Multi-Person Localization via RF Body Reflections

Fadel Adib, Zachary Kabelac, and Dina Katabi, *Massachusetts Institute of Technology*

Multi-Person Localization via RF Body Reflections

Fadel Adib   Zachary Kabelac   Dina Katabi
Massachusetts Institute of Technology

Abstract– We have recently witnessed the emergence of RF-based indoor localization systems that can track user motion without requiring the user to hold or wear any device. These systems can localize a user and track his gestures by relying solely on the reflections of wireless signals off his body, and work even if the user is behind a wall or obstruction. However, in order for these systems to become practical, they need to address two main challenges: 1) They need to be able to operate in the presence of more than one user in the environment, and 2) they must be able to localize a user without requiring him to move or change his position.

This paper presents WiTrack2.0, a multi-person localization system that operates in multipath-rich indoor environments and pinpoints users' locations based purely on the reflections of wireless signals off their bodies. WiTrack2.0 can even localize static users, and does so by sensing the minute movements due to their breathing. We built a prototype of WiTrack2.0 and evaluated it in a standard office building. Our results show that it can localize up to five people simultaneously with a median accuracy of 11.7 cm in each of the x/y dimensions. Furthermore, WiTrack2.0 provides coarse tracking of body parts, identifying the direction of a pointing hand with a median error of 12.5°, for multiple users in the environment.

1 INTRODUCTION

Over the past decade, the networking community has made major advances in RF-based indoor localization [5, 34, 21, 26, 38, 17, 21, 11, 22], which led to systems that can localize a wireless device with centimeter-scale accuracy. Recently however the research community has realized that it is possible to localize a user, without requiring him to wear or carry a wireless device [25, 3]. Such a leap from device-based to device-free indoor localization can enable ubiquitous tracking of people and their gestures. For example, it can enable a smart home to continuously localize its occupants and adjust the heating in each room according to the number of people in it. It would also enable a smart home to track our hand gestures so that we may control appliances by pointing at them, or turn the TV with a wave of our arm. Device-free tracking can also be leveraged in many applications where it is either inconvenient or infeasible for the user to hold/wear a device such as in gaming and virtual reality, elderly monitoring, intrusion detection, and search and rescue missions [3].

Past work has taken initial steps towards this vision [3, 18, 7]. However, these proposals have fundamental limitations that render them impractical for natural home environments. Specifically, they either require covering the entire space with a dense, surveyed grid of sensors [18, 7] or they fail in the presence of multiple users in the environment [3]. Additionally, these past proposals are also limited in their ability to detect the presence of users. In particular, they either require the user to continuously move to detect his presence [3], or they need to perform extensive prior calibration or training [18, 7].

In this paper, we introduce WiTrack2.0, a device free localization system that transcends these limitations. Specifically, WiTrack2.0 accurately localizes multiple users in the environment. It does so by disentangling the reflections of wireless signals that bounce off their bodies. Furthermore, it neither requires prior calibration nor that the users move in order to localize them.

To achieve its goal, WiTrack2.0 has to deal with multiple challenges. As with traditional device-based localization, the most difficult challenge in indoor environments is the multipath effect [34, 17]. Specifically, wireless signals reflect off all objects in the environment making it hard to associate the incoming signal with a particular location. To overcome this challenge, past work [3] focuses on motion to capture signal reflections that change with time. It then assumes that only one person is present in the environment, and hence all motion can be attributed to him. However, if multiple people move in the environment or if the person is static, then this assumption no longer works.

To address this challenge, we observe that the indoor multipath varies significantly when it is measured from different vantage points. Hence, one can address this problem by positioning multiple transmit and receive antennas in the environment, and measuring the time of flight from each of these transmit-receive antenna pairs. However, the signals emitted from the different transmitters will reflect off the bodies of the all the users in the environment, and these reflections interfere with each other leading to wireless collisions. In §5, we show how WiTrack2.0 disentangles these interfering reflected signals to localize multiple users in the presence of heavy indoor multipath.

A second challenge that WiTrack2.0 has to address is related to the near-far problem. Specifically, reflections off the nearest person can have much more power than distant reflections, obfuscating the signal from distant people, and preventing their detection or tracking. To address this issue, we introduce Successive Silhou-

ette Cancellation (SSC) an approach to address the near-far problem, which is inspired by successive interference cancellation. This technique starts by localizing the closest person, then eliminates his impact on the received signal, before proceeding to localize further users (who have weaker reflections). It repeats this process iteratively until it has localized all the users in a scene. Note, however, that each user is not a point reflector; hence, his wireless reflection has a complex structure that must be taken into account, as we describe in §6.

A third challenge that our system addresses is related to localizing static users. Specifically, past work that tracks human motion needs to eliminate reflections off static objects by subtracting consecutive measurements. However, this subtraction also results in eliminating the reflections off static users. To enable us to localize static users, we exploit the fact these users still move slightly due to their breathing. However, the breathing motion is fairly slow in comparison to body motion. Specifically, the chest moves by a sub-centimeter distance over a period of few seconds; in contrast, a human would pace indoors at 1 m/s. Hence, WiTrack2.0 processes the reflected signals at multiple time scales that enable it to accurately localize both types of movements as we describe in §7,

We have built a prototype of WiTrack2.0, using USRP software radios and an analog FMCW radio. We run experiments both in line-of-sight (LOS) scenarios and non-line-of-sight (NLOS) scenarios, where the device is in a different room and is tracking people's motion through the wall. Empirical results from over 300 experiments with 11 human subjects show the following:

- *Motion Tracking:* WiTrack2.0 accurately tracks the motion of up to four users simultaneously, without requiring the users to hold or wear any wireless device. In an area that spans 5 *m* × 7 *m*, its median error across all users is 12.1cm in the x/y dimensions.
- *Localizing Static People:* By leveraging their breathing motion, WiTrack2.0 accurately localizes up to five static people in the environment. Its median error is 11.2 cm in the x/y dimensions across all the users in the scene.
- *Tracking Hand Movements:* WiTrack2.0's localization capability extends beyond tracking a user's body to tracking body parts. We leverage this capability to recognize concurrent gestures performed in 3D space by multiple users. In particular, we consider a gesture in which three users point in different directions at the same time. Our WiTrack2.0 prototype detects the pointing directions of all three users with a median accuracy of 10.3°.

Contributions: This paper demonstrates the first device-free RF-localization system that can accurately localize multiple people to centimeter-scale in indoor multipath-rich environments. This is enabled by a novel transmission protocol and signal processing algorithms which allow isolating and localizing different users in the environ-ment. The paper also presents an evaluation of the system, showing that it can localize moving and static users in line-of-sight and through-wall settings with a median accuracy of 7-18 cm across all of these scenarios.

2 RELATED WORK

Indoor Localization. WiTrack2.0 builds on a rich networking literature on indoor localization [5, 34, 21, 26, 38, 17, 21, 11, 22] which has focused on localizing wireless devices. In comparison to all of these works, however, WiTrack2.0 focuses on localizing users by relying purely on the reflections of RF signals off their bodies.

WiTrack2.0 is also related to proposals for device-free localization, which deploy a sensor network and measure the signal strength between different nodes to localize users [36, 40]. However, in comparison to these past proposals, WiTrack2.0 neither requires deploying a network of dozens to hundreds of sensors [36, 18] nor does it require extensive calibration [25, 24, 40]. Furthermore, because it relies on time-of-flight measurements, it can achieve a localization accuracy that is 10× higher than state-of-the-art RSSI-based systems [25, 24, 18, 7, 19].

WiTrack2.0's design builds on our prior work, WiTrack [3], which also used time of flight (TOF) measurements to achieve high localization accuracy, and which did not require prior environmental calibration. In comparison to WiTrack, which could localize only a single person and only if that person is moving, WiTrack2.0 can localize up to five users simultaneously, even if they are perfectly static (by relying on their breathing motion).

WiTrack2.0 is also related to non-localization systems that employ RF reflections off the human body [4, 20, 33, 35]. These systems can detect the presence of people or identify a handful of gestures or activities. However, unlike WiTrack2.0, they have no mechanism for obtaining the location of a person.

Radar Systems. WiTrack2.0 builds on past radar literature. In particular, it uses the FMCW (Frequency Modulated Carrier Waves) technique to obtain accurate time-of-flight measurements [15, 23, 31, 9]. However, its usage of FMCW has a key property that differentiates it from all prior designs: it transmits from multiple antennas *concurrently* while still allowing its receivers to isolate the reflections from each of the Tx antennas.

More importantly, none of the past work on radar addresses the issue of indoor multipath. Specifically, past work on see-through-wall radar has been tailored for usage in strictly military settings. Hence, it mostly operates in an open field with an erected wall [23, 10], or it focuses on detecting metallic objects which have significantly higher reflection coefficients than furniture, walls, or the human body [13, 28, 27]. Work tested on human subjects in indoor environments has focused on detecting the presence of humans rather than on accurately localizing them [16, 39]; in fact, these techniques acknowledge

| (a) **FMCW provides TOF** | (b) **TOF Profile** | (c) **Background Subtraction** | (d) **TOF to Ellipse** |
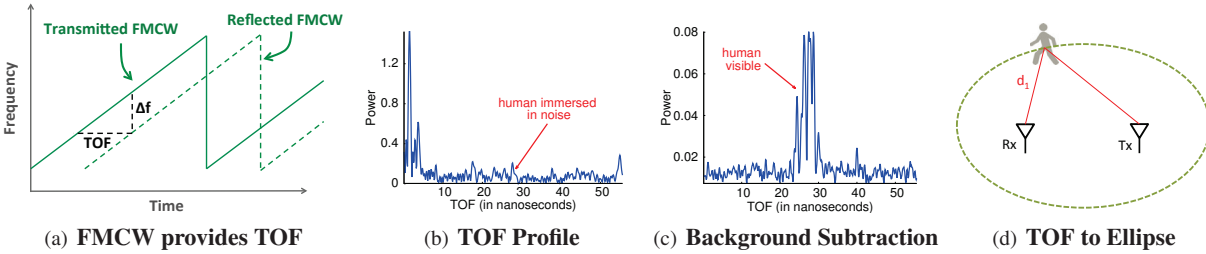
Figure 1—Measuring Distances using FMCW. (a) shows the transmitted FMCW signal and its reflection. The TOF between the transmitted and received signals maps to a frequency shift Δf between them. (b) shows the TOF profile obtained after performing an FFT on the baseband FMCW signal. The profile plots the amount of reflected power at each TOF. (c) shows that a moving person's reflections pop up after background subtraction. (d) shows how a TOF measurement maps to a round-trip distance, which may correspond to any location on an ellipse whose foci are Tx and Rx.

that multipath leads to the well-known "ghosting effect", but ignore these effects since they do not prevent detecting the presence of a human. In comparison to the above work, WiTrack2.0 focuses on accurate localization of humans in daily indoor settings, and hence introduces two new techniques that enable it to address the heavy multipath in standard indoor environments: multi-shift FMCW, and successive silhouette cancellation.

Iterative Cancellation Frameworks. The framework of iteratively identifying and canceling out the strongest components of a signal is widely used in many domains. Naturally, however, the details of how the highest power component is identified and is eliminated varies from one application to another. In the communications community, we refer to such techniques as successive interference cancellation, and they have been used in a large number of applications such as ZigZag [12], VBLAST [37], and full duplex [6]. In the radio astronomy community, these techniques are referred to as CLEAN algorithms and, similarly, have a large number of instantiations [14, 29, 30]. Our work on successive silhouette cancellation also falls under this framework and is inspired by these algorithms. However, in comparison to all the past work, it focuses on identifying the reflections of the humans in the environment and canceling them by taking into account the different vantage points from which the time-of-flight is measured as well as the fact that the human body is not a point reflector.

3 PRIMER

This section provides necessary background regarding single-person motion tracking via RF body reflections.

The process of localizing a user based on radio reflections off her body has three steps: 1) obtaining time-of-flight (TOF) measurements to various reflectors in the environment; 2) eliminating TOF measurements due to reflections of static objects like walls and furniture; and 3) mapping the user's TOFs to a location.

Step 1: Obtaining TOF measurements to various reflectors in the environment. A typical way for measuring the time-of-flight (TOF) is to use a Frequency-Modulated Carrier Waves (FMCW) radio. An FMCW transmitter sends a narrowband signal (e.g., a few KHz) but makes the carrier frequency sweep linearly in time,

as illustrated by the solid green line in Fig. 1(a). The reflected signal is a delayed version of the transmitted signal, which arrives after bouncing off a reflector, as shown by the dotted green line in Fig. 1(a). Because time and frequency are linearly related in FMCW, the delay between the two signals maps to a frequency shift Δf between them. Hence, the time-of-flight can be measured as the difference in frequency Δf divided by the slope of the sweep in Fig. 1(a):

$$TOF = \Delta f / slope \qquad (1)$$

This description generalizes to an environment with multiple reflectors. Because wireless reflections add up linearly over the medium, the received signal is a linear combination of multiple reflections, each of them shifted by some Δf that corresponds to its TOF. Hence, one can extract all these TOFs by taking an FFT of the received signal. The output of the FFT gives us the **_TOF profile_** which we define as the reflected power we obtain at each possible TOF between the transmit antenna and receive antenna, as shown in Fig. 1(b).[1]

Step 2: Eliminating TOFs of static reflectors. To localize a human, we need to identify his/her reflections from those of other objects in the environment (e.g., walls and furniture). This may be done by leveraging the fact that the reflections of static objects remain constant over time. Hence, one can eliminate the power from static reflectors by performing background subtraction – i.e., by subtracting the output of the TOF profile in a given sweep from the TOF profile of the signal in the previous sweep. Fig. 1(b) and 1(c) show how background subtraction eliminates the power in static TOFs from the TOF profile, and allows one to notice the weak power resulting from a moving person.

Step 3: Mapping TOFs to distances. Recall that the TOF corresponds to the time it takes the signal to travel from the transmitter to a reflector and back to the receiver. Hence, we can compute the corresponding round-trip distance by multiplying this TOF by the speed of light C:

$$round\ trip\ distance = C \times TOF = C \times \frac{\Delta f}{slope} \qquad (2)$$

---

[1] The FFT is performed on the baseband FMCW signal – i.e., on the signal we obtain after mixing the received signal with the transmitted FMCW.

(a) **Antenna** (b) **Antenna Setup**
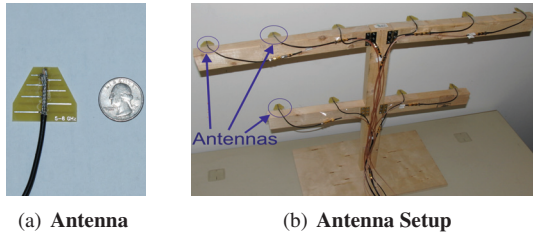
Figure 2—WiTrack2.0's Antennas and Setup. (a) shows one of WiTrack2.0's directional antenna ($3cm \times 3.4cm$) placed next to a quarter; (b) shows the antenna setup in our experiments, where antennas are mounted on a $2m \times 1m$ platform and arranged in a single vertical plane.

Knowing the round trip distance localizes the person to an ellipse whose foci are the transmit and receive antennas.

4 WITRACK2.0 OVERVIEW

WiTrack2.0 is a wireless system that can achieve highly accurate localization of multiple users in multipath-rich indoor environments, by relying purely on the reflections of wireless signals off the users' bodies. For static users, it localizes them based on their breathing, and can also localize the hand motions of multiple people, enabling a multi-user gesture-based interface.

WiTrack2.0 is a multi-antenna system consisting of five transmit antennas and five receive antennas, as shown in Fig. 2. The antennas are directional, stacked in a single plane, and mounted on a foldable platform as shown in Fig. 2(b). This arrangement is chosen because it enables see-through-wall applications, whereby all the antennas need to be lined up in a plane facing the wall of interest.

WiTrack2.0 operates by transmitting RF signals and capturing their reflections after they bounce off different users in the environment. Algorithmically, WiTrack2.0 has two main components: 1) Multi-shift FMCW, a technique that enables it to deal with multipath effects, and (2) Successive Silhouette Cancellation (SSC), an algorithm that allows WiTrack2.0 to overcome the near-far problem. The following sections describe these components.

5 MULTI-SHIFT FMCW

Multipath is the first challenge in accurate indoor localization. Specifically, not all reflections that survive background subtraction correspond to a moving person. This is because the signal reflected off the human body may also reflect off other objects in the environment before arriving at the receive antenna. As this person moves, this multipath reflection also moves with him and survives the background subtraction step. In single-user localization, one may eliminate this type of multipath by leveraging that these secondary reflections travel along a longer path before they arrive at the receive antenna. Specifically, by selecting the smallest TOF after background subtraction, one may identify the round-trip distance to the user.

However, the above invariant does not hold in multi-person localization since different users are at different distances with respect to the antennas, and the multipath

of a nearby user may arrive earlier than that of a more distant user, or even interfere with it. In this section, we explore this challenge in more details, and show how we can overcome it by obtaining time-of-flight measurements from different vantage points in the environment.

5.1 Addressing Multi-path in Multi-User Localization

To explore the above challenge in practice, we run an experiment with two users in a 5×7 m furnished room (with tables, chairs, etc.) in a standard office building. We study what happens as we successively overlay ellipses from different transmit-receive pairs. Recall from §3 that each transmit-receive antenna pair provides us with a *TOF profile* – i.e., it tells us how much reflected power we obtain at each possible TOF between the transmit antenna and receive antenna (see Fig. 1(c)) – and that each such TOF corresponds to an ellipse in 2D (as in Fig. 1(d)).

Now let us map all TOFs in a TOF profile to the corresponding round trip distances using Eq. 2, and hence the resulting ellipses. This process produces a heatmap like the one in Fig. 3(a), where the x and y axes correspond to the plane of motion. For each ellipse in the heatmap, the color in the image reflects the amount of received power at the corresponding TOF. Hence, the ellipse in red corresponds to a strong reflector in the environment. The orange, yellow, and green ellipses correspond to weaker reflections respectively; these reflections could either be due to another person, multi-path reflections of the first person, or noise. The blue regions in the background correspond to the absence of reflections from those areas.

Note that the heatmap shows a pattern of half-ellipses; the foci of these ellipses are the transmit and receive antennas, both of which are placed along the $y = 0$ axis. The reason we only show the upper half of the ellipses is that we are using directional antennas and we focus them towards the positive y direction. Hence, we know that we do not receive reflections from behind the antennas.

Fig. 3(a) shows the ellipses corresponding to the TOF profiles from one Tx-Rx pair. Now, let us see what happens when we superimpose the heatmaps obtained from two Tx-Rx pairs. Fig. 3(b) shows the heatmap we obtain when we overlay the ellipses of the first transmit-receive pair with those from a second pair. We can now see two patterns of ellipses in the figure, the first pattern resulting from the TOFs of the first pair, and the second pattern due to the TOFs of the second pair. These ellipses intersect in multiple locations, resulting in red or orange regions, which suggest a higher probability for a reflector to be in those regions. Recall that there are two people in this experiment. However, Fig. 3(b) is not enough to identify the locations of these two people.

Figs. 3(c) and 3(d) show the result of overlaying ellipses from three and four Tx-Rx pairs respectively. The figures show how the noise and multi-path from different

| (a) **One Tx-Rx pair** | (b) **Two Tx-Rx pairs** | (c) **Three Tx-Rx pairs** | (d) **Four Tx-Rx pairs** | (e) **Five Tx-Rx pairs** |
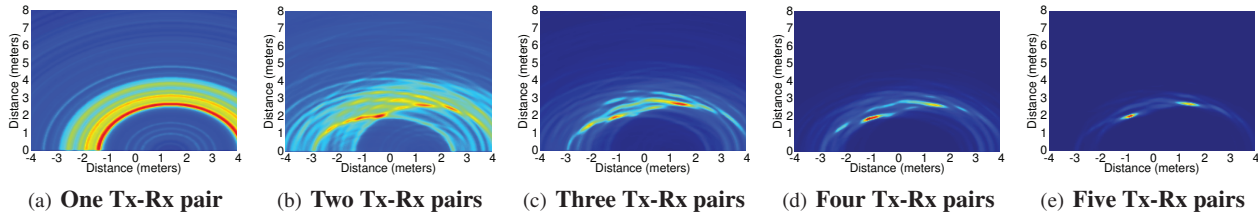
Figure 3—Increasing the Number of Tx-Rx pairs enables Localizing Multiple Users. The figure shows the heatmaps obtained from combining TOF profiles of multiple Tx-Rx antenna pairs in the presence of two users. The x/y axes of each heatmap correspond to the real world x/y dimensions.

antennas averages out to result in a dark blue background. This is because different Tx-Rx pairs have different perspectives of the indoor environment; hence, they do not observe the same noise or multi-path reflections. As a result, the more we overlay heatmaps from different Tx-Rx pairs, the dimmer the multipath effect, and the clearer the candidate locations for the two people in the environment.

Next, we overlay the ellipses from five transmit-receive pairs and show the resulting heatmap in Fig. 3(e). We can now clearly see two bright spots in the heatmap: one is red and the other is orange, whereas the rest of the heatmap is mostly a navy blue background indicating the absence of reflectors. Hence, in this experiment, we are able to localize the two users using TOF measurements from five Tx-Rx pairs. Combining these measurements together allowed us to eliminate the multipath effects and localize the two people passively using their reflections.

Summary: As the number of users increases, we need TOF measurements from a larger number of Tx-Rx pairs to localize them, and extract their reflections from multipath. For the case of two users, we have seen a scenario whereby the TOFs of five Tx-Rx pairs were sufficient to accurately localize both of them. In general, the exact number would depend on multipath and noise in the environment as well as on the number of users we wish to localize. These observations motivate a mechanism that can provide us with a large number of Tx-Rx pairs while scaling with the number of users in the environment.

5.2 The Design of Multi-shift FMCW

In the previous section, we showed that we can localize two people by overlaying many heatmaps obtained from mapping the TOF profiles of multiple Tx-Rx pairs to the corresponding ellipses. But how do we obtain TOFs from many Tx-Rx pairs? One option is to use one FMCW transmitter and a large number of receivers. In this case, to obtain N Tx-Rx pairs, we would need one transmitter and N receivers. The problem with this approach is that it needs a large number of receivers, and hence does not scale well as we add more users to the environment.

A more appealing option is to use multiple FMCW transmit and receive antennas. Since the signal transmitted from each transmit antenna is received by all receive antennas, this allows us to obtain N Tx-Rx pairs using only $\sqrt{N}$ transmit antennas and $\sqrt{N}$ receive antennas.
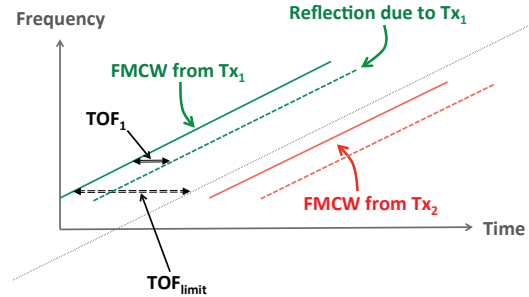

Figure 4—Multi-shift FMCW. WiTrack2.0 transmits FMCW signals from different transmit antennas after inserting virtual delays between them. Each delay must be larger than the highest time-of-flight (TOF_{limit}) due to objects in the environment.

However, the problem with this approach is that the signals from the different FMCW transmitters will interfere with each other over the wireless medium, and this interference will lead to localization errors. To see why this is true, consider a simple example where we want to localize a user, and we have two transmit antennas, Tx1 and Tx2, and one receive antenna Rx. The receive antenna will receive two reflections – one due to the signal transmitted from Tx1, and another due to Tx2's signal. Hence, its TOF profile will contain two spikes referring to two time-of-flight measurements TOF_1 and TOF_2.

With two TOFs, we should be able to localize a single user based on the intersection of the resulting ellipses. However, the receiver has no idea which TOF corresponds to the reflection of the FMWC signal generated from Tx1 and which corresponds to the reflection of the FMCW signal generated by Tx2. Not knowing the correct Tx means that we do not know the foci of the two ellipses and hence cannot localize. For example, if we incorrectly associate TOF_1 with Tx2 and TOF_2 with Tx1, we will generate a wrong set of ellipses, and localize the person to an incorrect location. Further, this problem becomes more complicated as we add more transmit antennas to the system. Therefore, to localize the user, WiTrack2.0 needs a mechanism to associate these TOF measurements with their corresponding transmit antennas.

We address this challenge by leveraging the structure of the FMCW signal. Recall that FMCW consists of a continuous linear frequency sweep as shown by the green line in Fig. 4. When the FMCW signal hits a body, it reflects back with a delay that corresponds to the body's TOF. Now, let us say TOF_{limit} is the maximum TOF
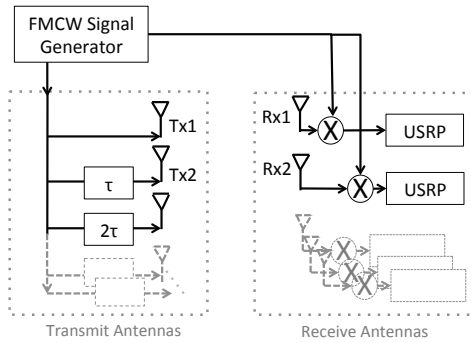
Figure 5—Multi-shift FMCW Architecture. The generated FMCW signal is fed to multiple transmit antennas via different delay lines. At the receive side, the TOF measurements from the different antennas are combined to obtain the 2D heatmaps.

that we expect in the typical indoor environment where WiTrack2.0 operates. We can delay the FMCW signal from the second transmitter by $\tau > TOF_{limit}$ so that all TOFs from the second transmitter are shifted by τ with respect to those from the first transmitter, as shown by the red line in Fig. 4. Thus, we can prevent the various FMCW signals from interfering by ensuring that each transmitted FMCW signal is time shifted with respect to the others, and those shifts are significantly larger than the time-of-flight to objects in the environment. We refer to this design as Multi-shift FMCW.

As a result, the receiver would still compute two TOF measurements: the first measurement (from Tx1) would be TOF_1, and the second measurement (from Tx2) would be $TOF_2' = TOF_2 + \tau$. Knowing that the TOF measurements from Tx2 will always be larger than τ, WiTrack2.0 determines that TOF_1 is due to the signal transmitted by Tx1, and TOF_2' is due to the signal transmitted by Tx2.

This idea can be extended to more than two Tx antennas, as shown in Fig. 5. Specifically, we can transmit the FMCW signal directly over the air from Tx1, then shift it by τ and transmit it from Tx2, then shift it by 2τ and transmit it from Tx3, and so on. At the receive side, all TOFs between 0 and τ are mapped to Tx1, whereas distances between τ and 2τ are mapped to Tx2, and so on.

Summary: Multi-shift FMCW has two components: the first component allows us to obtain TOF measurements from a large number of Tx-Rx pairs; the second component operates on the TOFs obtained from these different Tx-Rx pairs by superimposing them into a 2D heatmap, which allows us to localize multiple users in the scene.

6 SUCCESSIVE SILHOUETTE CANCELLATION

With multi-shift FMCW, we obtain TOF profiles from a large number of Tx-Rx pairs, map them to 2D heatmaps, overlay the heatmaps, and start identifying users' locations. However, in practice this is not sufficient because different users will exhibit the near-far problem. Specifically, reflections of a nearby user are much stronger than reflections of a faraway user or one behind an obstruction.
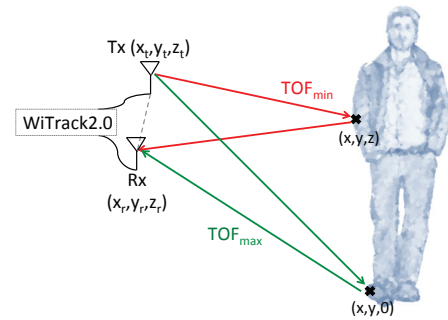
Figure 7—Finding TOF_{min} **and** TOF_{max}. TOF_{min} is determined by the round-trip distance from the Tx-Rx pair to the closest point on the person's body. Since the antennas are elevated, TOF_{max} is typically due to the round-trip distance to the person's feet.

Fig. 6(a) illustrates this challenge. It shows the 2D heatmap obtained in the presence of four persons in the environment. The heatmap allows us to localize only two of these persons: one is clearly visible at $(0.5, 2)$, and another is fairly visible at $(-0.5, 1.3)$. The other two people, who are farther away from WiTrack2.0, are completely overwhelmed by the power of the first two persons.

To deal with this near-far problem, rather than localizing all users in one shot, WiTrack2.0 performs Successive Silhouette Cancellation (SSC) which consists of 4 steps:

1. *SSC Detection:* finds the location of the strongest user by overlaying the heatmaps of all Tx-Rx pairs.

2. *SSC Re-mapping:* maps a person's location to the set of TOFs that would have generated that location at each transmit-receive pair.

3. *SSC Cancellation:* cancels the impact of the person on the TOF profiles of all Tx-Rx pairs.

4. *Iteration:* re-computes the heatmaps using the TOF profiles after cancellation, overlays them, and proceeds to find the next strongest reflector.
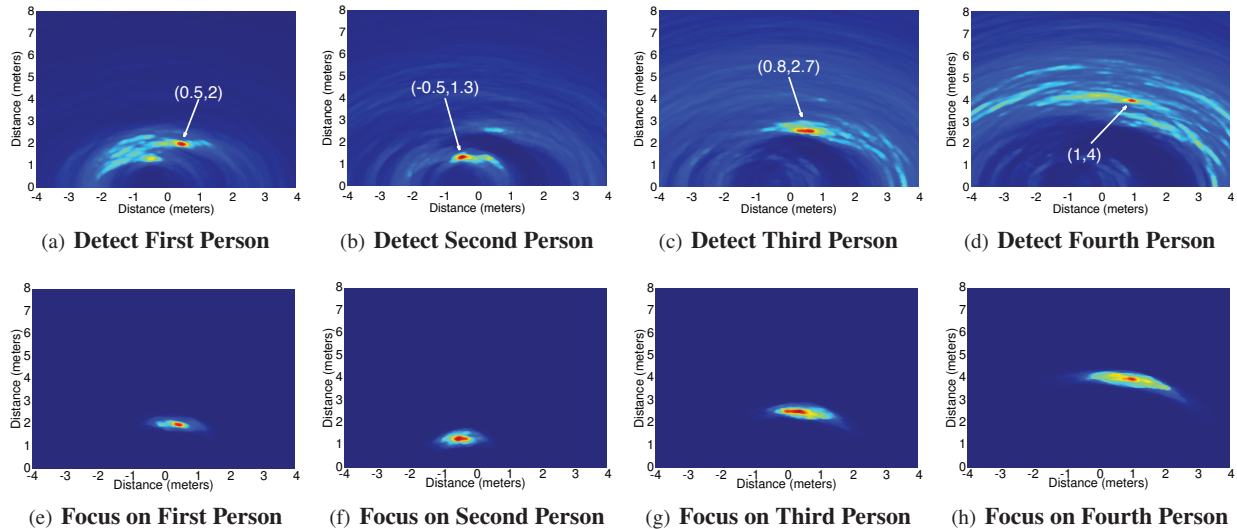
We now describe each of these steps in detail by walking through the example with four persons shown in Fig. 6.

SSC Detection. In the first step, SSC finds the location of the highest power reflector in the 2D heatmap of Fig. 6(a). In this example, the highest power is at $(0.5, 2)$, indicating that there is a person in that location.

SSC Re-mapping. Given the (x, y) coordinates of the person, we map his location back to the corresponding TOF at each transmit-receive pair. Keep in mind that each person is not a point reflector; hence, we need to estimate the spread of reflections off his entire body on the TOF profile of each transmit-receive pair.

To see how we can do this, let us look at the illustration in Fig. 7 to understand the effect of a person's body on one transmit-receive pair. The signal transmitted from the transmit antenna will reflect off different points on the person's body before arriving at the receive antenna. Thus, the person's reflections will appear between some TOF_{min} and TOF_{max} in the TOF profile at the Rx antenna.

Note that TOF_{min} and TOF_{max} are bounded by the closest and furthest points respectively on a person's body

(a) **Detect First Person**   (b) **Detect Second Person**   (c) **Detect Third Person**   (d) **Detect Fourth Person**

(e) **Focus on First Person**   (f) **Focus on Second Person**   (g) **Focus on Third Person**   (h) **Focus on Fourth Person**

Figure 6—Successive Silhouette Cancellation. (a) shows the 2D heatmap obtained by combining all the TOFs in the presence of four users. (b)-(d) show the heatmaps obtained after canceling out the first, second, and third user respectively. (e)-(h) show the result of the SSC focusing step on each of the users, and how it enables us to accurately localize each person while eliminating interference from all other users.

from the transmit-receive antenna pair. Let us first focus on how we can obtain TOF_{min}. By definition, the closest point on the person's body is the one that corresponds to the shortest round-trip distance to the Tx-Rx pair, where the round-trip distance is the summation of the forward path from Tx to that point and the path from that point back to Rx. Formally, for a Tx antenna at $(x_t, 0, z_t)$, an Rx antenna at $(x_r, 0, z_r)$,[2] we can compute d_{min} as:

$$\min_z \sqrt{(x_t - x)^2 + y^2 + (z_t - z)^2} + \sqrt{(x_r - x)^2 + y^2 + (z_r - z)^2} \quad (3)$$

where (x, y, z) is any reflection point on the user's body. One can show that this expression is minimized when:

$$\frac{z - z_t}{z - z_r} = -\sqrt{\frac{(x_t - x)^2 + y^2}{(x_r - x)^2 + y^2}} \quad (4)$$

Hence, using the detected (x, y) position, we can solve for z then substitute in Eq. 3 to obtain d_{min}.

Similarly, TOF_{max} is bounded by the round-trip distance to point on the person's body that is furthest from the Tx-Rx pair. Again, the x and y coordinates of the furthest point are determined by the person's location from the SSC Detection step. However, we still need to figure out the z coordinate of this point. Since the transmitter and receiver are both raised above the ground (at around 1.2 meters above the ground), the furthest point from the Tx-Rx pair is typically at the person's feet. Therefore, we know that the coordinates of this point are $(x, y, 0)$, and hence we can compute d_{max} as:

$$d_{max} = \sqrt{(x_t - x)^2 + (y)^2 + z_t^2} + \sqrt{(x_r - x)^2 + (y)^2 + z_r^2}.$$

Finally, we can map d_{min} and d_{max} to TOF_{min} and TOF_{max} by dividing them by the speed of light C.

[2]Recall that all the antennas are in the vertical plane $y = 0$, which is parallel to a person's standing height.

SSC Cancellation. The next step is to use TOF_{min} and TOF_{max} to cancel the person's reflections from the TOF profiles of each transmit-receive pair. To do that, we take a conservative approach and remove the power in all TOFs between TOF_{min} and TOF_{max} within that profile. Of course, this means that we might also be partially canceling out the reflections of another person who happens to have a similar time of flight to this Tx-Rx pair. However, we rely on the fact that multi-shift FMCW provides a large number of TOF profiles from many Tx-Rx pairs. Hence, even if we cancel out the power in the TOF of a person with respect to a particular Tx-Rx pair, each person will continue to have a sufficient number of TOF measurements from the rest of the antennas.

We repeat the process of computing TOF_{min} and TOF_{max} with respect of each Tx-Rx pair and cancelling the power in that range, until we have eliminated any power from the recently localized person.

Iteration. We proceed to localize the next person. This is done by regenerating the heatmaps from the updated TOF profiles and overlaying them. Fig. 6(b) shows the obtained image after performing this procedure for the first person. Now, a person at $(-0.5, 1.3)$ becomes the strongest reflector in the scene.

We repeat the same procedure for this user, canceling out his interference, then reconstructing a 2D heatmap in Fig. 6(c) using the remaining TOF measurements. Now, the person with the strongest reflection is at $(0.8, 2.7)$. Note that this heatmap is noisier than Figs. 6(a) and 6(b) because now we are dealing with a more distant person.

WiTrack2.0 repeats the same cancellation procedure for the third person and constructs the 2D heatmap in Fig. 6(d). The figure shows a strong reflection at $(1, 4)$. Recall that our antennas are placed along the $y = 0$ axis,
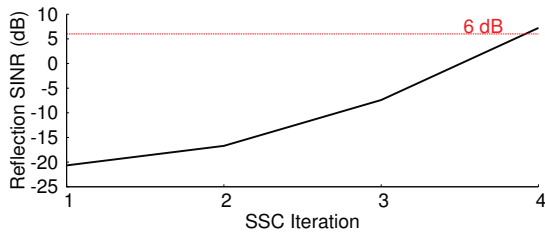
Figure 8—SINR of the Farthest User Throughout SSC Iterations. The figure shows how the SINR of the farthest user increases with each iteration of the SSC algorithm. After the first, second and third person are removed from the heatmap images in Figs.6(a)-(c), the SINR of the fourth person increases to 7dB, allowing us to detect his presence.
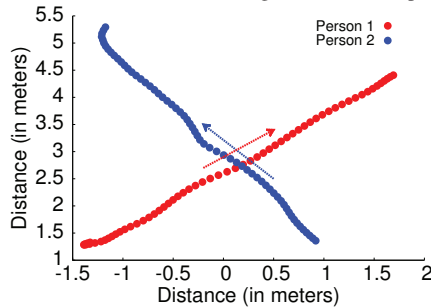
Figure 9—Disentangling Crossing Paths. When two people cross paths, they typically keep going along the same direction they were going before their paths crossed.

which means that this is indeed the furthest person in the scene. Also note that the heatmap is now even noisier. This is expected because the furthest person's reflections are much weaker. WiTrack2.0 repeats interference cancellation for the fourth person, and determines that the SNR of the maximum reflector in the resulting heatmap does not pass a threshold test. Hence, it determines that there are only four people in the scene.

We note that each of these heatmaps are scaled so that the highest power is always in red and the lowest power is in navy blue; this change in scale emphasizes the location of the strongest reflectors and allows us to better visualize their locations. To gain more insight into the power values and to better understand how SSC improves our detection of further away users, Fig. 8 plots the Signal to Interference and Noise Ratio (SINR) of the fourth person during each iteration of SSC. The fourth user's SINR initially starts at -21dB and is not visible in Fig. 6(a). Once the first and second users are removed by SSC, the SINR increases to -7dB and we can start detecting the user's presence in the back of Fig. 6(c). Performing another iteration raises the fourth person's SINR above the noise floor to 7dB. It also brings it above our threshold of 6dB – i.e., twice the noise floor – making him detectable.

We perform four additional steps to improve SSC:

- *Refocusing Step:* After obtaining the initial estimates of the locations of all four persons, WiTrack2.0 performs a focusing step for each user to refine his location estimate. This is done by reconstructing an interference-free 2D heatmap only using the range in the TOF profiles that corresponds to TOFs between TOF_{min} and TOF_{max}

for that Tx-Rx pair. Figs. 6(e)- 6(h) show the images obtained from this focusing step. In these images, the location of each person is much clearer,[3] which enables higher-accuracy localization.

- *Leveraging Motion Continuity:* After obtaining the estimates from SSC, WiTrack2.0 applies a Kalman filter and performs outlier rejection to reject impractical jumps in location estimates that would otherwise correspond to abnormal human motion over a very short period of time.

- *Disentangling Crossing Paths:* To disentangle multiple people who cross paths, we look at their direction of motion before they crossed paths and project how they would proceed with the same speed and direction as they are crossing paths. This helps us with associating each person with his own trajectory after crossing. Fig. 9 shows an example with two people crossing paths and how we were able to track their trajectories despite that. Of course, this approach does not generalize to every single case, which may lead to some association errors after the crossings but not to localization errors.

- *Extending SSC to 3D Gesture Recognition:* Similar to past work [3], WiTrack2.0 can differentiate a hand motion from a whole-body motion (like walking) by leveraging the fact that a person's hand has a much smaller reflective surface than his entire body. Unlike past work, however, WiTrack2.0 can track gestures even when they are simultaneously performed by multiple users. Specifically, by exploiting SSC focusing, it zooms onto each user individually to track his gestures. In our evaluation, we focus on testing a pointing gesture, where different users point in different directions at the same time. By tracking the trajectory of each moving hand, we can determine its pointing direction. Note that we perform these pointing gestures in 3D and track hand motion by using the TOFs from the different Tx-Rx pairs to construct a 3D point cloud rather than a 2D heatmap.[4] The results in §10.3 show that we can accurately track hand gestures performed by multiple users in 3D space.
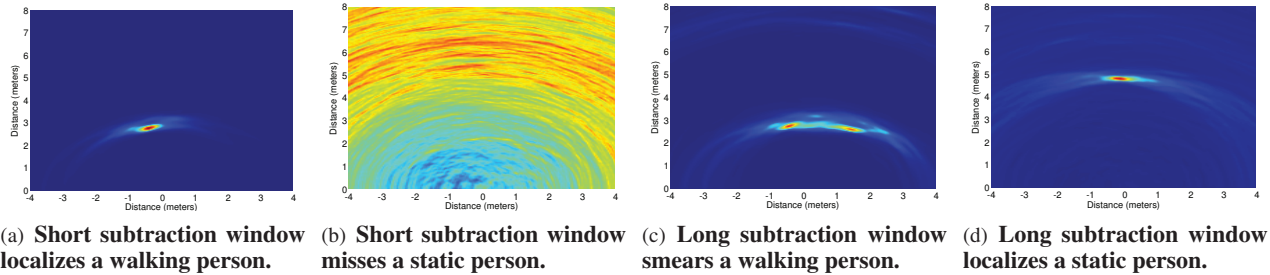
7 LOCALIZATION BASED ON BREATHING

We extend WiTrack2.0's SSC algorithm to localize static people based on their breathing. Recall from §3 that in order to track a user based on her radio reflections, we need to eliminate reflections off all static objects in the environment (like walls and furniture). This is typically achieved by performing a background subtraction step, i.e., by taking TOF profiles from adjacent time windows and subtracting them out from each other.[5]

---

[3]This is because all other users' reflections are eliminated, while, without refocusing, only users detected in prior iterations are eliminated.

[4]Recall from §3 that a given TOF maps to an ellipse in 2D and an ellipsoid in 3D. The intersection of ellipsoids in 3D allow us to track these pointing gestures.

[5]Recall that we obtain one TOF profile by taking an FFT over the received FMCW signal in baseband. Since the FMCW signal is repeatedly swept, we can compute a new TOF profile from each sweep.

(a) **Short subtraction window localizes a walking person.**   (b) **Short subtraction window misses a static person.**   (c) **Long subtraction window smears a walking person.**   (d) **Long subtraction window localizes a static person.**

Figure 10—Need For Multiple Subtraction Windows. The 2D heatmaps show that a short subtraction window accurately localizes a pacing person in (a) but not a static person in (b). A long subtraction window smears the walking person's location in (c) but localizes a breathing person in (d).

Whereas this approach enables us to track moving people, it prevents us from detecting a static person – e.g., someone who is standing or sitting still. Specifically, because a static person remains in the same location, his TOF does not change, and hence his reflections would appear as static and will be eliminated in the process of background subtraction. To see this in practice, we run two experiments where we perform background subtraction by subtracting two TOF profiles that are 12.5 milliseconds apart from each other. The first experiment is performed with a walking person and the resulting heatmap is shown in Fig. 10(a), whereas the second experiment is performed in the presence of a person who is sitting at $(0, 5)$ and the resulting heatmap is shown in Fig. 10(b). These experiments show how the heatmap of a moving person after background subtraction would allow us to localize him accurately, whereas the heatmap of the static person after background subtraction is very noisy and does not allow us to localize the person.

To localize static people, one needs to realize that even a static person moves slightly due to breathing. Specifically, during the process of breathing, the human chest moves by a sub-centimeter distance over a period of few seconds. The key challenge is that this change does not translate into a discernible change in the TOF of the person. However, over an interval of time of a few seconds (i.e., as the person inhales and exhales), it would result in discernible changes in the reflected signal. Therefore, by subtracting frames in time that are few seconds apart, we should be able to localize the breathing motion.

In fact, Fig. 10(d) shows that we can accurately localize a person who is sitting still by using a subtraction window of 2.5 seconds. Note, however, that this long subtraction window will introduce errors in localizing a pacing person. In particular, since typical indoor walking speed is around 1 m/s [8], subtracting two frames that are 2.5 seconds apart would result in smearing the person's location and may also result in mistaking him for two people as shown in Fig. 10(c).

Thus, to accurately localize both static and moving people, WiTrack2.0 performs background subtraction with different subtraction windows. To localize moving users, it uses a subtraction window of 12.5 ms. On the other hand, normal adults inhale and exhale over a period of 3–6 seconds [32] causing their TOF profiles to change over such intervals of time. Hence, we consider the first TOF profile during each 10-second interval, and subtract it from all subsequent TOF profiles during that interval. As a result, breathing users' reflections pop up at different instances, allowing us to detect and localize them.

8 IMPLEMENTATION

We built WiTrack2.0 using a single FMCW radio whose signal is fed into multiple antennas. The FMCW radio generates a low-power (sub-milliWatt) signal that sweeps 5.46-7.25 GHz every 2.5 milliseconds. The range and power are chosen in compliance with FCC regulations for consumer electronics [2].

The schematic in Fig. 5 shows how we use this radio to implement Multi-shift FMCW. Specifically, the generated sweep is delayed before being fed to directional antennas for transmission.[6] At the receive side, the signal from each receive antenna is mixed with the FMCW signal and the resulting signal is fed to the USRP. The USRP samples the signals at 2 MHz and transfers the digitized samples to the UHD driver. These samples are processed in software to localize users and recognize their gestures.[7]

The analog FMCW radio and all the USRPs are driven by the same external clock. This ensures that there is no frequency offset between their oscillators, and hence enables subtracting frames that are relatively far apart in time to enable localizing people based on breathing.

9 EVALUATION

Human Subjects. We evaluate the performance of WiTrack2.0 by conducting experiments in our lab with

---

[6]The most straightforward option to delay the signal is to insert a wire. However, wires attenuate the signal and introduce distortion over the wide bandwidth of operation of our system, reducing its SNR. Instead, we exploit the fact that, in FMCW, time and frequency are linearly related; hence, a shift τ in time can be achieved through a shift $\Delta f = slope \times \tau$ in the frequency domain. Hence, we achieve this delay by mixing FMCW with signals whose carrier frequency is Δf. This approach also provides us with the flexibility of tuning multi-shift FMCW for different TOF_{limit}'s by simply changing these carrier frequencies.

[7]Complexity-wise, WiTrack2.0's algorithms are linear in the number of users, the number of Tx antennas, and the number of Rx antennas.

eleven human subjects: four females and seven males. The subjects differ in height from 165–185 cm as well as in weight and build and span 20 to 50 years of age. In each experiment, each subject is allowed to move as they wish throughout the room. These experiments were approved by MIT IRB protocol #1403006251.

Ground Truth. We use the VICON motion capture system to provide us with ground truth positioning information [1]. It consists of an array of infrared cameras that are fitted to the ceiling of a $5\ m \times 7\ m$ room, and requires instrumenting any tracked object with infrared-reflective markers. When an instrumented object moves, the system tracks the infrared markers on that object and fits them into a 3D model to identify the object's location.

We evaluate WiTrack2.0's accuracy by comparing it to the locations provided by the VICON system. To track a user using the VICON, we ask him/her to wear a hard hat that is instrumented with five infrared markers. In addition, for the gestures experiments, we ask each user to wear a glove that is instrumented with six markers.

Experimental Setup. We evaluate WiTrack2.0 in a standard office environment that has standard furniture: tables, chairs, boards, computers, etc. We experiment with two setups: line-of-sight and through-the-wall. In the through-wall experiments, WiTrack2.0 is placed outside the VICON room with all transmit and receive antennas facing one of the walls of the room. Recall that WiTrack2.0's antennas are directional and hence this setting means that the radio beam is directed toward the room. The VICON room has no windows; it has 6-inch hollow walls supported by steel frames, which is a standard setup for office buildings. In the line-of-sight experiments, we move WiTrack2.0 inside the room. In all of these experiments, the subjects' locations are tracked by both the VICON system and WiTrack2.0.

Detection. Recall that WiTrack2.0 adopts iterative cancellation to detect different users in the scene. This limits the number of users it can detect because of residual interference from previous iterations. Therefore, we run experiments to identify the maximum number of people that WiTrack2.0 can reliably detect under various conditions. Detection accuracy is measured as the percentage of time that WiTrack2.0 correctly outputs the number of users present in the environment. The number of users in each experiment is known and acts as the ground truth. We run ten experiments for each of our testing scenarios, and plot the accuracies for each them in Fig. 11.

We make two observations from this figure. First, the accuracy of detection is higher in line-of-sight than in through-wall settings. This is expected because the wall causes significant attenuation and hence reduces the SNR of the reflected signals. Second, the detection accuracy for breathing-based localization is higher than that of the tracking experiments. While this might seem surprising,
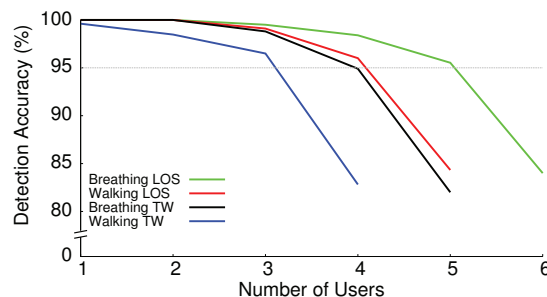
Figure 11—WiTrack2.0 's Detection Accuracy. The figure shows the percentage of time that WiTrack2.0 accurately determines the number of people in each of our evaluation scenarios.

it is actually due to the fact that the breathing experiments operate over longer subtraction windows. Specifically, the system outputs the number of people detected and their locations by analyzing the trace over windows of 10 seconds. In contrast, the tracking experiments require outputting a location of each person once every 12.5 ms,[8] and hence they might not be able to detect each person within such a small time window.

For our evaluation of localization accuracy, we run experiments with the maximum number of people that are reliably detectable, where "reliably detectable" is defined as detected an accuracy of 95% or higher. For reference, we summarize these numbers in the table below.

	Line-of-Sight	Through-Wall
Motion Tracking	4	3
Breathing-based Localization	5	4

Table 1—Maximum Number of People Detected Reliably.

10 PERFORMANCE RESULTS
10.1 Accuracy of Multi-Person Motion Tracking

We first evaluate WiTrack2.0's accuracy in multi-person motion tracking. We run 100 experiments in total, half of them in line-of-sight and the second half in through-wall settings. In each experiment, we ask one, two, three, or four human subjects to wear the hard hats that are instrumented with VICON markers and move inside the VICON-instrumented room. Each subject's location is tracked by both the VICON system and WiTrack2.0, and each experiment lasts for one minute. Since each FMCW sweep lasts for 2.5ms and we average 5 sweeps to obtain each TOF measurement, we collect around 5,000 location readings per user per experiment.

Figs. 12 and 13 plot the CDFs of the location error along the x and y coordinates for each of the localized persons in both line-of-sight and through-wall scenarios. The subjects are ordered from the first to the last as detected by SSC. The figures reveal the following findings:

[8]Since the user is moving, combining measurements over a longer interval smears his signal.
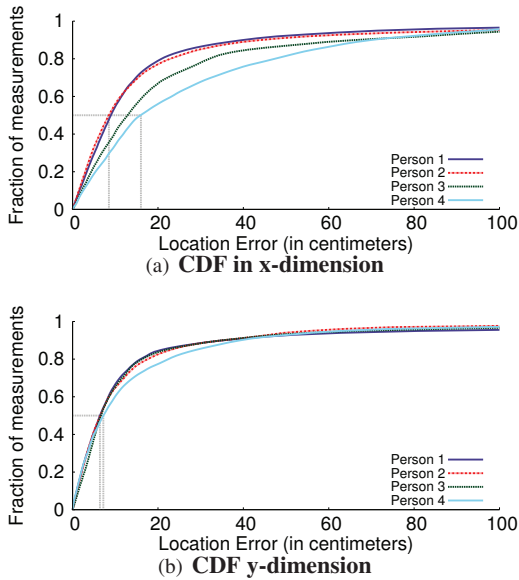
Figure 12—**Performance of WiTrack2.0's LOS Tracking.** (a) and (b) show the CDFs of the location error in x and y for each of the tracked users in LOS. Subjects are ordered from first to last detected by SSC.

- WiTrack2.0 can accurately track the motion of four users when it is in the same room as the subjects. Its median location error for these experiments is 8.5 cm in x and 6.4 cm in y for the first user detected, and decreases to 15.9 cm in x and 7.2 cm in y for the last detected user.
- In through-wall scenarios, WiTrack2.0 can accurately localize up to three users. Its median location error for these experiments is 8.4 cm and 7.1 cm in x/y for the first detected user, and decreases to 16.1 cm and 10.5 cm in x/y for the last detected user. As expected, the accuracy when the device is placed in the same room as the users is better than when it is placed behind the wall due to the extra attenuation (reduced SNR) caused by the wall.
- The accuracy in the y dimension is better than the accuracy in the x dimension. This discrepancy is due to the asymmetric nature of WiTrack2.0's setup, where all of its antennas are arranged along the $y = 0$ axis.
- The localization accuracy decreases according to the order the SSC algorithm localizes the users. This is due to multiple reasons: First, a user detected in later iterations is typically further from the device, and hence has lower SNR, which leads to lower accuracy. Second, SSC may not perfectly remove the reflections of other users in the scene, which leads to residual interference and hence lower accuracy.

10.2 Accuracy of Breathing-based Localization

We evaluate WiTrack2.0's accuracy in localizing static people based on their breathing. We run 100 experiments in total with up to five people in the room. Half of these experiments are done in line-of-sight and the other half
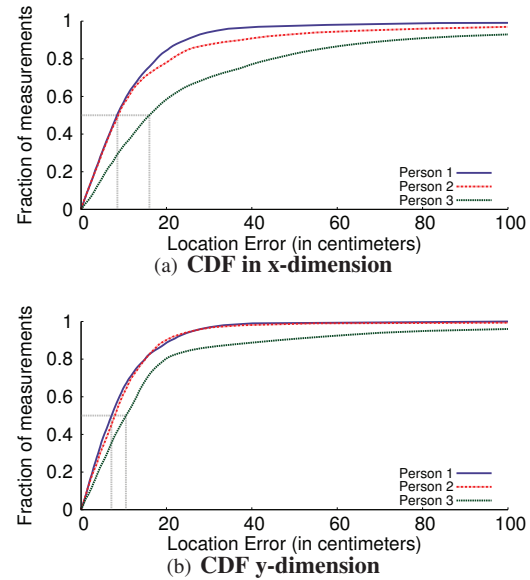
Figure 13—**Performance of WiTrack2.0's Through-Wall Tracking.** (a) and (b) show the CDFs of the location error in x and y for each of the tracked users. Subjects are ordered from first to last detected by SSC.
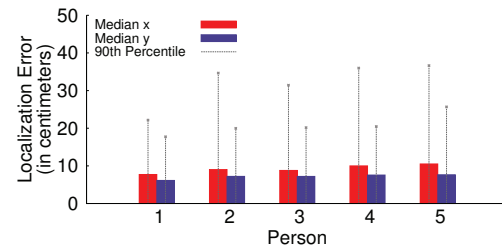
Figure 14—**Accuracy for Localizing Breathing People in Line-of-Sight..** The figure shows show the median and 90th percentile errors in x/y location. Subjects are ordered from first to last detected by SSC.

are through-wall. Experiments last for 3-4 minutes. Subjects wear hard hats and sit on chairs in the VICON room.

Figs. 14 and 15 plot WiTrack2.0's localization error in line-of-sight and through-wall settings as a function of the order with which the subject is detected by the SSC algorithm. The figures show the median and 90th percentile of the estimation error for the x and y coordinates of each of the subjects. The figures show the following results:

- WiTrack2.0's breathing-based localization accuracy goes from a median of 7.24 and 6.3 cm in x/y for the nearest user to 18.31 and 10.85 cm in x/y for the furthest user, in both line-of-sight and through-wall settings.
- Localization based on breathing is more accurate than during motion. This is because when people are static, they remain in the same position, providing us with a larger number of measurements for the same location.

10.3 Accuracy of 3D Pointing Gesture Detection

We evaluate WiTrack2.0's accuracy in tracking 3D pointing gestures. We run 100 experiments in total with one to three subjects. In each of these experiments, we ask each subject to wear a glove that is instrumented with
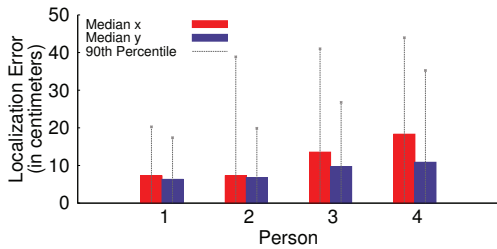
Figure 15—Accuracy for Localizing Breathing People in Through-Wall Experiments.. The figure shows show the median and 90th percentile errors in x/y location. Subjects are ordered from first to last detected by SSC.

(a) **Pointing Accuracy in ϕ**

(b) **Pointing Accuracy in θ**
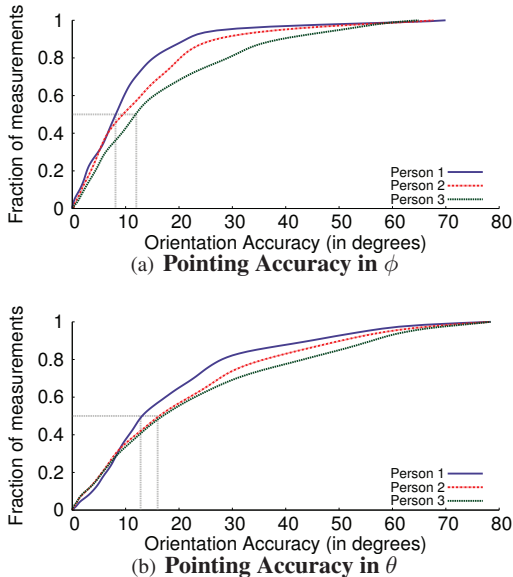
Figure 16—3D Gesture Accuracy. The figure shows the CDFs of the orientation accuracy for the pointing gestures of each participant. Subjects are ordered from first to last detected by the SSC algorithm.

infrared-reflective markers, stand in a different location in the VICON room, and point his/her hand in a random 3D direction of their choice – as if they were playing a shooting game or pointing at some household appliance to control it. In most of these experiments, all subjects were performing the pointing gestures simultaneously.

Throughout these experiments, we track the 3D location of the hand using the VICON system and WiTrack2.0. We then regress on the 3D trajectory to determine the direction in which each user pointed (similar to [3]). Fig. 16(a) and 16(b) plot the CDFs of the orientation error between the angles as measured by WiTrack2.0 and the VICON for the 1st, 2nd and 3rd participant (in the order of detection by SSC). Note that we decompose the 3D pointing gesture along two directions: azimuthal (in the $x - y$ plane), which we denote as ϕ, and elevation (in the $r-z$ plane), which we denote as θ. The accuracy along both of these angles is important since appliances which the user may want to control in a home environment (e.g., lamps, screens, shades) span the 3D space.

The figure shows that the median orientation error in

ϕ goes from 8.2 degrees to 12.4 degrees from the first to the third person, and from 12 degrees to 16 degrees in θ. Note that WiTrack2.0's accuracy in ϕ is slightly higher than its accuracy in θ. This is due to WiTrack2.0's setup, where the antennas are more spread out along the x than along the z, naturally leading to lower robustness to errors along the z axis, and hence lower accuracy in θ. These experiments demonstrate that WiTrack2.0 can achieve high accuracy in 3D tracking of a pointing gesture.

11 DISCUSSION & LIMITATIONS

WiTrack2.0 marks an important step toward enabling accurate indoor localization that does not require users to hold or wear any wireless device. WiTrack2.0, however, has some limitations that are left for future work.

1. *Number of Users:* WiTrack2.0 can accurately track up to 4 moving users and 5 static users. These numbers may be sufficient for in-home tracking. However, it is always desirable to scale the system to track more users.

2. *Coverage Area:* WiTrack2.0's range is limited to 10m due to its low power. To cover larger areas and track more users, one may deploy multiple devices and hand off the trajectory tracking from one to the next, as the person moves around. Managing such a network of devices, coordinating their hand-off, and arbitrating their medium access are interesting problems to explore.

3. *Lack of Identification:* The system can track multiple users simultaneously, but it cannot identify them. Additionally, it can track limb motion (e.g., a hand) but cannot differentiate between different body parts (a hand vs. a leg). We believe that future work can investigate this issue by identifying fingerprints of various reflectors.

4. *Limited Gesture Interface:* WiTrack2.0 focuses on tracking pointing gestures; however, the user cannot move other body parts while performing the pointing gesture. Extending the system to enable rich gesture-based interfaces is an interesting avenue for future work.

Overall, we believe WiTrack2.0 pushes the limits of indoor localization and enriches the role it can play in our daily lives. By enabling smart environments to accurately follow our trajectories, it paves way for these environments to learn our habits, react to our needs, and enable us to control the Internet of Things that revolves around our networked homes and connected environments.

References

[1] VICON T-Series. `http://www.vicon.com`. VICON.

[2] *Understanding the FCC Regulations for Low-power, Non-licensed Transmitters*. Office of Engineering and Technology Federal Communications Commission, 1993.

[3] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller. 3D Tracking via Body Radio Reflections. In *Usenix NSDI*, 2014.

[4] F. Adib and D. Katabi. See through walls with Wi-Fi! In *ACM SIGCOMM*, 2013.

[5] P. Bahl and V. N. Padmanabhan. Radar: An in-building RF-based user location and tracking system. In *IEEE INFOCOM*, 2000.

[6] D. Bharadia, E. McMilin, and S. Katti. Full duplex radios. In *SIGCOMM*. ACM, 2013.

[7] M. Bocca, O. Kaltiokallio, N. Patwari, and S. Venkatasubramanian. Multiple target tracking with RF sensor networks. *Mobile Computing, IEEE Transactions on*, 2013.

[8] R. Bohannon. Comfortable and maximum walking speed of adults aged 20-79 years: reference values and determinants. *Age and ageing*, 1997.

[9] P. K. Chan, W. Jin, J. Gong, and N. Demokan. Multiplexing of fiber bragg grating sensors using a fmcw technique. *IEEE Photonics Technology Letters*, 1999.

[10] G. Charvat, L. Kempel, E. Rothwell, C. Coleman, and E. Mokole. An ultrawideband (UWB) switched-antenna-array radar imaging system. In *IEEE ARRAY*, 2010.

[11] K. Chintalapudi, A. Padmanabha Iyer, and V. N. Padmanabhan. Indoor localization without the pain. In *ACM MobiCom*. ACM, 2010.

[12] S. Gollakota and D. Katabi. Zigzag decoding: combating hidden terminals in wireless networks. In *ACM SIGCOMM*, 2008.

[13] S. Hantscher, A. Reisenzahn, and C. Diskus. Through-wall imaging with a 3-d uwb sar algorithm. *Signal Processing Letters, IEEE*, 2008.

[14] J. Högbom. Aperture synthesis with a non-regular distribution of interferometer baselines. *Astronomy and Astrophysics Supplement Series*, 1974.

[15] Y. Huang, P. V. Brennan, D. Patrick, I. Weller, P. Roberts, and K. Hughes. Fmcw based mimo imaging radar for maritime navigation. *Progress In Electromagnetics Research*, 2011.

[16] Y. Jia, L. Kong, X. Yang, and K. Wang. Through-wall-radar localization for stationary human based on life-sign detection. In *IEEE RADAR*, 2013.

[17] K. Joshi, S. Hong, and S. Katti. Pinpoint: Localizing interfering radios. In *Usenix NSDI*, 2013.

[18] S. Nannuru, Y. Li, Y. Zeng, M. Coates, and B. Yang. Radio-frequency tomography for passive indoor multitarget tracking. *Mobile Computing, IEEE Transactions on*, 2013.

[19] N. Patwari, L. Brewer, Q. Tate, O. Kaltiokallio, and M. Bocca. Breathfinding: A wireless network that monitors and locates breathing in a home. *Selected Topics in Signal Processing, IEEE Journal of*, 2014.

[20] Q. Pu, S. Jiang, S. Gollakota, and S. Patel. Whole-home gesture recognition using wireless signals. In *ACM MobiCom*, 2013.

[21] A. Rai, K. K. Chintalapudi, V. N. Padmanabhan, and R. Sen. Zee: zero-effort crowdsourcing for indoor localization. In *ACM MobiCom*, 2012.

[22] S. Rallapalli, A. Ganesan, K. Chintalapudi, V. N. Padmanabhan, and L. Qiu. Enabling physical analytics in retail stores using smart glasses. In *ACM MobiCom*, 2014.

[23] T. Ralston, G. Charvat, and J. Peabody. Real-time through-wall imaging using an ultrawideband multiple-input multiple-output (MIMO) phased array radar system. In *IEEE ARRAY*, 2010.

[24] A. Saeed, A. Kosba, and M. Youssef. Ichnaea: A low-overhead robust wlan device-free passive localization system. *Selected Topics in Signal Processing, IEEE Journal of*, 2014.

[25] M. Seifeldin, A. Saeed, A. Kosba, A. El-Keyi, and M. Youssef. Nuzzer: A large-scale device-free passive localization system for wireless environments. *Mobile Computing, IEEE Transactions on*, 2013.

[26] S. Sen, B. Radunovic, R. R. Choudhury, and T. Minka. Spot localization using phy layer information. In *ACM MobiSys*, 2012.

[27] P. Setlur, G. Alli, and L. Nuzzo. Multipath exploitation in through-wall radar imaging via point spread functions. *Image Processing, IEEE Transactions on*, 2013.

[28] G. E. Smith and B. G. Mobasseri. Robust through-the-wall radar image classification using a target-model alignment procedure. *Image Processing, IEEE Transactions on*, 2012.

[29] A. R. Thompson, J. M. Moran, and G. W. Swenson Jr. *Interferometry and synthesis in radio astronomy*. John Wiley & Sons, 2008.

[30] J. Tsao and B. D. Steinberg. Reduction of side-lobe and speckle artifacts in microwave imaging: the clean technique. *Antennas and Propagation, IEEE Transactions on*, 1988.

[31] A. Vasilyev. *The optoelectronic swept-frequency laser and its applications in ranging, three-dimensional imaging, and coherent beam combining of chirped-seed amplifiers*. PhD thesis, 2013.

[32] H. K. Walker, W. D. Hall, and J. W. Hurst. *Respiratory Rate and Pattern*. Butterworths, 1990.

[33] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni. We can hear you with wi-fi! In *ACM MobiCom*, 2014.

[34] J. Wang and D. Katabi. Dude, where's my card? RFID positioning that works with multipath and non-line of sight. In *ACM SIGCOMM*, 2013.

[35] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu. E-eyes: device-free location-oriented activity identification using fine-grained wifi signatures. In *ACM MobiCom*. ACM, 2014.

[36] J. Wilson and N. Patwari. Radio tomographic imaging with wireless networks. In *IEEE Transactions on Mobile Computing*, 2010.

[37] P. W. Wolniansky, G. J. Foschini, G. Golden, and R. Valenzuela. V-blast: An architecture for realizing very high data rates over the rich-scattering wireless channel. In *IEEE ISSSE*, 1998.

[38] J. Xiong and K. Jamieson. ArrayTrack: a fine-grained indoor location system. In *Usenix NSDI*, 2013.

[39] Y. Xu, S. Wu, C. Chen, J. Chen, and G. Fang. A novel method for automatic detection of trapped victims by ultrawideband radar. *Geoscience and Remote Sensing, IEEE Transactions on*, 2012.

[40] M. Youssef, M. Mah, and A. Agrawala. Challenges: device-free passive localization for wireless environments. In *ACM MobiCom*, 2007.