

IRShield: A Countermeasure Against Adversarial Physical-Layer Wireless Sensing

Paul Staat¹, Simon Mulzer², Stefan Roth²,
Veelasha Moonsamy², Aydin Sezgin², and Christof Paar^{1,2}

¹Max Planck Institute for Security and Privacy, Bochum, Germany

²Ruhr University Bochum, Bochum, Germany

E-Mail: {paul.staat, christof.paar}@mpi-sp.org, {stefan.roth-k21, aydin.sezgin}@rub.de, simon@mulzer.eu, email@veelasha.org

Abstract—Wireless radio channels are known to contain information about the surrounding propagation environment, which can be extracted using established wireless sensing methods. Thus, today’s ubiquitous wireless devices are attractive targets for passive eavesdroppers to launch reconnaissance attacks. In particular, by overhearing standard communication signals, eavesdroppers obtain estimations of wireless channels which can give away sensitive information about indoor environments. For instance, by applying simple statistical methods, adversaries can infer human motion from wireless channel observations, allowing to remotely monitor premises of victims. In this work, building on the advent of intelligent reflecting surfaces (IRSs), we propose IRShield as a novel countermeasure against adversarial wireless sensing. IRShield is designed as a plug-and-play privacy-preserving extension to existing wireless networks. At the core of IRShield, we design an IRS configuration algorithm to obfuscate wireless channels. We validate the effectiveness with extensive experimental evaluations. In a state-of-the-art human motion detection attack using off-the-shelf Wi-Fi devices, IRShield lowered detection rates to 5% or less.

I. INTRODUCTION

Wireless connectivity drives many current digital innovations and is becoming increasingly ubiquitous. This trend manifests itself in a worldwide surge in the adoption of Internet of Things (IoT) devices with 75 billion connected devices projected by the year 2025 [20]. The IoT phenomenon is already heavily present in our daily lives: voice assistants, watches, locks, light bulbs, cameras, vacuum cleaners, sensors, and actuators have often turned into devices that we consider *smart*. However, virtually all of them rely on wireless connectivity, based on standards such as Wi-Fi, BLE, NB-IoT, or ZigBee [16].

A downside of having ubiquitous wireless communication is the burgeoning of new types of privacy concerns. Since wireless communication is based on an open medium, it is inherently shared with third, potentially adversarial, parties. Although cryptographic primitives are widely used for providing confidentiality and integrity of the application data, passive eavesdroppers are still able to exploit sensitive information from sniffed radio frequency (RF) signals. This is possible as the propagation of RF signals depends on the physical surroundings of devices, e. g., due to reflections off walls, objects, and individuals. Eavesdroppers can remotely observe such propagation effects and subsequently gather insights about the

physical environment of a legitimate transmitter, referred to as (adversarial) wireless sensing. As already demonstrated by existing work in the literature, this leads to an immediate privacy threat due to reconnaissance attacks, such as adversarial human motion sensing.

Zhu et al. [32] demonstrated the feasibility of carrying out fine-grained detection and tracking of human movements inside a building by only sniffing ordinary Wi-Fi signals. Their low-cost attack makes use of the Wi-Fi signal dynamics and variance of multipath signal propagation to track indoor movements without requiring any a priori knowledge on the adversary’s side. Similarly, Banerjee et al. [5] proposed a methodology to detect human motion within the line-of-sight (LOS), i. e., the direct path between the victim’s device and that of the passive sniffer. Generally, there is significant potential for (adversarial) wireless sensing on the physical layer, as highlighted by the large corpus of existing work surveyed by Ma et al. [15]. Besides exploiting physical-layer properties, an adversary can also monitor packet-level information to infer the state of IoT devices and actions carried out by users present in the network, cf. [2], [22], [31].

While inference attacks from packet-level wireless sniffing may be thwarted by means of phantom users [13], a versatile and easy to use countermeasure against wireless sensing on the physical layer is still lacking. One prominent approach was put forward by Qiao et al. who proposed PhyCloak [18]. By using a full-duplex radio, ambient wireless signals are re-transmitted with randomized modifications to obfuscate sensitive physical information. However, full-duplex is costly as it requires specialized and complex radio hardware. While eliminating the use of full-duplex equipment, the proposal by Yao et al. [29] requires several motorized moving hardware parts. Other approaches by Zhu et al. [32] and Wijewardena et al. [25] change the allocation of the channel. Therefore, they not only trade the quality of service of wireless communication against adversarial sensing capabilities, but also require integration with the wireless devices in the field.

In this paper, we aim to design a countermeasure against adversarial wireless sensing on the physical layer that overcomes the shortcomings of previous approaches and resolves the following challenges (C1–C3):

- **Device-agnostic countermeasure (C1):** There exists a

wide variety of wireless devices that are present in the real-world, and cannot be remotely updated or modified. Thus, we pursue a solution that is independent of already-deployed devices, the used wireless waveforms, and standards.

- **Maintain quality of service (C2):** The connectivity requirements of wireless applications are diverse and hard to predict, which renders a reduction in quality of service unacceptable. An effective countermeasure should therefore not affect the quality of the wireless link.

Inspired by a recent trend from the realm of wireless communication, we explore *smart radio environments* with intelligent reflecting surfaces (IRSs) to overcome these hurdles. To date, digitally controlled IRS are primarily used as adjustable elements in propagation environments, which improve wireless communication [19]. For wireless sensing, a key aspect of the IRS is that it directly affects the adversary’s observation hitherto used for privacy-violation attacks: the wireless channel. Based on this observation, we propose *IRShield* – a practical IRS-based wireless channel *obfuscation*. It greatly reduces adversarial capabilities for privacy violations from wireless sensing. Specific to the IRS, we address the third challenge:

- **Surface configuration (C3):** In a wireless communication context, IRS are typically configured based on channel information from legitimate receivers and careful integration into the wireless infrastructure [26]. However, in an adversarial wireless sensing scenario, the receiver (eavesdropper) is unknown and hostile and therefore the required channel information is not available. The challenge is thus to provide strong channel obfuscation, given the extremely large IRS configuration space of, e. g., 2^{256} .

IRShield addresses **C3** through a probabilistic surface configuration strategy specifically designed to achieve channel obfuscation as a standalone application. Further, IRShield affects the bare radio wave propagation, regardless of specific devices or waveforms, thus meeting the criteria for addressing **C1**. Finally, IRShield does not allocate the wireless channel itself but affects it randomly, hence solving **C2**.

Contribution. We are the first to propose IRS to be used as a hands-on countermeasure against wireless sensing attacks. We design a dedicated algorithm to generate randomized IRS configurations to achieve wireless channel obfuscation. Our scheme, which we refer to as IRShield, is laid out as a standalone privacy-preserving extension for plug-and-play integration with existing wireless infrastructure. Finally, we present an extensive experimental evaluation of the proposed technique, showing that it successfully prevents adversarial motion sensing. In particular, IRShield lowered detection rates to 5% or less in a state-of-the-art attack on Wi-Fi devices. Our measurement data will be available online in order to support research reproducibility.

TABLE I
TERMINOLOGY OVERVIEW

t	Time index
\mathcal{L}_t	Propagation paths at time t
K	Number of subcarriers
k	Subcarrier index
N_T	Number of transmit antennas
N_R	Number of receive antennas
M	Number of IRS elements
$r_m(t)$	IRS reflection coefficient
$\mathbf{H}(k, t)$	Channel at subcarrier k
$\bar{\sigma}^{(w)}(t)$	Average channel standard deviation

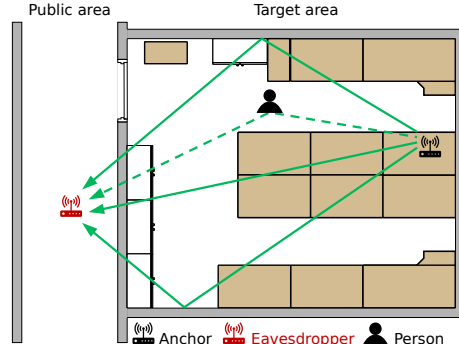


Fig. 1. Illustration of wireless signal propagation in indoor environments.

II. BACKGROUND

We now provide technical background information on wireless sensing methods using the physical layer of wireless communication systems. Afterwards, we give background on smart radio environments, including IRS. For the reader’s convenience, we have summarized the important terminology and symbols in Table I.

A. Wireless Sensing Methods

When wireless radio signals traverse from a transmitter to a receiver, they are distorted by the *channel response* which aggregates environment-dependent effects on the signal such as multipath propagation. In particular, the channel response between an arbitrary transmitter and receiver is determined by different propagation paths through the environment, as illustrated in Fig. 1. These are LOS paths and various non-LOS (NLOS) paths, e. g., from reflections off walls, interior objects, or potentially present humans. Particular to the case of human motion, new paths are created and existing paths are blocked, thus making the wireless signal propagation time-variant. At time t , we assume that the propagation paths \mathcal{L}_t are present.

In a wireless communication context, the (undesired) channel response is regularly estimated from known preamble sequences to be compensated subsequently from the received signal. This, likewise, happens for every received packet when considering IEEE 802.11n Wi-Fi communication based on orthogonal frequency division multiplexing (OFDM). OFDM

divides a wide bandwidth wireless channel into K parallel independent (i.e., orthogonal) narrowband channels, i.e., subcarriers, for data transmission. Thus, upon detection of a packet, a Wi-Fi receiver estimates the channel response for each subcarrier k using a known preamble to obtain

$$\mathbf{H}(k, t) = \underbrace{\sum_{l \in \mathcal{L}_t} \mathbf{G}_l(k, t)}_{\text{propagation paths through surrounding}} + \underbrace{\mathbf{N}(k, t)}_{\text{noise}}. \quad (1)$$

Considering a multiple-input multiple-output (MIMO) system typical for Wi-Fi in which the transmitter or receiver has N_T or N_R antennas, $\mathbf{G}_l(k, t)$ represents the $N_R \times N_T$ -dimensional signal propagation matrix of the l^{th} propagation path. This matrix then involves the complex channel gain and possible steering or response vectors at a certain time t . $\mathbf{N}(k, t)$ refers to measurement noise.

Building upon such channel estimates from conventional wireless communication systems, e.g., Wi-Fi [15], *wireless sensing* seeks to extract information on the physical propagation environment contained in the channel response. Applications such as human activity and gesture recognition, imaging, or vital sign monitoring [30], [15] impressively demonstrate the ability of commodity wireless devices to detect environmental conditions and provide novel use-cases for wireless communication equipment. However, despite its many legitimate uses, wireless sensing can also be deployed by adversaries: radio propagation relies on a shared medium, allowing passive eavesdroppers to receive packets and gain information about remote environments, leading to potential privacy violations of victim parties.

When it comes to the threat of adversarial human motion sensing, it is, first and foremost, important to understand the mechanisms underlying the attack. First, the amplitude of $\mathbf{G}_l(k, t)$ depends on whether a path is temporarily blocked by a person at time t or not. Additionally, human motion can also introduce temporary new propagation paths as the human body also reflects signals. Thus, the impact of human motion on the wireless channel is expected to vary over time, depending on the individual's position and the number of affected propagation paths. Further, due to distance-dependent path loss of radio wave propagation, shorter propagation paths, e.g., LOS or short NLOS paths, contribute most to the channel. This is an important aspect, since a person moving into the LOS affects the strongest propagation paths, yielding significant changes to the wireless channel.

B. Smart Radio Environments

The IRS is considered a promising technology for future wireless networks [14] as it enables *smart radio environments* [19]. More precisely, the IRS is a synthetic surface that has digitally reconfigurable reflection properties of radio waves. This rather new concept is rooted in physics research on metamaterials and metasurfaces [9], which recently saw drastic simplification that led to the IRS. Here, a large number of reflecting elements distributed across a surface are individually

and electronically adjustable, allowing dynamic manipulation of impinging radio waves. For instance, the IRS may be configured to reflect signals into a particular direction. Thus, the IRS allows partial reconfiguration of the radio propagation environment to alter the wireless channel.

We consider an IRS consisting of M reflecting elements with the m^{th} configurable reflection coefficient $r_m(t)$. The reflection coefficient can be selected among a discrete set of values and is changing the reflected signal, such that the incoming wave is multiplied by this factor. For each element, additional propagation paths become part of the channel. Incorporating the IRS contribution, we rewrite Eq. (1) as

$$\mathbf{H}(k, t) = \underbrace{\sum_{l \in \mathcal{L}_t} \mathbf{G}_l(k, t)}_{\text{propagation paths through surrounding}} + \underbrace{\sum_{m=1}^M r_m(t) \mathbf{G}'_m(k, t)}_{\text{propagation paths via IRS}} + \underbrace{\mathbf{N}(k, t)}_{\text{noise}}, \quad (2)$$

where $\mathbf{G}'_m(k, t)$ is the signal propagation via the m^{th} IRS element. Eq. (2) shows how the channel matrices can be modified by engineering the values of $r_m(t)$. On the other hand, randomizing the IRS configuration implies partial randomization of the channel.

III. RELATED WORK

In this section, we summarize relevant literature on privacy violations caused by sniffing wireless devices and report related work on wireless sensing. Furthermore, we briefly outline how our work differs from previous proposals for countermeasures.

With a particular focus on privacy violations, Banerjee et al. [5] and Zhu et al. [32] studied Wi-Fi-based physical-layer attacks to infer human motion from channel observations by eavesdropping packets. In both papers, the authors utilized a sliding-window standard deviation or variance as a measure of temporal channel variation. While [5] focused on detecting LOS crossings, [32] pursues monitoring of entire environments for motion, including differentiation of motion in specific parts of the target environment. The authors of [5] employed Wi-Fi routers with MIMO functionality to implement the attack and [32] presented a smartphone-based implementation.

To secure systems against physical-layer side-channel attacks, different types of countermeasures have been considered. Qiao et al. have proposed PhyCloak [18], a full-duplex relay node to introduce channel variation. Wijewardena et al. [25] put forward a game-theoretic framework to trade the privacy loss against the quality of service of the wireless communication. In particular, to prevent motion sensing from received signal strength variations, transmit power is reduced or devices are temporarily turned off.

Apart from motion sensing, there is a large body of work on Wi-Fi sensing applications as surveyed by Ma et al. [15]. For instance, Wi-Fi signals have already been employed for gesture recognition [1], identification of individuals [10], imaging [8], vital sign monitoring [12], or keystroke recognition [3]. Shifting the scope of wireless sensing, Camurati et al. [6] have

shown that RF signals from Bluetooth and Wi-Fi chipsets can be utilized for side-channel attacks against on-chip cryptographic hardware implementations from rather large distances.

On the packet level, Trimananda et al. [22], Zhang et al. [31], and Acar [2] presented side-channel attacks in smart home environments to infer events such as when individuals leave or come back. For such attacks, Liu et al. [13] recently proposed a countermeasure based on emulating phantom users.

Differentiation from previous work. Countermeasures against adversarial wireless physical-layer sensing have been pursued by previous works. A conceptually closely related work is PhyCloak introduced by Qiao et al. [18]. Using a full-duplex radio, legitimate RF signals are received and re-transmitted while applying randomized signal changes to hamper adversarial sensing. However, full-duplex requires specialized and highly complex radio hardware to achieve self-interference cancellation in order to allow for simultaneously receiving and transmitting on the same frequency. Another proposal by Yao et al. [29], likewise, receives and re-transmits legitimate signals but circumvents full-duplex by means of a directional transmit antenna. To introduce randomized changes to the adversarial observations, however, the setup employs several moving parts, e.g., a motor to rotate the antenna and a fast fan. While IRShield follows the general receive-and-re-transmit rationale, our approach works entirely different. Crucially, we utilize an IRS to unify reception, signal alteration, and re-transmission of signals into digitally adjustable reflection. That is, in contrast to [18] and [29], the IRS is inherently capable of full-duplex operation and does not require any moving parts.

Along with their attack implementation, Zhu et al. [32] proposed a countermeasure based on a fake access point to inject cover packets to mimic legitimate traffic. As a result, the eavesdropper observes a mix of packets transmitted by the victim devices and the fake access point, thus reducing the adversarial sensing capability. To this end, we believe the adversary can easily defeat this strategy since the channel from the fake access point is clearly distinguishable due to the location-dependence of wireless channels [7]. Another interesting approach put forward by Wijewardena et al. [25] is based on strategically randomizing transmit power or turning off devices completely. However, both strategies are based on changing the allocation of the channel and therefore, face a trade-off between the quality of service of wireless communication and reduction of adversarial sensing capabilities. We emphasize that IRShield rather changes the wireless channel itself instead of its allocation. Moreover, IRShield is independent of the existing equipment, i.e., it neither requires adjustments of transmit power or timing, nor does it need to be matched to the legitimate signals.

IV. THREAT MODEL

In this work, we consider a number of legitimate Wi-Fi devices which transmit packets. The devices, which we also refer to as *anchors*, are deployed within a normal environment,

such in a home or an office. The goal of our passive eavesdropper, Eve, is to infer human motion within the environment by analyzing eavesdropped packets. Therefore, we assume that Eve possesses a wireless receiver that is able to pick up and demodulate signals originating from the anchors. Eve does not have access to the environment but can position her receiver at arbitrary public locations outside the perimeter. After initial positioning of the receiver, Eve can then proceed to act remotely. Furthermore, we assume Eve is not able to break the applied cryptography, i.e., to read secured payload data. However, since the anchors employ standard-compliant Wi-Fi communication, Eve can obtain physical-layer channel estimations from the known packet preambles.

We assume that the owner (i.e. the *defender*) of the environment can position the anchor devices at will. Furthermore, we assume the defender can place one or multiple IRS within their space and apply customized configurations.

V. ATTACK SCENARIO

In this section, we outline the adversary's strategy, as described in the state-of-the-art attack in [32], for which we later propose a countermeasure, i.e. IRShield. We also elaborate on the experimental evaluation.

A. Eavesdropper Strategy as per [32]

As outlined in Section IV and illustrated in Fig. 1, we consider an adversary who leverages physical-layer channel observations to infer human motion. This is based on the rationale that the wireless channel available to the eavesdropper is coupled to environmental changes such as human motion. However, the exact relation between the observed channel and the environment is non-trivial, i.e., the channel response cannot be directly used to make elaborate claims on the physical state of the environment. Instead, several preprocessing steps shall be used in order to achieve motion detection. Here, we follow the attack strategy outlined in [32]. Similar to [5], [24], the authors describe an approach based on a sliding-window standard deviation over a time series of channel state information (CSI) magnitudes. The intuition behind this approach is rooted in the observation that a static channel response indicates a steady environment whereas a dynamic (time-varying) channel response indicates a varying environment. To illustrate this connection, we plot a set of Wi-Fi channel estimations with and without movement in Fig. 2 (a).

By eavesdropping Wi-Fi packets at time t from an anchor, Eve obtains a time series of complex-valued estimates of the wireless channel between Eve and the anchor node. At each time instant, Eve gathers a multitude of parallel channel estimates due to diversity from (i) the OFDM transmission on several subcarriers and (ii) several spatial channels for MIMO transmissions. Therefore, Eve can leverage the combined measurements

$$\mathbf{h}(t) = \left(\text{vec}(\mathbf{H}(1, t))^T, \dots, \text{vec}(\mathbf{H}(K, t))^T \right)^T, \quad (3)$$

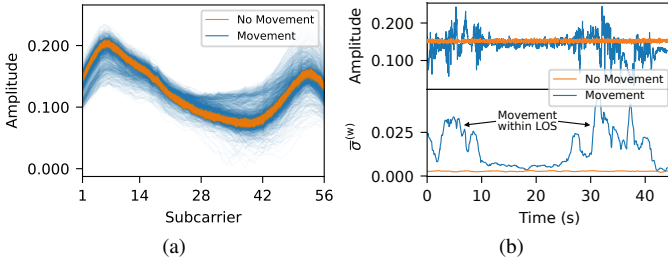


Fig. 2. (a) Ensembles of raw CSI amplitudes over subcarriers from a standard Wi-Fi connection without and with movement in the propagation environment. (b) Raw CSI time-series of a single subcarrier and sliding window standard deviation with window length 1 s with and without movement.

to obtain insights on the victim’s surroundings. In the following, we denote the n^{th} element of such a measurement vector as $h_n(t)$. Due to lacking synchronization, the signal phase of Wi-Fi channel estimates is subject to severe measurement imperfections and thus is not reliable. Therefore, we discard the phase and use the absolute value of $h_n(t) = |h_n(t)|$. Then, Eve uses a windowing technique with window size $w = 1$ s to calculate a moving standard deviation

$$\sigma_n^{(w)}(t) = \text{Std}_{\tau \in \{t-w+1, \dots, t\}}(h_n(\tau)). \quad (4)$$

To illustrate this step, Fig. 2 (b) depicts a single Wi-Fi subcarrier over time (top) and its corresponding sliding-window standard deviation (bottom). It can clearly be seen that the latter quantifies motion-induced temporal variations.

Finally, to take advantage of the available diversity domains [5], we average over $N_R N_T$ spatial channels and K subcarriers, i. e.,

$$\bar{\sigma}^{(w)}(t) = \frac{1}{K N_R N_T} \sum_{n=1}^{K N_R N_T} \sigma_n^{(w)}(t). \quad (5)$$

In the remainder of this paper, we also refer to this averaged sliding-window standard deviation, $\bar{\sigma}^{(w)}(t)$, as *adversarial observation*.

To decide whether $\bar{\sigma}^{(w)}(t)$ indicates motion, [32] utilizes a threshold-based detection, i. e., motion is detected if $\bar{\sigma}^{(w)}(t) > u$, where the threshold u is derived from a long-term reference measurement $\bar{\sigma}_{\text{ref}}^{(w)}(t)$ as

$$u = \text{median}_t \left\{ \bar{\sigma}_{\text{ref}}^{(w)}(t) \right\} + C \cdot \text{MAD}_t \left\{ \bar{\sigma}_{\text{ref}}^{(w)}(t) \right\}, \quad (6)$$

where $\text{median}_t \{\cdot\}$ and $\text{MAD}_t \{\cdot\}$ denote the median and median absolute deviation over t , respectively. According to the implementation details outlined in [32], we adopt the choice of the conservativeness factor $C = 11$. It should be noted that alternative threshold definitions are possible such as, e. g., based on comparison of long-term and short-term sliding-window variances [5].

B. Attack Setup

Since our goal is to develop a countermeasure, we aim for a strong adversarial configuration, using readily-available

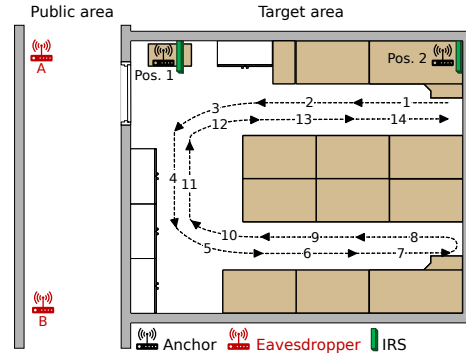


Fig. 3. The experimental setup with different positions of anchor nodes, eavesdroppers, and the motion path of a human.

off-the-shelf equipment. While [32] considered a smartphone-based eavesdropper with a single antenna, here we seek to improve this setup (by extension, strengthening the adversary capabilities) and therefore utilize multi-antenna Wi-Fi routers. In particular, we use TP-Link N750 routers as both anchor nodes and eavesdroppers. These are equipped with $N_R = N_T = 3$ antennas each and implement IEEE 802.11n Wi-Fi [21]. In our experimental setup, the routers run an OpenWrt operating system and provide CSI data from the ath9k-based Wi-Fi chipset [28] upon reception of a packet. The anchors transmit packets on a 20 MHz wide channel at 5320 MHz (Wi-Fi channel 64). Upon receipt of each packet, the eavesdropper obtains a complex-valued CSI vector, containing the channel estimations for each of the 9 spatial channels and 56 non-zero OFDM subcarriers. As in [32], out of the 56 subcarriers, we select $K = 28$ ones with the highest pair-wise correlation coefficient. Furthermore, while the smartphone-based implementation from [32] reports 8-11 packets/sec., our implementation is capable of approx. 70 packets/sec.

As the target environment, we consider an ordinary office within our institute’s building of size approx. 7.5x 5.5 m. A floor plan of the room (target area) is shown in Fig. 3.

C. Results

We evaluate the outlined attack implementation to passively sense motion within the target area shown in Fig. 3. Here, the anchor devices transmit Wi-Fi packets that are received by the eavesdropper from a publicly accessible area on the outside of the target area. The anchor devices transmit at a rate of approx. 70 packets/sec, thus the eavesdropper obtains high-rate and continuous channel observations.

In a real-world scenario, Eve would initially perform a long-term reference measurement to determine the motion-detection threshold u using Eq. (6), assuming that the environment conditions remain steady over a long period of time [32]. In our experiments, we grant the eavesdropper a 3 min reference measurement under optimal conditions, i. e., the target area is guaranteed to be without motion and completely steady.

Subsequently, we proceed with a measurement where a person walks along a pre-defined path through the target area. The exact course is indicated by the numbered arrows in Fig. 3.

The corresponding measurement results are shown in Fig. 4 where we plot the reference measurement and the derived threshold as well as the eavesdropper’s observation during human motion.

Fig. 4 (a) and (b) show the respective adversary observations by eavesdropping anchor 1 from positions A and B – as depicted in Fig. 3. For eavesdropper at position A , we can see from the motion-trigger region (shaded red) that motion is successfully detected when the person is in proximity to the anchor device (note the indication of the person’s position at the top of each figure). For eavesdropper at position B , motion is detected with greater spatial extent, i.e., not only close to the anchor but also within the LOS of the anchor and the adversary.

Firstly, this result confirms the attack scenario and the effectiveness of our implementation. Another key observation is the location-dependent attack performance: The motion detection works best when the person walking in the target environment is within or close to the LOS between the anchor and the eavesdropper. This observation is in accordance with the model from Section II-A: The strong LOS signal component and very direct NLOS signal components typically contribute most to the channel response. Thus, upon disturbance of these paths by a human, the eavesdropper observes stronger variations in the channel response. This also lines up with the results in Fig. 4 (c) and (d), showing the observations from eavesdropping anchor 2. The LOS between the anchor and the eavesdropper now covers a larger area in the target environment and is even crossed by the walking person several times. As depicted by the identical scaling of the y -axis for each figure, it can be clearly observed that the signals received from anchor 2 are more sensitive for human motion. Thus, this time the adversary successfully detects motion for the entire walk along the path.

VI. IRSHIELD: COUNTERMEASURE OUTLINE

In this section, we describe our IRS-based countermeasure against the previously outlined wireless sensing attack. We also briefly describe the IRS implementation used for our experiments and introduce a practical IRS-based channel obfuscation algorithm.

A. IRS-based Channel Obfuscation

As outlined in Section II-B, an IRS makes a portion of the wireless channel between a transmitter and a receiver adjustable. This technological novelty is our key building block to find an appropriate countermeasure against adversarial wireless physical-layer sensing: We introduce a randomly, time-varying IRS into the target area and thereby shape the eavesdropper’s channel observation and introduce randomness to hamper detection of human motion.

Crucially, the IRS is a powerful yet simple tool that the defender can use to invalidate the assumption underlying the attack: While variations of the channel previously allowed to conclude on environmental variations and human motion,

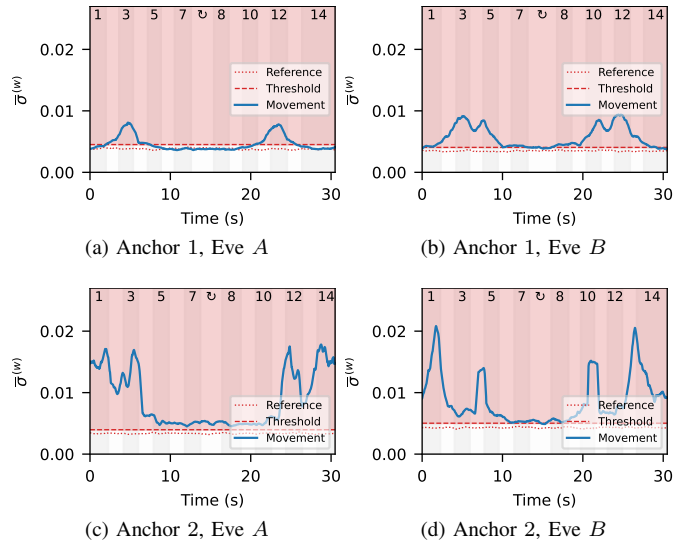


Fig. 4. Adversarial motion sensing results for eavesdropper positions A and B , receiving signals from anchor 1 and 2 with a person walking along a pre-defined path (see the numbering in Fig. 3). We plot the channel sliding-window standard deviation observed by the adversary with and without movement in the target area.

the IRS now introduces artificial variations, effectively obfuscating the mapping of channel variations to motion in the environment. In order to engineer a countermeasure from this, the defender can address either the eavesdropper’s threshold finding step or the actual motion sensing. From these two approaches, we can already formulate possible goals on desired effects of the countermeasure: (i) Make the adversary pick an overly high threshold such that environmental variation does not trigger detection. (ii) Let the adversary observe a strongly varying wireless channel such that the effect of human motion cannot be distinguished well. Additionally, since the adversary is passive, the defender is assumed to be unaware of the attack. Thus, we conclude that (iii) the defense needs to operate continuously.

A question yet to be answered is how the IRS should exactly vary the wireless channel to effectively hamper the adversarial detection. Recall from Section V that the attack is based on a sliding-window standard deviation of channel responses where temporal fluctuations in the strength of channel variation triggers detection. Therefore, to mimic the effect of human motion, we conclude that the strength of IRS-induced channel variation should exhibit randomized temporal changes.

B. Experimental Setup

For a first exploration of IRS-based channel obfuscation, we employ a prototypical IRS (see Fig. 5) with 256 binary-phase tunable identical unit cell elements arranged in a 16×16 array on a 43×35 cm standard FR4 PCB substrate. The unit cell consists of a rectangular reflector patch whose RF reflection coefficient can be electronically switched by means of a PIN diode to add a parasitic element to the reflector. Thus, by switching the bias voltage to the diode either on or off, the

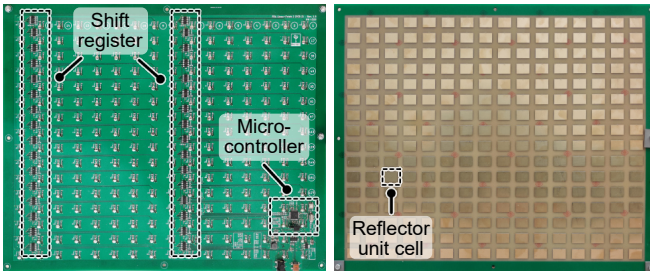


Fig. 5. Intelligent reflecting surface prototype (43×35 cm). Back side with control circuitry (left) and front side with reflecting elements (right).

reflection coefficient of each IRS element can be individually switched between values represented by a ‘0’ state (off) and a ‘1’ state (on). The corresponding voltages are generated by means of cascaded shift registers, digitally controlled using a microcontroller. Leveraging a standard USB connection for serial communication, the surface’s RF reflection behavior is elegantly programmed from a host computer via 256 bit words. The IRS used in our experiments is optimized to achieve a 180° phase shift at around 5.35 GHz in the reflected wave when switching between the ‘0’ and ‘1’ states, i.e., $r_m \in \{-1, 1\}$ in Eq. (2).

C. IRS Configuration Strategy

We now proceed with the design of a plug-and-play IRS configuration strategy to directly achieve such behavior in the field. Within the eavesdropper’s channel response, cf. Eq. (2), the contribution of the IRS is given as

$$\mathbf{H}_{\text{IRS}}(k, t) = \sum_{m=1}^M r_m(t) \mathbf{G}'_m(k, t). \quad (7)$$

Here we can see that the M individual IRS reflection coefficients $r_m(t)$ adjust the multipath components $\mathbf{G}'_m(k, t)$. The key observation to be made is that the signal amplitude from superposition of the multipath components in Eq. (7) depends on the IRS configuration. Thus, the amplitude of the eavesdropper’s observation, cf. Eq. (2), will likewise depend on the IRS. It should be noted that the mapping of specific IRS configurations to a particular channel response is non-trivial. As we here consider a binary-phase tunable IRS with 256 elements, there are 2^{256} possible realizations of $\mathbf{H}_{\text{IRS}}(k, t)$. In view of the large design space and the unknown channels to the eavesdropper $\mathbf{G}'_m(k, t)$ (recall that Eve is completely passive), it is not possible to distill or predict tailored IRS configurations.

Therefore, we pursue the following probabilistic approach where we gradually change the IRS configuration. Given an arbitrary initial set $\{r_m\}$, in the first step, we start by inverting a small amount of randomly chosen elements, e.g., 5% of all 256 elements. In the next step, we invert all 256 elements and go back to the beginning to repeat the procedure. Thereby, the IRS configuration will change gradually but random and similarly will the amplitude of the resulting IRS signal, thus

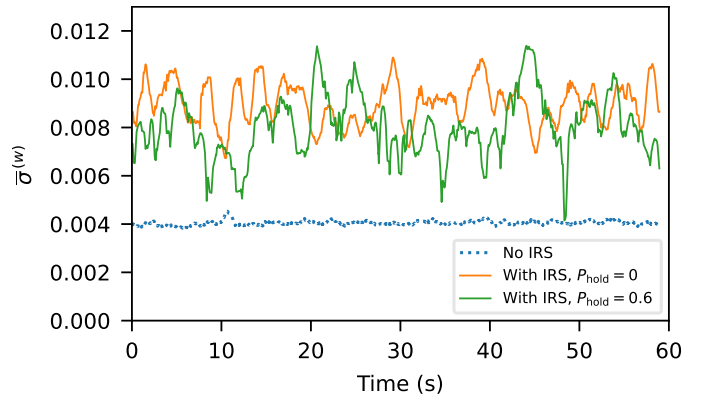


Fig. 6. Illustration of the effect of IRShield on the adversarial channel observation. We plot the channel sliding-window standard deviation with and without IRShield.

yielding smooth amplitude gradients. The intuition behind inverting the IRS configuration is the resulting 180° phase shift of $\mathbf{H}_{\text{IRS}}(k, t)$, i.e., when $r_m \in \{-1, 1\}$, we can factor out -1 in Eq. (7) such that $\mathbf{H}_{\text{IRS}}^{(0)}(k, t) = -\mathbf{H}_{\text{IRS}}^{(1)}(k, t)$. Thus, in combination with the non-IRS propagation in Eq. (2), a surface inversion will yield strong amplitude changes in the wireless channel.

We implement the outlined configuration strategy as further described by Algorithm 1. The algorithm leverages a state machine to switch back and forth between changing a small number of elements and inverting all elements. Additionally, the algorithm introduces randomized waiting times by skipping the state transition with a probability P_{hold} , with which the IRS configuration remains unchanged. In this way, randomized gaps are inserted into the IRS operation, causing short drops in channel variation to enhance the median absolute deviation in Eq. (6).

We test the effectiveness of the algorithm and place an IRS close, i.e., 30 cm, to the anchor position 2, and record channel observations of the eavesdropper at position A without any motion in the target environment. We first conduct a reference measurement with the IRS being disabled. Then, we use Algorithm 1 to configure the IRS at an update rate of 20 configurations/sec, and record channel measurements with $P_{\text{hold}} = 0$ and $P_{\text{hold}} = 0.6$. We plot the results as time series in Fig. 6. With the IRS being inactive, the channel remains static as expected. However, as can be seen from the plot, the IRS operation succeeds to produce strong time variations. Furthermore, the effect of P_{hold} is evident from pronounced excursions with low channel variation. As our results show, the IRS configuration algorithm meets the requirement to generate channel variations of varying strength.

VII. RESULTS

In this section, we experimentally investigate how IRShield affects adversarial motion detection. Further, we examine how size, distance, and orientation of the IRS play a role in the deployment of the countermeasure.

Algorithm 1: IRS configurations for IRShield.

```
random IRS configuration  $\text{cfg}[M] \in \{0, 1\}$ ;  
progression rate  $R = 0.05$ ;  
hold probability  $P_{\text{hold}} = 0.6$ ;  
nextState = 'RAND';  
while True do  
  if  $\text{getRandom}(0,1) < P_{\text{hold}}$  then  
     $\triangleright$  Remain in current state  
  else  
    currState = nextState;  
    if currState == 'RAND' then  
       $T \leftarrow \lceil R \cdot M \rceil$  random unique IRS  
      elements;  
      foreach  $m \in T$  do  
         $\text{cfg}[m] = \text{cfg}[m] \oplus 1$ ;  $\triangleright$  Flip  
        selected IRS elements  
      end  
      nextState = 'FLIP';  
    else if currState == 'FLIP' then  
       $\text{cfg} = \text{cfg} \oplus 1$ ;  $\triangleright$  Flip all elements  
      nextState = 'RAND';  
    end  
  end  
end  
write cfg to IRS;
```

A. Human Motion Sensing

In Section VI-A, we have already explained that the IRS provides a simplistic tool to vary wireless radio channels. Building upon this novel ability, we subsequently designed IRShield as a channel obfuscation scheme to counteract adversarial wireless sensing. Putting this scheme to test, we next investigate its impact on the detection of motion. Therefore, we repeat the human-motion sensing experiments from Section V, this time with IRShield in place. Here, we again use the experimental setup illustrated in Fig. 3, indicating a walking path and positions of the anchor device, the IRS, and the eavesdropper. For the IRS, we use the setup described in Section VI-B. Initially, the eavesdropper gathers a 3 min long reference measurement without any motion to determine a motion-detection threshold, however, observing not only the steady environment but also the time-varying IRS. Next, with the IRS is still active, the eavesdropper records channel measurements while a person walks along the pre-defined path.

In Fig. 7, we plot the results from four combinations of anchor-eavesdropper positions as time series, showing the reference measurement without motion, the resulting detection threshold according to [32], and the measurement with motion. As the baseline, we additionally plot the results without the IRS countermeasure, as presented in Fig. 4. When comparing results without IRShield, the first thing to note is that the adversary fails to detect motion in almost every case. This is because the adversary's reference measurement is degraded by the defender's IRS-induced channel variation, resulting in

an excessively high threshold value. For anchor position 1, Fig. 7 (a) and (b) show the results when Eve is located at positions *A* and *B*, respectively. In this case, the reference observation without motion strongly overlaps with the curve corresponding to motion. Thus, we conclude that our countermeasure renders adversarial wireless motion detection infeasible – even when adjusting the threshold. However, the situation is slightly different when considering less favorable anchor placement, e.g., anchor position 2 with a large LOS, for which we show the results in Fig. 7 (c) and (d). In this case, the IRS operation affects the threshold finding and thereby diminishes the adversary's success. However, it appears that LOS crossings may in principle still be detectable if the adversary is aware of the channel obfuscation and accordingly adjusts the threshold. Thus, we conclude that IRS-based channel obfuscation, just like the attack itself, is location-dependent. Therefore, we next systematically study spatial dependencies.

B. Systematic Coverage Tests

As we have outlined previously, due to the varying strength of different RF propagation paths, the adversarial motion sensing is subject to location-dependent performance and this likewise appears to be the case for our IRS-based countermeasure. Therefore, we now systematically study the attack and the countermeasure w.r.t. spatial performance. In particular, we seek to answer *where* in the target area motion can be detected.

For our experiments, we test at a total of 20 positions within the target area whether motion is detected. For each of the positions, distributed on a uniform grid throughout the target area (see Fig. 8), we place a rotating reflector as a point source of repeatable and RF-relevant motion. The reflector is a 50×50 cm aluminum sheet, mounted on a motor to rotate at approx. 40 rpm.

After an initial 3 min long reference measurement without any motion in the target area, we record the adversary's wireless channel observations from a particular anchor device with the rotating reflector at each of the 20 positions. We repeat this procedure with and without IRShield being active. To quantify the adversary's success, we use the motion-detection rate using a threshold found from the initial reference measurement. The detection rate is given as the fraction of total motion observations that lie above the threshold.

For anchor position 1 and two eavesdropper positions with a threshold according to [32], we show the spatial distribution of detection rates in Fig. 9 without ((a), (c)) and with the IRS being active ((b), (d)). Without the countermeasure, as expected, we can observe the highest detection rates within and around the LOS. Far off the LOS, the detection rate drops down to 0%. This effect is best visible in Fig. 9 (a): Due to the adversary's strong LOS signal, the sensitivity for weaker (environment-dependent) multipath signals is reduced. In contrast, when the IRS is active, the adversarial detection is completely suspended as evident from the 0% detection rate at all positions.

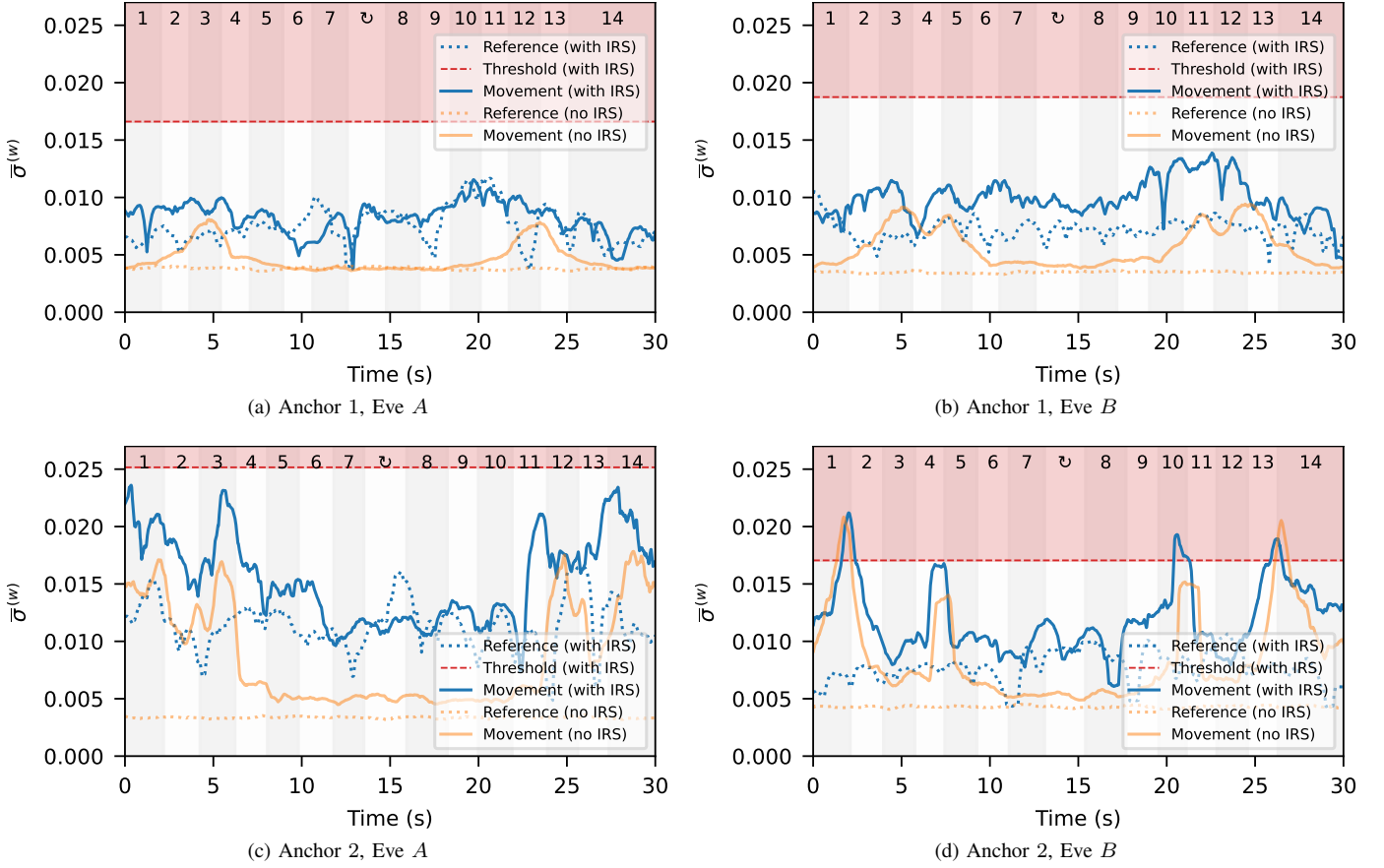


Fig. 7. Adversarial motion sensing results with IRShield active. We plot the channel sliding-window standard deviation observed by the eavesdropper during the reference measurement without motion and with a person walking along a pre-defined path (see the numbering in Fig. 3). Also, we indicate the motion-detection threshold according to [32] with $C = 11$. For comparison, we also plot the results when IRShield is not active, cf. Fig. 4.

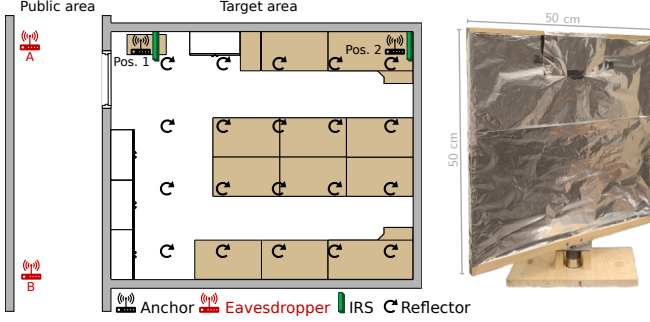


Fig. 8. Measurement setup of the coverage test, indicating the 20 rotating reflector positions (left) and the rotating reflector on a wooden frame (right).

For the same eavesdropper positions and threshold finding [32] for anchor position 2 (large LOS), we show the results in Fig. 10. While (a) again highlights the attack’s LOS-dependent coverage, the diagonal LOS through the target area in (c) allows the eavesdropper to obtain 100% detection rate on each position. Even for this challenging scenario, our countermeasure succeeds to reduce the detection rate to 0% on all positions ((b)) and 18 out of 20 positions ((d)).

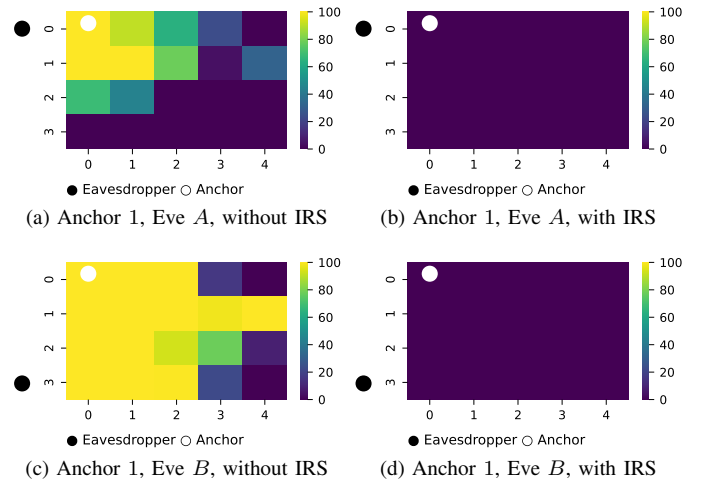


Fig. 9. Spatial distribution of detection rates for anchor position 1 with motion-detection threshold according to [32] with $C = 11$.

Threshold adjustments. We evaluated the specific threshold finding of the original attack as proposed by [32]. However, choosing the threshold less conservative will slightly improve

TABLE II
AVERAGE DETECTION RATES FOR VARIOUS THRESHOLDS, AND PLACEMENTS OF ANCHOR AND EAVESDROPPER.

	Anchor 1, eavesdropper A				Anchor 1, eavesdropper B				Anchor 2, eavesdropper A				Anchor 2, eavesdropper B			
	No IRS		With IRS		No IRS		With IRS		No IRS		With IRS		No IRS		With IRS	
	TPR	FPR	TPR	FPR	TPR	FPR	TPR	FPR	TPR	FPR	TPR	FPR	TPR	FPR	TPR	FPR
$\max_t \{\cdot\}$	0.64	0.00	0.03	0.00	0.92	0.00	0.08	0.00	0.94	0.00	0.28	0.00	1.00	0.00	0.39	0.00
$C = 1$	0.91	0.25	0.42	0.26	1.00	0.26	0.62	0.28	0.99	0.27	0.62	0.27	1.00	0.25	0.86	0.23
$C = 3$	0.75	0.02	0.11	0.04	0.98	0.04	0.23	0.04	0.98	0.04	0.41	0.04	1.00	0.03	0.55	0.02
$C = 5$	0.61	0.00	0.03	0.00	0.94	0.00	0.04	0.00	0.96	0.00	0.26	0.00	1.00	0.00	0.30	0.00
$C = 7$	0.51	0.00	0.00	0.00	0.88	0.00	0.00	0.00	0.91	0.00	0.16	0.00	1.00	0.00	0.15	0.00
$C = 9$	0.42	0.00	0.00	0.00	0.82	0.00	0.00	0.00	0.83	0.00	0.06	0.00	1.00	0.00	0.08	0.00
$C = 11$ [32]	0.35	0.00	0.00	0.00	0.76	0.00	0.00	0.00	0.78	0.00	0.00	0.00	1.00	0.00	0.05	0.00

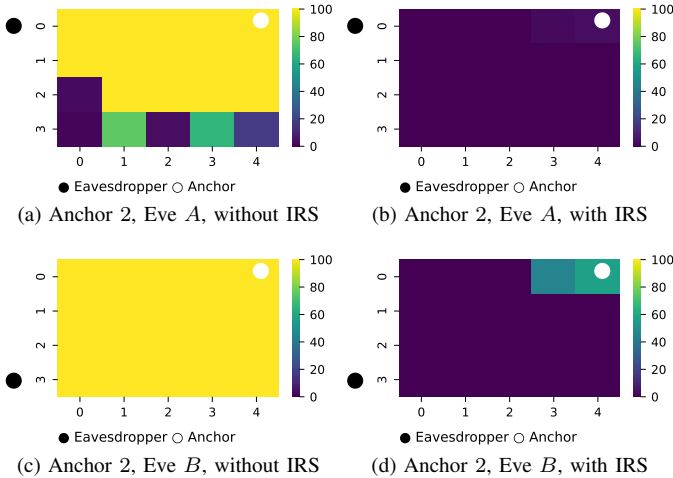


Fig. 10. Spatial distribution of detection rates for anchor position 2 with motion-detection threshold according to [32] with $C = 11$.

the detection rates. For each of the four combinations of anchor and eavesdropper positions, Table II lists the average detection rates for varying motion-detection thresholds. For the state-of-the-art attack [32], IRShield lowered detection rates to 5% or less, as can be seen from the last row, with $C = 11$ in Eq. (6), corresponding to Fig. 9 and Fig. 10. The first row, indicated by $\max_t \{\cdot\}$, gives the results with the maximum of the reference measurement being used as motion-detection threshold. Thus, the detection rates were obtained with the lowest threshold resulting in a false positive rate (FPR) of 0 (at least for the observed reference duration). The respective heatmaps are shown in Appendix A (Fig. 16 and Fig. 17).

For anchor position 2, an adapted threshold would allow the adversary to restore some detection capabilities. Regarding the spatial aspects of this result, it can be seen from Fig. 16 and Fig. 17 that motion with the LOS and directly adjacent to the anchor is still reliably detected. This result is a vivid example for a limitation of IRS-based channel obfuscation: Due to the finite size of the IRS, it reflects a limited amount of signal power. Thus, it is challenging for the IRS to produce channel variations as strong as LOS crossings. On the other hand, for anchor position 1, the eavesdropper does not observe

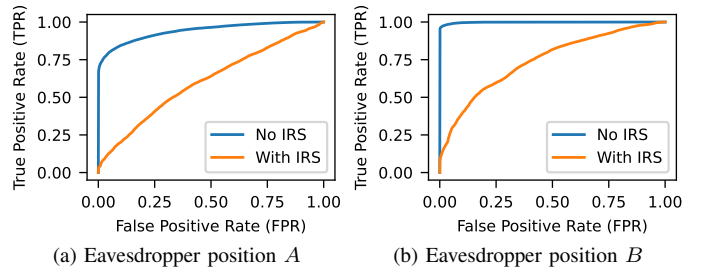


Fig. 11. ROC curves with and without IRShield for anchor position 1 and (a) eavesdropper position A, (b) eavesdropper position B.

a signal with significant LOS coverage. This is reflected in Table II, showing that the true positive rate (TPR) cannot be substantially increased without also increasing the FPR. To illustrate the separability of the eavesdropper's observations without motion and with motion, we plot the receiver operating curves (ROCs) in Fig. 11 for anchor position 2. The plots show the trade-off between the TPR and FPR, essentially visualizing all possible choices of detection thresholds. Especially for eavesdropper position A, the detection approaches a random classifier. This result highlights that IRS-based channel obfuscation under certain conditions provides near-ideal channel randomization, thus thwarting adversarial motion detection.

C. Impact of IRS

Thus far, we considered fixed conditions for the deployment of IRShield. Next, we seek to further characterize the effect of varying IRS parameters to provide a thorough guideline for practical deployments. Therefore, we change conditions such as size, distance, and orientation of the IRS w.r.t. an anchor device. For experimental simplicity, we test these conditions without considering human motion and leverage the strength of IRS-induced channel variation as a proxy for the effectiveness of channel obfuscation.

Effective IRS size. The IRS is passive w.r.t. RF signals, i. e., it can only reflect but not apply amplification. Thus, the physical size of the IRS is an important factor limiting the signal power that can be reflected [26]. We conclude that the IRS size is one parameter that affects the effectiveness of channel

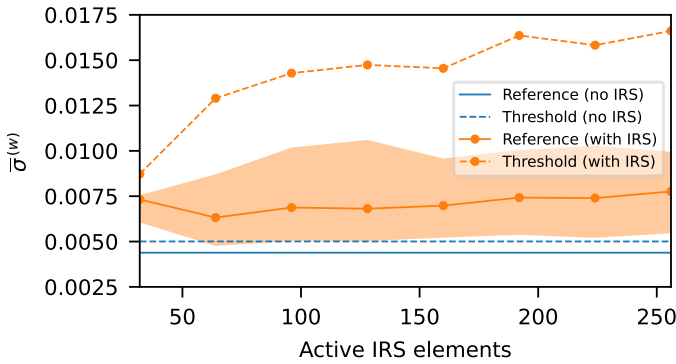


Fig. 12. Effect of varying number of active IRS elements. We plot the median and the 1st and 99th percentiles of the adversarial observations and the resulting motion-detection threshold according to [32] with $C = 11$.

obfuscation. Studying this connection experimentally, we vary the effective IRS size (since the physical size of the IRS is fixed). That is, we simply deactivate a certain number of IRS elements. Since IRShield is based on time-varying surface configurations, the deactivated elements will not contribute to the channel variation.

In our experiment, we increase the number of active IRS elements M from 32 to 256 in steps of 32 while we use Algorithm 1 to generate IRS configurations. The IRS elements to be added to the set of active elements are selected randomly. For each effective surface size, we then record 2 min of adversarial observations under channel obfuscation. We plot the median, the 1st and 99th percentiles of the channel variation together with the corresponding motion-detection threshold according to [32] over the number of active IRS elements in Fig. 12. We can see that the median and the threshold grow with the effective IRS size. This confirms our hypothesis that an increasing number of varying IRS elements enhances channel obfuscation. However, since the adversary’s motion-detection threshold is already substantially increased for 64 active elements, we conclude that IRShield does not necessarily require large IRS deployments.

IRS distance. Next, we investigate the connection between IRS-based channel obfuscation and the IRS distance to the anchor device. Signal propagation including an IRS is typically modeled by a multiplication of channels from the transmitter to the IRS and from the IRS to the receiver [26]. Thus, due to the inherent distance-dependent path loss of wireless radio channels, the received power from the IRS will be inversely proportional to the inverse squared product of the distances to and from the IRS [17]. That is, the IRS works best near the transmitter or receiver (since one of the IRS distances is minimized). However, specific to our channel obfuscation application for, e.g., living spaces, it is reasonable to assume that the IRS is located rather close to both anchor devices and the receiver (the eavesdropper).

To assess this experimentally, we vary the distance of the IRS to anchor 2. For each distance, the eavesdropper observes the wireless channel from position A for 3 min. We

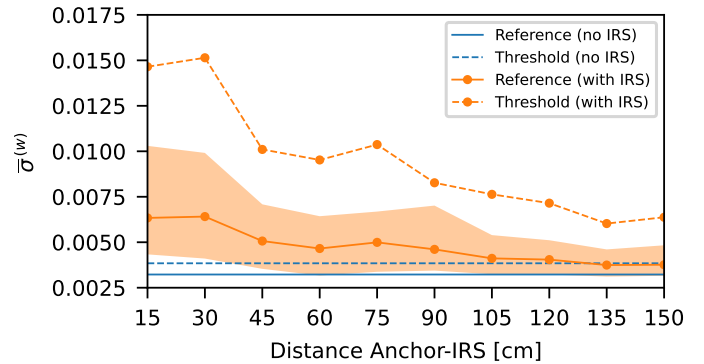


Fig. 13. Effect of varying distances of the IRS to the anchor device. We plot the median and the 1st and 99th percentiles of the adversarial observations and the resulting motion-detection threshold according to [32] with $C = 11$.

plot the median, the 1st and 99th percentiles of the channel variation and the corresponding motion-detection threshold according to [32] over IRS distances between 15 and 150 cm in Fig. 13. As expected, the channel obfuscation works best when the IRS is close to the anchor. Please note that the IRS generally can operate from higher distances as well, however, requiring optimal IRS configurations which the probabilistic Algorithm 1 is not designed for.

IRS orientation. Despite the distance, another relevant aspect of IRS placement is its orientation w.r.t. the eavesdropper. In the previous experiments, we placed the IRS facing towards the eavesdropper, i.e., the front of the IRS pointed towards the left side of the target area shown in Fig. 3. As this placement presumably is optimal for the defender, we additionally test less optimal orientations by rotating the IRS around an anchor device. For each IRS placement (at a constant distance of 30 cm), the adversary records channel measurements for 1 min. Fig. 14 shows the median, the 1st and 99th percentiles of the channel variation in a polar plot, where the angle indicates the IRS placement. For 0°, the IRS faces towards the eavesdropper. As expected, we can see that the channel obfuscation works best when the IRS is roughly directed towards the eavesdropper. Thus, when the environment to be protected has limited public accessibility (e.g., a hallway on one side of the property such as in Fig. 3), the defender should point the IRS towards this direction. However, despite that, even in the worst case where the IRS points in the opposite direction of the eavesdropper, the channel obfuscation is still effective, albeit with reduced impact.

Data throughput. Another important question is whether IRS-based channel obfuscation affects the wireless data throughput. Previous countermeasures, e.g., [32] and [25], are based on varying the allocation of the transmission medium and therefore face a trade-off between the quality of service of wireless communication and reduction of adversarial sensing capabilities. To this end, we emphasize that our channel obfuscation approach is based on changing the transmission medium itself instead of its allocation. Therefore, in short-range indoor wireless networks, we expect the channel obfuscation to have

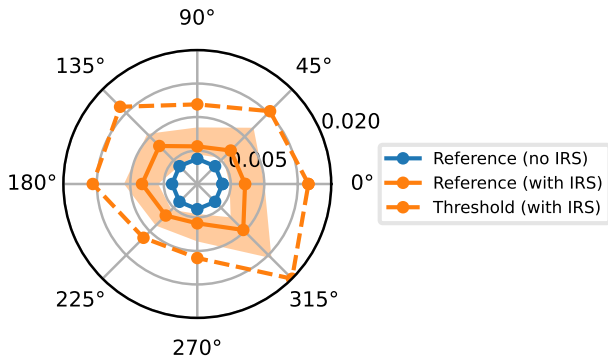


Fig. 14. Effect of IRShield on the adversarial reference measurement for varying IRS orientations. The IRS faces towards the eavesdropper at 0° . We plot the median and the 1st and 99th percentiles of the adversarial observations and the resulting motion-detection threshold according to [32] with $C = 11$.

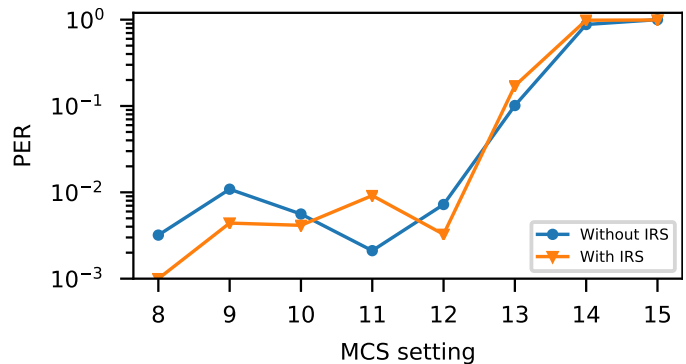


Fig. 15. Packet error rates over MCS setting without and with IRShield when a legitimate party receives packets from an anchor device. A higher MCS value indicates a higher data rate transmission.

a marginal effect on the data throughput.

We put this claim to test by measuring packet error rates (PERs) of a Wi-Fi connection with and without the channel obfuscation being active. For this, we deploy two legitimate IEEE 802.11n Wi-Fi devices in the eavesdropper’s target area and make one device transmit 200,000 packets while the other keeps track of the received packets. For the transmissions, we randomize the payload data and use a fixed modulation and coding scheme (MCS) [23]. As we use 2×2 MIMO devices, we use MCS values 8-15 (two spatial streams). Next to the transmitter, we place the IRS.

In Fig. 15, we plot the measured PER for each MCS setting with and without IRShield being active. The first thing to note is that the PER increases with the MCS setting for both cases. As higher MCS settings correspond to higher data rate transmissions with less immunity against noise, this indicates channel conditions with a limited signal-to-noise ratio. Apart from that, the key observation is that the channel obfuscation appears to cause only a negligible increase or decrease in PER. This is because our randomized IRS configuration approach will certainly both improve and degrade the wireless channel over time. Thus, as indicated by our experimental result, we believe that the average channel quality is not affected by channel obfuscation.

VIII. DISCUSSION

In the following, we discuss the experimental setup, our results and provide directions for future research.

A. Experimental Setup and Results

As our main goal was the evaluation of IRS-based channel obfuscation, we made some simplifications to the attack realization in order to facilitate experimentation. Notably, we thereby granted the adversary rather optimal conditions, which naturally is desirable for evaluation of a countermeasure.

First, we deployed a set of MIMO-capable Wi-Fi routers to act as the anchor and eavesdropper devices. Without loss of generality, the routers utilize standardized waveforms and can be considered as a substitute for any other Wi-Fi device.

The anchor routers transmit on 3 antennas, allowing the eavesdropper to estimate a total of 9 spatial wireless channels in parallel from a single anchor (the eavesdropper receives with 3 antennas). Thus, the adversary takes advantage of spatial diversity [7], i. e., multiple different wireless channels, each containing information on the target area to improve the adversarial motion detection [5].

Another important aspect is the transmission rate of the anchor devices. As outlined in [32], typical Wi-Fi devices for use in living spaces transmit at 3 packets/sec to 257 packets/sec during active operation. Likewise, in our setup, the anchors constantly transmit at approx. 70 packets/sec. Thus, we constantly allow the adversary to observe the target area, even with substantial oversampling since human motion is rather slow when sampled at 70 packets/sec. Further, the wireless traffic transmitted by our anchor devices does not carry meaningful application-driven data. However, this does not play a role for the attack which merely utilizes physical-layer observations, i. e., being independent of payload data and encryption.

For the adversarial motion detection, we considered channel variation as a proxy for motion and a threshold-based detection scheme from [32]. A key factor for the adversary’s success is a properly chosen threshold calculated from long-term channel observations, assuming that the target area typically does not exhibit motion [32]. For our experimental evaluation, other than in a realistic attack scenario, we granted the eavesdropper dedicated, well-behaved reference measurements without any movement in the target area.

Even under these clearly beneficial conditions for the adversary, our results show that IRS-based channel obfuscation almost completely defeats the motion-detection scheme from the literature [32]. Due to randomized artificial channel variation produced by the IRS, the eavesdropper picks an excessively high threshold. Thus, assuming the defense is known, we have also shown that the adversary in turn can adapt the threshold to slightly improve performance again. In this way, motion with strong impact on the wireless channel, e. g., within the LOS between the eavesdropper and an anchor, is still detected. This reveals a limitation of our channel obfuscation: It is difficult

for the probabilistically configured IRS to generate channel variations that are as strong as human LOS crossings. Instead, the defender could reduce such strong channel variations, e. g., by means of strategic anchor placement, preventing motion to take place within the eavesdropper’s LOS. Therefore, the defender could take advantage of limited public accessibility of the target area. As we have shown, the adversary’s observations of anchor devices close to the eavesdropper are less sensitive against motion.

B. IRS Deployment

The outlined IRS configuration strategy is a probabilistic approach, realizing channel obfuscation in a plug-and-play manner independently of the anchor devices. Thus, since the IRS directly interacts with wireless signals as they propagate, no integration efforts are required once the surface is deployed in proximity to an anchor device. Note that, however, the randomized IRS configurations are only sub-optimal, so that the IRS impact is likely to not be maximized, especially for higher distances of the IRS to the anchor devices. To this end, we emphasize that IRSs are being considered for future 6G wireless communication systems with potential for integration into walls and furniture [27]. Thus, the assumption of IRS deployments in proximity to anchor devices indeed is well justified.

The effectiveness of channel obfuscation is rooted in the used IRS hardware. The IRS is commonly considered as low-complexity device that can be built at low-cost, using standard PCB technology. The IRS deployed in our experiments is a prototypical device which leaves room for hardware improvements. For instance, currently only 5 GHz frequency bands are covered and signals are reflected with a certain loss. Thus, our results provide first insights to wireless channel obfuscation and can be further improved as the technology evolves.

C. Further Applications of IRShield

As Wi-Fi devices are particularly widespread in private and public spaces, we used IEEE 802.11n Wi-Fi to confirm the potential of IRS for channel obfuscation. However, in some cases, an adversary may also leverage signals from anchor devices using other wireless standards. Since the IRS operates directly on the physical layer and affects the wireless channel, it is independent of particular waveforms or standards. Thus, IRShield generally is not tied to Wi-Fi. The experimental verification of IRShield for other wireless standards is left for future work.

We examined channel obfuscation as a countermeasure particularly geared towards adversarial (human) motion detection. However, the applications of wireless sensing go beyond motion detection. For instance, gesture recognition, identification of individuals, imaging, vital sign monitoring, or keystroke recognition [15] have been reported and certainly can have far-reaching privacy implications. We believe that the intricate signal processing methods employed for these applications could be susceptible against signal variations induced by the IRS. Therefore, IRShield may also protect against other classes

of wireless sensing, which can be confirmed through further experimentation.

Finally, we would like to point out that the IRS could also be used by adversaries to corrupt legitimate wireless sensing applications, e. g., for motion-based intrusion detection [11]. However, details of such considerations are beyond the scope of this work.

D. Future Work

Our findings may serve as a starting point for future work that could investigate, for instance, the following aspects.

Optimization of IRS configuration. We have proposed a probabilistic IRS configuration algorithm to achieve channel obfuscation. While we already obtained satisfactory results, we believe that there is still room for improvements, e. g., by randomizing state transitions of the algorithm or by switching between predefined patterns. Currently, the IRS configuration procedure operates independently of the anchor device. The next logical step would be to incorporate information from the anchor devices and accordingly, adjust the IRS, e. g., by analyzing channel estimates.

Currently, research on IRS largely focuses on the improvement of wireless communication, e. g., by enhancing received signal power. Thus, future work could study whether IRShield could provide channel obfuscation while additionally improving communication of legitimate parties.

Beamforming. More recent Wi-Fi devices, e. g., for IEEE 802.11ac, additionally utilize beamforming. Thus, wireless signals are no longer transmitted to all directions (including the eavesdropper) but rather directed towards the intended receiver. Hence, an eavesdropper will receive weaker signals [4] which likely reduces the adversarial sensing capabilities. More work is required to investigate whether beamforming in conjunction with IRShield further degrades the adversary’s capabilities.

Improved adversary strategies. We have shown that IRShield successfully defeats existing adversarial motion-detection strategies. In some situations, channel obfuscation made the environment with motion virtually indistinguishable from that without motion. While, in our work, we considered linear separability, a more advanced adversary may apply elaborate signal processing and statistical classification methods. Future work could investigate whether such methods improve adversarial motion detection.

IX. CONCLUSION

In this paper, we introduced and experimentally examined IRShield as a plug-and-play countermeasure against adversarial motion detection from passive eavesdropping of wireless signals. By deploying an IRS, we partly randomize the wireless propagation environment, thus introducing random variations which obfuscate the adversary’s signal observations. To this end, we presented an extensive experimental investigation to characterize channel obfuscation. Beginning with human motion detection, we have shown that our scheme successfully

diminishes the adversary's success. We then systematically studied spatial aspects of the attack and the countermeasure and discussed how the adversary could adapt to the defense. Notably, we found that our scheme lowered detection rates to 5% or less for a state-of-the-art attack. In certain cases, it renders motion detection largely infeasible, regardless of the adversary strategy. Furthermore, we investigated IRS parameters such as size, distance, and orientation and showed that IRShield does not affect the wireless communication performance.

Due to its novel capabilities, IRS already receives enormous research interest in the context of wireless communication. Having shown that it can likewise be applied to counteract privacy-violations, we hope that our work stimulates further research on IRS in a security context.

ACKNOWLEDGEMENTS

We thank Simon Tewes for his help with setting up the Wi-Fi routers. This work was in part funded by the German Federal Ministry of Education and Research (BMBF) within the project MetaSEC (Grant 16KIS1234K) and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2092 CaSa - 390781972.

REFERENCES

- [1] H. Abdelnasser, M. Youssef, and K. A. Harras, "WiGest: A ubiquitous Wi-Fi-based gesture recognition system," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, ser. INFOCOM '15, 2015, pp. 1472–1480.
- [2] A. Acar, H. Fereidooni, T. Abera, A. K. Sikder, M. Miettinen, H. Aksu, M. Conti, A.-R. Sadeghi, and S. Uluagac, "Peek-a-boo: I see your smart home activities, even encrypted!" in *Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, ser. WiSec '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 207–218. [Online]. Available: <https://doi.org/10.1145/3395351.3399421>
- [3] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, "Keystroke Recognition Using WiFi Signals," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 90–102. [Online]. Available: <https://doi.org/10.1145/2789168.2790109>
- [4] D. Antonioli, S. Siby, and N. O. Tippenhauer, "Practical Evaluation of Passive COTS Eavesdropping in 802.11b/n/ac WLAN," in *Cryptology and Network Security*, S. Capkun and S. S. M. Chow, Eds. Cham: Springer International Publishing, 2018, vol. 11261, pp. 415–435. [Online]. Available: http://link.springer.com/10.1007/978-3-030-02641-7_19
- [5] A. Banerjee, D. Maas, M. Bocca, N. Patwari, and S. Kasera, "Violating Privacy through Walls by Passive Monitoring of Radio Windows," in *Proceedings of the 2014 ACM Conference on Security and Privacy in Wireless and Mobile Networks*, ser. WiSec '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 69–80. [Online]. Available: <https://doi.org/10.1145/2627393.2627418>
- [6] G. Camurati, S. Poeplau, M. Muench, T. Hayes, and A. Francillon, "Screaming Channels: When Electromagnetic Side Channels Meet Radio Transceivers," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 163–177. [Online]. Available: <https://doi.org/10.1145/3243734.3243802>
- [7] A. Goldsmith, *Wireless Communications*. USA: Cambridge University Press, 2005.
- [8] D. Huang, R. Nandakumar, and S. Gollakota, "Feasibility and Limits of Wi-Fi Imaging," in *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems*, ser. SenSys '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 266–279. [Online]. Available: <https://doi.org/10.1145/2668332.2668344>
- [9] N. Kaina, M. Dupré, G. Lerosey, and M. Fink, "Shaping complex microwave fields in reverberating media with binary tunable metasurfaces," *Scientific Reports*, vol. 4, no. 1, p. 6693, May 2015.
- [10] Y. Li and T. Zhu, "Gait-Based Wi-Fi Signatures for Privacy-Preserving," in *Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security*, ser. ASIA CCS '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 571–582. [Online]. Available: <https://doi.org/10.1145/2897845.2897909>
- [11] Linksys Holdings, Inc., "Linksys Aware," Accessed: December 01, 2021. [Online]. Available: <https://www.linksys.com/us/linksys-aware/>
- [12] J. Liu, Y. Wang, Y. Chen, J. Yang, X. Chen, and J. Cheng, "Tracking Vital Signs During Sleep Leveraging Off-the-Shelf WiFi," in *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, ser. MobiHoc '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 267–276. [Online]. Available: <https://doi.org/10.1145/2746285.2746303>
- [13] X. Liu, Q. Zeng, X. Du, S. L. Valluru, C. Fu, X. Fu, and B. Luo, "SniffMislead: Non-Intrusive Privacy Protection against Wireless Packet Sniffers in Smart Homes," in *24th International Symposium on Research in Attacks, Intrusions and Defenses*, ser. RAID '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 33–47. [Online]. Available: <https://doi.org/10.1145/3471621.3471856>
- [14] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. D. Renzo, and N. Al-Dhahir, "Reconfigurable Intelligent Surfaces: Principles and Opportunities," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1546–1577, 2021.
- [15] Y. Ma, G. Zhou, and S. Wang, "WiFi Sensing with Channel State Information: A Survey," *ACM Comput. Surv.*, vol. 52, no. 3, Jun. 2019. [Online]. Available: <https://doi.org/10.1145/3310194>
- [16] MarketsandMarkets, "Wireless Connectivity Market by Connectivity Technology (Wi-Fi, Bluetooth, NFC, ZigBee, GNSS, LTE CAT-M1, NB-IoT, LoRa, SigFox), Type (WLAN, WPAN, LPWAN), End-use (Wearables, Consumer Electronics, Healthcare), and Region - Global Forecast to 2025," 2020, <https://www.marketsandmarkets.com/Market-Reports/wireless-connectivity-market-192605963.html>, Accessed: December 01, 2021.
- [17] Ö. Özdoğan, E. Björnson, and E. G. Larsson, "Intelligent Reflecting Surfaces: Physics, Propagation, and Pathloss Modeling," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 581–585, May 2020.
- [18] Y. Qiao, O. Zhang, W. Zhou, K. Srinivasan, and A. Arora, "PhyCloak: Obfuscating Sensing from Communication Signals," in *Proceedings of the 13th Usenix Conference on Networked Systems Design and Implementation*, ser. NSDI'16. USA: USENIX Association, 2016, p. 685–699.
- [19] M. D. Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin, J. d. Rosny, A. Bounceur, G. Lerosey, and M. Fink, "Smart radio environments empowered by reconfigurable AI meta-surfaces: an idea whose time has come," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, p. 129, May 2019. [Online]. Available: <https://doi.org/10.1186/s13638-019-1438-9>
- [20] Statista, "Internet of Things - number of connected devices worldwide 2015-2025," 2016, <https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/>, Accessed: December 01, 2021.
- [21] TP-Link Corporation Limited, "TL-WDR4300 — N750 Wireless Dual Band Gigabit Router," Accessed: November 24, 2021. [Online]. Available: <https://www.tp-link.com/us/home-networking/wifi-router/tl-wdr4300/>
- [22] R. Trimananda, J. Varmarken, A. Markopoulou, and B. Demsky, "Packet-Level Signatures for Smart Home Devices," in *27th Annual Network and Distributed System Security Symposium, NDSS 2020, San Diego, California, USA, February 23-26, 2020*, ser. NDSS '20. The Internet Society, 2020. [Online]. Available: <https://www.ndss-symposium.org/ndss-paper/packet-level-signatures-for-smart-home-devices/>
- [23] F. Vergès, "MCS Index, Modulation and Coding Index 11n and 11ac," Accessed: November 21, 2021. [Online]. Available: <http://mcsindex.com/>

- [24] J. Wang, J. Xiong, H. Jiang, K. Jamieson, X. Chen, D. Fang, and C. Wang, "Low Human-Effort, Device-Free Localization with Fine-Grained Subcarrier Information," *IEEE Transactions on Mobile Computing*, vol. 17, no. 11, pp. 2550–2563, 2018.
- [25] P. M. Wijewardena, A. Bhaskara, S. K. Kasera, S. A. Mahmud, and N. Patwari, "A Plug-n-Play Game Theoretic Framework for Defending against Radio Window Attacks," in *Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, ser. WiSec '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 284–294. [Online]. Available: <https://doi.org/10.1145/3395351.3399368>
- [26] Q. Wu and R. Zhang, "Towards Smart and Reconfigurable Environment: Intelligent Reflecting Surface Aided Wireless Network," *IEEE Communications Magazine*, vol. 58, no. 1, pp. 106–112, 2020.
- [27] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent Reflecting Surface Aided Wireless Communications: A Tutorial," *IEEE Transactions on Communications*, vol. 69, no. 5, pp. 3313–3351, 2021.
- [28] Y. Xie, Z. Li, and M. Li, "Precise Power Delay Profiling with Commodity WiFi," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '15. New York, NY, USA: ACM, 2015, pp. 53–64, Paris, France.
- [29] Y. Yao, Y. Li, X. Liu, Z. Chi, W. Wang, T. Xie, and T. Zhu, "Aegis: An Interference-Negligible RF Sensing Shield," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, ser. INFOCOM '18, 2018, pp. 1718–1726.
- [30] S. Yousefi, H. Narui, S. Dayal, S. Ermon, and S. Valaee, "A Survey on Behavior Recognition Using WiFi Channel State Information," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 98–104, 2017. [Online]. Available: <http://ieeexplore.ieee.org/document/8067693/>
- [31] W. Zhang, Y. Meng, Y. Liu, X. Zhang, Y. Zhang, and H. Zhu, "HoMonit: Monitoring Smart Home Apps from Encrypted Traffic," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1074–1088. [Online]. Available: <https://doi.org/10.1145/3243734.3243820>
- [32] Y. Zhu, Z. Xiao, Y. Chen, Z. Li, M. Liu, B. Y. Zhao, and H. Zheng, "Et Tu Alexa? When Commodity WiFi Devices Turn into Adversarial Motion Sensors," in *27th Annual Network and Distributed System Security Symposium, NDSS 2020, San Diego, California, USA, February 23-26, 2020*, ser. NDSS '20. Internet Society, 2020.

APPENDIX A HEATMAPS WITH REDUCED THRESHOLD

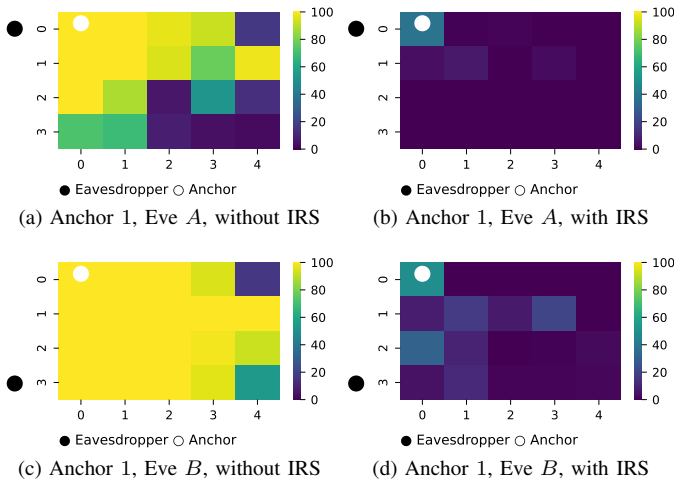


Fig. 16. Spatial distribution of detection rates for anchor position 1 with lowest motion-detection threshold yielding a false-positive rate of 0.

In Section VII-B, we systematically studied the spatial performance of adversarial motion sensing with and without IRShield. Complementing Fig. 9 and Fig. 10, Fig. 16 and

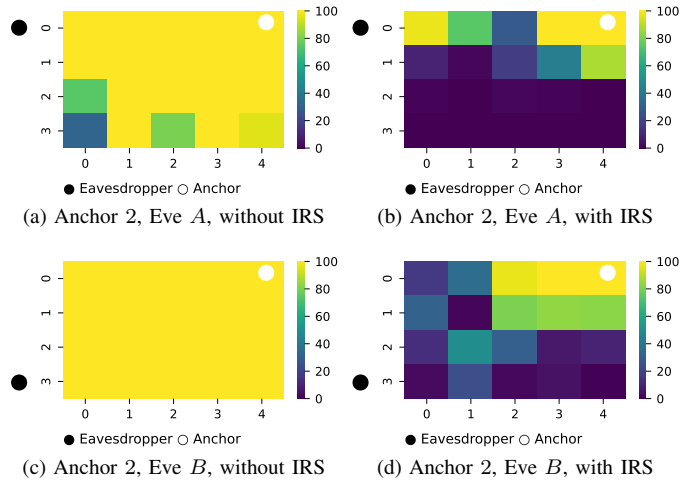


Fig. 17. Spatial distribution of detection rates for anchor position 2 with lowest motion-detection threshold yielding a false-positive rate of 0.

Fig. 17 show the spatial distribution of motion detection rates with the maximum of the reference measurement being used as motion-detection threshold. Thus, the detection rates were obtained with the lowest threshold resulting in a FPR of 0 (at least for the observed reference duration). This corresponds to the first row of Table II, indicated by $\max_t \{\cdot\}$.