Security Properties of Gait for Mobile Device Pairing

Arne Brüsch, Ngu Nguyen, *Member, IEEE,* Dominik Schürmann, *Member, IEEE,* Stephan Sigg, *Member, IEEE,* and Lars Wolf, *Member, IEEE*

Abstract—Gait has been proposed as a feature for mobile device pairing across arbitrary positions on the human body. Results indicate that the correlation in gait-based features across different body locations is sufficient to establish secure device pairing. However, the population size of the studies is limited and powerful attackers with e.g. capability of video recording are not considered. We present a concise discussion of security properties of gait-based pairing schemes including quantization, classification and analysis of attack surfaces, of statistical properties of generated sequences, an entropy analysis, as well as possible threats and security weaknesses. For one of the schemes considered, we present modifications to fix an identified security flaw. As a general limitation of gait-based authentication or pairing systems, we further demonstrate that an adversary with video support can create key sequences that are sufficiently close to on-body generated acceleration sequences to breach gait-based security mechanisms.

1 Introduction

With the proliferation of mobile devices and the upcoming Internet of Things, interaction between these devices will drastically increase [1]. In particular, smart textile and digital assistants are to generate dense body area networks [2]. This is extended by spontaneous pairings to other devices in the context of use [3]. In such environment where device pairings raise by n with each $n+1$st additional device, and device count and type changes on a sub-day schedule, manual pairing is impractical. Implicit pairing has been proposed e.g. based on acceleration [4], audio [5], magnetometer [6] and RF features [7]. Especially gait [8] is well suited in wearable settings as it is confined to a single person and can be read out at arbitrary body location [9].

Gait-based pairing schemes [10] (extended in [11]), [12], [13], [14] (extended in [15]), [16] share the common goal to protect against Man-in-the-Middle (MitM) attacks, where an attacker actively places herself between devices to modify intercepted communication. In contrast to other schemes such as Bluetooth Secure Simple Pairing (SSP) that typically requires the comparison of PINs, gait is leveraged for automatic MitM protection. However, no concise study of the security properties of quantization approaches for gait-based pairing has been presented to-date.

We close this gap by providing a comprehensive classification of attack surfaces for gait-based pairing and authentication. We analyse four recent quantization schemes covering protocol-specific attacks and potential security weaknesses, as well as distribution, statistical and entropy analysis of key sequences. Finally, we show that a sophisticated adversary using video can break gait-based pairing if executed in real-time. Our contributions are

- a concise investigation and comparison of popular quantization schemes for gait-based device-pairing,
- a comprehensive discussion of attack surfaces,
- an entropy, pattern and statistical analysis,
- an improved quantization for one of the approaches to mitigate an identified security weakness,
- the first ever empirical demonstration that video poses a significant threat to gait-based security.

We first introduce technical details of the four quantization schemes and their performance on a common dataset (Section 3). Then, to spot conceptual flaws, we analyse properties of generated keys in terms of bit distribution and statistical tests in Section 4, before identifying scheme-specific security weaknesses (Section 5). In particular, we consider the one-shot success probability, quantization-specific attacks, effects of error correction on security properties, gait mimicry, as well as, impersonation via video recording. In Section 6, we suggest improvements for specific schemes. Our work concludes in Section 7 with the main quantization differences and discusses the most promising scheme.

2 Related Work

We first discuss gait recognition approaches, before summarizing recent progress in gait-based authentication and gait-based pairing. In the remainder of the discussion, we then focus on using acceleration sequences from natural gait for device pairing.

2.1 Gait Recognition

Traditionally, gait recognition has been applied exploiting machine vision [17], [18], [19], [20]. Systems then comprise one or multiple cameras to capture natural gait and contain image recognition steps including background subtraction, feature extraction and recognition [21]. First work goes back to perception experiments on light point displays conducted in [22]. This work was further developed in [23] with

- *A. Brüsch, D. Schürmann and L. Wolf are with TU Braunschweig, Germany.*
- *N. Nguyen and S. Sigg are with Aalto University, Finnland.*

All authors contributed equally to this work and are listed in alphabetic order. *Manuscript received October 20, 2017, revised June 24, 2018.*

computer vision approaches to recognize people from gait. In preceding years drastic improvements have been made in gait recognition algorithms [24], [25]. Gait recognition approaches can be grouped into (1) temporal alignment-based, (2) static parameter-based and (3) silhouette shape-based approaches [19]. From these, [26] found that shape is more significant for person identification than kinematics.

Temporal alignment-based approaches emphasize both shape and dynamics and first extract silhouette features before aligning sequences of these e.g. with temporal correlation, dynamic time warping or hidden Markov models.

Static parameter-based approaches exploit gait dynamics such as stride length, cadence and stride speed [27]. However, they are least successful for gait-based identification due to their need for 3D calibration information.

Finally, silhouette shape-based approaches use silhouette shape similarity and disregard temporal information, often considering averaged silhouettes or treating silhouette shapes as collection without specific order [19]. For all above methods, gait recognition can be improved by combining statistical gait features from real and synthetic templates [18]

Due to the increasing availability of wearable sensors such as gyroscopes (rotation), accelerometers (acceleration) or force sensors (force during walking), gait recognition via such wearable sensors is increasingly investigated [28], [29], [30], [31], [32], [33], [34]. In these approaches, acceleration sequences are recorded from various body locations, most prominently at the waist. The acceleration signal is then denoised e.g. by applying wavelet transformation [33] and changes in walking speed are mitigated utilizing dynamic time warping [35] or similar approaches. Individual steps are identified from the resulting signal by searching for minima and by applying pattern or template matching [33]. Similarity can be estimated by the computation of cross-correlation [31]. Alternatively, machine learning classifiers are trained and applied [30].

Finally, a recent technique employed to acquire human gait is to monitor phase changes of an electromagnetic signal reflected from a subject walking towards a transceiver [36], [37]. The authors exploit changes in channel state information (CSI) from WiFi devices for the detection of gait. After generating spectograms from CSI measurements, similar to Doppler radars, and applying autocorrelation on the torso reflection to remove imperfection in these spectrograms, gait patterns are extracted.

Note that frequently, sensors installed in the floor such as pressure sensing mats are also mentioned as modalities for gait recognition [38], [39]. However, in these cases, not gait itself is extracted but other features such as footprints [39], ground reaction force [40] or heel-to-toe ratio [41].

2.2 Gait as a Biometric Pattern for Authentication

Biometric authentication systems comprise sensors converting analog stimuli to digital input that can then be quantized and compared to a database of previously stored biometric features. Gait as a discriminating feature was first studied in [22], [42]. It has been realized that characteristic features in gait enable identification of subjects also in larger gait databases [43], [44], [45], [46]. In addition, multiple studies have demonstrated that the success probability of an imposter trying to mimic a subjects gait are low [47] even when trained professionals with similar physical characteristics are employed [8]. For instance, Hoang et al. [48] generated a key fingerprint from the difference of a mean gait spanning the complete population to the individual's mean gait. In this way, the authors assured that the resulting sequence is well balanced and uniformly distributed. A good overview on gait-based user authentication is provided in [49], [50].

However, despite studies asserting that gait can be used as biometric feature [32], [51], [52], we remark that there is a lack of studies investigating the security features and entropy of gait as an authentication mechanism.

Several attacks though have been considered (cf. Table 1). For instance, Mjaaland et al [55] trained seven individuals to imitate one specific victim. Even after intensive training over two weeks (5 hours every day), it was not possible for the subjects to accurately imitate the walking pattern of the victim. Also, the provision of continuous visual feedback did not suffice to assist imitators in [56]. Furthermore, the authors of [47] investigated the success probability of an attacker towards a particular subject on a database of 100 subjects and concluded that it is unlikely for an adversary to mimic the subjects gait with sufficient accuracy. This result has been confirmed by [8] who employed professional actors to mimic the gait of 15 subjects with close physical properties. Indeed, the attempt to mimic gait incorporates the risk of asymmetric gait cycles and thus even lowers the chance of success. However, as indicated in [47], the probability of random matches significantly exceeds the expected probability in the birthday paradox.

This means that an attacker with knowledge of the template database can select persons that are close to him in terms of similarity as suitable victims. This poses a serious threat to gait-based authentication in general. This is confirmed in [49], [57] who report an equal error rate (EER)[1] of 20% for gait authentication. In addition, given the gait features of the victim and exploiting a treadmill to control speed, length of steps, thigh lift, hip movement and width of steps, the authors in [53] could reach a false acceptance rate (FAR) of 46.66%.

In addition, the high accuracy of video-based gait recognition systems also empowers an adversary to generate a database of gait information on multiple subjects unnoticed. Video-based attacks on gait-authentication systems are insufficiently investigated in the literature. In Section 5.5,

TABLE 1: Attacks on gait-based wearable authentication systems

Paper	Applications	Attacking
Muaaz et al. [8]	Gait recognition	Active imposter (imitation), 20% EER
Xu et al. [14]	Device pairing	Active imposter (imitation), passive imposter, MitM
Kumar et al. [53]	Gait recognition	Treadmill attack
Trippel et al. [54]	Injection of false acceleration	Poisoning acoustic injection attack
Derawi et al. [49]		Active imposter, 20% EER, significant random success probability
Mjaland et al. [55]	Gait biometrics	Active long-term trained impostors
Stang [56]	Gait biometrics	Training impostors with continuous visual feedback

1. Equal rates for false acceptance and false rejection

ShakeUnlock protocol
1) Record acceleration sequences
2) Remove gravity per axis, calculate magnitude and normalize to $[-1, 1]$
3) Share magnitude via secure channel
4) Slice magnitude segments; transform to frequency domain
5) Compute pairwise coherence via cross spectral- & power spectral density
6) Calculate the mean over all coherence values
7) Unlock IFF mean coherence exceeds threshold

Candidate Key protocol (SAPHE)
1) Extract features on devices
2) Hash feature values
3) Exchange hashes to identify matching values
4) When sufficient entropy collected (matching values), concatenate matching values to give secure key.

Walkie-Talkie protocol
1) Agree on heel-strike count. Then, record acceleration.
2) Use ICA for source separation; apply FFT on independent components
3) Low-pass filter (3Hz) in gravity direction (reduce noise and detect local maxima (heel-strikes))
4) Rotate acceleration data using gyroscope to same body coordinate system
5) Low pass filter (10Hz); normalize 3D acceleration to zero mean, unit length
6) Samples $\gtrless \mu + \alpha\sigma$ are interpreted as $1/0$ where μ and σ are computed per window
7) Matching samples chosen define key. IFF $\leq 0.5 + \varepsilon$ overlap, abort (counter impersonation)
8) XOR sequences between consecutive windows to obtain keys, axes are interleaved.

BANDANA protocol
1) Collect acceleration readings from the z-axis
2) Correct rotation wrt gravity (gyroscope)
3) Bandpass between 0.5Hz and 12Hz
4) Resampling (40 samples/gait) and gait detection
5) Compute mean gait
6) Difference between mean and instantaneous gait translates to binary sequence
7) Calculate reliability of bits, disregard least reliable
8) Share reliability ordering & create fingerprint
9) Fuzzy cryptography: Get key from fingerprint

Inter-Pulse-Interval (IPI) protocol
(1-4) Analog to the BANDANA Protocol
5) Detect left/right-foot-flat peaks from acceleration
6) $\overline{IPI}_{gray} = \text{Graycode}\left(\left\lfloor \frac{IPI}{m \cdot 1000/f_s} \mod 2^q \right\rfloor\right)$
7) Obtain key as first k bits in $\overline{IPI}_{gray}$

Fig. 1: Description of acceleration-based device-pairing protocols

we demonstrate that a sophisticated adversary with video support can estimate gait sufficiently accurate in order to break gait-based authentication and pairing schemes.

We conclude that gait-based authentication faces serious security threats and gait appears not feasible as sole basis for authentication, especially in systems where the adversary is targeting not a specific but any subject in the system. Furthermore, gait changes over time [21] and is affected by clothing, footwear, walking surface [17], walking speed [21] and emotion [58]. These effects are insufficiently studied and render gait-based authentication a challenging undertaking.

2.3 Acceleration-Based Pairing of Devices

Device pairing protocols execute quantization on one or more devices at the same time to generate similar bit sequences. In contrast to user authentication, these sequences are not matched against a template database. Instead they are used to authenticate a key agreement between all particpanting parties. Recently, several authors have considered acceleration or gait for the pairing of devices co-present on the same body [57], [59], [60]. In particular, these approaches exploit correlation in acceleration signals when devices are worn on the same body [61], [62] or shaken together [4], [63]. Note that, in contrast to exploiting gait for authentication, the existence of a unique and reproducible biometric gait sequence is not required for these approaches. Instead, the protocols exploit instantaneous, correlated acceleration sequences that can not be re-used at different time as the system can be restricted to single attempts [10]. The above described weaknesses for gait as biometric pattern therefore do not apply. Instead, the strength of the pairing approach is conditioned on the quantization used, what entropy that approach can guarantee and whether or not it leaks information to a powerful (realistic) attacker.

In [4], [63] the ShakeUnlock protocol is presented to unlock a mobile device when it is shaken simultaneously with a smartwatch. The individual steps of this protocol are briefly described in Figure 1. This approach requires the direct comparison of acceleration sequences in order to compute correlation and therefore needs an established secure channel to exchange this information.

However, other approaches that do not require an already established secure connection have been proposed recently. For authentication based on arbitrary co-aligned sensor data, Mayrhofer [64] proposes the candidate key protocol. An advanced variant that solves the known issue of low-entropy input data is implemented in SAPHE [13]. It interactively exchanges hashes from feature sequences as short secrets and concatenates the key from the secrets with matching hashes (cf. Figure 1).

Walkie-Talkie, an alternative approach conditioned on correlated acceleration sequences from a person's gait, is presented in [14]. The authors achieve a high bitrate by using individual samples for key bits if they deviate by at least α standard deviations from the mean (cf. Figure 1).

An extended version has been published as Gait-Key [15] providing a higher bitrate by applying multiple thresholds. Walkie-Talkie and Gait-Key use the acceleration values along gravity, walking and sideways direction. In another scheme by the same authors, movement on all three axis is used as a random source to generate a group key [16]. This group key is locked in a fuzzy vault using a secret set based on the acceleration along gravity only (cf. Walkie-Talkie). Other wearables can unlock the vault using a secret set, sufficiently similar to the original one, to retrieve the group key. When an attacker intercepts the locked fuzzy vault the security of this scheme solely depends on the secret set.

The BANDANA protocol [10] exploits acceleration along the z-axis only and conditions the gait fingerprint on the difference between instantaneous gait and mean gait at that body location. It thereby achieves normalization among acceleration sequences across body locations. Remaining

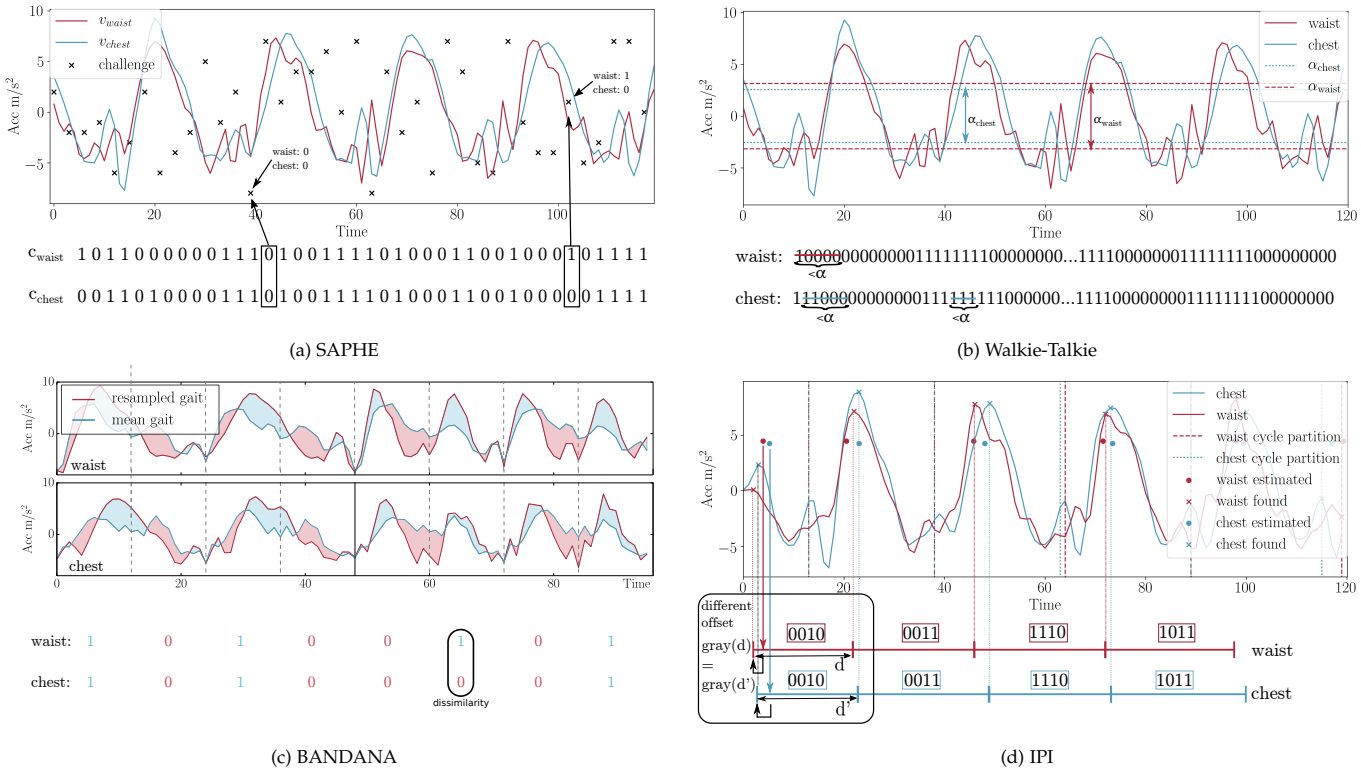


Fig. 2: Descriptive examples for the evaluated quantization schemes

dissimilarities in fingerprints are corrected with fuzzy cryptography exploiting BCH codes (cf. Figure 1).

In an extended version [11], the required key length has been reduced to 16 bit by using a Password Authenticated Key Exchange (PAKE).

Recently, the Inter-Pulse-Interval (IPI) between consecutive steps has been exploited for secure key generation from gait [12]. The protocol exploits the acceleration along the z-axis and concatenates the key sequence as gray-coded, scaled and rounded IPIs. As reported in [12] (cf. Table 2 in Section 5.1) the security and inter-class similarity depends on the speed of consecutive steps and steplength. The protocol was verified on gait captured from devices on the torso of subjects (lower back, upper right arm and right ear).

The quantization methods in these approaches diverge and result in different properties of the generated binary fingerprints, as described in Section 3.

An attack on acceleration-based pairing is described in [54]. An active adversary emitting modulated acoustic interference at the resonant frequency of materials in MEMS sensors can control or modify measured acceleration, and thus inject changes to acceleration sequences.

3 Comparison of Quantization Schemes

A crucial part in gait-based pairing is the quantization used. It has to preserve a *high similarity* between generated keys on different body parts, and generate *sufficiently unpredictable* bit sequences for the use as cryptographic keys that withstand a computationally unconstrained adversary.

In this section, we analyze the quantization of SAPHE, Walkie-Talkie, BANDANA and IPI and describe their working principles along Figure 2. In particular, we study the similarity of keys generated for pairs of devices on different body locations. Additionally, we evaluate how they fulfill the first requirement, i.e. to generate keys with *high similarity* between different locations on the same body (intra-body) and no similarity between different bodies (inter-body).

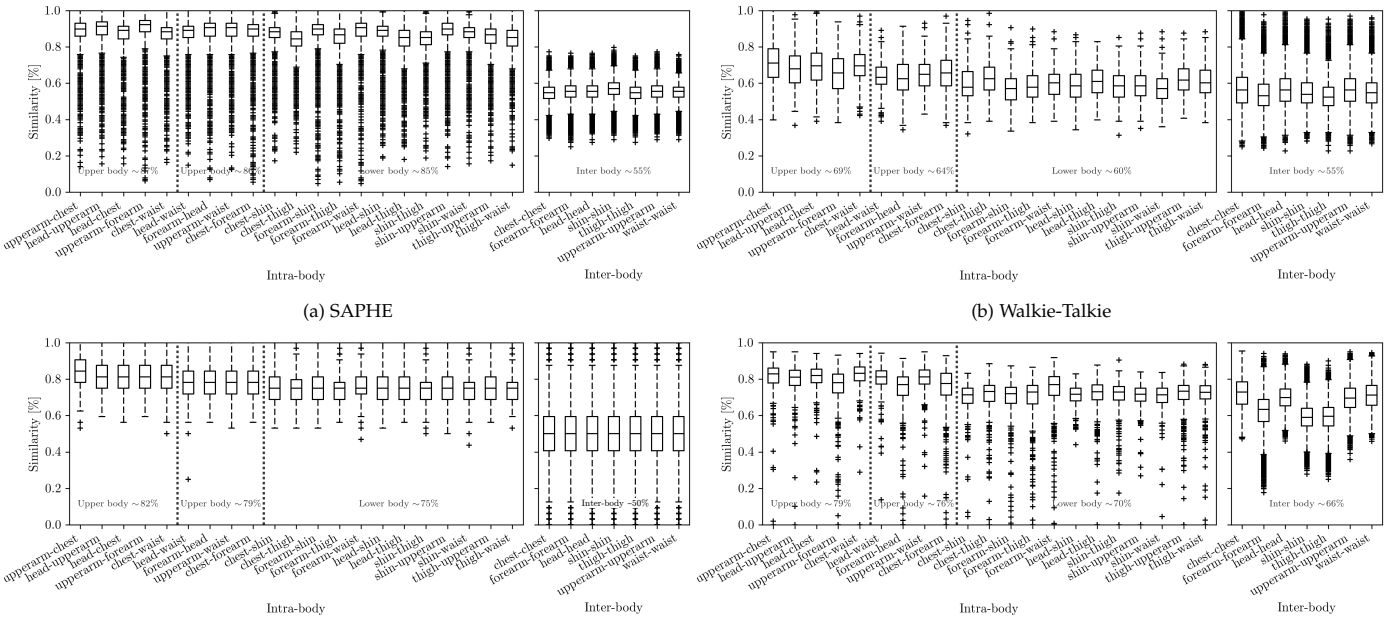
All quantization schemes have been analyzed using walking data recorded in [65][2], pre-processed by Madgwick's algorithm to correct accelerometer orientation. Each quantization scheme generated keys from same length walking data. Due to the protocols' different efficiency, key length may vary across schemes. The performance of the schemes to withstand adversaries is discussed in Section 4.

3.1 SAPHE

In the SAPHE [13] protocol, after generating and exchanging the hash $H(r_A)$ ($H(r_B)$) of the random seed r_A (r_B) to compute threshold values $\bar{t}_A$ ($\bar{t}_B$), as points in an Acceleration-time coordinate system $\mathbb{K}$, devices derive acceleration sequences $\overline{v}_A$ ($\overline{v}_B$) in $\mathbb{K}$. Challenges c_A (c_B) that describe whether $\bar{t}_A$ ($\bar{t}_B$) exceed $\overline{v}_A$ ($\overline{v}_B$) are exchanged together with r_A (r_B). The protocol does not disclose information on the acceleration during this communication.

We remark though, that the authors propose a second version which leaks information on the acceleration since, in addition, a distance ordering $\overline{o}_A$ ($\overline{o}_B$) between $\bar{t}_A$ ($\bar{t}_B$) and $\overline{v}_A$ ($\overline{v}_B$) is exchanged. The purpose of this distance ordering is to guard against a specific attack on the hash function (described in [13]). However, an adversary could exploit that

2. The dataset includes 15 subjects, 10 minutes walking each, acceleration sensors at 7 different body locations (50Hz) and is available at http://sensor.informatik.uni-mannheim.de



(a) SAPHE

(b) Walkie-Talkie

(c) BANDANA

(d) IPI

Fig. 3: Comparison of intra-body against inter-body similarity for the evaluated quantization schemes. Each value in the *intra-body* boxplot is defined by the similarity of two *different* sensor locations on the same subject (all possible combinations within each subject). For *inter-body*, each boxplot defines a different sensor location. Only *different* subjects are tested against each other with the *same* sensor location.

the threshold points $\overline{t}_A$ ($\overline{t}_B$) with small distance to $\overline{v}_A$ ($\overline{v}_B$) are good estimates of actual acceleration samples from $\overline{v}_A$ ($\overline{v}_B$). In addition, those threshold points $\overline{t}_A$ ($\overline{t}_B$) with large distance to $\overline{v}_A$ ($\overline{v}_B$) leak information on the probability of the resulting bit (0 or 1 for larger or smaller threshold).

As depicted in Figure 3a, although affected by outliers, SAPHE's generated key pairs match with high probability of 85% (lower body) to 86,87% (upper body) on average on devices worn on the same body (intra-body). The inter-body case matches on average with 55% i.e. is 10% higher than a random guess. Conclusively, SAPHE is able to generate keys that fulfill the requirement of a clear boundary between intra- and inter-body similarity.

3.2 Walkie-Talkie

The Walkie-Talkie protocol [14] is able to extract up to 1 key bit per acceleration sample.

Acceleration samples are interpreted as 0 or 1 conditioned on whether their acceleration is below or above a guard band, while samples that fall inside are ignored (cf. Figure 2b). Walkie-Talkie has also been utilized in [16] to lock a fuzzy vault containing a random key. The quantization is able to achieve higher bit rates by exploiting multiple thresholds [15]. We further discuss the impact of multiple thresholds in Section 6.2.

To mitigate hardware originated differences in acceleration strength, devices exchange and agree on samples in the acceleration sequence that shall constitute the key (*reconciliation*). The resulting sequence is thought to be biased towards alternating groups of 1-bits and 0-bits, which is addressed by applying an XOR between consecutive 30 bit long windows. We comment on this concern in 2.2.

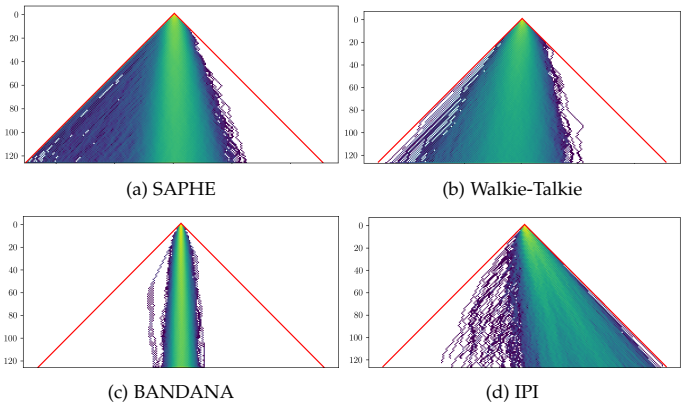
The protocol achieves 60-70% upper body bit-similarities and 55-65% for the lower body (cf. Figure 3b). This performance suggests further processing to provide reliable pairing among devices at different body location. Walkie-Talkie uses Independent Component Analysis as a preprocessing step in order to remove the arm swing (Figure 1).

In our implementation, the transformation to the body coordinate system was applied following Mohssen et al. [66]. Walkie-Talkie was then executed using the best performing parameters as mentioned in [14], such as an α of 0.8 and non-overlapping windows of size 10. We did not resample the input data as Walkie-Talkie applies a low pass filter with a cutoff frequency of 10Hz during the preprocessing. Independent Component Analysis was applied on the complete recording beforehand. We decided to only exclude arm swing components where they were clearly distinguishable.

3.3 BANDANA

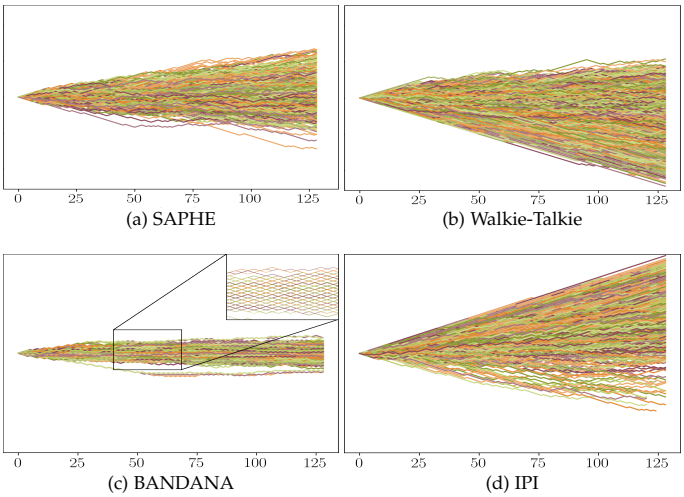
In BANDANA, key sequences are generated as a function of the difference between mean and instantaneous acceleration [11]. The approach of comparing to the mean at a particular body location serves as a normalisation procedure. The offset to the mean has a better correlation across various body locations than comparing absolute acceleration values. Furthermore, [67] argues that this approach might positively impact the distribution of bits in the key sequences towards uniformity as gait patterns are compared to their mean. To further amplify similarity of sequences generated at different body locations, bits with low difference between mean and instantaneous gait are disregarded.

The similarity between keys generated at different positions on the body is depicted in Figure 3c for the BANDANA protocol. The protocol achieves similarity results above 75% for all location-pairs and is able to render the chances of the adversary (inter-body) to random guess. The



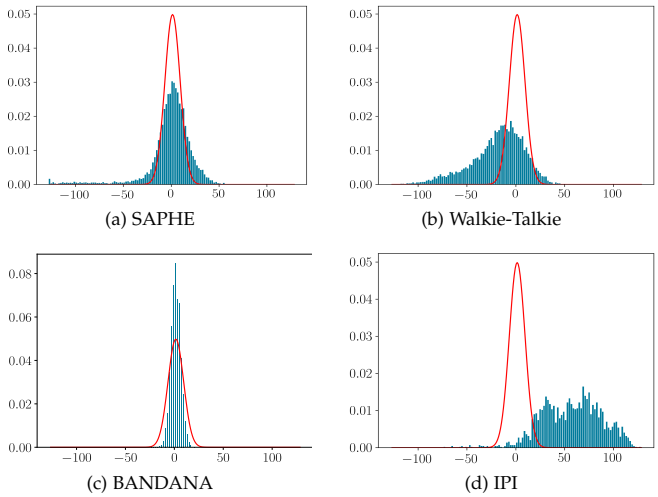
(a) SAPHE (b) Walkie-Talkie (c) BANDANA (d) IPI

Fig. 4: Heatmaps of random walks for 128 bit keys generated by the evaluated quantization schemes ($0 \rightarrow$ left; $1 \rightarrow$ right). The red lines depict the boundaries for any possible random walk.



(a) SAPHE (b) Walkie-Talkie (c) BANDANA (d) IPI

Fig. 5: Cumulative plot of random walks for 128 bit keys generated by the evaluated quantization schemes ($0 \rightarrow$ bottom; $1 \rightarrow$ top)



(a) SAPHE (b) Walkie-Talkie (c) BANDANA (d) IPI

Fig. 6: Cumulative sums distribution for 128 bit keys (distribution in the last rows in Figure 4). Expected biomial distribution in red.

protocol employs fuzzy cryptography in order to mitigate the remaining 25% of difference in the key sequences. We observe, however, a high variance for the inter-body case, which is due to a non-uniform distribution of the key sequences in the key space (cf. Section 4 and Section 5). In Section 6, we discuss how this problem can be addressed with a revised quantization approach.

3.4 IPI

The Inter-Pulse-Interval (IPI) protocol [12] exploits the random offset by which individual steps deviate from the mean gait cycle in time domain (cf. Figure 2d). The number of secret bits that can be extracted from the gait signal then depends on the sampling frequency as gait cycle estimation is more accurate with higher sampling rate. The authors report a standard deviation of 40.8 milliseconds for the IPI.

Figure 3d shows the similarity achieved for IPI between keys generated from devices located at different positions on the body. The similarity in the intra-body case is good. IPI also employs fuzzy cryptography to correct remaining bit-errors in the keys generated for devices across the same body. However, the figure also shows that the protocol does not prevent a remote adversary from paring with on-body devices, since inter-body similarities are as high as in the intra-body case. This is due to limited variation in the generated bit sequences. Inter-pulse intervals resemble a normal distribution centered around its mean. This variation around the mean is similar across subjects and the resolution employed is 4 bits only so that naturally similarity across generated bit sequences is high (cf. Section 5).

4 Randomness of Keys

In this section we investigate whether these keys are *sufficiently unpredictable* to withstand a computationally unconstrained adversary. For this, we analyze the randomness of keys and the results from the DieHarder and ENT Pseudo-random Number Sequence Tests.

4.1 Bit Distribution

To describe the randomness of keys, we compare their structure with random walks on a Galton board [68]. Plotting a sufficient amount of these sequences will eventually show a binomial distribution. Figure 4 shows heatmaps of random walks corresponding to the sequences generated by different quantization approaches. In addition, Figure 5 depicts each individual random walk such that specific patterns are observable. Based on the last row of each heatmap, Figure 6 depicts the cumulative sums distribution.

Key lengths of 128 bits are chosen for this study, which means that the acceleration sequence to generate a key varies between the different approaches.

SAPHE shows a close-to symmetric distribution centered around the mean.

The cumulative sums distribution is properly centered but shows deviations to include more '0's for a specific set of keys (cf. Figure6a). We explain this with the characteristic of acceleration readings in our data, which do not necessarily have zero-mean. With regard to the binomial distribution (depicted in red), SAPHE's key distribution is slightly stretched. Thus, while SAPHE shows good behaviour regarding similarity and usage of space in the Galton board, it carries some characteristics of the input

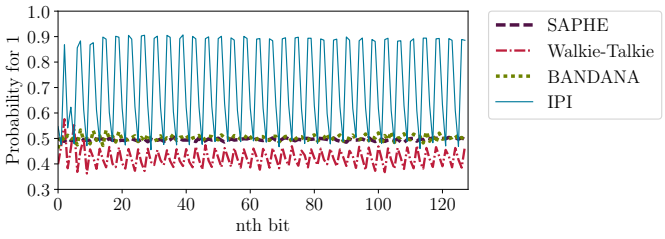


Fig. 7: Markov property: Probability of assigning 1 for the nth position in 128 bit keys

TABLE 2: Results for keys generated by the evaluated protocols after running the ENT Pseudorandom Number Sequence Test Program.

	SAPHE	Walkie-Talkie	BANDANA	IPI
Sequence size (bit)	1444864	3040848	113792	456104
Entropy (bits per bit)	0.9999	0.9855	0.9999	0.8929
Optimum compression rate	0%	1 %	0%	10%
Chi square distribution	6.91	61013.17	0.3586	65969.75
Arithmetic mean (random=.5)	0.501094	0.429175	0.5	0.690156
Monte Carlo Pi value (error)	3.122155	3.331471	3.642194	2.056830
Serial correlation coefficient (uncorrelated=0.0)	0.008204	0.055243	-0.644796	-0.002701

into the output data. Still, this does not pave the way for a successful attack. Assuming each bit position to be a state in a Markov chain, Figure 7 shows the resulting transition probabilities, aggregated over all sequences. SAPHE shows a good Markov property (cf. Figure 7).

The heatmap and distribution of Walkie-Talkie are depicted in Figure 4b and Figure 6b. The individual sequences do not show a bias (cf. Figure 5b). Walkie-Talkie, however, shows periodicity in the Markov property (cf. Figure 7). The BANDANA approach features symmetric behaviour but with low variance (cf. Figure 4c, 6c). We can observe from Figure 5c, that this weakness occurs since bit sequences consist of repetitive 'zig-zag' patterns. We discuss this problem in Section 5 and propose an improved quantization to mitigate it in Section 6. BANDANA shows a similar Markov property as SAPHE (cf. Figure 7). Finally, IPI shows good variance but a bias towards including more ones than zeros due to low variation in the quadruples generated as discussed above. IPI clearly deviates from a binomial distribution (cf. Figure 6d). We observed that consecutive 4-bit chunks repeat with a probability of 60%. This clearly shows in IPI's Markov property in Figure 7. Summarizing, while SAPHE and Walkie-Talkie exhibit reasonable randomness, BANDANA and IPI show biases in the generated keys.

4.2 Statistical Tests

To test the evaluated quantization schemes against bias in the produced random sequences, we ran the DieHarder statistical tests for each scheme. Figure 8 depicts the p-values computed from 20 runs of the DieHarder tests.

In SAPHE, the *dna* and *sts monobit* tests appear to be outliers. The *dna* test considers biases in the occurence of 10 letter words from an alphabet of 4 letters, determined by two designated bits in the sequence of random integers being tested. The *sts monobit* test counts the 1 bits in a long string of random entries and compares this to the expected number. Similar to SAPHE, Walkie-Talkie also shows a weakness in the *dna* test. In addition, the *rgb Kolmogorov-Smirnov* test falls out slightly and the *2D sphere* test features some outliers. The *kolmogorov-Smirnov* test applies a *Kuiper KS* test [69] and the *2D circle* test finds the minimum distance between pairs of randomly selected points to evaluate their randomness. BANDANA shows the most stable distribution of p-values. A slight bias might be associated with the *squeeze* test, which employs a *chi-square* test for cell frequencies on the number of multiplication with random integers that are required to reduce 2^{31} to 1. IPI shows potential weaknesses towards the *birthdays* test, the *Overlapping Quadruples Sparce Occupancy (oqso)* test, the *3D sphere* test as well as the *rgb permutation* and *rgb Kolmogorov Smirnov* test. The *rgb permutation* test counts the order of permutations of random numbers. *Birthdays* test determines the number of matching intervals from 512 'birthdays' drawn from a 24-bit 'year' while the *oqso* test, similar to the *dna* test, considers 4-letter words from an alphabet of 32 letters.

Additionally, we ran the *Ent Pseudorandom Number Sequence* Test[3]. The information density of bit sequences is computed together with reduction through optimal compression, chi square distribution, arithmetic mean of data bytes as well as serial correlation coefficient (cf. Table 2). We caution that these results are only showing the interdependence of single bits. Evaluating chunk instead of single bit interdependence, such as 4-bit chunks for BANDANA due to its 4 bit per gait cycle or 30-bit chunks for Walkie-Talkie's privacy amplification, heavily influences the test results.

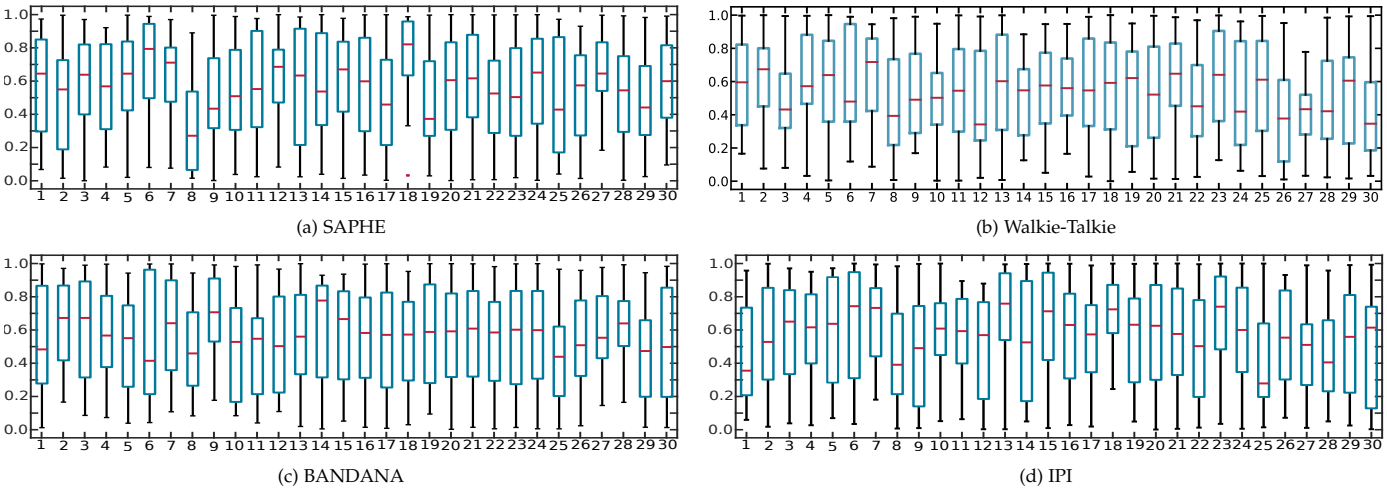
5 Security Analysis

As shown in the conceptional view in Figure 9, the pairing schemes follow a general design. Devices measure data, quantize it to bit strings after pre-processing, apply potentially error correction, and agree on a key.

Protection against MitM attacks is achieved only if all parts of a system are resilient. Our analysis follows the conceptual aproach proposed in [57], [70]. The discussed attacks are assigned to attack vectors A-G labeled in Figure 9.

An attack surface is exposed by the sensors (A). A device owner could be forced to behave in a certain way, e.g., by an adversary controlling stride speed with a treadmill. It could further be possible to bypass data acquisition (B) and reuse data from the past. With a biased quantization, a naïve brute force attack would become feasible (C). Some protocols employ communication before the actual key agreement (SAPHE: random seed and distance ordering, Walkie-Talkie: reconciliation, BANDANA: exchange reliability indices) which might potentially leak information (D). After error correction (e.g. in BANDANA and IPI), the key agreement is executed between both participants. Here, the risk of a Man-in-the-Middle (MitM) (E) or impersonation attack (G) must be considered. Finally, the key agreement could be weak or based on false assumptions, especially if it is not based on established standards (F). We do not discuss attack vector B as it assumes a compromised device, which falls outside the focus of this work.

3. http://www.fourmilab.ch/random/



(a) SAPHE (b) Walkie-Talkie

(c) BANDANA (d) IPI

Fig. 8: Distribution of p-values achieved for keys after 20 runs of the DieHarder set of statistical tests. Tests are: (1) birthdays (2) operm5 (3) rank32x32 (4) rank6x8 (5) bitstream (6) opso (7) oqso (8) dna (9) count-1s-str (10) count-1s-byt (11) parking (12) 2D circle (13) 3D sphere (14) squeeze (15) runs (16) craps (17) marsaglia (18) sts monobit (19) sts runs (20) sts serial [1-16] (21) rgb bitdistr. [1-12] (22) rgb min dist. [2-5] (23) rgb perm. [2-5] (24) rgb lagged sum [0-32] (25) rgb kstest (26) dab bytedistr. (27) dab dct (28) dab filltree (29) dab filltree 2 (30) dab monobit 2

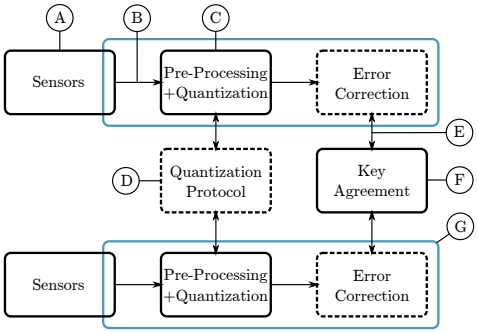


Fig. 9: Conceptual view of gait-based pairing protocols with attack vectors (blue line depicts device boundary, dashed parts are optional)

5.1 One-Shot Success Probability (E, G)

Without requiring additional knowledge about the victim's gait, an attacker may want to exhaust the keyspace $\mathcal{C}$ of all keys $\boldsymbol{k}$ to execute a MitM (E) or impersonation attack (G). However, in all discussed protocols, after each single try, a completely new authentication process (new $\boldsymbol{k}$ independent from the previous one) is started. Thus, it is impossible to exhaust $\mathcal{C}$, making this a one-shot attack. For comparison between protocols, we assume the same length of 16 bit for $\boldsymbol{k}$. The length of sequences sampled for a target key $\boldsymbol{k}$ of 16 bit may vary depending on the quantization scheme.

Note that 16 bit provide sufficient entropy since we suggest to implement a PAKE protocol as in [11], which prevents offline attacks and can thus provide a sufficiently large security margin even with short key lengths K.

5.1.1 Candidate Key Protocol Variants

The candidate key protocol is, for instance, realized in SAPHE [13], which resolves its original vunlnerability against MitM attacks. In particular, first, random challenges are chosen, as depicted in Figure 2a and committed by sharing their hashes. Afterwards, the acceleration sequence is challenged with respect to these random thresholds where an acceleration point with value lower (higher) than a threshold is interpreted as 0 (1). The success probability for a single randomly drawn key $\boldsymbol{k}$ in SAPHE is

$$\frac{1}{2^{16}} \approx 1.52588 \cdot 10^{-5} \tag{1}$$

5.1.2 Walkie-Talkie Protocol

The bits generated in the Walkie-Talkie protocol feature a high bit rate of 15–55 bits per second as reported in [14] (Figure 12(e)). However, high agreement rates are reached only for $\alpha > 0.8$ (Figure 12(d) and 12(f) in [14]), which corresponds to 15–25 bits per second. A 16 bit binary key can therefore be generated in approximately 1 second and the success probability of an adversary for a single randomly drawn $\boldsymbol{k}$ is then again $\frac{1}{2^{16}} \approx 1.52588 \cdot 10^{-5}$.

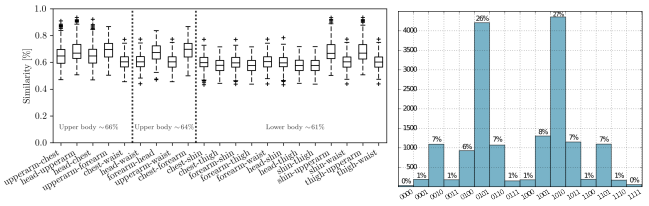
5.1.3 BANDANA Protocol

In the BANDANA protocol, $M = 48$ bit sequences are generated in about $12\,\mathrm{s}$. From each sequence, 16 bit are disregarded for reliability amplification. From the remaining 32 bit fingerprints, up to 8 bit are corrected by BCH codes, resulting in $|\boldsymbol{k}| = 16$ bit keys. The success probability of a single randomly drawn fingerprint is then (cf. Section 5.3)

$$\sum_{k=0}^{8} \binom{32}{k} / 2^{32} = \frac{\sum_{k=0}^{8} \left(\frac{32!}{(32-k)! \cdot k!} \right)}{2^{32}} \approx 0.0035 \tag{2}$$

5.1.4 IPI Protocol

In the IPI protocol, dependent on the sampling frequency, 2 to 20 secure bits are extracted from each gait cycle (cf. Table I in [12]). Depending on the sample rate of the accelerometer, the generation of 32 bits in the IPI protocol might therefore require from 2 to 16 seconds. Since the protocol also employs fuzzy cryptography for error correction, the same success probability as in the BANDANA protocol of 0.0035 applies for a single randomly drawn fingerprint.



(a) Walkie-Talkie: Overheared bit positions during reconciliation are used to generate bit sequences with similarities higher than 50%

(b) BANDANA: Non-uniform distribution of 4-bit chunks per gait cycles

Fig. 10: Increasing one-shot success probability due to bias in sequences

5.2 Quantization-Specific Attacks (C, D)

An attacker with insight to a quantization scheme might be able to exploit this knowledge in order to boost her one-shot success probability. We discuss our observations in the Walkie-Talkie, BANDANA and IPI protocols. For SAPHE, we did not identify any quantization-specific weakness.

5.2.1 Walkie-Talkie Protocol

As discussed in 3.2, Walkie-Talkie is thought to be biased towards generating alternating sequences of 1-bits and 0-bits, which should be mitigated by applying a *privacy amplification*. We note that if an adversary were able to reconstruct the pattern-prone sequence before the amplification step, she would also be able to compute the *amplification*-step.

Figure 10a shows key similarities achieved by this attack when guessed sequences are compared to actual acceleration-based sequences. However, this only works for large window sizes. For small window sizes such as 10, the consecutive runs of indices become very short. Even worse, they might change signs when running over window borders, due to the newly computed guard band. Thus, the concerns about obvious patterns in the generated sequence are effectively mitigated by a small window size.

5.2.2 BANDANA Protocol

As indicated in Section 3.3, we found that the random success probability for the BANDANA protocol exceeds random guess. Indeed, as observed in Section 4 (Figure 4c), the variance in generated sequences is low and, in particular, sequences follow specific patterns (cf. Figure 5c). As depicted in Figure 10b, we found as the reason for this weakness that in the 4-bit chunks, which are generated per gait cycle (and before throwing away bits for reliability amplification), sequences of alternating binary value are significantly more frequent than others. In particular, sequences 1111 or 0000, where the instantaneous acceleration constantly exceeds or deceeds the mean acceleration, are seldom. Consequently, the distribution of key sequences in the key space is not uniform, and an adversary could utlilize this knowledge to launch an attack (C). We propose an approach to mitigate this problem in Section 6.

5.2.3 IPI Protocol

As discussed in Section 3.4, the IPI protocol suffers from measurement noise in accurately capturing the inter pulse interval due to the limited sampling rate of accelerometers. Especially for lower sampling rates, this significantly restricts the size of the key space. For instance, with 50 Hz (500Hz) sampling rate, one sample is taken every 20 milliseconds (every 2 ms). Since devices are not synchronized, this translates to an unavoidable inaccuracy of up to 10ms (1ms) for the sampled gait on devices (cf. Figure 2d). This measurement noise, compared with only 40.8ms standard deviation for the IPI results in a small keyspace and, since gray codes are employed (modulo 16; $q = 4$), not all bits in the generated quadruples change. In particular, we investigated the variation in 4 bit chunks generated by the IPI protocol on the walking data from [65]. In about 63% of the consecutive 4 bit chunks, all bits are identical. Furthermore, in 24% of all cases, just one bit changed, with 11% 2 bits changed and with only 0.02%, 3 bits were different. An adversary with approximate information on the IPI can therefore boost her guessing success probability significantly beyond chance.

5.3 Benefits and Pitfalls in using Error Correction

In biometric authentication systems, noise of the biometric information is an intrinsic property (here: measurement noise in acceleration sensors). Fuzzy cryptography has been proposed in order to employ error correcting codes to mitigate such noise. Error correcting codes encode messages from a messagespace $m \in \mathcal{M}$ into codewords of the (larger) codespace $c \in \mathcal{C}$ introducing redundancies. This process allows to correct errors introduced to c by decoding it back to m. In fuzzy cryptography, the biometric information or fingerprints contain noise or errors that can be corrected after mapping into $\mathcal{C}$. The redundancy introduced in the encoding process, however, dictates that an adversary also does not have to guess all bits in the fingerprint correctly, but can be sloppy. For instance, assume a key length of K and an error correcting code able to correct a fraction of u bits from the total fingerprint length N. This means that the success probability of a single randomly drawn fingerprint is not 2^N, but instead only

$$\sum_{k=0}^{u} \binom{N}{k} / 2^N = \frac{\sum_{k=0}^{u} \left(\frac{N!}{(N-k)! \cdot k!} \right)}{2^N} \tag{3}$$

since up to u errors are allowed at arbitrary position in the fingerprint sequence. Careful choice of the parameters is therefore demanded to limit the advantage gained by an adversary through the use of fuzzy cryptography.

From the protocols we investigated, BANDANA [11] and the IPI-protocol [12] employ BCH codes for error correction. [16] integrates the fuzzy vault design that operates on order-invariant tuples generated by Walkie-Talkie. The GaitKey [15] variant, which is further discussed in Section 6.2, implements a scheme by Yan et al. [71].

5.4 Gait Mimicry (A)

As recently discussed in [72], it is unlikely that an attacker would be able to mimic natural gait of a victim to a degree where gait sequences are sufficiently similar to break gait-based authentication or pairing schemes. In particular, the authors employed professional actors to mimic the gait of victims with similar physical properties (age, weight, height, shoe size, upper leg length) and showed that after guided training and instructions, all actors failed to mimic the

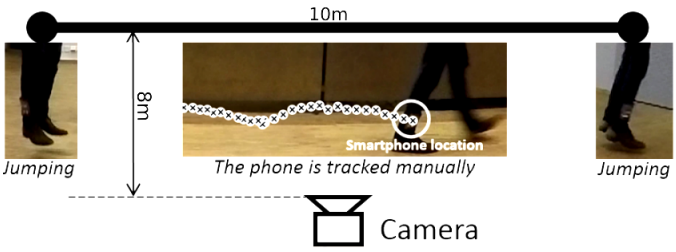


Fig. 11: Experimental setup for video-based attack on gait-based pairing

observed gait of victims. In a second test, by walking next to a victim one out of five attackers was able though to achieve sufficient similarity in the gait acceleration sequence. In particular, the authors assumed that the victim instinctively adapted her walking speed to the common step pattern with the adversary. This was, however, not further investigated.

5.5 Impersonation via Video Recording (G)

Cameras are omnipresent in these days, for instance as CCTV systems, personal camcorders, or mobile phones. The quality of captured videos is sufficient to discriminate subtle movements. An adversary with camera-support might therefore be able to extract pairing keys from recorded video (G). In this section, we investigate the threat of video-based side-channel attacks. In particular, we consider how accurate acceleration sequences describing gait can be estimated by tracking movement of body parts from video.

For our experiment, we captured movement of a subject both by a wearable inertial measurement unit (smartphone) and with a high-speed camera. The smartphone was attached to one leg. Five subjects (4 male; height: 1.63-1.95m; $\mu = 1.76$m) walked in a straight line in approximately 8m distance to the camera (1080p resolution; 90fps) mounted on a tripod (cf. Figure 11). Acceleration data was sampled at 50Hz. For synchronization between video and inertial sensor, a single jump both at the beginning and at the end framed the walking segment. Each subject conducted the experiment twice. We utilized Tracker[4] to manually track the location of the smartphone on the recorded video. Although human pose estimation [73] is able to estimate leg movements, we achieved higher accuracy by manually marking the location of the smartphone on the video frames.

For gait-based on-body pairing, the attacker is free to estimate gait according to the most easy to attack body location, since the protocols are inherently designed to pair acceleration sequences from arbitrary body location pairs. The Spearman's coefficient (1: perfect monotonically increasing relationship; 0: non monotonic relationship; -1: perfect monotonically decreasing relationship) [74] for gait sequences extracted at waist and shin in the dataset [65] is 0.44, which reflects their moderate increasing monotonic association. For instance, correlation between gaits extracted from these locations can be observed in Figure 12.

From the tracked trajectory we estimated the acceleration of the smartphone. We calculated the velocity in horizontal and vertical direction before computing the acceleration. The obtained result is smoothed by a Gaussian

4. http://physlets.org/tracker/

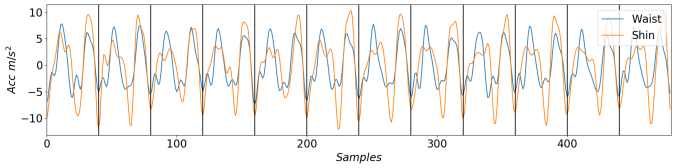
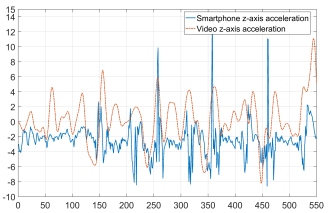
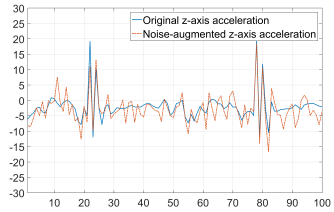


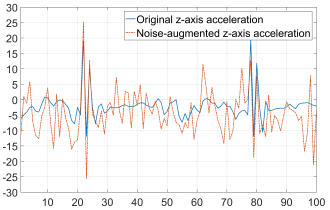
Fig. 12: Gait cycles extracted from shin and waist.



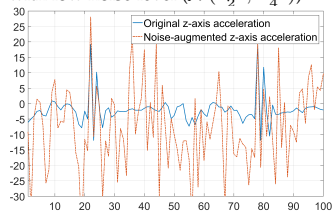
(a) Alignment of acceleration sequences from smartphone and video



(b) Acceleration sequence augmented with low noise level ($\mathcal{N}(\frac{\mu_v}{2}, \frac{\sigma_v^2}{4})$)



(c) Acceleration sequence augmented with video noise level ($\mathcal{N}(\mu_v, \sigma_v^2)$)



(d) Acceleration sequence augmented with high noise level ($\mathcal{N}(2\cdot\mu_v, 4\cdot\sigma_v^2)$)

Fig. 13: Acceleration signals featuring different noise levels

filter to reduce annotation noise. This estimated acceleration sequence is then re-sampled to match the 50Hz sampling rate of the inertial sensor. Note that we estimated movement orthogonal to ground since any rotation is implicitly corrected by the pairing scheme (Figure 13a).

To estimate the pairing performance and noise from video-extracted acceleration, in the dataset [65], we estimated the mean $\mu_v = 2.0921$[5] and standard deviation $\sigma_v = 6.0210$ of disparity values between optimally synchronized[6] gait acceleration sequences (estimated and recorded) in our experiment. These values were then used as parameters for noise distributions, which we added to the walking data recorded by the dataset in [65]. We generated Gaussian, Laplacian, and uniformly distributed noise[7].

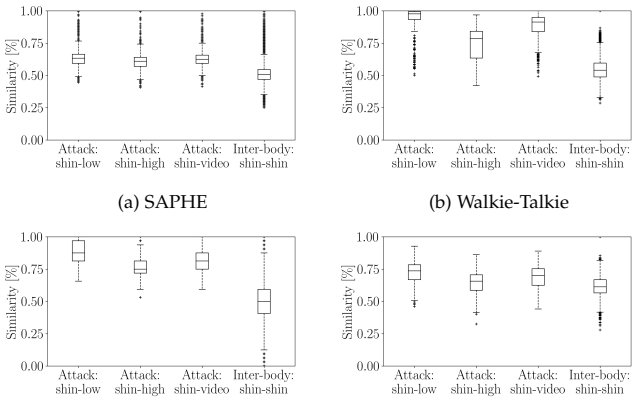
We then generated noisy acceleration signals with $\mathcal{N}(\mu_v, \sigma_v^2)$ (noise observed from video-based acceleration estimation), $\mathcal{N}(\frac{\mu_v}{2}, \frac{\sigma_v^2}{4})$ (low noise) and $\mathcal{N}(2 \cdot \mu_v, 4 \cdot \sigma_v^2)$ (high noise) as illustrated in Figure 13 for Gaussian additive noise. Other noise models are treated similar.

Figure 14 details the similarity for intra-body, inter-body, and video-based acceleration sequences with three noise levels. We assessed the effectiveness of video-based attacks on the four quantization schemes. Video-based acceleration is able to generate fingerprints which are sufficiently close to the actually recorded acceleration sequence, so that this

5. From the amplitude estimation error due to inaccurate distance measurement between camera and walking subject.

6. We refined the synchronization between the estimated and recorded acceleration sequences by shifting both sequences until a minimum root mean squared error is achieved

7. $p_n(n) = \frac{1}{\sqrt{\pi\sigma^2}} e^{\frac{(n-\mu)^2}{-\sigma^2}}$; $p_n(n) = \frac{1}{\sqrt{2}\sigma} e^{\frac{\sqrt{2}|n-\mu|}{-\sigma}}$; $p_n(n) = \frac{1}{2\sqrt{3}\sigma}$



(a) SAPHE (b) Walkie-Talkie

(c) BANDANA (d) IPI

Fig. 14: Attacks using Video-based impersonation: Similarity of gait-fingerprints with different noise levels over four pairing schemes

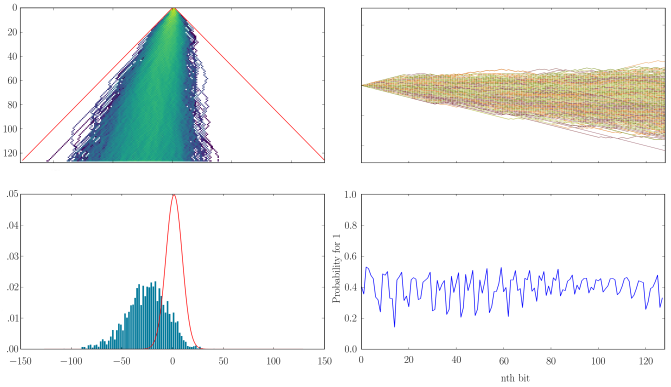


Fig. 15: Evaluations for Gait Key show an offset towards more zeros.

attack can break the gait-based pairing protocol for all three noise distributions considered. Walkie-Talkie [14] is the most vulnerable protocol under the video-based attacks. On the other hand, SAPHE [13] is the most secure protocol against video-based attacks (cf. Figure 14) .

6 Protocol Variants

In this section, we discuss improvements to SAPHE and BANDANA as well as a variant of Walkie-Talkie, exploiting n-ary quantization for higher bit-rate.

6.1 SAPHE

From the pairing schemes considered, SAPHE is the most promising as it introduces randomness instead of relying solely on gait-implicit randomness. As a potential improvement to our current implementation with a range of $1g$, we propose to implement a dynamic range. This would prevent outlier threshold values independent of the acceleration. Due to SAPHE's quantization, an attack, where a simple sinusoidal acceleration signal is artificially generated in alignment with the heel-strike, might then lead to a good estimate of the key. We propose to choose the threshold values as close to the acceleration reading as possible while still not revealing the actual unique gait features. This could be achieved by filtering out the dominant gait frequencies. Finally, instead of using hashed heuristic trees [13], we propose the usage of extensively studied cryptographic building blocks, such as fuzzy cryptography and a Password Authenticated Key Exchange.

6.2 Walkie-Talkie

In [15] Xu et al. present an evolved version of Walkie-Talkie, called Gait-Key. In contrast to Walkie-Talkie, multiple guard bands lead to several *quantization levels* and multiple bits.

We implemented this protocol and used four-ary quantization with $\alpha = 0.9$ as recommended in [15]. As a window size for quantization we chose 50 samples. We applied reconciliation and privacy amplification as in Walkie-Talkie. Figure 15 shows the randomness evaluation for Gait-Key. Similar to Walkie-Talkie, the key distribution is slightly shifted towards including more zeros. Remarkably, the Markov property shows periodic behaviour. The reasons for this are twofold: First, the quantization scheme calls for normal or equal distribution of acceleration samples. Biased distributions along a certain axis lead to unequal occurences of 1's and 0's. Due to the privacy amplification, exploiting XOR, this results in a larger amount of 0's in the final key (equal bits are mapped to 0). Second, slicing the acceleration space into several areas and assigning these with multiple bits per sample implies that consecutive samples generate bit sequences from identical or neighbouring areas. Hence, n-ary quantization achieves an improved, higher bit-rate but also introduces interdependence between bits while one guard band delivers the best performance.

6.3 BANDANA

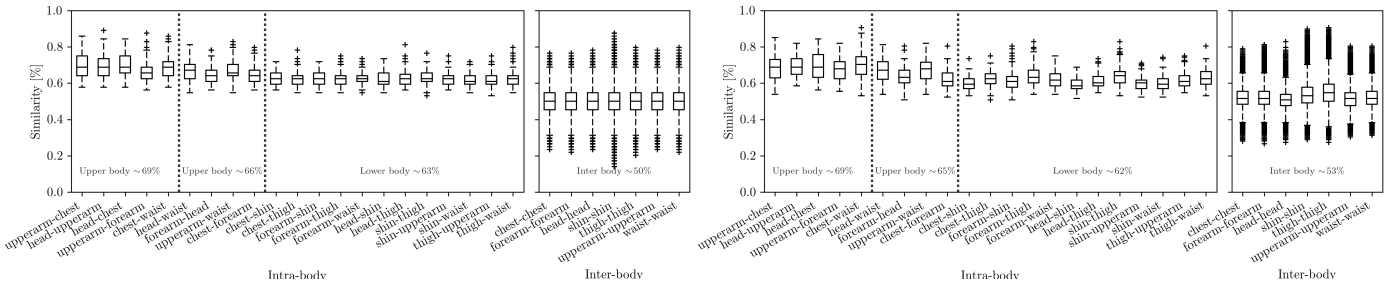
The quantization approach of BANDANA is biased towards specific patterns which are generated significantly more often than others (cf. Section 3). A straightforward solution is to disregard these 4-bit patterns with probability inverse to their occurrence frequency. However, due to the significant distortion of the histogram (cf. Figure 10b), this is not feasible. Since some patterns occur with a frequency of 1% or less, close to all frequent patterns would have to be discarded to arrive at a balanced random distribution.

Instead, we map each pair of consecutive bits in the generated key sequence to a single bit $(01, 11 \rightarrow 1, 10, 00 \rightarrow 0)$[8]. Figure 17a and 18a show the distribution of bit sequences after the mapping as well as the heatmap for fingerprints generated with the modified protocol.

The weakness described in Section 3 could be mitigated, however, due to the strong unbalancedness, some bias still remains even after the mapping as depicted in the histogram in Figure 17b. A further mapping can reduce this bias, however, this process also increases the time required to generate a particular key sequences as well as the similarity for intra-body pairings (cf Figure 16a).

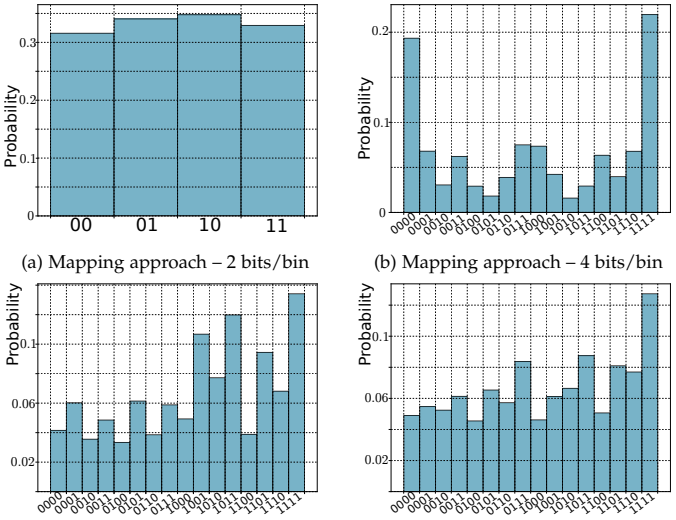
Another solution is to modify the comparison of gait sequences. The mean gait features an average amplitude with respect to the instantaneous gait sequences. Also, the acceleration peaks of the instantaneous gait fall with about

8. This does not leak information since 01 and 10 (11 and 00) are equally probable due to symmetry in the histogram in Figure 10b



(a) Mapping approach (b) Normalization approach

Fig. 16: BANDANA improvements: Comparison of intra-body against inter-body similarity for our proposed improvements

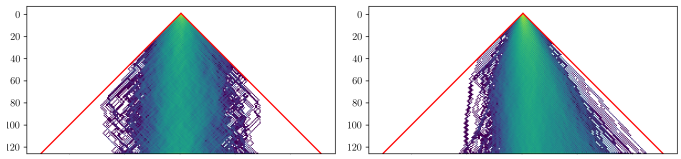


(a) Mapping approach – 2 bits/bin (b) Mapping approach – 4 bits/bin

(c) Normalization approach – normalizing acceleration amplitudes

(d) Normalization approach – additional disregarding of patterns according to inverse occurrence probabilities

Fig. 17: BANDANA improvements: Histograms generated from different improved versions of BANDANA



(a) Mapping approach (b) Normalization approach

Fig. 18: BANDANA improvements: Heatmaps of random walks for 128 bit keys generated by improved versions

equal probability to the left or right of the mean gait sequence. Consequently, the quantization, exploiting the difference between mean and instantaneous gait generates 0101 and 1010 patterns more often than other patterns. We suggest to normalize both mean and instantaneous gait prior to comparing them for gait generation. The heatmap and histogram for bit sequences generated with this modified versions are depicted in figures 17c and 18b.

The distribution is improved. Unfortunately, a bias towards including more '1'-s is introduced. However, since this bias is less severe than in the original BANDANA protocol, the effect can be damped by disregarding patterns with probability inverse to their observed occurrence frequency (cf. Figure 17d). We observe in Figure 16b that the similarity for intra-body pairing is slightly reduced.

7 Conclusion

We analyzed four acceleration-based pairing schemes. We have compared their quantization and discussed quantization-specific attacks. Their on-body pairing performance, statistical properties and entropy of generated key sequences were investigated based on walking data from 15 subjects and devices at 7 on-body locations.

Although not originally designed for the purpose of gait-based pairings, the SAPHE protocol achieved best results. We modified it towards gait-pairingby executing filtering and re-orientation before the pairing process. Still, room for improvement remains as shown in the randomness analysis.

The Walkie-Talkie protocol, which is able to generate the highest number of key bits achieves exact matching keys only across upper body locations and with low confidence. Together with SAPHE, it has the lowest one-shot success probability. This is, however, put into different perspective by a design flaw in the protocol. Even a naive adversary is able to boost her success probability to 0.125 by analysing the communication during the pairing process.

A similar quantization mechanism is utilized in the GaitKey implementation, which suffers from lack of randomness introduced by an n-ary quantization. The BANDANA protocol produces high similarity for different and also remote locations on the same body. However, the keys show a bias towards specific patterns. This problem originates from the quantization utilized and we proposed alternative mechanisms that address these issues.

Finally, the IPI protocol is also able to achieve high similarity across keys generated at different location on the same body. Our investigation revealed that the protocol suffers from a low variance in the generated binary patterns, so that similarity is also high for random gait sequences.

We further analyzed the threat of a video attack on gait authentication and pairing and found that a sophisticated attacker with video support and real-time gait estimation is able to break the studied gait-based pairing approaches.

Acknowledgment

We appreciate partial funding in the frame of an EIT Digital HII Active project, as well as from the Academy of Finland and from the German Academic Exchange Service (DAAD).

REFERENCES

[1] B. Guo, D. Zhang, Z. Wang, Z. Yu, and X. Zhou, "Opportunistic IoT: Exploring the harmonious interaction between human and the internet of things," *Journal of Network and Computer Applications*, vol. 36, no. 6, pp. 1531–1539, 2013.

[2] Z. Dawy, W. Saad, A. Ghosh, J. G. Andrews, and E. Yaacoub, "Toward massive machine type cellular communications," *IEEE Wireless Communications*, vol. 24, no. 1, pp. 120–128, 2017.

[3] S. Sigg, D. Schürmann, and Y. Ji, "PINtext: A framework for secure communication based on context," in *MobiQuitous 2011*, 2011.

[4] R. D. Findling, M. Muaaz, D. Hintze, and R. Mayrhofer, "Shake-Unlock: Securely Transfer Authentication States Between Mobile Devices," *IEEE Trans. on Mobile Computing*, vol. PP, no. 99, 2016.

[5] D. Schürmann and S. Sigg, "Secure communication based on ambient audio," *IEEE Trans. mobile computing*, vol. 12, no. 2, 2013.

[6] R. Jin, L. Shi, K. Zeng, A. Pande, and P. Mohapatra, "MagPairing: Pairing Smartphones in Close Proximity Using Magnetometers," *IEEE Trans. on Information Forensics and Security*, vol. 11, no. 6, 2016.

[7] S. Mathur, R. Miller, A. Varshavsky, W. Trappe, and N. Mandayam, "Proximate: proximity-based secure pairing using ambient wireless signals," in *Int. Conf. on Mobile systems, applications, and services*, 2011, pp. 211–224.

[8] M. Muaaz and R. Mayrhofer, "Smartphone-based Gait Recognition: From Authentication to Imitation," *IEEE Trans. on Mobile Computing*, 2017.

[9] K. Kunze, "Compensating for on-body placement effects in activity recognition," Ph.D. dissertation, Citeseer, 2011.

[10] D. Schürmann, A. Brüsch, S. Sigg, and L. Wolf, "BANDANA – Body Area Network Device-to-device Authentication using Natural gAit," in *IEEE PerCom*, Mar. 2017, pp. 190–196.

[11] D. Schürmann, A. Brüsch, N. Nguyen, S. Sigg, and L. Wolf, "Moves like jagger: Exploiting variations in instantaneous gait for spontaneous device pairing," *Pervasive and Mobile Computing*, vol. 47, pp. 1 – 12, 2018.

[12] Y. Sun, C. Wong, G.-Z. Yang, and B. Lo, "Secure key generation using gait features for Body Sensor Networks," in *IEEE BSN, 2017*, 2017, pp. 206–210.

[13] B. Groza and R. Mayrhofer, "SAPHE: simple accelerometer based wireless pairing with heuristic trees," in *10th Int. Conf. on Advances in Mobile Computing & Multimedia*. ACM, 2012, pp. 161–168.

[14] W. Xu, G. Revadigar, C. Luo, N. Bergmann, and W. Hu, "Walkie-Talkie: Motion-Assisted Automatic Key Generation for Secure On-Body Device Communication," in *ACM/IEEE Int. Conf. on Information Processing in Sensor Networks*, 2016.

[15] W. Xu, C. Javali, G. Revadigar, C. Luo, N. Bergmann, and W. Hu, "Gait-Key: A Gait-Based Shared Secret Key Generation Protocol for Wearable Devices," *ACM Trans. Sen. Netw.*, vol. 13, no. 1, pp. 6:1–6:27, Jan. 2017.

[16] G. Revadigar, C. Javali, W. Xu, A. V. Vasilakos, W. Hu, and S. Jha, "Accelerometer and fuzzy vault-based secure group key generation and sharing protocol for smart wearables," *IEEE Trans. on Information Forensics and Security*, vol. 12, no. 10, pp. 2467–2482, Oct. 2017.

[17] M. Nixon, J. Carter, D. Cunado, P. Huang, and S. Stevenage, "Automatic gait recognition," in *Biometrics*. Springer, 1996, pp. 231–249.

[18] J. Han and B. Bhanu, "Individual recognition using gait energy image," *IEEE Trans. on pattern analysis and machine intelligence*, vol. 28, no. 2, pp. 316–322, 2006.

[19] Z. Liu and S. Sarkar, "Improved gait recognition by gait dynamics normalization," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 6, pp. 863–876, 2006.

[20] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, and K. W. Bowyer, "The humanID gait challenge problem: data sets, performance, and analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 162–177, 2005.

[21] N. V. Boulgouris, D. Hatzinakos, and K. N. Plataniotis, "Gait recognition: a challenging signal processing technology for biometric identification," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 78–90, 2005.

[22] J. E. Cutting and L. T. Kozlowski, "Recognizing friends by their walk: Gait perception without familiarity cues," *Bulletin of the Psychonomic Society*, vol. 9, no. 5, pp. 353–356, 1977.

[23] S. A. Niyogi and E. H. Adelson, "Analyzing gait with spatiotemporal surfaces," in *IEEE Workshop on Motion of Non-Rigid and Articulated Objects*, 1994, pp. 64–69.

[24] J. E. Boyd, "Synchronization of oscillations for machine perception of gaits," *Computer Vision and Image Understanding*, vol. 96, no. 1, pp. 35–59, 2004.

[25] M. S. Nixon and J. N. Carter, "Advances in automatic gait recognition," in *IEEE Int. Conf. on Automatic Face and Gesture Recognition*, 2004, pp. 139–144.

[26] A. Veeraraghavan, A. K. Roy-Chowdhury, and R. Chellappa, "Matching shape sequences in video with applications in human movement analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, no. 12, pp. 1896–1909, 2005.

[27] A. Y. Johnson and A. F. Bobick, "A multi-view method for gait recognition using static body parameters," in *Int. Conf. on Audio-and Video-Based Biometric Person Authentication*, 2001, pp. 301–311.

[28] D. Gafurov, "Performance and security analysis of gait-based user authentication," Ph.D. dissertation, University of Oslo, 2008.

[29] S. J. Morris, "A shoe-integrated sensor system for wireless gait analysis and real-time therapeutic feedback," Ph.D. dissertation, Massachusetts Institute of Technology, 2004.

[30] B. Huang, M. Chen, P. Huang, and Y. Xu, "Gait modeling for human identification," in *IEEE Int. Conf. on Robotics and Automation*. IEEE, 2007, pp. 4833–4838.

[31] H. J. Ailisto, M. Lindholm, J. Mantyjarvi, E. Vildjiounaite, and S.-M. Makela, "Identifying people from gait pattern with accelerometers," in *Defense and Security*. Int. Society for Optics and Photonics, 2005, pp. 7–14.

[32] L. Rong, D. Zhiguo, Z. Jianzhong, and L. Ming, "Identification of individual walking patterns using gait acceleration," in *Int. Conf. on Bioinformatics and Biomedical Engineering*, 2007, pp. 543–546.

[33] L. Rong, Z. Jianzhong, L. Ming, and H. Xiangfeng, "A wearable acceleration sensor system for gait recognition," in *IEEE Conference on Industrial Electronics and Applications*, 2007, pp. 2654–2659.

[34] E. Vildjiounaite, S.-M. Mäkelä, M. Lindholm, R. Riihimäki, V. Kyllönen, J. Mäntyjärvi, and H. Ailisto, "Unobtrusive multimodal biometrics for ensuring privacy and information security with personal devices," in *Int. Conference on Pervasive Computing*. Springer, 2006, pp. 187–201.

[35] A. Kale, N. Cuntoor, B. Yegnanarayana, A. Rajagopalan, and R. Chellappa, "Gait analysis for human identification," in *Int. Conf. on Audio-and Video-Based Biometric Person Authentication*. Springer, 2003, pp. 706–714.

[36] W. Wang, A. X. Liu, and M. Shahzad, "Gait Recognition Using Wifi Signals," in *ACM Ubicomp 2016*, 2016, pp. 363–373.

[37] Y. Zeng, P. H. Pathak, and P. Mohapatra, "WiWho: WiFi-Based Person Identification in Smart Spaces," in *Int. Conf. on Information Processing in Sensor Networks*, April 2016.

[38] J. Jenkins and C. Ellis, *Using Ground Reaction Forces from Gait Analysis: Body Mass as a Weak Biometric*, 2007, pp. 251–267.

[39] K. Nakajima, Y. Mizukami, K. Tanaka, and T. Tamura, "Footprint-based personal recognition," *IEEE Trans. on Biomedical Engineering*, vol. 47, no. 11, pp. 1534–1537, 2000.

[40] R. J. Orr and G. D. Abowd, "The smart floor: A mechanism for natural user identification and tracking," in *CHI'00 extended abstracts on Human factors in computing systems*. ACM, 2000, pp. 275–276.

[41] L. Middleton, A. A. Buss, A. Bazin, and M. S. Nixon, "A floor sensor system for gait recognition," in *Workshop on Automatic Identification Advanced Technologies*. IEEE, 2005, pp. 171–176.

[42] G. Johansson, "Visual perception of biological motion and a model for its analysis," *Perception & Psychophysics*, vol. 14, no. 2, pp. 201–211, 1973.

[43] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, and K. W. Bowyer, "The humanid gait challenge problem: Data sets, performance, and analysis," *IEEE Trans. on pattern analysis and machine intelligence*, vol. 27, no. 2, pp. 162–177, 2005.

[44] Y. Wang, S. Yu, Y. Wang, and T. Tan, "Gait recognition based on fusion of multi-view gait sequences," *Advances in Biometrics*, pp. 605–611, 2005.

[45] T. Lam and R. Lee, "A new representation for human gait recognition: Motion silhouettes image," *Advances in Biometrics*, pp. 612–618, 2005.

[46] M. S. Nixon, T. Tan, and R. Chellappa, *Human identification based on gait*. Springer Science & Business Media, 2010, vol. 4.

[47] D. Gafurov, E. Snekkenes, and P. Bours, "Spoof attacks on gait authentication system," *IEEE Trans. on Information Forensics and Security*, vol. 2, no. 3, pp. 491–502, 2007.

[48] T. Hoang, D. Choi, and T. Nguyen, "Gait authentication on mobile phone using biometric cryptosystem and fuzzy commitment

scheme," *Int. Journal of Information Security*, vol. 14, no. 6, pp. 549–560, 2015.
[49] M. O. Derawi, C. Nickel, P. Bours, and C. Busch, "Unobtrusive user-authentication on mobile phones using biometric gait recognition," in *Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing*, 2010, pp. 306–311.
[50] M. O. Derawi, "Smartphones and biometrics: gait and activity recognition," 2012.
[51] D. Gafurov, "A survey of biometric gait recognition: Approaches, security and challenges," in *Annual Norwegian computer science conference*, 2007, pp. 19–21.
[52] A. Jain, P. Flynn, and A. A. Ross, *Handbook of biometrics*. Springer Science & Business Media, 2007.
[53] R. Kumar, V. V. Phoha, and A. Jain, "Treadmill attack on gait-based authentication systems," in *IEEE Int. Conf. on Biometrics Theory, Applications and Systems*, 2015.
[54] A. Kwong, C. Bolton, T. Trippel, W. Xu, and K. Fu, "Why Do You Trust Sensors? Analog Cybersecurity Attack Demos."
[55] B. B. Mjaaland, P. Bours, and D. Gligoroski, "Walk the walk: attacking gait biometrics by imitation," in *Int. Conf. on Information Security*. Springer, 2010, pp. 361–380.
[56] Ø. Stang, "Gait analysis: Is it easy to learn to walk like someone else?" Master's thesis, 2007.
[57] M. Muaaz and R. Mayrhofer, "An analysis of different approaches to gait recognition using cell phone based accelerometers," in *ACM Int. Conf. on Advances in Mobile Computing & Multimedia*, 2013, p. 293.
[58] L. Sloman, M. Berridge, S. Homatidis, D. Hunter, and T. Duck, "Gait patterns of depressed patients and normal subjects." *The American journal of psychiatry*, 1982.
[59] A. Srivastava, J. Gummeson, M. Baker, and K.-H. Kim, "Step-by-step detection of personally collocated mobile devices," in *ACM Int. Workshop on Mobile Computing Systems and Applications*, 2015, pp. 93–98.
[60] E. A. Heinz, K. S. Kunze, S. Sulistyo, H. Junker, P. Lukowicz, and G. Tröster, "Experimental evaluation of variations in primary features used for accelerometric context recognition," in *European Symposium on Ambient Intelligence*. Springer, 2003, pp. 252–263.
[61] J. Lester, B. Hannaford, and G. Borriello, *"Are You with Me?"–Using Accelerometers to Determine If Two Devices Are Carried by the Same Person*, Berlin, Heidelberg, 2004, pp. 33–50.
[62] C. Cornelius and D. Kotz, "Recognizing Whether Sensors Are on the Same Body," in *Pervasive'11*. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 332–349.
[63] R. D. Findling, M. Muaaz, D. Hintze, and R. Mayrhofer, "Shakeunlock: Securely unlock mobile devices by shaking them together," in *12th Int. Conf. on Advances in Mobile Computing and Multimedia*. ACM, 2014, pp. 165–174.
[64] R. Mayrhofer, "The candidate key protocol for generating secret shared keys from similar sensor data streams," in *European Workshop on Security in Ad-hoc and Sensor Networks*. Springer, 2007, pp. 1–15.
[65] T. Sztyler and H. Stuckenschmidt, "On-body Localization of Wearable Devices: An Investigation of Position-Aware Activity Recognition," in *IEEE PerCom 2016*, 2016, pp. 1–9.
[66] N. Mohssen, R. Momtaz, H. Aly, and M. Youssef, "It's the human that matters: Accurate user orientation estimation for mobile computing applications," in *MobiQuitous 2014*, 2014, pp. 70–79.
[67] T. Hoang, D. Choi, and T. Nguyen, "Gait authentication on mobile phone using biometric cryptosystem and fuzzy commitment scheme," *Int. Journal of Information Security*, vol. 14, no. 6, pp. 549–560, 2015.
[68] Y. Wang, "On Stochastic Security of Pseudorandom Sequences."
[69] N. Kuiper, "Tests concerning random points on a circle," in *Proceedings of the Koinklijke Nederlandse Akademie van Wetenschappen*, vol. Series a 63, 1962, pp. 38–47.
[70] R. M. Bolle, J. Connell, S. Pankanti, N. K. Ratha, and A. W. Senior, *Guide to biometrics*, 2013.
[71] L. Yang, W. Wang, and Q. Zhang, "Secret from muscle: Enabling secure pairing with electromyography," in *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems*, ser. SenSys '16. New York, NY, USA: ACM, 2016, pp. 28–41.
[72] M. Muaaz and R. Mayrhofer, "Smartphone-based Gait Recognition: From Authentication to Imitation," *IEEE Trans. on Mobile Computing*, vol. PP, no. 99, pp. 1–1, 2017.
[73] D. Mehta, S. Sridhar, O. Sotnychenko, H. Rhodin, M. Shafiei, H.-P. Seidel, W. Xu, D. Casas, and C. Theobalt, "VNect: Real-time 3D Human Pose Estimation with a Single RGB Camera," vol. 36, no. 4, 2017.
[74] C. Spearman, "The proof and measurement of association between two things," *The American Journal of Psychology*, vol. 15, no. 1, pp. 72–101, 1904. [Online]. Available: http://www.jstor.org/stable/1412159

Arne Brüsch received his B.Sc. in 2015 and his M.Sc. in 2017, both in computer science. Currently, he works as a research assistant at the Institute of Systems Security at TU Braunschweig. His research interests include quantization schemes for key generation in usable security as well as contextual security in general.

Ngu Nguyen is a doctoral student at Ambient Intelligence Group, Aalto University. He completed his bachelor's and master's degree at University of Science, Ho Chi Minh City, Vietnam. His research focuses on usable security and distributed machine learning.

Dominik Schürmann received the M.Sc. and Ph.D. degrees in 2014 and 2018 respectively, from TU Braunschweig. After working as a research fellow at the Institute of Operating Systems and Computer Networks at TU Braunschweig, he co-founded the company Confidential Technologies GmbH. His research interests include interaction-free security based on physical context and usable security in general.

Stephan Sigg received his M.Sc. degree in computer science from TU Dortmund, in 2004. and his Ph.D. degree from Kassel University, in 2008. Since 2015 he is an assistant professor at Aalto University, Finland. He has served as a TPC member of many conferences including IEEE PerCom, Ubicomp, etc. His research interests include Pervasive Computing, activity recognition, usable security and optimization of algorithms in mobile distributed systems.

Lars Wolf received his Ph.D. in 1995. In 1999 he was an associated professor at the computer science department of Universität Karlsruhe (TH). Since spring 2002 Lars Wolf is full professor for computer science at the TU Braunschweig where he is head of the Institute of Operating Systems and Computer Networks. His current research interests include wireless networking in general, sensor networks, vehicular networks, delay-tolerant networks, and mobile systems.