FormaTrack: Tracking People based on Body Shape

AVINASH KALYANARAMAN, University of Virginia, USA
DEZHI HONG, University of Virginia, USA
ELAHE SOLTANAGHAEI, University of Virginia, USA
KAMIN WHITEHOUSE, University of Virginia, USA

Knowledge of a person's whereabouts in the home is key to context-aware applications, but many people do not want to carry or wear a tag or mobile device in the home. Therefore, many tracking systems are now using so-called *weak biometrics* such as height, weight, and width. In this paper, we propose to use body shape as a weak biometric, differentiating people based on features such as head size, shoulder size, or torso size. The basic idea is to scan the body with a radar sensor and to compute the *reflection profile:* the amount of energy that reflects back from each part of the body. Many people have different body shapes even if they have the same height, weight, or width, which makes body shape a stronger biometric. We built a proof-of-concept system called *FormaTrack* to test this approach, and evaluate using eight participants of varying height and weight. We collected over 2800 observations while capturing a wide range of factors such as clothing, hats, shoes, and backpacks. Results show that *FormaTrack* can achieve a precision, recall, direction and identity accuracy (over all possible groups of 2 people) of 100%, 99.86%, 99.7% and 95.3% respectively. Results indicate that *FormaTrack* can achieve over 99% tracking accuracy with 2 people in a home with 5 or more rooms.

CCS Concepts: • **Computer systems organization** → **Sensors and actuators**;

Additional Key Words and Phrases: Smart Homes, Radar, Tracking Systems, Wireless Systems

1 INTRODUCTION

Knowledge of a person's whereabouts in the home is key to context-aware applications, such as personalized heating and cooling, entertainment, task assistance, and behavioral or health analysis. Most tracking systems today require people to wear tags, carry mobile devices, or have cameras or microphones but many people do not accept these technologies in the home [30], and particularly not the aging population that needs in-home care the most. As an alternative, many tracking systems are using so-called *weak biometrics*: physical characteristics of the body that can differentiate people, but that do not necessarily uniquely identify them. For example, several systems over the past several years have tracked people based on height [29, 63], weight [7, 58], and width [36]. Weak biometrics can be effective in a home environment because there are typically only a handful of residents and they often have different biometric features due to age, gender, and/or family role. However, weak biometrics

break down in environments with a large number of people (such as office buildings) or with people who have similar biometric features by chance.

In this paper, we propose a new weak biometric feature called *body shape* that differentiates people based on features such as head size, shoulder size, or torso size. The basic idea is to scan the body with a radar sensor and to compute the *reflection profile:* the amount of energy that reflects back from each part of the body. These energy levels are indicative of the relative size of each part of the body. Many people have different body shapes even if they have the same height, weight, or width, which makes body shape a stronger biometric. Even twins often have different body shapes. The most accurate body shape measurements would be achieved with sub-millimeter wave radars, such as those used as airport scanners [13], but we hypothesize that simpler, lower-power, and more compact radar sensors could adequately differentiate people in a home while still achieving higher accuracy than existing weak biometric sensors.

To test this hypothesis, we built a proof-of-concept system called *FormaTrack* using a monostatic pulse radar that measures total reflected energy at a given distance d, without collecting any information about angle of arrival. Thus, it produces a single dimensional measurement vector, as illustrated in Figure 1, that does not at all resemble the 3D imaging output produced by an airport scanner. We mounted the sensor at the top of a door frame pointed downward so that it measures the distance to the ground. When a person walks through the door frame, the sensor receives different levels of reflected energy from each part of their body. We use Doppler shift to detect the exact moment when the person is in the door frame, and collect the reflection profile at that moment to ensure the signature is repeatable. Then, we compare the signature to previous measurements in order to recognize the occupant. Additionally, we tilt the sensor in order to create an asymmetric sensing region, similar to prior work [29], in order to detect the walking direction. The identity and direction combined is sufficient for room-level tracking of occupants in the home.

The sensor that we use has a maximum sensing range of 1m and therefore does not pick up most of the torso size or any gait features observable in the arms and legs. It senses the person from above and therefore features of the face and neck are also occluded. Nevertheless, this sensor demonstrates that body shape can be an effective biometric feature. We evaluate *FormaTrack* using eight participants of varying height and weight who were asked to walk through a doorway in different ways every day for seven days. This process generated over 2800 crossing events while capturing a wide range of factors such as clothing, hats, shoes, and backpacks. Results show that *FormaTrack* can achieve a precision, recall, direction and identity accuracy (over all possible groups of 2 people) of 100%, 99.86%, 99.7% and 95.3% respectively. Additionally, on testing *FormaTrack* for 36 hours (over 12 days) when no one crossed the doorway, *FormaTrack* produced no false detections. To evaluate, how *FormaTrack* translates to a whole-home tracking accuracy, we use the empirical data to emulate 15 different floor plans varying from 3 to 9 rooms and test with data from 2 to 4 participants. We observe that *FormaTrack* can achieve over 99% tracking accuracy with 2 people in a home with 5 or more rooms. Even with 4 people in a home with only 3 rooms, room-level tracking accuracy is still above 92%. Finally, we demonstrate that simple techniques can be used to reduce average power consumption by over 70% while missing fewer than 1% of all doorway crossings.

2 RELATED WORK

Non-intrusive identification and tracking of people in the home environment has been an open problem for several years. Early work used smart-floor based sensing systems [7, 42, 58] to track people based on their weights. However, this approach would mis-identify somebody if they wear something heavy such as a backback. Subsequent work tracked people based on height [29, 63], but this approach would mis-identify people if they wore a hat or shoes. This approach was extended by complementing height with smartphone connectivity data [39], infrared sensors [27], or person width measurements [36], but these approaches are subject to compliance

issues or error based on body position, respectively. Another system attempts to identify people by sensing their shadows [12, 40] but requires a dense deployment of photodiodes on the floor, which would be difficult to install in homes. Several wireless-based device-free systems exist that attempt to identify people based on their gait [26, 31, 33, 68]. These systems typically analyze the variation of the received wireless signal to infer gait information (such as gait cycle time, walking speed, stride length, torso speed etc). However, these systems require a person to walk in a straight line for about 5m [33, 68], in order to collect enough information about the gait of the person, and gait can be affected by shoes, carried objects, the use of walking aids, and other factors. Vision based systems [18, 19, 32, 44] can also identify and track a person. For example, vision systems capture a facial image [44], the gait [18] or iris data [19] as biometrics. Other vision based systems [16, 22] try to extract the silhouette of a person from the video frames and identify them based on their body shape. One limitation of these techniques, however, is the need to deal with lighting issues such as darkness. Additionally, people often do not accept video based systems into their homes, especially given the possibility that the device could be hacked. *FormaTrack* avoids these problems by using wireless signals to sense the identity of people in a non-invasive way.

Other RF based sensors can perform much higher resolution imaging of humans to obtain their shape [13, 24]. These systems use sub-millimeter (for e.g. airport scanners) or millimeter waves, but are expensive and large in size [13], making it currently unsuitable for people tracking in homes. *RF-Capture* [8] is a centimeter-wave system similar to ours that senses the body shape of people in a room. However, it requires people to walk towards the device for shape sensing, and also has a more complicated hardware than *FormaTrack* involving 4 transmitters and 16 receivers. *FormaTrack* attempts to infer people's identity at the doorway with a single transceiver. Finally, depth imaging sensors such as the Microsoft Kinect can detect human body shape with high accuracy [47], but it is not yet a low power device and therefore would be difficult to use for body scanning. However, future versions of this sensor will be able to serve well for body shape sensing.

Aside from the systems described above, most indoor tracking systems either sense and track people but do not infer their identities [9, 41, 61, 62, 66, 70], or they infer identities by requiring wearable tags or electronic devices. So called device-based localization systems [45] include Cricket [52], Active Badges [69], RADAR [15], and Blue Sentinel [23] that use ultrasound, infrared, wireless fingerprinting, and Bluetooth Low Energy, respectively. Some other works have attempted to localize people based on FM-signal [21, 59], the powerline infrastructure [51] or GSM signal fingerprinting [50]. More recently, several device-based sensing techniques using WiFi attempt to localize the device (person) by measuring the angle-of-arrival or time of flight with respect to one or more access points [37, 38, 65, 72]. In addition, there are motion-capture systems, such as Vicon, Xsens, Zebra etc [4–6], that can capture the human figure and infer their identity, but requires the person to have several markers (sensors) on their body. However, all these systems require a person to carry/wear a device (or at times even rotate the device [38]) to be localized. While such systems may be practical in commercial settings, they are not accepted in homes due to the so-called *forget to wear, forget to charge* problem [30]. We refer the reader to Xiao et al. [71] for a more detailed survey of device-based and device-free localization systems.

Our system, *FormaTrack* is a radar-based system that leverages radar principles to perform identification and room-level tracking of people. Far from its initial domains of aviation, navigation or weather forecasting, radars have now entered into homes with applications such as non-intrusive vital signs measurement [10], gait analysis [31], and fall detection [11]. Furthermore, with the advent of Google Soli [43], millimeter-wave based radar systems have been designed for hand-gesture recognition [67], object and material identification [73], or even to be a musical instrument interface [17]. However, *FormaTrack* is different from these systems in that it uses a radar for identification and room-level tracking of people via body shape.

Finally, *FormaTrack* uses the Doppler principle to identify when the person walks from one side of the doorway to the other. The principle of Doppler has been used for several applications such as inferring hand gestures [28, 43, 53], sleep-sensing [54], connecting mobile devices [14], and tracking phone position [74]. They all rely on the phenomenon that the frequency of the signal changes when the transmitter or receiver or a nearby

reflecting object (*virtual transmitter)* moves. *FormaTrack* builds on the same principle but uses it in a different context.

3 APPROACH

Any doorway tracking system should perform three main tasks : (i) detect when a person crosses through the doorway, (ii) estimate their direction of movement, and (iii) infer the identity of the person who made the doorway transition. Our system *FormaTrack* is a radar-based system mounted atop the doorway that analyzes the signals reflected back from the environment to perform the aforementioned tasks. Since *FormaTrack* is radar-based, we first provide a quick overview of key radar terminology and concepts.

A pulse radar is a device that emits a pulse of electromagnetic energy towards an expected target area, where a portion of this radiated energy is reflected back by a target. The radar device analyzes this reflected signal to obtain information about the target. For example, the amount of reflected energy indicates the target's radar cross section (size, curvature, reflectivity, etc) and the frequency of the reflected signal indicates the targets velocity, as per the Doppler principle. A radar device is said to be *monostatic* when its transmitter and receiver are collocated. In *FormaTrack*, a *monostatic* radar is mounted atop the doorway. The transmitter of our monostatic radar radiates pulses towards the target area, while the receiver measures the reflected energy from the pulse interfering with targets.

The output of a pulse radar is a *radar frame* that contains information about the total reflected power at varying distances from the radar. These distances are quantized into *bins* called *range bins*, and the size of each bin corresponds to the *range resolution* of the radar (i.e. the minimum distance between two distinguishable targets). This *range resolution* of a radar depends on the sampling rate of the receiver. For example, a receiver that samples at 40GS/s can provide a range resolution of about 4mm. The total number of bins in a *radar frame* depends on the *maximum detectable range* of the radar (i.e. how far can a target be positioned to be detected by the radar). Figure 1 shows an example of a *radar frame* with two dominant reflectors at about 0.4m and 0.6m respectively. The size of each *range bin* in this frame is 4mm. There are a total of 256 *range bins* in this frame resulting in a *maximum detectable range* of 1.024m.

Radars typically deal with two time dimensions - *fast time* and *slow time* [56]. *Fast time* refers to the time between two pulses - i.e. it represents the range bins for a given pulse (frame), and is a function of the sampling rate of the radar. On the other hand, *slow time* refers to the time dimension in the granularity of a pulse. It is a function of the pulse repetition frequency (frame rate) of the radar. For a more detailed review of radar and its principles, we point the readers to [56, 60]. The following sub-sections explain how we use the radar to detect crossings, direction and body shape.

3.1 Crossing Detection:

Having seen the basic working of a radar system, we next explain how we use such a system to perform doorway tracking. The first step in any doorway tracking system is to detect when a person actually walks through the doorway. For this, we leverage the intuition that the total amount of energy reflected back on to the radar when a person is crossing the doorway is much higher than when no one walks through.

To capture this intuition, we first pass the captured *radar frame* through a 3-frame moving target indication (MTI) filter [56], to eliminate any unwanted static clutter. Next, we calculate the total power in each radar frame as the sum of the absolute reflected power at each distance in the frame. More formally, the power of the ith radar frame denoted as $RFP(i)$ (read as the radar frame power of the ith frame) is given by:
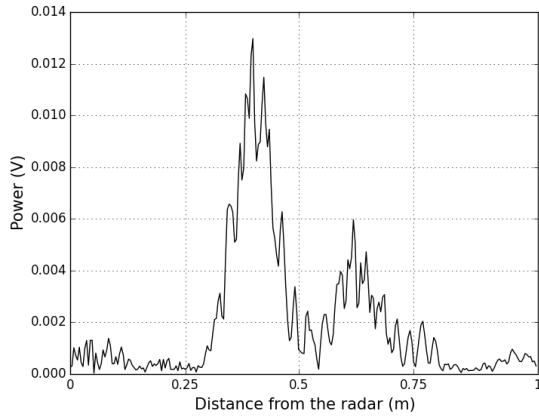
Fig. 1. An example of a radar frame: the radar measures the reflected power at various distances.
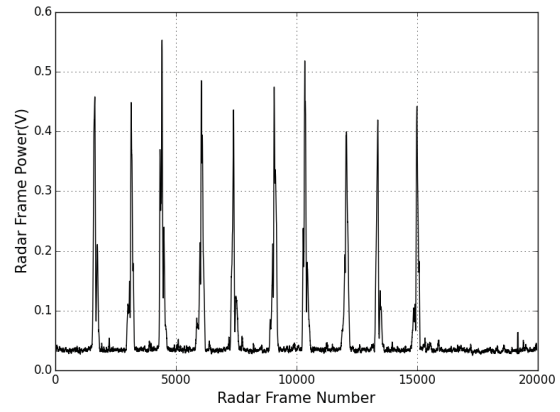

Fig. 2. The total energy reflected back on to the radar during a crossing (a peak in the figure) is much larger than that in the absence of a crossing.

$$RFP(i) = \sum_{d=0}^{N-1} |P_d(i)|, \tag{1}$$

where N is the total number of range bins, and $P_d(i)$ is the reflected power at the dth bin, in the ith radar frame.

We next filter the radar frame power values via a two-stage discrete FIR filter [25] to eliminate any impulse noise. Figure 2 shows the result of this filtering, for 10 doorway crossings. We can clearly see that the radar frame power during a crossing is much larger than during a crossing absence. Consequently, we detect a crossing when the total radar frame power exceeds a threshold Th_{cross} (set as 4 times the noise floor). This technique of using the radar frame power to detect a crossing is only the first step towards crossing detection. A second filtering step is needed to eliminate false positives - i.e. to filter out those non-crossing events (events where no one is crossing the doorway) that have a high radar frame power because of transient device noise.

We eliminate these false positive (FP) crossing events by leveraging on the intuition that since there is no one crossing the doorway, the height estimates (or the lack of it) during a potential crossing event, can be used to filter out FP events. We next explain how these height estimates are obtained. With *FormaTrack* mounted atop the doorway, the head is the dominant reflector during a doorway crossing. Consequently, the distance to the head at radar frame t corresponds to the range-bin with the maximum reflected power. Formally,

$$Distance\ to\ Head(t) = \arg\max_{d} |P_d(t)|. \tag{2}$$

Figure 3a shows the distance to the head from the radar during a crossing and non-crossing event (FP). We observe that the distance to head follows the expected V-shaped pattern during a crossing, but occupies the noise floor during a non-crossing. Hence by looking at the range of the distance to head estimates in a small window (we use 100 frames) around the actual doorway crossing frame (obtained in Section 3.2), we can filter out false
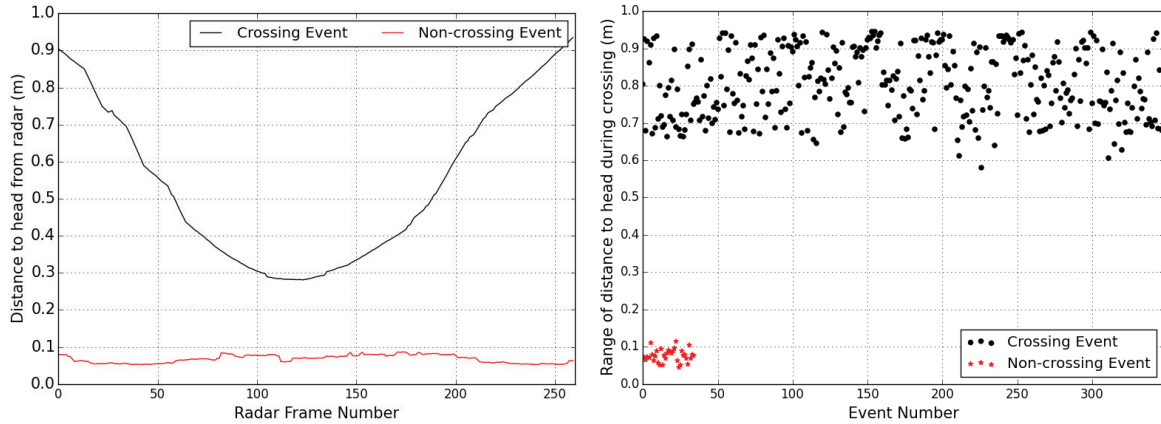
(a) The distance to the head from the radar follows a V-shaped pattern during a crossing, but occupies the noise-floor during a non-crossing event.

(b) The range of the distance to head estimates during a crossing and non-crossing event are vastly different.

Fig. 3. The distance to head estimate from *FormaTrack* can be used to filter out false positives.

positives[1]. Figure 3b shows this range estimate calculated for 350 doorway crossing events of a person and for all FP events detected in one of the days of our study (35 non crossing events passed the radar frame power test). We can see a clear separation between true positive and false positive events. We consider an event to be a false positive if it has a range estimate less than a threshold Th_{head_range} (0.24m; determined to be 50% of the 1st percentile of 900 crossing events collected during a testing period).

3.2 Direction Estimation

Having detected a crossing event, the next step of a doorway tracking system is to determine the direction of target motion (i.e. from which room to which room did the person move). For this, we leverage two key factors: (i) the principle of Doppler to identify when the person is exactly at the doorway, and (ii) tilting the radar towards one of the rooms so that the total reflected power from one side of the doorway is greater than the other, during a crossing. From classical physics, a target moving towards the receiver of a transmission causes a positive Doppler shift at the receiver, while a target moving away from the receiver induces a negative Doppler shift. Consequently, we determine the direction of motion by identifying the point of transition from positive to negative Doppler shift, and then comparing the total radar frame power in a small window around this point of Doppler transition. Since the radar is tilted towards one of the rooms, by comparing the reflected power on both sides of the Doppler transition point, we can determine the direction of motion. Given this overview, we next describe the details of our direction estimation - i.e. how we go from the radar frames described in Section 3.1 to a direction estimate.

Our first step in Doppler-based direction estimation is to convert a sequence of radar frames into a *range-Doppler* matrix that shows the Doppler shift at varying distances from the radar. As shown in Figure 4, this is done by horizontally stacking $K(= 32)$ consecutive radar frames and performing a K-point FFT over each of the $N(= 256)$ range bins in a frame [56]. The outcome of this transformation is the range-Doppler matrix having N

---

[1]On looking at a sufficiently large window around the doorway crossing radar frame for a crossing event, the distance to head measure actually takes an M-shape - i.e. rising from the noise floor when the person comes in the vicinity of the device, and dropping to the noise floor when the person goes outside its vicinity. The noise floor is due to a low-power static path that remains after MTI filtering.
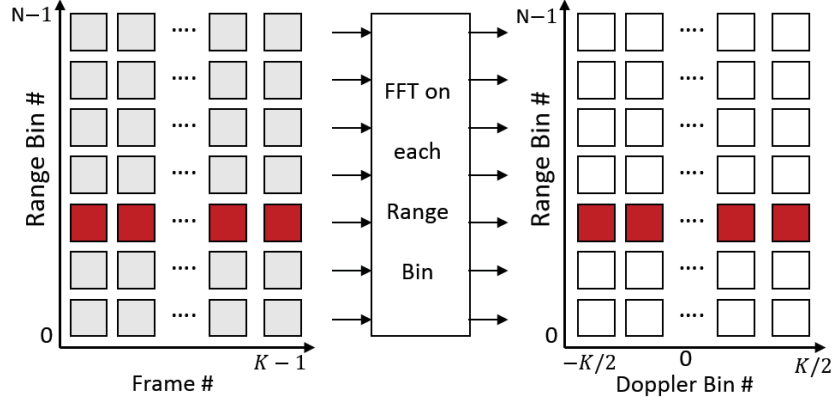
Fig. 4. The range-Doppler matrix is formed by aggregating K radar frames, and then performing an FFT over each range-bin.

rows and K columns. Each row of the range-Doppler matrix corresponds to a range bin, while each column of the range-Doppler matrix corresponds to a Doppler-shift bin. The size of each Doppler bin equals *Radar Frame Rate/K* Hz [64], where *Radar Frame Rate* is the total number of emitted pulses (radar frames) per second. For example, with a *Radar Frame Rate* of 170 Hz and a horizontal stacking of $K = 32$ radar frames, each Doppler bin has a resolution of 5.31 Hz.

The left half of the range-Doppler matrix corresponds to negative Doppler frequencies, and the right half corresponds to positive Doppler shifts. Each cell (i, j) in the matrix represents the Doppler power for the ith range bin and j^{th} frequency bin. Essentially, the range-Doppler matrix captures the amount of frequency shift at each distance. Figure 5 shows an example of a range-Doppler matrix captured by *FormaTrack* when a person is approaching the doorway. As the person approaches the doorway, she causes a positive Doppler shift, resulting in higher power in the right half of the matrix.

The Doppler shift caused by a target moving at velocity v at an angle θ relative to the receiver, is given by [20]

$$\Delta f = \frac{2 * f_c * v * cos\theta}{c} \tag{3}$$

where f_c is the transmitter's center frequency and c is the speed of light in the transmission medium.

With a radar frame rate of F frames per second, the X-dimension of the range-Doppler matrix spans from $-F/2$ Hz to $+F/2$ Hz. In other words, we can measure Doppler shifts induced by targets up to $F/2$ Hz. At higher speeds, the frequency "wraps around" the edge of the matrix, leading to *Doppler aliasing* [49]. Given our radar center frequency of 7.2 GHz and a radar frame rate of 170 Hz, from Equation 3, we can measure targets moving at speeds up to 1.77 ms^{-1}. This is beyond the average human walking speed of 1.2 to 1.3 ms^{-1} [48].

For a crossing detection frame t_{cross}, identified in Section 3.1, we compute RD_{cross}, the set of range-Doppler matrices for w_{rd} frames around t_{cross}. More formally,

$$RD_{cross} = \{RD(t, d, f) \mid t_{cross} - w_{rd} \leq t \leq t_{cross} + w_{rd}\}. \tag{4}$$

In our system, we use $w_{rd} = 200$ frames (translating to around 1.2 s). To capture the intuition that a person causes positive Doppler shift while approaching the doorway, and negative while exiting it, we transform RD_{cross} into the *dominant Doppler matrix DD (t,d)* - a matrix that contains the Doppler frequency with the highest reflected power at every distance over time. We perform this transformation by calculating the dominant Doppler
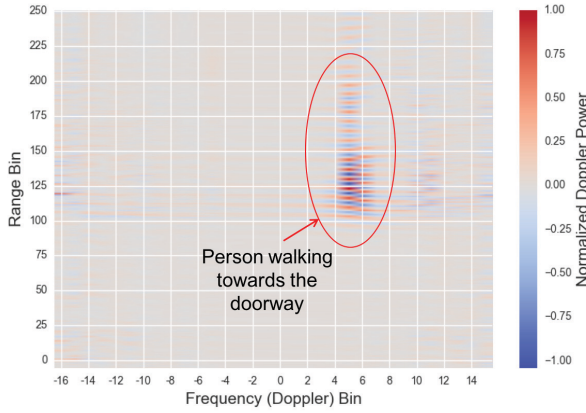
Fig. 5. Range Doppler Matrix: As the target moves towards the radar, she induces a positive Doppler shift. This is seen by higher power in the right half of the matrix.
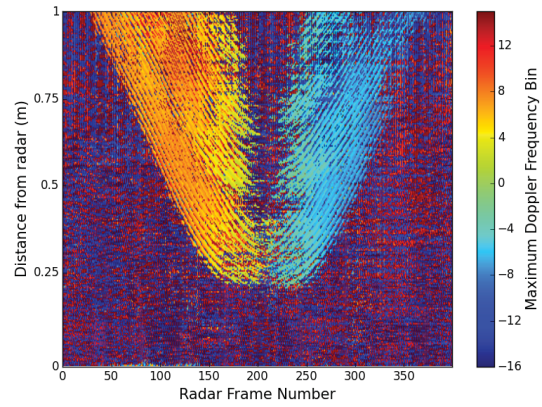

Fig. 6. [Figure Best Viewed in Color] Doppler during a doorway crossing: A person causes a positive Doppler shift while walking towards the doorway, and a negative Doppler shift while walking away.

frequency for every range-bin, for each range-Doppler matrix in RD_{cross}. Said more formally,

$$DD(t, d) = \arg \max_{f} \ RD(t, d, f), \text{ where } RD(t, d, f) \in RD_{cross}. \tag{5}$$

In other words, the relationship between $RD(t, d, f)$ and $DD(t, d)$ is as follows: the former captures at each radar frame t, the amount of frequency shift at each distance, while the latter takes it one step further and gives the frequency with the largest reflected power at each distance, over multiple radar frames. Figure 6 shows an example of the dominant Doppler matrix *DD (t,d)* during a crossing. Here, we can clearly see positive Doppler frequencies dominating when a person approaches the doorway, and negative frequencies dominating during exit. Our next step is to identify the Doppler transition point, which corresponds to when the person is actually at the doorway.

For this we define another term called the *Approach Away Power (AA)*. This term weights the *sign* of the dominant Doppler (+1 for positive Doppler and -1 for negative Doppler) by the corresponding Doppler Power, for each range bin in the dominant Doppler Matrix. Formally,

$$AA(t) = \sum_{d=0}^{N-1} sign(f) * RD(t, d, f), \tag{6}$$

$$\text{where } f = DD \ (t, d), \ sign(f) = \begin{cases} +1, f > 0 \ (Positive \ Doppler) \\ -1, f < 0 \ (Negative \ Doppler) \end{cases}. \tag{7}$$

Intuitively, *AA* captures the cumulative dominant Doppler power over the covered range of radar, factoring in the direction of movement. Via this measure, a person approaching the doorway causes positive *Approach Away Power*, while a person exiting the doorway causes negative *Approach Away Power*. Hence, if we calculate *AA (t)* over the entire crossing window, then capturing the zero-crossing would give us the Doppler transition point. Figure 7 shows the normalized *Approach Away Power* during a doorway crossing. We can clearly see our zero
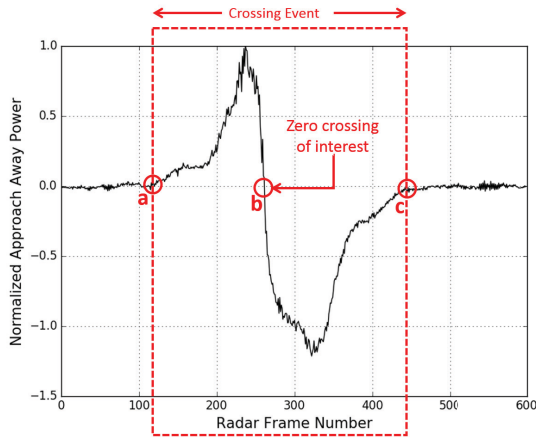
Fig. 7. The *Approach Away* (AA) Power is used to identify the point of transition from positive Doppler to negative Doppler. There is positive AA power when the person approaches the doorway, and negative AA power when the person leaves. Identifying the zero-crossing tells us when the person crosses the sensor.
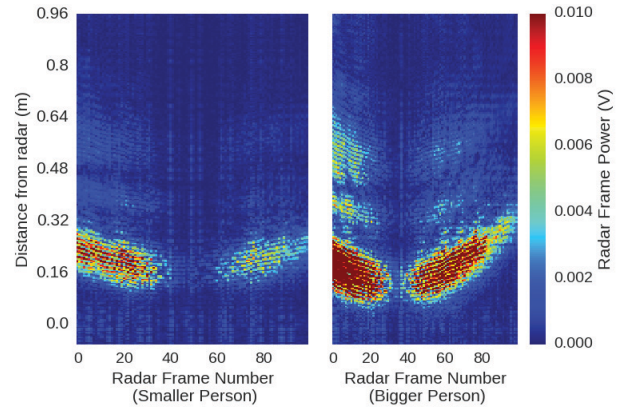
Fig. 8. Two people of similar height reflect very differently as they walk through the doorway. The bigger person reflects back a stronger signal to *FormaTrack*.

crossing of interest around frame 240. However, we also notice other zero-crossings due to noise cropping up (i.e. the *Approach Away Power* oscillates about 0 during a non-crossing).

To identify the zero-crossing of interest, we use the observation that the zero-crossing during a doorway crossing is much more widely spaced than during a non-crossing (i.e. the frequency of oscillations about 0 during a non-crossing is much higher than during a crossing). Hence to identify our zero-crossing of interest (corresponding to the person being at the doorway), we first compute the widest zero-crossing interval *(zc1, zc2)*. In Figure 7, this could correspond to the *(a,b)* interval or *(b,c)* interval (i.e. the approach or the exit part of a doorway crossing). For the chosen widest zero-crossing interval *(zc1, zc2)*, we next check the *Approach Away Power* values in the interval (zc1, zc2). If they are all positive then *(zc1,zc2)* corresponds to a doorway entry, and we choose the right-end (zc2) as our zero-crossing of interest. On the other hand, if they are all negative, then *(zc1,zc2)* corresponds to a doorway exit, and we choose the left-end (zc1) as our zero-crossing of interest. As mentioned earlier, this computed zero-crossing of interest (referred to as $ZC_{interest}$) corresponds to the person actually being at the doorway.

Now that we know the exact point of doorway crossing, we can determine direction. As mentioned earlier, *FormaTrack* is tilted towards one of the rooms adjoining the doorway. This tilt creates asymmetric reflected power during a crossing; reflected power from one room is always greater than the other. As a result, by comparing the total radar frame power in a small window of $w_{dir}(= 200)$ frames on either side of the $ZC_{interest}$, we can determine direction. If we define motion from tilted side to the non-tilted side as being *IN*, and vice versa as being *OUT*, then we can formally define our direction estimate as:

$$Direction = \begin{cases} \text{IN, if } \sum_{t1=ZC_{interest}-w_{dir}}^{ZC_{interest}} RFP(t1) \; > \; \sum_{t2=ZC_{interest}}^{ZC_{interest}+w_{dir}} RFP(t2) \\ \\ \text{OUT, otherwise} \end{cases}. \tag{8}$$

3.3 Body Shape Sensing

The third piece of any doorway tracking system is to infer the identity of the crossing individual. For this, *FormaTrack* uses knowledge of when the person is exactly at the doorway (Section 3.2), and the distance to head estimate (Section 3.1), to sense the shape of the person.

According to the radar range equation [56], the total power reflected back on to the radar depends not just on the distance to the target, but also the radar cross section. This radar cross section is a property of the target [43] and is a measure of the target's ability to reflect signals back on to the radar. In the case of a person walking through a doorway, this becomes a function of the body shape of the person. Moreover, as a bigger person occupies more area, he would also reflect power at more distances than the smaller person. Figure 8 shows an example of two people of similar height but very different shapes as they cross the doorway. We see that the bigger person (on the right) reflects back more energy towards the sensor than the smaller person (on the left). *FormaTrack* leverages this intuition and computes a *reflection profile (RP)* to sense the body shape when the person is at the doorway.

FormaTrack computes the *reflection profile* by capturing the intuition that two differently shaped people will reflect different power at different distances while at the doorway. To compute the *reflection profile*, *FormaTrack* first aggregates a small number of $w_{id}(= 50)$ frames around the doorway crossing point ($ZC_{interest}$). Next for each frame t, *FormaTrack* first computes the distance to the head d_{head} (*Distance to Head (t)* in Section 3.1). This distance estimate, which corresponds to the start of the person's body in a frame, is used as an *anchor point*. Next, we collect the power values at the remaining distances relative to this anchor point, and integrate over the $2*w_{id}$ crossing frames to obtain the *reflection profile* as the following:

$$RP = \left\{ \sum_{t=ZC_{interest}-w_{id}}^{ZC_{interest}+w_{id}} P_d(t) \right\}, \text{ for } d \in [d_{head}, N-1]. \tag{9}$$

We then pass this computed reflection profile of the person (a measure of how the person's body reflects while at the doorway), which is a vector of aggregated reflected power at various distances, and the distance to head estimate at $ZC_{interest}$, to a support vector machine classifier with the RBF kernel (C=0.5, gamma=100) to infer the identity of the person.

4 EXPERIMENTAL SETUP

To test our hypothesis, we mounted the Salsa Ancho kit [3] atop the doorway of a home, as shown in Figure 9a. The Salsa Ancho kit uses the Novelda Xethru X2 ultra-wideband impulse radar transceiver chip [2], which operates at a center frequency of about 7.2 GHz, and provides a range resolution of 4mm (i.e. a sampling rate of 39 GS/s). The kit outputs raw radar baseband frames as shown in Figure 1, which are then transferred via a BeagleBone Black to a host computer for analysis. We used a 6dBi sinuous directional antenna on the kit which provided a horizontal and vertical beamwidth of 85°and 65°respectively. We also modified the X2 radar registers to output an average of 170 frames per second to prevent any Doppler aliasing. There are no health concerns
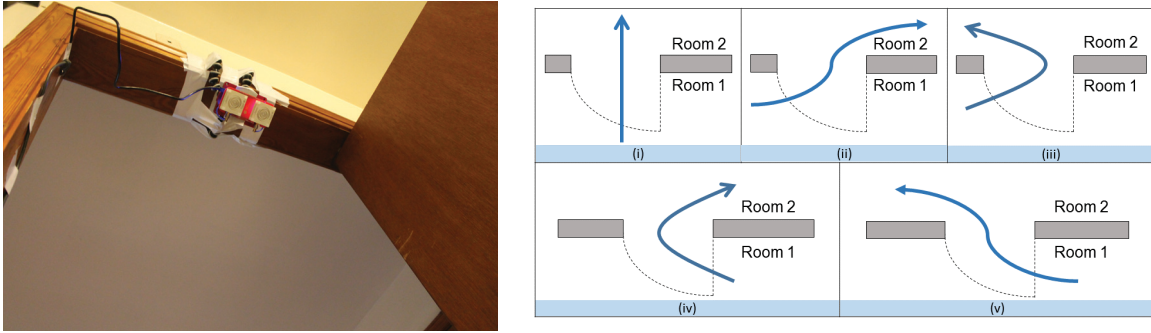
(a) The Xethru radar is mounted atop the doorway tilted towards one side.

(b) Participants were asked to walk through the instrumented doorway in 5 different ways.

Fig. 9. Experimental Setup.

Participant #	P1	P2	P3	P4	P5	P6	P7	P8
Height (cm)	161	165	167	169	170	172	175	183
Weight (kg)	60.8	58.1	73.5	77.1	78.0	56.7	97.9	68.0

Table 1. The height and weight of the 8 participants who walked through an instrumented doorway for 7 days generating 2800 doorway crossing events

associated with *FormaTrack* as the average *transmit power* from the *Salsa Ancho kit* is -13dBm (50μW) at 7GHz[2]. This is over 450 times less than the regulations set by the International Commission on Non-Ionizing Radiation Protection (ICNIRP)[3][1]. For comparison, the maximum FCC permitted transmit power for an indoor 5GHz WiFi access point is 1W[4].

We asked 8 participants to walk for 7 days through the instrumented doorway (i.e. one session of 25 back and forth crossings per day for 7 days to capture variations due to speed and clothing). The height and weight of these participants are shown in Table 1. On each day, as shown in Figure 9b, they were asked to walk a total of 25 times - 5 times back and forth through the doorway in each of the following ways : (i) straight from one room to the other (henceforth referred to as room1 and room2 respectively), (ii) anywhere from the left-side of room1 to the right-side of room2, (iii) anywhere from the left-side of room1 to the left-side of room2, (iv) anywhere from the right-side of room1 to the right-side of room2, and (v) anywhere from the right-side of room1 to the left-side of room2. No restrictions were imposed on the type of clothing the participants wore, or the time of experiment. Our study had participants walking shortly after waking up, after a gym workout, on the way to/from work, during an illness etc. The crossings of the participants were recorded by a video camera, which were then manually analyzed. In all, we collected a total of 2800 doorway crossings.

We evaluate *FormaTrack* via four metrics:

- *Recall*: The fraction of actual doorway crossings that were correctly detected by *FormaTrack*.

[2]https://www.xethru.com/chips-salsa-uwb-radar-development-kit.html/
[3]http://www.icnirp.org/
[4]https://www.ecfr.gov/cgi-bin/text-idx?SID=eed706a2c49fd9271106c3228b0615f3&mc=true&node=pt47.1.15&rgn=div5

Metric	Accuracy (%)
Precision	100
Recall	99.86
Direction Accuracy	99.17
Identity Accuracy	90.3 (95.3)

Table 2. *FormaTrack* achieves over 90% accuracy in all metrics of interest for 2800 doorway crossings. The identity accuracy increases from 90.3% with 1 day of training to 95.3% with 6 days of training.

$$Recall = \frac{\text{\# Correctly detected crossings by } FormaTrack}{\text{\# Ground truth crossings}} \qquad (10)$$

- *Precision*: Amongst the doorway crossings detected by *FormaTrack*, the fraction that actually occurred.

$$Precision = \frac{\text{\# Correctly detected crossings by } FormaTrack}{\text{Total \# crossings detected by } FormaTrack} \qquad (11)$$

- *Direction Accuracy*: The fraction of correctly detected doorway crossings having the correct direction.

$$Direction\ Accuracy = \frac{\text{\# Correct direction crossings by } FormaTrack}{\text{\# Correct crossings detected by } FormaTrack} \qquad (12)$$

- *Identity Accuracy*: The fraction of correctly detected doorway crossings classified to be the right person.

$$Identity\ Accuracy = \frac{\text{\# Correctly classified crossing person's identity by } FormaTrack}{\text{\# Correct crossings detected by } FormaTrack} \qquad (13)$$

In this study, we did not ask multiple people to walk through the doorway one behind the other in a platoon fashion. This is because such platooning is not a common scenario in homes (for e.g., in a study, the median time difference between two different people walking through the same doorway in an 8-room home was 10 minutes [27]).

5 EVALUATION

5.1 System Accuracy

Of the 2800 crossings collected, only 4 of them were missed due to the conservative threshold (Th_{cross}), and there were no spurious crossings detected. Among the 2796 crossings correctly detected, 8 of them had an incorrect direction. We evaluated the identity accuracy by testing all possible combinations of 2 people from the 8 participants (8 choose 2), for a total of 28 combinations. For each combination, each person is trained on all crossings collected from a day (one day at a time), and tested on the remaining 6 days. This gives us an overall precision, recall, direction and identity accuracy of 100%, 99.86%, 99.7%, 90.3% respectively. Table 2 shows the average accuracy of *FormaTrack*. Using more training data through 7-fold cross validation (i.e. training on 6 days and testing on 1) would give 95.3% identification accuracy, as illustrated in Figure 15. The identity accuracy of *FormaTrack* increases with the number of days used for training, and we evaluate this in Section 5.6.

5.2 Effect of False Positives

We next test *FormaTrack* for false positives. We do this by collecting 36 hours of data (3 hours for 12 days) from the instrumented doorway on an empty room, where no person crossed the doorway. We compare our *distance*
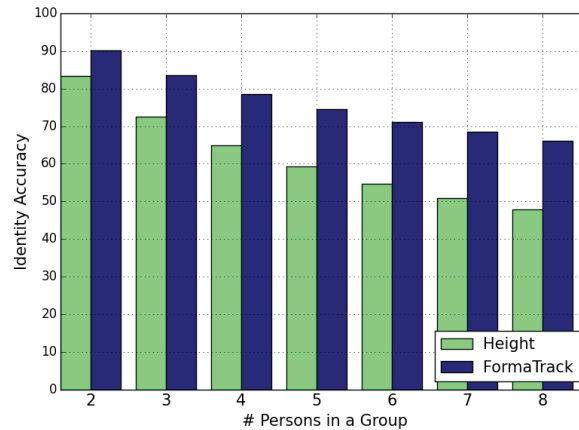
Fig. 10. *FormaTrack's* reflection-profile based technique consistently out-performs the height biometric baseline. *FormaTrack* is also more robust - i.e. accuracy decreases more slowly with people. *FormaTrack* achieves nearly 80% identity accuracy even while sensing a group size of 4.

filter algorithm against the no filtering technique (i.e. simply relying on radar frame power alone to detect a crossing). In the absence of any filtering, 152 energy peaks were observed over the 36-hour period. This is due to transient noises on the radar device, as mentioned earlier. However, our *distance filter* technique filtered them all out, resulting in 0 false positives over the study period.

5.3 Effect of Number of People

We next evaluate how *FormaTrack's reflection profile technique* performs as we start increasing the number of people being tracked. For an *N*-person group, we calculate the average *identity accuracy* by generating all possible *N*-person combinations from the 8 participants. Each combination is trained on a day (one day at a time), and tested on the remaining 6 days. We compare our approach of using *reflection profile* against a baseline algorithm that used *heights* alone. The height estimate is obtained as the distance to the dominant reflector (i.e. the head), as explained in Section 3.1, when the person is at the doorway (i.e. the radar frame corresponding to the Doppler sign change).

From Figure 10, we notice that *FormaTrack* not just consistently out-performs the *height* alone approach, but is also more robust, i.e. the accuracy drops more slowly as more people are added to the system. We see that for 2, 4 and 8 people, the reflection profile technique is 7%, 14% and 18% better than the height-alone approach. As expected, as the number of people being sensed increases, the *identity accuracy* starts to drop. However, *FormaTrack* still manages to achieve nearly 80% accuracy while sensing a 4-person group.

5.4 System-level Tracking Accuracy

Individual doorway accuracy numbers do not directly translate to a room-level tracking accuracy of occupants [29, 34]. This is because errors made on a single doorway can typically be corrected by subsequent observations made on other doorways of a home. However, for cost reasons, we were unable to measure room-level tracking accuracy via a full-home deployment. As an alternative, we perform a tracking simulation by considering 15 different floor plans (obtained from Google Images) of homes varying from 3 to 9 rooms. For each home under test, we consider 2 to 4 inhabitants (limited by the computational complexity of the tracking algorithm [34, 35, 55]).
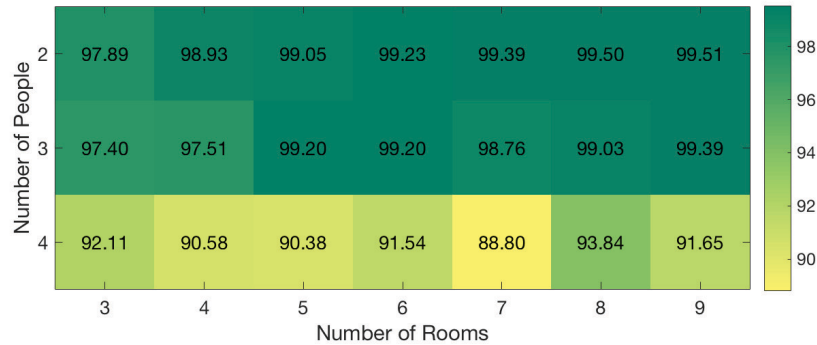
Fig. 11. Whole home tracking simulation of *FormaTrack* with 15 different floorplans obtained from Google Images and various number of people: *FormaTrack* can achieve an average room-level tracking accuracy of over 99%, 98% and 91% for 2, 3 and 4 people respectively.

These inhabitants were chosen at random from our pool of 8 participants. For each person we model their room transitions via a uniform distribution, and model their dwell time (in a room) based on 6 days of in-situ data [34]. For each simulated crossing of a person, we pick a random crossing event performed by the corresponding participant on our instrumented doorway. We repeat the entire process 10 times, and obtain a total of 223,982 simulated crossings. We pass these crossings through *TransTrack* [34], a multiple-hypothesis tracking (MHT) [55] algorithm for doorways, and measure the *transition accuracy* [34] (F-score of precision and recall of crossings).

Figure 11 shows the average transition accuracy for different number of rooms and people. We see that even though the individual identity accuracies for 2, 3 and 4 person groups (from Figure 10) are 90%, 84% and 79% respectively, they translate to an average of over 99%, 98% and 91% of transition accuracy, respectively. This increase is because the tracking algorithm leverages future doorway crossing events to help rectify mistakes in prior crossing events.

5.5 Energy Analysis

We next explore the potential of a low-power variant of *FormaTrack*. In particular, we explore the following options to save energy:

- Using a motion sensor to duty-cycle the radar
- Lowering the frame rate of the radar (down-sampling in Slow Time)
- Lowering the sampling rate of the radar (down-sampling in Fast Time)

5.5.1 Duty-cycling via motion sensor: There are many hours in the day when no occupant of a home makes any doorway transition (for e.g. when the residents are asleep or out at work). Similarly, there are many rooms in the home which are very infrequently visited (for e.g. the bathroom is occupied for around 6% of the day [27]). Under these circumstances, the *FormaTrack* device(s) can effectively be turned off for long periods of time to save energy. In order to perform such duty-cycling, we explore the possibility of using low-power motion sensors on the doorway to trigger the radar on.

For this, we analyze the data from one of our prior in situ studies [34]. The 6-day study contained 1756 doorway crossing events from 8 doors in a home. Each doorway was equipped with two PIR motion sensors facing each of the adjoining rooms. Figure 12 shows the *recall* (i.e. the fraction of crossings detectable by a system like *FormaTrack* that would be triggered by a motion sensor) compared against the fraction of time a system like
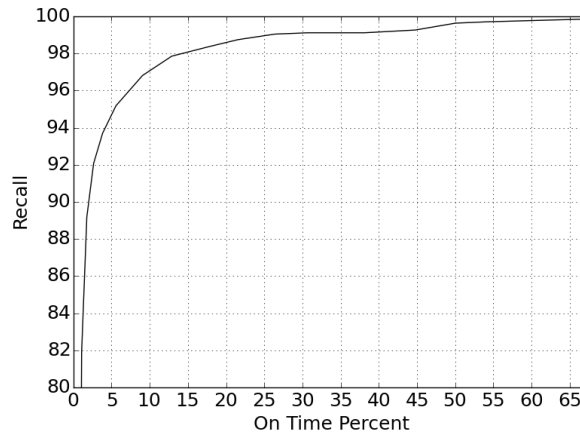
Fig. 12. By augmenting *FormaTrack* with a motion sensor, *FormaTrack* can achieve 99% recall by remaining ON for less than 30% of the time.

FormaTrack would be ON. We see that we can achieve 99% recall by keeping *FormaTrack* ON for just 30% of the day. (i.e. *FormaTrack* can be turned OFF for 70% off the time). In other words, if the *FormaTrack* system is turned off when nobody is moving in an adjacent room, its average energy consumption can be reduced by 70%. The typical power consumption of the Xethru X2 radar which *FormaTrack* uses is about 120mW[5]. This is comparable to the power consumption of the ultrasonic sensors used in Doorjamb (150mW) [29] or a thermal camera (150mW)[6], and far less than other alternatives such as a LIDAR (8-60W)[7] or a Kinect (15W) [29, 47]. Mohammadmoradi et al. [46] used an infrared array sensor that consumed an average power of just 15mW. However, unlike other aforementioned systems, they do not have a notion of identity (i.e. person-counting application). Furthermore, by being ON for just 30% of the time, the average power consumption of *FormaTrack* drops to just 46.6mW (36mW from duty-cycled radar and 10.6mW[8] from two PIR motion sensors). This saving can be further increased by using motion sensors with only a limited view (e.g. 1m) on either side of the doorway.

5.5.2 Down sampling in Slow Time: We next explore the possibility of lowering the frame rate of the radar to save energy (i.e. the radar can start to sleep between every pulse transmission). We do this by considering every N^{th} radar frame, and measuring our metrics of interest. As before, for identity accuracy, we consider all possible 2 person groups. Every group is trained one day at a time, and tested on the remaining 6 days. Figure 13 shows that we can down-sample by a factor of 2, and still achieve comparable performance to the non down-sampled case. At greater down-sampling factors, *Doppler aliasing* dominates, affecting the calculation of our zero-crossing of interest, resulting in direction and identity performance loss.

5.5.3 Down sampling in Fast Time: A third approach to save energy in *FormaTrack* is by lowering the sampling rate of the radar (resulting in lesser stress on the samplers of the radar). This effectively translates to a decrease in the range resolution of the radar. We simulate a fast-time down sampling rate of N, by averaging every N range bins of a radar frame. Figure 14 shows that similar to the slow-time case, we can down-sample by a factor

---

[5]https://www.xethru.com/community/resources/x2-datasheet.24/

[6]https://cdn.sparkfun.com/datasheets/Sensors/Infrared/FLIR_Lepton_Data_Brief.pdf

[7]http://velodynelidar.com/docs/datasheet/LiDAR%20Comparison%20chart_Rev-A_Web.pdf

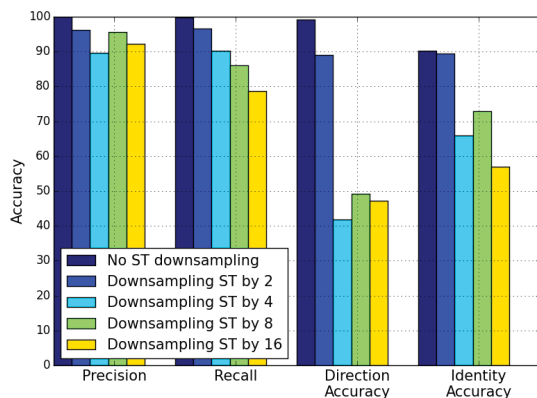[8]https://www.sparkfun.com/products/13285

Fig. 13. Down sampling in slow time (ST) for energy savings: The performance of *FormaTrack* remains comparable even after lowering the frame rate by a factor of 2. At even lower frame rates, *Doppler aliasing* impacts accuracy.
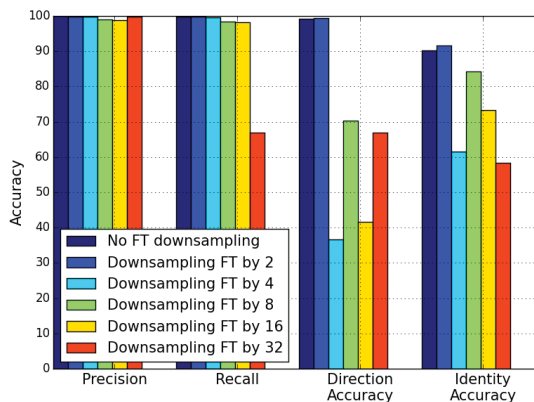
Fig. 14. Down sampling in fast time (FT) for energy savings: *FormaTrack's* performance remains unaffected even after lowering the sampling rate of the radar by a factor of 2.

of 2 with no performance loss. Even at greater down sampling rates (up to a factor of 16), the precision and recall remains over 98%. However, the direction and identity accuracies start to suffer. This is because the range-Doppler matrix, as seen in Figure 5 is not smooth (as the target is moving). An aggregation of range bins means we no longer see a clear transition from approach to away, thus affecting the determination of our desired zero-crossing. As a result, both direction and identity accuracies suffer.

We conclude that the greatest amount of energy savings can be observed by duty-cycling *FormaTrack* with an augmented motion sensor. Furthermore, even while ON, we see that lowering the frame rate and the sampling rate of the sensor by a factor of 2 does not cause any significant performance degradation. In the future, we plan to explore the option of building an energy-harvesting *FormaTrack* that stores energy for most of the day, and uses them during the active times of the day.

5.6 Effect of Training Size

Figure 15 shows the effect of the number of training samples (back and forth doorway crossings) on *identity accuracy* for all combinations of two people. We see that the identity accuracy increases as we increase the number of training samples with nearly 85% accuracy from just 10 training samples. This shows that *FormaTrack's* training would not be very cumbersome to the users. These training samples can potentially be collected without controlled walk-throughs, by using any events when the person is known to be home alone.

5.7 Effect of Objects

We next analyze the effect of carried objects on *FormaTrack's* reflection profile technique. For this, we asked five of the participants (P1, P2, P5, P6 and P7), to walk through the instrumented doorway, 25 times as mentioned in Section 4, but under different scenarios - (i) with a mobile phone near the ear (simulating talking on the phone), (ii) with a baseball cap, (iii) wearing a backpack with a laptop inside, (iv) carrying a 5lb dumbbell in each hand, (v) wearing flip-flops.

We calculate the identity accuracy for each scenario, by considering all possible 2-person groups involving the participant. For every group, we train the two people on each day (one day at a time). We test the object mounted
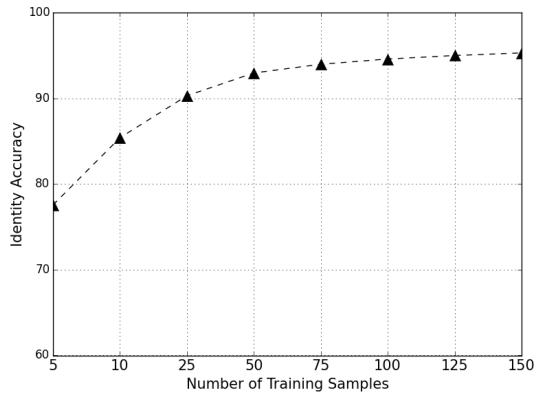
Fig. 15. Effect of training size on identity accuracy: the identity accuracy increases with the number of training samples. With just 10 back and forth crossings for training, *FormaTrack* can achieve an identity accuracy of over 85%.
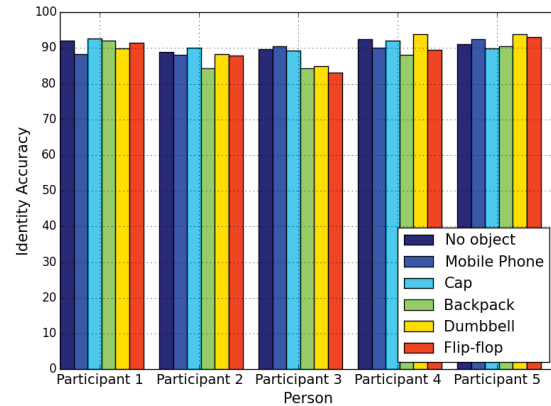

Fig. 16. The identity accuracy of *FormaTrack* remains comparable even in the presence of some commonly used objects.

person, with the object mounted walk, and test the non-object mounted participant with each of the remaining 6 days (one at a time).

From Figure 16, we see that the identity accuracy is not affected greatly by objects such as mobile phone, cap, backpack and dumbbells, because these objects only occupy a small additional area on the body, and hence do not affect the reflection profile greatly. We also noticed that in the presence of heels, for participants P1 and P7 (two participants comfortable with heels), the identity accuracy was 66% and 89% respectively. We believe this accuracy drop is because heels not just translate the height significantly but also changes the posture of the person while underneath the sensor. This issue can potentially be fixed via more training data.

6 LIMITATIONS AND FUTURE WORK

6.1 In-situ study

FormaTrack is tested on over 2800 crossings in different crossing angles, and a handful of scenarios in Section 5.7 to provide a proof-of-concept of sensing the shape of the person when the person is at the doorway. As a future work, we plan to do a more long-term in-situ study after establishing a feasible ground-truth collection method [27, 29].

6.2 Effect of Doors

While we do not explicitly consider door interactions in this paper, we hypothesize that *FormaTrack* can be made to handle these interactions by filtering out the effect of doors. For direction and identity estimation, since the door is taller than an average human, we can filter the door by only considering the dominant Doppler shifts and radar frame power at distances corresponding to the average human height range. An alternate way to address doors would be a data-driven approach that is trained on the door actions. We leave it as a future work to test *FormaTrack* in the presence of door movements.

6.3 Effect of walking up to the doorway and turning back

One potential limitation of *FormaTrack* is the case of a person walking all the way to the threshold of the door, turning around and walking back. This motion shows all signs of a crossing event - the radar frame power (*RFP*) would cross Th_{cross}, there would be positive and negative Doppler, and the *distance to head filter* would see the expected V-shaped pattern. We believe this can be filtered by measuring the total *RFP* before and after the *Doppler crossing*, and observing them to be similar. An alternate solution is to place a small antenna array that is triggered during a crossing. This array which measures the angle of arrival (-90°to +90°) of incident paths [57], can filter out such events by observing the angle of the reflected path to be on one-side of 0°.

6.4 Effect of running through the doorway

With drastic speed changes such as running through the doorway, *FormaTrack's* direction estimation algorithm can obtain the correct direction so long as the person's velocity is less than $1.77\text{ms}^{-1}\cos\theta$ (Section 3.2). However, the identity algorithm could be impacted as a clear shape signature might not be obtainable. This can potentially be handled by deriving the velocity from the Doppler shift, and using it as a feature or falling back to using height when the speed of the target is significantly different from what was trained. Since this is not a common occurrence, we leave the velocity incorporation as a future work.

7 CONCLUSION

In this paper, we present *FormaTrack*, a privacy-preserving radar-based doorway tracking system that estimates the room-location of people in a home. *FormaTrack* determines the direction of room transition of people via the Doppler effect, and infers the identity of people by sensing their shape. We evaluate the system on an instrumented doorway in a home using 2800 doorway crossings. Our results indicate that *FormaTrack* can achieve over 90% tracking accuracy even while tracking 4 people. We also show that by simply augmenting *FormaTrack* with a low-power motion sensor, *FormaTrack* can be heavily duty-cycled. Our analyses indicate that *FormaTrack* can then achieve nearly 99% recall by staying ON for just 30% of the day, showing positive signs for a future low-power energy harvesting *FormaTrack*.

8 ACKNOWLEDGEMENTS

REFERENCES

[1] 2017. Novelda Xethru Sensor Emissions. https://www.xethru.com/community/resources/xethru-sensor-emissions.53/. (2017).
[2] 2017. Novelda Xethru X2 SoC. https://www.xethru.com/x2-uwb-radar-chip.html/. (May 2017).
[3] 2017. Salsa Ancho Kit. https://www.xethru.com/chips-salsa-uwb-radar-development-kit.html/. (May 2017).
[4] 2017. VICON Motion Capture System. http://www.vicon.com. (May 2017).
[5] 2017. XSens 3D Motion Capture. https://www.xsens.com. (May 2017).
[6] 2017. Zebra Sports Solutions. https://www.zebra.com/us/en.html. (May 2017).
[7] M.D. Addlesee, A. Jones, F. Livesey, and F. Samaria. 1997. The ORL active floor [sensor system]. *Personal Communications, IEEE* 4, 5 (1997), 35–41.
[8] Fadel Adib, Chen-Yu Hsu, Hongzi Mao, Dina Katabi, and Frédo Durand. 2015. Capturing the human figure through a wall. *ACM Transactions on Graphics (TOG)* 34, 6 (2015), 219.
[9] Fadel Adib, Zachary Kabelac, and Dina Katabi. 2015. Multi-Person Localization via RF Body Reflections.. In *NSDI*. 279–292.

[10] Fadel Adib, Hongzi Mao, Zachary Kabelac, Dina Katabi, and Robert C Miller. 2015. Smart homes that monitor breathing and heart rate. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. ACM, 837–846.

[11] Moeness G Amin, Yimin D Zhang, Fauzia Ahmad, and KC Dominic Ho. 2016. Radar signal processing for elderly fall detection: The future for in-home monitoring. *IEEE Signal Processing Magazine* 33, 2 (2016), 71–80.

[12] Chuankai An, Tianxing Li, Zhao Tian, Andrew T Campbell, and Xia Zhou. 2015. Visible light knows who you are. In *Proceedings of the 2nd International Workshop on Visible Light Communications Systems*. ACM, 39–44.

[13] Roger Appleby and Rupert N Anderton. 2007. Millimeter-wave and submillimeter-wave imaging for security and surveillance. *Proc. IEEE* 95, 8 (2007), 1683–1690.

[14] Md Tanvir Islam Aumi, Sidhant Gupta, Mayank Goel, Eric Larson, and Shwetak Patel. 2013. DopLink: using the doppler effect for multi-device interaction. In *Proceedings of Ubicomp 2013*. ACM, 583–586.

[15] Paramvir Bahl and Venkata N Padmanabhan. 2000. RADAR: An in-building RF-based user location and tracking system. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, Vol. 2. Ieee, 775–784.

[16] Alexandru Bălan and Michael Black. 2008. The naked truth: Estimating body shape under clothing. *Computer Vision–ECCV 2008* (2008), 15–29.

[17] Francisco Bernardo, Nicholas Arner, and Paul Batchelor. 2017. O Soli Mio: Exploring Millimeter Wave Radar for Musical Interaction. (2017).

[18] Imed Bouchrika, Michaela Goffredo, John Carter, and Mark Nixon. 2011. On using gait in forensic biometrics. *Journal of forensic sciences* 56, 4 (2011), 882–889.

[19] Kevin W Bowyer, Karen Hollingsworth, and Patrick J Flynn. 2008. Image understanding for iris biometrics: A survey. *Computer vision and image understanding* 110, 2 (2008), 281–307.

[20] Walter G Carrara, Ron S Goodman, and Ronald M Majewski. 2007. *Spotlight synthetic aperture radar.*

[21] Yin Chen, Dimitrios Lymberopoulos, Jie Liu, and Bodhi Priyantha. 2012. FM-based indoor localization. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*. ACM, 169–182.

[22] Robert T Collins, Ralph Gross, and Jianbo Shi. 2002. Silhouette-based human identification from body shape and gait. In *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on*. IEEE, 366–371.

[23] Giorgio Conte, Massimo De Marchi, Alessandro Antonio Nacci, Vincenzo Rana, and Donatella Sciuto. 2014. BlueSentinel: a first approach using iBeacon for an energy efficient occupancy detection system.. In *BuildSys@ SenSys*. 11–19.

[24] Ken B Cooper, Robert J Dengler, Nuria Llombart, Tomas Bryllert, Goutam Chattopadhyay, Erich Schlecht, John Gill, Choonsup Lee, Anders Skalare, Imran Mehdi, et al. 2008. Penetrating 3-D imaging at 4-and 25-m range using a submillimeter-wave radar. *IEEE Transactions on Microwave Theory and Techniques* 56, 12 (2008), 2771–2778.

[25] Douglas F Elliott. 2013. *Handbook of digital signal processing: engineering applications*. Academic press.

[26] Jonathon L Geisheimer, William S Marshall, and Eugene Greneker. 2001. A continuous-wave (CW) radar for gait analysis. In *Signals, Systems and Computers, 2001. Conference Record of the Thirty-Fifth Asilomar Conference on*, Vol. 1. IEEE, 834–838.

[27] Erin Griffiths, Avinash Kalyanaraman, Juhi Ranjan, and Kamin Whitehouse. 2017. An Empirical Design Space Analysis of Doorway Tracking Systems for Real World Environments. *ACM Transactions on Sensor Networks (TOSN)* 13, 4 (2017).

[28] Sidhant Gupta, Daniel Morris, Shwetak Patel, and Desney Tan. 2012. Soundwave: using the doppler effect to sense gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1911–1914.

[29] Timothy W Hnat, Erin Griffiths, Ray Dawson, and Kamin Whitehouse. 2012. Doorjamb: unobtrusive room-level tracking of people in homes using doorway sensors. In *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*. ACM, 309–322.

[30] Timothy W Hnat, Vijay Srinivasan, Jiakang Lu, Tamim I Sookoor, Raymond Dawson, John Stankovic, and Kamin Whitehouse. 2011. The hitchhiker's guide to successful residential sensing deployments. In *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems*. ACM, 232–245.

[31] Chen-Yu Hsu, Yuchen Liu, Zachary Kabelac, Rumen Hristov, Dina Katabi, and Christine Liu. 2017. Extracting Gait Velocity and Stride Length from Surrounding Radio Signals. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2116–2126.

[32] Amit Kale, AN Rajagopalan, Naresh Cuntoor, and Volker Kruger. 2002. Gait-based recognition of humans using continuous HMMs. In *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on*. IEEE, 336–341.

[33] Kaustubh Kalgaonkar and Bhiksha Raj. 2007. Acoustic doppler sonar for gait recoginalation. In *Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conference on*. IEEE, 27–32.

[34] Avinash Kalyanaraman, Erin Griffiths, and Kamin Whitehouse. 2016. TransTrack: Tracking Multiple Targets by Sensing Their Zone Transitions. In *Distributed Computing in Sensor Systems (DCOSS), 2016 International Conference on*. IEEE, 59–66.

[35] Avinash Kalyanaraman and Kamin Whitehouse. 2015. An event-based data fusion algorithm for smart cities. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*. ACM, 1575–1582.

[36] Nacer Khalil, Driss Benhaddou, Omprakash Gnawali, and Jaspal Subhlok. 2016. Nonintrusive Occupant Identification by Sensing Body Shape and Movement. In *Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments*. ACM, 1–10.

[37] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. 2015. Spotfi: Decimeter level localization using wifi. In *ACM SIGCOMM Computer Communication Review*, Vol. 45. ACM, 269–282.

[38] Swarun Kumar, Stephanie Gil, Dina Katabi, and Daniela Rus. 2014. Accurate indoor localization with zero start-up cost. In *Proceedings of the 20th annual international conference on Mobile computing and networking*. ACM, 483–494.

[39] Seungwoo Lee, Daye Ahn, Sukjun Lee, Rhan Ha, and Hojung Cha. 2014. Personalized energy auditor: Estimating personal electricity usage. In *Pervasive Computing and Communications (PerCom), 2014 IEEE International Conference on*. IEEE, 44–49.

[40] Tianxing Li, Chuankai An, Zhao Tian, Andrew T Campbell, and Xia Zhou. 2015. Human sensing using visible light communication. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 331–344.

[41] Xiang Li, Shengjie Li, Daqing Zhang, Jie Xiong, Yasha Wang, and Hong Mei. 2016. Dynamic-music: accurate device-free indoor localization. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 196–207.

[42] Wen-Hau Liau, Chao-Lin Wu, and Li-Chen Fu. 2008. Inhabitants tracking system in a cluttered home environment via floor load sensors. *IEEE Transactions on Automation Science and Engineering* 5, 1 (2008), 10–20.

[43] Jaime Lien, Nicholas Gillian, M Emre Karagozler, Patrick Amihood, Carsten Schwesig, Erik Olson, Hakim Raja, and Ivan Poupyrev. 2016. Soli: Ubiquitous gesture sensing with millimeter wave radar. *ACM Transactions on Graphics (TOG)* 35, 4 (2016), 142.

[44] Zongyi Liu and Sudeep Sarkar. 2007. Outdoor recognition at a distance by fusing gait and face. *Image and Vision Computing* 25, 6 (2007), 817–832.

[45] Dimitrios Lymberopoulos, Jie Liu, Xue Yang, Romit Roy Choudhury, Vlado Handziski, and Souvik Sen. 2015. A realistic evaluation and comparison of indoor location technologies: Experiences and lessons learned. In *Proceedings of the 14th international conference on information processing in sensor networks*. ACM, 178–189.

[46] Hessam Mohammadmoradi, Sirajum Munir, Omprakash Gnawali, and Charles Shelton. 2017. Measuring People-Flow Through Doorways using Easy-to-Install IR Array Sensors. In *Distributed Computing in Sensor Systems (DCOSS), 2017 International Conference on*.

[47] Sirajum Munir, Ripudaman Singh Arora, Craig Hesling, Juncheng Li, Jonathan Francis, Charles Shelton, Christopher Martin, Anthony Rowe, and Mario Berges. 2017. Real-Time Fine Grained Occupancy Estimation using Depth Sensors on ARM Embedded Platforms. In *Real-Time and Embedded Technology and Applications Symposium (RTAS), 2017 IEEE*. IEEE, 295–306.

[48] Tommy Oberg, Alek Karsznia, and Kurt Oberg. 1993. Basic gait parameters: reference data for normal subjects, 10-79 years of age. *Journal of rehabilitation research and development* 30, 2 (1993), 210.

[49] Alan V Oppenheim. 1999. *Discrete-time signal processing*. Pearson Education India.

[50] Veljo Otsason, Alex Varshavsky, Anthony LaMarca, and Eyal De Lara. 2005. Accurate GSM indoor localization. In *International conference on ubiquitous computing*. Springer, 141–158.

[51] Shwetak N Patel, Khai N Truong, and Gregory D Abowd. 2006. Powerline positioning: A practical sub-room-level indoor location system for domestic use. In *UbiComp*, Vol. 6. Springer, 441–458.

[52] Nissanka B Priyantha, Anit Chakraborty, and Hari Balakrishnan. 2000. The cricket location-support system. In *Proceedings of the 6th annual international conference on Mobile computing and networking*. ACM, 32–43.

[53] Qifan Pu, Sidhant Gupta, Shyamnath Gollakota, and Shwetak Patel. 2013. Whole-home gesture recognition using wireless signals. In *Proceedings of the 19th annual international conference on Mobile computing & networking*. ACM, 27–38.

[54] Tauhidur Rahman, Alexander T Adams, Ruth Vinisha Ravichandran, Mi Zhang, Shwetak N Patel, Julie A Kientz, and Tanzeem Choudhury. 2015. Dopplesleep: A contactless unobtrusive sleep sensing system using short-range doppler radar. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 39–50.

[55] Donald Reid. 1979. An algorithm for tracking multiple targets. *IEEE transactions on Automatic Control* 24, 6 (1979), 843–854.

[56] Mark A Richards, James A Scheer, William A Holm, et al. 2010. *Principles of modern radar*. Citeseer.

[57] Ralph Schmidt. 1986. Multiple emitter location and signal parameter estimation. *IEEE transactions on antennas and propagation* 34, 3 (1986), 276–280.

[58] Yu-Lin Shen and Chow-Shing Shin. 2009. Distributed sensing floor for an intelligent environment. *IEEE Sensors Journal* 9, 12 (2009), 1673–1678.

[59] Shuyu Shi, Stephan Sigg, and Yusheng Ji. 2013. Joint localization and activity recognition from ambient FM broadcast signals. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*. ACM, 521–530.

[60] Merrill Ivan Skolnik. 1970. Radar handbook. (1970).

[61] Elahe Soltanaghaei, Avinash Kalyanaraman, and Kamin Whitehouse. 2017. Peripheral WiFi Vision: Exploiting Multipath Reflections for More Sensitive Human Sensing. In *Proceedings of the 4th International on Workshop on Physical Analytics*. ACM, 13–18.

[62] Elahe Soltanaghaei and Kamin Whitehouse. 2016. WalkSense: Classifying Home Occupancy States Using Walkway Sensing. In *Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments*. ACM, 167–176.

[63] Vijay Srinivasan, John Stankovic, and Kamin Whitehouse. 2010. Using height sensors for biometric identification in multi-resident homes. In *International Conference on Pervasive Computing*. Springer, 337–354.

[64] G Trunk and S Brockett. 1993. Range and velocity ambiguity resolution. In *Radar Conference, 1993., Record of the 1993 IEEE National*. IEEE, 146–149.

[65] Deepak Vasisht, Swarun Kumar, and Dina Katabi. 2016. Decimeter-level localization with a single wifi access point. In *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*. USENIX Association, 165–178.

[66] Ju Wang, Hongbo Jiang, Jie Xiong, Kyle Jamieson, Xiaojiang Chen, Dingyi Fang, and Binbin Xie. 2016. LiFS: Low Human Effort, Device-Free Localization with Fine-Grained Subcarrier Information. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*. ACM, 243–256.

[67] Saiwen Wang, Jie Song, Jaime Lien, Ivan Poupyrev, and Otmar Hilliges. 2016. Interacting with soli: Exploring fine-grained dynamic gesture recognition in the radio-frequency spectrum. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, 851–860.

[68] Wei Wang, Alex X Liu, and Muhammad Shahzad. 2016. Gait recognition using wifi signals. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 363–373.

[69] Roy Want, Andy Hopper, Veronica Falcao, and Jonathan Gibbons. 1992. The active badge location system. *ACM Transactions on Information Systems (TOIS)* 10, 1 (1992), 91–102.

[70] Daniel H Wilson and Chris Atkeson. 2005. Simultaneous tracking and activity recognition (STAR) using many anonymous, binary sensors. In *International Conference on Pervasive Computing*. Springer, 62–79.

[71] Jiang Xiao, Zimu Zhou, Youwen Yi, and Lionel M Ni. 2016. A survey on wireless indoor localization from the device perspective. *ACM Computing Surveys (CSUR)* 49, 2 (2016), 25.

[72] Jie Xiong and Kyle Jamieson. 2013. ArrayTrack: A Fine-Grained Indoor Location System.. In *NSDI*. 71–84.

[73] Hui-Shyong Yeo, Gergely Flamich, Patrick Schrempf, David Harris-Birtill, and Aaron Quigley. 2016. Radarcat: Radar categorization for input & interaction. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, 833–841.

[74] Sangki Yun, Yi-Chao Chen, and Lili Qiu. 2015. Turning a mobile device into a mouse in the air. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 15–29.