accepted to appear in Proceedings of IEEE International Conference on Communications (ICC) 2021, Virtual, Montreal.

Adversarial Occupancy Monitoring using One-Sided Through-Wall WiFi Sensing

Steven M. Hernandez and Eyuphan Bulut
Department of Computer Science, Virginia Commonwealth University
401 West Main St. Richmond, VA 23284, USA
{hernandezsm, ebulut}@vcu.edu

***Abstract*—Through-wall sensing systems can aid in performing building security, and collecting analytics or more ominously be leveraged for surveillance. With the pervasive nature of WiFi routers and devices in our office buildings and homes, we essentially place an unencrypted (at the frame level) transmitting source directly in our buildings which can then be leveraged for surveillance by adversaries. In this work, we study such a device-free WiFi sensing system for occupancy monitoring and crowdcounting and evaluate it in a number of through-wall conditions. We demonstrate that with a proper analysis of Channel State Information (CSI) collected from the WiFi signals, we can recognize both the presence of targets as well as their moving direction in a hallway environment which can be leveraged to track and count the flow of traffic throughout a building. We specifically demonstrate through real world experiments how an adversary with very limited physical access to a building can still successfully collect surveillance data of a target area through the wall.**

***Index Terms*—WiFi sensing, occupancy monitoring, privacy, through-wall sensing.**

I. Introduction

Occupancy monitoring and crowdcounting offers the ability to collect analytics and insights into traffic within indoor spaces for use in intelligent energy efficient heating and air conditioning control systems [1], building security through intruder detection [2] and crowd safety [3]. Human target surveillance in public and private scenarios can also benefit from both occupancy monitoring and crowdcounting techniques. For private locations such as businesses or homes, typical surveillance devices such as cameras or microphones would require an adversary to have full access to the target areas. This of course is not always possible when considering private residences. Further, even when access is possible, the device payload will likely attract attention by the presence of features such as a camera lens.

In this work, we propose the use of relatively new WiFi sensing techniques which use traits of WiFi signals to understand actions occurring in a physical environment. Because WiFi is designed to penetrate walls, device payloads no longer need to be placed directly in the target area. Instead, the WiFi receiving device can be placed on the outer perimeter of a room or building to then sense through the walls. Furthermore, because of the ubiquity of WiFi devices and routers in environments such as residential homes and commercial buildings, a WiFi sniffer device can also leverage the natural ambient radio traffic from the existing devices in the environment to detect human presence or activities. In this study, we consider the case where an adversary places a WiFi transmitter and a WiFi receiver in a non-line-of-sight (NLOS) location that is behind the wall of the target hallway area.

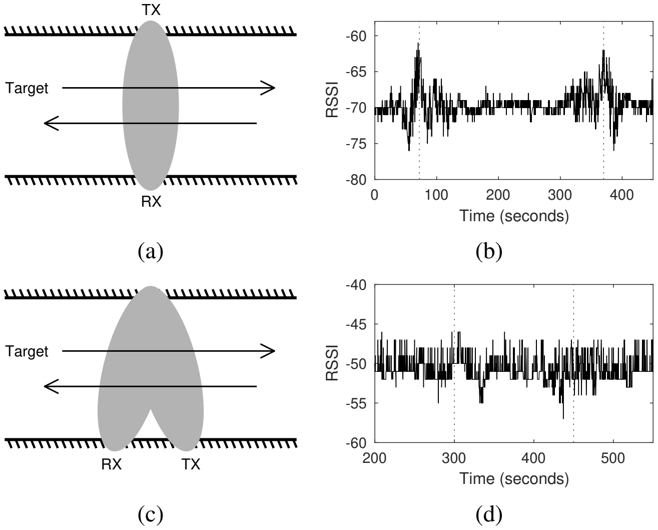


Fig. 1: Through-wall hallway experiment diagrams. Dark lines indicate the walls of the hallway while the gray areas indicate the multi-path propagation of the radio signals from transmitter (TX) to receiver (RX) as the target walks through the hallway. (a) LOS experiment setup, (b) RSSI for LOS experiment, (c) NLOS experiment setup, (d) RSSI for NLOS experiment.

While there exist studies which consider through-wall crowdcounting such as in [4]; in our work, we take these efforts further and investigate if it is possible to perform crowdcounting successfully when the access to the monitored area is much more limited. That is, in existing through wall research, it is assumed that both transmitter (TX) and receiver (RX) can be placed across from one another as illustrated in Fig. 1a. In an adversarial scenario however, an attacker may not be able to place the devices in both areas to perform LOS sensing. Instead, an attacker may only have access to a single room in a building or to the exterior of the building resulting in a limited monitoring ability from only one side. However, this limits the attacker to NLOS conditions as shown in Fig. 1c. Indeed, WiFi sensing based recognition has been shown to be successful in NLOS scenarios [5] in a hallway environment, however the radio is considered to be in the center of the hallway rather than being hidden behind a wall. Also note that these methods require extensive training phases with labeled data for each new target environment before successful results are achieved. Thus, any person tasked with tracking targets must have full access to the environment and then must perform time-consuming setup training before each new deployment. Our proposed system in this study instead leverages signal features common to all environments for prediction. While there are several existing studies that look at the through-wall occupancy detection and crowdcounting problem through device-free WiFi sensing, to the best of our knowledge, this NLOS scenario has not been considered yet, while it is a more practical scenario for an adversary.

The rest of the paper is organized as follows. We first review related work in WiFi sensing and through-wall sensing in Section II and provide the motivation for this study. In Section III, we then provide the proposed method in which we describe the environmental setting where the signals are monitored and the pre-processing steps for the signal analysis. Next, in Section IV, we elaborate on the proposed detection framework that is used to understand both human presence and walking direction followed by our experimental results. Finally, we provide our concluding remarks and discuss on future work in Section V.

II. Background

Existing studies in occupancy monitoring and crowdcounting have typically been accomplished through the use of video camera streams. For example, in [6], the authors use an approach to count the number of pedestrians passing some designated line in the view of a static camera. In densely crowded environments, head-counting techniques [7] have been shown to be sufficient when full bodies are not visible to the camera. Mobile cameras have also been considered, with the use of drones [8] which can follow crowds as they change shape and size such as during protest marches. In indoor environments [9], smaller scales must be considered because of the restrictions inherited from the inclusion of walls blocking the sight of camera devices. In commercial buildings, existing surveillance camera systems can be further leveraged to collect occupancy analytics over time [10].

In addition to camera-based systems, other building-specific data sources have been used to predict occupancy monitoring. Data sourced from electrical power meter usage [11] has been shown to reveal occupancy from residential smart meters. Collecting data from additional sensors placed throughout an office building such as CO_2 sensors, temperature sensors and light sensors have been demonstrated to allow for high accuracy occupancy detection [12], however, these sensors are not usually placed within typical indoor environments. Similarly, the deployment of new radio based occupancy detection systems using Radio Frequency Identification (RFID) tags [13] and Bluetooth Low Energy (BLE) beacons [14] have shown to provide promising results. However, these systems not only require deployment of new hardware into the environment but also require individuals to carry additional hardware on the body for tracking (i.e., not device-free). The requirement to deploy additional hardware has been avoided in studies (e.g., [15]) that leverage the existing WiFi infrastructure in commercial buildings to track the connected devices such as smartphones of users through their MAC addresses. However, such an approach will not work as some devices may not be connected to the WiFi router and further some people may not even have a transmitting device with them.

WiFi sensing [16]–[18] has recently gathered interest in allowing for wireless, device-free sensing of environments using standard WiFi packet transmissions. WiFi sensing is possible through the use of Channel State Information (CSI) from the WiFi devices in orthogonal frequency-division multiplexing (OFDM) systems [19]. CSI is a metric composed of signal amplitude and phase across N subcarrier frequencies used in the process of link adaption. In link adaption each subcarrier is able to transmit symbols at adaptable data rates and power levels in parallel across each subcarrier [19] to allow for multiple-input multiple-output (MIMO). OFDM transmits shared pilot symbols interleaved within the data frame which can then be used to estimate a shared value for CSI H for the pair of devices as described in the equation:

$$y^{(i)} = H^{(i)}x^{(i)} + \eta^{(i)} \tag{1}$$

where i indicates the subcarrier index, $y^{(i)}$ indicates the signal characteristics received, $x^{(i)}$ is the actual transmitted signal (shared pilot) and $\eta^{(i)}$ is a noise vector. The CSI vector H consists of complex numbers with the combination of both real and imaginary numbers representing the attributes of the received signal. With $H_r^{(i)}$ as the real value of H at subcarrier i and $H_{im}^{(i)}$ as the imaginary value, we can then compute amplitude ($A^{(i)}$) and phase ($\phi^{(i)}$) for each subcarrier i through the following equations:

$$A^{(i)} = \sqrt{(H_{im}^{(i)})^2 + (H_r^{(i)})^2} \tag{2}$$

$$\phi^{(i)} = atan2(H_{im}^{(i)}, H_r^{(i)}) \tag{3}$$

For each frame, we receive values for 64 subcarriers where 52 of them contain actual CSI data.

WiFi sensing has previously been used to identify individual targets through the analysis of unique effects of walking movements such as torso speed and gait on the WiFi signals [20]. Further works have considered the crowdcounting problem using WiFi sensing such as in [21]–[23] where some number of targets (up to 10) are contained within a room and asked to walk in random paths around a given area. The transmitters and receivers in each of these works are co-located in the same room with the targets. The targets must continuously walk within the area for upwards of an hour before the model can successfully make any predictions. Furthermore, if any target stops walking or targets are added or removed, then the model will not be able to make accurate predictions. Beyond these issues, in a surveillance situation, radios co-located with targets would arouse suspicion. Instead, keeping radios out of visual sight of targets by performing through-wall WiFi sensing was demonstrated in [4]. The system described requires that the transmitter is placed behind one wall while the receiver is placed behind the opposite wall as illustrated in Fig. 1a. As such, target counting is accomplished similarly to existing non-through-wall works [21] in which targets are only recognized when they pass the LOS of the transmitter and receivers.

III. Proposed Method

We begin explaining our through-wall occupancy monitoring system by first performing empirical experiments in a real-world hallway environment. For the first experiment, we record radio signal data in LOS conditions as illustrated in Fig. 1a as performed in previous works [21], [23]. Then we move transmitter and receiver into NLOS positions shown in Fig. 1c.

RSSI has previously been used as a simple and more easily obtained signal metric because of its immediate availability on smartphones and other consumer radio-enabled devices. In LOS, RSSI works well to recognize targets as demonstrated in our experiment result in Fig. 1b where the vertical dotted lines indicate the time when the target is passing the LOS. We can see directly that as the target passes, more RSSI variation occurs. However, if we perform a NLOS experiment as illustrated in Fig. 1c, we find that RSSI no longer reveals any signal variations when the target passes as we can see in our NLOS experiment result in Fig. 1d. Instead, in this work, we evaluate the use of the CSI signal metric in similar situations which gives much more fine grained details compared to RSSI.

A. CSI Pre-Processing

Channel State Information varies in new environments because of the unique multi-path conditions of each location. Thus, received $A^{(i)}$ and the change of $A^{(i)}$ over time will vary uniquely when similar actions are performed but in unique environmental conditionals. To combat this for our occupancy monitoring problem, we suggest the following signal pre-processing steps be applied to $A^{(i)}$. We begin the pre-processing by applying a windowed outlier filter. That is, we find

$$\bar{A}_t^{(i)} = \begin{cases} A_t^{(i)}, & \frac{\left|A_t^{(i)} - \mu\left(A^{(i)}\{t-w_1:t\}\right)\right|}{\sigma\left(A^{(i)}\{t-w_1:t\}\right)} < \lambda \\ A_{t-1}^{(i)}, & \text{otherwise} \end{cases} \tag{4}$$

where

$$\mu(\mathbf{x}) = \frac{1}{|\mathbf{x}|}\sum_{j=1}^{|\mathbf{x}|} \mathbf{x}^{(j)} \tag{5}$$

is the mean function which is applied to the received signal from time $t-w_1$ until the current time t on $A^{(i)}$, where w_1 represents the window length parameter. Similarly,

$$\sigma(\mathbf{x}) = \sqrt{\frac{\sum_{j=1}^{|\mathbf{x}|}\left(\mathbf{x}^{(j)} - \mu(\mathbf{x})\right)^2}{|\mathbf{x}|}} \tag{6}$$

is the standard deviation function applied to the same window of $A^{(i)}$. The goal of (4) is to replace any outlier samples (those which are greater than λ standard deviations from the mean) with the most recent valid/normal sample. Outlier filtering is applied to each subcarrier independent of one another.

After filtering outliers, we want to gather aggregated statistics across all subcarriers independently across another time window of size w_2. Here, while we can use different values for both w_1 and w_2, for simplicity, we keep $w_1 = w_2$. For this, we apply some windowed statistical aggregation function $\Phi(\mathbf{x})$ on each subcarrier independently,

$$\tilde{A_t}^{(i)} = \Phi\left(\bar{A}^{(i)}_{(t-w_2:t)}\right). \tag{7}$$

For our experiments, we consider $\Phi(\mathbf{x}) \equiv \sigma(\mathbf{x})$ because our goal is to understand how noisy each subcarrier is independently, however $\Phi(\mathbf{x})$ can be replaced with any other statistical function as needed.

After collecting a noise metric for all time instances on each subcarrier, we want to find if the noise present in one subcarrier is similar to the noise present in other subcarriers. Again, we apply a new statistical aggregation function $\Psi(\mathbf{x})$, this time on all subcarriers for a single time instance t,

$$A_{CSI,t} = \Psi\left(\tilde{A_t}^{(1:|A_t|)}\right). \tag{8}$$

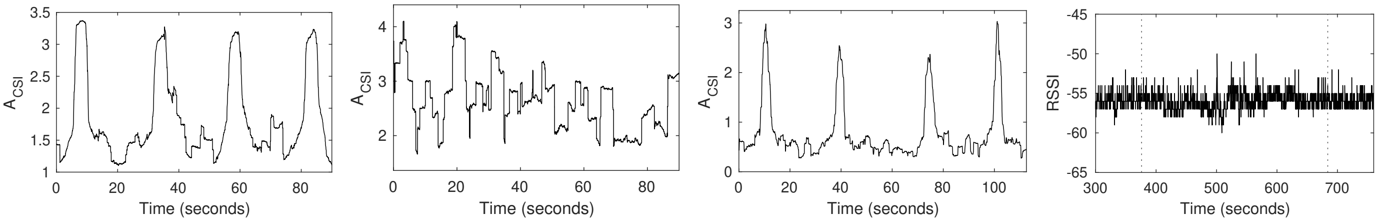
For our purposes, $\Psi(\mathbf{x}) \equiv \mu(\mathbf{x})$ with the intuition that if all subcarriers are high in noise, then $A_{CSI,t}$ will be larger than it is when only a small subset of subcarriers are affected by noise. This is important because the noise resulting from the environment may cause subcarriers to randomly produce noise which is not represented across any other subcarriers. Instead, when a target is present, noise will be present across more subcarriers which will more consistently produce a higher value for $A_{CSI,t}$. On the other hand, when no target is present, any noise anomalies present on a single subcarrier will be filtered out because of disagreement across subcarriers. For notation simplicity, we will denote $A_{CSI} \equiv A_{CSI,t}$ with the understanding that A_{CSI} is a scalar metric for some time instance t.

B. Standard LOS Through-Wall

As shown in many previous experiments in WiFi sensing, recognizing targets as they pass through the LOS between a transmitter and a receiver is a trivial task and can then be used to estimate the number of targets in an area [4], [23]. We perform our first experiment with one target passing through the LOS. The results in Fig. 2a show when a target passes the LOS four separate times, A_{CSI} gives distinct peaks, indicating that the target has passed by the receiver four separate times. In between these passing events, A_{CSI} returns to some lower noise-floor level.

C. NLOS Through-Wall

As discussed, in certain environments it may not be possible to place a transmitter and a receiver to create such LOS conditions. For example, when rooms are not available on both sides of a hallway or if access is restricted for these adjacent rooms. In these cases, it would be most advantageous



(a) (b) (c) (d)

Fig. 2: A_{CSI} for (a) LOS experiment setup, (b) NLOS experiment setup without directional shielding, (c) NLOS experiment with directional shielding, and (d) RSSI with directional shielding. Target is still not detectable with RSSI even with shielding.

for both the transmitter and the receiver to be placed in a single room together against one wall. This is particularly useful in adversarial conditions where an attacker has access to only a single location because they can then keep an eye on the radios as they perform sensing tasks. Fig. 1c illustrates this setup. To the best of our knowledge, such adversarial placement of the transmitter and the receiver has not been attempted by any of the device-free WiFi sensing studies in the literature.

In this case, the target will no longer be in the LOS of the devices, thus a NLOS monitoring will be required. For our first experiment with this NLOS placement, we position both a transmitter and a receiver 6 meters apart, both 50 centimeters away from the wall. The resulting A_{CSI} in Fig. 2b shows that the target passing times are not clearly visible. This is because the direct LOS between the transmitter and the receiver dominates the received signal path which travels through the wall and comes back. As a result, the targets passing through the hallway does not cause a distinguishable change on the received signal amplitude collected. Thus, an update to the setting is needed in order to make the effect of such through-wall NLOS signals prominent.

Typically, WiFi antennas are designed to transmit omnidirectionally, but unidirectional antennas allow for signals to be focused in more specific areas. Unidirectional antennas however must be aimed with great accuracy to ensure that signals are eventually received by the receiver. This may not be an easy task without knowing the characteristics of the environment on the other side of the wall. Our solution is to shield both transmitter and receiver with an aluminum tin placed pointing the wall. This prevents the direct LOS signal from dominating the NLOS signal while still allowing for partial omnidirectional propagation in the target area. This will be additionally useful if multiple receivers are used for through-wall sensing with a single transmitter. After adding the directional shielding, we can see in Fig. 2c that we can again identify distinct peaks when the target passes through the hallway environment. Note that RSSI still cannot be used to recognize the passing target in this directional setting as shown in Fig. 2d.

IV. Detection Framework and Evaluation

We now move on to defining our full framework for target detection. For all of the experiments in this work, we use our ESP32-CSI-Toolkit[1] [17] to collect CSI which uses two ESP32 WiFi-enabled microcontrollers for our transmitter and receiver, respectively. Using these small, low-cost microcontrollers demonstrates how an adversary could both implement and distribute large numbers of adversarial devices with less fear of discovery because of their small size and without fear of loss because of the low-cost of each stand-alone ESP32 module. The ESP32 devices are set to send and receive CSI at a packet rate of 100Hz. The entire framework is designed such that a low resource device such as the ESP32 can perform all tasks in real time without requiring additional external computation power such as a server of laptop as it is often required in WiFi sensing literature. From an adversarial perspective, this is important to ensure that the devices remain small and easy to conceal.

A. Human Presence

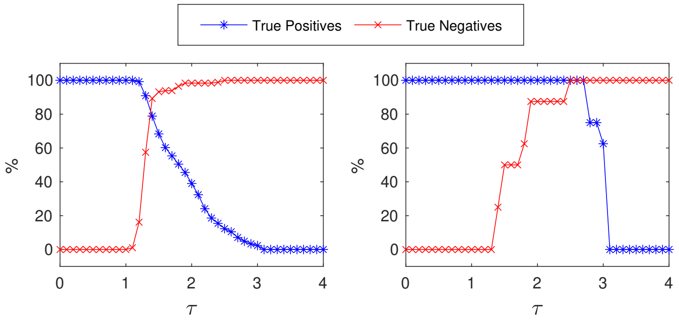
As shown in Section III, we see that our A_{CSI} metric can be used visually to detect activity whenever a target passes. For our model to predict the binary presence of a target, we designate a threshold parameter τ. When $A_{CSI} \geq \tau$, then the model predicts the presence of the target. We perform our experiment with a target passing the monitored area five times. For each time instance, our model predicts whether a target is present. To evaluate how well different thresholds work in predicting the class of our samples, we define a class prediction probability metric $P^{(c)}_{samples}$ for samples of a given class $c \in \{\text{‘target’}, \text{‘no target’}\}$. The class for a sample at time t is denoted as $\mathbf{C}^{(t)}$. We thus define $T^{(c)} = \{t \in \boldsymbol{T}\ s.t.\ \mathbf{C}^{(t)} = c\}$ to be the set of time instances labeled as class c, where $\boldsymbol{T}$ is the set of all time instances. Further, $N^{(c)} = \left|T^{(c)}\right|$ is the number of CSI frame samples marked as class c. From this, we define $P^{(c)}_{samples}$ as:

$$P^{(\mathrm{c})}_{samples} = \frac{1}{N^{(\mathrm{c})}} \sum_{t \in \boldsymbol{T}^{(\mathrm{c})}} Y(t, c) \tag{9}$$

where

$$Y(t,c) = \begin{cases} 1 & \textit{if } A_{CSI,t} \geq \tau \text{ and } c = \text{‘target’} \\ 1 & \textit{if } A_{CSI,t} < \tau \text{ and } c = \text{‘no target’} \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

[1] https://github.com/StevenMHernandez/ESP32-CSI-Tool



(a) (b)

Fig. 3: Prediction accuracy as threshold parameter τ changes. (a) With all recorded samples using $P_{samples}$. (b) With all independent action segments using $P^{(c)}_{segments}$.

We can interpret $P^{(c)}_{samples}$ as the percentage of CSI frame samples which are truly class c and are predicted as class c; or put simply, the true-positive and true-negative rates. In Fig. 3a we see the results of our model. As it is expected, as τ increases, $P^{(\text{'no target'})}_{samples}$ increases and $P^{(\text{'target'})}_{samples}$ decreases. Specifically, when $\tau < 1.0$, $P^{(\text{'no target'})}_{\text{samples}} = 0.0$ and $P^{(\text{'target'})}_{samples} = 1.0$, this is because there are no samples where $A_{CSI} < 1.0$. In addition to this, we can see that there is no value for τ where we are able to achieve perfect accuracy on predicting both the true positives and true negatives. However, our goal is not to predict the action for all time instances individually. Instead, we are only interested in correctly classifying each action segment overall.

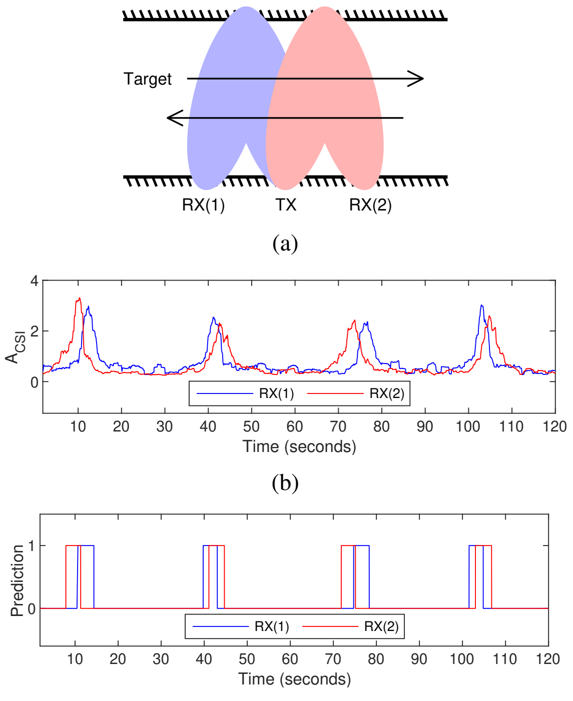
To define action segments, we first collect a set of time instances ($\mathbf{I}$) which indicate the beginning and ending of different actions. To determine these indices, we apply the following:

$$\mathbf{I} = \{0\} \cup \left\{ t \in \{2:\mathbf{T}\} \textit{ where } C^{(t-1)} \neq C^{(t)} \right\} \cup \{\mathbf{T}\}. \quad (11)$$

With this, we can describe the number of action segments recorded $N_{seg} = |\mathbf{I}| - 1$. We say that the target is predicted as present ($P^{(i)}_{\text{'target'}}$) during some action segment i if $\exists t \in \{\mathbf{I}^{(i)} : \mathbf{I}^{(i+1)}\}$ *s.t.* $A_{CSI,t} \geq \tau$. Action segments containing no target on the other hand are denoted $P^{(i)}_{\text{'no target'}}$ which is simply the negation of $P^{(i)}_{\text{'target'}}$. The number of segments for a given class is described as $N^{(c)}_{seg}$. To evaluate our predictions on all segments, we define:

$$P^{(c)}_{segments} = \frac{1}{N^{(c)}_{seg}} \sum_{i=1}^{N_{seg}} \begin{cases} 1 & \textit{if } P^{(i)}_{(c)} \text{ and } \mathbf{C}^{(t)} = c \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

which is described as the percentage of segments correctly labeled as class c. With this, our final goal is to find a value for τ such that we maximize the number of segments where a target was present and minimize the number of time segments predicted as containing a target when no target was present. When considering this segmented approach, we see in Fig. 3b that when $\tau \in [2.5, 2.7]$ both the true positive and true negative rate reach 1.0, indicating a range of perfect predictions.



(a)

(b)

(c)

Fig. 4: Using two receivers we are able to identify the directional movement of the human target based on which receiver sees an increase in A_{CSI} first. (a) Experiment setup with all adversarial ESP32 devices on one side of the wall: TX at the center, RX(1) to the left of TX and RX(2) to the right. (b) Raw A_{CSI} showing four peaks when the target moves back and forth within the hallway environment, (c) After applying binary human detection algorithm, we can even more clearly identify the human target direction.

B. Human Direction

Our next task is to recognize the moving direction of the target in the hallway environment. While we show in Section III that A_{CSI} reveals human presence, moving direction of the target is not directly revealed by the metric. To address this, we use two receiving devices, one located to the left of the transmitter and the other to the right as shown in Fig. 4a. Both receivers again use directional shielding so that when the central transmitter sends radio signals, they are both able to receive the signals for different areas within the hallway environment. The expectation is that when a target moves from left to right, the device located to the left-most side will recognize the target first, then later on, the right-most device will recognize the target. Afterwards, we would expect the left-most device will stop recognizing the target before the right-most device. In Fig. 4b we see A_{CSI} for both RX(1), which is placed to the left-hand side of the transmitter (TX), and RX(2), which is placed to the right. As the target moves back and forth through the hallway environment, A_{CSI} for RX(2) increases before A_{CSI} for RX(1) indicating that the user moved in the direction of right-to-left. In the second pair of peaks at around 40 seconds, A_{CSI} for RX(1) increases first, indicating that the target moved back to the starting point from left-to-right. By applying the binary human detection algorithm from Section IV-A, we can see this relationship even more clearly as shown in Fig. 4c.

V. Conclusion

In this work, we study the use of Channel State Information for adversarial through-wall occupancy monitoring in hallway environments. We demonstrate through real world experiments how an attacker could perform surveillance of a building if given access to a single room or even from a single exterior wall. Through the use of our previously developed WiFi sensing toolkit [17], we demonstrate how this sort of attack is very low cost and much easier to conceal compared to camera-based surveillance methods. Using the signal pre-processing steps proposed in this work, we are able to demonstrate that the two components required for tracking humans, namely, presence and moving direction, can be successfully predicted even in one-sided through-wall scenarios which can then be used for crowdcounting by adversaries.

In our future work, we will look at person detection and tracking with multiple targets and analyze the impact of different parameters such as the speed of targets, distance of the targets from the wall as well as overlapping of target movement within the target hallway area. Moreover, because this work shows that it is possible for an adversary to identify human presence and direction information which could be used for crowdcounting purposes, we will investigate defense mechanisms to prevent such detection scenarios in passive WiFi receiving mode.

Acknowledgements

This work is supported in part by National Science Foundation (NSF) Graduate Research Fellowship Program (GRFP) under Grant No. 1744624, Virginia Commonwealth University Presidential Research Quest Fund (PeRQ) and Commonwealth Cyber Initiative (CCI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agencies.

References

[1] C. Wang, K. Pattawi, and H. Lee, "Energy saving impact of occupancy-driven thermostat for residential buildings," *Energy and Buildings*, vol. 211, p. 109791, 2020.

[2] G. Aravamuthan, P. Rajasekhar, R. K. Verma, S. V. Shrikhande, S. Kar, and S. Babu, "Physical intrusion detection system using stereo video analytics," in *Proceedings of 3rd International Conference on Computer Vision and Image Processing*, B. B. Chaudhuri, M. Nakagawa, P. Khanna, and S. Kumar, Eds. Singapore: Springer Singapore, 2020, pp. 173–182.

[3] Z. Zhang, M. Wang, and X. Geng, "Crowd counting in public video surveillance by label distribution learning," *Neurocomputing*, vol. 166, pp. 151–163, 2015.

[4] S. Depatla and Y. Mostofi, "Crowd counting through walls using WiFi," in *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2018, pp. 1–10.

[5] S. Fang, R. Alterovitz, and S. Nirjon, "Non-Line-of-Sight around the corner human presence detection using commodity WiFi devices," in *Proceedings of the 1st ACM International Workshop on Device-Free Human Sensing*. ACM, 2019, pp. 22–26.

[6] Z. Ma and A. B. Chan, "Crossing the line: Crowd counting by integer programming with local features," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2013.

[7] M. B. Shami, S. Maqbool, H. Sajid, Y. Ayaz, and S. S. Cheung, "People counting in dense crowd images using sparse head detections," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 9, pp. 2627–2636, Sep. 2019.

[8] M. Küchhold, M. Simon, V. Eiselein, and T. Sikora, "Scale-adaptive real-time crowd detection and counting for drone images," in *25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 943–947.

[9] D. Ryan, S. Denman, C. Fookes, and S. Sridharan, "Scene invariant multi camera crowd counting," *Pattern Recognition Letters*, vol. 44, pp. 98–112, 2014.

[10] I. Mutis, A. Ambekar, and V. Joshi, "Real-time space occupancy sensing and human motion analysis using deep learning for indoor air quality control," *Automation in Construction*, vol. 116, p. 103237, 2020.

[11] R. Razavi, A. Gharipour, M. Fleury, and I. J. Akpan, "Occupancy detection of residential buildings using smart meter data: A large-scale study," *Energy and Buildings*, vol. 183, pp. 195–208, 2019.

[12] L. M. Candanedo and V. Feldheim, "Accurate occupancy detection of an office room from light, temperature, humidity and co2 measurements using statistical learning models," *Energy and Buildings*, vol. 112, pp. 28–39, 2016.

[13] N. Li, G. Calis, and B. Becerik-Gerber, "Measuring and monitoring occupancy with an RFID based system for demand-driven HVAC operations," *Automation in Construction*, vol. 24, pp. 89 – 99, 2012.

[14] P. Barsocchi, A. Crivello, M. Girolami, F. Mavilia, and F. Palumbo, "Occupancy detection by multi-power bluetooth low energy beaconing," in *IEEE International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2017, pp. 1–6.

[15] E. Vattapparamban, B. S. Çiftler, I. Güvenç, K. Akkaya, and A. Kadri, "Indoor occupancy tracking in smart buildings using passive sniffing of probe requests," in *IEEE International Conference on Communications Workshops (ICC)*, 2016, pp. 38–44.

[16] Y. Ma, G. Zhou, and S. Wang, "WiFi Sensing with Channel State Information: A Survey," *ACM Comput. Surv.*, vol. 52, no. 3, pp. 46:1–46:36, 2019. [Online]. Available: https://doi.org/10.1145/3310194

[17] S. M. Hernandez and E. Bulut, "Lightweight and Standalone IoT based WiFi Sensing for Active Repositioning and Mobility," in *21st International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*, Cork, Ireland, Jun. 2020.

[18] S. M. Hernandez and E. Bulut, "Performing WiFi Sensing with Off-the-shelf Smartphones," in *IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, 2020, pp. 1–3.

[19] T. Hwang, C. Yang, G. Wu, S. Li, and G. Y. Li, "OFDM and Its Wireless Applications: A Survey," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 1673–1694, May 2009.

[20] B. Korany, C. R. Karanam, H. Cai, and Y. Mostofi, "XModal-ID: Using WiFi for Through-Wall Person Identification from Candidate Video Footage," in *25th Annual International Conference on Mobile Computing and Networking*. ACM, 2019, p. 36.

[21] O. T. Ibrahim, W. Gomaa, and M. Youssef, "CrossCount: A Deep Learning System for Device-Free Human Counting Using WiFi," *IEEE Sensors Journal*, vol. 19, no. 21, pp. 9921–9928, 2019.

[22] S. Liu, Y. Zhao, and B. Chen, "WiCount: A Deep Learning Approach for Crowd Counting Using WiFi Signals," in *IEEE International Symposium on Parallel and Distributed Processing with Applications and IEEE International Conference on Ubiquitous Computing and Communications (ISPA/IUCC)*, 2017, pp. 967–974.

[23] H. Zou, Y. Zhou, J. Yang, W. Gu, L. Xie, and C. J. Spanos, "FreeCount: Device-Free Crowd Counting with Commodity WiFi," in *IEEE Global Communications Conference (GLOBECOM), Singapore, December 4-8, 2017*, pp. 1–6.