Pistis: Replay Attack and Liveness Detection for Gait-Based User Authentication System on Wearable Devices Using Vibration

Wei Song, Hong Jia, Min Wang, *Member, IEEE*, Yuezhong Wu, Wanli Xue, Chun Tung Chou, *Senior Member, IEEE*, Jiankun Hu, *Senior Member, IEEE*, and Wen Hu, *Senior Member, IEEE*

***Abstract*—Wearable devices-based biometrics has become mainstream in the biometric domain, especially in mobile computing, due to its convenience, flexibility, and potentially high user acceptance. Among various modalities, wearable devices-based gait recognition has been recognized as an effective user authentication method and employed in various applications, such as automated entry systems for home, school, work, vehicles, and automated ticket payment/validation for public transport. However, how secure wearable gait remains an open research question. In this study, we conduct a comprehensive security analysis of the wearable gait. Then, we demonstrate that gait itself is not robust against some attacking methods, such as spoofing or forgery. Therefore, we argue that an anti-spoofing mechanism is important for enhancing the security of wearable gait biometric systems. To this end, we proposed a novel authentication protocol called *Pistis* that embedded gait biometrics and a liveness detection mechanism that is aiming to detect various attacks of gait authentication systems. Our extensive experiments based on 50 subjects demonstrate that *Pistis* is effective in liveness detection and authentication performance enhancement, providing 100% accuracy for human and nonhuman detection, and 99.53% accuracy for user authentication. Pistis can be used as a liveness detection method for wearable devices-based biometrics, significantly for wearable gait.**

***Index Terms*—Biometric authentication, gait, liveness detection, wearable device security.**

I. Introduction

WEARABLE devices have become very popular and indispensable in people's lives. In particular, they benefit people in various ways, such as convenient payments and continuous healthcare monitoring health services. To protect the sensitive data in wearable devices, traditional personal identification number (PIN), and some other secure patterns have been developed [97]. However, existing methods are not completely secure, especially when under various attacks, such as spoof attacks and shoulder surfing attacks. To this end, researchers have studied a number of biometrics-based approaches as authentication methods to improve the security of the wearable devices, such as through fingerprint [81], face [82], vein [83], Iris [84], voice [85], electrocardiography (ECG) [86], [87], and gestures [88], [89], [90], [91], [92]. However, most of these methods require dedicated sensors to capture the biometrics, increasing the cost and price of wearable devices. Besides, these static authentication methods are vulnerable to replay attacks, for example, replay attacks to face, fingerprint, and veins using 3-D printing technologies [93], [94], [95], [96]. Therefore, there is still a paucity to design an inexpensive and reliable authentication method for ubiquitous wearable devices.

Recently, human gait has shown to be a unique characteristic among individuals and is hard to be attacked through reproducing techniques [3], [37], [52]. In the wearable-based gait authentication approach, human gait (i.e., walking motion dynamics) is usually obtained by the accelerometer, e.g., within an inertial measurement unit (IMU), which is widely embedded in many wearable devices. Specifically, as an indicator, human gait (i.e., walking motion dynamics) is usually captured by the accelerometer in the wearable-based gait authentication approach [1], [2]. The measured acceleration, which contains the motion dynamics of walking, can be used for user authentication and identification. Nevertheless, the security of wearable gait authentication systems has always been a concern in the research community [7]. To date, the challenge of wearable gait security has been mainly focusing on imitation attacks [8], [54], [55], in which imitators/attackers imitate how a victim walks. The imitation attack scarcely compromised the gait-based authentication system since gait is a behavioral biometric that embedded many user-dependent unique dynamics [37], [52]. Despite some

Manuscript received 26 October 2022; accepted 9 December 2022. Date of publication 22 December 2022; date of current version 25 April 2023. This work was supported by the Australian Research Council under Linkage Grant LP180100663. *(Corresponding author: Wei Song.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the University of New South Wales Ethical Committee of DVCR RECS Human Ethics under Application No. HC210467

Wei Song, Yuezhong Wu, Chun Tung Chou, and Wen Hu are with the School of Computer Science and Engineering, The University of New South Wales, Sydney, NSW 2052, Australia (e-mail: wei.song1@unsw.edu.au; yuezhong.wu@student.unsw.edu.au; c.t.chou@unsw.edu.au; wen.hu@unsw.edu.au).

Hong Jia is with the Department of Computer Science and Technology, University of Cambridge, CB2 1TN Cambridge, U.K. (e-mail: hj359@cam.ac.uk).

Min Wang and Jiankun Hu are with the School of Engineering and Information Technology, The University of New South Wales, Canberra, ACT 2612, Australia (e-mail: maggie.wang1@adfa.edu.au; j.hu@adfa.edu.au).

Wanli Xue is with the Security Lab of Sophos, Sophos, Sydney, NSW 2060, Australia (e-mail: wanli.xue@sophos.com.au).

Digital Object Identifier 10.1109/JIOT.2022.3231381

research on imitation attacks, studies on the overall security features of wearable-based gait authentication are still not been closely investigated [43]. Hence, the need for a further investigation of wearable-based gait user authentication security becomes increasingly crucial.

In this work, we comprehensively analyze the security of wearable gait authentication and propose a novel *vibration replay attack* based on the mechanism of the existing wearable gait authentication system by using stand-alone vibration motors to recreate a victim's gait pattern. Through a comprehensive analysis of the gait features of 50 subjects, we find that their gait patterns can be well approximated by the weighted sum of three or fewer single-frequency sinusoids. Therefore, we propose to control the amplitude, frequency, and phase of vibration signals produced by the stand-alone vibration motors to recreate these sinusoidal components so that the resulting signal sampled by the accelerometer will approximate the victim's gait pattern. In this way, we can attack the wearable-based gait authentication into accepting the resulting accelerometer signal as the victim's gait signal. To demonstrate the feasibility of the vibration attack, we developed a hardware-based attack prototype using the stand-alone vibration motors, which are able to accurately reproduce victims' gait signals and be used to compromise the wearable gait-based authentication systems.

To defend against such a vibration replay attack, we propose *Pistis*,[1] a novel liveness detection mechanism, which exploits the *bio-vibrometry* of the human body. *Pistis* utilizes the wearable device's *on-board* vibration motor to generate a vibration signal as an excitation. This vibration signal will reach the human body (i.e., human wrist and hand) and reflect. Then we collect the reflected vibration signal, which is considered the vibration response signal using the *on-board* accelerometer of the same wearable device. The collected vibration response signal depends on the physical characteristics—such as the mass, geometrical features for example bone size, and body-fat ratio—of the wearer's hand and wrist, and can be exploited to differentiate humans from nonhuman entities. Finally, we integrate *Pistis* with the wearable gait authentication to obtain a novel secure gait-based authentication system with liveness detection.

To reach the goal of liveness detection for wearable-based gait user authentication, we face multiple technical challenges in practice. First, bio-vibrometry features of the human body are affected by various elements, such as body mass, geometry features, and body-fat ratio, such that it is challenging to model and quantify the bio-vibrometry of the human body together with the wearable device. To address this challenge, we model the wearable device-human body connection as a *spring-mass-damper* vibration system, in which the mass m, stiffness coefficient k, and damping coefficient c are exploited to model the bio-vibrometry features embedded in the vibration signal. Second, for the current COTS mobile and wearable devices, the sampling rates of the accelerometer are limited (e.g., up to 100 Hz), while the frequency of vibration signals generated by the vibration motors is in the range of 130–180 Hz. According to the Nyquist sampling theorem, the vibration signal is undersampled, such that part of the hand biometrics information would be lost which may influence the liveness detection performance. Thus, *Pistis* should be able to reconstruct the undersampled vibration signal and obtain the full information. To overcome the vibration undersampling problem in COTS wearable devices, we develop a signal mapping method by using ℓ_1 minimization for signal reconstruction in compressive sensing. Third, since the current COTS wearable devices are computational limited, the liveness detection is expected to be lightweight, and no requirement for any extra hardware.

Our main contributions are summarized as follows.

1) We comprehensively analyze the accelerometer sensor-based human gait signal of 50 subjects and show that these signals can be well approximated by a weighted sum of three or fewer sinusoids.
2) We propose a novel wearable gait replay attack based on stand-alone COTS vibration motors, and develop an attack model prototype. We evaluate the attack performance of various wearable gait authentication models with 50 subjects, showing that our attack model can compromise all the wearable gait authentication models and systems.
3) We design and develop a novel liveness detection mechanism, *Pistis*, for the existing wearable-based gait authentication systems. *Pistis* exploits on-board vibration motors and accelerometers of COTS wearable devices without any additional hardware. To the best of our knowledge, we are the first to fuse bio-vibrometry and gait biometrics for user authentication and liveness detection purposes. We evaluate *Pistis* with a data set of 50 subjects, and extensive experiments show that *Pistis* produces 100% accuracy in liveness detection and the authentication accuracy of 99.53%.

The remainder of this article is organized as follows. We first discuss the related work in Section II and present our comprehensive gait analysis in Section III. Followed by a description of our proposed vibration attack model in Section IV. Then, we demonstrate the design of our proposed liveness detection model in Section V. After that, we evaluate the performance of our attack model and liveness detection mechanism in Section VI, and finally, Section VII concludes this article.

II. Related Work

A. Gait Authentication

Gait, as identification biometrics, has been actively studied for many years. There are mainly three approaches: 1) computer vision-based gait authentication methods [70], [71]; 2) floor sensor-based gait authentication models [72], [73], [74]; and 3) wearable gait authentication systems [1], [2], [7], [10], [36], [40], [41], [42], [47], [48], [49], [50], [51], [53]. Among these three approaches, the accelerometer-based gait authentication system is the most popular method. This first work of gait authentication with the accelerometer sensor was proposed by Ailisto et al. [40].

[1] In Greek mythology, *Pistis* is the personification of good trust, reliability, and faith.

In their experiment, the dedicated accelerometer sensor is attached to the subject's waist. With walking data from 26 subjects, they achieved an equal error rate (EER) of 7% using a signal correlation approach. Gafurov et al. [50] developed gait-based authentication systems further with a 3-D accelerometer sensor attached to the subject's leg right above the ankle. They assumed that the acceleration magnitude of the subject's leg is larger than those in other body parts such as the waist, which would provide better walking information. With 3-D acceleration signals, they could reach an authentication EER of 5%. With the prevalence of mobile wearable devices (e.g., smartphones, smartwatches, and smart wristbands), researchers proposed multiple gait authentication systems based on smartphones, smartwatches, and smart wristbands because these wearable devices were well equipped with various motion sensors, such as accelerometers and gyroscopes, which enabled collecting gait acceleration signals ubiquitously. Hong et al. proposed an unobtrusive gait recognition system, which is based on mobile smartphones in [51]. Then, Primo et al. investigated the impacts of different body locations (e.g., hand and pocket) on the accuracy of the gait-based authentication systems [53]. Previous work [40], [50], [51], [53] developed several advanced algorithms to extract the gait signal from acceleration data and recognize gait patterns. However, their experiments are based on the ideal scenario, i.e., walking on flat even ground. In reality, different walking activities, e.g., walking upstairs, walking downstairs, and different types of ground, e.g., walking on grass and walking on the carpet, would significantly affect the accuracy of gait recognition. To this end, Xu et al. [1] highlighted the limitation of gait authentication systems without activity context before further developing a context-aware gait authentication system that can reach an authentication accuracy of 95.3%.

B. Security of Gait

There have been several studies investigating how to attack sensor-based gait authentication systems. Gafurov et al. [54] asked attackers to imitate victims' walking manners without any feedback. They collected 760 gait sequences from 100 subjects, and their experiments were composed of two stages. In the first experiment stage, subjects walked in their usual manner, reaching an EER of 13%. In the second experiment stage, subjects attempted to walk into victims' gait patterns. Unfortunately, the impersonation attack failed to attack the gait authentication systems, which demonstrated their resilience to the imitation attack. In comparison, Mjaaland et al. [55] trained the attackers when the attackers were imitating the victims with feedback. They assumed that such training with feedback would increase the probabilities of successful attacks. However, their results showed that imitation attacks were very difficult to launch because it is challenging for the attackers to imitate the victims' walking style accurately. Ren et al. [8] also verified that mimicry attack on gait authentication systems was very difficult based on their long-term experiment. Unlike the previous gait attacking works, Zhu et al. [7] proposed a one-cycle attack model to attack the gait authentication systems that only used one gait cycle for authentication. They assumed that independent individuals might have similar gait cycles and applied the K-means algorithm to the gait cycles of different subjects. They found that some of the gait cycles could be grouped into the same cluster. The gait cycles (of different subjects) from the same cluster could bypass seven classifiers used by gait authentication systems. Zhu et al. evaluated their one-cycle attacking model with both a public data set with 744 subjects and their own data set with 20 subjects. Their results showed that the one-cycle attack model could compromise most victims within a limited number of attempts.

However, the one-cycle attack cannot successfully attack the gait authentication systems, which utilize multiple gait cycles. Therefore, the proposed liveness detection gait-based authentication protocol takes three gait cycles as the input.

Different from the above imitation attacks, our proposed vibration replay attack exploits COTS vibration motors to directly reproduce the victim's gait template signal. We have demonstrated that it can successfully attack the wearable gait authentication systems.

C. Vibration Authentication

There have been few studies that have investigated body vibration for user authentication [80], [97], [98], [99]. First, researchers proposed to utilize the standalone vibration motor and accelerometer which are attached to the user's wrist to acquire body vibration features for user authentication purposes [80]. However, their work can only work with a small group of users, i.e., less than 10. Besides, their work is not based on commercial wearable devices which are inconvenient for ubiquitous deployments. Similarly, researchers also investigated how to use human body vibration to authenticate users [97]. Specifically, they proposed a vibration authentication model that exploits Apple Watches to authenticate users directly based on models, including Random Forest and support vector machine (SVM). However, their authentication model is also based on a small group of users, i.e., less than 6. In comparison, we implemented the models proposed [45] and evaluated their performance on our data set and show that they can produce approximately 98% (similar to that presented in [45]), 91%, and 53% authentication accuracy for 6, 10, and 20 subjects, respectively. Meanwhile, we demonstrate that the vibration-only authentication model is not scalable to a large number of users for authentication purposes. To solve this, Thuraisingham et al. [98] proposed a two-factor authentication method in which a user needs to input the PIN or draw a pattern on a dedicated vibrating panel. Similarly, Li et al. [99] proposed VELODY which is based on a dedicated device, i.e., a panel embedded with speakers and accelerometers. It generates vibrations at a randomly selected frequency and authenticates the user by verifying the response to vibration when the user places their palm on the vibrating panel.

In contrast, instead of using it to authenticate the user, we first leverage vibration to perform a replay attack to the wearable gait biometric, and then we embed the human body vibration into our authentication protocol framework as a liveness detection method to defense various attacks.

III. Gait Analysis

A. Primitives for Gait Analysis

This section provides background on wearable-based gait authentication systems and mechanical vibration produced by the vibration motors, which are used for both attack and liveness detection models. We also explain the phenomenon of signal aliasing due to undersampling, which is used in our proposed replay attack.

1) Wearable Gait Authentication Systems: Human gait is a sequence of movements produced by a walking person. Usually, we use gait cycles to describe walking because this kind of motion exhibits cyclicity. One gait cycle comprises six phases: 1) heel strike, 2) foot flat; 3) mid-stance; 4) heel-off; 5) toe-off; and 6) mid-swing [56]. These phases together will create the human walking style, which is different from person-to-person due to their unique height, muscle and bone structure, body mass index, etc. Gait authentication systems utilize such a human walking style to authenticate users.

The unique properties in the gait cycle (e.g., walking speed, length of steps, thigh lift, hip movement, the swing of hands, and width of steps) make it feasible to use the human gait pattern for authentication purposes. Depending on how the human gait raw data are measured, gait-based authentication systems are divided into three groups: 1) vision based; 2) floor sensor based; and 3) wearable based [1]. In a vision-based gait authentication system, a video camera placed at an appropriate distance records a subject's walking movement sequences as video/image frames. After that, silhouettes will be extracted from the image frames with well-designed video/image processing algorithms. Finally, the subject's gait cycles and features will be extracted for further authentication purposes. In a floor sensor-based gait authentication system, floor sensors (e.g., force plates) are installed under the floor to collect gait features, such as stride length, stride cadence, and toe-to-heel ratio [4]. These features can then be fed to machine learning algorithms for authentication purposes.

In a *wearable gait authentication system*, human gait is measured by an accelerometer, e.g., within an IMU. Compared with vision-based gait measurements, walking acceleration can describe gait dynamics more directly and accurately [1]. In the past, body-worn accelerometers were attached to the different human body locations, e.g., waist, leg, and arm, to collect the acceleration data of human walking. With the miniaturization of accelerometers and the prevalence of wearable devices, e.g., smartphones and smart wristbands, recent works of gait authentication are mainly based on wearable devices, which have onboard IMUs. As shown in Fig. 1, The wearable device-based gait authentication system is composed of two components: 1) enrollment and 2) authentication.

Enrollment: When the user activates the enrollment, the acceleration will be examined by utilizing the attached accelerometer at a specific body position. Since the energy of the human gait signal is concentrated in the frequency range of 0.5–3 Hz, a band-pass filter will be applied to the raw acceleration data to remove both high and low-frequency noise. After filtering, we can obtain good-quality acceleration data of the subject's gait cycles. Next, the gait cycle segmentation algorithm is applied to extract the individual gait cycles (individual gait cycles are searched by minima of the gait signal after the band-pass filter). This algorithm may be implemented by simply searching for the minima in the acceleration signal. For example, we may take the samples between every two consecutive minima as a step, and one gait cycle is obtained by combining two consecutive steps. After that, abnormal gait cycles will be eliminated based on the length of gait cycles. For humans, specifically, most of the gait cycles last for 0.8–1.3 s [1]; therefore, given a specific sampling rate (e.g., 100 Hz), most of the gait cycles length should be in the range of 80–130, and for gait cycles which are in this range will be eliminated. After abnormal gait cycle elimination, linear interpolation and normalization algorithms will be applied to reduce the differences caused by different accelerometer attachment locations, i.e., different body locations. Finally, the gait cycles will be input into the deep learning model to train a user-specific authentication model.

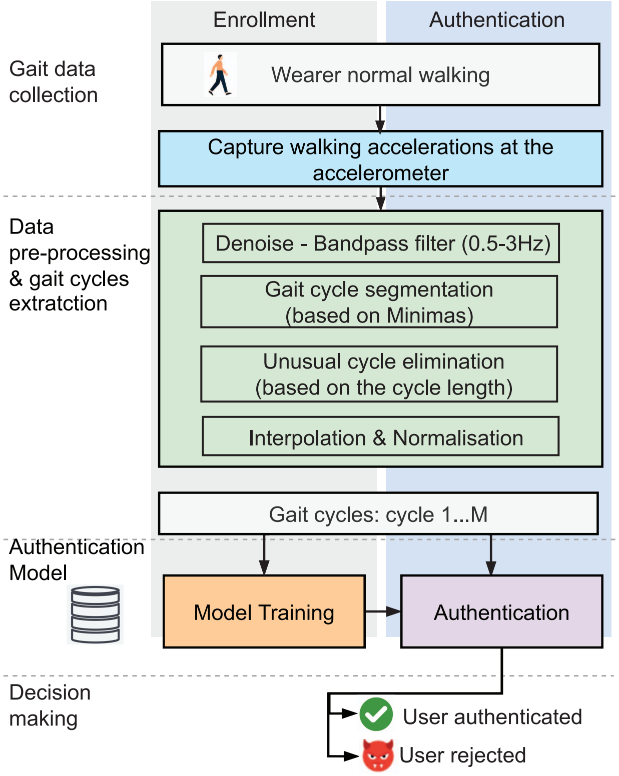

Enrollment
Authentication
Gait data collection
Wearer normal walking
Capture walking accelerations at the accelerometer
Data pre-processing & gait cycles extratction
Denoise - Bandpass filter (0.5-3Hz)
Gait cycle segmentation (based on Minimas)
Unusual cycle elimination (based on the cycle length)
Interpolation & Normalisation
Gait cycles: cycle 1...M
Authentication Model
Model Training
Authentication
Decision making
User authenticated
User rejected

Fig. 1. Typical wearable gait authentication system.

Authentication: After obtaining the authentication model from the enrollment, the gait authentication system is ready for user authentication purposes when the user attempts to access the wearable devices. Please note that the gait sample data collection and preprocessing are the same as what we did in the enrollment phase, i.e., denoise, gait cycle segmentation, unusual cycle elimination, interpolation, and normalization. Finally, the query gait cycles will be put into the well-trained model for authentication until an authenticated user to match.

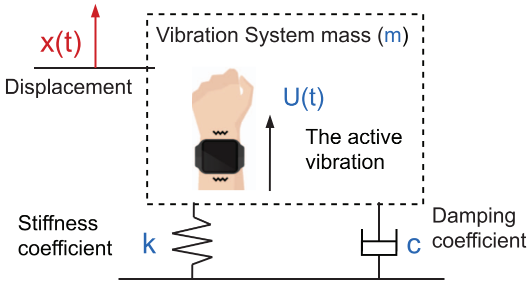


Fig. 2. Spring-mass-damper vibration model.

2) Modeling Vibration: We use vibration motors for two purposes: 1) realizing replay attack (Section IV) and 2) liveness detection (Section V). Vibration motors are low-cost devices extensively used in various scenarios and applications, e.g., message alerting in wearable devices, vibration diagnostic tools [32], [33], and vibration tactile feedback applications [34], [35]. Generally, vibration motors produce vibration movements by using an unbalanced mass inside their enclosure [13]. There are two main types of vibration motors: 1) eccentric rotating mass (ERM) and 2) linear resonant actuator (LRA). An ERM motor uses an off-center mass to generate a centripetal force when the mass rotates at high speed. This motion causes the entire vibration motor to vibrate. In contrast, an LRA motor uses magnetic fields and electrical currents to create a force to drive the mass up and down. This causes the whole vibration motor to displace and vibrate.

The vibration generated by a vibration motor can be divided into two states: 1) transient state and 2) steady state. The transient state occurs when the vibration motor is initially activated and typically lasts for a short duration. In comparison, the steady state is the stage where the vibration becomes oscillatory [15]. Our work is based on a steady state. Unless explicitly specified, our discussion below assumes the steady state of vibration.

Mechanical vibration can often be modeled by the well-known spring-mass-damper model, as illustrated in Fig. 2. The model presents a mechanical system with mass m, the spring stiffness coefficient k, the damper damping coefficient c, and displacement $x(t)$ at time t under the external excitation force $U(t)$. The external excitation force here is the centripetal force generated by the mass rotation/movement, which is usually modeled as follows:

$$U(t) = U_0 \cdot \sin(2\pi f t) \tag{1}$$

where f is the external excitation frequency, t is the time, and U_0 is the amplitude of the centripetal force. The mass displacement satisfies the following ordinary differential equation when the excitation force is applied to the mass:

$$m\ddot{x}(t) + c\dot{x}(t) + kx(t) = U(t) \tag{2}$$

where $\dot{x}(t)$ and $\ddot{x}(t)$ is first and second derivatives of $x(t)$, i.e., the velocity and acceleration, respectively. At steady state, the mass displacement $x(t)$ is the following sinusoid:

$$x(t) = \frac{U_0}{\sqrt{(k - f^2 m)^2 + (2\zeta f)^2}} \cdot \sin(2\pi f t + \phi) \tag{3}$$

where ζ is damping ratio, which is used to characterize the amount of damping in a mechanical vibration system, and it normally will cause the signal attenuation, and ϕ is phase shift. The formulas for ζ and ϕ are

$$\zeta = \frac{c}{2 \cdot \sqrt{km}} \quad \text{and} \tag{4}$$

$$\phi = \arctan\left(\frac{c \cdot f}{k - mf^2}\right). \tag{5}$$

For liveness detection in Section V, the m, c, and k refer to the mechanical properties of human skin or nonhuman entities. We find that we can use vibration response to *distinguish* between humans and nonhumans because they have distinct responses to vibration.

When we use the vibration motor to realize the gait replay attack in Section IV, we aim to find a force $U(t)$ that will create an accelerometer output, which will mimic the gait of the victim. Our experience shows that both damping and phase shift can significantly influence the attack success rates. To prevent vibration attenuation caused by damping and to eliminate the impact of the phase shift, we suspend the victim's wearable device in the air. In this way, all constraints in the vibration direction are removed, such that the damping coefficient (c) for this case is negligible and may be considered to be zero. If c is 0, then our vibration model can be simplified to a *spring-mass-undamped* model whose displacement can be written as follows:

$$x(t) = \frac{U_0}{k - f^2 m} \cdot \sin(2\pi f t). \tag{6}$$

3) Signal Aliasing: To realize our proposed replay attack on the wearable gait, we need to reproduce a gait acceleration pattern using the vibration motors. However, there is a mismatch between the frequency produced by the COTS vibration motors (which is at 10 Hz or above) and the gait frequency (which is 0.5–3 Hz). We see that from (3) that a vibration motor oscillating at frequency f will produce an oscillatory movement of the same frequency f in an object that the motor is attached to. Therefore, it is not possible to directly produce a 0.5–3 Hz using a vibration motor. However, we can overcome this problem by using signal aliasing as a result of undersampling.

We consider the case where a signal of frequency f is sampled at frequency F_s where $f > (F_s/2)$. This means F_s is below the Nyquist frequency of f. Standard results from digital signal processing says that the sampled signal will have a frequency $\varepsilon \in [-(F_s/2), (F_s/2)]$ where f, F_s, and ε are related by

$$f = n \cdot F_s + \varepsilon \tag{7}$$

where n is a positive integer. The sampling frequencies of the accelerometers on wearable devices for gait-based authentication systems can be considered public domain knowledge, and we will assume that F_s is 100 Hz, which is the maximum sampling frequency supported by most Android wearable devices. Therefore, by using a vibration motor that oscillates at $f = 102$ Hz, the above result says that the output of the accelerometer will have a frequency of 2 Hz, which is inside the gait frequency range.

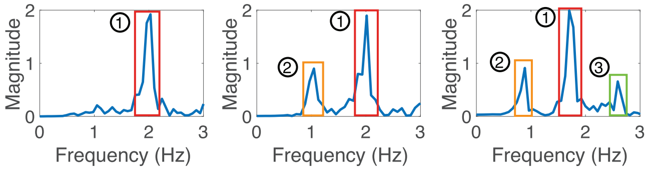


Fig. 3. Spectrograms of the gait signal from three subjects. Gait signal with one (a), two (b), and three (c) dominant frequency peaks.

Although accelerometers can measure accelerations in three directions, previous gait authentication systems show that it is sufficient to use the acceleration in one direction to authenticate users [1]. Therefore, in our study, we use the gravity direction acceleration data collected from users' wrists using a smartwatch as the raw gait data.

B. Data Set

Our data collection campaign[2] has 50 subjects with 20 males and 30 females altogether. Their ages range from 18–s45, heights ranging from 150 to 190 cm, and weights ranging from 45 to 120 kg. All of them are healthy and able to walk naturally. Subjects are asked to wear a smartwatch (Samsung Gear Live Smart Watch) on their right hand during the data collection and then walk normally on an even flat lane. Simultaneously, the onboard vibration motor will be activated, and vibration response will be collected while the subjects are walking. For all the data collection experiments, we ensure that the vibration and walking start simultaneously. The sampling frequency of the internal accelerometer of the smartwatches is set as 100 Hz. We collect 10 min of walking and related vibration response data for each subject.

C. Gait Analysis Results

Since human gait is naturally a repeated and regular motion, gait signals are quite sparse in the frequency domain, which will be exploited in our attack model. We find that the number of dominant frequencies in the range of 0.5–3 Hz for our 50 subjects is three or less. Fig. 3 shows the spectrograms of the gait signal of three different subjects, who have, respectively, 1, 2, and 3 dominant frequency peaks in their gait. In the range of 1.5–2.5 Hz we can find the one with the largest magnitude and we define it as dominant-1 frequency for all subjects. Similarly in the range of 0.5–1.5 and 2.5–3 Hz we usually can find dominant-2 and dominant-3 frequencies, respectively, which have lower magnitudes than dominant-1 frequency.

With data from 50 subjects, we find that approximately 20% of them have one prominent frequency in their gait signal frequency spectrum, 60% of them have two, and 20% of them have three. Therefore, we can model the human gait as a sum of n sinusoids, where n is the number of dominant frequency components, and the ith dominant frequency component has amplitude a_i, frequency f_i and phase ϕ_i. With this notation, the gait signal s_t can be written as follows:

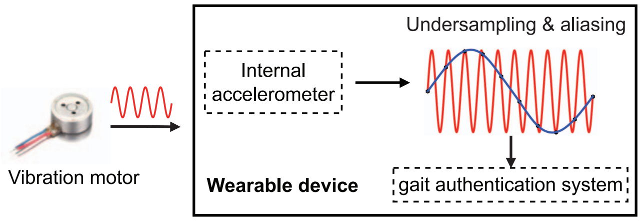


Fig. 4. Illustration of the attack model. The victim's wearable device is taken as a black box. The external vibration motor presents a vibration signal; it will be sampled by the internal accelerometer of a wearable device before reaching the gait authentication system.

$$s_t = \sum_{i=1}^{n} a_i \cdot \sin(2\pi f_i t + \phi_i). \tag{8}$$

Since the gait signals are normalized in the gait authentication signal processing pipeline, we can replace the absolute amplitudes $a_1, \ldots, a_n$ with relative amplitudes. As for phase shifts, we developed a phase shift searching algorithm, which searches values of phase shift from 0 to π with an interval of $(\pi/10)$. We select the phase shift values that can reach the best decomposition (i.e., with a minimum Euclidean distance to the original signal).

IV. Gait Attack Model

In this section, we introduce our replay attack using vibration motors. Our discussion includes attack scenario and model, prototype implementation, and attack model configuration.

A. Model Description

We assume that the attackers have the victim's gait template data and wearable devices, e.g., smartphones, smart watches, and smart wristbands. Here, the attackers may gain access to the victim's gait template signal in various ways. For example, video-based side-channel attack [43], [44] can efficiently extract gait acceleration sequences that are close to on-body measured acceleration sequences. This kind of video-based attack for pilfering gait template sequences requires a high-speed camera that currently is omnipresent (e.g., personal camcorders and smartphones). During the video-based attack, an attacker will first make a short video of the victim's walking (e.g., 10 s) before extracting the acceleration sequences of gait from the video frames. A video-based side-channel attack is an efficient way to obtain the victim's gait template sequences, which forms the basis of our vibration replay attack.

In our attack model, as illustrated in Fig. 4, attackers take the victim's wearable device as a *black box* and attempt to compromise the authentication algorithm on the victim's device by using the signal generated by *external vibration motors*. Moreover, we assume that the attacker knows the accelerometer's sampling rate either from public knowledge or by brute-force attempts, e.g., Android supports four different accelerometer sampling rates (FASTEST, GAME, UI, and NORMAL) only.

[2] Data collection was approved by the Ethics Committee of XX (anonymized during double-blind review process), reference number HC210467.

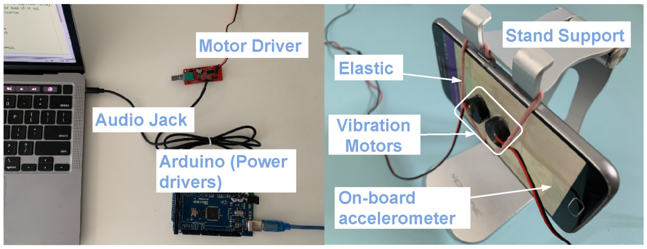


Fig. 5. Implementation of the attack model with external vibration motors, and undersampling theory.

When the sampling rate is insufficient, with signal undersampling and aliasing phenomenon, the purposely designed vibration signals generated by the vibration motors can create a sampled acceleration signal, which is similar to the victim's gait template signal. This signal will be fed into the wearable gait authentication system to attack the authentication model. We note that our attack model does not require knowledge of the authentication algorithm because our attack is based on creating a sampled acceleration signal, which mimics the victim's gait template signal.

We build a prototype for our attack model using vibration motors from Core Electronics [6] and a wearable device (Samsung Galaxy S6) shown in Fig. 5. The vibration motors are attached to the top of this victim device, and this victim device is suspended in the air with the help of a stand support and an elastic, which ensures the victim device can vibrate freely with the attached vibration motors in the vibration direction. Therefore, the vibration in this prototype can be modeled as a *spring-mass-undamped* model as discussed in Section III-A2. Each vibration motor is driven with an audio driver [6] from a laptop. Well-designed vibration signals based on analyzing the victim's gait template signals are generated in MATLAB, which are then sent to the corresponding vibration motors using the audio drivers through the laptop's audio jack. The number of vibration motors required to perform the attack equals the number of dominant frequencies in the victim's gait signal. If multiple vibration motors are utilized, we synchronize the multiple laptops with network time protocol (NTP) to ensure that the motors start vibrating simultaneously without significant phase shifts between multiple vibration signals.

B. Model Configuration

To efficiently attack the wearable gait authentication system, we configure our attack model based on the number of dominant frequencies in the victim's gait signal. We divide them into *two configuration cases*: one dominant frequency configuration and two or more dominant frequency peaks configuration. For the first case, victims' gait signals only have dominant-1 frequency in their gait signals. We only need one vibration motor and set the vibration frequency the same as dominant-1 frequency. In comparison, for victims whose gait signals have two or more dominant frequency peaks, we utilize two or more synchronized vibration motors, and each vibration motor generates a sinusoidal signal with a corresponding frequency peak. Specifically, we first decompose the gait template signal (see Section III-C), and obtain multiple single sinusoidal signals. Then, we reproduce these single signals using multiple vibration motors. Finally, these synchronized single-vibration signals will be sampled by the internal accelerometer in the wearable device (e.g., Galaxy S6) and produce a composite undersampled vibration signal. The performance of the attack model will be extensively evaluated in Section VI-A.

V. Liveness Detection-Embedded Defense Model

Biometrics, such as faces, fingerprints, and irises are widely used to identify users because they are convenient to use. However, biometric security is attracting increasing public concerns. Because of the rapid development in the recording and reproducing technologies, 3-D printing, wireless eavesdropping, and malware, biometric authentication is under various replay attacks, such as spoofing the user's face performed by [57], [97], [101], [102], and [103] reproducing the user's fingerprint [104], and recording and replaying the user's voice [59], [105]. An active research direction for preventing relay attacks is liveness detection. Specifically, in biometrics, liveness detection refers to a system's capacity to examine whether a presented biometric is genuine, i.e., taken from a live person present at the time of capture or not. To perform liveness detection in biometric authentication, user's motions or interactions are required. For example, Pan et al. [58] proposed a real-time face authentication liveness detection method by capturing eye blinks to prevent photograph-based attacks. Ray et al. exploited the head movements measured by the onboard IMU as the liveness detection feature in face-based authentication in [59]. Tang et al. [25] proposed to randomly display an image (i.e., a challenge) using the smartphone screen to a user during the authentication procedure and simultaneously use a camera to collect the light reflection (i.e., a response) and the face from the user for authentication and liveness detection purpose.

However, in our study, *Pistis* does not require the user's specific motion or interaction, instead, it utilizes bio-vibrometry of the human body. In this mechanism, we exploit onboard vibration motors of wearable devices to generate a vibration signal as the excitation, which will be propagated through the contacted medium. Then, it determines if the contacted medium is part of the human body (i.e., hand and wrist in the context of *Pistis*) or a nonliving medium. By doing so, our system can detect if the presented query biometric (i.e., gait) is genuine or reproduced by the vibration motor.

Specifically, we exploit *on-board* vibration motors of wearable devices to generate a vibration signal as the excitation, which will be propagated through part of the human body (i.e., hand and wrist in the context of *Pistis*) if the device is worn on a human or through a nonliving medium (e.g., the air in the set up of Fig. 5). The response of this excitation will be captured by the *on-board* accelerometer on the same device. The measured response depends on the mechanical characteristics, i.e., mass (m), stiffness (k), and damping (c), of the materials next to the wearable device, as discussed in Section III-A2.

We postulate that the physical features of humans are distinctive from nonhuman objects because the mechanical

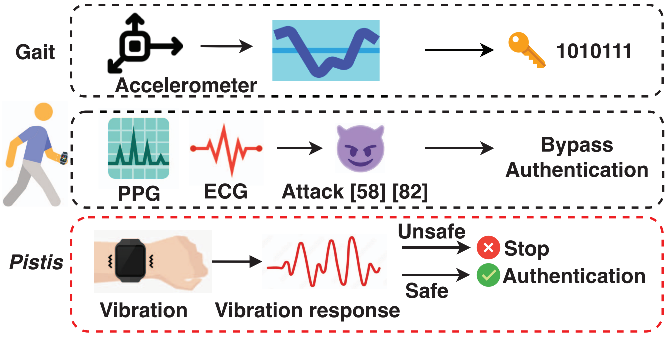


Fig. 6. Illustration of the proposed vibration-based liveness detection model, and comparison with other liveness detection methods, such as ECG and PPG.

properties of human body tissue are different from nonhuman entities. Specifically, the human skin comprises three primary layers: 1) the superficial epidermis; 2) the dermis; and 3) the deepest hypodermis [27], which have their unique viscosity, elasticity, extensibility, and structure that are different from other materials [28], [29]. Due to the unique viscosity and elasticity of live human skin tissue, the reflected vibration signals from the human skin differ widely that from nonhuman entities because of the significant difference in the properties of m, k, and c.

Compared with other liveness detection mechanisms for wearable-based gait authentication shown in Fig. 6, *Pistis* does not require any additional hardware or sensors and does not open extra potential attack surfaces. For example, the fusion of ECG and gait [67] not only requires an ECG monitor to be attached to the breast for ECG data collection but also brings new security concerns, e.g., spoof ECG [57]. We also considered other methods, e.g., photoplethysmography (PPG), which is an optical technique to detect blood volume changes in the microvascular bed of tissue, and skin electrical resistance, but they normally require additional sensors that may be not available in COTS wearable devices and PPG itself also has potential security issues [78]. To this end, the vibration motor and accelerometer being used in *Pistis* are ubiquitously available. Furthermore, the vibration response signal and gait signal are collected by the shared onboard accelerometer at the same time, which does not produce extra sampling overhead to wearable devices.

A. Integration of Pistis and Gait

As illustrated in Fig. 7, *Pistis*, as a liveness detection module, is integrated with the existing wearable gait authentication system to protect it from *vibration replay attack*s. Therefore, it is *a novel and more secure authentication protocol*, which comprises three primary phases: 1) data acquisition; 2) data separation; and 3) liveness detection and authentication.

Phase 1 (Data Acquisition): When the user activates the authentication application, the accelerometer on the user's wearable device will start sampling. The sampled acceleration is the combined effect of the user's walking and the response to the excitation due to the onboard vibration motor. Similar to prior work [2], our authentication protocol samples the accelerometer at 100 Hz, and the sampling window is set as 5 s to collect at least three complete gait cycles because the duration of a gait cycle is usually in the range of 0.8–1.3 s.

Phase 2 (Data Separation): Since the collected acceleration signal is a superimposition of gait and vibration excitation responses, we need to separate them before further processing. Our gait analysis in Section III shows that the human gait signal is usually in the frequency range of 0.5–3 Hz.

However, the vibration produced by the internal vibration motor frequency is in the range of 130–180 Hz [17], which requires a sampling rate of more than 100 Hz, and is very challenging for the resource-constraint COTS wearable devices [30]. Based on the under-sampling model, the vibration signal will have an aliasing problem. Specifically, after sampling at 100 Hz in our prototype, the signal's frequency will be in the range of 20–50 Hz. Nevertheless, we can exploit the difference in the frequency range to separate these two different signals. In our protocol design, we utilize two bandpass filters to the raw acceleration data to separate gait and vibration signals: one operates between 20 and 50 Hz for the vibration signal and the other operates between 0.5 and 3 Hz for the gait signal.

Phase 3 (Liveness Detection and Authentication): Since the vibration motors in COTS wearable devices usually operate from 130 to 180 Hz and the maximum sampling rate of the accelerometer in these devices is approximately 100 Hz, the vibration signal is under sampled. Our experience of using the sampled signal directly for liveness detection is that the accuracy is low. Instead, we propose to use an ℓ_1 *minimization* technique to map the sampled vibration signal to a higher frequency range and we find that this method produces much better accuracy. ℓ_1 *minimization* is a signal processing technique to recover high-dimension information from low-dimension measurements, and we will discuss the details of ℓ_1 *minimization* in Section V-B.

After obtaining the full vibration signal, the straight path vibration will be eliminated from the reconstructed vibration signal, before passing into the liveness detection module, where both the time domain and frequency domain features as shown in Table I will be extracted. Finally, an SVM model is applied to detect whether the vibration response signal comes from a live human or nonhuman entity.

To train the SVM classifier, we first develop a feature extractor to extract the features we discussed in Table I from full vibration signals which are recovered from ℓ_1 minimization. After that, we take the features as the input to the SVM model to do training and evaluation. To balance the training data set, we set the ratio of the number of vibration signals from humans and nonhumans as 1:1, and the ratio of samples used for training and testing as 8:2. We employ 5-fold cross-validation to validate our SVM model. If the live human is detected from *Pistis*, our protocol will proceed to the authentication stage with the corresponding gait data.

B. Key Techniques

ℓ_1 Minimization: Our proposed protocol uses the sampled vibration signal to perform liveness detection. The vibration

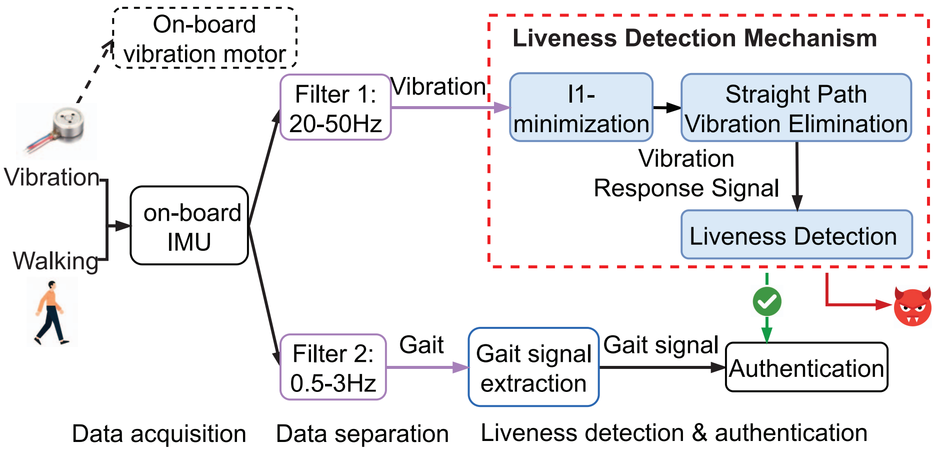


Fig. 7. Architecture of integrating gait authentication with liveness detection *Pistis*. Vibration and gait acceleration data will be measured simultaneously, then they will be separated and used for liveness detection and authentication, respectively.

TABLE I
FEATURES USED IN THE PROPOSED APPROACH

Features	Definition
RMS	Root mean square
Std	Standard deviation
Var	Variance
Skewness	Skewness of frequency domain
MAD	Median absolute deviation
Kurtosis	Frequency signal kurtosis
IQR	Interquartile range (i.e., 25%, 75%)
IAV	Integral absolute value
Energy	Average sum of the squares
Entropy	Signal entropy
Zero Crossing	Zero crossing points in the signal
Frequency Peaks	Peaks in the frequency domain
SMA	Signal magnitude area

motor produces a vibration signal in the range of 130–180 Hz [17], [30]. Since the sampling rates of the accelerometer in wearable devices are limited (e.g., up to 100 Hz [30]), the vibration signal is undersampled. To overcome the vibration undersampling problem in COTS wearable devices, we develop a signal mapping method by using ℓ_1 minimization for signal reconstruction in compressive sensing. Specifically, we propose to use ℓ_1-minimization to map the vibration signal from 20–50 Hz to either 100–150 or 150–200 Hz range.

One can work out that the sampled vibration signal is in the 20–50 Hz range. We attempted to use this sampled vibration signal directly for liveness detection but the performance is poor. We found that we could obtain 100% liveness detection accuracy by first *mapping* this signal to either the 100–150 or 150–200 Hz range via ℓ_1-minimization. This mapping step enables us to remove some spurious low-frequency components which may be present in the sampled vibration signal and cause poor liveness detection performance.

Our signal mapping method is inspired by using ℓ_1 minimization for signal reconstruction in compressive sensing. We will first explain the method by assuming the signal generated by the vibration motor is in the range 130–150 Hz. This is because we can interpret our mapping method as signal reconstruction if this is the case. We will explain what happens if the vibration motor signal is in the 130–180 Hz range later on.

In order to describe the mapping method in a concrete way, we assume the vibration motor signal $s(t)$ is a sum of n sinusoids with amplitudes a_i, frequencies f_i and phases ϕ_i

$$s(t) = \sum_{i=1}^{n} a_i \ \sin(2\pi f_i t + \phi_i) \tag{9}$$

where $f_i \in [130, 150]\ \forall i$. Let F_s be the sampling frequency of accelerometer which is 100 Hz, and $T_s = 1/F_s$ is the sampling time. Let M be the number of samples collected. The sampled (discrete time) signal y_k $(k = 0, \ldots, M-1)$ is therefore $y_k = s(kT_s)$. It can be shown that the undersampled signal y_k has frequency components $\tilde{f}_i = f_i - F_s$

$$y_k = \sum_{i=1}^{n} \tilde{a}_i \ \sin\Big(2\pi \tilde{f}_i k T_s + \tilde{\phi}_i\Big) \tag{10}$$

where $\tilde{a}_i = a_i$ and $\tilde{\phi}_i = \phi_i$. Note that even though undersampling has caused aliasing, we can use y_k to reconstruct $s(t)$. An intuitive method is to estimate $\tilde{a}_i$, $\tilde{f}_i$, and $\tilde{\phi}_i$ from y_k, and then mapped them to the corresponding parameter in $s(t)$. Here, we will use ℓ_1-optimization, which is a more robust method. Our method will reconstruct a sampled version of $s(t)$ at a higher sampling frequency F_h, where F_h is at or above the Nyquist rate of $s(t)$. For simplicity, we assume F_h is h times of F_s, where h is an integer. Let also $T_h = 1/f_h$. If $s(t)$ is sampled at rate F_h, the sampled signal is

$$z_k = \sum_{i=1}^{n} a_i \ \sin(2\pi f_i k T_h + \phi_i) \tag{11}$$

where $k = 0, \ldots, (M-1)h$. The signal z_k is the desired output of our ℓ_1-minimization. A key part of our ℓ_1 formulation is the optimization constraint. Since we will use the discrete cosine transform (DCT) basis in ℓ_1-minimization, we will first examine how we can determine the DCT coefficients of the vector signal $\vec{z}$, which is formed by stacking z_k into a column vector. Let D_{full} be the inverse DCT matrix of dimension M_h-by-M_h, where $M_h = (M-1)h + 1$. The DCT coefficients $\vec{c}$ of the signal $\vec{z}$ can be computed from $D_{\text{full}}\vec{c} = \vec{z}$. We will use the last equality to help us to arrive at a constraint in ℓ_1 minimization.

Recall that the rows of D_{full} correspond to the time samples and the columns of D_{full} correspond to frequencies. Since the 1st, $(h+1)$th, $(2h+1)$th samples of z_k correspond to the samples of y_k, we will use only the corresponding rows in D_{full}. In addition, we want the reconstructed signal to be in the 100–150 Hz range, we will only select those columns in D_{full} which corresponds to this frequency range. We will use D_{sub} to denote the matrix that is formed by the elements in the intersection of the selected rows and columns of D_{full}. Let θ be a vector decision variable, where the number of elements in θ is equal to the number of columns in D_{sub}. The ℓ_1-minimization problem for reconstructing the original $s(t)$ is

$$\hat{\theta} = \arg\min \|\theta\| \text{ subject to } D_{\text{sub}}\theta = \vec{y} \tag{12}$$

where $\vec{y}$ is a column vector formed by stacking the samples y_k in (10) in a column. Once we have computed $\hat{\theta}$, we can obtain an estimate of z_k as follows. Let $D_{100:150}$ be formed by extracting the columns of D_{full} that correspond to the frequencies in 100–150 Hz, then the estimate for $\vec{z}$ is $D_{100:150}\hat{\theta}$. Our numerical experiments show that this reconstruction is accurate if the number of frequency components in $s(t)$ is small.

We now return to the case, where the frequency components f_i in $s(t)$ in (9) are in the range 130–180 Hz. In this case, the sampled signal is also of the form (10) but we have $\tilde{f}_i$ equals to $f_i - F_s$ if $f_i \in [130, 150]$ and $2F_s - f_i$ if $f_i \in [150, 180]$, $\tilde{\phi}_i$ equals to ϕ_i if $f_i \in [130, 150]$ and $\pi - \phi_i$ if $f_i \in [150, 180]$, and $\tilde{a}_i = a_i$. This means that undersampling will cause the frequency components below 150 Hz to cancel out those above 150 Hz. This is certainly undesirable but fortunately we did not observe serious annihilation in our signal because the vibration signal produced by the motor seems to be concentrated in a very narrow band below or above 150 Hz. We note that even with some signal annihilation taking place, the ℓ_1-minimization will map a sampled signal y_k in (10) to a unique $\hat{\theta}$ provided that the restricted isometry property (RIP) holds. Unfortunately, RIP is computationally demanding to verify but our numerical experience shows that an $s(t)$ is mapped to an accurate estimate of $\hat{\theta}$. Note that, instead of reconstructing the signal in the 100–150 Hz range, it is also possible to reconstruct the signal in 150–200 Hz range. We found that, with real-life data, liveness detection on either one of these frequency ranges produces 100% accuracy (see Section VI-B later for the details).

Based on the above theory, we develop a ℓ_1-minimization algorithm. To verify the performance of the ℓ_1-minimization algorithm, a standalone accelerometer (MPU9250) has been utilized to produce the original vibration signal that is not undersampled (i.e., ground truth).

When the smartphone on-board vibration motor starts vibrating, MPU9250, and smartphone on-board accelerometer will measure the vibration signal simultaneously with a sampling frequency of 500 and 100 Hz, respectively. We take the measurements from MPU9250 as the ground truth of the original vibration signal, while the measurements from the smartphone onboard accelerometer as the undersampled vibration signal which is to be reconstructed. By using 60 randomly produced vibration signals, our reconstructed signals have an average Root Mean Square Error is 6.69% only. Fig. 8 shows an example reconstruction.

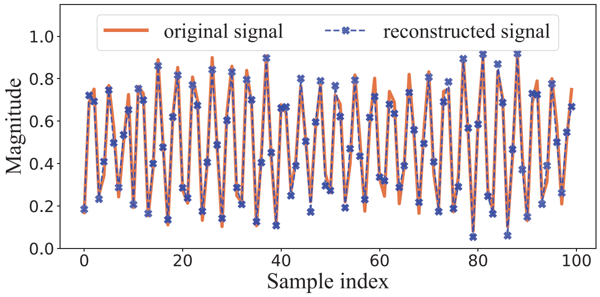


Fig. 8. Comparison of the reconstructed vibration signal and the original vibration signal.

Elimination of Vibration Straight Path Interference: When the vibration motor in a wearable device is operating, part of the generated vibration will reach the human body and bounce back. It will be captured by the accelerometer in the wearable device, which can be considered as the vibration response signal. However, a significant part of the generated vibration will directly reach the internal accelerometer, which is known as vibration straight path interference. Since the energy of the straight path signal is significantly higher than the vibration response signal from the human body, we utilize the straight path interference elimination method in [30] to reduce the energy of the straight path signal.

VI. Evaluation and Results

In this section, we present results on evaluating the proposed vibration attack model, the liveness detection module *Pistis*, and the combination of gait authentication and *Pistis*. We also study the viability of using vibration for authentication and present a security analysis of our proposed authentication-and-liveness detection model.

A. Attack Model Evaluation

We first explain the existing wearable gait recognition and authentication models. Subsequently, we present the performance of our vibration replay attack on these models.

Many machine learning methods have been proposed for gait patterns/signals-based authentication in recent years. They can be generally divided into two main groups: one is to construct a gait template profile from the accelerometer data directly and to score the similarity between the template and test samples [7], e.g., Pearson correlation coefficient (PCC) and dynamic time warping (DTW). The other uses feature engineering methods to produce informative features from a particular user's gait pattern and build classifiers, e.g., SVM, convolutional neural network (CNN), and long short-term memory (LSTM) neural networks.

To demonstrate the effectiveness of our attack model, we implemented SVM, CNN, LSTM, and their optimization strategies using state-of-the-art methods. Specifically, we implemented the SVM-based model and its optimization strategy in [8], which authenticates users by walking acceleration cycles. Additionally, we follow [42] to implement their LSTM-based gait authentication model and use the same parameters and their values.

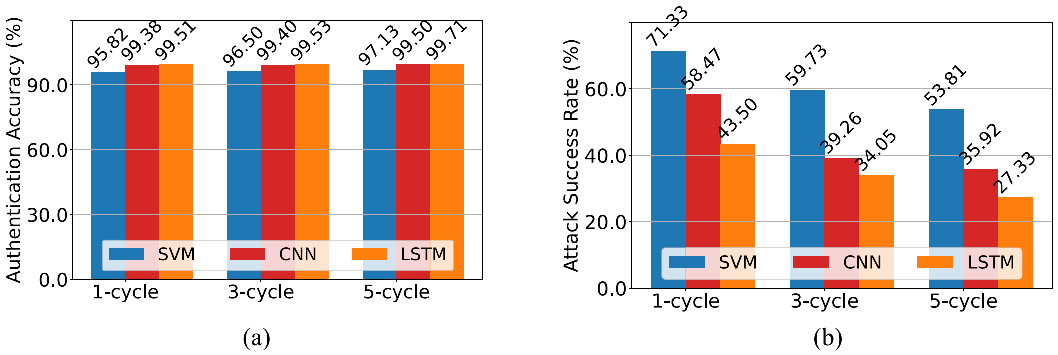


Fig. 9. Authentication accuracy of SVM-based, CNN-based, and LSTM-based gait authentication models, and the success rates of our proposed replay attack. (a) Model performance. (b) Attack success rate.

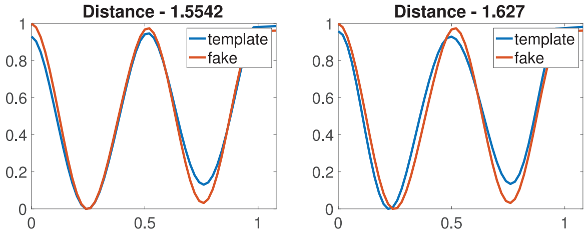


Fig. 10. Successful attack cases. The well-modulated vibration produced fake gait cycles which are very similar to the real gait template cycles.

With our 50 subjects' data set, all the models can reach a high-level authentication accuracy as shown in Fig. 9(a) (e.g., the minimum accuracy of 96.5% when three gait cycles are used). The authentication performance of our models is similar to those of state-of-the-art approaches [41], [42].

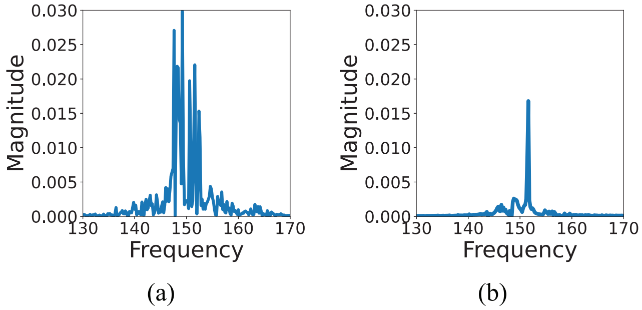
We set up the attack model prototype as illustrated in Fig. 5. We first extract some of the gait cycles of a "victim" as the gait template signals from our data set before obtaining the major frequency peaks. If the user's gait template signal has multiple major frequency peaks, we then utilize a signal decomposition algorithm (i.e., iterative filtering (IF), which is a signal decomposition algorithm that decomposes a signal into multiple simpler components sequentially [79]) to find its sinusoidal components. Finally, we attempt to modulate and generate the vibration signal, which is similar to a victim's gait signal, using the external vibration motors.

We take the 50 subjects in our data set as victims and apply the proposed vibration attack model to them. We create 60 "fake" cycles for each subject and assess the attack performance. Here, we find that most of the reproduced cycles are similar to gait template signals (Fig. 10 shows two examples). Furthermore, we input all the reproduced signals to gait authentication systems/models to assess how many fake signals will be misidentified as the genuine user. Here, we study the attack success rates when the attack model is used on different authentication models and the impact of the number of gait signal cycles on the attack success rates.

1) *Different Models:* We input the fake gait cycles generated above to one cycle SVM [8], CNN, and LSTM-based authentication systems, and assess the number of fake gait cycles that can fool the systems. Compared with the SVM-based authentication system with an attack success rate of 71.33%, the attack success rates for CNN and LSTM-based authentication systems reduce moderately to 58.47% and 43.5%, respectively. Our conjecture is that the features generated in the CNN and LSTM-based authentication models automatically are more difficult to attack than those used by the SVM-based models, which affects the attack success rates.
2) *Different Number of Gait Cycles:* Intuitively, a longer gait signal (more number of gait cycles) contains more contextual information and may make it more difficult to be attacked. Here, we investigate the attack success rate for *one, three, and five* gait cycle-based authentication models. As shown in the results in Fig. 9(b), for the SVM-based authentication system, the attack success rate of the one-cycle-based model can reach 71.33%, while for three and five-cycle-based systems, they decrease to 59.73% and 53.81%, respectively. Similarly, for CNN and LSTM-based authentication models, the attack success rates decrease when we increase the number of gait cycles used for authentication. When we generate fake cycles using vibration motors, all the fake cycles are the same, which provides no connection contextual information between two neighboring gait cycles. The gait-based authentication models may make use of this contextual information to distinguish different users, which makes them more difficult to be attacked than those without the contextual information (i.e., one-cycle gait authentication models). Nevertheless, the minimum attack success rate is still very high (i.e., 27.33% with a five-cycle LSTM model), and justified the motivation of our liveness detection module, i.e., *Pistis*, which will be evaluated next.

B. Liveness Detection Evaluation

In this section, we evaluate the performance of the proposed liveness detection mechanism (i.e., *Pistis*). The goals for this evaluation are threefold: 1) to evaluate the accuracy of *Pistis* in differentiating live humans from nonhuman entities; 2) if the integrated authentication protocol can reach high authentication accuracy and low EER; and 3) if the



(a)

(b)

Fig. 11. Differences in the spectrum of human and nonhuman entities. (a) Typical vibration response from humans. (b) Typical vibration responses from nonhumans include various meats (e.g., beef, chicken, and mutton), rubber, silicone, carpet, soft mats, etc.

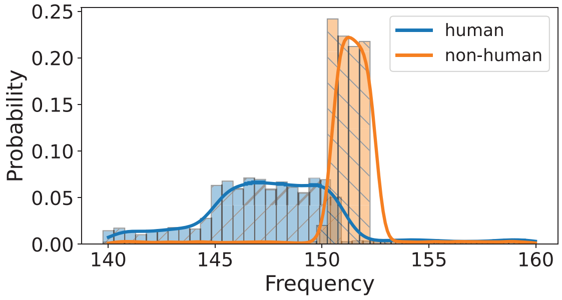


Fig. 12. Frequency distribution of the responses to vibration from the human group (50 subjects) and nonhuman group. The frequencies of human response signals are mainly in the range of 145–150 Hz, but those of nonhuman entities' response signals are mainly in the range of 150–152 Hz.

proposed authentication protocol can efficiently counter the vibration replay attack and other attacks (i.e., security analysis). Additionally, we use 5-fold cross-validation with our 50 subjects' gait data set, and the following metrics to evaluate our authentication protocol.

1) *Authentication Accuracy:* It represents the percentage of the correct number of authentication sessions to the total number of all authentication sessions.
2) *Equal Error Rate:* It represents the intersection in the decision error tradeoff (DET) curve, where false positive rates (FPRs) and false negative rates (FNRs) are equal. In general, the lower the EER value is, the better the security and usability of the authentication system will be. When denote $\text{Sensitivity} = (\text{TP}/\text{TP} + \text{FN})$, $\text{Specificity} = (\text{TN}/\text{TN} + \text{FP})$, the EER can be obtained by

$$\text{EER} = 1 - 0.5 \times (\text{sensitivity} + \text{specificity}). \quad (13)$$

To investigate the performance of *Pistis*, we consider air, various tables, metals, various types of meats, rubbers, carpets, soft mats, and other daily materials that we can access as the nonhuman group, while the 50 subjects as the human group. The number of samples we collected from the human group and nonhuman group are 2000 and 1000, respectively. As shown in Fig. 11 and 12, the vibration responses and signal frequency distributions from humans and nonhumans are quite distinctive. There are multiple frequency peaks with different magnitudes in the frequency spectrum for the human group [Fig. 11(a)], which represent the damping effect when the vibration propagates through human bodies. On the contrary, there is only one prominent frequency peak for the nonhuman group [Fig. 11(b)]. Furthermore, the frequency range of vibration signal responses from human bodies (i.e., 145–150 Hz in Fig. 12) is different from that of nonhuman items (i.e., 150–152 Hz in Fig. 12). Therefore, we implement a binary SVM-based model with an RBF kernel to evaluate the human and nonhuman differentiation with time-domain and frequency-domain features (i.e., variance, median, entropy, and root mean square) extracted from vibration responses, and our results show that it can reach 100% differentiation accuracy.

In addition, we also study the impact of wearable devices' location on the human body, e.g., wrist for smartwatch users. We collect the vibration responses by randomly shifting the smartwatch within ±1 cm to the baseline location where the smartwatch was first worn. Even with this shift, our liveness detection mechanism can still differentiate humans from nonhumans with 100% accuracy. Therefore, we conclude that vibration response is a suitable identifier used to differentiate humans from nonhumans accurately.

C. Overall Authentication Performance

As discussed earlier in Fig. 7, the authentication part in the proposed authentication protocol will consider outputs from both liveness detection (i.e., *Pistis*) and gait verification modules. For the gait verification, we follow the-state-of-art gait recognition implementation as discussed in Fig. 1. Here, we utilize the SVM, CNN, and LSTM to train the model based on 50 subjects. Additionally, we consider the intrinsic inter-relationships of two continuous gait cycles and connect contextual information between them. Specifically, every *three* continuous gait cycle is utilized as the input. The optimal values of parameters of the SVM model are learned in the training phase, and the neural network model parameters and their values for CNN and LSTM are the same as [42]. With 50 subjects, we first train the authentication models for each individual and then calculate the overall average authentication accuracy and EER. Our results show that the overall classification/recognition accuracy is 96.5% (SVM), 99.4% (CNN), and 99.53% (LSTM) as shown in Fig. 9(a). Finally, the average EER of the proposed protocol with LSTM is 6.91% (which is similar to that (7.27%) of the state-of-the-art [7]).

D. Vibration Authentication Evaluation

Recent work proposed a vibration authentication model that exploits Apple Watches to authenticate users directly based on Random Forest and SVM, and manually chosen features such as the root mean square, variance, and mean absolute value of vibration signals [45]. Their evaluation of six subjects showed that they can achieve approximately 99% accuracy. We implemented the models proposed [45] and evaluated their performance with our data set. Our results show that they can produce approximately 98% (similar to that presented in [45]), 91%, and 53% authentication accuracy for 6, 10, and 20 subjects, respectively. Similarly, vibration work in [80] also can only work with a small group of users, e.g., less than 10. Our results also show that the vibration-only authentication model is not scalable to a large number of users; therefore, we use it for the liveness detection module (i.e., *Pistis*) only.

E. Security Analysis and Evaluation

Pistis in essential can be considered as a challenge-response protocol. Here, *the Challenge* is the vibration excitation generated from the wearable device's *on-board* vibration motor, and *the Response* is the vibration signal reflected from contacted entities. Its security mechanism is guaranteed by two elements: 1) difficult to forge vibration response and 2) difficult to break the synchronization between vibration response and corresponding gait signal. The vibration response from human bodies contains mechanical and biological features, which are difficult to be forged by other materials. In addition, in the proposed authentication protocol, vibration response, and gait signals are collected simultaneously when there is an authentication request, such that vibration response and gait signals are synchronized. Therefore, the adversary may reproduce the gait signal with the external vibration motors, but it is difficult for the adversary to compromise the corresponding synchronized vibration response signal from the victim.

We try to consider all the existing attack methods for the gait authentication system. To the best of our knowledge, *Imitation attack* [7], [8], [54], [55], *in-wear vibration attack* [97], and our proposed *vibration attack* are all the existing attack methods that may threaten the gait authentication system and liveness detection mechanism. To answer how secure is the proposed scheme, we analyze the resistance of liveness detection to all the possible attacks.

Imitation Attack: Gafurov et al. [54] asked attackers to imitate victims' walking manner without any feedback. Specifically, they collected 760 gait sequences from 100 subjects, and conduct two stages for their experiments. In the first experiment stage, subjects walked in their usual manner; while in the second experiment stage, subjects attempted to walk in victims' gait patterns for imitation attack. In comparison, Mjaaland et al. [55] trained the attackers when the attackers were imitating the victims with feedback. However, their results presented that imitation attacks were very difficult to launch because it is challenging for the attackers to imitate the victims' walking style accurately. Therefore, based on experiments from previous studies, we conclude that the proposed *Pistis* can resist the imitation attack.

In-Wear Vibration Attack: For this attack scenario, an attacker wears the wearable device on his/her body (e.g., wrist) and maintains static posture (i.e., not walking), so that *Pistis* can obtain vibration response from a human body part. In the meantime, the attacker attaches vibration motors on top of the wearable device to generate the gait patterns of the victim. However, our vibration attack model exploits the *spring-mass-undamped* model, in which the damping coefficient c and damping ratio ζ are equal to zero. To accurately realize the spring-mass-undamped model, the wearable device must be suspended in the air to remove the constraints and damping in the vibration direction of vibration motors. Otherwise, the vibration signal will be significantly distorted and it will be difficult to produce gait signals that are close to those of the victim. Fig. 13(a) shows an example that the (human body) constraints significantly attenuate the vibration and make it significantly different from the victim's gait signal. Here, the similarity between the vibration attack signal and the victim's gait signal is quantified by the Euclidean distance calculated from the DTW algorithm. We have also produced 200 attack signals randomly, and the average distance between vibration signals in the in-wear attacks and the victim's gait signals is very large (8.25). These observations show that it is very difficult to launch a successful in-wear vibration attack.

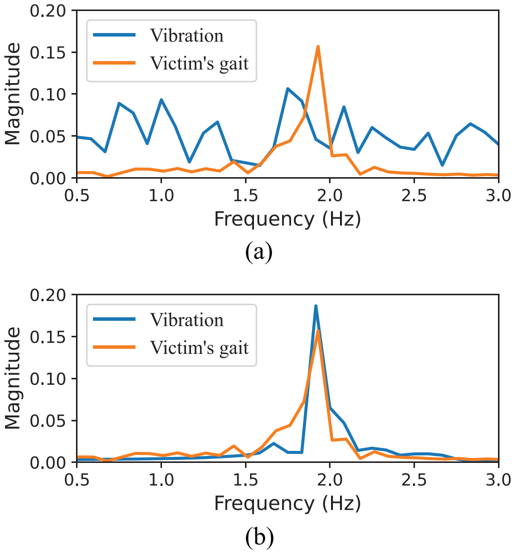


Fig. 13. Impact of positions to place the smartphone to the vibration attack model. Our attack is based on the spring-mass-undamped vibration model which requires the damping coefficient c and damping ratio ζ to be equal to zero. When the vibration motors and victim's smartphone are placed on an attacker, the constraints from the attacker's wrist/hand will distort the vibration signal. (a) Smartphone place on a static attacker. The distance between vibration attack signal and that of victims gait is 7.309. (b) Smartphone suspend in the air. The distance between the vibration attack signal and that of the victims gait is 1.431.

Vibration Attack: We apply the proposed vibration attack on our liveness detection protocol with the device suspended in the air. Since the reflected vibration signal comes from the external vibration motors and the wearable device itself instead of the human body, the vibration response signal is similar to Fig. 11(b). Intuitively, *Pistis* can efficiently counter this type of attack, and our results show that all the fake cycles are indeed rejected by the liveness detection mechanism with 100 attack sessions.

To the best of our knowledge, these three attacks are all the possible attack methods, and based on the above analysis we can see that the proposed authentication methods can efficiently protect the gait to be attacked.

VII. Conclusion and Future Work

In this article, we propose an efficient *vibration replay attack* model by exploiting the widely available COTS vibration motors and fundamental signal processing theories. We demonstrate that it is feasible to reproduce the gait template of a victim and attack wearable gait authentication systems. Furthermore, to defend against such vibration attacks, we introduce *Pistis*, which is a novel liveness detection module with onboard vibration motors ubiquitously available in wearable devices by exploiting human body tissue's unique bio-vibrometry. We integrate *Pistis* with gait biometrics as a novel authentication protocol and systematically evaluate it with a data set of 50 subjects, which shows it is both robust against different attacks and highly accurate. Specifically, the accuracy of human and nonhuman detection with vibration response is 100%, and the authentication accuracy is 99.53%.

Limitations and Future Work: There are mainly two limitations in our current work.

1) despite the great efficiency of the attack model, we must suspend the victim's wearable device in the air when practising attacking, which could cause small inconvenience for attacking. In detail, our attack model exploits the spring-mass-undamped model, in which the damping coefficient c and damping ratio ζ are all equal to zero. Thus, to accurately implement the spring-mass-undamped model, the wearable device must be suspended in the air to remove the constraints and damping in the vibration direction of vibration motors. Otherwise, the vibration signal will be quite noisy and hard to produce gait similar signals as shown in Fig. 13, which is the resulting signal when the vibration motor and smartphone are placed on a table ($\zeta \simeq 1$). Also, the table will significantly attenuate the vibration and make the vibration signal quite noisy, which challenges the successful vibration attack.
2) Our work explored the vibration attack method for the gait authentication system which is based on z direction accelerations. However, some gait authentication works may use accelerations of multiple directions at the same time. The attacking performance, in this case, however, is unclear to us. Therefore, how to improve the currently proposed vibration attack (i.e., not only working when the suspend victim's wearable device in the air) and how to apply the vibration attack on the gait authentication systems based on multiple direction accelerations is important and we will leave these questions as our future works.

REFERENCES

[1] W. Xu, Y. Shen, Y. Zhang, N. Bergmann, and W. Hu, "Gait-Watch: A context-aware authentication system for smart watch based on gait recognition," in *Proc. IoTDI*, 2017, pp. 59–70. [Online]. Available: https://doi.org/10.1145/3054977.3054991

[2] W. Xu, Y. Shen, C. Luo, J. Li, W. Li, and A. Y. Zomaya, "Gait-Watch: A Gait-based context-aware authentication system for smart watch via sparse coding," *Ad Hoc Netw.*, vol. 107, Oct. 2020, Art. no. 102218. [Online]. Available: https://doi.org/10.1016/j.adhoc.2020.102218

[3] A. Ferlini, D. Ma, R. Harle, and C. Mascolo, "EarGate: Gait-based user identification with in-ear microphones," in *Proc. 27th Annu. Int. Conf. Mobile Comput. Netw.*, 2021, pp. 337–349, doi: 10.1145/3447993.3483240.

[4] L. Middleton, A. A. Buss, A. Bazin, and M. S. Nixon, "A floor sensor system for gait recognition," in *Proc. IEEE Workshop Autom. Identification Adv. Technol.*, 2005, pp. 171–176. [Online]. Available: https://doi.org/10.1109/autoid.2005.2

[5] Y. Luo, S. M. Coppola, P. C. Dixon, S. Li, J. T. Dennerlein, and B. Hu, "A database of human gait performance on irregular and uneven surfaces collected by wearable sensors," *Sci. Data*, vol. 7, p. 219, Jul. 2020. [Online]. Available: https://doi.org/10.1038/s41597-020-0563-y

[6] "Bone conduction kit," Accessed: Oct. 20, 2021. [Online]. Available: https://core-electronics.com.au/bone-conduction-kit.html

[7] T. Zhu, L. Fu, Q. Liu, Z. Lin, Y. Chen, and T. Chen, "One cycle attack: Fool sensor-based personal gait authentication with clustering," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 553–568, 2021. [Online]. Available: https://doi.org/10.1109/tifs.2020.3016819

[8] Y. Ren, Y. Chen, M. C. Chuah, and J. Yang, "User verification leveraging gait recognition for smartphone enabled mobile healthcare systems," *IEEE Trans. Mobile Comput.*, vol. 14, no. 9, pp. 1961–1974, Sep. 2015, [Online]. Available: https://doi.org/10.1109/tmc.2014.2365185

[9] N. Al-Naffakh, N. Clarke, F. Li, and P. Haskell-Dowland, "Unobtrusive gait recognition using smartwatches," in *Proc. Int. Conf. Biometr. Special Interest Group*, 2017, pp. 1–5. [Online]. Available: https://doi.org/10.23919/biosig.2017.8053523

[10] D. Ma, G. Lan, W. Xu, M. Hassan, and W. Hu, "Simultaneous energy harvesting and gait recognition using piezoelectric energy harvester," *IEEE Trans. Mobile Comput.*, vol. 21, no. 6, pp. 2198–2209, Jun. 2022.

[11] T. T. Ngo, Y. Makihara, H. Nagahara, Y. Mukaigawa, and Y. Yagi, "The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication," *Pattern Recognit.*, vol. 47, no. 1, pp. 228–237, 2014.

[12] "Dynamic time warping." Accessed: Oct. 21, 2021. [Online]. Available: https://en.wikipedia.org/wiki/Dynamic_time_warping

[13] "Vibrator (mechanical)." Accessed: Oct. 21, 2021. [Online]. Available: https://en.wikipedia.org/wiki/Vibrator_(mechanical)

[14] "Damping." Accessed: Nov. 30, 2021. [Online]. Available: https://en.wikipedia.org/wiki/Damping

[15] X. Xu et al., "TouchPass: Towards behavior-irrelevant on-touch user authentication on smartphones leveraging vibrations," in *Proc. 26th Annu. Int. Conf. Mobile Comput. Netw.*, 2020, pp. 1–13.

[16] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge, U.K.: Cambridge Univ. Press, 2019.

[17] J. Yim, R. Myung, and B. Lee, "The mobile phone's optimal vibration frequency in mobile environments," in *Usability and Internationalization. HCI and Culture* (LNCS 4559). Heidelberg, Germany: Springer, 2007, pp. 646–652, doi: 10.1007/978-3-540-73287-7_75.

[18] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[19] J. Wang, S. Kwon, and B. Shim, "Generalized orthogonal matching pursuit," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6202–6216, Dec. 2012.

[20] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4680–4688, Jul. 2011.

[21] K. Schnass, "Average performance of orthogonal matching pursuit (OMP) for sparse approximation," *IEEE Signal Process. Lett.*, vol. 25, no. 12, pp. 1865–1869, Dec. 2018.

[22] D. Needell and J. A. Tropp, "CoSaMP: iterative signal recovery from incomplete and inaccurate samples," *Commun. ACM*, vol. 53, no. 12, pp. 93–100, 2010.

[23] S. Vujović, M. Daković, and L. Stanković, "Comparison of the L1-magic and the gradient algorithm for sparse signals reconstruction," in *Proc. IEEE 22nd Telecommun. Forum Telfor (TELFOR)*, Nov. 2014, pp. 577–580.

[24] E. E. Alaa, A. S. Ashour, Y. Guo, and H. M. Kasem, "A novel weighted compressive sensing using L1-magic recovery technique in medical image compression," *Health Inf. Sci. Syst.*, vol. 8, no. 1, pp. 1–10, 2020.

[25] D. Tang, Z. Zhou, Y. Zhang, and K. Zhang, "Face flashing: A secure liveness detection protocol based on light reflections," in *Proc. Netw Distrib. Syst. Security Symp.*, 2018, pp. 1–19. [Online]. Available: https://doi.org/10.14722/ndss.2018.23176

[26] W. E. Siri, "The gross composition of the body," *Adv. Biol. Med. Phys.*, vol. 4, nos. 239–279, p. 513, 1956.

[27] L. Zhou, S. Wang, J. Zhang, J. Wang, and C. Li, "In vivo measurement of the anisotropic mechanical properties of human skin by indentation test," *Mech. Mater*, vol. 158, Jul. 2021, Art. no. 103851. [Online]. Available: https://doi.org/10.1016/j.mechmat.2021.103851

[28] V. R. Sherman, Y. Tang, S. Zhao, W. Yang, and M. A. Meyers, "Structural characterization and viscoelastic constitutive modeling of skin," *Acta Biomater.*, vol. 53, pp. 460–469, Apr. 2017. [Online]. Available: https://doi.org/10.1016/j.actbio.2017.02.011

[29] Y. A. Kvistedal and P. M. F. Nielsen, "Estimating material parameters of human skin in vivo," *Biomech. Model. Mech.*, vol. 8, no. 5, pp. 1–8, 2009. [Online]. Available: https://doi.org/10.1007/s10237-007-0112-z

[30] Y. Huang, K. Chen, Y. Huang, L. Wang, and K. Wu, "Vi-liquid: Unknown liquid identification with your smartphone vibration," in *Proc. 27th Annu. Int. Conf. Mobile Comput. Netw.*, 2021, pp. 174–187. [Online]. Available: https://doi.org/10.1145/3447993.3448621

[31] "Analog." Accessed: Oct. 28, 2021. [Online]. Available: https://www.analog.com/en/products/adxl335.html#product-overview,

[32] T. Tomko, M. Puskar, M. Fabian, and R. Boslai, "Procedure for the evaluation of measured data in terms of vibration diagnostics by application of a multidimensional statistical model," *Sci. J. Silesian Univ. Technol. Transp.*, vol. 91, no. 1, pp. 125–131, 2016. doi: 10.20858/sjsutst.2016.91.13.

[33] S. Fábry and M. Češkovič, "Aircraft gas turbine engine vibration diagnostics," *Mag. Aviat. Dev.* vol. 5, no. 4, pp. 24–28, 2017. [Online]. Available: https://doi.org/10.14311/mad.2017.04.04

[34] T. Hachisu, G. Cirio, M. Marchal, A. Lecuyer, and H. Kajimoto, "Pseudo-haptic feedback augmented with visual and tactile vibrations," in *Proc. IEEE Int. Symp. VR Innov.*, 2011, pp. 327–328. [Online]. Available: https://doi.org/10.1109/isvri.2011.5759662

[35] M. B. Rosson, D. Gilmore, S. Brewster, F. Chohan, and L. Brown, "Tactile feedback for mobile interactions," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2007, pp. 159–162. [Online]. Available: https://doi.org/10.1145/1240624.1240649

[36] "Practical Applications for Frictionless Authentication." Accessed: Oct. 21, 2021. [Online]. Available: https://unify.id/use-cases/

[37] M. De Marsico and A. Mecca, "A survey on gait recognition viawearable sensors," *ACM Comput. Surveys*, vol. 52, no. 4, pp. 1–39, 2019.

[38] M.-S. Axente, C. Dobre, R.-I. Ciobanu, and R. Purnichescu-Purtan, "Gait recognition as an authentication method for mobile devices," *Sensors*, vol. 20, no. 15, p. 4110, 2020.

[39] A. H. Johnston and G. M. Weiss, "Smartwatch-based biometricgait recognition," in *Proc. IEEE 7th Int. Conf. Biometr. Theory Appl. Syst. (BTAS)*, 2015, pp. 1–6.

[40] H. J. Ailisto, M. Lindholm, J. Mantyjarvi, E. Vildjiounaite, and S. M. Makela, "Identifying people from gait pattern with accelerometers," in *Proc. Biometr. Technol. Human Identification II*, vol. 5779, Mar. 2005, pp. 7–14.

[41] M. Gadaleta and M. Rossi, "IDNet: Smartphone-based gait recognition with convolutional neural networks," *Pattern Recognit.*, vol. 74, pp. 25–37, Feb. 2018.

[42] Q. Zou, Y. Wang, Q. Wang, Y. Zhao, and Q. Li, "Deep learning-based gait recognition using smartphones in the wild," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 3197–3212, 2020.

[43] A. Brüsch, N. Nguyen, D. Schürmann, S. Sigg, and L. Wolf, "Security properties of gait for mobile device pairing," *IEEE Trans. Mobile Comput.*, vol. 19, no. 3, pp. 697–710, Mar. 2020.

[44] Y. Wu et al., "Poster abstract: Using deep learning to classify the acceleration measurement devices," in *Proc. 19th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, 2020, pp. 351–352, doi: 10.1109/IPSN48710.2020.00-11.

[45] Y. Kim et al., "Usable user authentication on a smartwatch using vibration," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, 2021, pp. 304–319, doi: 10.1145/3460120.3484553.

[46] L. Havasi, Z. Szlavik, and T. Sziranyi, "Detection of gait characteristics for scene registration in video surveillance system," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 503–510, Feb. 2007, doi: 10.1109/TIP.2006.888339.

[47] L. Rong, Z. Jianzhong, L. Ming, and H. Xiangfeng, "A wearable acceleration sensor system for gait recognition," in *Proc. 2nd IEEE Conf. Ind. Electron. Appl.*, 2007, pp. 2654–2659.

[48] C. Nickel, T. Wirtl, and C. Busch, "Authentication of smartphone users based on the way they walk using k-NN algorithm," in *Proc. 8th Int. Conf. Intell. Inf. Hiding Multimedia Signal Process.*, 2012, pp. 16–20.

[49] M. O. Derawi, P. Bours, and K. Holien, "Improved cycle detection for accelerometer based gait authentication," in *Proc. 6th Int. Conf. Intell. Inf. Hiding Multimedia Signal Process.*, 2010, pp. 312–317.

[50] D. Gafurov, K. Helkala, and T. Sandrol, "Biometric gait authentication using accelerometer sensor," *J. Comput.* to be published. [Online]. Available: https://doi.org/10.4304/jcp.1.7.51-59

[51] H. Lu, J. Huang, T. Saha, and L. Nachman, "Unobtrusive gait verification for mobile phones," in *Proc. ACM ISWC*, 2014, pp. 91–98.

[52] C. Wan, L. Wang, and V. V. Phoha, "A survey on gait recognition," *ACM Comput. Surveys*, vol. 51, no. 5, pp. 1–35, 2018.

[53] A. Primo, V. V. Phoha, R. Kumar, and A. Serwadda, "Context-aware active authentication using smartphone accelerometer measurements," in *Proc. CVPR*, 2014, pp. 98–105.

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

[55] B. B. Mjaaland, P. Bours, and D. Gligoroski, "Information security," in *Proc. 13th Int. Conf. ISC*, Boca Raton, FL, USA, Oct. 2010, pp. 361–380. [Online]. Available: https://doi.org/10.1007/978-3-642-18178-8_31

[56] S. J. Shultz, P. A. Houglum, and D. H. Perrin, *Examination of Musculoskeletal Injuries*. Warsaw, Poland: Human Kinetics, 2015.

[57] S. Eberz, N. Paoletti, M. Roeschlin, A. Patani, M. Kwiatkowska, and I. Martinovic, "Broken hearted: How to attack ECG biometrics," *Proc. Netw. Distrib. Syst. Security Symp.*, 2017, p. 7. [Online]. Available: https://doi.org/10.14722/ndss.2017.23408

[58] G. Pan, L. Sun, Z. Wu, and S. Lao, "Eyeblink-based anti-spoofing in face recognition from a generic webcamera," in *Proc. IEEE 11th Int. Conf. Comput. Vis.*, 2007, pp. 1–8. [Online]. Available: https://doi.org/10.1109/iccv.2007.4409068

[59] I. Ray et al., "Proceedings of the 22nd ACM SIGSAC conference on computer and communications security, CCS 15," in *Proc. 22nd ACM SIGSAC Conf. Comput. Commun. Security*, 2015, pp. 1558–1569.

[60] L. Zhang, S. Tan, J. Yang, and Y. Chen, "VoiceLive: A phoneme localization based liveness detection for voice authentication on smartphones," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, 2016, pp. 1080–1091. [Online]. Available: https://doi.org/10.1145/2976749.2978296

[61] B. Thuraisingham et al., "Proceedings of the 2017 ACM SIGSAC conference on computer and communications security, CCS 17," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, 2017, pp. 57–71.

[62] E. Weippl et al., "Accessorize to a crime," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, 2016, pp. 1528–1540. [Online]. Available: https://doi.org/10.1145/2976749.2978392

[63] C. M. Tey, P. Gupta, and D. Gao, "I can be you: Questioning the use of keystroke dynamics as biometrics," in *Proc. NDSS*, 2013, pp. 1–8.

[64] T. Trippel, O. Weisse, W. Xu, P. Honeyman, and K. Fu, "Walnut: Waging doubt on the integrity of MEMS accelerometers with acoustic injection attacks," in *Proc. IEEE Eur. Symp. Security Privacy*, 2017, pp. 3–18.

[65] Y. Tu, Z. Lin, I. Lee, and X. Hei, "Injected and delivered: Fabricating implicit control over actuation systems by spoofing inertial sensors," in *Proc. 27th USENIX Security Symp. (USENIX Security)*, 2018, pp. 1545–1562.

[66] N. Zheng, K. Bai, H. Huang, and H. Wang, "You are how you touch: User verification on smartphones via tapping behaviors," in *Proc. IEEE 22nd Int. Conf. Netw. Protocols*, 2014, pp. 221–232.

[67] M. Derawi and I. Voitenko, "Fusion of gait and ECG for biometric user authentication," in *Proc. Int. Conf. Biometr. Special Interest Group (BIOSIG)*, 2014, pp. 1–4.

[68] 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 *Proc. 15th ACM IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, 2016, pp. 1–12. [Online]. Available: https://doi.org/10.1109/ipsn.2016.7460726

[69] W. Xu et al., "KEH-Gait: Towards a mobile healthcare user authentication system by kinetic energy harvesting," in *Proc. Netw Distrib. Syst. Security Symp.*, 2017, p. 19. [Online]. Available: https://doi.org/10.14722/ndss.2017.23023

[70] J. Han and B. Bhanu, "Individual recognition using gait energy image," *IEEE Trans. Patern Anal. Mach. Intell.*, vol. 28, no. 2, pp. 316–322, Feb. 2006.

[71] S. A. Shaikh and J. R. Rabaiotti, "Characteristic trade-offs in designinglarge-scale biometric-based identity management systems," *J. Netw. Comput. Appl.*, vol. 33, no. 3, pp. 342–351, 2010.

[72] L. Middleton et al., "A foor sensor system for gait recognition," in *Proc. 4th IEEE Workshop Autom. Identification Adv. Technol.*, 2005, pp. 171–176.

[73] R. J. Orr and G. D. Abowd, "The smart foor: A mechanism fornatural user identification and tracking," in *Proc. ACM CHI*, 2000, pp. 275–276.

[74] K. A. Sidek, V. Mai, and I. Khalil, "Data mining in mobile ECG based biometric identification," *J. Netw. Comput. Appl.*, vol. 44, pp. 83–91, Sep. 2014.

[75] Y. Ren, Y. Chen, M. C. Chuah, and J. Yang, "Smartphonebased user verification leveraging gait recognition for mobile healthcare systems," in *Proc. InSecon*, 2013, pp. 149–157.

[76] S. Eberz, G. Lovisotto, A. Patané, M. Kwiatkowska, V. Lenders and I. Martinovic, "When your fitness tracker betrays you: Quantifying the predictability of biometric features across contexts," in *Proc. IEEE Symp. Security Privacy (SP)*, 2018, pp. 889–905, doi: 10.1109/SP.2018.00053.

[77] "VibrationEffect." Accessed: Oct. 21, 2021. [Online]. Available: https://developer.android.com/reference/android/os/VibrationEffect

[78] L. Li, C. Chen, L. Pan, J. Zhang, and Y. Xiang, "Video is all you need: Attacking PPG-based biometric authentication," 2022, *arXiv:2203.00928v1*.

[79] A. Cicone, J. Liu, and H. Zhou, "Adaptive local iterative filtering for signal decomposition and instantaneous frequency analysis." *Appl. Comput. Harmon. Anal.*, vol. 41, no. 2, pp. 384–411, 2014.

[80] L. Yang, W. Wang, and Q. Zhang, "VibID: User identification through bio-vibrometry," in *Proc. 15th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, 2016, pp. 1–12.

[81] "Samsung Patent Illustrates Continued Work on Under-Display Fingerprint Scanning for Future Smartphones and Galaxy Watch." Accessed: Sep. 11, 2021. [Online]. Available: https://www.patentlymobile.com/2018/11/samsung-patent-illustrates-continued-work-on-under-display-fingerprint-scanning-for-future-smartphones-and-galaxy-watch.html

[82] L. de Sousa Britto Neto, V. R. M. L. Maike, F. L. Koch, M. C. C. Baranauskas, A. de Rezende Rocha, and S. K. Goldenstein, "A wearable face recognition system built into a smartwatch and the blind and low vision users," in *Proc. Int. Conf. Enterprise Inf. Syst.*, 2015, pp. 515–528.

[83] "Patentlymobile." Accessed: May 23, 2022. [Online]. Available: https://www.patentlymobile.com/2018/11/samsung-patent-illustrates-continued-work-on-under-display-fingerprint-scanning-for-future-smartphones-and-galaxy-watch.html

[84] FiDELYS—The World's First Iris Recognition Enabled Smartwatch. Accessed: May 23, 2022. [Online]. Available: https://www.youtube.com/watch?v=sw-OCo48F9s&ab_channel=IriTechInc

[85] T. Nguyen and N. D. Memon, "Smartwatches locking methods: A comparative study," in *Proc. SOUPS*, 2017, pp. 1–9.

[86] "ECG Smartwatches Explained: How They Work and the Best on the Market." Accessed: May 23, 2022. [Online]. Available: https://www.wareable.com/health-and-wellbeing/ecg-heart-rate-monitor-watch-guide-6508

[87] "NYMI." Accessed: May 23, 2022. [Online]. Available: https://www.nymi.com/

[88] A. Buriro, B. Crispo, M. Eskandri, S. Gupta, A. Mahboob, and R. Van Acker, "SnapAuth: A gesture-based unobtrusive smartwatch user authentication scheme," in *Proc. Int. Workshop Emerg. Technol. Authorization Authentication*, 2018, pp. 30–37.

[89] A. Buriro, R. V. Acker, B. Crispo, and A. Mahboob, "AirSign: A gesture-based smartwatch user authentication," in *Proc. Int. Carnahan Conf. Security Technol. (ICCST)*, 2018, pp. 1–5.

[90] A. H. Johnston and G. M. Weiss, "Smartwatch-based biometric gait recognition," in *Proc. IEEE 7th Int. Conf. Biometr. Theory Appl. Syst. (BTAS)*, 2015, pp. 1–6.

[91] C. X. Lu, B. Du, X. Kan, H. Wen, A. Markham, and N. Trigoni, "VeriNet: User verification on smartwatches via behavior biometrics," in *Proc. 1st ACM Workshop Mobile Crowdsensing Syst. Appl.*, 2017, pp. 68–73.

[92] T. Nguyen and N. Memon, "Tap-based user authentication for smartwatches," *Comput. Security*, vol. 78, pp. 174–186, Sep. 2018.

[93] "Wired." Accessed: May 25, 2022. [Online]. Available: https://www.wired.com/story/cheap-3d-printer-trick-smartphone-fingerprint-locks/

[94] "Hackers Just Broke the iPhone X's Face ID Using a 3D-Printed Mask." Accessed: May 25, 2022. [Online]. Available: https://www.wired.co.uk/article/hackers-trick-apple-iphone-x-face-id-3d-mask-security

[95] "Hackers Make a Fake Hand to Beat Vein Authentication." Accessed: May 25, 2022. https://www.vice.com/en/article/59v8dk/hackers-fake-hand-vein-authentication-biometrics-chaos-communication-congress

[96] "A Look at How Easily 3D-Printed Heads Can Hack Facial Recognition." Accessed: May 25, 2022. [Online]. Available: https://interestingengineering.com/innovation/a-look-at-how-easily-3d-printed-heads-can-hack-facial-recognition

[97] Y. Kim et al., "Usable user authentication on a smartwatch using vibration," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, 2021, pp. 304–319.

[98] B. Thuraisingham et al., "VibWrite," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, 2017, pp. 73–87.

[99] J. Li, K. Fawaz, and Y. Kim, "VELODY: Nonlinear vibration challenge-response for resilient user authentication," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security (CCS)*, 2019, pp. 1201–1213. [Online]. Available: https://doi.org/10.1145/3319535.3354242

[100] "What Is Liveness Detection? Types and Benefits of Liveness Detection." Accessed: May 25, 2022. [Online]. Available: https://www.idcentral.io/article/what-is-liveness-detection-types-and-benefits-of-liveness-detection/

[101] D. F. Smith, A. Wiliem, and B. C. Lovell, "Face recognition on consumer devices: Reflections on replay attacks," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 4, pp. 736–745, Apr. 2015, doi: 10.1109/TIFS.2015.2398819.

[102] Y. Xu, T. Price, J. M. Frahm, and F. Monrose, "Virtual U: Defeating face liveness detection by building virtual models from your public photos," in *Proc. 25th USENIX Security Symp. (USENIX Security)*, 2016, pp. 497–512.

[103] X. Song, X. Zhao, L. Fang, and T. Lin, "Discriminative representation combinations for accurate face spoofing detection," *Pattern Recognit.*, vol. 85, pp. 220–231, Jan. 2019.

[104] D. Menotti et al., "Deep representations for iris, face, and fingerprint spoofing detection," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 4, pp. 864–879, Apr. 2015, doi: 10.1109/TIFS.2015.2398817.

[105] G. O. Karame and A. Stavrou, "CCSW'17: 2017 ACM cloud computing security," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security (CCS)*, New York, NY, USA, 2017, pp. 2627–2628. [Online]. Available: https://doi.org/10.1145/3133956.3137050

[106] A. Makhzani, J. Shlens, N. Jaitly, I. J. Goodfellow, and B. Frey, "Adversarial autoencoders," 2015, *arXiv:1511.05644v2*.

Wei Song received the master's degree from The University of New South Wales, Sydney, NSW, Australia, in 2021, where he is currently pursuing the Ph.D. degree with the School of Computer Science and Engineering. He is supervised by Prof. W. Hu and Prof. C. T. Chou.

His research interests include cyber security and biometric authentication.

Hong Jia received the B.S. degree from Harbin Engineering University, Harbin, China, in 2012, the M.S. degree from Northeastern University, Shenyang, China, in 2015, and the Ph.D. degree in computer science from The University of New South Wales, Sydney, NSW, Australia, in 2021.

He is currently an Research Associate with the University of Cambridge, Cambridge, U.K. His research interests include mobile computing, sensing, and efficient machine learning.

Dr. Jia served as a TPC in TheWebConf and reviewers in EWSN, IPSN, and TMC.

Min Wang (Member, IEEE) received the Ph.D. degree in computer science from The University of New South Wales, Canberra, ACT, Australia, in 2019.

She is a Postdoctoral Research Fellow with the School of Engineering and Information Technology, The University of New South Wales. Her research interests include biometrics, privacy and security, pattern recognition, machine learning, and bio-cryptography.

Yuezhong Wu received the master's degree from the University of Science and Technology of China, Hefei, China, in 2016. He is currently pursuing the Ph.D. degree with the School of Computer Science and Engineering, The University of New South Wales, Sydney, NSW, Australia. He is supervised by Prof. W. Hu and Prof. M. Hassan.

His research interests include cyber security and novel applications for the Internet of Things.

Wanli Xue received the Ph.D. degree from The University of New South Wales, Sydney, NSW, Australia, in 2019.

He is a Threat Intelligence Researcher with Sophos, Sydney, NSW, Australia. Before that, he was a Research Fellow with Cyber Security CRC and the School of Computer Science and Engineering, The University of New South Wales. His research interests are mainly in security and privacy issues in cyber–physical systems and IoT, including highly efficient privacy-preserving techniques for IoT as well as IoT-related sensing systems and data analytic services.

Chun Tung Chou (Senior Member, IEEE) received the B.S. degree in engineering science from the University of Oxford, Oxford, U.K., in 1988, and the Ph.D. degree in control engineering from the University of Cambridge, Cambridge, U.K., in 1994.

He is currently an Associate Professor with the School of Computer Science and Engineering, The University of New South Wales, Sydney, NSW, Australia. His research interests are in the communications and computing aspects of natural and synthetic bio-molecular systems.

Dr. Chou is on the editorial board of IEEE WIRELESS COMMUNICATIONS LETTERS, IEEE TRANSACTIONS ON MOLECULAR, BIOLOGICAL AND MULTI-SCALE COMMUNICATIONS, and *Nano Communication Networks*.

Jiankun Hu (Senior Member, IEEE) received the bachelor's degree in industrial automation from Hunan University, Changsha, China, in 1983, the Ph.D. degree in engineering from Harbin Institute of Technology, Harbin, China, in 1993, and the master's degree in computer science and software engineering from Monash University, Melbourne, VIC, Australia, in 2000.

He is a Full Professor of Cyber Security with the School of Engineering and Information Technology, The University of New South Wales, Canberra, ACT, Australia. His main research interest is in the field of cyber security, including biometrics security, where he has publications at top venues including the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE.

Dr. Hu has served on the editorial boards of up to seven international journals, including the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY.

Wen Hu (Senior Member, IEEE) received the Ph.D. degree in computer science and engineering from The University of New South Wales, Sydney, NSW, Australia, in 2006.

He is currently an Associate Professor with the School of Computer Science and Engineering, The University of New South Wales. He has published regularly in the top-rated sensor network and mobile computing venues, such as IPSN, SenSys, MobiCom, UbiComp, TOSN, TMC, TIFS, and the PROCEEDINGS OF THE IEEE. His research interests focus on the novel applications, low-power communications, security and compressive sensing in sensor network systems, and the Internet of Things (IoT).

Prof. Hu is an Associate Editor of *ACM Transactions on Sensor Networks*. He is the General Chair of CPS-IoT Week 2020, and serves on the organizing and program committees of networking conferences, including IPSN, SenSys, MobiSys, MobiCom, and IoTDI. He is a Senior Member of ACM.