IndoTrack: Device-Free Indoor Human Tracking with Commodity Wi-Fi

XIANG LI and DAQING ZHANG, Peking University
QIN LV, University of Colorado Boulder
JIE XIONG, Singapore Management University
SHENGJIE LI, YUE ZHANG, and HONG MEI, Peking University

Indoor human tracking is fundamental to many real-world applications such as security surveillance, behavioral analysis, and elderly care. Previous solutions usually require dedicated device being carried by the human target, which is inconvenient or even infeasible in scenarios such as elderly care and break-ins. However, compared with device-based tracking, device-free tracking is particularly challenging because the much weaker reflection signals are employed for tracking. The problem becomes even more difficult with commodity Wi-Fi devices, which have limited number of antennas, small bandwidth size, and severe hardware noise.

In this work, we propose IndoTrack, a device-free indoor human tracking system that utilizes only commodity Wi-Fi devices. IndoTrack is composed of two innovative methods: (1) Doppler-MUSIC is able to extract accurate Doppler velocity information from noisy Wi-Fi Channel State Information (CSI) samples; and (2) Doppler-AoA is able to determine the absolute trajectory of the target by jointly estimating target velocity and location via probabilistic co-modeling of spatial-temporal Doppler and AoA information. Extensive experiments demonstrate that IndoTrack can achieve a 35cm median error in human trajectory estimation, outperforming the state-of-the-art systems and provide accurate location and velocity information for indoor human mobility and behavioral analysis.

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

Additional Key Words and Phrases: Wi-Fi, Doppler, Device-free indoor tracking

ACM Reference format:
Xiang Li, Daqing Zhang, Qin Lv, Jie Xiong, Shengjie Li, Yue Zhang, and Hong Mei. 2017. IndoTrack: Device-Free Indoor Human Tracking with Commodity Wi-Fi. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 72 (September 2017), 22 pages.
DOI: http://doi.org/10.1145/3130940

This work is supported by National Key Research and Development Plan under Grant No. 2016YFB1001200 and Peking University Information Technology Institute (Tianjin Binhai).
Author's addresses: X. Li, D. Zhang, S. Li, Y. Zhang and H. Mei, Key Laboratory of High Confidence Software Technologies (Ministry of Education), School of Electronics Engineering and Computer Science, Peking University, Beijing, China 100871; email: {lixiang13, dqzsei, lishengjie, zy.zhangyue, meih}@pku.edu.cn. Q. Lv, Department of Computer Science, University of Colorado Boulder, Boulder, CO 80309 USA; email: qin.lv@colorado.edu. J. Xiong, School of Information Systems, Singapore Management University, Singapore; email: jxiong@smu.edu.sg. Corresponding Author: Daqing Zhang; email: dqzsei@pku.edu.cn.

2474-9567/2017/9-ART72 $15.00
DOI: http://doi.org/10.1145/3130940

1 INTRODUCTION

Indoor human tracking is a fundamental component required by a wide range of real-world applications including security surveillance, physical/behavioral analysis, elderly/patient monitoring, indoor navigation, workspace interaction, indoor space design, etc. Over the years, various solutions have been proposed for indoor human tracking. Most of them require dedicated devices to be carried by the human target, such as RFID tags [44], mobile phones [4, 8, 43], or wearable devices [18, 19], which is inconvenient and even infeasible under certain scenarios. For instance, the elderly are reluctant to carry devices all the time [33]; and terrorists would avoid carrying any device that can be tracked. As such, device-free indoor human tracking is urgently desirable and has attracted attentions from researchers.

Existing camera [10] or sound-based [27, 39, 47] solutions are mostly device-free, and have been shown to be effective in certain scenarios such as gesture recognition. However, camera-based systems require dense deployment and raise severe privacy concerns in indoor environments. Sound-based methods have a very small coverage area and the performance degrades significantly in noisy environments, limiting their real-life applications. In contrast, Wi-Fi based solutions are promising as Wi-Fi is ubiquitous and no extra infrastructure is required. However, the existing Wi-Fi based tracking systems [41, 42] are mostly device-based and there are very few device-free indoor tracking systems. In this work, our goal is to achieve accurate passive indoor human tracking with only commodity Wi-Fi devices.

Device-free tracking with commodity Wi-Fi devices is challenging. Compared with device-based tracking, the reflection signals employed for device-free tracking are much weaker than the direct-path signal so it is more difficult to extract useful and accurate tracking information from the reflection signals. Furthermore, most commodity Wi-Fi cards only have three antennas and limited bandwidth, making the angle-of-arrival (AoA) and time-of-arrival (ToA) based passive tracking too coarse [26]. Therefore, in this work, we aim to tackle the human tracking problem with Doppler information, whose resolution is not limited by the number of antennas or bandwidth size.

The basic idea and theoretical underpinning of our Wi-Fi based device-free tracking system are as follows: in an indoor environment with one transmitter and two receivers, human movements change the path length of the Wi-Fi signal reflected from the human body for each pair of transceivers, resulting in Doppler frequency shift of the signal. Both amplitude and direction of the Doppler frequency shift can be estimated with the phase information of the Channel State Information (CSI) available at the commodity Wi-Fi receiver. The amount of Doppler frequency shift is dependent on how fast the reflection path length changes, which is thus related to the human target's velocity and location. Furthermore, from CSI, we can obtain the AoA spectrum, which represents the location (angle) probability of the target. By combining Doppler frequency shift with the AoA spectrum, we are then able to estimate the human velocity, location, and thus moving trajectory, achieving our goal of accurate device-free human tracking with commodity Wi-Fi devices.

A number of challenges need to be addressed before the idea can be realized on commodity Wi-Fi devices. First, the amount of Doppler frequency shift caused by human movement is very small (e.g., less than 100 Hz for a 5 GHz Wi-Fi channel) compared with the carrier frequency (5 GHz). Second, commodity Wi-Fi receivers are not tightly synchronized with the transmitter in terms of carrier frequency and time, causing a random phase offset in each CSI sample. If this random phase offset is not handled properly, the accuracy of Doppler frequency shift estimation is greatly affected. Third, even with accurate estimation of the Doppler frequency shift of the reflection signal, there is no direct translation from Doppler frequency shift to target velocity, since the target's velocity also depends on the current location of the target which is unknown. Last, even with target velocity estimated, without knowing the accurate starting point, the absolute trajectory still can not be obtained.

In this work, we propose IndoTrack, a device-free tracking system that utilizes only three commodity Wi-Fi devices: one transmitter and two receivers. IndoTrack addresses all the above mentioned challenges with two innovative methods. Specifically, this work makes the following contributions:

- We propose **Doppler-MUSIC** to estimate the Doppler velocity (i.e., the change speed of the path length) from Doppler frequency shift with CSI samples collected from commodity Wi-Fi devices. Our method intelligently utilizes information from two colocated antennas on a same Wi-Fi card to remove the random CSI phase offset introduced by sampling frequency offset, packet detection delay, and the carrier frequency offset between the transmitter and receiver. We also remove the strong direct path signal to increase the accuracy of Doppler estimation of the weak reflection signal. Together, these strategies allow us to capture accurate CSI phase changes in time domain, and then process the CSI phase information with the MUSIC algorithm to obtain the Doppler velocity.
- We propose **Doppler-AoA**, a probabilistic, space-time joint trajectory estimation method. The Doppler-AoA method combines Doppler velocity and AoA spectrum of the target reflection path to estimate the absolute location and project Doppler velocity into target velocity, thus determining the absolute starting point and achieving accurate human tracking in indoor environments.
- We design and implement IndoTrack on commodity Intel 5300 Wi-Fi cards and conduct comprehensive experiments in different indoor environments covering typical human movement patterns. Moreover, we conduct an in-situ experiment to test the feasibility of IndoTrack for daily human trajectory recording. Our results demonstrate that IndoTrack can achieve highly accurate trajectory estimations (35cm median error), and present accurate velocity and location information for high-level physical and behavioral analysis of human movements in indoor environments.

The rest of the paper is organized as follows. Section 2 discusses related work and Section 3 gives an overview of our system design. Section 4 and Section 5 describe in detail the Doppler-MUSIC method for estimating Doppler velocity and the Doppler-AoA method for space-time joint estimation of absolute human trajectory. Section 6 presents the evaluation results. Section 7 discusses some limitations and future directions followed by a conclusion in Section 8.

2 RELATED WORK

Our work is broadly related to research in the areas of indoor human tracking, indoor localization and gesture recognition. Over the years, many different location and tracking technologies have been developed including radio frequency [2, 44], inertial sensors [7, 40], camera [10], sound [14, 27, 39], infrared [16, 17], and visible light [20, 24]. Given the large amount of work in these areas, an exhaustive survey is beyond the scope of this paper. Inertial sensor-based methods fall in the device-based category. Camera-based solutions require good lighting conditions and raise privacy concerns. Sound-based solutions are vulnerable to environmental noise and the coverage area is very small. Infrared systems require dedicated infrastructure to be installed which are expensive. Visible light-based solutions require strict line-of-sight (LoS) condition and do not work in dark environment. In this work, we focus on device-free human tracking using commodity Wi-Fi devices and we only discuss the most related research works here.

2.1 Indoor Human Tracking

Indoor human tracking has been employed to monitor and study people's behaviours, such as elderly and disabled people's daily routines [11, 12, 22], people's habits and interactions in workspace [9, 28]. Earlier solutions require either a special purpose device to be carried by the user, or extensive instrumentation of the environment. For example, Au et al. proposed a tracking and navigation system using RSS (Received Signal Strength) and

compressive sensing on a mobile device [4]. Hardegger et al. utilized mobile and wearable sensors to track locations and associate them with human activities [18, 19].

On the other hand, there is very limited work on Wi-Fi-based device-free human tracking. WiDeo [21] is implemented on the WARP software-defined radio platform and achieves a fine-grained motion tracing. However, hardware modification is required to implement it on commodity Wi-Fi devices. Wi-Vi [3] relies on a special device that can capture the reflections of its own transmitted signals reflected off moving objects behind a wall in order to track them, and it tracks only the relative movement but not be able to obtain the absolute location. The most relevant work is WiDar [30], which also tracks human movement by estimating the Doppler frequency shift. WiDar estimates the Doppler frequency shift with CSI amplitude information using FFT method, which is similar to [37, 38]. However, the CSI amplitude is known to be coarse with a large variation [35]. Furthermore, the CSI amplitudes do not provide direction information of the Doppler velocity. As such, WiDar requires multiple Wi-Fi links with significantly different spatial features in order to determine the moving direction and trajectory location. Thus, for WiDar to work with only two receivers, the antennas of the Wi-Fi receiver need to be separated far away from each other, which is difficult for commodity Wi-Fi devices. Adding a feeder line to the antenna to increase the distance would significantly decrease the quality of the received signal and affect the data communication. In comparison, our system employs the more stable phase information of CSI and proposes an innovative Doppler-MUSIC method to estimate the accurate velocity with direction information. Moreover, IndoTrack proposes a Doppler-AoA method that combines the Doppler velocity and AoA spectrum to estimate the absolute starting position and absolute trajectory. IndoTrack can work with only two receivers and does not require the antennas to be separated by a large distance which is fully compatible with commodity Wi-Fi devices.

2.2 Gesture Recognition and Gesture Tracking

Human gesture recognition and gesture tracking have attracted a lot of attention in the research community of mobile communication and human-computer interaction. While the former focuses on recognizing a set of pre-defined gestures [1, 25, 29], the latter aims to track hand movements in the air [34, 36]. For instance, WiSee achieves whole-home gesture recognition using software-defined radios (USRP), and uses FFT to detect Doppler shifts [29]. RF-IDraw builds a virtual touch screen using RFID and utilizes AoA information for fine-grained tracking [36]. WiDraw utilizes AoA information and needs dense Wi-Fi transmitters in the environment to cover all directions [34]. A number of acoustic solutions have also been proposed for gesture recognition and gesture tracking [27, 39, 47]. They generally leverage the Doppler shift in the acoustic frequency band, and augment that with inertial sensors such as accelerometer and gyroscope to further improve the accuracy.

2.3 Indoor Localization

Indoor localization has been an area of active research in the last decade. For Wi-Fi-based localization, both device-based and device-free solutions have been developed [5, 23, 26, 32, 35, 41, 45, 46]. Earlier solutions are mostly based on RSSI (Received Signal Strength Indicator) and achieve meter-level localization accuracies [5]. More recent solutions utilize CSI (Channel State Information) and AoA to localize a target within tens of centimeters. Most of these solutions either require the target to hold a transmitting device, or require dense deployment in the environment, and their localization accuracies deteriorate significantly when the target is moving. Our recent work Dynamic-MUSIC addresses the moving target localization problem by separating incoherent dynamic path signals from coherent static path signals [26]. Still, dense AP deployment is needed and the system is not able to track continuous movement of human target. Instead, in this work, we aim to tackle the human tracking problem by jointly estimating human velocity and location at each timestamp continuously.

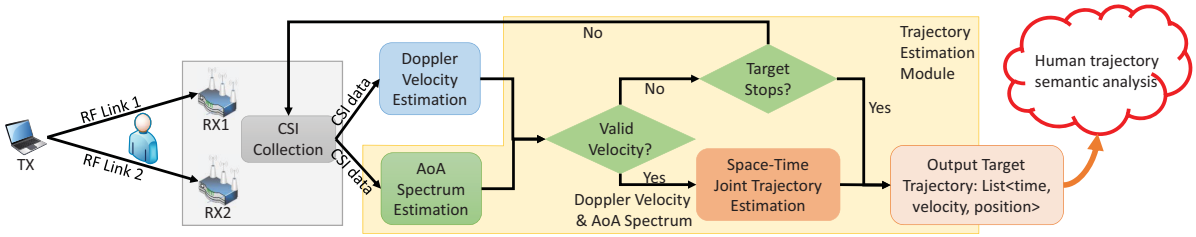


Fig. 1. System overview: to accomplish device-free human tracking with commodity Wi-Fi devices, we propose two innovative methods: Doppler-MUSIC estimates the Doppler velocity (i.e., the change speed of path length) employing CSI phase information; Doppler-AoA couples Doppler velocity and AoA spectrum for probabilistic space-time joint estimation of the human absolute trajectory.

3 SYSTEM OVERVIEW

Figure 1 illustrates the overall framework of IndoTrack. IndoTrack accurately estimates the human velocity at each time point, and then couples the velocity information with angle information for human tracking. IndoTrack only employs one Wi-Fi transmitter (TX) and two receivers (RX), and leverages Wi-Fi CSI samples available at commodity Wi-Fi devices for estimation. IndoTrack achieves the objective with a two-step design employing novel methods proposed in this work. Here is a brief description of the workflow for IndoTrack:

(1) The receivers receive Wi-Fi signals from the transmitter and record the CSI of each packet.
(2) With the recorded CSI samples, IndoTrack employs the proposed Doppler-MUSIC method to estimate the speed of the moving human's reflection path length change. The Doppler-MUSIC method addresses the varying intervals issue of CSI samples and the random CSI phase offsets problem due to the unsynchronized Wi-Fi transmitter and receivers. It includes a number of novel steps to accurately measure the subtle Doppler shift introduced by human movements.
(3) When a non-zero Doppler velocity is estimated (i.e., human is moving), IndoTrack leverages Doppler-AoA, another method proposed in this work, to map the Doppler velocity to human velocity and obtain velocity and location information at the same time. Doppler-AoA works via probabilistic joint estimation of human velocity and location with space-time constraints of Doppler velocity and AoA spectrum. The AoA spectrum is estimated from the recorded CSIs with method proposed in our previous work [26] and indicates the moving target's angles to the two receivers.
(4) The system continuously estimates the human trajectory with the methods above. When the target stops moving, such as sitting on the sofa or sleeping on the bed, no valid velocity will be detected by IndoTrack. In such a scenario, IndoTrack will output the estimated trajectory and prepare for the estimation of next trajectory segment.

The output trajectory generated by IndoTrack is a series of $< time, location, velocity >$ tuples, which can be used for comprehensive physical/behavioral analysis of high-level semantics. By knowing if and when a person is moving, how fast a person is moving, and the locations where a person is moving from/to, we can gain useful insights about a target's activities of daily living, identify routine vs. abnormal behaviours (e.g., scheduled meal times vs. delayed meals, missed activities such as taking medication or exercising, level of activeness, any deteriorating physical mobility, etc.).

In the next two sections, we will present in detail how to estimate the Doppler velocity with commodity Wi-Fi CSI samples (Doppler-MUSIC in Section 4), and how to jointly estimate human location and velocity to obtain the trajectory information (Doppler-AoA in Section 5).

4 DOPPLER-MUSIC FOR DOPPLER VELOCITY ESTIMATION

In this section, we describe in detail how we estimate Doppler velocity (i.e., the change speed of the path length) using CSI samples collected from commodity Wi-Fi devices. We first introduce how the subtle Doppler frequency shift caused by human movements can be estimated. Then, we propose a MUSIC-based algorithm to estimate the Doppler velocity from CSI samples with varying intervals. Meanwhile, we combine information from many subcarriers to address the influence of frequency selective fading. At last, we discuss in detail how to handle the random CSI phase offsets on commodity Wi-Fi devices.

4.1 Extracting Doppler Frequency Shift from CSI

As illustrated in Figure 2, in a typical indoor environment with a pair of Wi-Fi transmitter and receiver, the Wi-Fi signal is not only propagated along the direct path, but also reflected by other objects such as the human body and walls. The signal received at the receiver is thus a superposition of signals from all the paths. This phenomenon is called multipath propagation. When a person moves in the environment, the path length of the human reflection signal changes accordingly, which introduces a Doppler frequency shift to the carrier frequency of the reflected signal:

$$f_{Doppler} = f\frac{v_{path}}{c},$$

where f is the original carrier frequency of the signal, v_{path} is the speed of path length change, and c is the propagation speed of the Wi-Fi signal in the air. Note that when the human is moving, the introduced Doppler frequency shift to the Wi-Fi signal is only tens of hertz for a 5 GHz channel [38]. Detecting such a subtle Doppler frequency shift from a much large carrier frequency is difficult.

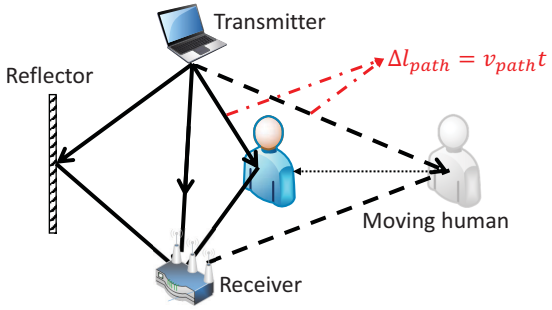


Fig. 2. Human movement leads to changes in the length of the human refection path, and the speed of this path length change can be estimated by Doppler frequency shift.

Fortunately, the CSI information available on commodity Wi-Fi devices presents us a chance to estimate the Doppler frequency shift with only a several hundred hertz of sampling rate. For a Wi-Fi signal, the CSI represents the amplitude attenuation and phase change in each subcarrier due to signal propagation from the transmitter to the receiver. Considering only one signal, the CSI of the signal at time t_0 is $x(f, t_0) = A_0 e^{-j2\pi f \tau_0}$, where A_0 is the attenuation and τ_0 is the propagation delay. If the propagation path length changes at a speed of v and we ignore the attenuation change, after a short time period t, the path length change is $\Delta l_{path} = vt$ and the propagation delay change is $\Delta\tau = \frac{vt}{c}$. So the CSI of the signal is $x(f, t_0 + t) = A_0 e^{-j2\pi f(\tau_0 + \frac{vt}{c})} = x(f, t_0)e^{-j2\pi f \frac{vt}{c}}$. The phase change rate of CSI represents the Doppler frequency shift of the signal. When multipath reflections in the real

world are considered, the CSI recorded for each packet can be represented as follows:

$$x(f, t_0 + t) = \sum_{i=1}^{L} A_i e^{-j2\pi f(\tau_i + \frac{v_i t}{c})}, \tag{1}$$

where L is the number of propagation paths, τ_i is the propagation delay of the i^{th} path signal at time t_0, and v_i is the i^{th} path's length change speed. Based on Equation 1, the CSI can be treated as the superposition of L "signals" and the phase change rate of the i^{th} "signal" is the Doppler frequency shift ($f\frac{v_i}{c}$) of the i^{th} path signal. Therefore, if the sampling rate of CSI is larger than twice of the max Doppler frequency shift, we could obtain the Doppler frequency shift of each path signal. This shift includes both the amplitude and direction information. For the walking speed, the required sampling rate of CSI is below several hundred hertz, which is easily achievable on a commodity Wi-Fi device.

4.2 MUSIC-based Doppler Estimation

Although it is possible to estimate Doppler frequency shift of the Wi-Fi signal with only several hundred hertz of sampling rate, two problems still exist in order to obtain accurate Doppler frequency shift estimation using commodity Wi-Fi devices. First, only one CSI sample is obtained for each Wi-Fi packet, so the sample interval depends on when the packets arrive. In reality, commodity Wi-Fi devices are not able to send and receive packets with fixed precise intervals due to the packet loss/delay caused by environmental noise and interference. Therefore, the adjacent CSI samples usually have varying intervals between them. Second, the Wi-Fi signal is transmitted on multiple subcarriers. Due to frequency selective fading, different subcarriers have different signal-to-noise ratios (SNR). If we select a subcarrier with low SNR to process, the Doppler information estimated is less accurate.

To address these problems, we propose a MUSIC-based algorithm to obtain the accurate Doppler frequency shift estimates. We first consider the scenario with only one path signal. Assume we receive M CSI samples, each is timestamped at a microsecond-level precision (supported by commodity Wi-Fi cards such as Intel 5300). The first CSI sample is collected at t_0, and the sampling interval of each sample with respect to the first sample is $[0, \Delta t_2, \ldots, \Delta t_M]$, where $\Delta t_1 = 0$. In a short sampling window, we can ignore the attenuation difference across different CSI samples and treat the path change speed as a constant. So the phase difference between the i^{th} CSI sample and the first sample is $e^{-j2\pi f\frac{v\Delta t_i}{c}}$, where f is the original carrier frequency of signal. Thus, the phase differences between the M^{th} CSI samples and the first CSI sample can be expressed as follows:

$$\vec{a}(v) = [1, e^{-j2\pi f\frac{v\Delta t_2}{c}}, e^{-j2\pi f\frac{v\Delta t_3}{c}}, \ldots, e^{-j2\pi f\frac{v\Delta t_M}{c}}]^\top \tag{2}$$

We call $\vec{a}(v)$ the *Doppler vector* and the CSI sample matrix with M samples is represented as:

$$\begin{aligned} \mathbf{X}(f) &= [x(f, t_0), x(f, t_0 + \Delta t_2), \ldots, x(f, t_0 + \Delta t_M)]^\top \\ &= [1, e^{-j2\pi f\frac{v\Delta t_2}{c}}, e^{-j2\pi f\frac{v\Delta t_3}{c}}, \ldots, e^{-j2\pi f\frac{v\Delta t_M}{c}}]^\top x(f, t_0) + n(f) \\ &= \vec{a}(v)x(f, t_0) + n(f) \end{aligned} \tag{3}$$

where $n(f)$ is the noise. When only one path signal exists, we can calculate the Doppler frequency shift easily from the phase measurements across CSI samples. In real-life situation with multipath, L path signals would

arrive at the receiver. Based on Equation 1 and Equation 3, the CSI sample matrix is now expressed as:

$$\begin{aligned}\mathbf{X}(f) &= \sum_{i=1}^{L} \vec{a}(v_i)s_i(f,t_0) + n_i(f) \\ &= [\vec{a}(v_1),\vec{a}(v_2),\ldots,\vec{a}(v_L)][s_1(f,t_0),s_2(f,t_0),\ldots,s_L(f,t_0)]^\top + \mathbf{N}(f) \\ &= \mathbf{AS}(f) + \mathbf{N}(f) \end{aligned} \tag{4}$$

where v_i is the path change speed of the i^{th} path signal, $\mathbf{A} = [\vec{a}(v_1),\vec{a}(v_2),\ldots,\vec{a}(v_L)]$ is an $M \times L$ matrix with all Doppler vectors, $s_i(f,t_0)$ is the CSI of the i^{th} path signal that is measured at the first sampling time (t_0), matrix $\mathbf{S}(f) = [s_1(f,t_0),s_2(f,t_0),\ldots,s_L(f,t_0)]^\top$ is the *signal matrix*, and $\mathbf{N}(f)$ is the noise matrix. If we can resolve matrix $\mathbf{A}$, we could get the Doppler frequency shift of each path signal.

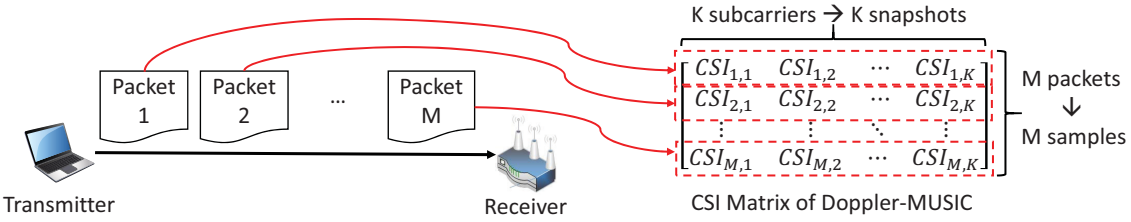


Fig. 3. K CSI snapshots across K subcarriers for each sample and M samples can be extracted from M packets.

To resolve matrix $\mathbf{A}$, we apply the MUSIC algorithm [31], which was previously applied to estimate the AoA of each path signal. However, one more condition needs to be satisfied before we could apply the MUSIC algorithm: each CSI sample requires multiple snapshots to average out the random noise. If we only employ one snapshot for a sample, the estimate may have a large error due to random noise. To address this problem, we take multiple snapshots for each CSI sample in the frequency domain as shown in Figure 3. The commodity Wi-Fi device provides CSI on multiple subcarriers.[1] Let K be the number of subcarriers with CSI, we have K snapshots of CSI for a CSI sample. For the i^{th} CSI sample, it can be expressed as:

$$\vec{x}(f,t_0+\Delta t_i) = [x(f_1,t_0+\Delta t_i),x(f_2,t_0+\Delta t_i),\ldots,x(f_K,t_0+\Delta t_i)] \tag{5}$$

where f_k is the original carrier frequency of the k^{th} subcarrier. The i^{th} path signal in the signal matrix $\mathbf{S}(f)$ can be expressed as:

$$\vec{s_i}(f,t_0) = [s_i(f_1,t_0),s_i(f_2,t_0),\ldots,s_i(f_K,t_0)] \tag{6}$$

where $s_i(f_k,t_0)$ is the CSI of the k^{th} snapshot of the i^{th} path signal at the first sampling time (t_0). We take CSIs from all subcarriers, thus avoiding choosing a particular subcarrier with low SNR which may lead to less accurate Doppler shift estimation. With Equation 5 and Equation 6, Equation 4 can be expressed as:

$$\begin{aligned}\mathbf{X}(f) &= [\vec{x}(f,t_0),\vec{x}(f,t_0+\Delta t_2),\ldots,\vec{x}(f,t_0+\Delta t_M)]^\top \\ &= [\vec{a}(v_1),\vec{a}(v_2),\ldots,\vec{a}(v_L)][\vec{s_1}(f,t_0),\vec{s_2}(f,t_0),\ldots,\vec{s_L}(f,t_0)]^\top + \mathbf{N}(f) \\ &= \mathbf{AS}(f) + \mathbf{N}(f) \end{aligned} \tag{7}$$

We are now ready to apply the MUSIC algorithm to estimate the Doppler frequency shift using the CSI samples. The basic idea of the MUSIC algorithm is eigenstructure analysis of an $M \times M$ correlation matrix $\mathbf{R_X}$ of the M

[1]The Intel 5300 Wi-Fi card provides the CSI of 30 subcarriers for each packet.

CSI samples. From Equation 7, we express $\mathbf{R_X}$ as:

$$\begin{aligned} \mathbf{R_X} &= \mathbb{E}[\mathbf{X}\mathbf{X}^{\mathrm{H}}] \\ &= \mathbf{A}\mathbb{E}[\mathbf{S}\mathbf{S}^{\mathrm{H}}]\mathbf{A}^{\mathrm{H}} + \mathbb{E}[\mathbf{N}\mathbf{N}^{\mathrm{H}}] \\ &= \mathbf{A}\mathbf{R_S}\mathbf{A}^{\mathrm{H}} + \sigma^2\mathbf{I} \end{aligned} \tag{8}$$

where $\mathbf{R_S}$ is the correlation matrix of the signal matrix, $\mathbf{I}$ is an identity matrix and σ^2 is the variance of noise. The correlation matrix $\mathbf{R_X}$ has M eigenvalues. The smallest $M-L$ eigenvalues correspond to the noise and the other L eigenvalues correspond to the L path signals. The eigenvectors corresponding to the smallest $M-L$ eigenvalues construct a noise subspace $\mathbf{E_N} = [\vec{e}_1, \ldots, \vec{e}_{M-L}]$. The signal and the noise subspace are orthogonal so the Doppler velocity spectrum function can be expressed as:

$$P(v)_{MUSIC} = \frac{1}{\mathbf{a}^{\mathrm{H}}(v)\mathbf{E_N}\mathbf{E_N}^{\mathrm{H}}\mathbf{a}(v)} \tag{9}$$

in which sharp peaks correspond to the Doppler velocity of the path signals appear. Since the original carrier frequency and propagation speed of Wi-Fi signals are known, we can then obtain the Doppler velocity.

4.3 Addressing Random CSI Phase Offset

On commodity Wi-Fi devices, there is one more challenge before we can apply the MUSIC-based algorithm proposed above for Doppler velocity estimation. For commodity Wi-Fi devices, the receivers are not tightly time synchronized with the transmitter, resulting in a time-variant random phase offset $e^{-j\theta_{offset}}$ [38, 42] in each CSI sample as shown in Equation 10:

$$\begin{aligned} x(f, t_0 + t) &= e^{-j\theta_{offset}} \sum_{i=1}^{L} A_i e^{-j2\pi f(\tau_i + \frac{v_i t}{c})} \\ &= e^{-j\theta_{offset}} (x_s(f, t_0) + \sum_{i \in G_m} A_i e^{-j2\pi f(\tau_i + \frac{v_i t}{c})}) \end{aligned} \tag{10}$$

where $x_s(f, t_0)$ represents the static path component in the CSI sample and G_m is the set of mobile path components. Here mobile path components refer to those signals reflected from the moving target while the static path component refers to the direct path signal and the reflection signal from static objects such as walls. The time-variant random phase offsets distort the phase changes in time domain, thus affecting the Doppler velocity estimation. To remove these phase offsets, we propose the following novel steps:

(1) **Conjugate multiplication.** A commodity Wi-Fi card is usually equipped with multiple antennas (e.g., 3 antennas for Intel 5300 Wi-Fi card). A CSI sample contains the CSI from all antennas. The key property we utilize here is that the time-variant random phase offsets are the same across different antennas on a Wi-Fi card [23, 26] as they share the same RF oscillator. Therefore, on a Wi-Fi card, we can apply conjugate

multiplication between the CSI of two antennas to remove the time-variant random phase offsets:

$$\begin{aligned} x_{cm}(f, t_0 + t) &= x_1(f, t_0 + t)\overline{x}_2(f, t_0 + t) \\ &= (x_{1,s}(f, t_0) + \sum_{i \in G_{m1}} A_i e^{-j2\pi f(\tau_i + \frac{v_i t}{c})})(\overline{x}_{2,s}(f, t_0) + \sum_{k \in G_{m2}} A_k e^{j2\pi f(\tau_k + \frac{v_k t}{c})}) \\ &= \overbrace{x_{1,s}(f, t_0)\overline{x}_{2,s}(f, t_0)}^{①} + \overbrace{\sum_{i \in G_{m1}, k \in G_{m2}} A_{1,i} A_{2,k} e^{-j2\pi f((\tau_i - \tau_k) + \frac{(v_i - v_k)t}{c})}}^{②} \\ &\quad + \overbrace{\overline{x}_{2,s}(f, t_0) \sum_{i \in G_{m1}} A_{1,i} e^{-j2\pi f(\tau_i + \frac{v_i t}{c})}}^{③} + \overbrace{x_{1,s}(f, t_0) \sum_{k \in G_{m2}} A_{2,k} e^{j2\pi f(\tau_k + \frac{v_k t}{c})}}^{④} \end{aligned} \tag{11}$$

where $x_{cm}(f, t_0 + t)$ is the output after conjugate multiplication, $x_1(f, t_0 + t)$ is the CSI of the first antenna, $\overline{x}_2(f, t_0 + t)$ is the conjugate of the CSI of the second antenna, G_{m1} and G_{m2} are the sets of mobile paths at the first antenna and second antenna, respectively.

(2) **Remove static component.** In Equation 11, the product of the static-path components of two antennas, which is marked with ①, can be treated as a constant in a short time period. It does not contain the Doppler velocity information we care about. However, the power of the static component can be very high, since it contains the strong direct path signals. To avoid the static component's interference on Doppler velocity estimation, we remove the static component by subtracting the mean value from the conjugate multiplication.

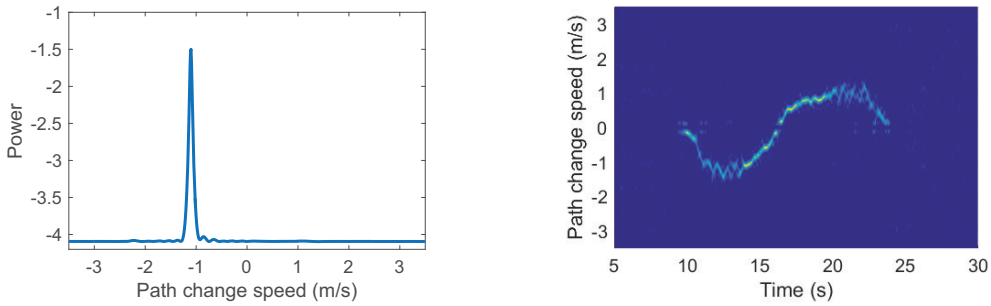
(3) **Adjust the power of each antenna.** In Equation 11, the product of mobile path components, which is marked with ②, has a very small value so we can ignore it. The remaining terms, which are marked with ③ and ④, are the two products of the static paths component of one antenna and the mobile paths component of another antenna. These two terms contain the Doppler velocity information that we care about. Since the two close-by antennas have similar multipaths, the Doppler velocity information in these two terms have similar value but opposite direction. We want the Doppler velocity obtained from the term that is the product of the first antenna's mobile paths component and the second antenna's static paths component, which is marked with ③. Thus, we reduce the power of the static paths component on the first antenna by subtracting a value α and increase the power of static paths component on the second antenna by adding a value β.[2] With the power adjustment step described above, the term containing the correct Doppler velocity information has much higher power in the multiplication output and can be identified in the spectrum.

In IndoTrack, we employ the above steps to remove the random CSI phase offset and then use the MUSIC-based algorithm to estimate the Doppler velocity. Figure 4(a) shows an example of the Doppler velocity spectrum obtained. The position of the highest peak corresponds to the target's Doppler velocity, which is introduced by the human's movement [37]. Figure 4(b) shows the speed-time plot we obtain when a person walks towards the ligature of the pair of transceivers and then walks away.

5 DOPPLER-AOA FOR HUMAN TRAJECTORY ESTIMATION

The Doppler-MUSIC method proposed in previous section allows us to estimate the Doppler velocity of the human reflection path. To obtain absolute trajectory estimation, more work is needed. First, there is no direct translation between Doppler velocity and human velocity, since it also depends on the human's location at each

[2]In our implementation, in each estimation window, we choose α so the minimum amplitude of CSI across all the samples within the window at the first antenna is reduced to zero, and we set beta as 1000α.



(a) Doppler velocity spectrum (b) Speed-time spectrum

Fig. 4. (a) is a Doppler velocity spectrum based on CSI samples collected in $0.3s$ with a $200Hz$ sampling rate; (b) is a speed-time spectrum which we estimate the Doppler velocity every $0.05s$.

timestamp. Furthermore, the velocity information does not contain the absolute location information so the starting point of the trajectory is still unknown.

To address the problems, we observe that (1) although AoA-based location estimation with only 3 antennas is very coarse, the AoA spectrum does help to restrict the trajectory in spatial domain; (2) although different from human velocity, the Doppler velocity does contain the velocity information of the target and thus restrict the trajectory in time domain; and (3) by jointly considering Doppler velocity and AoA spectrum, we can continuously estimate the absolute human location and velocity at the same time and thus obtain accurate absolute trajectory estimate. In this section, we first analyze the geometrical relationship among Doppler velocity and human velocity with respect to the human target's location. We then present the detailed design of space-time joint trajectory estimation.

5.1 Geometrical Relationship: From Doppler Velocity to Human Velocity

For simplicity of explanation, we treat the moving human target as a point in this section. When the human movement changes the length of the reflection path, it will introduce a Doppler frequency shift on the Wi-Fi signal. Given a pair of Wi-Fi transceivers and a human target, there exists an ellipse with foci at the transmitter and receiver, and the human target is on the ellipse, as shown in Figure 5(a). If the human target walks along the tangent direction of the ellipse, the movement will not change the reflection path length and we can not detect any Doppler frequency shift. However, if there is a velocity component along the normal direction of the ellipse, the movement will change the reflection path length and cause a Doppler frequency shift.

Suppose we know the human's position $\vec{p}_h =< x_h, y_h >$**.** Given the positions of the transmitter ($\vec{p}_{tx} =< x_{tx}, y_{tx} >$) and receiver ($\vec{p}_{rx} =< x_{rx}, y_{rx} >$), at the human's position $\vec{p}_h$, we can calculate the normal vector ($\vec{l}_n =< x_n, y_n >$) of the ellipse easily. As illustrated in Figure 5(b), we obtain the following equations regarding to the human's normal velocity component ($\vec{v}_n =< v_{nx}, v_{ny} >$):

$$\begin{cases} \vec{v}_n = k\frac{\vec{l}_n}{\|\vec{l}_n\|} \\ \frac{\vec{v}_n \cdot (\vec{p}_h - \vec{p}_{tx})}{\|\vec{p}_h - \vec{p}_{tx}\|} = \frac{v_{path}}{2} \\ \frac{\vec{v}_n \cdot (\vec{p}_h - \vec{p}_{rx})}{\|\vec{p}_h - \vec{p}_{rx}\|} = \frac{v_{path}}{2} \end{cases} \tag{12}$$

where v_{path} is the path change speed (Doppler velocity) we estimate from the Doppler-MUSIC method. Based on these equations, we can calculate the human's normal velocity component. As shown in Figure 5(c), if we have

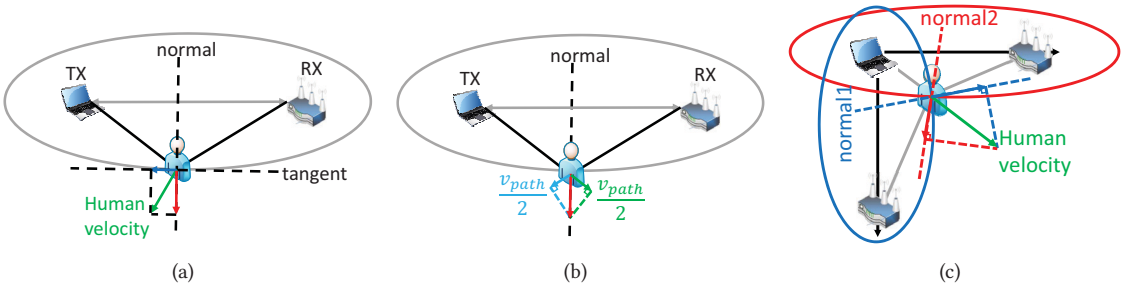


Fig. 5. Geometrical relationship between human velocity and Doppler velocity.

another receiver, we could obtain another group of equations for the human's velocity ($\vec{v}_h =< v_{hx}, v_{hy} >$):

$$\begin{cases} \frac{\vec{v}_h \cdot \vec{l}_{n1}}{\|\vec{l}_{n1}\|} = \|\vec{v}_{n1}\| \\ \frac{\vec{v}_h \cdot \vec{l}_{n2}}{\|\vec{l}_{n2}\|} = \|\vec{v}_{n2}\| \end{cases} \tag{13}$$

where $\vec{l}_{ni} =< x_{ni}, y_{ni} >$ is the normal vector and $\vec{v}_{ni} =< v_{nxi}, v_{nyi} >$ is the normal velocity component of the human for the i^{th} pair of transceivers at the human's current position. Therefore, we can calculate the human's velocity $\vec{v}_h$ based on Equation 13.

5.2 Doppler-AoA Method

The discussion above assumes that we know the starting position of the human target. In this section, we describe how to determine the absolute starting position. Subsequent human locations and velocities can be determined based on the current location and velocity. We apply joint space and time constraints to obtain the probabilistic estimation of the starting position and trajectory. As illustrated in Figure 6, let $< x_1, y_1 >$ be the starting position. Based on the geometrical relationship, we can calculate the human velocity from the Doppler velocities of the two pairs of transceivers. Thus, we could know the human's position at the next time point. Similarly, based on the human's position and detected Doppler velocities at time point 2, we can calculate the human velocity and

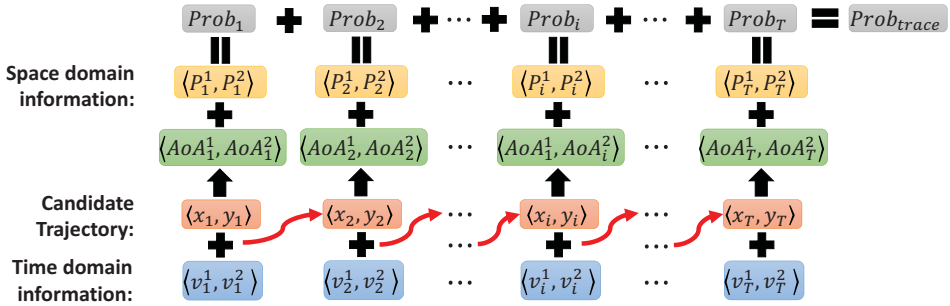


Fig. 6. Calculate the confidence value for a candidate trajectory: a trajectory contains T points, $< x_1, y_1 >$ is the starting position, $< v_i^1, v_i^2 >$ is the Doppler velocity estimated by the Doppler-MUSIC method at point i, $< AoA_i^1, AoA_i^2 >$ is the expected AoA to the two receivers, $< P_i^1, P_i^2 >$ is the estimated AoA spectrum, $Prob_i$ is the confidence value of the position at point i and $Prob_{trace}$ is the overall confidence value of the whole candidate trajectory.

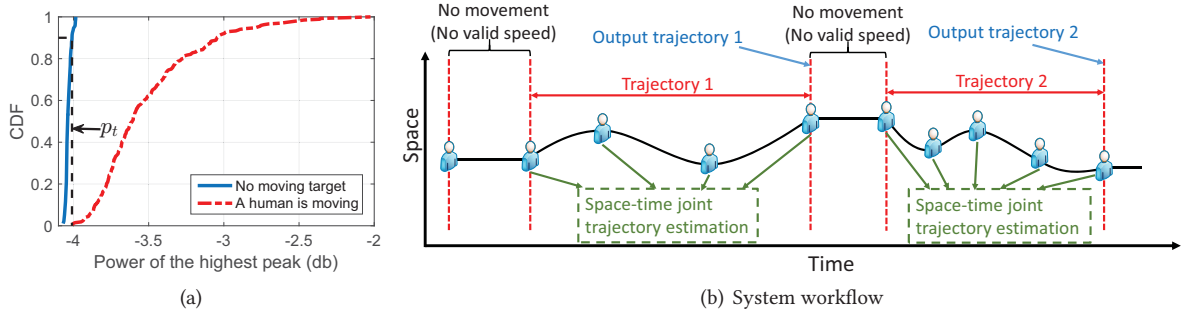
obtain the human's position at time point 3. Finally, we can determine the whole trajectory. On the other hand, at time point i of the trajectory, based on the CSI measurements, we can obtain the AoA spectrum ($< P_i^1, P_i^2 >$), which represents the angle information of the human to the two receivers. We can also calculate the expected human angle ($< AoA_i^1, AoA_i^2 >$) to the two receivers if the human's position is ($< x_i, y_i >$). Thus, we can now compare the AoA spectrum and the human angle to calculate the confidence value of the human's position at time point t_i: $Prob_i = P_i^1(AoA_i^1) \cdot P_i^2(AoA_i^2)$. We compute the confidence value of the whole trajectory with the starting position at $< x_1, y_1 >$ by adding the confidence values at all positions along the trajectory:

$$Prob_{trace}(< x_1, y_1 >) = \sum_{i=1}^{T} Prob_i \tag{14}$$

where T is the number of points along the trajectory. Using this method, we can calculate the confidence values for all candidate starting positions and choose the position with the highest overall confidence value as the starting position. As a result, we obtain the absolute trajectory. Moreover, to speed up the starting point estimate, we narrow down the search space by checking if the whole trajectory is still within the sensing zone. If we find the trajectory exceeds the sensing zone, the starting position is removed from the candidate list. Removing such infeasible staring positions helps reduce the computation and speed up the estimation process.

5.3 Continuous Human Trajectory Tracking

Our system is designed to track human movement continuously and automatically output segmented trajectories when the human target is stationary, such as sitting on the sofa, sleeping in the bed, or moving outside of the monitoring area. We use a threshold-based method to determine when the human target stops moving. In the Doppler-MUSIC method, if there is no moving target in the environment, mobile path signal does not exist. Thus the highest peak's power in the Doppler velocity spectrum would be very low. In contrast, when there is human movement, the power is much higher. Figure 7(a) shows the cumulative distribution function (CDF) of the highest peak's power of Doppler velocity spectrum in the static and dynamic environments. There is a clear gap between the two curves. Based on comprehensive experiments, we choose the $90th$ percentile of the highest power in the static environment as the threshold p_t. For each spectrum, we choose the highest peak as the target peak and its power is p_{peak}. If $p_{peak} < p_t$, we conclude that there is no moving target and the target's Doppler velocity is zero. Due to environmental noise, this threshold-based method may detect few false movements in static environment. However, such false movement does not last long, so it can be easily filtered out. Based on the detected Doppler



(a) (b) System workflow

Fig. 7. (a) is the CDF of the power of the highest peak on the Doppler velocity spectrum; (b) shows the system workflow for continuous human trajectory tracking.

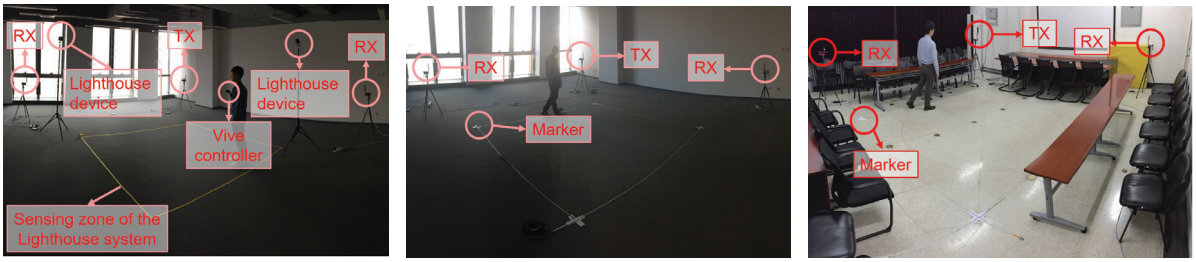
velocity, if the person stays stationary for a period of time, we conclude the person has completed a trajectory and output this trajectory. Figure 7(b) shows the full process of continuous human trajectory tracking.

6 EVALUATION

In this section, we first describe our system implementation and experimental setup. We then present the detail experimental results and the performance of IndoTrack in terms of both of velocity estimation accuracy and tracking accuracy. We also conduct a feasibility study in a home environment to analyze high-level semantics that can be supported via IndoTrack's indoor human trajectory tracking capabilities.

6.1 Implementation

IndoTrack requires one Wi-Fi transmitter and two receivers. We employ GIGABYTE miniPCs equipped with cheap off-the-shelf Intel 5300 Wi-Fi cards as the transmitter and receivers. Three antennas are attached to each receiver and the antenna space is half-wavelength of the Wi-Fi signal. The CSI tool [15] developed by Halperin is installed on the miniPCs to collect the CSI sample for each received packet. The sampling rate of CSI in our experiments is set to 200 Hz to ensure that all Doppler frequencies caused by human movements can be detected as the human velocity in indoor environment is usually less than $3m/s$ [6]. For each packet, the Intel 5300 Wi-Fi card provides CSI on 30 subcarriers and thus, we have 30 snapshots of CSI for our Doppler-MUSIC method. For each Doppler velocity estimation, we employ CSI samples collected in a time window of 0.3 seconds. IndoTrack can be hosted on any Wi-Fi channel in the 2.4 GHz and 5 GHz bands. In order to avoid interference, our experiments are conducted in an unused 5 GHz frequency band with a 20 MHz bandwidth. There is no restriction on the type of packets, so any Wi-Fi packet including the beacons can be employed for the tracking system. Therefore, IndoTrack has a minimum impact on existing Wi-Fi data communication. When we deploy the IndoTrack system, we carefully measure the positions of Wi-Fi transmitter and receivers with a laser range meter.



(a) (b) (c)

Fig. 8. (a) is the setup for the evaluation of Doppler-MUSIC method; (b) and (c) show the two experimental environments for the evaluation of tracking performance: an empty room and a meeting room.

To evaluate the performance of IndoTrack, we first conduct experiments to evaluate the accuracies of target velocity estimation, in terms of both velocity amplitude and direction. In order to obtain the ground truth of human moving velocity, we use the Lighthouse system of HTC Vive, which is a real-time localization and motion tracking system for virtual reality applications. The size of the sensing zone of the Lighthouse system is $3.5m \times 3.5m$ and the tracking accuracy is $1.5mm$. We fix a Vive controller on the chest of the human target to measure the ground truth of his moving velocity. Figure 8(a) shows the setup of the experiment.

Then we evaluate the tracking performance of IndoTrack by conducting experiments in two indoor environments, as shown in Figure 8(b) and 8(c): one is an empty room and the other is a meeting room equipped with

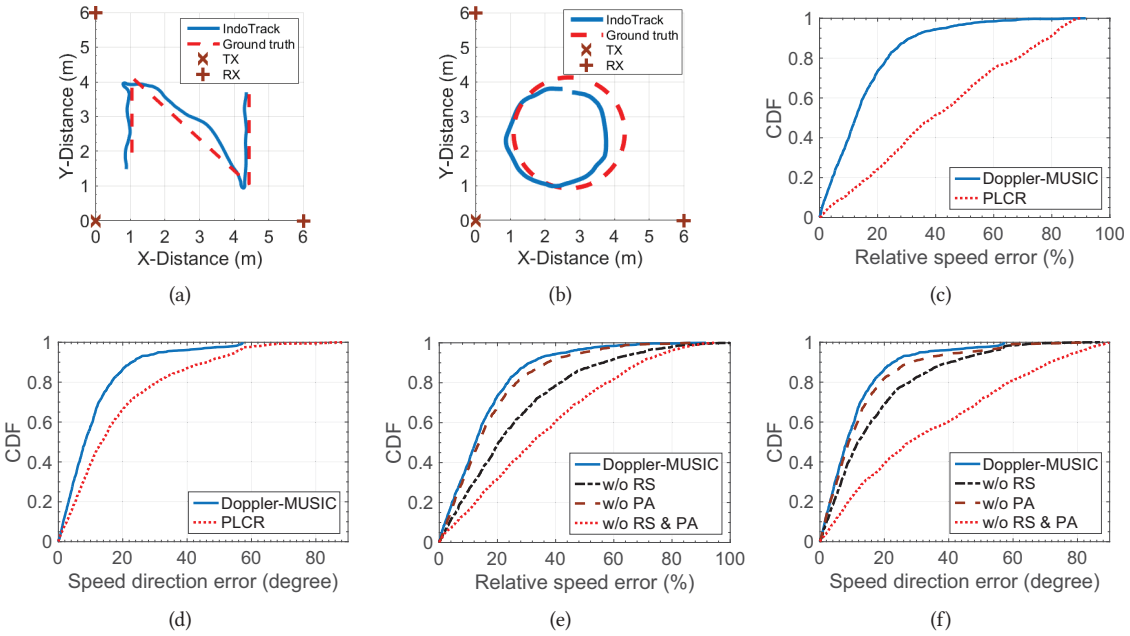


Fig. 9. (a) and (b) are two examples of the tracking results of IndoTrack; (c) and (d) show the overall performance of Doppler-MUSIC method: (c) shows the accuracy of estimated human velocity amplitude, (d) shows the accuracy of estimated human velocity direction; (e) and (f) show the benefit of adjusting power of antennas and removing static component .

furniture and electrical appliances with rich multipath. The size of sensing zone of our tracking is $6m \times 6m$. We place some markers on the floor and use video cameras to record when the person walks over the markers, which is the ground truth of each trajectory. In our experiments, 5 volunteers (targets) walk along different shapes of trajectories such as line, rectangular, circle, etc over a time period of 3 weeks. We collect 270 trajectories for each person and we report the tracking error[3] to show the performance of IndoTrack. Figures 9(a) and 9(b) show two examples of the tracking results of IndoTrack.

Finally, we also conduct an in-situ experiment to test the feasibility of IndoTrack to record human daily indoor trajectories and extract the high-level semantics based on the trajectories.

6.2 Performance of Doppler-MUSIC

Velocity estimation. As shown in Figure 9(c) and 9(d), the median relative speed (velocity amplitude) error is as small as 12% and the median velocity direction error is only 8°. We also compare our Doppler-MUSIC method with PLCR, which was proposed in WiDar for Doppler frequency shift estimation [30]. Since PLCR estimates Doppler velocity from CSI amplitudes, it can not provide the direction information. In the comparison, we only use the PLCR method to estimate the amplitude of Doppler velocity and use our Doppler-MUSIC method to provide the direction information for PLCR. As shown in Figure 9(c) and 9(d), even with our direction information added to improve the performance, the median relative speed error of PLCR is 38% and the median direction error is 13°. The reason is that our Doppler-MUSIC method estimates Doppler velocity with CSI phase, which is more

[3]We treat the human body as a cylinder with a width of 50cm as the reflection points on the body are always changing. If the estimated location is within the zone of the cylinder, we consider there is no tracking error.

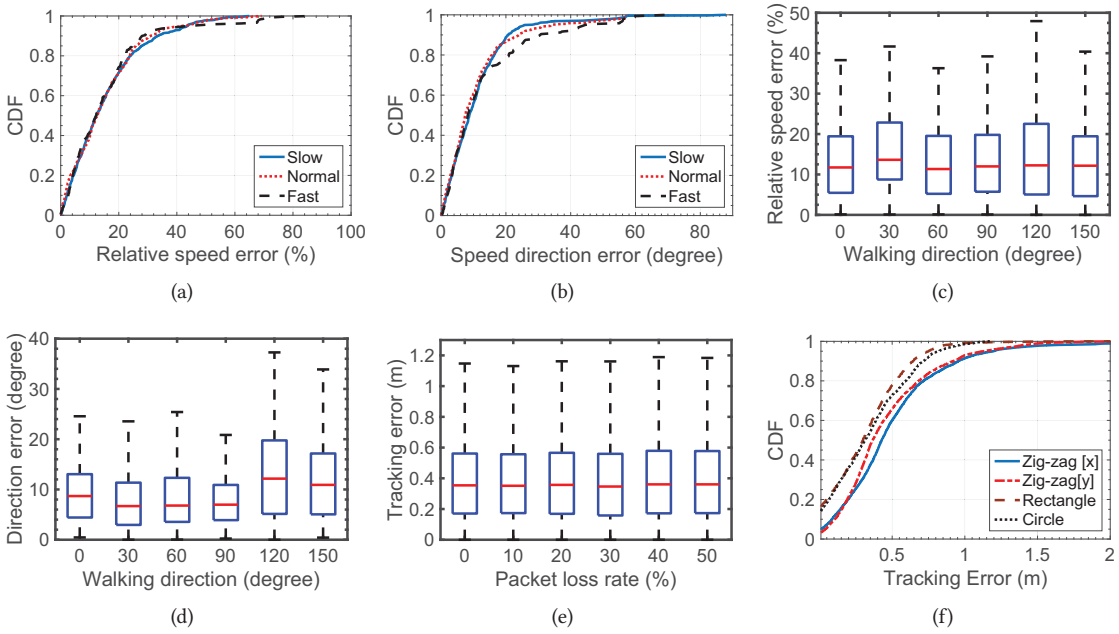


Fig. 10. (a) and (b) show the estimated human velocity accuracy for different walking speeds; (c) and (d) show the estimated human velocity accuracy for different walking direction; (e) shows the overall tracking performance and the impact of packet loss/delay; (f) shows the CDF of tracking error for four different types of trajectories.

stable than CSI amplitude used in PLCR. Therefore, we can make more accurate Doppler velocity estimation than WiDar and achieve a smaller velocity error. In addition, compared with the PLCR method, our proposed method could estimate both the amplitude and direction of Doppler velocity at the same time.

Benefit of adjusting power of antennas (PA) & removing static component (RS). To improve the accuracy of Doppler velocity estimation as described in section 4.3, we adjust the power of the two antennas before the conjugate multiplication, and remove the static component after the multiplication. We show the effect of these two steps in Figure 9(e) and 9(f). If we neither adjust the power of antennas nor remove the static component, the median relative speed error is 32% and the median direction error is 28°. If we only adjust the power of antennas, the relative speed error is 20% and the median direction error is 12°. If we only remove the static component, relative speed error is 14% and the median direction error is 9°. Thus, we can see that after we remove the strong interfering static component, the performance of IndoTrack is improved significantly. The power of the two antennas is adjusted to ensure we can detect the correct direction information of the Doppler velocity. With this step, the tracking performance is further improved.

Impact of moving speed diversity. According to the literature [6, 13], the comfortable walking speed of a human is $1 - 1.5m/s$. In order to show the impact of human moving speed, we let the human target walk at three different of speed levels: slow (below $1m/s$), normal ($1 - 1.5m/s$) and fast ($1.5 - 3m/s$). As shown in Figure 10(a) and 10(b), the speed does not affect the performance of our Doppler-MUSIC method and thus IndoTrack achieves similar velocity estimation errors.

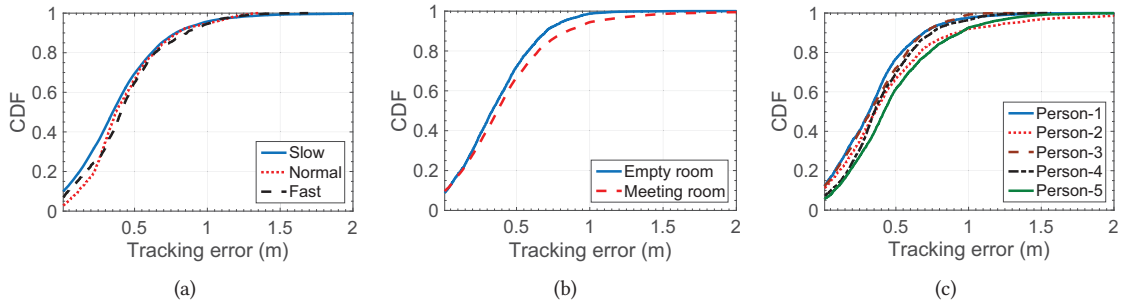


Fig. 11. (a) shows the CDF of tracking error for different walking speeds; (b) shows the CDF of tracking error for different environments; (c) shows the CDF of tracking error for different persons.

Impact of moving direction diversity. Moreover, we also evaluate the impact of moving direction diversity. We let the human walk along straight lines at different directions and then estimate the velocity with the proposed Doppler-MUSIC method. As shown in Figure 10(c) and 10(d), the Doppler-MUSIC method achieve very similar performance at all directions so the orientation does not affect the performance of IndoTrack.

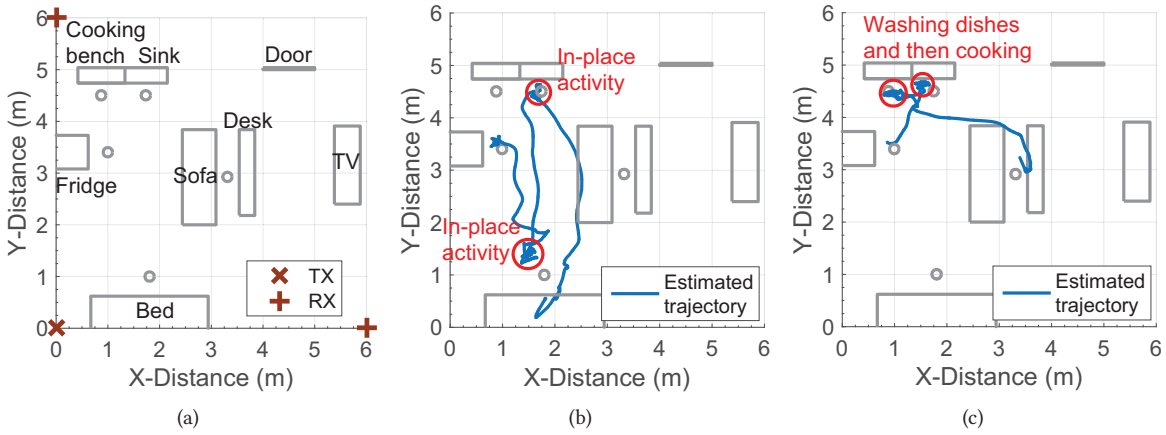
6.3 Tracking Performance

Impact of packet loss/delay. As shown in Figure 10(e), without packet loss/delay, the median tracking error of IndoTrack is $35cm$. In real environment, commodity Wi-Fi devices will suffer from the packet loss/delay due to environmental noise and interference. We thus randomly discard some Wi-Fi packets to emulate the packet loss/delay scenario. As shown in Figure 10(e), even with a 50% packet loss, IndoTrack still achieves a similar high tracking accuracy as there is no packet loss.

Impact of trajectory diversity. We choose 4 different typical shapes of trajectories to show the impact of trajectory diversity: two types (vertical & horizontal) of zig-zag paths that look like 'Z', rectangle, and circle. As shown in Figure 10(f), the median tracking error of the horizontal zig-zag is $40cm$ and the median tracking error of the vertical zig-zag is $36cm$. Both the rectangle and circle trajectories have a median tracking error of $30cm$. So IndoTrack is robust against different shapes of trajectories.

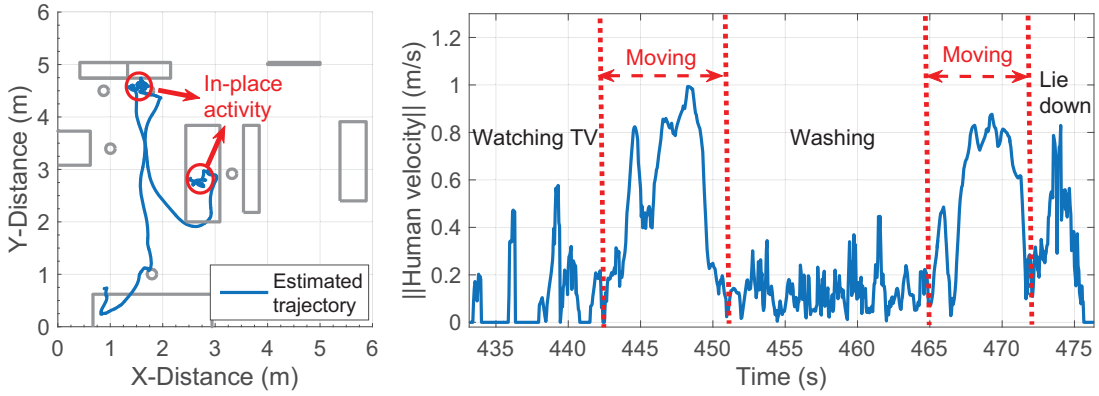
Impact of human speed. For the tracking performance of IndoTrack, we also conduct experiments to show the influence of human speed. We let the human target walk at three different levels of speeds: slow (below $1m/s$), normal ($1 - 1.5m/s$) and fast ($1.5 - 3m/s$). As shown in Figure 11(a), even with very different walking speeds, IndoTrack still achieves consistent high tracking accuracies.

Impact of environment. In order to show the influence of multipath, we conduct experiments in two different environments. The sensing zones in the two environments have the same size: $6m \times 6m$. Figure 11(b) shows the CDF of tracking errors in the two indoor environments. In the empty room, the median error is $33cm$ while in the meeting room, the median error is $37cm$. Even though there are more objects in the meeting room with more reflection paths, IndoTrack achieves similar tracking performance. This demonstrates that the static multipaths have little impact on IndoTrack's performance as they are removed during the process and we only care about the mobile multipath reflected from the human body.



(a) (b) (c)

Fig. 12. (a) shows the layout of the in-situ experiment; (b) is an estimated trajectory, the sequence of activities is: getting up–washing face at the sink–making the bed–taking food from fridge; (c) is an estimated trajectory, the sequence of activities is: taking food from fridge–washing the dishes in the sink–cooking–eating on the sofa.



(a) (b)

Fig. 13. (a) is an estimated trajectory, the sequence of activities is: watching TV on the sofa–washing face at the sink–going to sleep; (b) is the human velocity amplitudes that correspond to the trajectory shown in Figure 13(a).

Impact of human diversity. Five volunteers are employed as tracking targets including 2 females and 3 males of different ages, heights, and weights. Figure 11(c) shows the CDF of tracking errors of IndoTrack for the 5 targets. We can clearly see that, IndoTrack achieves consistent accuracies across different targets.

6.4 Feasibility Study of Daily Human Trajectory Recording

As shown in Figure 12(a), we emulate a typical studio flat with furniture placed in the meeting room (Figure 8(c)) for an in-situ experiment. The small grey circles in the figure represent the positions the human target stops to perform some location related activities. We ask the volunteer to do some daily activities (e.g., cooking, washing

face and eating, etc.) based on a script without any constraint on the trajectories. Figure 12 and Figure 13 show the trajectories estimated. From the experiment, we can see that IndoTrack could successfully track the trajectory and obtain the information such as the starting point and stopping point, for any random trajectory.

Moreover, from the recorded trajectory, we could extract the positions with in-place activities. As show in Figure 12(b), when the target is performing in-place activities, his location changes in a very small range. The human velocity, during the in-place activities, also exhibits random directions and small values. In addition, IndoTrack is capable to distinguish two nearby in-place activities. As shown in Figure 12(c), the target washes the dishes in the sink and then cooks at the cooking bench. Although the distance between the sink and the cooking bench is less than $1m$, the estimated trajectory could still reflect when the target moves from the sink to the cooking bench. Besides the location semantic information, IndoTrack could also provide the velocity semantic information of the target's daily activities. Figure 13(b) shows the estimated velocity values of the target, which corresponds to the trajectory in Figure 13(a). IndoTrack not only provides the velocity semantic information when the target walks, but also the velocity information of the in-place activities, which may reflect the type of activities and the degree of activeness. To summarize, the output of IndoTrack can be used for future physical/behavioral analysis by extracting the high-level semantics.

7 DISCUSSIONS

In this work, we focus on designing an indoor human tracking system that is device-free and utilizes only commodity Wi-Fi devices. Our current design supports single-person tracking. There are several directions to further extend our work, which we discuss below.

Multiple-person tracking. Passive or device-free tracking of multiple people is known to be challenging. Multi-person tracking requires distinguishing the reflection path signals from different human targets. However, a person may generate more than one reflection path signal and these signals may be similar to other people's reflection path signals. As such, it is challenging to identify each person's reflection path signal accurately in a mixed signal with reflection paths from multiple persons, and this remains an important topic for further research.

Real-time tracking. To estimate the absolute starting point, IndoTrack needs to calculate the confidence value of each candidate starting position and choose the starting position with the maximum confidence value. Due to the large number of candidate positions and the dynamic nature of human trajectories, the computational load is high which impedes the real-time tracking. However, since the confidence value calculation of each candidate starting position is independent from each other, it is possible to parallelize the computation for real-time tracking.

High-level semantics. The output of IndoTrack is timestamped human location, velocity, and trajectory information, from which various types of high-level semantics can be extracted for further analysis in real-world applications. For instance, for elderly care, the moving patterns captured by the trajectories offer rich information about how the target moves around, his/her overall activeness and any abnormal movements for detecting illness, missed tasks, etc.

More real-world evaluations. In this paper, we analyze the performance of IndoTrack based on evaluations in two controlled environments. We also carry out a feasibility study about the capability of IndoTrack detecting daily indoor movements and activities with an in-situ experiment. In the future, we plan to deploy IndoTrack in a real home environment with people living and perform more evaluations about daily human activity and trajectory tracking for quantitative analysis to facilitate real-world applications.

8 CONCLUSION

In this paper, we propose a device-free human tracking system IndoTrack hosted on cheap commodity Wi-Fi devices. We propose a novel Doppler-MUSIC method to accurately estimate the Doppler velocity introduced by the human movements. We innovatively utilize information from two antennas to remove the random CSI phase offsets to ensure accurate Doppler velocity estimation. We also propose the Doppler-AoA method to combine Doppler velocity and AoA spectrum information to jointly estimate the absolute trajectory. IndoTrack is able to track a single person at a median tracking error of 35cm. In addition, IndoTrack is able to monitor the daily activities of a human target and provide timestamped human location, velocity, and trajectory information for high-level mobility and behavioral analysis.

REFERENCES

[1] Heba Abdelnasser, Moustafa Youssef, and Khaled A. Harras. 2015. WiGest: A ubiquitous WiFi-based gesture recognition system. In *Proceedings of 2015 IEEE Conference on Computer Communications (INFOCOM '15)*. 1472–1480.

[2] Fadel Adib, Zach Kabelac, Dina Katabi, and Robert C. Miller. 2014. 3D Tracking via Body Radio Reflections. In *Proceedings of the 11th USENIX Symposium on Networked Systems Design and Implementation (NSDI '14)*. USENIX Association, Seattle, WA, 317–329.

[3] Fadel Adib and Dina Katabi. 2013. See Through Walls with Wi-Fi!. In *Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM (SIGCOMM '13)*. ACM, New York, NY, USA, 75–86.

[4] Anthea Wain Sy Au, Chen Feng, Shahrokh Valaee, Sophia Reyes, Sameh Sorour, Samuel N. Markowitz, Deborah Gold, Keith Gordon, and Moshe Eizenman. 2013. Indoor Tracking and Navigation Using Received Signal Strength and Compressive Sensing on a Mobile Device. *IEEE Transactions on Mobile Computing* 12, 10 (Oct. 2013), 2050–2062.

[5] Paramvir Bahl and Venkata N. Padmanabhan. 2000. RADAR: An In-Building RF-based User Location and Tracking System. In *Proceedings of 2000 IEEE Conference on Computer Communications (INFOCOM 2000)*. IEEE, Atlanta, GA, USA, 775–784.

[6] RICHARD W. BOHANNON. 1997. Comfortable and maximum walking speed of adults aged 20-79 years: reference values and determinants. *Age and Ageing* 26, 1 (1997), 15.

[7] Agata Brajdic and Robert Harle. 2012. Scalable Indoor Pedestrian Localization Using Inertial Sensing and Parallel Particle Filters. In *Proceedings of 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN '12)*. IEEE, 1–10.

[8] Yuntian Brian Bai, Tao Gu, and Andong Hu. 2016. Integrating Wi-Fi and Magnetic Field for Fingerprinting Based Indoor Positioning System. In *Proceedings of 2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN '12)*. IEEE, 1–6.

[9] Chloë Brown, Christos Efstratiou, Ilias Leontiadis, Daniele Quercia, and Cecilia Mascolo. 2014. Tracking Serendipitous Interactions: How Individual Cultures Shape the Office. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '14)*. ACM, New York, NY, USA, 1072–1081.

[10] Q. Cai and J. K. Aggarwal. 1998. Automatic Tracking of Human Motion in Indoor Scenes across Multiple Synchronized Video Streams. In *Proceedings of the Sixth International Conference on Computer Vision (ICCV '98)*. IEEE, 356–362.

[11] Ke-Yu Chen, Mark Harniss, Justin Haowei Lim, Youngjun Han, Kurt L. Johnson, and Shwetak N. Patel. 2013. uLocate: A Ubiquitous Location Tracking System for People Aging with Disabilities. In *Proceedings of the 8th International Conference on Body Area Networks (BodyNets '13)*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), ICST, Brussels, Belgium, Belgium, 173–176.

[12] Mohamed Eldib, Francis Deboeverie, Wilfried Philips, and Hamid Aghajan. 2016. Behavior Analysis for Elderly Care Using A Network of Low-Resolution Visual Sensors. *J. Electron. Imaging* 25, 4 (Mar 2016), 17.

[13] Stacy Fritz and Michelle Lusardi. 2009. White paper: "walking speed: the sixth vital sign". *Journal of Geriatric Physical Therapy* 32, 2 (2009), 46.

[14] 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 (CHI '12)*. ACM, New York, NY, USA, 1911–1914.

[15] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2011. Tool Release: Gathering 802.11N Traces with Channel State Information. *SIGCOMM Comput. Commun. Rev.* 41, 1 (Jan. 2011), 53–53.

[16] Ju Han and Bir Bhanu. 2005. Human Activity Recognition in Thermal Infrared Imagery. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05) - Workshops - Volume 03 (CVPR '05)*. IEEE, Washington, DC, USA, 17–17.

[17] Qi Hao, Fei Hu, and Jiang Lu. 2010. Distributed Multiple Human Tracking with Wireless Binary Pyroelectric Infrared (PIR) Sensor Networks. (Nov 2010), 946–950.

[18] Michael Hardegger, Daniel Roggen, and Gerhard Tröster. 2015. 3D ActionSLAM: Wearable Person Tracking in Multi-floor Environments. *Personal Ubiquitous Comput.* 19, 1 (Jan. 2015), 123–141.

[19] Michael Hardegger, Gerhard Tröster, and Daniel Roggen. 2013. Improved actionSLAM for Long-term Indoor Tracking with Wearable Motion Sensors. In *Proceedings of the 2013 International Symposium on Wearable Computers (ISWC '13)*. ACM, New York, NY, USA, 1–8.

[20] Pan Hu, Liqun Li, Chunyi Peng, Guobin Shen, and Feng Zhao. 2013. Pharos: Enable Physical Analytics Through Visible Light Based Indoor Localization. In *Proceedings of the Twelfth ACM Workshop on Hot Topics in Networks (HotNets-XII)*. ACM, New York, NY, USA, Article 5, 7 pages.

[21] Kiran Joshi, Dinesh Bharadia, Manikanta Kotaru, and Sachin Katti. 2015. WiDeo: Fine-grained Device-free Motion Tracing Using RF Backscatter. In *Proceedings of the 12th USENIX Conference on Networked Systems Design and Implementation (NSDI '15)*. USENIX Association, Berkeley, CA, USA, 189–204.

[22] Soo-Cheol Kim, Young-Sik Jeong, and Sang-Oh Park. 2013. RFID-based Indoor Location Tracking to Ensure the Safety of the Elderly in Smart Home Environments. *Personal Ubiquitous Comput.* 17, 8 (Dec. 2013), 1699–1707.

[23] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. 2015. SpotFi: Decimeter Level Localization Using Wi-Fi. In *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication (SIGCOMM '15)*. ACM, New York, NY, USA, 269–282.

[24] Ye-Sheng Kuo, Pat Pannuto, Ko-Jen Hsiao, and Prabal Dutta. 2014. Luxapose: Indoor Positioning with Mobile Phones and Visible Light. In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking (MobiCom '14)*. ACM, New York, NY, USA, 447–458.

[25] Hong Li, Wei Yang, Jianxin Wang, Yang Xu, and Liusheng Huang. 2016. WiFinger: Talk to Your Smart Devices with Finger-grained Gesture. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 250–261.

[26] 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 (UbiComp '16)*. ACM, New York, NY, USA, 196–207.

[27] Wenguang Mao, Jian He, and Lili Qiu. 2016. CAT: High-precision Acoustic Motion Tracking. In *Proceedings of the 22Nd Annual International Conference on Mobile Computing and Networking (MobiCom '16)*. ACM, New York, NY, USA, 69–81.

[28] Akhil Mathur, Marc Van den Broeck, Geert Vanderhulst, Afra Mashhadi, and Fahim Kawsar. 2015. Tiny Habits in the Giant Enterprise: Understanding the Dynamics of a Quantified Workplace. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 577–588.

[29] 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 and Networking (MobiCom '13)*. ACM, New York, NY, USA, 27–38.

[30] Kun Qian, Chenshu Wu, Zheng Yang, Chaofan Yang, and Yunhao Liu. 2016. Decimeter Level Passive Tracking with Wi-Fi. In *Proceedings of the 3rd Workshop on Hot Topics in Wireless (HotWireless '16)*. ACM, New York, NY, USA, 44–48.

[31] R. Schmidt. 1986. Multiple Emitter Location and Signal Parameter Estimation. *IEEE Transactions on Antennas and Propagation* 34, 3 (Mar 1986), 276–280.

[32] Moustafa Seifeldin, Ahmed Saeed, Ahmed E. Kosba, Amr El-keyi, and Moustafa Youssef. 2013. Nuzzer: A Large-Scale Device-Free Passive Localization System for Wireless Environments. *IEEE Transactions on Mobile Computing* 12, 7 (July 2013), 1321–1334.

[33] Robert Steele, Amanda Lo, Chris Secombe, and Yuk Kuen Wong. 2009. Elderly persons' perception and acceptance of using wireless sensor networks to assist healthcare. *International Journal of Medical Informatics* 78, 12 (2009), 788–801. Mining of Clinical and Biomedical Text and Data Special Issue.

[34] Li Sun, Souvik Sen, Dimitrios Koutsonikolas, and Kyu-Han Kim. 2015. WiDraw: Enabling Hands-free Drawing in the Air on Commodity Wi-Fi Devices. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (MobiCom '15)*. ACM, New York, NY, USA, 77–89.

[35] 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 (MobiCom '16)*. ACM, New York, NY, USA, 243–256.

[36] Jue Wang, Deepak Vasisht, and Dina Katabi. 2014. RF-IDraw: Virtual Touch Screen in the Air Using RF Signals. In *Proceedings of the 2014 ACM Conference on SIGCOMM (SIGCOMM '14)*. ACM, New York, NY, USA, 235–246.

[37] Wei Wang, Alex X. Liu, and Muhammad Shahzad. 2016. Gait Recognition Using Wi-Fi Signals. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 363–373.

[38] Wei Wang, Alex X. Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2015. Understanding and Modeling of Wi-Fi Signal Based Human Activity Recognition. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (MobiCom '15)*. ACM, New York, NY, USA, 65–76.

[39] Wei Wang, Alex X. Liu, and Ke Sun. 2016. Device-free Gesture Tracking Using Acoustic Signals. In *Proceedings of the 22Nd Annual International Conference on Mobile Computing and Networking (MobiCom '16)*. ACM, New York, NY, USA, 82–94.

[40] Oliver Woodman and Robert Harle. 2009. RF-Based Initialization for Inertial Pedestrian Tracking. In *Proceedings of the 7th International Conference on Pervasive Computing (Pervasive '09)*. Springer-Verlag, Berlin, Heidelberg, 238–255.

[41] Jie Xiong and Kyle Jamieson. 2013. ArrayTrack: A Fine-grained Indoor Location System. In *Proceedings of the 10th USENIX Conference on Networked Systems Design and Implementation (NSDI '13)*. USENIX Association, Berkeley, CA, USA, 71–84.

[42] Jie Xiong, Karthikeyan Sundaresan, and Kyle Jamieson. 2015. ToneTrack: Leveraging Frequency-Agile Radios for Time-Based Indoor Wireless Localization. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (MobiCom '15)*. ACM, New York, NY, USA, 537–549.

[43] Han Xu, Zheng Yang, Zimu Zhou, Longfei Shangguan, Ke Yi, and Yunhao Liu. 2016. Indoor Localization via Multi-modal Sensing on Smartphones. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 208–219.

[44] Lei Yang, Yekui Chen, Xiang-Yang Li, Chaowei Xiao, Mo Li, and Yunhao Liu. 2014. Tagoram: Real-time Tracking of Mobile RFID Tags to High Precision Using COTS Devices. In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking (MobiCom '14)*. ACM, New York, NY, USA, 237–248.

[45] Zheng Yang, Zimu Zhou, and Yunhao Liu. 2013. From RSSI to CSI: Indoor Localization via Channel Response. *ACM Comput. Surv.* 46, 2, Article 25 (Dec. 2013), 32 pages.

[46] Moustafa Youssef, Matthew Mah, and Ashok Agrawala. 2007. Challenges: Device-free Passive Localization for Wireless Environments. In *Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking (MobiCom '07)*. ACM, New York, NY, USA, 222–229.

[47] 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 (MobiSys '15)*. ACM, New York, NY, USA, 15–29.

Received February 2017; revised May 2017; accepted July 2017