AudioGuard: Omnidirectional Indoor Intrusion Detection Using Audio Device

TIANBEN WANG, ZHANGBEN LI, HONGHAO YAN, XIANTAO LIU, and BOQIN LIU,
Northwest A&F University, China
SHENGJIE LI, JD.com, Inc.,China
ZHONGYU MA, Northwest Normal University, China
JIN HU, Northwest A&F University, China
DAQING ZHANG, Peking University, China
TAO GU, Macquarie University, Australia

Indoor intrusion detection is a critical task for home security. Previous works in intrusion detection suffer from the problems such as blind spots in non-line-of-sight (NLOS) areas, restricted device locations, massive offline training required, and privacy concern. In this article, we design and implement an omnidirectional indoor intrusion detection system, named *AudioGuard*, using only a pair of speaker and microphone. *AudioGuard* is able to detect both line-of-sight (LOS) and NLOS intrusions. Our observation of acoustic signal propagation in an indoor environment shows that there exist abundant multipath reflections and human movement introduces Doppler shift in echo signals. We hence capture periodical Doppler shift caused by intruder's walking motion to detect intrusion. Specifically, we first extract the Doppler shift embedded in echo signals, and we then propose a periodicity polarization method to cancel out the impact of the change of radial angle and the distance on periodicity of Doppler shift. Finally, we detect intrusion by measuring periodicity of Doppler shift over time. Extensive experiments show that *AudioGuard* achieves a miss report rate of 0% and 1.75% for LOS and NLOS intrusion, respectively, and a false alarm rate of 4.17%.

4

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing systems and tools**;

Additional Key Words and Phrases: Indoor intrusion detection, acoustic sensing, periodic doppler shift

ACM Reference format:
Tianben Wang, Zhangben Li, Honghao Yan, Xiantao Liu, Boqin Liu, Shengjie Li, Zhongyu Ma, Jin Hu, Daqing Zhang, and Tao Gu. 2023. AudioGuard: Omnidirectional Indoor Intrusion Detection Using Audio Device. *ACM Trans. Internet Things* 5, 1, Article 4 (December 2023), 22 pages.
https://doi.org/10.1145/3625305

This work is supported in part by the Key Research and Development Project in Shaanxi Province of China (2023-YBGY-257), Shaanxi Key Industry Innovation Chain Project (2023-ZDLNY-69) and Yangling Livestock Industry Innovation Center Double-chain Fusion Project (2022GD-TSLD-46).

Authors' addresses: T. Wang, Z. Li, H. Yan, X. Liu, B. Liu, and J. Hu, Northwest A&F University, China; e-mails: {wangtb, lzb996}@nwafu.edu.cn, yanhonghao9@163.com, liuxt@nwafu.edu.cn, boqinliu_lbq@163.com, hujin007@nwsuaf.edu.cn; S. Li, JD.com, Inc., China; e-mail: lishengjie@pku.edu.cn; Z. Ma, Northwest Normal University, China; e-mail: mazybg@nwnu.edu.cn; D. Zhang, Peking University, China; e-mail: dqzhang@sei.pku.edu.cn; T. Gu, Macquarie University, Australia; e-mail: tao.gu@mq.edu.au.

2577-6207/2023/12-ART4 $15.00
https://doi.org/10.1145/3625305

1 INTRODUCTION

Indoor intrusion (someone enters the room without permission) detection plays a crucial role in home security such as protecting assets and preventing personal attacks. A recent report in SafeAtLast [36] shows that there are 2.5 million burglaries per year, 66% of which are home break-ins, causing more than US$3.1 billion in damages every year. More than 25% of those who interrupt burglars become victims of violent crimes. Homes without a security system have a 300% more chance of getting broken into. The high incidence of home burglary demands effective intrusion detection in home settings.

Video-based surveillance [7–10] has been widely used to detect intrusion in public places. However, the video-based approach may present a severe privacy concern when applied to a home setting. Additionally, video-based systems fail to detect **non-line-of-sight (NLOS)** intrusion. Infrared-based approaches [11–14] have been well studied over a decade. However, these systems typically have a limited sensing range, since sensors are deployed in each typical entrance such as the main entrance and window. In addition, these sensors need to be properly installed by well-trained professionals due to the strict requirement of sensor direction. Ultrasonic sensor approaches [33, 34] may suffer the same problem.

Radar-based approaches [15–17] have been proposed in recent years. Although these approaches can achieve accurate intrusion detection, they typically require expensive human efforts in offline training, and radar hardware is usually costly, hence limiting its applicability in home settings. For a cost-effective solution, WiFi devices have been used to build intrusion detection systems. These approaches [2, 18–22, 26–29] share the same idea of extracting **Receive Signal Strength Indicator (RSSI)** or **Channel State Information (CSI)** variation pattern and applying machine learning algorithms for pattern matching. These approaches rely heavily on massive data for offline training. To overcome data dependency, studies in References [23, 24] detect intrusion by comparing RSSI variance to specific threshold. However, RSSI variance varies significantly with respect to distance, location, and walking direction, and hence setting an accurate threshold is not feasible. Li et al. [30–32] detect intrusion by identifying the transient moment of an intruder entering house with accurate estimation WiFi sensing boundary. However, WiFi-based approaches require the transmitter and receiver placed at two sides of an intruder, or the performance declines significantly.

The audio devices embedded in smartphone have been used to detect intrusion by detecting door opening and closing events [35]. Microphone array has been used to detect intrusion [1, 3]. However, like radar- and WiFi-based approaches, they rely heavily on massive data for offline training to cover all the possible conditions. Ultrasonic sensors also have been used to build intrusion detection systems [33, 34]. Due to strong directionality, ultrasonic sensors-based approaches suffer the same problem as in the infrared-based approaches. Zieger et al. [4] and Zu et al. [38] proposed to extract various time and domain features of the data received by microphone array to identify intrusion. However, the experience-based method suffers from poor environment adaptation.

In this article, we design and implement *AudioGuard*, an omnidirectional indoor intrusion detection system using only a speaker and microphone. The system is able to detect both **line-of-sight (LOS)** and NLOS intrusions. *AudioGuard* can be implemented on different audio device and is robust against interference and transceiver's location and orientation. We observe that there exist abundant acoustic reflections in an indoor environment, and an intruder's walking motion always introduces a periodic Doppler shift. We hence discover our basic idea to capture an intruder's walking motion that is inevitable during intrusion. By measuring the periodic Doppler shift embedded in echo signals, we can detect intrusion. Designing and implementing such an omnidirectional intrusion detection system, however, entails two main challenges as follows:

(1) Different from the shift of main peak of echo spectrum as explained in the classical Doppler effect theory, due to multipath reflections in indoor environments, the Doppler shift caused by walking motion appears as sidelobes of an echo spectrum without an obvious peak. This raises a problem of how to quantify the Doppler shift.
(2) Although limbs swing during walking is periodic, as the change of radial angle and the distance from intruder to device, the Doppler shift over time embeds with nonlinear trend and its amplitude varies nonlinearly. It seriously decreases the periodicity of Doppler shift and makes it difficult to distinguish intrusion and interference, e.g., curtain fluttering.

To address the aforementioned challenges, we first propose to capture Doppler shift using an echo **power spectrum density (PSD)** difference vector. We then propose a periodicity polarization method to cancel out the impact of the change of radial angle and the distance on Doppler shift periodicity. Finally, we detect intrusion by measuring the periodicity of Doppler shift sequence. The demo video is available at https://tinyurl.com/4y44pdbk or https://youtu.be/iI-Pk4st75o.

The main contributions of this article are summarized as follows:

(1) We design and implement *AudioGuard*, an omnidirectional intrusion detection system using only a speaker and a microphone. It captures both LOS and NLOS intrusions. We propose to capture Doppler shift using PSD difference vector and cancel out the impact of the change of radial angle and the distance on Doppler shift over time using a periodicity polarization algorithm. It enlarges the Doppler shift periodicity difference between walking and other movement interference.
(2) We conduct extensive experiments to evaluate *AudioGuard* with a variety of indoor settings. Experiments show that *AudioGuard* can be implemented on different audio device and robust against the variation of transceiver's location and orientation. *AudioGuard* achieves a miss report rate of 0% and 1.75% for LOS and NLOS intrusion, respectively. The false alarm rate under interference is 4.17%.

2 RELATED WORK

In this section, we briefly review the most relevant works in indoor intrusion detection that can be grouped into three categories: video- and infrared-based approaches, RF-based approaches, and audio-based approaches.

2.1 Video- and Infrared-based Approaches

Video-based approaches [7–10] have been widely used to detect intrusion in public places. Cameras are installed to capture images or video for intruder recognition. Compared with other intrusion detection approaches, video-based approaches can detect intrusion and also retain full evidence. However, limited by visual angle, multiple cameras have to work together from different positions to cover an entire room, and it fails if an intruder is in an NLOS area. In addition, video-based approaches are usually sensitive to light conditions and may present a severe privacy concern when applied to home settings.

Infrared-based approaches [11–14] have been a mature intrusion detection solution for many years. These approaches can be further grouped into two categories. The first category leverages a pyroelectric infrared sensor to capture infrared signals released by an intruder. Limited by small sensing range, pyroelectric infrared sensors are usually deployed to monitor the small area around entrance. The second category leverages directional infrared sensor to detect transient moment when intruder block the line-of-sight between sender and receiver. To avoid underreporting, multiple sensors have to be deployed in every possible entrance, such as door and window, to form a wireless sensor network. In addition, due to strong directionality, thes sensors are required to be carefully installed by well-trained professionals. The complex deployment may prevent these systems from large-scale deployment in home settings.

2.2 RF-based Approaches

Radar-based approaches [15–17] share the same basic idea of extracting various features from the Doppler effect caused by walking and then detecting intrusion by matching feature variation pattern using machine learning algorithms. These approaches rely heavily on massive data for offline training. In addition, expensive hardware prevents large-scale deployment of these systems in home settings.

To build an intrusion detection system that is friendly for the home environment, researchers have turned their attention to widely available commercial WiFi devices. In the early stage of WiFi sensing, RSSI was explored to detect intrusion [1, 18, 19, 22–25, 39]. References [23, 24] detect intrusion using a threshold to identify whether movement occurs. These approaches are sensitive to movement interference (e.g., curtain fluttering). References [1, 18, 19, 22, 25, 39] share the same basic idea as radar-based approaches. They detect intrusion by extracting RSSI variation pattern from offline RSSI data using machine learning algorithms. These approaches heavily rely on massive offline data and suffer from poor environment adaptation ability.

Similarly, References [2–4, 20-21, 26-29] detect intrusion by exacting the CSI variation pattern from offline CSI data using different machine learning algorithms. These approaches suffer from similar problems as RSSI-based and radar-based approaches. To overcome the limitations of the above approaches, Li et al. [30] propose to detect intrusion using a relatively robust feature CSI ratio, i.e., the ratio between dynamic CSI component and static CSI component. Furthermore, Li et al. [31] propose to detect intrusion by measuring Doppler shift embedded in CSI [40]. However, due to lack of mechanism to avoid interference, these two approaches are sensitive to movement interference such as object falling and curtain fluttering. The approach in Reference [32] accurately detects intrusion by identifying the transient moment of an intruder entering a house with an accurate estimation of WiFi sensing boundary. Lin et al. [5] propose a CSI-EIH model to describe the effect of moving object's height to CSI amplitudes. Based on this model, the system can detect intrusion and avoid false alarm caused by pets. However, the system may raise false alarms if moving object is higher than the given height threshold. In addition to these shortcomings, CSI-based approaches have restriction on device location. They require that transmitter and receiver placed at two sides of intruder, else the performance declines significantly.

2.3 Audio-based Approaches

Audio-based approaches have recently attracted researchers' attention. Ultrasonic sensors are leveraged to build intrusion detection systems [33, 34]. Due to strong directionality, ultrasonic sensors-based approaches suffer the same problem as in infrared-based approaches. Dissanayake et al. [35] use the speaker and microphone embedded in smartphones to detect intrusion by identifying door opening and closing events based on Doppler shift. However, it is sensitive to the location and orientation of the smartphone, because different locations and orientations of the smartphone may result in completely different Doppler shifts. In addition, like both radar- and RSSI-based approaches, it also heavily relies on massive data for offline training. Microphone arrays have been used to detect intrusion [1, 3]. However, these methods also heavily rely on massive data for offline training to cover all the possible conditions. Zieger et al. [4] and Zu et al. [38] proposed to extract various time and domain features of the data received by microphone array to identify intrusion. These approaches are all experience-based approaches lacking explainable theory. It leads to poor environment adaptation ability.

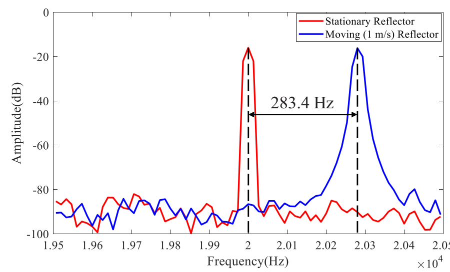


Fig. 1. Classical Doppler shift caused by moving reflector.

Differently, in this article, we design and implement *AudioGuard*, an omnidirectional indoor intrusion detection system using only a pair of speaker and microphone. *AudioGuard* detects intrusion by fully leveraging abundant multipath reflection in indoor environment to capture periodical Doppler shift caused by intruder's walking. It is able to detect both LOS and NLOS intrusions.

3 DOPPLER SHIFT CAUSED BY INTRUSION IN INDOOR ENVIRONMENT

3.1 Classical Doppler Effect

Doppler frequency shift is caused by relative movement between transmitter and receiver. In general, when a receiver moves toward a signal source, the frequency of the received signal increases and vice versa. Mathematically, the frequency of the received signal can be described as follows:

$$f' = \frac{c \pm v_r}{c \mp v_s} f, \tag{1}$$

where f' is the received frequency, f is the transmitted frequency, c is the velocity of the wave in propagation medium, v_r is the radial velocity of the receiver relative to the medium (positive if the receiver is moving toward the source and negative otherwise), and v_s is the radial velocity of the source relative to the medium (positive if the source is moving away from receiver and negative otherwise).

If the signal source and receiver are integrated to form an acoustic radar, then $v_r = v_s$. Plugging $\pm v_s = |\boldsymbol{v}| \cdot cos(\theta)$ ($\boldsymbol{v}$ and θ denote the walking speed of reflector and the angle between $\boldsymbol{v}$ and signal source, respectively) into Equation (1), the Doppler frequency shift Δf can be represented as

$$\Delta f = f' - f = \left(\frac{2\,|\boldsymbol{v}| \cdot cos\,(\theta)}{c \mp |\boldsymbol{v}| \cdot cos\,(\theta)} \right) \cdot f\ . \tag{2}$$

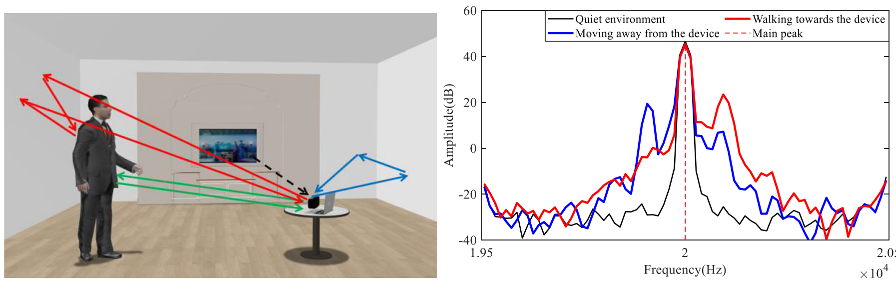
When $\boldsymbol{v}$ is much less than the velocity of sound c, we have $c \mp v_s \approx c$. Δf can be further simplified as

$$\Delta f = \frac{2\,|\boldsymbol{v}|\,cos\,(\theta)}{c} f. \tag{3}$$

The ideal model above assumes that the receiver only receives the signal reflected from the front of the moving reflector. Thus, the Doppler shift appears as the peak shift echo spectrum. As shown in Figure 1, when the reflector moves toward the transceiver with a speed of 1 m/s, the Doppler shift Δf apears as the peak shift of 283.4 Hz (the frequency of transmitting signal is 20 kHz and the sampling frequency is 48 kHz).

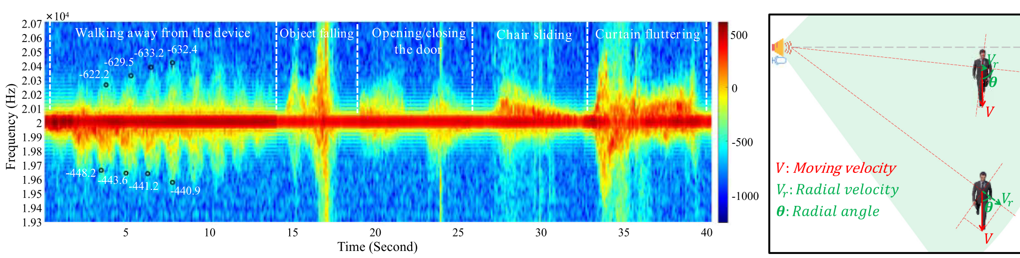
3.2 Doppler Shift Caused by Walking in Spectrum

Comparing with the human body, the reflection area of a static environment is much larger. In other words, the power of the multipath signal reflected from intruder is much lower than the



(a) (b)

Fig. 2. Comparing the frequency resolution of FFT and the proposed method.



(a) (b)

Fig. 3. Time–frequency spectrum of the echo when walking and other disturbance events happen.

power of the multipath signal reflected from the static environment. However, only the signals reflected from the moving object embed with the Doppler shift. The result is that rather than changing the main peak of echo spectrum, the Doppler shift caused by intruder's walking appears as sidelobes of echo spectrum. Additionally, as shown in Figure 2(a), the change of the reflection path length of the signal reflected from the front and the back of intruder are always opposite, so they always introduce opposite Doppler shift (positive Doppler shift versus negative Doppler shift). Based on the above analysis, we can finally derive that a Doppler shift caused by walking always appears as two different shapes of sidelobes at two sides of the main peak of echo. Figure 2(b) shows the PSDs of echo (0.1 second) when an intruder is approaching and walking away from the device (the frequency of transmitting signal is 20 kHz and the sampling frequency is 48 kHz). So finding a way to quantify the Doppler shift caused by the walking is challenging.

We consider a scenario where there are two people in a room and they are out of the sight of each other. For example, one is in bedroom and the other one is outside bedroom. One can still hear what the other says, since the voice signals may be reflected from the static environment and propagate through multipath. Similarly, due to abundant acoustic multipath reflections in indoor environments, even if the intruder is in NLOS areas, the microphone can still receive the signal indirectly reflected from the intruder, and, thus, the echo is still embedded with the Doppler shift caused by the intruder's walking.

3.3 Periodicity of the Doppler Shift Caused by Walking over Time

Figure 3(a) shows the time–frequency spectrum of the echo when intruder is walking away from the device (as shown in Figure 3(b)) and other interference (e.g., object falling, curtains fluttering, opening and closing door, and chair sliding) happen.

From Figure 3(a), we observe that due to the periodic limbs swing during walking, the Doppler shift caused by walking shows some periodicity, while a Doppler shift caused by interference is

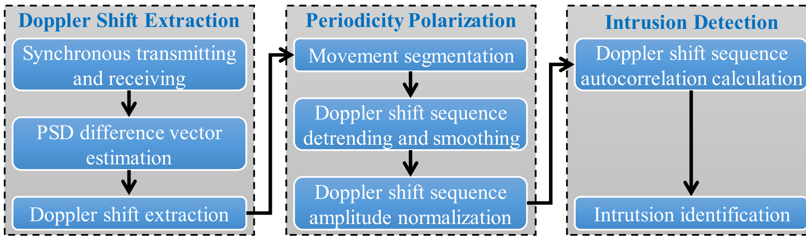


Fig. 4. System framework.

aperiodic. So, our insight to detect intrusion is capturing the periodic Doppler shift sequence over time. However, from Figure 3(a) we also observe that the Doppler shift varies nonlinearly over time as radial angle and the distance varies during walking (shown in Figure 3(b)). Except for last two steps before a stop, when the intruder is close to the device, the Doppler shift appears as a low-frequency shift but with high power. When the intruder is far from the device, the Doppler shift appears as a high-frequency shift but is low in power. As shown in Figure 3(b), it is caused by the fact that when an intruder is close to the device, the energy of the signal reflected from intruder is relative larger, while large radial angle leads to small radial velocity, finally resulting in a low-frequency shift. When an intruder is far from the device, the energy of the signal reflected from intruder is low, while a small radial angle makes large radial velocity, finally resulting in a high-frequency shift. So finding a way to cancel out the impact of radial angle and distance variation on the periodicity of the Doppler shift is challenging.

4 SYSTEM DESIGN AND IMPLEMENTATION

4.1 System Framework

As shown in Figure 4, *AudioGuard* has three modules, namely the Doppler effect extraction module, the periodicity polarization module, and the intrusion detection module. The Doppler effect extraction module extracts a Doppler shift using a PSD difference vector. The periodicity polarization module first retains the segments that contain moving events and then polarizes the periodicity of the Doppler shift sequence. The intrusion detection module detects intrusion by measuring the periodicity of the Doppler shift sequence over time.

4.2 Doppler Shift Extraction

AudioGuard continuously transmits a 20-kHz single frequency acoustic signal, which is inaudible for a human, through a player. The microphone receives the echo synchronously with the sampling rate $f_s = 48$ kHz. The length of the received frame is 0.1 seconds. To remove environmental noise in echo, a band-pass filter is adapted. Since the maximum walking velocity of a human is approximately 4.3 m/s, according to the Doppler shift formula [41], the maximum Doppler shift caused by an intruder walking is about 500 Hz. Therefore, the pass band of the filter is set as $[f_c - 500, f_c + 500]$.

As mentioned in Section 3, the Doppler shift caused by an intruder's walking motion appears as two different shapes of sidelobes at two sides of the main peak of the echo spectrum. The traditional method, i.e., extracting a main peak shift of PSD of echo, cannot be applied to capture the Doppler shift under this condition. In this article, we capture the Doppler shift using a PSD difference vector.

PSD can be estimated using the Welch algorithm [6], which can significantly improve the variance characteristics of the power spectrum via overlap mechanism, also effectively reducing spectrum leakage via the window function. Specifically, the PSD of one echo frame $x_r(n)$ is estimated as

$$P(\omega_k) = \frac{1}{LN} \sum_{l=0}^{L-1} \left| \sum_{n=0}^{N-1} x_r(n + l \cdot M)\, w(n)\, e^{-jn\omega_k} \right|^2, \tag{4}$$

where ω_k is the digital angular frequency, which is defined as $\omega_k = k\Delta\omega = 2\pi k/N$, L is the number of overlapped segments, N is the length of each segment, M is the hop size, and w is a window function with a length of N. As mentioned, the Doppler effect caused by walking appears at frequency band $f_c - 500 \leq f_k \leq f_c + 500$. According to the relationship between physical frequency f_k and digital angular frequency ω_k, i.e., $\omega_k = 2\pi f_k/f_s$, where f_s is the sampling frequency, we can obtain the range of ω_k,

$$\frac{2\pi(f_c - 500)}{f_s} \leq \omega_k \leq \frac{2\pi(f_c + 500)}{f_s}. \tag{5}$$

Plugging $\omega_k = k\Delta\omega = 2\pi k/N$ into Equation (5), we obtain the range of k,

$$\frac{N(f_c - 500)}{f_s} \leq k \leq \frac{N(f_c + 500)}{f_s}.$$

The minimum and maximum of k is

$$k_1 = \left\lceil \frac{N(f_c - 500)}{f_s} \right\rceil, \; k_H = \left\lfloor \frac{N(f_c + 500)}{f_s} \right\rfloor$$

Then the PSD difference vector of the echo frame at time t_i is defined as

$$\boldsymbol{PD}_i = \left(diff^i_{k_1}, \; diff^i_{k_2}, \ldots, \; diff^i_{k_m}, \ldots, d^i_{iffk_H} \right), \; k_1 < k_2 < \cdots < k_H, \tag{6}$$

where $diff^i_{k_m}$ is defined as

$$diff^i_{k_m} = P_i(\omega_{k_m}) - P_{ref}(\omega_{k_m}), \tag{7}$$

where P_i denotes the PSD of the echo frame at time t_i (refer to Equation (4)) and P_{ref} denotes the reference PSD, which is the average PSD of the echo collected from current environment without any movement event. Every time *AudioGuard* starts, it first runs the initializer to obtain a reference PSD of the current environment. Specifically, the initializer continuously estimates the echo PSD for 20 seconds and then calculates the average PSD value as a reference PSD. During this process, any movement event is forbidden. If movement happens while the PSD is being referenced, then the initializer automatically tries again to obtain the reference PSD. Specifically, if the PSDs show obvious fluctuation, which can be easily observed from the variances of points in PSD over time, then the initializer runs again.

We randomly record 20 seconds to get reference PSD for five days in the same room without movement interference. Figure 5 shows the recorded reference PSD of each day. It can be observed that the reference PSDs are very similar. It indicates that if the environment does not change, then we only need to get reference PSD for only one time.

Figure 6(a) show the reference PSD and the PSD of one echo frame when there is no moving object. Figure 6(b) shows the PSD difference vector derived from Figure 6(a). Figure 6(c) shows the reference PSD and the PSD of one echo frame when one subject is walking. Figure 6(d) shows the PSD difference vector derived from Figure 6(c). Figure 7 shows all the samples of PSD difference vectors during one walking event containing seven steps. Even though during walking the limb swing is periodic, the PSD difference vectors during walking do not show obvious periodicity. It is consistent with our analysis in Section 3.3.

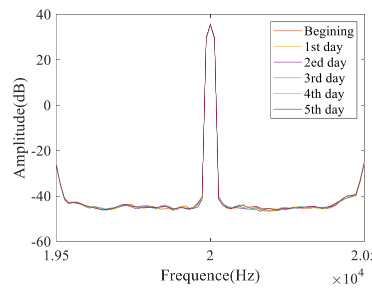


Fig. 5. Reference PSDs over five days.

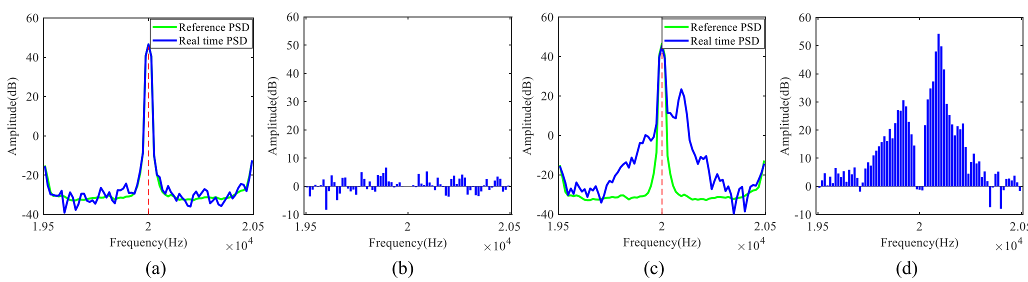


Fig. 6. PSD difference vector.

Finally, we quantify Doppler shift of the echo frame at time t_i as the first normal form of $\boldsymbol{PD}_i$, i.e.,

$$d_i = \|\boldsymbol{PD}_i\|_1 = \sum_{j=1}^{H} \left| diff_{k_j}^{i} \right|. \tag{8}$$

4.3 Periodicity Polarization

Due to abundant acoustic multipath reflection in a room environment, almost all the movement events (e.g., walking, door opening and closing, object falling, object sliding, and curtain fluttering) will introduce a Doppler shift in echo. We have to first identify all the movement events and then identify intrusion among these movement events. According to Equation (8), compared with the condition where there is no moving object, when a moving event occurs, the value of d_i will be obviously larger due to a Doppler shift. So, we can simply truncate the segments containing movement using a threshold. Specifically, to avoid the interference caused by system jitter, if d_i is larger than the threshold for 3 continuous times, we start to record d_i. Similarly, if d_i is smaller than the threshold for 3 continuous times, then we stop to record d_i. Thus, each movement event can be segmented out, and d_i over time is recorded. We call the recorded d_i over time as Doppler shift sequence. Figure 8 shows the Doppler shift sequence of different movement events.

We observe that the Doppler shift sequences caused by curtains fluttering and objects sliding are aperiodic. Even though walking is periodic, the Doppler shift sequence caused by walking (shown in Figure 8(c)) shows weak periodicity. It is caused by the change of radial angle and the distance from intruder to device during walking (refer to Section 3.2). Too-weak periodicity of a Doppler shift sequence caused by walking will result in intrusion detection error. In order to improve intrusion detection accuracy, we intend to enhance the Doppler shift with weak periodicity while keeping the periodicity of an aperiodic Doppler shift sequence. In other words, we enlarge the periodicity difference between a weak periodic Doppler shift sequence and an

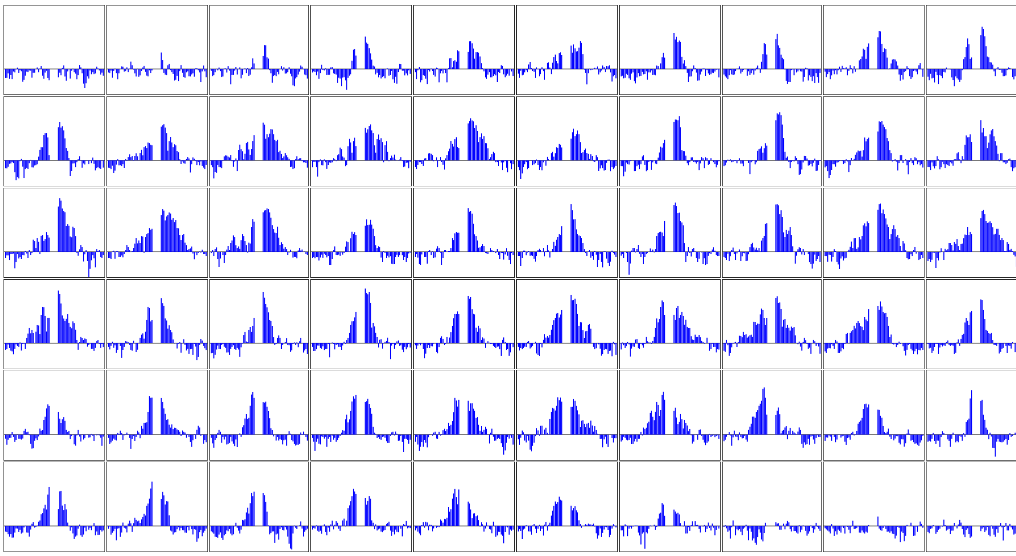


Fig. 7. Samples of PSD difference vector during a walking event.

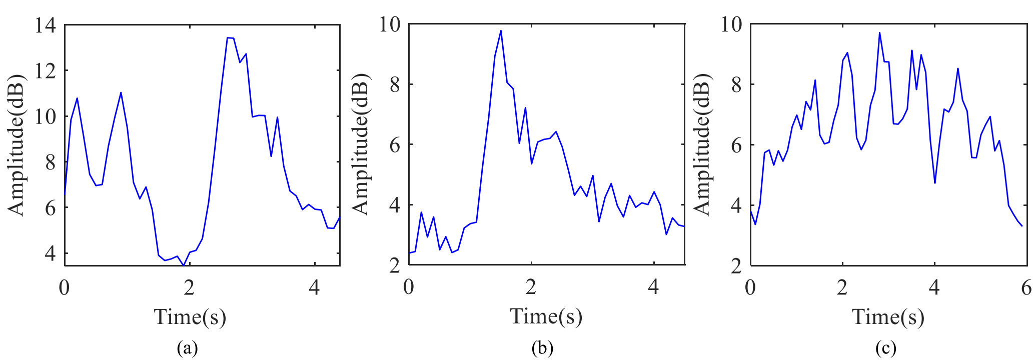
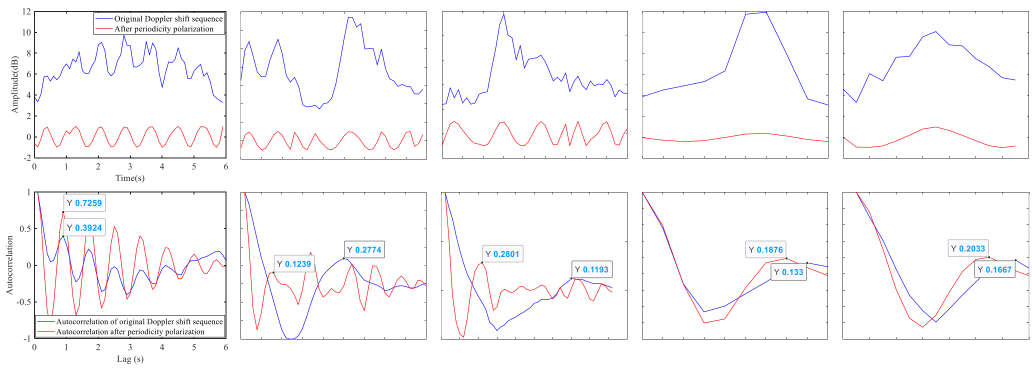


Fig. 8. Doppler shift sequence of different movement events. (a) Curtain fluttering. (b) Object sliding. (c) Walking.

aperiodic Doppler shift sequence. Thus, the intrusion can be easily detected as the Doppler shift sequence with strong periodicity.

It is well known that strong periodicity simultaneously requires the following: (i) the shapes of the sequence during all periods are similar, and (ii) the length of all the periods is almost constant. If any one of the requirements is not met, then the signal will show weak even no periodicity. From Figure 6, we clearly see that a Doppler shift sequence caused by walking only meets the first requirement, while Doppler shift sequences caused by other aperiodic movements do not meet both requirements. If we can enhance the similarity of sequence shape during each period while keep the length of original periods, then a Doppler shift sequence caused by walking will then meet both requirements, while Doppler shift sequences caused by other aperiodic movements only meet the first requirement. Thus, the periodicity of a Doppler shift sequence caused by walking will be significantly enhanced while the periodicity of Doppler shift sequence caused by other aperiodic movement will not be enhanced.



(a) Walking (b) Curtain fluttering (b) Object sliding (d) Object falling (e) Opening or closing door

Fig. 9. Periodic polarization of the Doppler shift sequence of different moving events.

Periodicity polarization is designed to eliminate the difference of fluctuation amplitude in each period. Specifically, the periodicity polarization process is composed of detrending, smoothing, and amplitude normalization. First, we extract the polynomial trend of a Doppler shift sequence and subtract it from Doppler shift sequence. Then, we smooth the detrended Doppler shift sequence to eliminate small burrs. Finally, we normalize the amplitude of Doppler shift sequence by eliminating the envelope of the sequence using a Hilbert transform [37]. For a detrended and smoothed Doppler shift sequence $\boldsymbol{d} = d_1, d_2, \ldots d_i, \ldots$, a discrete Hilbert transform can be calculated by leveraging Discrete Fourier Transform and Inverse Discrete Fourier Transform,

$$\hat{\boldsymbol{d}} = IDFT\left(\hat{\boldsymbol{D}}\right), \tag{9}$$

where $\hat{\boldsymbol{D}}$ is defined as

$$\hat{\boldsymbol{D}}(k) = \begin{cases} -jD(k), \ k = \begin{cases} 1, 2, \ldots, N/2 - 1, \ when\ N\ is\ even \\ 1, 2, \ldots, (N-1)/2, when\ N\ is\ odd \end{cases} \\ jD(k), k = \begin{cases} N/2 + 1, \ldots, \ N - 1, when\ N\ is\ even \\ (N+1)/2, \ \ldots, N - 1, when\ N\ is\ odd \end{cases} \end{cases}, \tag{10}$$

where $\boldsymbol{D} = DFT(\boldsymbol{d})$. With $\hat{\boldsymbol{d}}$, the envelope of $\boldsymbol{d}$, i.e., the amplitude of $\boldsymbol{d}$ over time can be calculated as

$$\boldsymbol{A}(n) = \sqrt{\boldsymbol{d}(n)^2 + \hat{\boldsymbol{d}}(n)^2}\ . \tag{11}$$

We then normalize $\boldsymbol{d}$ as

$$\boldsymbol{d}'(\boldsymbol{n}) = \frac{\boldsymbol{d}(\boldsymbol{n})}{\boldsymbol{A}(\boldsymbol{n})}\ . \tag{12}$$

The subfigures in first line of Figure 9 show the original Doppler shift sequences (blue line) and the Doppler shift sequences after periodicity polarization (red line) of different moving events. Subfigures in the second line of Figure 9 show the corresponding autocorrelations. A higher peak of the autocorrelation function means stronger periodicity. Comparing the autocorrelation of original Doppler shift sequences and that of the Doppler shift sequences after periodicity polarization, we can see that the periodicity of Doppler shift sequence of walking is significantly enhanced while the periodicities of other Doppler shift sequences are not enhanced obviously. It indicates that periodicity polarization is able to enlarge the periodicity difference between weak periodic Doppler shift sequence and aperiodic Doppler shift sequence.

4.4 Intrusion Detection

After periodicity polarization, intrusion can be easily detected as the Doppler shift sequence with strong periodicity. We measure the periodicity of Doppler shift sequence using autocorrelation function. The autocorrelation of the Doppler shift sequence after periodicity polarization $\boldsymbol{d'}$ is given by

$$R_x(k) = \frac{c_k}{c_0}, \tag{13}$$

where c_k is the auto-covariance of $\boldsymbol{S_i}$,

$$c_k = \frac{1}{N}\sum_{n=1}^{N-k}\left(\boldsymbol{d'}(n) - \overline{\boldsymbol{d'}}\right)\left(\boldsymbol{d'}(n+k) - \overline{\boldsymbol{d'}}\right), \quad k = 0, 1, \ldots, N-1. \tag{14}$$

From Figure 8, we observe that the autocorrelation function of periodic Doppler shift sequence looks like a sinusoid, but its amplitude decreases gradually, while the autocorrelation function of aperiodic Doppler shift sequence varies irregularly. Based on this characteristic, the rules are built as follows to judge whether intrusion happens. Suppose the coordinate value of first three peaks of the autocorrelation are (p_1, l_1), (p_2, l_2), (p_3, l_3), respectively; then the peak intervals are $PkIntevs = [l_1, l_2 - l_1, l_3 - l_2]$. If the following rules are satisfied, then we judge intrusion happening,

$$\begin{cases} p_1 > PkThrd, \ and \ p_1 > p_2 > p_3, \ and \\ \frac{max(PkIntvs) - min(PkIntvs)}{mean(PkIntvs)} > IntvThrd \end{cases}$$

The first rule is designed to ensure that the shapes of the Doppler shift sequence within all periods are similar enough. The second rule ensures that the lengths of all the periods are almost constant. The combination of the above two rules ensures that the detected Doppler shift is strong periodic. According to the experimental results in Section 5.6, the threshold *PkThrd* and *PkIntvs* are suggested to set as 0.3–0.4 and 0.1–0.2, respectively.

5 EVALUATION

In this section, we conduct comprehensive experiments to evaluate *AudioGuard*. First, we evaluate *AudioGuard* with LOS intrusion in office, laboratory, and home environments. We then test the performance for NLOS intrusion detection with five different settings in a real home environment. Finally, we evaluate its robustness. We test the impact of different transceivers, the variation of transceiver's location and orientation, different walking speeds, and various interference. Finally, we discuss the limitations of *AudioGuard*. The demo video of *AudioGuard* is available at https://tinyurl.com/4y44pdbk and https://youtu.be/iI-Pk4st75o.

5.1 Prototype Implementation

We implement *AudioGuard* on two different acoustic transceivers, shown in Figure 10. Two transceivers have the same commercial microphone (SAMSON MeteorMic, 16 bit, 48 kHz, 96 dB, 20 Hz–20 kHz) but different speakers. The speaker in transceiver 1 is a commercial speaker (JBL Jembe, 6 W, 80 dB, 80 Hz–20 kHz), while the speaker in transceiver 2 is a customized speaker (50 W, 96 dB, 1 kHz–40 kHz). We use transceiver 2 by default and compare the sensing range of transceiver 1 and transceiver 2 in Section 5.4.1. The acoustic transceiver is connected to a Lenovo laptop (Intel Core i7-7500UCPU, 8 GB RAM). The intrusion detection algorithm is implemented in MATLAB and runs in real time. The frequency of transmitted signal is 20 kHz. The length of echo frame is 0.1 second.

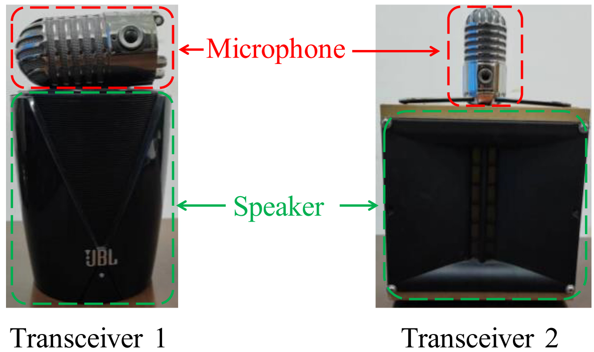


Fig. 10. Two different transceivers.

Fig. 11. LOS intrusion evaluation in three different rooms.

Table 1. Experimental Result of LOS Intrusion Detection

Room	Number of experiments	Number of miss report	Miss report rate
Office	105	0	0%
Laboratory	121	0	0%
Home	110	0	0%

5.2 Evaluation for LOS Intrusion

As shown in Figure 11, we conduct an LOS intrusion detection experiment in three different rooms, i.e., office (5.2 m × 3 m), laboratory (8.4 m × 5.8 m), and home (irregular shape, 148 m^2). In each environment, four subjects are recruited to enter the room or walk freely in the room for at least 100 times to test the miss report rate (i.e., False Negative Rate). Note that in the home setting, i.e., the third setting, to ensure the intrusion is in LOS, the subjects are required to walk in the areas highlighted as yellow.

Table 1 shows the experimental result. We can see that *AudioGuard* accurately captures all the intrusion. Because there is no restriction on subjects' walking path, subjects change walking path dynamically during walking in part of the experiments. So, the result also indicates that within its sensing range (more than 120 m^2, refer to Section 5.5) *AudioGuard* is robust to intruder's location and walking direction.

5.3 Evaluation with NLOS Intrusion

We conduct NLOS intrusion detection experiment in a real home environment. Figure 12 shows the settings. The red circles mark the location of the transceiver. The areas highlighted as yellow are the places where intrusion happens. The settings ensure that the intrusion happens in NLOS area for transceiver.

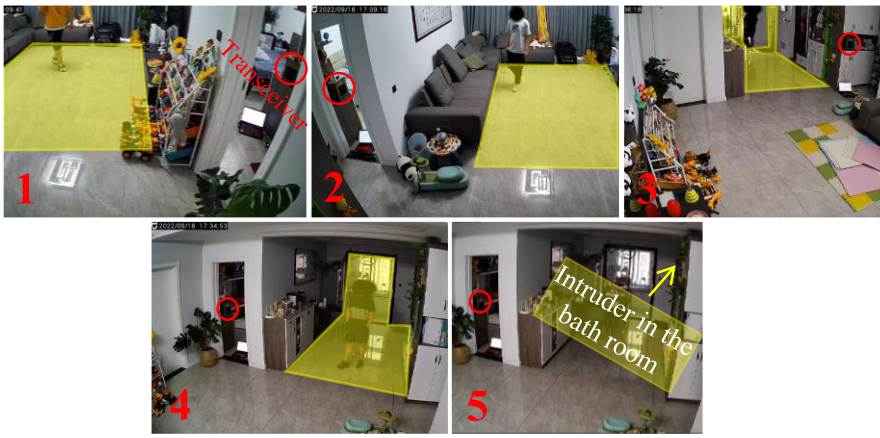


Fig. 12. NLOS intrusion evaluation settings.

Table 2. Experimental Result of NLOS Intrusion Detection

Setting Num.	Number of experiments	Number of miss report	Miss report rate
1	105	1	0.95%
2	121	2	1.65%
3	110	0	0
4	114	2	1.75%
5	100	87	87%

In each setting, three subjects are recruited to walk in the yellow area freely for at least 100 times to test the miss report rate. Table 2 shows the experimental result. We can see that in the first four settings *AudioGuard* can accurately capture intrusions. However, in setting 5, *AudioGuard* fails because of the lack of effective reflected signal (i.e., the received multipath signal reflected from intruder). Similar results occur if the door is closed in settings 1, 2, and 4. In other words, if the effective reflected signal is blocked (e.g., door is closed) or the propagation path is too complex, requiring too many times of reflection (e.g., setting 5), then *AudioGuard* fails to detect intrusion.

5.4 Robustness Evaluation

5.4.1 Impact of Different Transceiver. The sensing range is related to transmitting power. In the same environment, larger transmitting power results in larger sensing range. We conduct experiments in a laboratory (8.4 m × 6 m × 3.4 m) and a large lobby (38 m × 8 m × 9 m) to compare the sensing range of two different transceiver mentioned in Section 5.1. Specifically, as shown in Figures 13(a) and 13(b), we divide the space of laboratory and lobby into 1 × 1 m^2 squares. To make full use of the space, we first place the transceiver at one corner of the laboratory and lobby to measure the sensing range in front of the transceiver. Two subjects are asked to walk away from the transceiver at each square. The blue arrows and the dots in Figure 13(a) and (b) denote the walking direction and the start point of walking. The red circles linked with arrow denote the location and orientation of transceiver in Figure 13(a)–(f). If *AudioGuard* successfully detects the walking events, then the square where the subject starts walking from is detectable, else the square is undetectable. To estimate the detectable area in the back of transceiver, the transceiver is placed facing toward the wall and the subject walks in the area in the back of transceiver.

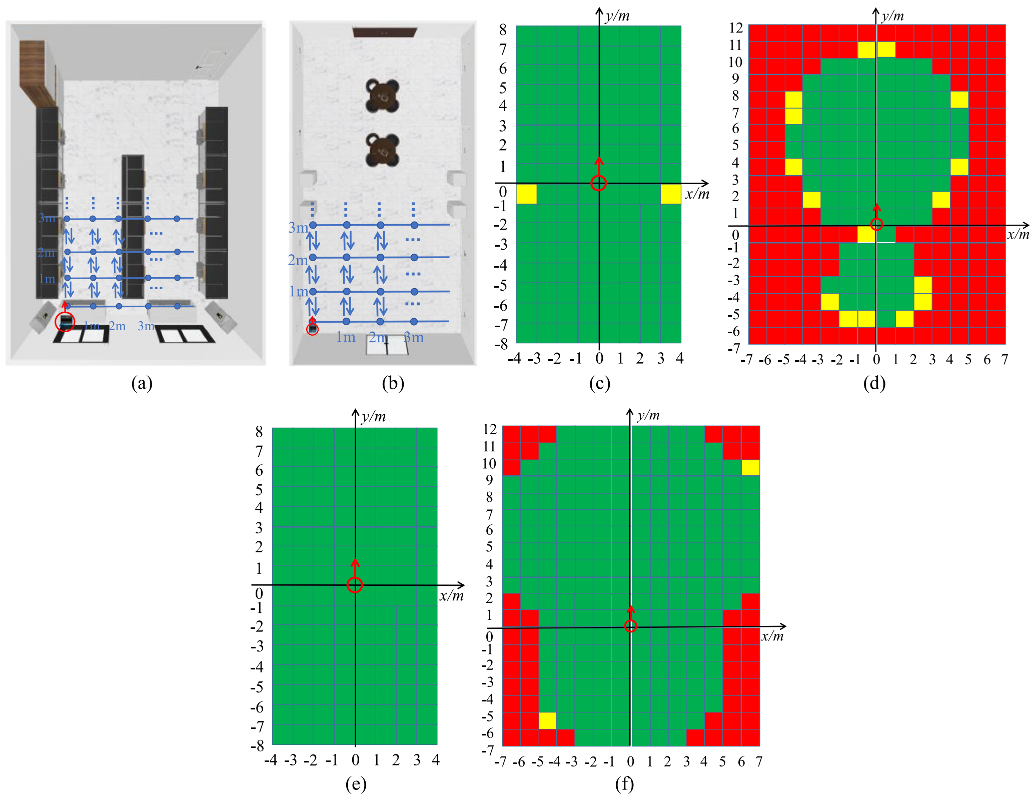
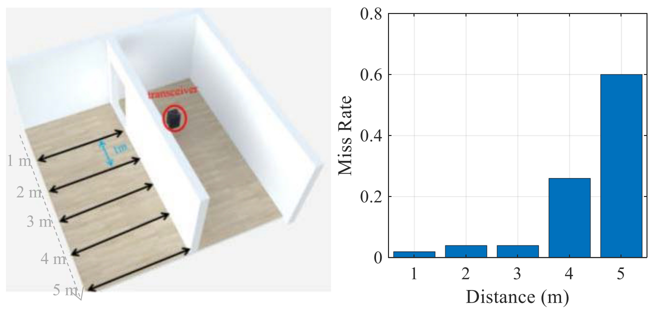


Fig. 13. Sensing range of different transceivers in different environment.

The sensing range of transceiver 1 in the laboratory and lobby are shown in Figures 13(c) and 13(d). The sensing range of transceiver 2 is shown in Figures 13(e) and 13(f). The green squares denote that *AudioGuard* successfully captures the walking events of both subjects in these squares. The yellow squares colored denote that *AudioGuard* only captures the walking event of one subject and fails to detect the walking event of another subject in these squares. The red squares denote that *AudioGuard* fails to capture the walking events of both subjects in these squares.

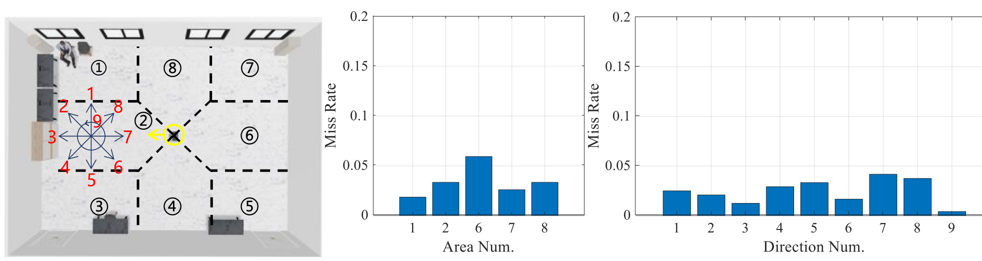
We observe some interesting phenomena. First, the sensing range of transceiver 2 is larger than that of transceiver 1, especially in the lobby. It is reasonable that the larger transmitting power of the transceiver results in a larger sensing range. Second, comparing the sensing range of transceiver 1 in the laboratory and lobby (as shown in Figures 13(c) and 13(d)), we find that the space of the lobby is much larger than that of the laboratory, but the detectable area in the laboratory is larger than that in the lobby. However, the detectable area of transceiver 2 in the lobby (as shown in Figure 13(f)) is larger than that in the laboratory (as shown in Figure 13(e)). It indicates that the sensing range is related to both the transmitting power and the space of the environment. Larger rooms may have weaker multipath reflections and result in a smaller sensing range.

5.4.2 Impact of NLOS Distance. From the experimental results in Section 5.3, we know that the detectability of NLOS intrusion depends on the complexity of signal propagation path, which is closely related to NLOS distance (i.e., the distance along wall that blocks LOS between transceiver and intruder). In this section, we evaluate the impact of NLOS distance on intrusion detecting accuracy. Specifically, the evaluation environment is shown in Figure 14(a). Two adjoining rooms



(a) Experimental settings (b)Miss report rate of each path

Fig. 14. Impact of NLOS distance.



(a) The areas and walking directions (b) Missing report rate in each area (c) Miss report rate of each direction

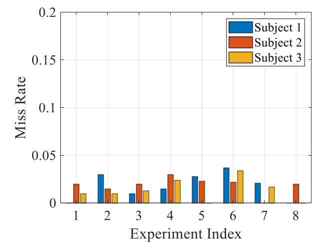
Fig. 15. Evaluation with various location and walking direction of intruder.

(6 m × 3 m) are connected by a door. The transceiver is placed in one room, and the intrusion happens in the other room. The black arrows denote five walking paths that have different distances from the wall, 1–5 m. For each path, the one subject is recruited to walk 50 times. Figure 14(b) shows the miss report rates when the subject walks along different paths. When the distance is larger than 3 m, the miss report rate increases significantly. Though increasing transmitting power may mitigate this problem, as NLOS distance increases, failing to detect intrusion is inevitable.

5.4.3 Impact of the Location and Walking Direction of Intruder. Adapting to the variation of intruder's location and walking direction is necessary for intrusion detection in real application environment. We now conduct experiments to test the impact of the location and walking direction of intruder on intrusion detection. As shown in Figure 15(a), we divide the room (10 m × 8 m × 3.4 m) into eight areas with different shapes. The transceiver is placed at the center facing left (highlighted as the yellow circle with an arrow). Fifteen participants (including 5 females and 10 males) are recruited to walk in each area toward nine different directions (direction 9 is clockwise circle) to test the missing report rate. Each subject walks for two times toward each direction in each area. From the layout of the areas, areas 1 and 3, areas 7 and 5, areas 8 and 4 are all symmetric, respectively. So, we only need to test *AudioGuard* in areas 1, 2, 6, 7, and 8.

We summarize the experimental results from two aspects. Figure 15(b) shows the missing report rate in each area. We observe that except for area 6, the missing report rates of other areas are lower than 5%. The missing report rates of area 6 are slightly higher, because area 6 is directly behind the transceiver. From the result of experiments in Section 5.4.2, the detectable area behind the transceiver is smaller than other orientations. In other words, the detectability of area 6 is relatively lower than other areas. Figure 15(c) shows the statistics of missing report rate of each

(a) Location and orientation of transceiver in each experiment.



(b) Miss report rate of each experiment.

Fig. 16. Evaluation with various location and orientation of transceiver.

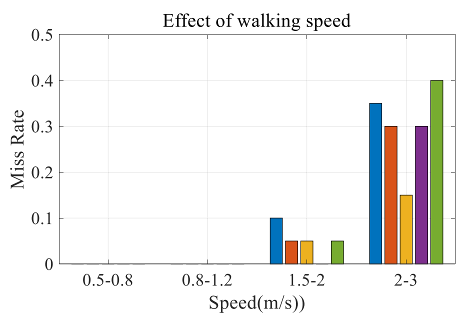


Fig. 17. Effect of walking speed.

direction. We see that all the missing report rates are lower than 4%. In summary, the results indicate that *AudioGuard* is robust against intruder's location and walking direction.

5.4.4 Impact of the Location and Orientation of Transceiver. Robustness to the variation of transceiver's location and orientation is important for practical application. We place the transceiver randomly to test the missing report rate.

Figure 16(a) shows the transceiver's location and orientation in eight experiments. In each experiment, three subjects are recruited to enter the room or walk freely in the room (8.4 m × 5.8 m). Figure 16(b) shows the miss report rate in each experiment. We observe that the miss report rates are all lower than 4%. It indicates that *AudioGuard* is robust against the variation of transceiver's location and orientation.

5.4.5 Impact of Walking Speed. Different people have different walking speeds. It is necessary to test the robustness against different speeds. Five subjects are recruited to evaluate the miss report rate of *AudioGuard*. Subjects are required to walk at four speed levels: 0.5–0.8 m/s, 0.8–1.2 m/s, 1.5–2 m/s, and 2–3 m/s. There is no restriction for walking path. The experimental result is shown in Figure 17. We observe that there is no miss report when the walking speed is lower than 1.2 m/s, which is a normal walking speed. When walking fast, i.e., the speed reaches 1.5–2 m/s, the miss report rate is still lower than 10%. When running, i.e., the speed reaches 2–3 m/s, the miss report rate increases significantly. It is reasonable that when running the stride frequency is about 2–4 steps per second, i.e., 2–4 Hz. As mentioned in Section 4.2, the length of the received echo frame is 0.1 second. It means that the sampling rate of Doppler shift is $1/0.1 = 10$ Hz, which is too low to clearly depict the periodicity of Doppler shift sequence caused by running. Thus, when running, the miss report rate increases significantly.

Table 3. Interference

Interference	Number of experiments
Talking	20
Knocking door	20
Playing music	20
Opening or closing door and window	20
Object falling	20
Curtain fluttering	20

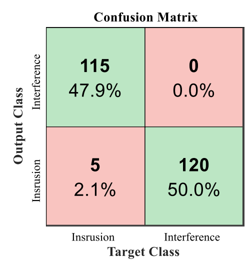


Fig. 18. Confusion matrix.

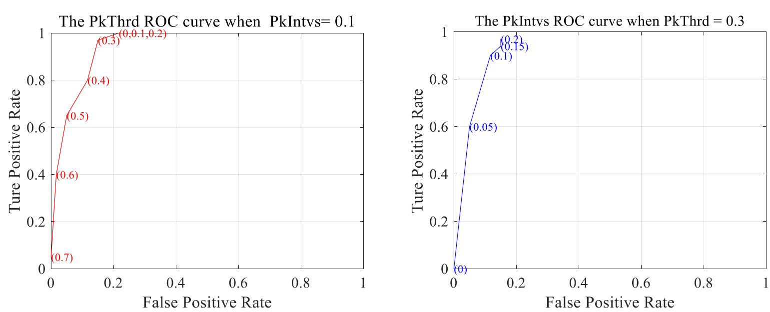


Fig. 19. Impact of two key thresholds, $PkThrd$ and $PkIntvs$.

5.4.6 *Impact of Interference.* We now conduct an experiment to test the robustness against noise interference (talking, knocking door, playing music) and movement interference (opening or closing door and window, object falling, curtain fluttering).

As shown in Table 3, there are six different interferences, and each interference is repeated for 20 times. A total of 120 times interference randomly mixed with an equal number of walking are used to test both the miss report rate and false alarm rate (i.e., False Positive Rate, FPR). Figure 18 shows the confusion matrix.

5.4.7 *Impact of Key Thresholds.* We now evaluate the impact of two key thresholds, i.e., $kThrd$ and $PkIntvs$ (refer to Section 4.4), on miss report rate and false alarm rate. We know that common ambient noise cannot incur false alarms, as they do not incur Doppler shift. So, in this experiment, we only add three kinds of movement events (including opening or closing door and window, object falling, and curtain fluttering) as interference. Specifically, intrusion and movement interference randomly happen for 100 times and 60 times respectively. Figure 19(a) shows the ROC curve when $PkIntvs$ = 0.1 and $kThrd$ varies from 0.1 to 0.6 with a step of 0.1. Figure 19(a) shows

Table 4. System Runtime and Latency

	Number of Iterations	Mean	Variance	Minimum	Maximum
No any movement	1,000	0.027	0.000	0.025	0.036
Movement interference happens	1,000	0.027	0.000	0.025	0.041
Intrusion happens	1,000	0.028	0.003	0.026	0.049

the ROC curve when $kThrd = 0.3$ and $PkIntvs$ varies from 0 to 0.2 with a step of 0.05. According to above results, $PkThrd$ and $PkIntvs$ are suggested to set as 0.3–0.4 and 0.1–0.2, respectively.

5.5 Consuming Time Evaluation

In this section, we conduct an experiment to test the computational overhead of *AudioGuard*. As presented in Section 4.2, the length of the received audio frame by the microphone is 0.1 second. The iteration cycle of *AudioGuard* is equal to the length of audio frame. To ensure that *AudioGuard* runs in real time, the consuming time of signal processing have to be shorter than 0.1 second, or else frame drop occurs. We test the consuming time of the signal processing of *AudioGuard* under three different settings: (i) no movement occurs, (ii) movement interference occurs, and (ii) intrusion occurs. For each setting,

AudioGuard iterates for 1,000 times, and we record the consuming time. Table 4 shows the statistics of the consuming time under three settings. We observe that the maximums of consuming time of signal processing under three settings are all smaller than 0.049 seconds, which is far less than the upper bound 0.1 second.

5.6 Discuss

We now discuss limitations in the current implementation.

(1) *AudioGuard can only detect the intrusion caused by single intruder.* It cannot detect intrusion caused by multiple intruders. *AudioGuard* detects intrusion by measuring the periodicity of Doppler shift. When multiple intruders walk simultaneously, their different stride frequencies (non-multiple) incur aperiodic Doppler shift, which will be regarded as interference by *AudioGuard*. A promising way to tackle this problem is decomposing Doppler shift sequence using the method such as time-frequency domain analysis and independent component analysis.

(2) *AudioGuard is sensitive to periodical movement interference.* As mentioned in Section 3.3, the basic idea of *AudioGuard* is to capture the periodical Doppler shift sequence; when periodical movement interference happens, *AudioGuard* may experience false alarms. It is worth noting that the periodical sound interference (such as the sound of knocking door, the sound of footsteps outside the room, the sound of periodical music, etc.) will not result in false alarm, because pure sound will not introduce a Doppler shift. To avoid periodical movement interferences, other features of Doppler shift need to be extracted to distinguish the category of moving object.

(3) *AudioGuard may miss the intrusion when intruder is running.* As mentioned in Section 5.4.5, running will blur the periodicity of Doppler shift sequence, which may finally lead to miss report.

(4) *AudioGuard may miss the intrusion in NLOS condition if the reflected signal is seriously blocked.* As mentioned in Section 5.3, if the reflected signal is completely blocked by door, then *AudioGuard* fails. Additionally, if the propagation path between the area where the transceiver locates and the area where the intruder locates is too complex, requiring too

many times of reflection, then *AudioGuard* fails. This problem may be solved by exploiting room **Channel Impulse Response (CIR)** estimation, which is able to quantify both the time delay and amplitude attenuation of the multipath signals. CIR is more sensitive to movement than Doppler effect in indoor environment.

ACKNOWLEDGMENTS

The authors express their special appreciation to all the volunteers for participating in our experiments.

REFERENCES

[1] Young-Keun Choi, Ki-Man Kim, Ji-Won Jung, Seung-Yong Chun, and Kyu-Sik Park. 2005. Acoustic intruder detection system for home security. *IEEE Trans. Cons. Electr.* 51, 1 (Feb. 2005), 130–138. http://doi.org/10.1109/TCE.2005.1405710

[2] Yanni Yang, Jiannong Cao, Xiulong Liu, and Xuefeng Liu. 2020. Door-monitor: Counting in-and-out visitors with COTS WiFi devices. *IEEE IoT J.* 7, 3 (Mar. 2020), 1704–1717. http://doi.org/10.1109/JIOT.2019.2953713

[3] Yuepeng Li, Jun Yang, Xiaodong Li, and Jing Tian. 2006. Ultrasonic intruder detection system for home security. In *Intelligent Control and Automation, Vol.* 344. Springer, Berlin. http://doi.org/10.1007/978-3-540-37256-1_143

[4] Zieger Christian, Alessio Brutti, and Piergiorgio Svaizer. 2009. Acoustic based surveillance system for intrusion detection. In *Proceedings of the 6th IEEE International Conference on Advanced Video and Signal Based Surveillance.* IEEE, 314–319. http://doi.org/10.1109/AVSS.2009.49

[5] Yuxiang Lin, Y. Gao, Bingji Li, and Wei Dong. 2020. Revisiting indoor intrusion detection with WiFi Signals: Do not panic over a Pet!. *IEEE IoT J.* 7, 10 (Oct. 2020), 10437–10449. http://doi.org/10.1109/JIOT.2020.2994101

[6] Tianben Wang, Daqing Zhang, Leye Wang, Yuanqing Zheng, Tao Gu, Bernadette Dorizzi, and Xingshe Zhou. 2019. Contactless respiration monitoring using ultrasound signal with Off-the-Shelf audio devices. *IEEE IoT J.* 6, 2 (Apr. 2019), 2959–2973. http://doi.org/10.1109/JIOT.2018.2877607

[7] G. Milanesi, A. Sarti, and S. Tubaro. 2002. Real-time video analysis for intrusion detection in indoor environments. In *Proceedings of the 11th European Signal Processing Conference.* IEEE, 1–4.

[8] Manoranjan Paul, Shah M. E. Haque, and Subrata Chakraborty. 2013. Human detection in surveillance videos and its applications-a review. *EURASIP J. Adv. Sign. Process.* 176 (Nov. 2013), 1–16. http://doi.org/10.1186/1687-6180-2013-176

[9] Bo-Wei Chen, Chen-Yu Chen, and Jhing-Fa Wang. 2013. Smart homecare surveillance system: Behavior identification based on state-transition support vector machines and sound directivity pattern analysis. *IEEE Trans. Syst. Man Cybernet.: Syst.* 43, 6 (Nov. 2013), 1279–1289. http://doi.org/10.1109/TSMC.2013.2244211

[10] Rashmiranjan Nayak, Mohini Mohan Behera, Umesh Chandra Pati, and Santos Kumar Das. 2019. Video-based Real-time intrusion detection system using deep-learning for smart city applications. In *Proceedings of the IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS'19).* IEEE, 1–6. http://doi.org/10.1109/ANTS47819.2019.9117960

[11] Ju Han and B. Bhanu. 2005. Human activity recognition in thermal infrared imagery. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) Workshops.* IEEE. http://doi.org/10.1109/CVPR.2005.469

[12] Yun Li, Yong Song, Yufei Zhao, Shangnan Zhao, Xu Li, Lin Li, and Songyuan Tang. 2017. An infrared target detection algorithm based on lateral inhibition and singular value decomposition. *Infrared Phys. Technol.* 85 (Sep. 2017), 238–245. http://doi.org/10.1016/j.infrared.2017.07.005

[13] Khirod Chandra Sahoo and Umesh Chandra Pati. 2017. IoT based intrusion detection system using PIR sensor. In *Proceedings of the 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT'17).* IEEE, 1641–1645. http://doi.org/10.1109/RTEICT.2017.8256877

[14] Sami Aldalahmeh, Amer Hamdan, Mounir Fhogho, and Des McLernon. 2016. Enhanced-range intrusion detection using pyroelectric infrared sensors. In *Proceedings of the Sensor Signal Processing for Defence (SSPD'16).* IEEE, 1–5. http://doi.org/10.1109/SSPD.2016.7590597

[15] Michael Otero. 2005. Application of a continuous wave radar for human gait recognition. In *Signal Processing, Sensor Fusion, and Target Recognition XIV* 5809, 538–548. http://doi.org/10.1117/12.607176

[16] Milenko S. Andrić, Boban P. Bondžulić, Dimitrije M. Bujaković, and Srđan T. Mitrović. 2011. Analysis of radar doppler echoes from various ground moving targets. In *Proceedings of the International Conference on Aerospace Sciences and Aviation Technology.* 1–11. http://doi.org/10.21608/ASAT.2011.23243

[17] Fioranelli Francesco, Matthew Ritchie, and Hugh Griffiths. 2015. Multistatic human micro-Doppler classification of armed/unarmed personnel. *IET Radar Sonar Navig.* 9, 7 (Aug. 2015), 857–865. http://doi.org/10.1049/iet-rsn.2014.0360

[18] Mu Zhou, Yaoping Li, Liangbo Xie, and Wei Nie. 2019. Maximum mean discrepancy minimization based transfer learning for indoor WLAN personnel intrusion detection. *IEEE Sens. Lett.* 3, 8 (Aug. 2019), 1–4. http://doi.org/10.1109/LSENS.2019.2932099

[19] Mu Zhou, Yaoping Li, Xiaoge Huang, Qianlin Pu, and Hui Yuan. 2019. Indoor WLAN intrusion detection using intra-class transfer learning with low effort. In *Proceedings of the IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'19)*. IEEE, Istanbul, Turkey, 1–6. http://doi.org/10.1109/PIMRC.2019.8904445

[20] Jiguang Lv, Dapeng Man, Wu Yang, Xiaojiang Du, and Miao Yu. 2017. Robust WLAN-based indoor intrusion detection using PHY layer information. *IEEE Access* 6 (Dec. 2017), 30117–30127, http://doi.org/10.1109/ACCESS.2017.2785444

[21] Jiguang Lv, Wu Yang, Liangyi Gong, Dapeng Man, and Xiaojiang Du. 2016. Robust WLAN-Based indoor fine-grained intrusion detection. In *Proceedings of the IEEE Global Communications Conference (GLOBECOM'16)*. IEEE, 1–6. https://doi.org/10.1109/GLOCOM.2016.7842238

[22] Mu Zhou, Yixin Lin, Nan Zhao, Qing Jiang, Xiaolong Yang, and Zengshan Tian. 2020. Indoor WLAN intelligent target intrusion sensing using ray-aided generative adversarial network. *IEEE Trans. Emerg. Top. Comput. Intell.* 4, 1 (Feb. 2020), 61–73. https://doi.org/10.1109/TETCI.2019.2892748

[23] Yue Jin, Zengshan Tian, Mu Zhou, Ze Li, and Zhenyuan Zhang. 2018. A whole-home level intrusion detection system using WiFi-enabled IoT. In *Proceedings of the 14th International Wireless Communications & Mobile Computing Conference (IWCMC'18)*. IEEE, 494–499. https://doi.org/10.1109/IWCMC.2018.8450442

[24] Mohamed Hadi Habaebi, Mahamat Mahamat Ali, M. M. Hassan, M. S. Shoib, A. A. Zahrudin, A. A. Kamarulzaman, W. S. WanAzhan, and Md. RafiqulIslam. 2015. Development of physical intrusion detection system using Wi-Fi/ZigBee RF signals. *Proc. Comput. Sci.* 76 (2020), 547–552. https://doi.org/10.1016/j.procs.2015.12.342

[25] Zengshan Tian, Xiangdong Zhou, Mu Zhou, Shuangshuang Li, and Luyan Shao. 2015. Indoor device-free passive localization for intrusion detection using multi-feature PNN. In *Proceedings of the 10th International Conference on Communications and Networking in China (ChinaCom'15)*. IEEE, 272–277. https://doi.org/10.1109/CHINACOM.2015.7497950

[26] Enjie Ding, Xiansheng Li, Tong Zhao, Lei Zhang, and Yanjun Hu. 2018. A robust passive intrusion detection system with commodity WiFi devices. *J. Sens.* 2018 (June. 2018), 1–12. https://doi.org/10.1155/2018/8243905

[27] Chong Han, Qingqing Tan, Lijuan Sun, Hai Zhu, and Jian Guo. 2018. Csi frequency domain fingerprint-based passive indoor human detection. *Information* 9, 4 (Apr. 2018), 95–108. https://doi.org/10.3390/info9040095

[28] Dan Wu, Youwei Zeng, Ruiyang Gao, Shengjie Li, Yang Li, Rahul C. Shah, Hong Lu, and Daqing Zhang. 2021. WiTraj: Robust indoor motion tracking with WiFi signals. *IEEE Trans. Mobile Comput.* (Dec. 2021), 1–1. https://doi.org/10.1109/TMC.2021.3133114

[29] Zengshan Tian, Yong Li, Mu Zhou, and Ze Li. 2018. WiFi-Based adaptive indoor passive intrusion detection. In *Proceedings of the IEEE 23rd International Conference on Digital Signal Processing (DSP'18)*. IEEE, 1–5. https://doi.org/10.1109/ICDSP.2018.8631613

[30] Shengjie Li, Xiang Li, Kai Niu, Hao Wang, Yue Zhang, and Daqing Zhang. 2017. Ar-alarm: An adaptive and robust intrusion detection system leveraging csi from commodity wi-fi. *International Conference on Smart Homes and Health Telematics*. Springer, Cham, 211–223. https://doi.org/10.1007/978-3-319-66188-9_18

[31] Shengjie Li, Zhaopeng Liu, Yue Zhang, Xiaopeng Niu, Leye Wang, and Daqing Zhang. 2019. A real-time and robust intrusion detection system with commodity wi-fi. In *Adjunct Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the ACM International Symposium on Wearable Computers (UbiComp/ISWC'09)*. 316–319. https://doi.org/10.1145/3341162.3343789

[32] Shengjie Li, Zhaopeng Liu, Yue Zhang, Qin Lv, Xiaopeng Niu, Leye Wang, and Daqing Zhang. 2020. WiBorder: Precise Wi-Fi based boundary sensing via through-wall discrimination. *Proc. ACM Interact. Mobile Wear. Ubiq. Technol.* 4, 3 (Sep. 2020), 1–30. https://doi.org/10.1145/3411834

[33] Omar Sonbul and Alexander N. Kalashnikov. 2013. Low cost ultrasonic wireless distributed security system for intrusion detection. In *Proceedings of the IEEE 7th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS'13)*. IEEE, 235–238. https://doi.org/10.1109/IDAACS.2013.6662679

[34] R. Unni and U. C Pati. 2018. PC based ultrasonic intrusion detection system. In *Proceedings of the International Conference on Communication and Signal Processing (ICCSP'18)*. IEEE, 942–947. https://doi.org/10.1109/ICCSP.2018.8524262

[35] Thilina Dissanayake, Takuya Maekawa, Daichi Amagata, and Takahiro Hara. 2018. Detecting door events using a smartphone via active sound sensing. *Proc. ACM Interact. Mobile Wear. Ubiq. Technol.* 2, 4 (Dec. 2018), 1–26. https://doi.org/10.1145/3287038

[36] https://safeatlast.co/blog/burglary-statistics/

[37] Michael Feldman. 2011. *Hilbert Transform Applications in Mechanical Vibration: Feldman/Hilbert Transform Applications in Mechanical Vibration.* John Wiley & Sons.

[38] Xingshui Zu, Feng Guo, Jingchang Huang, Qin Zhao, Huawei Liu, Baoqing Li, and Xiaobing Yuan. 2017. Design of an acoustic target intrusion detection system based on small-aperture microphone array. *Sensors* 17, 3 (Mar. 2017), 514. https://doi.org/10.3390/s17030514

[39] Chia-How Lin and Kai-Tai Song. 2013. Probability-based location aware design and on-demand robotic intrusion detection system. *IEEE Trans. Syst. Man Cybernet.: Syst.* 44, 6 (Jun. 2013), 705–715. https://doi.org/10.1109/TSMC.2013.2277691

[40] Kai Niu, Fusang Zhang, Xuanzhi Wang, Qin Lv, Haitong Luo, and Daqing Zhang. 2021. Understanding WiFi signal frequency features for position-independent gesture sensing. *IEEE Trans. Mobile Comput.* 21, 11 (Mar. 2021), 4156–4171. https://doi.org/10.1109/TMC.2021.3063135

[41] Tianben Wang, Daqing Zhang, Leye Wang, Yuanqing Zheng, Tao Gu, Bernadette Dorizzi, and Xingshe Zhou. 2018. Contactless respiration monitoring using ultrasound signal with off-the-shelf audio devices. *IEEE IoT J.* 6, 2 (Apr. 2018), 2959–2973. https://doi.org/10.1109/JIOT.2018.2877607

Received 31 October 2022; revised 7 April 2023; accepted 25 July 2023