




# User's Authentication Using Information Collected by Smart-Shoes

Luca Brombin, Margherita Gambini, Pietro Gronchi, Roberto Magherini,  
Lorenzo Nannini, Amedeo Pochiero, Alessandro Sieni,  
and Alessio Vecchio<sup>(✉)</sup> 

University of Pisa, Pisa, Italy  
alessio.vecchio@unipi.it

**Abstract.** In the last years, smart-shoes moved from the medical domain, where they are used to collect gait-related data during rehabilitation or in case of pathologies, to the every-day life of an increasing number of people. In this paper, a method useful to effortlessly authenticate the user during gait periods is proposed. The method relies on the information collected by shoe-mounted accelerometers and gyroscopes, and on the distance between feet collected by Ultra-WideBand (UWB) transceivers. Experimental results show that a balanced accuracy equal to 97% can be achieved even when information about the possible impostors is not known in advance. The contribution of the different information sources, accelerometer, gyroscope, and UWB, is also evaluated.

**Keywords:** Gait · Authentication · Biometrics · Wearable device · Smart-shoe

## 1 Introduction

Wearable devices gained widespread popularity during the last years. Smart-watches and smart-wristbands are daily used by a large fraction of people to track their activities, estimate the amount of burnt calories, and as an unobtrusive means for receiving notifications [2, 20]. More recently, also smart-shoes started being adopted by the general public. In fact, smart-shoes were initially used in the e-health domain, to collect data about gait-related pathologies [7, 9]. Now, they are increasingly used by sports professionals and amateur athletes to track their sessions and obtain detailed information about their performance. Several major brands operating in the footwear sector now include smart-shoes in their catalogs.

Smart-shoes are generally equipped with an Inertial Measurement Unit (IMU) and a transceiver. The former is used to capture the movements of the user, the latter to transmit the data to an external device, such as a smartphone. In some cases, pressure sensors may be available as well [18]. Data collected by means of smart-shoes are not only highly informative about the running/walking

style of the user, they are also able to provide abundant information about the identity of the user himself. Several studies demonstrated that accelerometric information collected during gait periods can be used to identify or authenticate the user [8, 14, 16]. In authentication, the goal is to automatically understand if the current user is the legitimate one or not. In identification, the goal is to automatically recognize the current user among a set of known ones [5, 25]. Both possibilities can be extremely useful: authentication, to reduce the burden required from the user of mobile devices, who is frequently asked to confirm his/her identity through pins and/or passwords; identification, to customize the parameters of operations of devices shared among a set of people.

In this paper, we focus on an authentication method based on smart-shoes. Information provided by accelerometers and gyroscopes is used to understand if the user is the legitimate one or not. Besides the information provided by IMUs, the method also relies on distance information collected by means of Ultra-WideBand (UWB) transceivers. The possibility of collecting distance information via UWB was considered because of the increasing diffusion of IEEE 802.15.4-2011 in the wearable domain [23, 24]. IEEE 802.15.4-2011 is a standard for low-rate personal area networks that also includes a UWB physical layer. Results demonstrate that reliable authentication of the legitimate user is possible also when the learning phase does not make use of other users' gait samples.

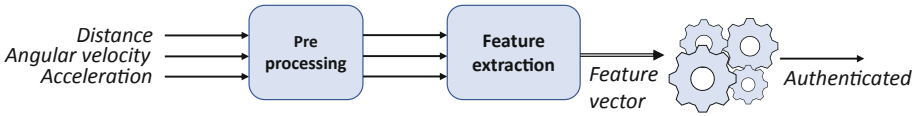
## 2 Related Work

The limited input interfaces of wearable devices and their personal nature gave rise to new challenges in the security domain. For this reason, the possibility of using a person's gait as an authentication behavioral biometrics has been explored in recent years [17].

Some gait-based authentication methods relied on the acceleration signal collected by a smartphone attached to the hip [6, 15, 16]. In other cases, acceleration was collected using a wrist-worn device, as this position can be more comfortable for the end users [4, 10]. The security strength of a smartphone-based authentication system against zero-effort and impersonator attacks was studied in [13], where professional actors tried to mimic the gait style of other users. Results show that mimicking does not increase the chances of obtaining a false positive, i.e. the erroneous recognition of another user as the legitimate one.

More recently, Fangmin et al. [21] proposed a speed-adaptive gait cycle segmentation method and an individualized method for setting the threshold used to distinguish the legitimate user from possible impostors. These mechanisms make easier to identify gait cycles even in the presence of changes in gait speed. In addition, adapting the threshold contributes to reducing the authentication error. Finally, the proposed adaptive methods were compared with the ones obtained by other state-of-the-art techniques. Results show improvements both in gait recognition and user authentication.

Other authors studied the possibility of using One-Class Classification (OCCs) to achieve biometrics-based continuous authentication [12]. Consistently with the OCC philosophy, the approach relies on the availability of a sufficient



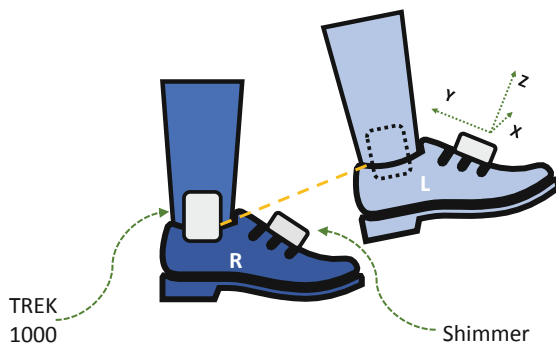
**Fig. 1.** Overview of the authentication method.

number of positive (genuine) behavioral samples only, while ruling out the negative (impostor) ones. Four methods – Elliptic Envelope (EE), Vector Machine (SV1C), Local Outlier Factor (LOF), and Isolation Forest (IF) – along with their fusions were investigated. The performance was assessed on four distinct behaviorimetric datasets, which comprised both motion and touch gesture patterns. SV1C and LOF achieved the best results in terms of error rates. The performance of OCC methods was also compared with the performance of eight well-known multi-class classifiers. SV1C and LOF outperformed half of the investigated more traditional classifiers, therefore proving the feasibility of OCC for continuous authentication.

Identification by means of inertial sensors attached to users’ feet was studied in [11]. Gait data were collected in terms of 3-axial acceleration and 3-axial angular velocity at both feet. Features were extracted using discrete cosine transform restricted to the low frequencies. Then, identification was carried out by means of a random forest classifier, with a group of eight users.

Besides user’s authentication and identification, gait has been proposed as a method for sharing a secret between devices worn by the same subject. In particular, BANDANA is an authentication scheme that allows two wearable devices, placed at random body location, to pair in a secure way through a fresh secure shared secret extracted from user’s gait [19]. First, the data produced by each sensor is rotated so that each z-axis is oriented in the opposite direction to gravity, then the signal is denoised by a bandpass filter. A quantization process produces fingerprint bits evaluating the energy difference between  $Z_i$  and  $A$ , where  $Z_i$  is the  $i$ th gait cycle and  $A$  is the average gait cycle. The higher is this difference, the more reliable the related bit is, then the least reliable ones are discarded. Each device is then able to reach the same key using fuzzy cryptography. Being the gait style unique, only devices on the same body are authenticated.

Differently from most of the above-mentioned works, we study the effectiveness of gait-based authentication using information collected by means of smart-shoes. We believe that this category of wearable devices will be even more popular in the next future. In addition, our study does not only consider data generated by IMUs, but also includes the distance between feet collected by means of UWB. Finally, the proposed method does not rely on the availability of possible impostors’ gait samples, but operates according to a realistic usage scenario where only the data produced by the legitimate user is available during the training phase.



**Fig. 2.** Position of devices on users' feet and orientation of axes of Shimmer devices.

**Table 1.** Volunteers' physical characteristics.

ID	Age	Gender	Height (cm)	Weight (kg)
1	24	M	180	95
2	24	M	174	63
3	24	M	165	58
4	27	M	183	85
5	25	M	180	90
6	24	M	186	78
7	25	F	159	57
8	24	M	180	65
9	28	F	165	59
10	23	M	178	75

### 3 Method

Users' gait is observed in terms of acceleration and angular velocity of feet, and distance between feet. Data are segmented into non-overlapping windows and pre-processed. Then, from each window, a set of features is extracted. A one-class classifier is trained using only the data originated from the legitimate owner. The trained system is evaluated against previously unseen users (the possible impostors). An overview of the proposed method is depicted in Fig. 1.

#### 3.1 Data Acquisition and Pre-processing

Data were collected from ten volunteers, two females and eight males, having the physical characteristics shown in Table 1. The equipment consisted of two Shimmer3 IMU devices - used to acquire inertial data from each foot - and two devices from the DecaWave TREK1000 kit - used to acquire the distance

**Table 2.** Configuration parameters of the Shimmer devices.

Sampling rate	102.4 Hz
Accelerometer range	$\pm 8$ g
Gyroscope range	$\pm 500$ dps

**Table 3.** Configuration parameters of the TREK1000 devices.

Sampling rate	10 Hz
Data rate	6.8 Mbps
Power source	Tag: battery powered Anchor: connected to portable PC via USB

between feet. Shimmer devices were configured to collect acceleration and angular velocity according to the parameters shown in Table 2. The four devices were attached to volunteers’ feet and ankles as shown in Fig. 2. The orientation of the axes of the accelerometer and of the gyroscope, with respect to the device case, are also shown in Fig. 2. TREK1000 devices are equipped with a transceiver compatible with the IEEE 802.15.4-2011 UWB standard. Each device is able to estimate the distance towards other devices by using a technique based on two-way ranging time-of-arrival. One of the devices from the DecaWave TREK1000 kit was configured as an anchor (right foot) and the other one as a tag (left foot). Table 3 provides the other operational parameters.

Volunteers were asked to walk for five minutes keeping their normal pace. For each volunteer, thirteen signals were collected: the acceleration and the angular velocity along the three axes for each foot, and the distance between feet. An example of such signals - acceleration and gyroscope for just one foot and the distance between feet - is shown in Fig. 3.

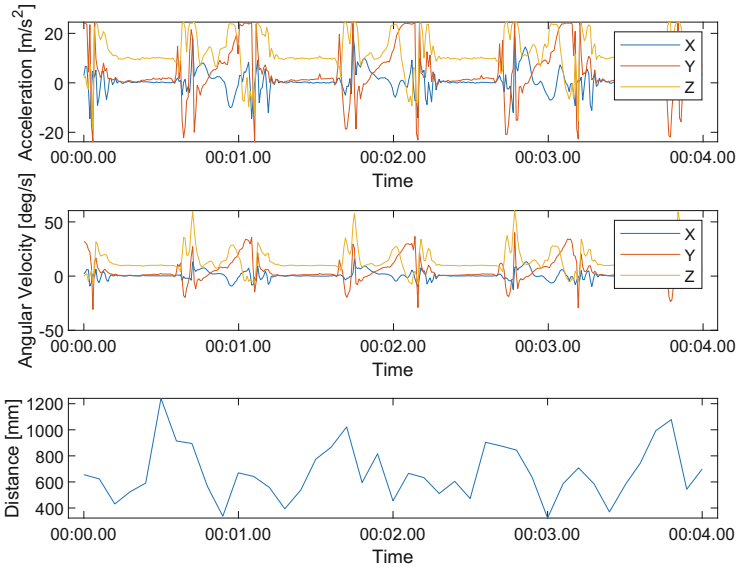
The dataset is available at:

<http://vecchio.iet.unipi.it/vecchio/data/>.

The traces produced by the Shimmer and TREK devices were first synchronized and trimmed. This last step was carried out to remove non-walking data at the beginning and at the end of each trace. Signals were then filtered by applying a low-pass Butterworth filter with a cut-off frequency of 15 Hz for inertial signals and 4.9 Hz for the distance signal. In the end, 4 min and 30 s of clean, filtered walking data were available for each user.

### 3.2 Feature Extraction

Traces were divided into fixed-duration windows. For each window, a set of features commonly used in similar domains was extracted from all the thirteen signals. The set of features is: *mean*, *standard deviation*, *max-min*, *median absolute deviation*, *average absolute variation* [1], and *mean crossing rate*.



**Fig. 3.** Acceleration, angular velocity, and distance data when walking.

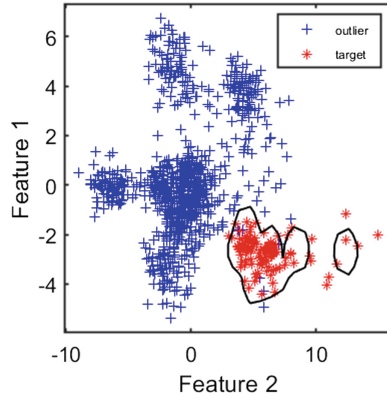
A vector containing 78 features is then produced for each window (the above-indicated six features computed on the 13 signals). Feature vectors are used to train the one-class classifier and for its evaluation, as described later. Figure 4 shows the gait samples of one of the volunteers in the feature space, restricted to two dimensions to make the image readable, against the other volunteers. The user's model produced by a trained OCC method is represented too. The training phase was performed by using the samples coming from the examined user only (the samples of the genuine/target user); the samples collected from the other users (i.e. the impostors' samples) were added to the scatterplot only later.

## 4 Results

The performance of the proposed approach was evaluated by training an OCC method using a portion of the data of one of the volunteers and then testing the trained system on previously unseen data (produced by the same user, to test the capability of the model to recognize the legitimate owner, and produced by other volunteers, to test the capability of the model to reject the possible impostors).

### 4.1 Impact of Window Size on Authentication Accuracy

As previously mentioned, gait data were divided into fixed-duration windows. To understand the influence of the duration of windows on authentication results,

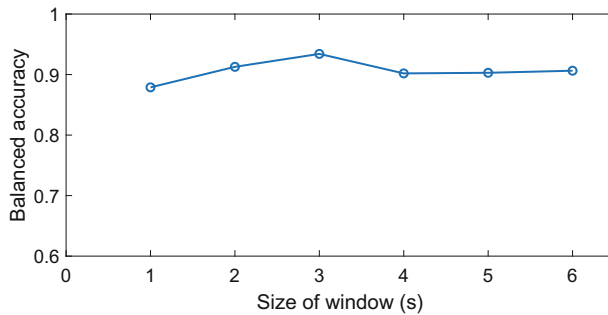


**Fig. 4.** A user’s model in a two-dimensional space, and gait instances of other users.

window sizes in the range from 1 s to 6 s were evaluated. Figure 5 shows the Balanced Accuracy (BA) - a performance index which works well with imbalanced data [3] - when varying the window size. In particular, the BA values depicted in Fig. 5 are the average of the BA values obtained when using five different OCC methods, operating according to different principles (to better understand how this parameter impacts the classification accuracy in general, i.e. not relatively to a single OCC method). The five considered methods are the OCC version of the following classifiers: Gaussian, Minimum Spanning Tree (MST), k-Means, k-Nearest Neighbors (kNN), and Auto Encoder [22]. The Gaussian OCC method models the target class according to a Gaussian distribution; in MST, the distance from a minimum spanning tree derived from the training instances is used as an indicator of distance from the class to be recognized; in k-Means, the class to be recognized is modeled as  $k$  clusters; in kNN, the distance from the  $k$  nearest neighbors is used to classify new instances; Auto Encoders are neural networks trained to reconstruct the input at the output, then the difference between the input and the output is used to identify the target class.

The performance was evaluated under the assumption that the gait model of the legitimate user is learned automatically during the initial 2 min of walk. The remaining part of the user’s trace (2.5 min) was used for testing the performance of the trained classifier on previously unseen data. The capability of the trained classifier to reject impostors was evaluated using the same amount of data (2.5 min) extracted from every other user’s trace. Finally, results were averaged across all users.

The best result is obtained when using windows with a duration of 3 s. Therefore, this value is used for computing the results presented in the next sections.



**Fig. 5.** Balanced accuracy when varying the windows size.

**Table 4.** Average balanced accuracy of five OCC methods when varying the number of dimensions.

# features	BA
5	90%
10	93%
15	92%
20	88%

## 4.2 Reduction of the Feature Space

As mentioned, the number of features used to describe the gait style of users is 78. The feature space was reduced to a smaller number of dimensions in order to avoid overfitting and improve the efficiency of the classification model. Reduction of dimensionality was obtained using Principal Component Analysis (PCA). This method maps the entire features space in a new space that has a smaller number of dimensions.

Table 4 shows how the BA varies when the number of dimensions of the PCA space is changed. The BA value is the average of the BA obtained by the same five OCC methods. The best results are obtained when setting the PCA space to 10–15 dimensions.

## 4.3 Techniques for One-Class Classification

The accuracy that can be achieved by some popular OCC techniques was evaluated. In particular, besides the already introduced five OCC methods, also the following classifiers were assessed: MCD Gaussian, where a minimum covariance determinant density is fit onto the data; Naive Parzen, where Gaussian kernels are centered on training instances and used for estimating the probability density [22].

For each classifier, the evaluation was carried out again using 2 min of walking for training and 2.5 min for the evaluation, with a preliminary reduction to a



**Table 5.** Balanced accuracy obtained with different OCC methods.

Method	BA (% , <i>mean</i> $\pm$ <i>std.dev.</i> )	FPR (%)	FNR (%)
Gaussian	93.0 $\pm$ 4.6	$\sim$ 0	14.0
MCD Gaussian	94.1 $\pm$ 3.6	5.2	6.6
k-NN	90.1 $\pm$ 9.2	16.9	3.0
MST	97.0 $\pm$ 1.4	2.4	3.6
k-Means	92.3 $\pm$ 5.7	9.0	6.4
Naive Parzen	92.4 $\pm$ 3.4	6.3	9.0
Auto Encoder	93.8 $\pm$ 2.8	0.4	12.0

**Table 6.** Balanced accuracy obtained by MST when using the different sources of information.

Sensor	BA
Accelerometer	92.8%
Gyroscope	89.7%
UWB	77.8%

feature space with 10 and 15 dimensions. In addition to BA, the OCC methods were evaluated also in terms of False Positive Rate (FPR) - an impostor being incorrectly classified as the legitimate user - and False Negative Rate (FNR) - the legitimate user being incorrectly classified as an impostor.

Table 5 reports the BA values obtained by averaging the results across all the users in the dataset. For every user, the remaining ones were used as possible impostors. Only the best BA value obtained with 10 and 15 dimensions is reported.

MST is the method that provides the best results, with a BA of 97%. Also, the standard deviation of BA is small, this means that the method operates consistently across all users.

#### 4.4 Contribution of the Different Sensors

As stated in Sect. 3, data were acquired from three different typologies of sensors: accelerometers (both feet), gyroscopes (both feet), and UWB transceivers (to estimate the distance between feet).

We evaluated the contribution of the different information sources to the authentication process. To this purpose, the best OCC method found in the previous section - MST - was evaluated again on data originated from a single information source at a time. PCA feature selection was applied to the set of accelerometric features to reduce the number of dimensions from 36 to 10. The same was done to the features extracted from the data produced by the gyroscope. For UWB, all 6 features were used. Results are shown in Table 6.

A BA of 92.8% can be achieved - by an MST classifier - when using the data produced by the accelerometers only. The BA that can be obtained when using information produced by the gyroscopes is relatively close. Distance collected via UWB seems to be less useful as, in the absence of the two other information sources, is able to reach a BA value of 77.8%. However, it is important to note that distance is collected at 10 Hz (the maximum frequency allowed by the adopted hardware solution), whereas acceleration and angular velocity are collected at a much higher rate (102.4 Hz). It is thus possible that the limited sampling rate is unable to capture all the details of an individual's gait style.

## 5 Conclusion

Smart-shoes, which are increasingly used by common users, can be extremely useful to achieve passive, effortless authentication. Experimental results show that a balanced accuracy as high as 97% can be reached when using IMUs and UWB transceivers as sources of information and adopting a one-class classification approach.

It is important to note that, differently from most of the existing literature on authentication based on smart-shoes, the training phase of the system has been carried out using only the data of the legitimate user. This makes the training phase simpler, as there is no need for other users' data to create a model of the possible impostors.

Of the three considered sensors, accelerometers proved to be the most useful information sources for authentication purposes. However, the contribution of gyroscopes and UWB transceivers is not negligible, as they make possible to increase the balanced accuracy from 92.8% to 97%. This highlights the benefits achieved by approaches based on sensor fusion.

Our study considered only users walking at a normal pace, but it would be interesting to evaluate the performance of the proposed technique when varying the walking speed or in case of changes in the physical condition of the user (for example after intense fatigue).

**Acknowledgment.** This work was partially funded by the Italian Ministry of Education and Research (MIUR) in the framework of the CrossLab project (Departments of Excellence).

## References

1. Abbate, S., Avvenuti, M., Cola, G., Corsini, P., Light, J., Vecchio, A.: Recognition of false alarms in fall detection systems. In: Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC), pp. 23–28, January 2011. <https://doi.org/10.1109/CCNC.2011.5766464>
2. Avila, L., Bailey, M.: The wearable revolution. *IEEE Comput. Graph. Appl.* **35**(2), 104–104 (2015). <https://doi.org/10.1109/MCG.2015.44>

3. Brodersen, K.H., Ong, C.S., Stephan, K.E., Buhmann, J.M.: The balanced accuracy and its posterior distribution. In: Proceedings of the 20th International Conference on Pattern Recognition, pp. 3121–3124. IEEE (2010)
4. Cola, G., Avvenuti, M., Musso, F., Vecchio, A.: Gait-based authentication using a wrist-worn device. In: Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, MOBIQUITOUS 2016, pp. 208–217. ACM, New York (2016). <https://doi.org/10.1145/2994374.2994393>
5. Cola, G., Avvenuti, M., Vecchio, A.: Real-time identification using gait pattern analysis on a standalone wearable accelerometer. *Comput. J.* **60**(8), 1173–1186 (2017). <https://doi.org/10.1093/comjnl/bxw111>
6. Derawi, M.O., Nickel, C., Bours, P., Busch, C.: Unobtrusive user-authentication on mobile phones using biometric gait recognition. In: Proceedings of the Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 306–311, October 2010. <https://doi.org/10.1109/IIHMSP.2010.83>
7. Eskofier, B.M., et al.: An overview of smart shoes in the Internet of health things: gait and mobility assessment in health promotion and disease monitoring. *Appl. Sci.* **7**(10) (2017). <https://doi.org/10.3390/app7100986>, <http://www.mdpi.com/2076-3417/7/10/986>
8. Gafurov, D., Sneekenes, E., Bours, P.: Gait authentication and identification using wearable accelerometer sensor. In: Proceedings of the IEEE Workshop on Automatic Identification Advanced Technologies, pp. 220–225, June 2007. <https://doi.org/10.1109/AUTOID.2007.380623>
9. Howell, A.M., Kobayashi, T., Hayes, H.A., Foreman, K.B., Bamberg, S.J.M.: Kinetic gait analysis using a low-cost insole. *IEEE Trans. Biomed. Eng.* **60**(12), 3284–3290 (2013). <https://doi.org/10.1109/TBME.2013.2250972>
10. Johnston, A.H., Weiss, G.M.: Smartwatch-based biometric gait recognition. In: Proceedings of the IEEE International Conference on Biometrics Theory, Applications and Systems (BTAS), pp. 1–6, September 2015. <https://doi.org/10.1109/BTAS.2015.7358794>
11. Kim, J., Lee, K.B., Hong, S.G.: Random forest based-biometric identification using smart shoes. In: Proceedings of the Eleventh International Conference on Sensing Technology (ICST), pp. 1–4. IEEE (2017)
12. Kumar, R., Kundu, P.P., Phoha, V.V.: Continuous authentication using one-class classifiers and their fusion. In: Proceedings of the IEEE International Conference on Identity, Security, and Behavior Analysis (ISBA), pp. 1–8. IEEE (2018)
13. Muaaz, M., Mayrhofer, R.: Smartphone-based gait recognition: from authentication to imitation. *IEEE Trans. Mob. Comput.* **16**(11), 3209–3221 (2017). <https://doi.org/10.1109/TMC.2017.2686855>
14. Ngo, T.T., Makihara, Y., Nagahara, H., Mukaigawa, Y., Yagi, Y.: The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication. *Pattern Recogn.* **47**(1), 228–237 (2014)
15. Nickel, C., Busch, C.: Classifying accelerometer data via Hidden Markov Models to authenticate people by the way they walk. *IEEE Aerosp. Electron. Syst. Mag.* **28**(10), 29–35 (2013). <https://doi.org/10.1109/MAES.2013.6642829>
16. Nickel, C., Wirtl, T., Busch, C.: Authentication of smartphone users based on the way they walk using k-NN algorithm. In: Proceedings of the Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 16–20, July 2012. <https://doi.org/10.1109/IIH-MSP.2012.11>

17. Oak, R.: A literature survey on authentication using behavioural biometric techniques. In: Bhalla, S., Bhateja, V., Chandavale, A.A., Hiwale, A.S., Satapathy, S.C. (eds.) *Intelligent Computing and Information and Communication*. AISC, vol. 673, pp. 173–181. Springer, Singapore (2018). [https://doi.org/10.1007/978-981-10-7245-1\\_18](https://doi.org/10.1007/978-981-10-7245-1_18)
18. Ramirez-Bautista, J.A., Huerta-Ruelas, J.A., Chaparro-Cárdenas, S.L., Hernández-Zavala, A.: A review in detection and monitoring gait disorders using in-shoe plantar measurement systems. *IEEE Rev. Biomed. Eng.* **10**, 299–309 (2017). <https://doi.org/10.1109/RBME.2017.2747402>
19. Schürmann, D., Brüsch, A., Sigg, S., Wolf, L.: BANDANA - body area network device-to-device authentication using natural gait. In: *Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 190–196. IEEE (2017)
20. Seneviratne, S., et al.: A survey of wearable devices and challenges. *IEEE Commun. Surv. Tutorials* **19**(4), 2573–2620 (2017). <https://doi.org/10.1109/COMST.2017.2731979>
21. Sun, F., Mao, C., Fan, X., Li, Y.: Accelerometer-based speed-adaptive gait authentication method for wearable iot devices. *IEEE Internet Things J.* **6**(1), 820–830 (2019)
22. Tax, D.: *DDtools, the Data Description Toolbox for Matlab*, January 2018, version 2.1.3
23. Vecchio, A., Cola, G.: Fall detection using ultra-wideband positioning. In: *2016 IEEE Sensors*, pp. 1–3, October 2016. <https://doi.org/10.1109/ICSENS.2016.7808527>
24. Vecchio, A., Mulas, F., Cola, G.: Posture recognition using the interdistances between wearable devices. *IEEE Sens. Lett.* **1**(4), 1–4 (2017). <https://doi.org/10.1109/LENS.2017.2726759>
25. Vecchio, A., Cola, G.: A method based on UWB for user identification during gait periods. *Healthcare Technol. Lett.* (2019). <https://digital-library.theiet.org/content/journals/10.1049/htl.2018.5050>