped down to 2.14%, 1.89%, 5.65% and 14.81% for individual activities (i.e. normal, fast, down and upstairs). These results demonstrated the negative impact of outliers and the importance of removing outliers during the pre-processing stage of the biometric process (as is typically recommended in practice). With the exception of the activities involving the stairs, performance rates across the board are on average in the sub-5% range which is more than sufficient to support the viability of the approach within a *TAS (Transparent Authentication System) (*Clarke, 2011*)*. With respect to the number of samples required to achieve this, whilst specific best results are achieved from 17 samples (2 minutes 50 seconds) onwards, the results generally support the use of as many samples as possible. With the additional samples helping to provide the additional confidence. In practice the decision over the time-window would be left to the wider *TAS* and the risk management processes – which would manage the samples and biometric modalities within the system to provide comprehensive and continuous identify verification.

Table 8: Majority Voting Results for each Number of Samples across All Gait Activities

Activity Type # Samples /Time (second)		Normal	Fast	Down Stairs	Upstairs	Normal And Fast	All Activities
3 (30)	Median	7.93	5.50	15.65	24.72	8.02	8.53
	Mean	9.90	9.77	18.97	26.48	9.99	11.99
5 (50)	Median	6.41	4.19	13.24	21.08	7.39	7.50
	Mean	8.52	7.95	15.85	23.41	8.82	10.60
7 (1:10)	Median	5.38	3.38	10.57	18.69	6.64	7.08
	Mean	7.78	7.96	15.08	23.62	8.07	9.95
9 (1:30)	Median	5.29	3.38	8.74	18.98	5.48	6.27
	Mean	7.30	7.35	13.95	23.05	7.54	9.52
11 (1:50)	Median	4.52	2.91	7.02	19.05	5.10	4.93
	Mean	7.06	7.11	12.71	23.10	7.37	9.04
13 (2:10)	Median	3.56	2.96	6.32	16.13	4.70	5.24
	Mean	6.62	6.96	12.05	21.28	6.99	8.72
15 (2:30)	Median	4.65	3.30	6.65	16.64	4.57	4.70
	Mean	6.44	6.79	12.75	21.65	6.79	8.24
17 (2:50)	Median	3.46	**1.89**	5.71	16.18	4.26	4.71
	Mean	6.34	6.40	13.24	21.79	6.80	8.33
19 (3:10)	Median	3.13	2.24	6.88	**14.81**	3.92	4.10
	Mean	6.05	6.89	13.67	**20.55**	6.37	8.04
21 (3:30)	Median	3.06	2.48	6.48	15.12	4.05	3.96
	Mean	5.79	7.16	13.69	20.49	6.35	7.74
23 (3:50)	Median	3.21	3.25	5.94	16.58	4.06	3.95
	Mean	6.08	7.59	12.54	21.59	6.26	7.84
25 (4:10)	Median	3.32	2.48	**5.65**	18.42	3.82	3.52
	Mean	5.80	7.04	12.12	22.27	6.11	7.61
27 (4:30)	Median	2.48	3.48	6.59	19.02	3.67	**3.36**
	Mean	5.37	7.05	12.03	23.22	6.19	7.48
29 (4:50)	Median	2.90	3.15	6.15	16.25	3.99	3.49
	Mean	5.55	7.09	11.43	21.89	6.34	7.69
31 (5:10)	Median	**2.14**	2.89	6.06	16.76	**3.50**	3.38
	Mean	5.31	6.43	11.91	22.33	5.87	7.45

5 Discussion

The study sought to consider four research question revolving around evaluating the viability of gait recognition using uncontrolled real world data – the first such study that has both removed the controlled capturing of gait-based samples but also introduced the use of an *Activity Identification Model* to process and label the data so that a multi-algorithmic approach to classification could be examined.

Overall, the results from the classification stage demonstrate average performances that fall within the scope of viability. The results also demonstrate that the use of individual classifiers for each activity, does provide for an improvement in overall performance. However, examining the performance difference from normal, fast and the combined normal and fast walking show only a marginal improvement. Given the current activities that are identified by the *Activity Identification Model,* the results suggest that removing the stair-based activities (whose performance is not good) from being classified at all and then subsequently applying a single classifier for all walking based activities identified would result in a similar performance than using individual classifiers for normal and fast walking. So the evidence for a multi-algorithmic approach based upon the currently identified activities is less strong. However, if a desire exists to include stair-based activities and in the future the *Activity Identification Model* can differentiate a wider set of activities, the data does suggest a multi-algorithmic approach would be a better performance classification strategy.

It is worth recognising and highlighting the user variance that exists within the results. Whilst a good proportion have achieved error rates that would make the approach viable for them, a proportion of the population have not – as per the majority of behavioural based biometric modalities. Whilst the wider *TAS* should still be able to support these users through the use of other biometric modalities to which they can achieve a good recognition performance against, it also worth acknowledging that the individual techniques need to establish a mechanism for self-evaluating the viability of individual users, so it is able to determine whether the approach is viable or not.

Exploring the composition and length of the feature vector raised an interesting result over the prior art and previous studies undertaken by the authors. The prior art had suggested there was value in prioritising and reducing the feature vector – with the consequential advantage of reducing classification complexity (and the curse of dimensionality). However, using real-life data and consequently the increased variance in the gait signal data that existed, the study has demonstrated that using a feature dense sample is critical to achieving a more reliable and consistent performance. Whilst this will of course have a practical impact, longer training times and more complex classifiers, the size and nature of these samples and classifiers is arguably not beyond the majority of modern Smartphones. Furthermore, *TAS* are often implemented in a cloud-based infrastructure to permit off-boarding computational complex actions into the cloud (Al Abdulwahid, 2017).

Table 9 compares the controlled experiment results achieved from the authors' prior work (Al-Obaidi et al., 2018), which itself is amongst the best performing results from the prior art, with the performance achieved using the real world data. The results do show there is a significant difference in performance achieved using real-life versus controlled data, which raises some questions over the viability of the approach and the conclusions that prior studies have drawn. However, the table also shows that the addition of the majority voting decision

scheme, for use in a continuous authentication scheme, does subsequently reduce this error rate down to a more acceptable level of performance for normal/fast walking based activities.

Table 9: Comparing Controlled and Realistic System Performance

Activity Type	Controlled Dataset (Cross Day)	Realistic System Without Voting	Realistic System Best Voting		Decision Time
		EER (%)			
Normal	2.09	11.38	Median	2.14	5:10s
			Mean	5.31	5:10s
Fast	3.91	11.32	Median	1.89	2:50s
			Mean	6.43	5:10s
Down Stairs	23.45	24.52	Median	5.65	4:10s
			Mean	11.43	4:50s
Upstairs	23.32	27.33	Median	14.81	3:10s
			Mean	20.55	3:10s
			Mean	5.87	5:10s
All Activities	6.58	15.08	Median	3.50	4:30s
			Mean	7.45	5:10s

Finally, it is worth reflecting upon a number of limitations and restrictions associated with the study in order to appreciate the opportunities for further research:

- The collected dataset was acquired using a single type of mobile device (Samsung Galaxy S6). Whilst different devices will likely exhibit differing sensor sensitivities, as the devices are personal to individuals (i.e. both training and test data will be captured using the same device), any differences that might exist should not impeded recognition performance. However, further investigation of this would be worth while as would the wider collection of a larger volume of participants to aid in better generalising the results.

- The *Activity Identification Model* is able to provide sorting of raw signal data into a series of simple walking activities. For future work, it would be useful for this system is recognise a wider set of activities and in particular introduce context into into the recognition process. For example, recognising the difference of when an individual is walking uphill, downhill or on the flat using signal data from the GPS sensor. Or when an individual is walking with someone else and whose gait might adapt accordingly – for example when a parent is walking with a young child. The addition of this context sensitive approach will likely enable a more refined set of classifiers, remove variability within each and subsequently lead to stronger classification performance.

- The evaluation of this study involved a large number of iterative tests to evaluate differing feature vectors, classification approaches and decision schemes. What this highlighted was that individual performances differed across configurations. It is therefore important when looking to deploy this practically to define a process by which a user's individual classification configuration can be optimised. This is likely to have a computational overhead which will need to be explored in further depth to appreciate where such functionality should reside – on the device or in the cloud. Once the classifiers have been created, it is anticipated that the authentication process will be easier achieved on the mobile device.

6 Conclusions and Future Work

The evaluation of real-world smartphone-based gait recognition has revealed that it could provide a secured and appropriate user authentication system for a significant proportion of individuals. The use of real-world data does introduce an increased variability into the gait-based signal data and based upon classification results alone, does introduce some questions over it's general variability. However, incorporated within a decision-based majority voting scheme, population-based performances are very encouraging and aligned to existing behavioural-based biometric modalities.

The use of additional sensor data, both within the gait-based system and within the wider *TAS* will arguably provide a useful resource of additional context sensitive information which can help inform the decision logic and provide more refined classification strategies. The combination of which will provide for more robust techniques that are able to adapt and better manage a wider set of circumstances the user might find themselves in.

Finally, future research should also focus upon how robust the technique is against targeted attacks. The study utilised a standard research methodology for determining the rate at which impostors are accepted onto the system, but this approach does not incorporate targeted attacks (e.g. when individuals specifically look to mimic an authorised users behaviour) to understand how susceptible the technique is to this type of attack. Indeed, there is little work to date looking at the susceptibility of behavioural based biometrics that are used in a continuous authentication schemes.

Credit Author Statement

Hind Alobaidi: Methodology, software, validation, writing – original draft

Nathan Clarke: Conceptualization, methodology, writing – original draft, writing – review & editing, supervision

Fudong Li: Software, data curation, writing – review & editing, supervision

Abdulrahman Alruban: Software, validation, writing – review & editing

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Al AAbdulwahid, A.A. "Federated Authentication using the Cloud (Cloud Aura)," University of Plymouth, Plymouth, 2017. http://hdl.handle.net/10026.1/9596

Al-Obaidi, H., Li, F., Clarke, N., Ghita, B. and Ketab, S., 2018, June. A multi-algorithmic approach for gait recognition. In ECCWS 2018 17th European Conference on Cyber Warfare and Security, pp. 20-28

Alruban, A., Alobaidi, H., Clarke, N., Li, F. "Physical activity recognition by utilising smartphone sensor signals" In 8th International Conference on Pattern Recognition Applications and Methods, 2019, pp. 342-351 SciTePress.

AndroSensor, "AndroSensor." [Online]. Available:

https://play.google.com/store/apps/details?id=com.fivasim.androsensor&hl=en_GB.

Antos, S. A., Albert, M. V. and Kording, K. P. "Hand, belt, pocket or bag: Practical activity tracking with mobile phones," J. Neurosci. Methods, vol. 231, pp. 22–30, 2014, doi: 10.1016/j.jneumeth.2013.09.015.

Capela, N. A., Lemaire, E. D.,Baddour, N., Rudolf, M., Goljar, N. and Burger, H. "Evaluation of a smartphone human activity recognition application with able-bodied and stroke participants," Journal of Neuroeng. Rehabil., vol. 13, no. 1, pp. 1–10, 2016, doi: 10.1186/s12984-016-0114-0.

Clarke, N. "Transparent user authentication: biometrics, RFID and behavioural profiling". London: Springer, 2011.

Derawi, M. O. "Smartphones and Biometrics: Gait and Activity Recognition", Doctoral dissertations at Gjøvik University College, 2012. https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/144368

Derawi, M. and Bours, P. "Gait and activity recognition using commercial phones," Comput. Secur., vol. 39, no. July 2015, pp. 137–144, 2013, doi: https://doi.org/10.1016/j.cose.2013.07.004

Derawi, M. O., Nickel, C., Bours, P. and Busch, C. "Unobtrusive User-Authentication on Mobile Phones Using Biometric Gait Recognition," in 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2010, pp. 306–311, doi: 10.1109/IIHMSP.2010.83.

Frank, J., Mannor, S. and Precup, D. (2010a)"Activity and Gait Recognition with Time-Delay Embeddings Time-Delay Embeddings," in AAAI Conference on Artificial Intelligence, 2010, pp. 407–408, doi: 10.1.1.224.7630.

Frank, J., Mannor, S. and Precup, D.(2010 b) "A novel similarity measure for time series data with applications to gait and activity recognition," UbiComp'10 - Proc. 2010 ACM Conf. Ubiquitous Comput., pp. 407–408, 2010, doi: https://doi.org/10.1145/1864431.1864460

Gadaleta, M. and Rossi, M. "IDNet: Smartphone-based gait recognition with convolutional neural networks," Pattern Recognit., vol. 74, pp. 25–37, 2018, doi: 10.1016/j.patcog.2017.09.005.

Gafurov, D. "Performance and Security Analysis of Gait-based User Authentication", Doktoravhandling, University of Oslo, 2008 http://urn.nb.no/URN:NBN:no-19706

Ganti, R. K., Srinivasan, S. and Gacic, A. "Multisensor Fusion in Smartphones for Lifestyle Monitoring," in 2010 International Conference on Body Sensor Networks, 2010, pp. 36–43, doi: 10.1109/BSN.2010.10.

Hestbek, M. R., Nickel, C. and Busch, C. "Biometric gait recognition for mobile devices using wavelet transform and support vector machines," 2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP), 2012, pp. 205-210.

Hoang, T., Choi, D., Vo, V., Nguyen, A. and Nguyen, T. "A Lightweight Gait Authentication on Mobile Phone Regardless of Installation Error", in Security and Privacy Protection in Information Processing Systems, pp: 83-101 2013

Hoang, T., Nguyen, T., Luong, C., Do, S. and Choi, D. "Adaptive cross-device gait recognition using a mobile accelerometer," J. Inf. Process. Syst., vol. 9, no. 2, pp. 333–348, 2013, doi: 10.3745/JIPS.2013.9.2.333.

Hocking, C.G., Furnell, S.M., Clarke, N.L. and Reynolds, P.L., 2013. Co-operative user identity verification using an Authentication Aura. Computers & security, 39, pp.486-502

Karatzouni, S. "Non-Intrusive Continuous User Authentication for Mobile Devices", University of Plymouth, 2014, available: https://pearl.plymouth.ac.uk/handle/10026.1/3225

Kwapisz, J. R., Weiss, G. M. and Moore, S.A. "Cell Phone-Based Biometric Identification," Biometrics Theory Appl. Syst. (BTAS), 2010 Fourth IEEE Int. Conf., pp. 1–7, 2010. doi: 10.1109/BTAS.2010.5634532.

Kwapisz, J. R., Weiss, G. M. and Moore, S.A. "Activity recognition using cell phone accelerometers," ACM SIGKDD Explorations Newsletter, Volume 12, Issue 2, December 2010, pp 74–82, https://doi.org/10.1145/1964897.1964918

Muaaz, M. "A Transparent and Continuous Biometric Authentication Framework for User-Friendly Secure Mobile Environments," in Adjunct Proceedings the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2013), pp. 4–7

Muaaz, M. and Mayrhofer, R. "An Analysis of Different Approaches to Gait Recognition Using Cell Phone Based Accelerometers," Proceedings of International Conference on Advances in Mobile Computing & Multimedia, December 2013 Pages 293–300, doi:https://doi.org/10.1145/2536853.2536895

Muaaz, M. and Nickel, C. "Influence of different walking speeds and surfaces on accelerometer-based biometric gait recognition," 2012 35th International Conference on Telecommunications and Signal Processing (TSP), 2012, pp. 508-512,

doi: 10.1109/TSP.2012.6256346.

Nakano, K. and Chakraborty, B. "Effect of dynamic feature for human activity recognition using smartphone sensors," 2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST), 2017, pp. 539-543, doi: 10.1109/ICAwST.2017.8256516.

Nickel, C., Brandt, H. and Busch, C., (2011a). "Classification of acceleration data for biometric gait recognition on mobile devices". In: Brömme, A. & Busch, C. (Hrsg.), BIOSIG 2011 – Proceedings of the Biometrics Special Interest Group. Bonn: Gesellschaft für Informatik e.V.. (S. 57-66).

Nickel, C., Brandt, H. and Busch, C. (2011b) "Benchmarking the performance of SVMs and HMMs for accelerometer-based biometric gait recognition," 2011 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2011, pp. 281-286, doi: 10.1109/ISSPIT.2011.6151574.

Nickel, C., Brandt, H. and Busch, C. (2011c) "Classification of Acceleration Data for Biometric Gait Recognition on Mobile Devices," in Proceedings of the Special Interest Group on Biometrics and Electronic Signatures, 2011, pp. 57–66.

Nickel, C. and Busch, C. "Classifying accelerometer data via hidden Markov models to authenticate people by the way they walk," in IEEE Aerospace and Electronic Systems Magazine, vol. 28, no. 10, pp. 29-35, Oct. 2013, doi: 10.1109/MAES.2013.6642829.

Nickel, C., Busch, C., Rangarajan, S. and Möbius, M. (2011d), "Using Hidden Markov Models for accelerometer-based biometric gait recognition," 2011 IEEE 7th International Colloquium on Signal Processing and its Applications, 2011, pp. 58-63, doi: 10.1109/CSPA.2011.5759842.

Nickel, C., Derawi, M. O., Bours, P. and Busch, C. (2011e) "Scenario test of accelerometer-based biometric gait recognition," 2011 Third International Workshop on Security and Communication Networks (IWSCN), 2011, pp. 15-21, doi: 10.1109/IWSCN.2011.6827712.

Nickel, C., Wirtl, T. and Busch, C. "Authentication of Smartphone Users Based on the Way They Walk Using k-NN Algorithm," 2012 Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2012, pp. 16-20, doi: 10.1109/IIH-MSP.2012.11.

Ngo, T. T., Makihara, Y., Nagahara, H., Mukaigawa, Y. and Yagi, Y. "The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication," Pattern Recognit., vol. 47, no. 1, pp. 228–237, Jan. 2014, doi: https://doi.org/10.1016/j.patcog.2013.06.028.

Phan DT., Dam N.NT., Nguyen MP., Tran MT., Truong TT. "Smart Kiosk with Gait-Based Continuous Authentication". In: Streitz N., Markopoulos P. (eds) Distributed, Ambient, and Pervasive Interactions. DAPI 2015. Lecture Notes in Computer Science, vol 9189. Springer, Cham. https://doi.org/10.1007/978-3-319-20804-6_18

Rana, J.R. "Obtrusive and Unobtrusive Authentication of Portable Devices for Security," International Journal of Research, Volume 2, Issue 4, 2015. page:766-771

Saevanee, H., Clarke, N., Furnell, S., Biscione, V. "Continuous user authentication using multi-modal biometrics", Computers & Security, Volume 53, 2015, Pages 234-246, ISSN 0167-4048, DOI: https://doi.org/10.1016/j.cose.2015.06.001.

Shoaib, M., Bosch, S., Incel, O. D., Scholten, H. and Havinga, P. J. M. "Complex human activity recognition using smartphone and wrist-worn motion sensors," Sensors (Switzerland), vol. 16, no. 4, pp. 1–24, 2016, doi: https://doi.org/10.3390/s16040426

Sprager, S. and Juric MB. Inertial Sensor-Based Gait Recognition: A Review. Sensors (Basel). 2015 Sep 2;15(9):22089-127. doi: 10.3390/s150922089. PMID: 26340634; PMCID: PMC4610468.

Sprager, S. and Zazula, D. "A cumulant-based method for gait identification using accelerometer data with principal component analysis and support vector machine," WSEAS Trans. Signal Process., vol. 5, no. 11, pp. 369–378, 2009.

Statista (2021) "Number of smartphone subscriptions worldwide from 2016 to 2026", available at: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

Thang, H. M., Viet, V. Q., Dinh Thuc, N. and Choi, D. "Gait identification using accelerometer on mobile phone," 2012 International Conference on Control, Automation and Information Sciences (ICCAIS), 2012, pp. 344-348, doi: 10.1109/ICCAIS.2012.6466615.

Watanabe, Y. "Influence of Holding Smart Phone for Acceleration-Based Gait Authentication," 2014 Fifth International Conference on Emerging Security Technologies, 2014, pp. 30-33, doi: 10.1109/EST.2014.24.

Watanabe, Y. "Toward Application of Immunity-based Model to Gait Recognition Using Smart Phone Sensors: A Study of

Various Walking States," Procedia Comput. Sci., vol. 60, pp. 1856–1864, 2015, doi: https://doi.org/10.1016/j.procs.2015.08.296.

- **Biography**

Dr Hind Alobaidi

Dr Al-Obaidi is a lecturer in computer science department, college of education, Baghdad University, Iraq. Dr Hind received a bachelor's and MSc degree in computer science from University of Technology, Iraq. She is awarded a PhD in Cyber security from the University of Plymouth, UK. Her research interests in information security, biometrics, and gait recognition.

Professor Nathan Clarke

Professor Clarke is a Professor in Cyber Security and Digital Forensics at the University of Plymouth. He is also an adjunct Professor at Edith Cowan University in Australia. His research interests reside in the areas of cyber security, biometrics, and digital forensics. Prof Clarke has over 200 outputs consisting of journal papers, conference papers, books, edited books, book chapters and patents. He created and co-chairs the IFIP 11.12 Symposium on the *Human Aspects of Information Security & Assurance.* Prof Clarke is a chartered engineer, a fellow of the British Computing Society (BCS), a senior member of the IEEE and a full member of the Chartered Institute of Information Security. He is the author of *Transparent Authentication: Biometrics, RFID and Behavioural Profiling* published by Springer. Prof Clarke has been involved in a variety of national and international research projects, to the value of £20 million. He has graduated over 42 doctoral students.

Dr Fudong Li

Fudong Li is a senior lecturer in Cyber Security and course leader of BSc Cyber Security and Forensic Computing in the School of Computing at University of Portsmouth, Portsmouth, UK. Dr Li has over 10 years working experience in the cyber security domain. His research interests are in the areas of information security, biometric authentication, and digital forensics. He has published over 50 conference papers and journal articles in the aforementioned areas and ten PhD completions. Fudong worked on several national and international research projects, such as an EPSRC project (UK) in the area of biometrics and network security and the EU FP7 GERYON project in the domain of network security for real-time multimedia

communications, as work package leaders. Dr Li is the publicity chair of the HAISA conference; also he regularly reviews papers for conferences and journals (including Computers and Security). Apart from research, Dr Li also enjoys teaching various cyber security related topics and he is a Fellow of the Higher Education Academy

Dr Abdulrahman Alruban

Abdulrahman holds a Ph.D. from the University of Plymouth, UK, in Computer Science, and a master's degree from the University of Glamorgan, UK, in Computer Systems Security. A Bachelor degree from Southampton Solent University, UK, in Information Technology. He published more than fifteen scientific papers as a senior researcher and participant in several scientific conferences and magazines in the field of artificial intelligence, digital forensics and cybersecurity.

As an inventor, Abdulrahman invented a digital forensics system which patented in the British and American Patent Offices in 2017. He obtained several honours and awards such as the Golden Medal of the 12th International Invention Fair in the Middle East 2020, Second Place of the GCC Patent Office Award 2020, and First Place Award of the Ministry of Economy, UAE Hackathon 2020. As a recognition, Abdulrahman was chosen in 2018 to represent outstanding Saudi students in the United Kingdom and delivered a speech on their behalf at the Saudi Embassy in London. Also, as a recognition of his scientific distinction, he received the Scientific Creativity Award from Majmaah University in 2018.