mD-Track: Leveraging Multi-Dimensionality in Passive Indoor Wi-Fi Tracking

Yaxiong Xie
Nanyang Technological University

Jie Xiong
UMass Amherst

Mo Li
Nanyang Technological University

Kyle Jamieson
Princeton University

ABSTRACT

Wi-Fi localization and tracking face accuracy limitations dictated by antenna count (for angle-of-arrival methods) and frequency bandwidth (for time-of-arrival methods). This paper presents mD-Track, a device-free Wi-Fi tracking system capable of jointly fusing information from as many dimensions as possible to overcome the resolution limit of each individual dimension. Through a novel path separation algorithm, mD-Track can resolve multipath at a much finer-grained resolution, isolating signals reflected off targets of interest. mD-Track can localize human passively at a high accuracy with just a single Wi-Fi transceiver pair. mD-Track also introduces novel methods to greatly streamline its estimation algorithms, achieving real-time operation. We implement mD-Track on both WARP and cheap off-the-shelf commodity Wi-Fi hardware, and evaluate its performance in different indoor environments.

CCS CONCEPTS

• **Networks → Wireless access points, base stations and infrastructure**;

KEYWORDS

Localization, AoA, ToF, Doppler, Multi-dimensional Estimation, Super-resolution Algorithm, Diversity

Figure 1: Joint estimation in the (a) Time-of-Flight (ToF) and Angle-of-Arrival (AoA) dimensions, (b) ToF, AoA, and Doppler shift or Angle-of-Departure (AoD) dimensions, which can separate incoming signals more effectively.

1 INTRODUCTION

Passive motion tracking without any device carried by or attached to a person has been an exciting area of recent interest, with important applications including security surveillance [28], elderly care [4], and retail business [32]. Diverse technologies have been proposed for localization and tracking including ultrasound [25], infrared [43, 44], cameras [14], and LED visible light [15, 21, 22]. Among these technologies, Wi-Fi based systems stand out as particularly promising due to the pervasive availability of Wi-Fi access points (APs). For device-free passive tracking, Wi-Fi based systems rely on signal reflections off targets to extract essential motion and location information. By its nature, passive tracking is more challenging than localization of active wireless transmitters, because reflected signals of interest are typically orders of magnitude weaker than the direct-path signal, and are typically superimposed with the strong direct-path signal as well as signals reflected from walls, furniture, and other nearby clutter. How to accurately resolve and identify the weak signal reflected off a target of interest becomes a major challenge for these systems.

Recent progress in this area has explored many ways of extracting different *parameters* of the wireless signal, such as angle-of-arrival (AoA) and time-of-flight (ToF) [12, 30, 51] for localization and tracking. These approaches rely on accurately estimating the AoA or ToF of each signal path, and so when multiple paths have similar AoAs or ToFs, these systems face fundamental difficulties resolving the paths to obtain accurate parameter estimates.

Resolvability is determined by the number of antennas (for AoA) and the transmission frequency bandwidth (for ToF), respectively. So a straightforward way to improve AoA resolution is to increase the number of antennas and radios, which results in higher hardware cost. Improving ToF resolution is harder, as Wi-Fi standards fix channel bandwidth. Recent attempts to overcome these limitations include creating virtual antenna arrays by physically moving the antenna and combining adjacent channels to form a larger bandwidth with channel hopping [20, 47, 50, 53]. However, these methods may impose constraints on ongoing data communication that we seek to avoid here, such as the use of a wider bandwidth channel in a situation where use of a narrower bandwidth channel would be more favorable from a communications standpoint. While other recent systems [18, 19] jointly estimate AoA and ToF, they are inherently limited, by design, to the two signal dimensions they can explore, and have high computational complexity, so are not easily scalable to higher dimensions.

In this work, rather than adding more antennas or increasing channel bandwidth, we explore more dimensions of the wireless signal itself. We illustrate the intuition behind this idea in Fig. 1. Three signals arrive at the receiver simultaneously. Signals S_1 and S_2 are close in time and therefore cannot be resolved by employing ToF. However, S_1 and S_2 can be easily resolved with AoA, since their AoA difference is large. Similarly, signals S_1 and S_3 cannot be separated with AoA, but are resolvable with ToF. This concept extends to higher dimensions as shown in Fig. 1(b), where the receiver jointly estimates ToF, AoA, and a third signal parameter (Doppler shift or Angle-of-Departure–AoD). Here S_4 and S_5 are close to each other in both AoA and ToF, but since S_4 is a signal from a moving source, it exhibits a non-zero Doppler shift that separates it from S_5, a reflected signal from a static object with zero Doppler shift. There is thus an opportunity to improve signal resolvability by jointly exploiting information from more signal dimensions, without changing the resolution limit of any individual dimension.

This paper leverages the foregoing opportunity, describing *multi-Dimensional Track* (*mD-Track*), a passive Wi-Fi tracking system that fuses information from multiple signal dimensions, significantly improving resolvability without requiring a wider frequency bandwidth or a larger number of antennas. mD-Track jointly estimates multi-dimensional parameters simultaneously so that the respective parameters corresponding to a single path can be easily associated with that path, improving passive localization and tracking of the motions of multiple targets simultaneously with only a single transmitter-receiver pair. mD-Track makes the following contributions:

1. Multi-dimensional signal estimator. mD-Track introduces a signal processing structure (the *multi-dimensional estimator*, shown in Fig. 3), that combines sources of information from all available signal parameters into a single metric. This structure serves as the building block in our next algorithmic contribution.

2. Iterative path parameter refinement. Separating signals of close-by paths is challenging: reflected signals are much weaker compared to the direct path signal, so it is difficult to accurately estimate their parameters in the presence of interference from a strong direct path. To separate incoming signals, mD-Track employs an iterative path parameter refinement method during which the signals are iteratively re-estimated, more accurately reconstructed, and then subtracted from the received signal with refined parameters in multiple rounds of estimation. This design recalls the structure of the Turbo Decoder [34] and also can be shown to be an expectation maximization estimation algorithm.

3. Bounding computation for real time operation. While multi-dimensional joint estimation helps to improve signal resolvability and parameter estimation accuracy, the computation required by the joint estimator increases exponentially with the number of signal dimensions, making the required amount of computation intractable for a real-time system design. To address this challenge, we design a linear-time estimator by exploiting a coordinate descent method together with the expectation maximization algorithm to reduce computational complexity significantly, making the design practical for real-time operation (§3.3).

We implement mD-Track on both WARP and commercial off-the-shelf (COTS) Wi-Fi APs. Our experimental evaluation begins with a sensitivity analysis that analyses each signal parameter, measuring the relative ability of different parameters to resolve signals. Head-to-head comparisons in parameter estimation and passive localization demonstrate 3.5× accuracy improvements over the state-of-the-art SpotFi [19] system. Further experiments measure the effect of adding another signal dimension (Doppler shift) to mD-Track, showing approximately a 3× accuracy improvement, in contrast with a marginal 20% improvement from doubling the frequency bandwidth. Experiments show that with the three antennas available on the COTS Wi-Fi card, mD-Track can resolve more than 10 signals and estimate the parameters of each signal path accurately.

2 THE WIRELESS CHANNEL

Localization or motion tracking of a target relies on separating superimposed signals and accurately estimating each signal's parameters. A wireless signal is characterized by multi-dimensional parameters, each parameter providing us a piece of location or motion information about the target. Fig. 2 summarizes the parameters that can be retrieved from a wireless signal reflected from a human target. We only consider the path with one or less reflections, since the signal experiencing two or more reflections during the transmission, has extremely low signal strength and can hardly be captured by wireless receiver.
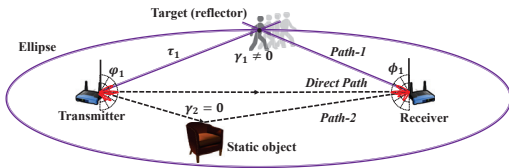

Figure 2: The multi-dimensional parameters of signal paths related to location and motion tracking.

2.1 Parameters of a signal path

Time of flight (τ). The propagation time the signal takes to travel along a particular path from the transmitter to the receiver is referred as the *time of flight* (ToF) τ. As Fig. 2 depicts, ToF estimation of a reflected signal defines an ellipse (with the transmitter and receiver as the two focal points) where the reflector is located. The resolution of ToF estimation is inversely proportional to the channel bandwidth [47].

Angle of arrival (ϕ) and angle of departure (φ). The *angle of arrival* (AoA) ϕ indicates the direction of the signal arriving at the receiver, and the *angle of departure* (AoD) φ indicates the direction of the signal departure from the transmitter, as shown in Fig. 2. The number of antennas at sender/receiver determines the resolution of the AoA/AoD estimates, respectively.

Doppler shift (γ). Movement of the transmitter, receiver, or reflectors all introduce frequency shifts to the carrier frequency of the signal which is referred as *Doppler shift* γ. Doppler resolution is related to the observation interval: the longer the interval, the finer the resolution.

Complex attenuation (α). The signal is attenuated by α when propagating from the sender towards the receiver.

2.2 Wireless signal model

The mD-Track transmitter has an array of N antennas, and receiver has an array of M antennas. They are linear arrays with a uniform spacing d between adjacent antennas, as shown in Fig. 3. If we denote the transmitted signal as $U(t) = [u_1(t), u_2(t), \ldots, u_N(t)]$, then we can use the above signal parameters to express the signal reaches the receiver through a single path as follows:

$$\mathbf{s}(t; v) = \alpha e^{j2\pi\gamma t} \mathbf{c}(\phi) \mathbf{g}(\varphi)^T \mathbf{U}(t - \tau), \tag{1}$$

where $v = [\phi, \varphi, \tau, \gamma, \alpha]^T$ is the *parameter vector* containing the parameters that characterize the signal. The N-element *transmit array steering vector* $\mathbf{g}(\varphi)$ characterizes the phase relationship of the signal coming out of the N transmitting antennas while the M-element *receive array steering vector* $\mathbf{c}(\phi)$ characterizes the phase relationship of the signal arriving at the M receiving antennas [51]. The signal received at the receive antenna due to this single path is then:

$$\mathbf{y}(t) = \mathbf{s}(t) + \mathbf{W}(t), \tag{2}$$

where $\mathbf{W}(t) = [w_1(t), w_2(t), \ldots, w_M(t)]^T$ is M-dimensional complex Gaussian noise capturing the background noise.[1]

3 PARAMETER ESTIMATION

We first describe the design of mD-Track's parameter estimation algorithm in Section 3.1 in a simplified scenario where there is only one path in the environment, and describe in detail how different signal parameters (ToF, AoA, AoD, Doppler shift and attenuation) of that single path can be estimated. Then in Section 3.2 we introduce our joint estimation method to handle multiple signals arriving through different propagation paths. Section 3.3 discusses our approach to make computational complexity tractable.

3.1 Multi-dimensional estimator

This section first considers an environment where there is only one path from sender to receiver. We assume perfect transceivers without phase offsets across radio chains. Such an assumption is justified in Section 4.1 with phase calibration. Our design factors into different modules, each corresponding to one of the above parameters. Multiple modules are employed together to jointly estimate all the parameters.

Channel estimation. Before estimating the parameters, we need to estimate the wireless channels $h_{i,j,k}$ of subcarrier k between each transmit antenna i and receive antenna j. For explanatory purposes, we will consider a running example with N subcarriers and two transmit and two receive antennas as shown in Fig. 3. The task is to estimate:

$$\mathbf{H}_k = \begin{bmatrix} h_{11,k} & h_{12,k} \\ h_{21,k} & h_{22,k} \end{bmatrix}. \tag{3}$$

for all subcarriers $k = [1, 2, \ldots, N]$. To avoid sudden signal level changes, the 802.11n standard [1] multiplies the preamble or *High Throughput Long Training Field* (HT-LTF[2]) with

---

[1] We use $s(t)$ and $s(t; v)$ interchangeably throughout the paper.
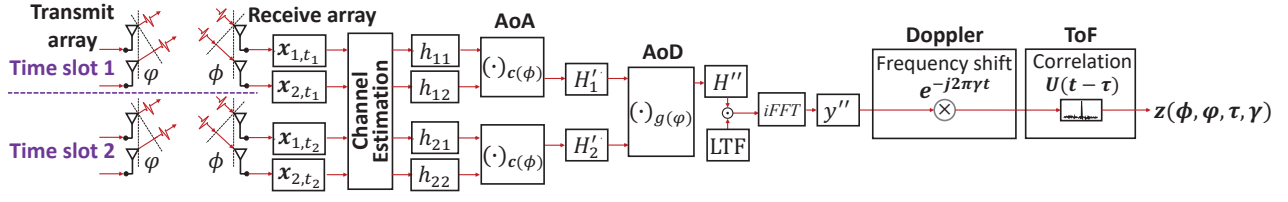[2] We use HT-LTF and LTF interchangeably throughout the paper.

Figure 3: mD-Track's four-dimensional estimator that estimates parameters AoA, AoD, Doppler, and ToF of a wireless signal as it propagates along a single path. Receive and transmit antenna steering vector $\mathbf{c}(\phi)$ and $\mathbf{g}(\varphi)$ are defined in Section 2.2.
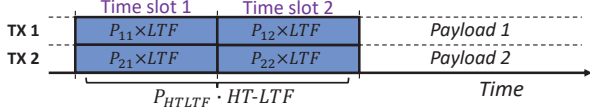

Figure 4: 802.11n (High Throughput) preamble.

the *HT-LTF mapping matrix*:

$$\mathbf{P}_{HTLTF} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \tag{4}$$

whose columns correspond to each of the two time slots and whose rows correspond to each of the two transmit antennas. This results in the frequency domain transmitted signal $\mathbf{P}_{HTLTF} \times \mathbf{LTF}(k)$, as shown in Fig. 4. When both antennas transmit simultaneously, the received frequency domain signal from subcarrier k at the two receive antennas during the two time slots t_1 and t_2 is given by:

$$\begin{bmatrix} \mathbf{x}_{1,k,t_1} & \mathbf{x}_{1,k,t_2} \\ \mathbf{x}_{2,k,t_1} & \mathbf{x}_{2,k,t_2} \end{bmatrix} = \begin{bmatrix} h_{11,k} + h_{21,k} & -h_{11,k} + h_{21,k} \\ h_{12,k} + h_{22,k} & -h_{12,k} + h_{22,k} \end{bmatrix} \times \mathbf{LTF}(k)$$

Since $\mathbf{P}_{HTLTF}$ is known and constant, the receiver estimates the wireless channel by multiplying the received signal with $\mathbf{P}^*_{HTLTF}$, and obtains:

$$\begin{bmatrix} \mathbf{x}_{1,k,t_1} & \mathbf{x}_{1,k,t_2} \\ \mathbf{x}_{2,k,t_1} & \mathbf{x}_{2,k,t_2} \end{bmatrix} \times \mathbf{P}^*_{HTLTF} = 2 \times \begin{bmatrix} h_{11,k} & h_{21,k} \\ h_{12,k} & h_{22,k} \end{bmatrix} \times \mathbf{LTF}(k).$$

Via the above process, the receiver decouples the two simultaneously transmitted preambles to estimate the wireless channel. We note that 802.11's cyclic time delay across different antennas, introduces known linear phase shifts across subcarriers on $h_{21,k}$ and $h_{22,k}$, which are removed by mD-Track after extracting the channel estimates.

AoA estimator. The AoA estimator, *i.e.*, the *AoA* box in Fig. 3, is implemented by multiplying the estimated channel $\mathbf{H}$ with the receive antenna array steering vector $\mathbf{c}(\phi)$ defined in Section 2.2. Specifically, on subcarrier k, we obtain:

$$\begin{bmatrix} h'_{1,k}(\phi) \\ h'_{2,k}(\phi) \end{bmatrix} = \begin{bmatrix} h_{11,k} & h_{21,k} \\ h_{21,k} & h_{22,k} \end{bmatrix} \mathbf{c}^*(\phi). \tag{5}$$

For the wideband Wi-Fi channel with N subcarriers, we apply Eq. 5 to the estimated channels of all subcarriers and

obtain the combined channel:

$$\mathbf{H}'(\phi) = \begin{bmatrix} \mathbf{H}'_1(\phi) \\ \mathbf{H}'_2(\phi) \end{bmatrix} = \begin{bmatrix} h'_{1,1}(\phi) & \cdots & h'_{1,k}(\phi) & \cdots & h'_{1,N}(\phi) \\ h'_{2,1}(\phi) & \cdots & h'_{2,k}(\phi) & \cdots & h'_{2,N}(\phi) \end{bmatrix} \tag{6}$$

Differing AoAs ϕ yield differing steering vectors, so once the correct ϕ is applied to the channel $\mathbf{H}$, the channel between the transmit antenna and the two receive antennas align and add constructively, so the strength of the AoA estimator output is maximized. Therefore, we estimate the AoA by searching for the ϕ^* that maximizes the sum power of the combined channel:

$$\phi^* = \arg\max_\phi \sum_{i=1}^{2} \sum_{k=1}^{N} \left\| h'_{i,k}(\phi) \right\|^2. \tag{7}$$

where $h'_{i,k}(\phi)$ is the combined channel of the kth subcarrier on the ith antenna.

AoD Estimator. When received, the second transmit antenna's signal travels an additional distance to reach the receive array, and hence introduces an extra phase shift. The mD-Track AoD estimator takes $\mathbf{H}'(\phi^*)$ as input and correlates this matrix with the transmit antenna array steering vector $\mathbf{g}(\varphi)$ defined in Section 2.2:

$$\mathbf{H}''(\varphi; \phi^*) = \mathbf{g}^H(\varphi)\mathbf{H}'(\phi^*), \tag{8}$$

where the combined channel $\mathbf{H}''(\varphi; \phi^*)$ is a $1 \times N$ vector. We estimate the AoD by searching for the φ^* that maximizes the sum power of the combined channel:

$$\varphi^* = \arg\max_\varphi \sum_{k=1}^{N} \left\| h''_k(\varphi; \phi^*) \right\|^2, \tag{9}$$

where $h''_k(\varphi; \phi^*)$ is the k-th element of $\mathbf{H}''(\varphi; \phi^*)$, *i.e.*, the channel of the k-th subcarrier. As with the AoA estimator, once the AoD guess is correct, the phase differences due to AoD are removed. The channels thus constructively combine, maximizing the output magnitude.

Doppler and ToF Estimator. We reconstruct the received time domain signal using the channel matrix $\mathbf{H}''(\varphi^*, \phi^*)$ and the known transmitted frequency domain $\mathbf{LTF}$ by *iFFT*:

$$\mathbf{y}''(t; \phi^*, \varphi^*) = \mathcal{F}^{-1}\{\mathbf{H}''(\varphi^*, \phi^*) \odot \mathbf{LTF}\}, \tag{10}$$

where the operator $\odot$ is the Hadamard product, *i.e.,* element wise multiplication of matrix. The signal $\mathbf{y}''(t; \phi^*, \varphi^*)$ that arrives at the Doppler and ToF estimators is a frequency-shifted and delayed version of the transmitted signal $\mathbf{U}(t)$. To estimate that frequency shift and the delay, our approach is to cancel the frequency shift and reverse the delay of the signal and then correlate it with the transmitted signal $\mathbf{U}(t)$. If the correct frequency shift is canceled and correct delay is reversed, there will be a peak in the correlation result.

Specifically, the Doppler shift causes a $2\pi\gamma t$ phase shift to the received time domain signal according to Eq. 1. We therefore cancel such a frequency shift by multiplying the signal by $e^{-j2\pi\gamma t}$. To reverse the delay and obtain $\mathbf{y}''(t + \tau; \phi^*, \varphi^*)$ is practically difficult since $\mathbf{y}''(t; \phi^*, \varphi^*)$ is the received signal. Instead of reversing the delay, we correlate the received signal with the delayed version of transmitted signal $\mathbf{U}(t - \tau)$. For a given ϕ^* and φ^* the correlation is computed as:

$$z(\tau, \gamma; \phi^*, \varphi^*) = \int_T e^{-j2\pi\gamma t}\mathbf{y}''(t; \phi^*, \varphi^*)\mathbf{U}^*(t - \tau)dt, \quad (11)$$

where T is the signal duration of $\mathbf{y}''(t; \phi^*, \varphi^*)$. A peak appears when $\mathbf{U}(t - \tau)$ and $\mathbf{y}''(t; \phi^*, \varphi^*)$ align with each other in both time and frequency.

mD-Track combines the above individual modules into a four-dimensional estimator shown in Fig. 3, for a two-antenna transmitter and receiver. We first estimate AoA at each timeslot, before feeding the output of the AoA beamformer to the AoD estimator. The output of the AoD estimator is then passed to the Doppler estimator to estimate and remove any Doppler shift, and then correlated with a delayed transmit signal $\mathbf{U}(t - \tau)$ to estimate the ToF. The careful reader will notice that mD-Track's processing in this part of the design is sequential. However, the order of processing is carefully chosen: Doppler shift affects all antennas equally, and so the AoA and AoD estimators, which rely on measuring phase differences across different antennas, are unimpaired by arbitrary Doppler shifts. Furthermore, the search in parameter space described in the next section allows us to jointly estimate all parameters. For the general four-dimensional estimator given in Fig. 3, its output is:

$$z(\phi, \varphi, \tau, \gamma) = \int_T e^{-j2\pi\gamma t}\mathcal{F}^{-1}\{\mathbf{g}^H(\varphi)\mathbf{Hc}(\phi)^* \odot \mathbf{LTF}\}\mathbf{U}^*(t - \tau)dt,$$

where the $z(\phi, \varphi, \tau, \gamma)$, denoted as the *z-function*, is calculated for all possible $[\phi, \varphi, \tau, \gamma]$ combinations. The estimation of parameters $[\phi, \varphi, \tau, \gamma]$ is then obtained by:

$$(\phi, \varphi, \tau, \gamma)_{est} = \arg\max_v |z(\phi, \varphi, \tau, \gamma)|. \quad (12)$$

Combining Eq. 1 and Eq. 12 with Eq. 2 then yields α:

$$(\alpha)_{est} = \frac{1}{M \cdot T \cdot P}z((\phi, \varphi, \tau, \gamma)_{est}), \quad (13)$$

where M is the antenna number, T is the signal duration and P is the transmit power. The received signal $s(t)$ is now fully characterized by the parameter vector $v = [\phi, \varphi, \tau, \gamma, \alpha]^T$.

Our design is module-based with a high flexibility. It can be adapted to a three-dimensional $[\phi, \tau, \gamma]$ estimator for the single-antenna transmitter case, and one-dimensional ToF estimator for the single antenna transceiver case.

3.2 Resolving multiple paths

The discussion in the preceding section has considered a single wireless propagation path. We now extend our design to handle multipath propagation. We assume the antenna array receives signals from L distinct paths, denoting the signal from the lth path as $s(t; v_l)$ where $v_l = [\phi_l, \varphi, \tau_l, \gamma_l, \alpha_l]^T$. The received signal is thus the superposition of the signals from all L paths:

$$\mathbf{Y}(t) = \sum_{l=1}^L \mathbf{s}_l(t; v_l) + \mathbf{W}(t). \quad (14)$$

The goal is to estimate the path parameters:

$$\mathbf{V} = [v_1, v_2, \ldots, v_L], \quad (15)$$

for all L paths in $\mathbf{Y}(t)$. Note that the number of multipaths L is also unknown and thus must be estimated.

Resolvability is the premise of accurate path estimation. Non-resolvable signals are merged and the estimated parameters of the combined signal lead to a non-existing path which deviates from the true ones [51, 53]. Although our multi-dimensional estimator in Section §3.1 improves the resolvability by higher dimensional parameters $[\phi_l, \varphi, \tau_l, \gamma_l, \alpha_l]^T$ of each path, the energies of nearby signal paths may still affect the estimation accuracy, e.g., a weaker signal may be overwhelmed by a nearby strong signal and thus not detected. We perform a controlled experiment (with the settings in Section §6.1.1) to demonstrate such an effect, where two signals with AoAs ($60.7°$ and $73.4°$) and ToFs (20.6 and 28.1 *ns*) are created. The signal with ToF 20.6 *ns* is 10 dB stronger than the other. We input the superimposed signal into our 2-dimensional estimator (AoA and ToF) and obtain the z-function as depicted in Fig. 5 (a). It is expected that two peaks will appear in Fig. 5 (a), but the weaker signal, in practice, is dominated by the stronger signal and cannot be detected from the z-function. Similar results are obtained in previous studies using 1-dimensional MUSIC or 2-dimensional SpotFi [19]. A main reason is that those approaches identify the signal paths and estimate the parameters in a single round. No iterative refining is included to accurately determine energy shares between different signal paths.

mD-Track employs iterative parameter refinement during which the path signals are iteratively re-estimated and more accurately reconstructed from refined parameters in multiple
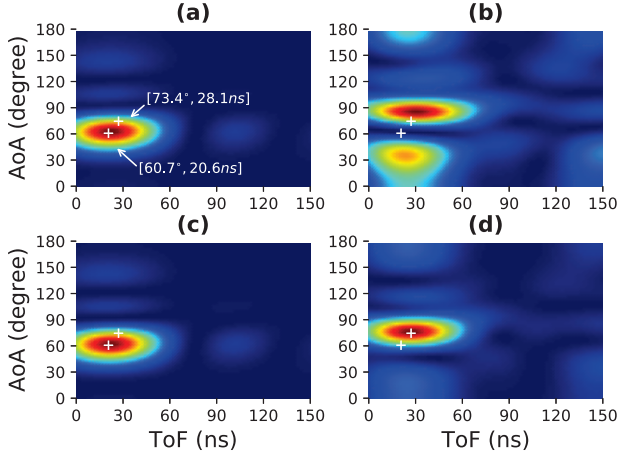
Figure 5: Output (z-function) of the multi-dimensional estimator with input: (a) superimposed signal of a stronger signal with AoA and ToF $[60.7°, 20.8\ ns]$ and a weaker signal with $[73.4°, 28.1\ ns]$; (b) the residual signal after canceling the stronger signal; (c) reconstructed stronger signal and (d) reconstructed weaker signal after three iterations. The peaks of z-function appear at (a)$[63.5°, 22.5\ ns]$, (b) $[85.9°, 30.5\ ns]$, (c) $[61.2°, 21.3\ ns]$, and (d) $[76.7°, 27.0\ ns]$.

rounds. We borrow the idea of successive interference cancellation (SIC) technique from data communications [13, 31] to stepwise estimate and subtract the signal of each path from the received signal and derive the initial estimates as a starting point of the refinement process.

3.2.1 Initial estimation. To process the L superimposed signals $[s_1(t), s_2(t), ..., s_L(t)]$ in decreasing order of signal strength, we first treat all signals but the strongest, *i.e.*, $s_1(t)$, as noise, applying the multi-dimensional estimator to the raw received signal $Y(t)$ to obtain the output z-function, just as in Fig. 5(a). We estimate the parameters v_1 of signal $s_1(t)$ as the highest peak in the z-function. After this estimate, we reconstruct signal $s_1(t)$ using v_1 and cancel it from $Y(t)$. We then calculate the *residual* signal as:

$$\mathbf{y}_2(t) = \mathbf{Y}(t) - \mathbf{s}_1(t) = \mathbf{s}_2(t) + \cdots + \mathbf{s}_L(t) + \mathbf{W}(t). \quad (16)$$

In this residual signal, the second strongest signal $\mathbf{s}_2(t)$ dominates. We then iterate the process, passing $\mathbf{y}_2(t)$ to the multi-dimensional estimator, which now treats $[\mathbf{s}_3(t), ..., \mathbf{s}_L(t)]$ as noise. Now $\mathbf{s}_2(t)$ results in the highest peak in z-function, as shown in Fig. 5 (b). We continue to iterate until all the signals are separately estimated, stopping when the residual power in $\mathbf{y}_L(t)$ is below the dynamic range of the radio. The number of paths L is also obtained. We note that, the final

residual signal:

$$\widehat{\mathbf{W}}(t) = \mathbf{Y}(t) - \sum_{l=1}^{L} \mathbf{s}_l(t), \quad (17)$$

is our estimation of the background noise.

3.2.2 Iterative path parameter refinement. At this point, mD-Track has obtained an estimate of the number of signal paths L and a coarse estimate of each path's parameters. Remaining error arises from two sources. First, although small, interference from weaker signals still affects the estimation of stronger signals. As shown in Fig. 5(a), the peak of the z-function, *i.e.*, $[63.5°, 22.5ns]$, is the estimation result for stronger signal, which deviates slightly from the ground truth, *i.e.*, $[60.7°, 20.8ns]$, due to the existence of the nearby weaker signal. Second, the energy leakage from stronger signal during the cancellation introduces errors to the estimation of the weaker signal. The stronger signal we have reconstructed and canceled is therefore not identical to the true signal due to estimation errors. Energy leaks in the residual signal have a large effect on the estimation of weaker signal and can result in a large estimation error, as shown in Fig. 5(b). To remove the interference, we reconstruct the stronger signal $\mathbf{y}_l(t)$ from l^{th} path with its estimated parameters and the estimated noise as:

$$\mathbf{y}_l'(t) = \mathbf{s}_l(t; v_l) + \widehat{\mathbf{W}}(t), \quad (18)$$

where the $\mathbf{s}_l(t; v_l)$ is reconstructed using the parameters v_l obtained in the first round. The signal $\mathbf{y}_l'(t)$ is now refined, containing less interference compared to the raw signal $\mathbf{y}_l(t)$ in which weaker signals are also considered as noise. We then re-estimate the path parameters of $\mathbf{y}_l'(t)$, v_l', with the multi-dimensional estimator given in §3.1. To reduce the leakage, we keep calculating and updating the noise. Supposing we have obtained new parameters $[v_1', v_2', ..., v_{l_s}']$ for l_s stronger paths, we can update the noise of the $(l_s + 1)^{\text{th}}$ path as:

$$\widehat{\mathbf{W}}(t) = \mathbf{Y}(t) - \sum_{l=1}^{l_s} \mathbf{s}_l'(t; v_l') - \sum_{l=l_s+1}^{L} \mathbf{s}_l(t; v_l), \quad (19)$$

where the $\mathbf{s}_l'(t; v_l')$ is reconstructed using refined parameters v_l' and is closer to the true received signal so that less leakage is contained in the noise. The signal of $(l_s + 1)^{\text{th}}$ path is then reconstructed using updated noise according to Eq. 18.

The path re-estimation and reconstruction is performed for every path, and so we obtain the estimate $\mathbf{V}' = [v_1', ..., v_L']$ for all paths. Now we have obtained a better estimate V', we feed it back into Eq. 18 to start another round of estimation. The above steps of multipath signal re-estimation and reconstruction are iterated, and stop when the difference of each path parameter between two consecutive iterations is smaller than a predefined threshold. The threshold can be flexibly
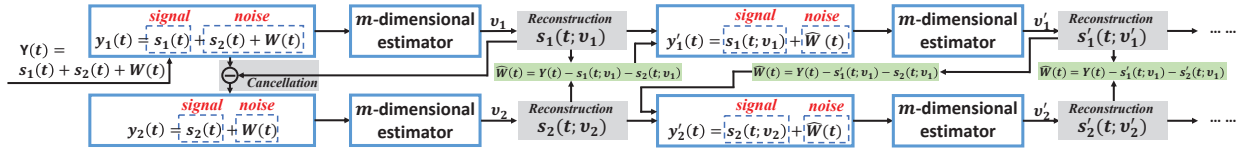
Figure 6: The framework of the iterative path parameter estimation algorithm ($L = 2$ wireless propagation paths). SIC gives the coarse estimation v_1 and v_2. The iterative process keep refining the estimation till convergence.

tuned to meet different levels of accuracy requirements. In our current implementation, the threshold is set as the step size of the parameter search in §3.3, yielding accurate results for passive localization and motion tracking (§6).

Fig. 6 illustrates the workflow of the iterative process based on an example of superimposed signal from two paths. The initial signal cancellation and estimation steps obtain the coarse estimates v_1 and v_2 for the two paths. From there, the iterative process reconstructs the path signals $\mathbf{y}_1'(t)$ and $\mathbf{y}_2'(t)$, re-estimates the path parameters v_1' and v_2', and updates the residual signal $\widehat{\mathbf{W}}(t)$ in each iteration. Fig. 5s (c) and (d) present the z-function of our multi-dimensional estimator after three rounds of iteration. It is clear that the peak in Fig. 5s (c) and (d) is much closer to the true path parameters compared with Fig. 5 (a) and (b).

Convergence. A common issue for iterative algorithms is their convergence. It can be proved[3] that the above iterative process belongs to the EM family of algorithms [9, 10] with the expectation step given by Eq. 18 and maximization step given by Eq. 12, and so convergence is guaranteed [9]. One problem with EM is that it may converge to a local instead of the global maximum if the EM algorithm is not properly initialized. Comprehensive experiments in §6 empirically show that our cancellation-based initialization is able to provide an accurate initial start, ensuring that our iterative algorithm almost always converges to the global maximum.

3.3 Reducing computational complexity

From Fig. 6, we see that the multi-dimensional estimator accounts for a significant portion of mD-Track's computation complexity. Each multi-dimensional estimator solves am ML problem by an exhaustive search. Specifically, the total number of possible combinations that a 4-dimensional estimator that estimates $[\phi, \varphi, \tau, \gamma]$ has to access is given by:

$$\eta = p_\phi \times p_\varphi \times p_\tau \times p_\gamma, \tag{20}$$

where $p_\phi, p_\varphi, p_\tau$ and p_γ are the number of steps for each path parameter, respectively (*e.g.*, $p_\phi = 100$ if the range for AoA estimation is $[0, \pi]$ with a step size of 0.01π). The number of combination η increases exponentially with the number of dimensions, causing unbearable overhead. From Fig. 6, we

---

[3]The proof is not included due to page limitation.

see that the estimator is applied $N_{\text{iter}} \times L$ times, supposing we have L paths and the system iterates N_{iter} rounds.

During each iteration, the input to each multi-dimensional estimator is actually the signal from one path plus the noise, as shown in Fig. 6. Normally, the power level of the signal is tens of dB higher than noise, so we observe a dominant peak in the output z-function of the estimator as Figures 5 (c) and (d) show. Based on this observation, we use a *coordinate descent* method [45] to approach the maximum. We replace the four-dimensional search in Eq. 12 with four one-dimensional searches. Specifically, we fix three of the four parameters (*e.g.* φ, τ, and γ) and search for the value of the fourth parameter(*e.g.*, ϕ) that can generate a maximal output:

$$\phi'' = \arg\max_\phi |z(\phi, \varphi', \tau', \gamma')|, \tag{21}$$

where φ', τ' and γ' are the estimates from the previous iteration. We repeat this process for the other three parameters:

$$\varphi'' = \arg\max_\varphi |z(\phi'', \varphi, \tau', \gamma')|, \tag{22}$$

$$\tau'' = \arg\max_\tau |z(\phi'', \varphi'', \tau, \gamma')|, \tag{23}$$

$$\gamma'' = \arg\max_\gamma |z(\phi'', \varphi'', \tau'', \gamma)|, \tag{24}$$

and obtain an update of all four parameters. By doing so, we reduce the search space from $p_\phi \times p_\varphi \times p_\tau \times p_\gamma$ to $p_\phi + p_\varphi + p_\tau + p_\gamma$, a significant reduction. The trade-off is that the global maximum of each multi-dimensional estimator in Fig. 6 can not be guaranteed [11, 24]. We can only guarantee an increase in the output z-function compared with the start point instead of maximizing it.

Convergence of mD-Track. From Fig. 6 we see that L multi-dimensional estimators are needed in each round of iterations. If every multi-dimensional estimator in this round maximizes its z-function, then the expectation is maximized and mD-Track walks a big step towards the optimal. We relax the requirement of maximization to simply increasing the expectation, which is called generalized EM (GEM) [11, 24]. Coordinate descent method satisfies the increasing of z-function and mD-Track becomes a GEM. For a GEM algorithm, the convergence to maximum (either local or global maximum) is still guaranteed but more iterations may be needed to

achieve the final maximization of expectation [11, 24]. Nevertheless, since we carefully initialize the algorithm and select a reasonably good starting point with the initial cancellation and estimation steps, mD-Track still converges fast. Our experimental results show that mD-Track converges in less than 9 rounds in 90% of cases, even with 10 signal paths. The number of iterations typically reduces to as small as five when there are five signal paths. Therefore, the overall computational overhead is significantly reduced.

4 CHANNEL MEASUREMENT

In this section, we introduce how we estimate the wireless channel using Wi-Fi transceivers. Channel measurement errors are inevitable due to imperfect hardware, and so we also describe how we handle channel measurement uncertainty.

In order to estimate one channel parameter, we have to sample the wireless channel in its corresponding sampling domain. For example, to estimate ToF, we must sample the channel in the frequency domain. Similarly, to estimate a frequency shift that has been added to the signal (*e.g.*, a Doppler frequency shift), we need to sample the channel in the time domain. Channel sampling in the spatial domain is required to estimate signal angle (including AoA and AoD). Table 1 summarizes the channel parameter and its corresponding sampling domain.
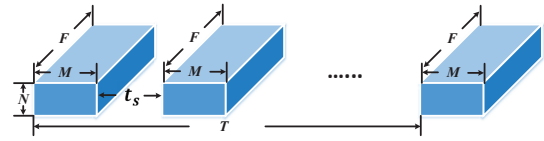
Channel parameter:	ϕ, φ	τ	γ
Sampling domain:	Space	Frequency	Time

Table 1: Channel parameters and their corresponding channel sampling domains.

To estimate multiple parameters, we need to sample the channel in multiple domains. Fig. 7 depicts the channel sampling results for estimating AoA + AoD + ToF + frequency shift. Where N and M are the numbers of transmitting/receiving antennas (spatial domain), F is the subcarrier number (frequency domain), and T is the number of channel samples in time (time domain). If the interval between two channel sampling is t_s, then we sample the channel for the duration of $T \cdot t_s$. The channel sampling results is a matrix of size $N \times M \times F \times T$. Any sub-matrix can be used to estimate a subset of parameters. For example, a channel sampling matrix of size $M \times F$ can be used to estimate AoA + ToF. We note that our parameter estimation algorithm, described in the previous section, can take any dimensional matrix as input and estimate any combination of parameters.

4.1 Measurement Error Handling

Since the hardware we use to measure the wireless channel is imperfect, channel measurement errors are introduced inevitably. For example, it has been observed that a constant,

Figure 7: Channel sampling results for jointly estimating four channel parameters (AoA + AoD + ToF + frequency shift).

time-invariant phase offsets exist for the signals transmitted or received at sender and receiver [51], which affect estimates of AoD and AoA, respectively. Sampling frequency offset (SFO) and symbol timing offset (STO) affect the estimation of ToF. Carrier frequency offset (CFO) itself is a frequency shift that added to the signal, which appears as an actual Doppler frequency shift. Technically speaking, CFO is not an error but will affect the estimation of Doppler shift without proper handling. Table 2 summarizes error sources and the estimation of each parameter they affect.

Error source:	Phase offsets of TX-chains	Phase offsets of RX-chains	SFO, STO	CFO
Parameter:	AoD	AoA	ToF	Doppler

Table 2: Channel measurement error source and the corresponding channel parameter it affects.

Phase offset across radio chains. The wireless signal transmitted or received by multiple antennas experiences a phase shift introduced by the radio chains [51]. The phase offsets across radio chains must be eliminated to provide accurate angle estimation (AoA and AoD). Such a phase offset is constant across time (T samples) and frequency (F subcarriers), and can be measured by connecting the transmit chain and receive chain of WARP or COTS Wi-Fi devices via a coaxial cable, as Fig. 8(a) illustrates, where three transmit chains and receive chains are connected. We, however, observe that spatial multiplexing fails and no packets can be correctly received with such a setup as the channel matrix becomes singular. To break the singularity, we use the setups shown in Fig. 8 (b) and (c). In Fig. 8 (b), we measure the phase offset between: 1) transmit chain 2 and transmit chain 3, denoted as α_2; 2) receive chain 1 and receive chain 2, denoted as β_1. Similarly, phase offsets α_1 and β_2 are measured with the setups shown in Fig. 8(c). The phase offsets are then canceled with the measured value.

SFO and STO. Due to the lack of tight time synchronization between Wi-Fi transceivers, SFO and STO introduce phase errors in the frequency domain and thus affect the estimation of ToF [47, 53]. The SFO and STO are the same across antennas of the same channel sample (one packet) but
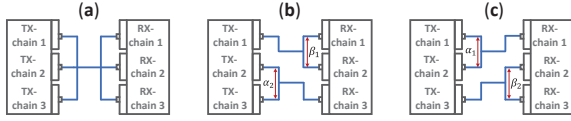
Figure 8: Three TX chains and RX chains are connected together with power splitter/combiner in (a). One TX or RX radio chains are connected to two TX or RX radio chains in (b) and (c).

vary across samples, which results in different phase errors added to the T channel samples of Fig. 7. Estimation based on such a matrix won't give us meaningful ToF results. We thus propose to use the phase of the first channel sample as an anchor and then align the phase of the rest $T-1$ samples to it. Specifically, the phase error introduced by SFO and STO can be modeled as $e_k = k \cdot \lambda$, where k is the index of the subcarrier. The error is added linearly to the phase: $P_k = \theta_k + e_k$, where P_k and θ_k are the measured and the real phase of subcarrier k, respectively. The exact value of slope λ is decided by STO and SFO and is hard to measure. We thus calculate the slope difference between ith channel sample and the first channel sample by $P_{k,i} - P_{k,1} = k \cdot (\lambda_i - \lambda_1)$. We then remove the difference to align the phase errors in the frequency domain.
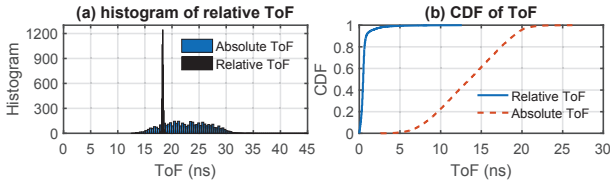
Figure 9: A microbenchmark study of two paths: a) the histogram of the absolute ToF and the relative ToF between two paths; (b) the CDF of the ToF error.

The phase error introduced by SFO and STO is aligned but not removed, which makes it practically difficult to obtain the absolute ToF [19]. We observe that the time delay added by SFO and STO is the same for all the paths. Therefore, the ToF difference between a pair of two paths is invariant even in the presence of SFO and STO. Fig. 7 shows the results from our controlled experiment by connecting radio chains of two QCA9558 together using two RF coaxial cables of different lengths to generate different ToFs ($t_1 = 9.2 \; ns$ and $t_2 = 27.4 \; ns$). We then estimate the ToF of the received signal. Fig. 9 (a) plots the histogram of the estimated absolute ToF values t_1 of the signal from the shorter cable, which exhibits a large variation across different transmissions. On the other hand, we see from this figure that the variation of the relative ToF $|t_2 - t_1|$ between the two paths is minimal. Fig. 9 (b)

presents the CDF of the ToF estimation error. We see that even when the absolute ToF measurements are inaccurate (a median error of 13 ns), the relative ToF estimates can be very accurate (a median error of 0.48 ns).

Based on the above observation, we use the ToF measurement of the direct path as a reference basis to calibrate the ToF estimations of other reflection paths. We first calculate the ground-truth ToF, AoA, AoD of the direct path based on the locations of the transceiver pair. From all signal paths output from mD-Track, we identify the path with shortest ToF and largest amplitude, as the direct path and derive the $\Delta\tau$ between the measured and ground-truth ToF. We then use the derived $\Delta\tau$ to correct the ToF measurements for all reflection paths. In this way, we eliminate the impact of SFO and STO, and thus obtain accurate ToF information.

CFO. Both CFO and Doppler shift are frequency shifts added to the signal and so our estimator is unable differentiate the two. There are, however, two major differences between CFO and Doppler: 1) the magnitude, *i.e.*, CFO is on the order of 100s of Hz and Doppler introduced by a human is only a few Hz; 2) CFO is added to all the multipaths, but Doppler is only introduced to mobile paths. Two important parameters determine the performance of frequency shift estimation. First, the channel sampling interval t_s as Fig. 7 shows, which gives the maximum frequency shift f_m that can be estimated by $f_m = 1/t_s$. Second, the total sampling period $T \cdot t_s$, which gives the minimum frequency difference the estimation can differentiate by $1/(T \cdot t_s)$. The standard Wi-Fi protocol uses two consecutive LTF of one preamble to estimate the CFO, which has small sampling interval (and thus a large frequency range to estimate CFO of 100s of Hz) and small total sampling period (and thus coarse frequency resolution). We reuse the CFO estimation of standard Wi-Fi to get a coarse estimation of CFO and then remove it from the channel sampling results before the actual estimation so that it becomes much smaller. We then sample the channel with multiple packets, just as Fig. 7 shows, with a larger sampling interval $t_s = 25ms$ (smaller estimation range of 40Hz) and larger sampling period 1s (higher frequency resolution of 1Hz). With such a high frequency resolution, the small frequency shift introduced by human and the small residual CFO can be estimated. As CFO is constant for all paths, we subtract the frequency shift of the direct path (a static path) from the frequency estimation of all the paths. A residual frequency shift from Doppler effect remains.

5 IMPLEMENTATION

We implement the core estimation algorithms of mD-Track at a backend server which is a desktop workstation. We use both WARP v3 boards [29] and commodity Wi-Fi routers to sample the channel and send the results back to server.
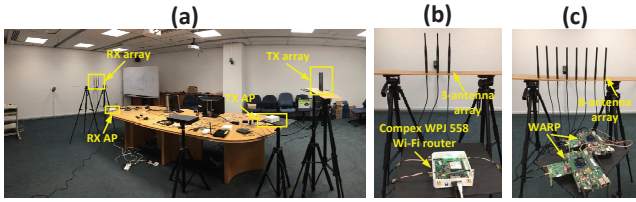
Figure 10: (a) Experimental set up in a indoor meeting room. We use both (b) COTS Wi-Fi AP with three antennas and (c) WARP with at most eight antennas as radio frontend to collect channel measurements.

Version	mD-Track 2D	mD-Track 3D	mD-Track 4D	SpotFi
Parameter	AoA + ToF	AoA + ToF + Doppler	AoA + ToF + AoD + Doppler	AoA + ToF
Input	$M \times 64$	$M \times 64 \times T$	$N \times M \times 64 \times T$	$M \times 56$

Table 3: We implement three different versions of mD-Track as well as SpotFi for comparisons.

WARPs. WARP v3 boards collect time domain IQ samples and send them to a PC connected to the WARP via an Ethernet cable. The PC detects the begining of the packet, extracts the LTF and sends the LTF to the server.

Commercial APs. All of our APs are equipped with Atheros Wi-Fi NIC or SoC (AR9340, AR9580, and QCA9558), and run a customized OpenWRT system. We modify the Wi-Fi driver in kernel space of OpenWRT to enable CSI collection [48]. We build a user-space application that retrieves the received frequency domain LTF by multiplying the transmitted LTF (defined in 802.11 standard) with received CSI (empty subcarriers of CSI and LTF are padded with zeros). We transform the frequency domain LTF into time domain sample via the IFFT, which are sent to the server over the backhaul network.

6 EVALUATION

We conduct experiments in different indoor environments including labs (600 and 420 m^2), meeting rooms (54 and 32 m^2) and corridors. Fig. 10 (a) shows the experimental setup in one of our indoor meeting rooms. Channel measurements collected by commercial APs (Fig. 10 (b)) and WARP (Fig. 10 (c)) are sent to the server via wired backhaul, where parameters are estimated. All APs work in MIMO mode when using multiple antennas.

Multiple versions of mD-Track. We implement three versions of mD-Track, as shown in Table 3. mD-Track 2D uses just AoA and ToF, like SpotFi, while mD-Track 3D adds Doppler and mD-Track 4D adds both Doppler and AoD. The input channel measurement to each version of mD-Track is different. According to Fig. 7 and Table 1, a channel measurement matrix with size $M \times 64$ (20 MHz) is used for mD-Track 2D. The matrix size increases to $M \times 64 \times T$ and $M \times N \times 64 \times T$ for mD-Track 3D and 4D respectively.

SpotFi. We compare the passive localization performance with SpotFi, which jointly estimates the AoA plus ToF for localization[4] Spatial and time smoothing is applied to SpotFi. We note that SpotFi is actually able to estimate the parameters of all the paths, including the reflection paths. We therefore make use of the parameters of the reflection path that SpotFi estimates to passively localize objects. To provide a fair comparison, we use the same input channel measurement matrix of $M \times 64$ for 2D SpotFi and 2D mD-Track. Keeping increasing the matrix for SpotFi, *e.g.*, $N \times 64 \times T$, only provides a marginal improvement for SpotFi since the time domain sampling works for estimation of frequency shift not the AoA or ToF, as Fig. 7 and Table 1 shows.

6.1 Resolving multipaths

In this section, we first present a resolvability analysis of mD-Track which studies its capability to resolve two paths that has similar parameters. We then study the estimation accuracy of the path parameters in each dimension.

6.1.1 Resolvability. We demonstrate that mD-Track has a significantly better ability than previous algorithms to resolve two signals. We conduct controlled experiments by connecting the transmitting and receiving RF chains of two WARPs with coaxial cables, varying the length of cable to provide different propagation times. We rotate the transmitted signal phases to emulate different AoAs, AoDs and Doppler shifts. We vary the ToF, AoA, AoD and Doppler differences between the two signals by a multiple of the parameter's *basic resolution*: the resolution of ToF at a 20 MHz bandwidth is $\Delta\tau = 50$ *ns*, and Doppler at a 1 *s* observation time is $\Delta\gamma = 1$ Hz. Angular resolution of AoA and AoD is $\Delta\phi = 14.2°$ for eight antennas in the array [17]. When two paths are close to each other, they may not be able to be resolved and may be estimated as one path deviated from both the two paths. For each parameter configuration, we transmit 1,000 packets and resolve the two signals with mD-Track 1,000 times using received packets.

Fig. 11 plots the results. The x-axis and y-axis are the path parameter differences relative to basic resolution in that dimension, *e.g.*, 0.2 of ToF means 0.2×50 *ns* = 10 *ns* and the color depth (from blue to red) indicates the probability that two path signals are resolvable in 1,000 estimations. We compare the resolvability of mD-Track with SpotFi and MUSIC. We observe that all algorithms perform better than the basic resolution limit so that two signals with parameter difference smaller than its basic resolution can still be resolved. Furthermore, increasing the number of dimensions significantly increases resolvability. One dimensional MUSIC

---

[4]Wi-Deo [18] is another passive localization system using estimated path parameters. As we do not have the source codes of Wi-Deo, we do not directly compare with it.
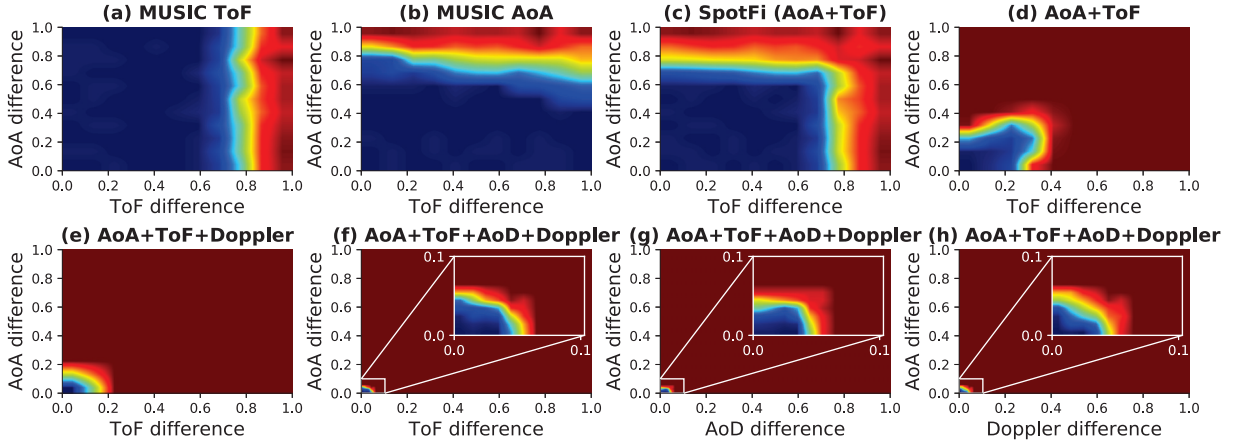
Figure 11: The color indicates the probability of resolving two signals: blue indicates "non-resolvable" and red means "fully resolvable". We compare the resolvability of: (a) MUSIC 1D to estimate ToF; (b) MUSIC 1D to estimate AoA; (c) SpotFi 2D to jointly estimate AoA+ToF; (d) mD-Track 2D to jointly estimate AoA+ToF; (e) mD-Track 3D to jointly estimate AoA+ToF+Doppler with Doppler unchanged; (f) mD-Track 4D to jointly estimate AoA+ToF+Doppler+AoD, with AoD and Doppler unchanged; (g) mD-Track 4D, with ToF and Doppler unchanged; and (h) mD-Track 4D with ToF and AoD unchanged.
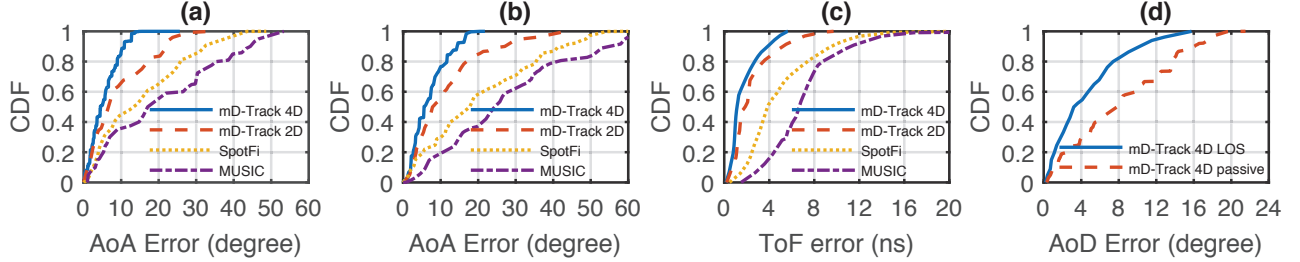

Figure 12: mD-Track runs on COTS AP. CDF of (a) AoA estimation errors of the direct path signal; (b) AoA estimation errors of reflection path signals; (c) ToF estimation errors; (d) AoD estimation errors, for mD-Track, SpotFi, and MUSIC.

is not able to resolve two signals with AoA/ToF difference smaller than 0.75 of its basic resolution. The fraction can be reduced to about 0.4 if applying 2D mD-Track, 0.1 if applying 3D mD-Track and 0.04 if applying 4D mD-Track. At last, even with the same dimension, 2D mD-Track still outperforms 2D SpotFi in Fig. 11 (c) and (d). The improvement stems from the iterative interference cancellation and path refinement process in mD-Track. **Takeaway.** *Increasing the dimension improves the resolvability significantly.*

6.1.2 Estimation Accuracy. We demonstrate that mD-Track can estimate path parameters at high accuracies even using commodity Wi-Fi APs with only three antennas in this section. We collect channel readings using Compex WPJ558 router. We calculate the ground truth of AoA, AoD, and ToF based on the actual positions of the transceivers. We compare AoA and ToF estimation with MUSIC and SpotFi and

favor MUSIC and SpotFi by selecting the path with the AoA estimate closest to the ground truth as their estimate.

AoA and ToF estimation. AoA estimation results are presented in Figs. 12 (a) and (b). mD-Track provides significantly more accurate AoA estimates: with only three antennas, mD-Track's median AoA error for the direct path is as small as 4.4° and 6.2° on 4D and 2D versions respectively, compared to 13.4° and 17.1° from SpotFi and MUSIC. The AoA estimates of the weaker reflection paths are less accurate. Despite that, mD-Track still achieves a median AoA error of 5.6° and 7.3° on 4D and 2D versions respectively, compared with 16.9° and 24.2° for SpotFi and MUSIC. From Fig. 12 (c), we see that mD-Track 4D and 2D are able to estimate ToF with a median error of 1.23 *ns* and 1.85 *ns*, while SpotFi and MUSIC achieve a median error of 3.8 *ns* and 6.7 *ns*.

AoD estimation. Fig. 12(d) gives mD-Track 4D's AoD estimation results for LoS path and reflection path. The median
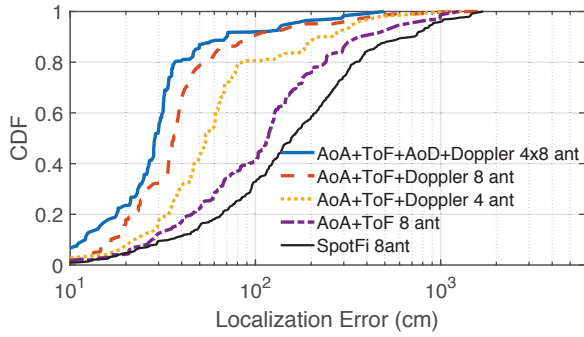
Figure 13: Passive localization error of mD-Track and SpotFi when using the WARP and a 40 MHz channel.
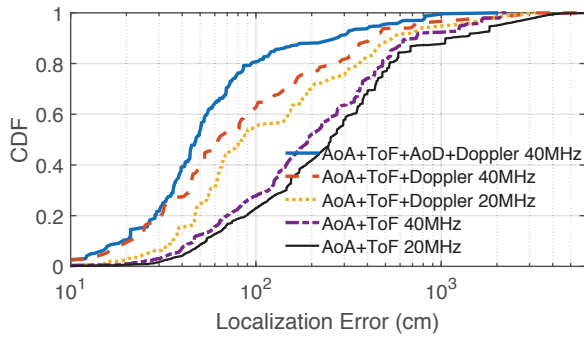
Figure 14: Localization error of mD-Track when using commercial Wi-Fi devices with only 3 antennas.

errors are 3.3° and 7.6° respectively. According to our knowledge, mD-Track is the first system based on COTS Wi-Fi hardware that can estimate AoD at such a high accuracy.

6.2 Passive Localization

In this section, we evaluate the device-free passive localization performance of mD-Track with a single pair of wireless terminals (WARP or COTS Wi-Fi APs).

WARP performance. The results obtained using WARP are presented in Fig. 13. mD-Track 3D and 4D are respectively able to achieve median localization errors of 0.36 m and 0.28 m, with eight antennas and a 40 MHz bandwidth. To the best of our knowledge, no prior Wi-Fi based systems have demonstrated such a high accuracy with a single transceiver pair for passive localization. We also compare mD-Track with SpotFi. For a fair comparison, we implement an AoA and ToF joint estimator based on SpotFi's algorithm and modify it to work with reflected signals. The median error of SpotFi is 1.56 m with eight antennas and 40 MHz. We can see that even mD-Track 2D already achieves significantly better localization performance than SpotFi.

**COTS AP performance.** The results obtained using commodity Wi-Fi APs are presented in Fig. 14. The COTS AP is equipped with only three compared with eight on the WARP platform. We use all three antennas of the AP but vary the size of the bandwidths and the number of dimensions of mD-Track to evaluate performance. As a comparison with WARP, mD-Track 3D and 4D are still able to achieve a median error of 0.67 m and 0.48 m with only three antennas.

Impact of dimensionality. Our system relies on the estimated parameters to localize and track the target. Therefore, accurate parameter estimation is the premise of accurate localization and tracking. The parameter estimation is, however, severely affected by the separability or resolvability of paths. We evaluate how the dimensionality and resolution of each signal domain, *i.e.*, time, angle and space, affect the resolvability, and hence the localization accuracy.

Fig. 13 and 14 depict the localization error with varying signal bandwidth (20 and 40 MHz), antenna count (3, 4 and 8) and dimensionality (2D, 3D and 4D). We clearly see a trend where higher numbers of dimensions lead to higher accuracy. With the same bandwidth and antenna count (*e.g.* 8 antennas and 40 MHz), mD-Track 4D achieves a much smaller median error (0.28 m) compared with mD-Track 3D (0.36 m) and 2D (1.16 m). We see that increasing the number of antennas or bandwidth can improve the performance. By fixing the bandwidth (40 MHz) and doubling the antenna count from 4 to 8 using WARP, mD-Track 3D reduces the median error from 0.57 m to 0.36 m. By fixing the antenna count (3), and doubling the bandwidth from 20 MHz to 40 MHz using COST Wi-Fi AP, mD-Track 3D reduces the error from 0.89 m to 0.67 m. We emphasize that increasing the number of antennas (or radio chains) and bandwidth significantly increases the hardware overhead, especially for COTS device. Increasing the dimensionality, however, does not incur any extra hardware overhead.

Multi-target localization. Passively localizing multiple targets in the same space is a well-known and challenging problem, especially when the targets are close to each other. The capability of resolving signals from two nearby reflectors (which we have demonstrated above), plays a key role in multi-target localization. We ask two human to stand 0.5 m, 1 m, 2 m, and 3 m apart, and wave their hands. It is likely that two targets even physically close-by, have different AoA, ToF, AoD or Doppler, so that 4D mD-Track is able to resolve signals from such two targets and successfully localize both of them, achieving a median error of 0.51 m. As shown in Fig. 15, mD-Track running on the WARP platform is able to locate four targets simultaneously with a median error of 0.47 m when they are 3 m apart from each other. The error is 0.94 m even when they are 0.5 m apart.

Detecting different types of motions. mD-Track is not only capable of tracking the location, but also the motion of the target. In this section, we present the performance of
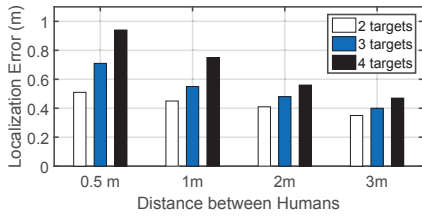
Figure 15: Localization error of mD-Track 4D when tracking multiple targets separated with different distances, using WARP equipped with 8 antennas and transmitting using 40 MHz signals.

Figure 16: Doppler shifts introduced by different motions: (a) the transmitter moves; (b) a human walks; (c) a human stands still and waves his hand; (d) a human stands still and moves his finger.

Figure 17: Estimated ToFs, AoAs and Doppler shifts. mD-Track resolves three static paths in the environment and four mobile paths caused by four hands.

4D mD-Track on motion tracking using WARP. We test the capability of detecting four types of motions with mD-Track. In scenario one, we move the transmitter. In scenario two, the locations of the transceivers are fixed and we let a human target move towards and away from the receiver. In the third and fourth scenario, the human target stands still and only waves his hand and fingers back and forth, respectively. mD-Track estimates the Doppler shift introduced by different motions and the results are depicted in Fig. 16. In Fig.16 (a), we see that mD-Track clearly resolves three multipath signals, and all three multipaths have non-zero Doppler shifts. This is because when the transmitter moves, all the signal paths have Doppler shifts. In Fig. 16 (b), when the human target is moving, mD-Track resolves multiple paths and only one has a non-zero Doppler shift, which is the signal reflected from the human body. Similarly, in Fig. 16 (c), mD-Track resolves multiple paths and only one has a non-zero Doppler shift caused by the hand movements. Comparing Fig. 16 (b) and (c), we observe that the Doppler shift caused by hand movements changes more quickly than the one caused by human body movement. In Fig. 16 (d), we can see that even the Doppler shift introduced by slight finger movements can still be detected.

Multi-motion tracking. We demonstrate that mD-Track can detect multiple the occurrence of multiple things moving. We let two people stand and wave their hands at the same time. mD-Track estimates the ToF, AoA, AoD and Doppler shift of all the resolvable multipaths – the AoA, ToF and Doppler results are shown in Fig. 17. We see three static paths since their Doppler shifts are zero (one of them is the direct path). There are four paths whose Doppler shifts are non-zero – their ToFs and AoAs are different from each other, which correspond to the four waving hands. The four reflected signals from the hands are clearly resolved and the parameters of each path are accurately estimated. mD-Track can use Doppler shifts to detect, as well as ToF, AoA and AoD to locate multiple motions at the same time. For each
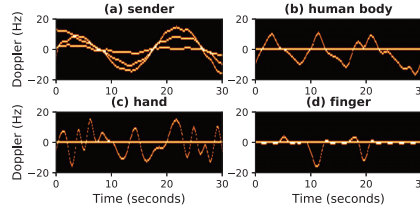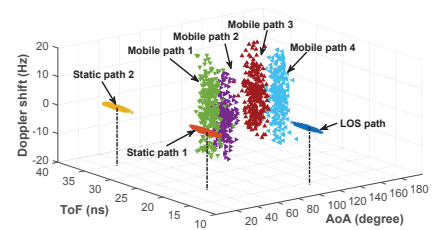
mobile path, we can use their unique time domain Doppler patterns (similar to Fig. 16), to classify pre-defined motions, *e.g.*, gesture recognition [26], activity tracking [40, 42], *etc.*

6.3 Computational complexity

According to our analysis on computational complexity in Section 3.3, four parameters affect the computational complexity of mD-Track: the number of channel paths, the number of iterations mD-Track takes to converge, the dimensionality of estimation, and the number of search steps in each dimension. We run our estimation algorithm using channel measurements collected from different experimental settings and record the number of iterations, number of propagation paths, and the execution time for each trace. Fig. 18 (a) plots the number of iterations mD-Track performs before converging. We see that mD-Track also converges within four iterations 90% of the time for traces with three dominant paths. mD-Track converges quickly within five, eight and nine iterations even when there are five, eight and ten paths, 90% of the time. Fig. 18 (b) shows the overall end-to-end execution time and we see that even with eight multipaths, the median system latency is as small as 130*ms*.

Dimensionality and step size affect the number of combinations for searching. We fix the step sizes of frequency to 0.1*Hz*, vary the step size of angle and time, and evaluate the computational complexity of mD-Track 2D, 3D, and 4D, in comparison with SpotFi. mD-Track 2D estimates AoA and ToF jointly, as SpotFi does. mD-Track conducts multiple one-dimensional searches and thus runs much faster than SpotFi that conducts multi-dimensional search, as depicted in Fig. 18 (c). In our current implementation, we use a step size of [0.02 *rad*, 0.5 *ns*, 0.1 *Hz*] and as a result mD-Track 2D, 3D, and 4D run 30×, 20× and 14× faster than SpotFi.

7 RELATED WORK

Indoor active localization. RSSI based indoor localization techniques [7, 54] provide coarse localization accuracies.
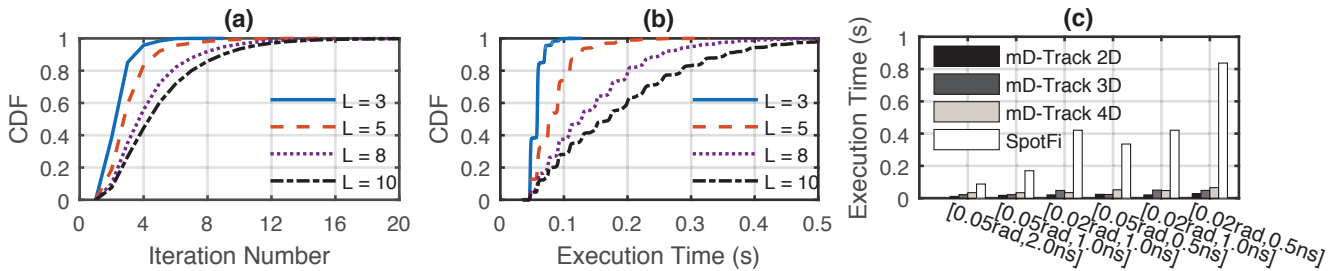
Figure 18: (a) Number of iterations mD-Track takes to converge with different number of multipaths. (b) End-to-end execution time of mD-Track. (c) End-to-end execution time (seconds) of SpotFi and mD-Track with varied granularities (step sizes).

More advanced techniques [12, 51] make use of an antenna array of APs [33, 56] to estimate the AoA of a signal. These systems can achieve accuracies of decimeters, which, however, are vulnerable to multipath propagation because of the poor resolvability due to the limited number of antennas. Moving the antenna mechanically to emulate a large antenna array is proposed [20], but the locations of most APs today are still fixed and the antennas cannot move freely. The latest ToF based systems [38, 47, 52, 53] also provide high accuracies by combining channels to form a virtual wider bandwidth for finer resolution. However, the channel hopping those systems depend on, does affect data communication and there is only a total of 70 MHz frequency bandwidth in the 2.4GHz spectrum. Prior systems in the field tries to increase the resolution in one particular dimension. mD-Track on the other hand jointly estimates multi-dimensional information to achieve a much finer resolution.

Passive localization and motion tracking. A lot of attentions has been paid to passive localization, gesture recognition and Wi-Fi imaging in recent years. Acoustic signal based solutions [36, 41, 57] can only work within a few decimeters. Wi-Fi RSSI/CSI signature based passive localization systems [2, 6, 39, 40, 42, 46, 55] usually rely on high-density deployments which are not realistic for large-scale deployments. Wi-Vi [5] employs a signal nulling technique to cancel signals from static objects and then estimate the AoA of the signal reflected off the human, which can track the direction of human's movement, but no location information can be obtained. WiTrack [3, 4] and other systems [8, 16, 23, 37, 58, 59] use dedicated hardware with more than 1 GHz bandwidth to achieve a high resolution ToF for human tracking which is not possible with Wi-Fi. Although WiTrack 2.0 [3] uses similar signal cancellation techniques as mD-Track, mD-Track takes signal cancellation a step further by integrating it with our iterative path refinement mechanism, maximizing its efficacy.

WiSee [26] recognizes the human gestures by estimating the Doppler shifts introduced to Wi-Fi signal by human gestures. The key intuition of mD-Track and WiSee to address Doppler estimation is the same – we need a long observation time interval to extract a small Doppler shift. Specifically, WiSee transmits multiple consecutive OFDM symbols (for one second) and takes a large FFT over the received OFDM symbols to estimate the small Doppler shift. Similarly, mD-Track sends multiple packets and use the received LTFs to estimate the Doppler shift. We note that mD-Track not only provides motion related information, *i.e.*, Doppler, but also estimated location related information, *i.e.*, AoA, AoD and ToF, for every resolvable multipath.

While some recent systems [18, 19, 27, 35, 49] consider the use of multiple parameters for object tracking, their systems are designed to estimate fixed number of parameters, and cannot be generalized easily to estimate parameters of more dimensions, which mD-Track enables. Furthermore, the spatial smoothing techniques adopted by some prior work increases computational complexity significantly [18, 19]

8 CONCLUSION

mD-Track is a system that incorporates information from as many dimensions as possible to advance the accuracy of passive wireless sensing in a multipath environment. Our experiments demonstrate greatly improved performance for passive multi-target localization and motion tracking. We have also demonstrated how mD-Track can support applications including gesture recognition and Wi-Fi imaging.

REFERENCES

[1] IEEE Standard for Information technology– Local and metropolitan area networks– Specific requirements– Part 11: Wireless LAN Medium Access Control (MAC)and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput. 2009.

[2] H. Abdel-Nasser, R. Samir, I. Sabek, and M. Youssef. Monophy: Mono-stream-based device-free wlan localization via physical layer information. In *IEEE WCNC*, 2013.

[3] F. Adib, Z. Kabelac, and D. Katabi. Multi-person localization via RF body reflections. In *NSDI*, 2015.

[4] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller. 3d tracking via body radio reflections. In *NSDI*, 2014.

[5] F. Adib and D. Katabi. See through walls with WiFi! In *ACM SIGCOMM*, 2013.

[6] H. Aly and M. Youssef. New insights into WiFi-based device-free localization. In *ACM UbiComp Adjunct*, 2013.

[7] P. Bahl and V. N. Padmanabhan. RADAR: an in-building rf-based user location and tracking system. In *IEEE INFOCOM*, 2000.

[8] S. Bartoletti and A. Conti. Passive network localization via uwb wireless sensor radars: The impact of TOA estimation. In *IEEE ICUWB*, 2011.

[9] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B*, 1977.

[10] M. Feder and E. Weinstein. Parameter estimation of superimposed signals using the EM algorithm. *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 1988.

[11] J. A. Fessler and A. O. Hero. Space-alternating generalized expectation-maximization algorithm. *IEEE Trans. on Signal Processing*, 1994.

[12] J. Gjengset, J. Xiong, G. McPhillips, and K. Jamieson. Phaser: Enabling phased array signal processing on commodity WiFi access points. In *ACM MobiCom*, 2014.

[13] D. Halperin, T. Anderson, and D. Wetherall. Taking the sting out of carrier sense: Interference cancellation for wireless LANs. In *ACM MobiCom*, 2008.

[14] H. Hile and G. Borriello. Positioning and orientation in indoor environments using camera phones. *IEEE Computer Graphics and Applications*, 2008.

[15] P. Hu, L. Li, C. Peng, G. Shen, and F. Zhao. Pharos: Enable physical analytics through visible light based indoor localization. In *HotNets*, 2013.

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

[17] D. H. Johnson and D. E. Dudgeon. *Array Signal Processing: Concepts and Techniques*. Simon & Schuster, 1992.

[18] K. Joshi, D. Bharadia, M. Kotaru, and S. Katti. WiDeo: Fine-grained device-free motion tracing using RF backscatter. In *NSDI*, 2015.

[19] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti. SpotFi: Decimeter level localization using WiFi. In *ACM SIGCOMM*, 2015.

[20] S. Kumar, S. Gil, D. Katabi, and D. Rus. Accurate indoor localization with zero start-up cost. In *ACM MobiCom*, 2014.

[21] L. Li, P. Hu, G. Shen, C. Peng, and F. Zhao. Epsilon: A visible light based positioning system. In *NSDI*, 2014.

[22] T. Li, C. An, Z. Tian, A. T. Campbell, and X. Zhou. Human sensing using visible light communication. In *ACM MobiCom*, 2015.

[23] N. Maaref, P. Millot, C. Pichot, and O. Picon. A study of UWB fm-cw radar for the detection of human beings in motion inside a building. *IEEE Trans. on Geoscience and Remote Sensing*, 2009.

[24] R. M. Neal and G. E. Hinton. A view of the em algorithm that justifies incremental, sparse, and other variants. In *Learning in graphical models*, pages 355–368. Springer, 1998.

[25] N. Priyantha, A. Chakraborty, and H. Balakrishnan. The Cricket location-support system. In *MobiCom*, 2000.

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

[27] K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, and Y. Liu. Widar2.0: Passive human tracking with a single wi-fi link. In *ACM MobiSys*, 2018.

[28] Y. Qiao, O. Zhang, W. Zhou, K. Srinivasan, and A. Arora. PhyCloak: Obfuscating sensing from communication signals. In *NSDI*, 2016.

[29] Rice Univ. Wireless Open Access Research Platform (WARP). http://warpproject.org.

[30] S. Sen, J. Lee, K.-H. Kim, and P. Congdon. Avoiding multipath to revive inbuilding WiFi localization. In *MobiSys*, 2013.

[31] S. Sen, N. Santhapuri, R. R. Choudhury, and S. Nelakuditi. Successive interference cancellation: A back-of-the-envelope perspective. In *ACM HotNets*, 2010.

[32] L. Shangguan, Z. Zhou, X. Zheng, L. Yang, Y. Liu, and J. Han. Shop-Miner: Mining customer shopping behavior in physical clothing stores with COTS RFID Devices. In *ACM SenSys*, 2015.

[33] C. Shepard, A. Javed, and L. Zhong. Control channel design for many-antenna mu-mimo. In *ACM MobiCom*, 2015.

[34] B. Sklar. A primer on Turbo Code concepts. *IEEE Comms. Magazine*, pages 94–102, Dec. 1997.

[35] E. Soltanaghaei, A. Kalyanaraman, and K. Whitehouse. Multipath triangulation: Decimeter-level wifi localization and orientation with a single unaided receiver. In *ACM MobiSys*, 2018.

[36] K. Sun, W. Wang, A. X. Liu, and H. Dai. Depth aware finger tapping on virtual displays. In *ACM MobiSys*, 2018.

[37] S. Sur, X. Zhang, P. Ramanathan, and R. Chandra. Beamspy: Enabling robust 60 ghz links under blockage. NSDI, 2016.

[38] D. Vasisht, S. Kumar, and D. Katabi. Decimeter-level localization with a single WiFi access point. In *NSDI*, 2016.

[39] J. Wang, H. Jiang, J. Xiong, K. Jamieson, X. Chen, D. Fang, and B. Xie. Lifs: Low human-effort, device-free localization with fine-grained subcarrier information. In *ACM MobiCom*, 2016.

[40] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu. Understanding and modeling of WiFi signal based human activity recognition. In *ACM MobiCom*, 2015.

[41] W. Wang, A. X. Liu, and K. Sun. Device-free gesture tracking using acoustic signals. In *ACM MobiCom*, 2016.

[42] 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*, 2014.

[43] R. Want, A. Hopper, V. Falcao, and J. Gibbons. The active badge location system. *ACM Trans. on Information Systems*, 1992.

[44] A. Ward, A. Jones, and A. Hopper. A new location technique for the active office. *IEEE Personal Communications*, 1997.

[45] S. J. Wright. Coordinate descent algorithms. *Mathematical Programming*, 2015.

[46] J. Xiao, K. Wu, Y. Yi, L. Wang, and L. M. Ni. FIMD: Fine-grained device-free motion detection. In *IEEE ICPADS*, 2012.

[47] Y. Xie, Z. Li, and M. Li. Precise power delay profiling with commodity Wi-Fi. In *ACM MobiCom*, 2015.

[48] Y. Xie, Z. Li, and M. Li. Precise power delay profiling with commodity wifi. In *ACM MobiCom*, 2015.

[49] Y. Xie, J. Xiong, M. Li, and K. Jamieson. xd-track: Leveraging multi-dimensional information for passive wi-fi tracking. In *ACM HotWireless*, 2016.

[50] Y. Xie, Y. Zhang, J. C. Liando, and M. Li. Swan: Stitched wi-fi antennas. In *ACM MobiCom*, 2018.

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

[52] J. Xiong, K. Jamieson, and K. Sundaresan. Synchronicity: Pushing the envelope of fine-grained localization with distributed MIMO. In *ACM HotWireless*, 2014.

[53] J. Xiong, K. Sundaresan, and K. Jamieson. ToneTrack: Leveraging frequency-agile radios for indoor wireless localization. In *MobiCom*, 2015.

[54] M. Youssef and A. Agrawala. The horus WLAN location determination system. In *ACM MobiSys*, 2005.

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

[56] H. Yu, O. Bejarano, and L. Zhong. Combating inter-cell interference in 802.11ac-based multi-user mimo networks. In *ACM MobiCom*, 2014.

[57] S. Yun, Y.-C. Chen, H. Zheng, L. Qiu, and W. Mao. Strata: Fine-grained acoustic-based device-free tracking. In *ACM MobiSys*, 2017.

[58] Y. Zhou, C. L. Law, Y. L. Guan, and F. Chin. Localization of passive target based on UWB backscattering range measurement. In *IEEE ICUWB*, 2009.

[59] Y. Zhu, Y. Zhu, B. Y. Zhao, and H. Zheng. Reusing 60ghz radios for mobile radar imaging. In *ACM MobiCom*, 2015.