Widar2.0: Passive Human Tracking with a Single Wi-Fi Link

Kun Qian[1], Chenshu Wu[2], Yi Zhang[1], Guidong Zhang[3], Zheng Yang[1], Yunhao Liu[1,4]

[1]Tsinghua University
[2]University of Maryland, College Park
[3]University of Science and Technology of China
[4]Michigan State University
{qiank10,wucs32,zhangyithss,hmilyyz,yunhaoliu}@gmail.com,zgd@mail.ustc.edu.cn

ABSTRACT

This paper presents Widar2.0, the first WiFi-based system that enables passive human localization and tracking using a single link on commodity off-the-shelf devices. Previous works based on either specialized or commercial hardware all require multiple links, preventing their wide adoption in scenarios like homes where typically only one single AP is installed. The key insight underlying Widar2.0 to circumvent the use of multiple links is to leverage multi-dimensional signal parameters from one single link. To this end, we build a unified model accounting for Angle-of-Arrival, Time-of-Flight, and Doppler shifts together and devise an efficient algorithm for their joint estimation. We then design a pipeline to translate the erroneous raw parameters into precise locations, which first finds parameters corresponding to the reflections of interests, then refines range estimates, and ultimately outputs target locations. Our implementation and evaluation on commodity WiFi devices demonstrate that Widar2.0 achieves better or comparable performance to state-of-the-art localization systems, which either use specialized hardwares or require 2 to 40 Wi-Fi links.

CCS CONCEPTS

• **Human-centered computing** → **Ubiquitous and mobile computing**;

1 INTRODUCTION

Recent years have witnessed the ever-fast development of WiFi-based localization and tracking. Fine-grained localization has been achieved with sub-meter accuracy and applicability in non-line-of-sight (NLOS) scenarios [13, 24, 29, 39]. Yet these systems mostly require instrumentation of human body, meaning that some wireless device is carried by a person in order to be localized. This limits
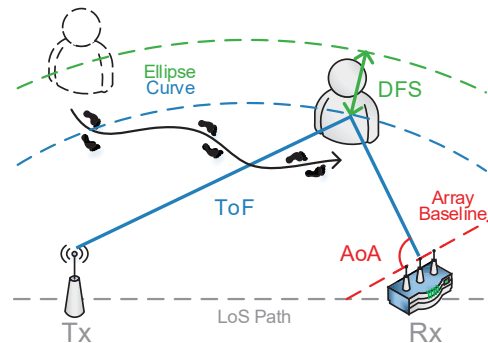
Figure 1: High-level design of Widar2.0. One receiver (e.g. laptop) overhears packet transmission from an AP and calculates ToF, AoA and DFS of the reflection path, for location estimate.

their applications in important scenarios like elderly care, security monitoring, retail analytics, etc. As such, device-free passive localization without any device attached to the user attracts increasing research interests recently [3, 5, 16, 17].

Techniques based on visible light [44] and depth imaging [1] have been proposed and commercialized, yet they only track motions within the vicinity of directional LOS. In contrast, WiFi-based approaches become more popular thanks to the unique advantages of WiFi in its ubiquitous deployment and non-line-of-sight coverage. Generally, passive WiFi tracking works by capturing and analyzing the signals reflected off human body and thus imprinted with a signature of the body motions. Building a passive tracking system, however, is difficult because reflected signals are orders of magnitude weaker than directly received powers and are superimposed with reflections off other objects at the receiver.

Early attempts are devoted to radio tomography [41] and mapping [6] based on coarse-grained RSS or fine-grained CSI, which involve a dense deployment of cooperative WiFi devices and labor-intensive fingerprinting for localization and thus are deficient in practical applicability and accuracy. Recent innovations circumvent the burdensome deployment and training with geometric interpretation of channel parameters such as angle-of-arrival (AoA), Time-of-Flight (ToF), Doppler frequency shift (DFS). In practice, it is a challenging task to precisely estimate any of these parameters due to fundamental limits in frequency bandwidth and antenna number of commercial WiFi as well as noisy channels. To achieve localization in complicated multipath environments indoors, previous works either rely on specialized hardware or software-defined radios [2, 3, 13], rendering them not ubiquitously applicable, or

Kun Qian, Chenshu Wu, Yi Zhang, Guidong Zhang, Zheng Yang, Yunhao Liu

Table 1: A comparison of state-of-the-art works for passive WiFi tracking

Properties	WiTrack [3]	WiDeo [11]	Widar [21]	Dy. Music [16]	IndoTrack [17]	LiFS [32]	**Widar2.0**
Technique	FMCW	FD Wi-Fi	Wi-Fi	Wi-Fi	Wi-Fi	Wi-Fi	Wi-Fi
Parameter	ToF	ToF, AoA	DFS	AoA	AoA, DFS	Attenuation	All
#(Tx,Rx)/Link	(1, 1)/2	(1, 1)/1	(1, 2)/6[1]	(2, 2)/4	(1, 2)/3	(4, 7)/40	(1, 1)/1
#Rx Antenna	1×2	4×1	6×1	3×2	3×3	-	3×1
Tracking range	9 m	10 m	4 m	8 m	6 m	12 m	8 m
Accuracy	0.3 m	0.7 m	0.35 m	0.6 m	0.48 m	0.7 m	0.75 m

require multiple links [16, 32], making them less favorable for practical situations especially in smart home where people would like not to deploy many devices for sensing. Table 1 compares recent RF-based passive tracking systems from aspects of deployment cost and performance. The latest approaches still require at least two links with one AP and two clients for localization and tracking.

As an early attempt of passive tracking, Widar [21] aggregates DFS of CSIs from multiple links, but leaves richful features within CSI unaddressed. In this paper, we present a subsequent work - Widar2.0-, aimed at passive human tracking with only a single pair of COTS WiFi devices (e.g., one access point plus one client), without support from any additional infrastructure or inertial sensors. To avoid use of multiple links, we resort to leverage multidimensional parameters including AoA, ToF, DFS and attenuation from one single link (Figure 1). However, it is non-trivial to gain multidimensional parameter estimation from noisy CSI measurements and translate imperfect parameter estimates into precise locations. In particular, great challenges need to be addressed: How to simultaneously and jointly estimate multiple parameters with limited bandwidth and small antenna array (typically only three on COTS WiFi devices)? How to clean unpredictable phase noises contained in CSI measurements? How to derive fine-grained locations from the low-resolution parameters of a single link, which could be very noisy and erroneous? We overcome all these challenges and propose Widar2.0, a system that realizes the above goal.

First, we build a unified model for simultaneous and joint estimation of multiple parameters including AoA, ToF, DFS, and attenuation. Most of previous works exploit only signal power and DFS for motion sensing [17, 21, 32] since they are the easiest to accurately obtain on commodity off-the-shelf (COTS) WiFi, compared to AoA that is fundamentally limited by antenna number and ToF limited by frequency bandwidth. While deriving any single parameter is challenging enough, we aim to exploit multiple parameters, which will provide multi-dimensional orthogonal and complimentary information, allowing an opportunity to avoid the need of multiple links for location tracking. Towards this goal, we devise a model involving all parameters of interests to quantify the relationships between user motion and CSI, formulate the problem of multiple parameter estimation as a maximum likelihood estimation problem, and employ an Expectation Maximization (EM) solver to efficiently derive accurate estimates.

Second, we eliminate random phase noises between packets via conjugate multiplication of CSI from collocated antennas on the same radio (e.g., the receiver). It is well-known that CSI measurements on COTS WiFi devices suffer from severe phase noises stemming from timing offset and carrier frequency offset. To handle

the noises, previous works proposed to apply linear fitting over multiple subcarriers for calibration [13, 25]. However, this method is not applicable in passive tracking because reflection signals are orders of magnitudes weaker than static signals, and the residual error after calibration is still strong enough to obfuscate reflection signals of interest. In contrast, we propose a novel method using the conjugate multiplication of CSI measurements, which renders an offset-free form of CSI. The rationale lies in that a pair of collocated antennas on one radio device undergo identical phase noises from channel, which can thus be removed out by conjugate multiplication. The resulted conjugate multiplied version of CSI is not polluted with phase noises while can still be leveraged for multiple parameter estimation using the proposed algorithm.

Third, we propose novel algorithms to derive precise locations from erroneous parameter estimates. Our parameter algorithm outputs cluttering parameters of multiple reflections. To locate the target of interest, we devise a novel graph-based algorithm to accurately identify the parameters corresponding to the target's reflection from cluttering multipath parameters. In principle, tracking can then be simply enabled by geometric deduction based on AoA and ranges derived from ToF. However, due to fundamental limits in bandwidth, ranges obtained from inaccurate ToF are not accurate enough for fine-grained tracking. Potential range precision gains lie in combining ToF estimates with relative distance changes calculated from DFS. By doing this, we improve the range estimation resolution from 0.3 m to 0.05 m, underlying the solid foundation for precise localization.

We have implemented Widar2.0 on commodity WiFi devices, i.e., one single antenna AP as transmitter and a laptop with three antennas as receiver, both equipped with Intel 5300 NICs for CSI collection. We evaluate Widar2.0 in three different indoor scenarios and compare it with two state-of-the-art methods namely IndoTrack [17] and DynamicMusic [16]. Experimental results show that, using a single WiFi link, Widar2.0 achieves a median accuracy of about 0.75 m in a 6 m × 5 m tracking area, better than DynamicMusic and comparable with IndoTrack. We believe Widar2.0 brings passive tracking to practical applications especially in scenarios like smart homes and small stores where typically only one AP would be installed and mobile environments where one can rapidly set up a tracking system with simply a laptop and a smartphone, without any fixed infrastructure.

In summary, our key contribution is the first device-free tracking system Widar2.0 that works with only one single WiFi link.

---

[1]Widar puts Rx antennas distantly separate and treats each antenna as an individual link.
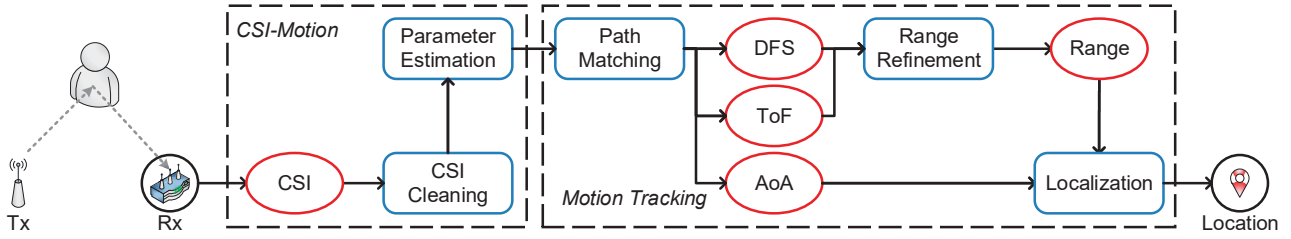
Figure 2: System overview of Widar2.0.

Widar2.0 enables a single COTS WiFi device connected to an existing AP to passively localize a moving target at sub-meter level accuracy without any extra hardware support, be it additional infrastructure or inertial sensors. Widar2.0 also contributes novel multiple signal parameters estimation and fusion algorithms, which underlie precise passive localization with a single link.

The rest of the paper is organized as follows. First an overview is presented in Section 2. We introduce multiple signal parameter estimation in Section 3 and describe tracking in Section 4. Experiments and evaluation are provided in Section 5. We review related works in Section 7 and conclude in Section 8.

2 OVERVIEW

The core of Widar2.0 is to enable sub-meter level passive tracking of moving human by using only a single WiFi link on COTS devices. A key insight underlying the possibility is to seek for multiple signal parameters from a single link rather than using single parameter from multiple links. As shown in Figure 2, we achieve this in Widar2.0 by two key modules, namely *CSI-Motion module* and *Motion Tracking module*.

The CSI-Motion module conducts joint multiple parameters estimation of multipath signals from noisy CSIs. Upon receiving CSI measurements, Widar2.0 first cleans them. The objectives of *CSI cleaning* are two-fold: 1) eliminating random CSI phase noises caused by asynchronous transceivers, and 2) surpassing strong signals from static paths, e.g. the LoS path, and amplifying signals reflected by moving target. Thereafter, Widar2.0 applies the proposed *parameter estimation* algorithm to obtain multidimensional parameters (ToF, AoA, DFS, and attenuation) of multipath signals.

The Motion-Tracking module calculates the target locations from multiple signal parameters estimates. First, Widar2.0 identifies parameters of interest corresponding to the signal reflected by the moving target from cluttering estimates via *path matching* module. Then, the ToF and DFS are fused through a Kalman Smoother to refine distance estimates in *range refinement* module. Finally, the estimated distance and AoA are fed into the *localization* framework to derive the location of the moving target.

In principle, Widar2.0 only needs one single pair of Wi-Fi transmitter and receiver (e.g., an AP and a client). Nevertheless, in case more receivers are available, further accuracy gains would be obtained by combining the location estimations of different receivers. Our design enables lightweight and rapid deployment of passive tracking systems in ubiquitous environments. In particular, not only in traditional environments like homes and offices, Widar2.0 can also be easily deployed in mobile scenarios with, e.g., a smartphone set as AP and a laptop (equipped with three antennas) as a receiver.

Such capability opens up passive tracking to new applications for emergency use or interactive exergames anywhere.

3 MOTION IN CSI

When a target moves, all parameters (i.e. ToF, AoA, DFS and attenuation) of the signal reflected by the target are likely to change. However, as shown in Table 1, existing approaches using commercial Wi-Fi only explore one or two parameters but fail to capture all information of target movements, therefore requiring multiple links for tracking.

A key to circumvent the use of multiple links resorts to inferring multiple parameters from a single link. While DFS and attenuation are relatively easy to obtain on CSI, it becomes much more difficult to estimate ToF and AoA, not to mention all of them simultaneously. In this section, we attempt to jointly derive all of these parameters from CSI measurements, accurately and efficiently. We achieve this by three key components: 1) a novel unified CSI model (§3.1) and its a maximum likelihood formulation, coming together with an efficient solving algorithm (§3.2) for multidimensional parameter estimation; 2) a CSI cleaning technique that removes random phase noises caused by timing offset and carrier frequency offset (§3.3).

3.1 CSI-Motion Model

The wireless channel, due to multipath effect, has the following measurement at time t, frequency f and sensor (antenna) s:

$$
\begin{aligned}
H(t, f, s) &= \sum_{l=1}^{L} P_l(t, f, s) + N(t, f, s) \\
&\triangleq \sum_{l=1}^{L} \alpha_l(t, f, s) e^{-j2\pi f \tau_l(t, f, s)} + N(t, f, s)
\end{aligned}
\tag{1}
$$

where L is the total number of multipath components, P_l is the signal of the l-th path. α_l and τ_l are the complex attenuation factor and propagation delay of the l-th path respectively. N is the complex white Gaussian noise capturing the background noise.

Wi-Fi NICs measure channel discretely in time (packet), frequency (subcarrier) and space (sensor) [10]. Denoting the discrete measurement at the i-th packet, j-th subcarrier and k-th sensor as $H(i, j, k)$, and taking $H(0, 0, 0)$ as reference, the signal phase (divided by 2π) of the l-path in $H(i, j, k)$ is transformed as:

$$
\begin{aligned}
f\tau_l(i, j, k) &= (f_c + \Delta f_j)(\tau_l - \frac{f_{D_l}}{f_j}\Delta t_i + \Delta s_k \cdot \boldsymbol{\phi}_l) \\
&\approx f_c\tau_l + \Delta f_j\tau_l + f_c\Delta s_k \cdot \boldsymbol{\phi}_l - f_{D_l}\Delta t_i
\end{aligned}
\tag{2}
$$

where f_c is the carrier frequency of the channel; Δt_i, Δf_j, Δs_k are differences of time, frequency and spatial position between $H(i, j, k)$ and $H(0, 0, 0)$; And τ_l, ϕ_l and f_{D_l} are the ToF, (unit direction vector of) AoA and DFS of the l-th path in $H(0, 0, 0)$ respectively. The term $-\frac{f_{D_l}}{f}\Delta t_i$ reflects the change of ToF caused by target movement, and the term $\Delta s_k \cdot \phi_l$ is the ToF difference between sensors. The second-order terms are omitted, since they are orders of magnitudes smaller than the linear and constant terms.

Within short time window, narrow bandwidth and small aperture size, the signal attenuation α_l is assumed to be constant for all measurements. In addition, the term $f_c \tau_l$ in Equation 2 is the same for all measurements and can be merged into the complex attenuation α_l, in the view of parameter estimate. Denoting the signal parameter of the l-th path as $\theta_l = (\alpha_l, \tau_l, \phi_l, f_{D_l})$, the first step of tracking is to estimate the multidimensional parameter θ of the signal reflected by the target.

3.2 Joint Multiple Parameter Estimation

We formulate the problem of joint multiple parameter estimation as a maximum likelihood estimation (MLE) problem and designs algorithm for parameter estimate with CSI. Our approach is fundamentally different from previous algorithms like MUSIC and FFT. For brevity, we denote $m = (i, j, k)$, $i = 0, 1, \cdots, T-1, j = 0, 1, \cdots, F-1$, $k = 0, 1, \cdots, S-1$, as the hyper-domain for CSI measurement $H(i, j, k)$, where T, F, S are the number of packets, subcarriers and sensors respectively.

Given measurement observation $h(m)$, our objective is to obtain the MLE of multidimensional multipath signal parameters of $\Theta = (\theta_l)_{l=1}^L$. The log-likelihood function of Θ is [19]:

$$\Lambda(\Theta; h) = -\sum_m |h(m) - \sum_{l=1}^L P_l(m; \theta_l)|^2 \qquad (3)$$

And the MLE of Θ is the solution that maximizes Λ:

$$\hat{\Theta}_{ML} = \text{argmax}_\Theta \{\Lambda(\Theta; h)\} \qquad (4)$$

Obviously, the function is non-linear and no closed-form solution exists. Furthermore, the direct search of Θ_{ML} is computationally prohibitive due to the high dimension of Θ (i.e. $4L$) for large L.

In order to solve the problem efficiently, we apply the Space Alternating Generalized Expectation Maximization (SAGE) algorithm [8] that reduces the overall search space. The SAGE algorithm is an extension of the Expectation Maximization (EM) algorithm [7], where each iteration of the algorithm only re-estimates a subset of the components of Θ while keeping the estimations of the other components fixed. Thus, we can divide the estimate of Θ into multiple estimates of individual parameters.

Parameters of each path are optimized in turn. Specifically, for the l-th path, the expectation step is to decompose the CSI and calculate the signal P_l of the l-th path:

$$\hat{P}_l(m; \hat{\Theta}') = P_l(m; \hat{\theta}_l') + \beta_l\big(h(m) - \sum_{l'=1}^L P_l(m; \hat{\theta}_{l'}')\big) \qquad (5)$$

where $\hat{\Theta}'$ is the parameter estimated in the last iteration. β_l is a non-negative coefficient that controls the convergence rate of the algorithm, and is set as 1 by default.

Algorithm 1 Parameter estimation algorithm

Input: $h(m)$
Output: $\Theta = (\theta_l)_{l=1}^L$
 1: Initialization. $\Theta = 0$;
 2: **while** $||\Theta'' - \Theta'|| > \epsilon$ **do**
 3:     **for** $l = 1$ to L **do**
 4:         Expectation Step, Equation 5
 5:         Maximization Step, Equation 6
 6:     **end for**
 7: **end while**

Then, the maximization step is carried out sequentially as:

$$\hat{\tau}_l'' = \text{argmax}_\tau \{|z(\tau, \hat{\phi}_l', \hat{f}_{D_l}'; \hat{P}_l(m; \hat{\Theta}'))|\}$$

$$\hat{\phi}_l'' = \text{argmax}_\phi \{|z(\hat{\tau}_l'', \phi, \hat{f}_{D_l}'; \hat{P}_l(m; \hat{\Theta}'))|\}$$

$$\hat{f}_{D_l}'' = \text{argmax}_{f_D} \{|z(\hat{\tau}_l'', \hat{\phi}_l'', f_D; \hat{P}_l(m; \hat{\Theta}'))|\} \qquad (6)$$

$$\hat{\alpha}_l'' = \frac{z(\hat{\tau}_l'', \hat{\phi}_l'', \hat{f}_{D_l}''; \hat{P}_l(m; \hat{\Theta}'))}{TFA}$$

where

$$z(\tau, \phi, f_D; P_l) = \sum_m e^{2\pi\Delta f_j \tau_l} e^{2\pi f_c \Delta s_k \cdot \phi_l} e^{-2\pi f_{D_l} \Delta t_i} P_l(m) \qquad (7)$$

Since the MLE of α_l can be derived in a closed form as a function of τ_l, ϕ_l and f_{D_l}, it is calculated at the end of each iteration.

Algorithm 1 summarizes the parameter estimation algorithm. Besides the main E-step and M-step, Θ is initialized as 0, and the iteration ends when the estimates of Θ converge, namely, the difference between successive estimations is within a pre-defined threshold ϵ.

3.3 CSI Cleaning

Unfortunately, the above algorithm is not directly applicable to CSI measurements on commercial WiFi due to their significant noises. As the original purpose of CSI is to equalize channel for data demodulation, the CSI contains not only the channel response, but also various phase noises caused by asynchronization between transceivers and hardware imperfection. Specifically, the erroneous version of CSI measurement $H(m)$ is:

$$\tilde{H}(m) = H(m) e^{2\pi(\Delta f_j \epsilon_{t_i} + \Delta t_i \epsilon_f) + \zeta_{s_k}} \qquad (8)$$

where ϵ_{t_i} and ϵ_f are the timing offset (TO) and carrier frequency offset (CFO) between transceivers, and ζ_{s_k} is the initial phase of the receiver sensor. Initial phases ζ_{s_k} are constant every time the receiver starts up, and can thus be manually calibrated [39]. In contrast, ϵ_{t_i} and ϵ_f vary between packets, and need to be estimated per packet. Thus, it is impossible to directly estimate signal parameters from raw CSIs.

Observing that the phase noises are linear across subcarriers, SpotFi [13] proposes a ToF sanitization algorithm using linear fitting method. However, it is inapplicable to passive tracking with single device, where the problems are two-fold.

First, SpotFi localizes active devices and only needs to estimate the LoS signal between transceivers, which is usually much stronger than human-induced reflections. Figure 3 shows the strength ratio between the reflection signal and the static LoS signal from experiment. Specifically, the transmitter and the receiver are placed 2 m
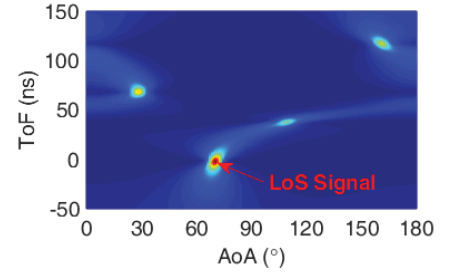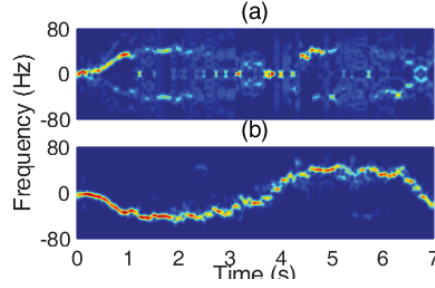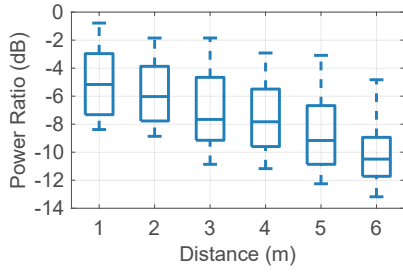
Figure 3: Power ratio between signal re-flected by moving human and static signal.

Figure 4: Spectrograms of CSI with different sanitization methods. (a) SpotFi; (b) Widar2.0.

Figure 5: Pseudo-spectrum of joint AoA and ToF estimations with SpotFi.

apart, and a person walks away from the link, from 1 m to 6 m. As shown, the average ratio is smaller than −5 dB and decays exponentially with the increase of the distance between the person and the link. As a result, the residue of noises after sanitization is still comparable to the reflection signal, leading to severe obfuscation. To validate the concern, we further let a person walk away from and then back towards a Wi-Fi link, which theoretically causes negative DFS first and then positive DFS. Figure 4a shows the normalized spectrograms of CSIs calibrated by SpotFi. Obviously, the reflection signal, as indicated by DFS, is highly interfered by phase noises.

Second, SpotFi's ToF sanitization algorithm removes the absolute ToF of the LoS signal, since the phases contributed by ToF are also linear across subcarriers (recall the term $-2\pi\Delta f_j \tau_l$ in Equation 2). As a result, the absolute ToF of any path cannot be obtained, and the target cannot be localized with only AoA of single device. As an example, we simulate multipath signal with four paths in AWGN channel, where ToFs of the four paths are 50, 90, 120 and 170 ns respectively. Figure 5 shows the corresponding pseudo-spectrum of the signal. While SpotFi accurately captures the difference of ToFs between paths, it fails to obtain the absolute ToF of the target LoS path.

Instead, to filter out irrelevant noises and retain only channel responses of interest, we carefully analyze the structure of noisy CSI and propose the CSI Cleaning algorithm. The basis of the algorithm is that CSI phase noises caused by TO and CFO only vary in time and frequency, but not space. That is, all sensors of the same NIC experience the same unknown phase noises at the same time. Thus, Widar2.0 selects a sensor as the reference sensor (e.g. k_0-th sensor), and calculates the conjugate multiplications $C(m)$, between CSIs of each sensor and the reference sensor:

$$C(m) = \tilde{H}(m) * \tilde{H}^*(m_0) = H(m) * H^*(m_0) \quad (9)$$

where $m_0 = (i, j, k_0)$.

By classifying multipath signals into static ($f_D = 0$) group P_s and dynamic ($f_D \neq 0$) group P_d, the conjugate multiplication can be divided as:

$$C(m) = \sum_{n_1, n_2 \in P_s} P_{n_1}(m) P_{n_2}^*(m_0)$$
$$+ \sum_{l \in P_d, n \in P_s} P_l(m) P_n^*(m_0) + P_n(m) P_l^*(m_0) \quad (10)$$
$$+ \sum_{l_1, l_2 \in P_d} P_{l_1}(m) * P_{l_2}^*(m_0)$$

On one hand, since static signals are constant over time, the first summation term in Equation 10 can be removed via high-pass filter. On the other hand, since static signals are much stronger than signals reflected by moving objects, the third summation term is orders weaker than the first two terms, and can be omitted.

As for the second summation term, for any $l \in P_d$ and $n \in P_s$, according to Equation 1 and 2, we have:

$$P_l(m)P_n^*(m_0) = \alpha_l \alpha_n^* e^{-2\pi\Delta f_j(\tau_l - \tau_n) - 2\pi f_c \Delta s_k \cdot \phi_l + 2\pi f_{D_l} \Delta t_i}$$
$$P_n(m)P_l^*(m_0) = \alpha_n \alpha_l^* e^{-2\pi\Delta f_j(\tau_n - \tau_l) - 2\pi f_c \Delta s_k \cdot \phi_n - 2\pi f_{D_l} \Delta t_i} \quad (11)$$

Note that we omit the term $e^{-2\pi f_c \Delta s_{k_0} \cdot \phi_n}$ in $P_l^{(k)} P_n^{(k_0)*}$ and the term $e^{-2\pi f_c \Delta s_{k_0} \cdot \phi_l}$ in $P_n^{(k)} P_l^{(k_0)*}$, which are the same for all measurements and do not impact parameter estimation.

The first term $P_l(m)P_n^*(m_0)$ has the same phase structure as Equation 2, except that the ToF is $(\tau_l - \tau_n)$. Suppose P_l is the target reflection path and P_n is the LoS path, since transceivers are fixed and their locations are available, the ToF τ_n can be directly calculated from the link distance, and the ToF τ_l can further be derived.

However, conjugate multiplication produces the by-product term $P_n(m)P_l^*(m_0)$, which has fake ToF, AoA and DFS as $(\tau_n - \tau_l)$, ϕ_n and $-f_{D_l}$ respectively. To eliminate the by-product term, we coarsely remove the static responses by subtracting a constant value β from CSI amplitudes of all sensors, and adding a constant value γ to CSI amplitudes of the reference sensor, as in [17]. When $m \neq m_0$, we have

$$|P_n(m)P_l^*(m_0)| = (|\alpha_n| - \beta)|\alpha_l|$$
$$<< |\alpha_l|(|\alpha_n| + \gamma) = |P_l(m)P_n^*(m_0)| \quad (12)$$

Thus, the by-product term can be omitted. As comparison, Figure 4b shows the spectrogram of CSIs calibrated by our conjugate multiplication based method, where DFS is accurately recovered from noisy CSI. In addition, even if the by-product term in $C(m_0)$ cannot be mitigated, we still keep $C(m_0)$ in order for sufficient number of antennas for AoA estimation. $C(m_0)$ can be viewed as the sum two symmetric multipath terms and also applies to the parameter estimation algorithm.

The method of conjugate multiplication of CSIs for phase calibration is first proposed in WiDance [22], and then used in IndoTrack [17] for tracking. However, these works only focus on estimate of DFS, while Widar2.0 addresses all signal parameters
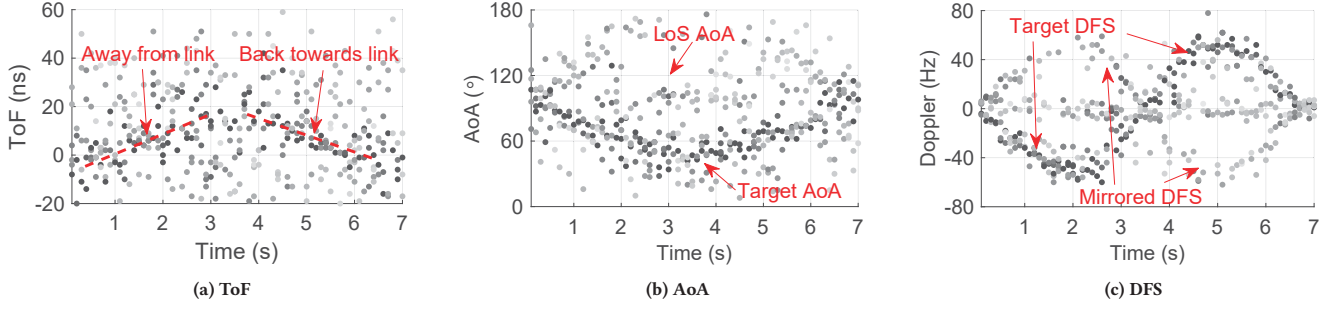
Figure 6: Examples of multipath parameter estimations.
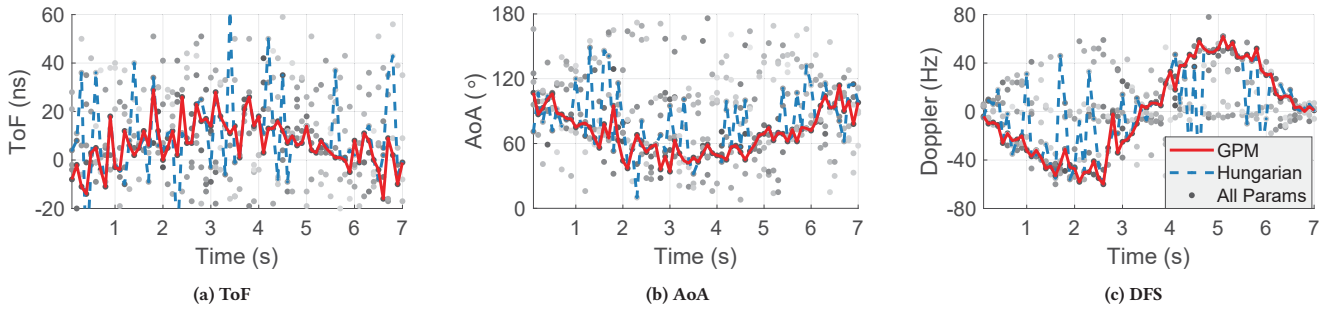
Figure 7: Examples of path matching.

for tracking. With amplitude adjustment and high-pass filtering, the conjugate multiplication $C(\boldsymbol{m})$ only contains significant terms $P_l(\boldsymbol{m})P_n^*(\boldsymbol{m}_0)$, where $l \in P_d$ and $n \in P_s$, and thus has the same structure as CSI measurements $H(\boldsymbol{m})$. Then, we apply the parameter estimation algorithm (Algorithm 1) to $C(\boldsymbol{m})$ to estimate signal parameters. Specifically, Widar2.0 divides CSI collection to 0.1 s segments, where signal parameters are assumed as static. It then applies the parameter estimate algorithm to each segment and obtain an estimation instance of signal parameters.

4 LOCALIZATION

While the parameter estimation algorithm yields multidimensional signal parameters from the cleaned CSI, these estimates are typically erroneous due to multipath effects and low resolution. More precise parameters are desired for accurate tracking. Hence tn this section, we propose to identify the parameters corresponding to the reflection paths of interests from the cluttering estimates (§4.1) and improve the resolution by leveraging parameters in orthogonal dimensions (§4.2). After that, we present a framework (§4.3) based on ToF and distance estimated from single link to localize the target.

4.1 Path Matching

The output of the parameter estimate algorithm consists of parameters of multipath signals. A screening procedure is needed to identify parameters of interests. However, it is not straightforward to select parameters of target's reflection, since all multipath parameters are cluttering together. For example, Figure 6 shows typical multipath signal parameters of a trace, where the target

first walks away from the link and then back to the link. For better illustration, we weight the color of the parameters according to attenuations of corresponding paths. While the target parameters can be indistinctly spotted from the scatter plots, they are confused with parameters of other paths and noises. For example, Figure 7b and 7c show the parameter corresponding to the by-product term in Equation 11. Furthermore, since we filter out most static signals, the remaining part may contain more significant noises, as indicated by spreading outliers.

To select target parameters out of cluttering estimations, WiDeo [11] proposes to use Hungarian algorithm [15] to match parameters of adjacent estimations. However, this algorithm cannot be directly applied to Widar2.0, as CSIs from COTS Wi-Fi devices contain much more noises than signals from backscatters with self-interference cancellation. Figure 7 shows the matching results (blue dash lines) of only using Hungarian algorithm. The results suffer from noises severely, and may lead to large tracking errors.

In order to overcome noises in the cluttering multipath parameters, we propose a graph-based path matching (GPM) algorithm that simultaneously matches successive multiple segments. Formally, as shown in Figure 8, suppose estimations of N CSI segments are considered and each estimation contains parameters of L paths, we build a weighted N-partite graph $G = (V, E, W)$, where $v_{ij} \in V$ represents the parameters of j-th path in the i-th estimation $\boldsymbol{\theta}_{ij}$; $e_{i_1 j_1}^{i_2 j_2} \in E$ represents the edge between $v_{i_1 j_1}$ and $v_{i_2 j_2}$; and $w_{i_1 j_1}^{i_2 j_2} \in W$ represents the weight of the edge $e_{i_1 j_1}^{i_2 j_2}$. The weight is defined as the
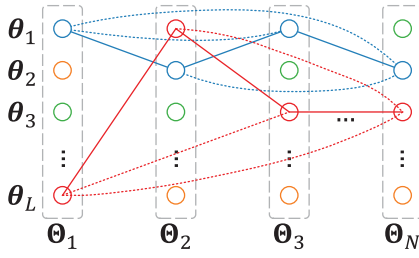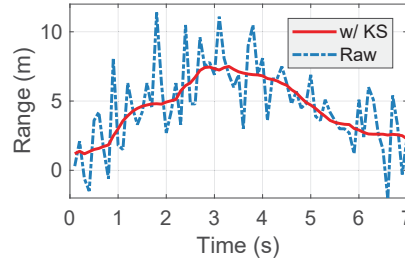
Figure 8: Principle of graph-based path matching.

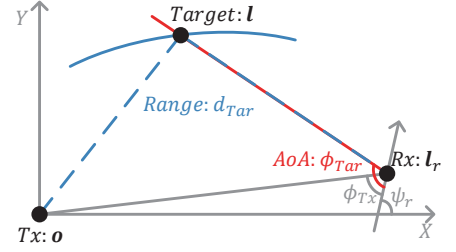**Figure 9: Ranging refinement with Kalman smoother.**

Figure 10: Localization using ToF and AoA.

distance between parameters:

$$w_{i_1 j_1}^{i_2 j_2} = w_{i_2 j_2}^{i_1 j_1} = ||c^T(\theta_{i_1 j_1} - \theta_{i_2 j_2})|| \tag{13}$$

where c is the vector of coefficients that normalize different parameters ToF, AoA, DFS and attenuation.

We denote $x_{i_1 j_1}^{i_2 j_2}$ as binary variable that indicates whether the edge $e_{i_1 j_1}^{i_2 j_2}$ is selected for matching. Thus, the objective function is:

$$x_{\text{opt}} = \text{argmin}_x w^T x \tag{14}$$

where w and x are vectorized weights and variables respectively.

To make sure that selected edges forms L N-order complete graphs, several constraints should be fulfilled.

(i) Edges within the i-th part must not be selected:

$$x_{i j_1}^{i j_2} = 0 \tag{15}$$

(ii) Number of selected edges in the i_1-th part must be equal to the number of vertices in the rest parts:

$$\sum_{j_1=1}^{L} \sum_{i_2=1, i_2 \neq i_1}^{N} \sum_{j_2=1}^{L} x_{i_1 j_1}^{i_2 j_2} = L(N-1) \tag{16}$$

(iii) Number of selected edges between any vertex $v_{i_1 j_1}$ and vertices in the i_2-th part must be no more than 1:

$$\sum_{j_2=1}^{L} x_{i_1 j_1}^{i_2 j_2} \leq 1 \tag{17}$$

(iv) If $e_{i_1 j_1}^{i_2 j_2}$ and $e_{i_2 j_2}^{i_3 j_3}$ are selected, then $e_{i_1 j_1}^{i_3 j_3}$ must be selected:

$$x_{i_1 j_1}^{i_2 j_2} + x_{i_2 j_2}^{i_3 j_3} \leq 1 + x_{i_1 j_1}^{i_3 j_3} \tag{18}$$

This constraint is to ensure that selected edges form complete graphs.

The constraints in Equation 15, 16, 17, 18, together with the objective function in Equation 13, form a binary integer program (BIP) problem, and can be solved via BIP solvers like YALMIP [18].

Since BIP is NP-complete, and signal parameters are likely to change with time, we only calculate the optimal matching of a small group of estimations (e.g. $N = 6$), and concatenate matchings with boundary estimations of groups. Specifically, given successive $2M + 1$ estimations $\Theta_1, \Theta_2, \cdots, \Theta_{2M+1}$ (e.g. $M = 5$), we select $M + 1$ estimations with odd indices and search their optimal matching with BIP. Then, we match the rest estimations (with even indices) with their adjacent estimations using Hungarian algorithm. After completing path matching in one group, we calculate the median

parameters of paths and use it as the new estimation Θ_1 for path matching in next group. Figure 7 shows the matching results (red solid lines) of Widar2.0, which only has fewer outliers. To remove the outliers, Hampel filter is further applied to the matching result.

4.2 Range Refinement

Theoretically, the relative range between the reflection path and the LoS path can be calculated by multiplying estimated ToF with the speed of light. However, the estimated range may suffer from strong noises and low resolution of ToF. Specifically, we adopt a resolution of 1 ns for ToF estimation, which corresponds to that of 0.3 m for range. Thus, a ToF error of a few ns may lead to meters of ranging error. Figure 9 shows the range calculated directly from ToF. Unfortunately, the range estimation is fluctuating and thus cannot be used for localization.

To refine range estimation, we combine absolute yet coarse-grained ToF and fine-grained yet relative DFS, and propose an efficient smoothing algorithm. Specifically, DFS is equivalent to the change rate of path ranging v [21]:

$$f_D = -\frac{v}{\lambda} \tag{19}$$

where λ is the wavelength of the signal, and is about 0.05 m for 5.8 GHz Wi-Fi signal. Thus, a resolution of 1 Hz for DFS estimation corresponds to that of only 0.05 m/s for path range change rate.

Based on this observation, we adopt a Kalman Smoother (KS) to refine ranges from ToF estimations with the change rates of path range from DFS estimations. The process noise and observation noise are initialized as the variance of the first 2 seconds data respectively. Figure 9 shows the ranges refined by Kalman Smoother, which is smoother than the raw estimation. With the relative range, we can further derive the absolute range of reflection path, by adding the constant distance between the transmitter and the receiver.

4.3 Localization Model

Finally, Widar2.0 localizes targets with derived range and AoA. Figure 10 shows the localization framework. Without loss of generality, we denote the locations of the transmitter, receiver and target as $o = (0, 0)$, $l_r = (x_r, y_r)$ and $l = (x, y)$ respectively. The AoA of LoS signal, ϕ_{Tx}, can be calculated from the original CSI measurements using CSI-SAGE algorithm. Further, the orientation of the receiver array ψ_r can be calculated from ϕ_{Tx} and (x_r, y_r). Denoting the range and AoA of the reflection path as d_{Tar} and ϕ_{Tar} respectively,
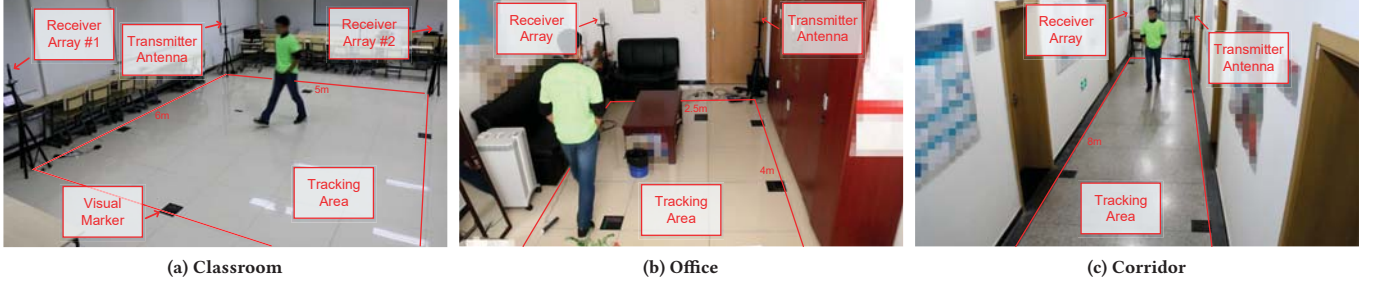
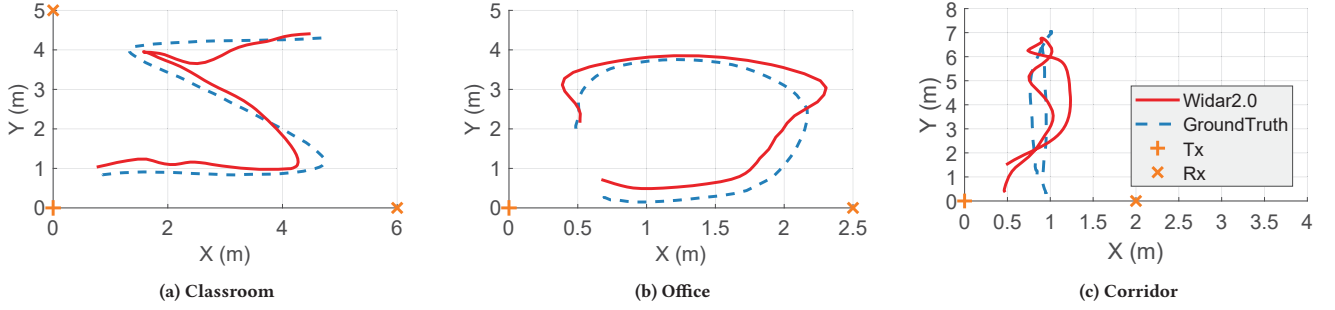**Figure 11: Experimental setup in different scenarios.**

Figure 12: Examples of tracking results of Widar2.0.

we have the following equation system:

$$\begin{cases} \sqrt{x^2 + y^2} + \sqrt{(x - x_r)^2 + (y - y_r)^2} = d_{Tar} \\ (y - y_r)\cos(\psi_r - \phi_{Tar}) = (x - x_r)\sin(\psi_r - \phi_{Tar}) \end{cases} \quad (20)$$

Assuming that the tracking area is at one side of the link, the unique solution can be derived from the equation system, which is the intersection of the semi-ellipse determined by range and the semi-line determined by AoA. Suppose the receiver is on the X boundary of the tracking area, the close-form solution of Equation 20 is:

$$\begin{cases} x = \frac{1}{2}\frac{d_{Tar}^2 + 2s_r d_{Tar} x_r \sec\varphi + x_r^2 \sec^2\varphi - (x_r \tan\varphi - y_r)^2}{x_r + y_r \tan\varphi + s_r d_{Tar}\sec\varphi} \\ y = \tan\varphi(x - x_r) + y_r \end{cases} \quad (21)$$

where $\varphi = \psi_r - \phi_{Tar}$, and $s_r = \text{sgn}\{(x - x_r)\cos\varphi\}$. Since we know the boundary of tracking area, the sign s_r can be calculated by replacing x with an arbitrary X value within the tracking range. The solution for receiver on the Y boundary is dual to the solution in Equation 21, and is omitted for brevity.

In practice, there may exist multiple receivers around the monitoring area. To fully utilize indoor Wi-Fi infrastructure, we further fuse the location results from multiple R receivers:

$$l = \sum_{i=1}^{R} u_i l_i \quad (22)$$

where u_i is the weight for the location estimation of the i-th receiver. Observing that larger DFS f_D leads to more accurate location estimation, we assign the weights heuristically:

$$u_i = \frac{1 + |f_{D_i}|}{R + \sum_{i=1}^{R} |f_{D_i}|} \quad (23)$$

Note that this integration step is only for practical consideration of further improvements in case more than one receivers are available. Widar2.0 itself works gracefully with only one single link.

5 EVALUATION

5.1 Experiment Methodology

Implementation. We implement Widar2.0 using a pair of off-the-shelf laptops equipped with Intel 5300 NIC. The transmitter has one antenna and broadcasts packets into the air. The receiver has three antennas, which forms a uniform linear array. Linux 802.11n CSI Tool [10] is installed in devices to collect CSI measurements. Devices are set to work with monitor mode, on channel 165 at 5.825 GHz. The transmission rate of packets is set to 1000 Hz. The processing computer uses a Intel i7- 7700 3.6GHz CPU, and processes CSI data using MATLAB.

Evaluation setup. To fully evaluate the performance of Widar2.0, we conduct experiments in 3 indoor environments: a large empty classroom, a small office room with various furnitures and a narrow corridor. Figure 11 shows the deployment of devices and tracking areas in different scenarios. In particular, in the scenario of classroom, we deploy an additional receiver to demonstrate the performance of Widar2.0 with existence of multiple devices. The two links are in orthogonal with each other. In total, 6 volunteers (4 males and 2 females) participate in the experiment, and walk along different shapes of trajectories such as line, rectangular, circle, etc. Figure 12 show examples of tracking results of Widar2.0. Code and data samples are available at our official website [2].
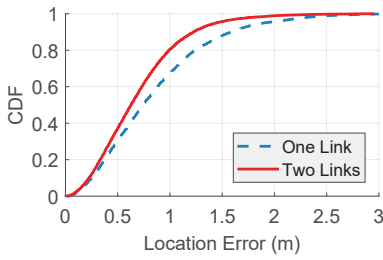
---
[2] http://tns.thss.tsinghua.edu.cn/wifiradar/Widar2.0Project.zip
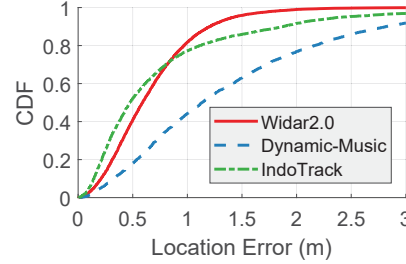
Figure 13: Overall localization accuracy.

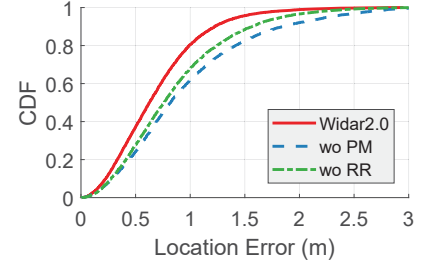**Figure 14: Performance comparison.**

Figure 15: Benefits of individual modules.

Ground truth. We obtain ground truth via video-based tracking solution. Specifically, a digital camera is installed to capture walking videos. Meanwhile, volunteers are asked to wear a light-green T-shirt for easy identification and tracking. The tracking process first calculates the projection matrix between the pixel frame and the world frame with the markers in the field of view. Then, it converts the pixel location of the target point into horizontal 2D location in real world, given the constant height of the target point.

5.2 System Performance

Localization accuracy. We first report the overall performance of Widar2.0. As shown in Figure 13, Widar2.0 achieves an average localization error of 0.75 m, with single Wi-Fi link only. As comparison, using two links improves the performance, with an average localization error of 0.63 m. The performance improvement attributes to not only increasing number of measurements, but also the orthogonal deployment of the two links. Specifically, with two orthogonal links, at least one link is able to capture reflection signal with sufficient large DFS [21], leading to clear extraction of reflections with bandpass filter, during the CSI calibration process (Section 3.3). Noting that walking direction of the target do influence the system performance, we further analyse its impact in Section 5.3.

Comparative study. We compare Widar2.0 with the state-of-the-arts, DynamicMusic [16] and IndoTrack [17]. Specifically, DynamicMusic uses JADE to estimate AoAs of signals reflected by the target at receivers, and pinpoints the intersection of AoAs as target location. IndoTrack further incorporates DFS with AoA for tracking. It recursively calculates target velocity from instantaneous DFS and target location, and then updates target location with newly estimated target velocity. Meanwhile, AoA is used for spotting the initial location of the target, and computing the confidence level of the trace calculated from DFS. Both DynamicMusic and IndoTrack estimate AoAs of reflection signals directly from CSI, which however fails to yield stable AoA estimations as the target is away from links. As an alternative, we use AoAs estimated by Widar2.0 as the input of the two approaches.

Figure 14 shows the system performance of three approaches. First, Widar2.0 significantly outperforms DynamicMusic, which has average localization deviation of 1.1 m. It is mainly because that DynamicMusic only uses AoA and thus cannot compensate AoA errors with movement continuity indicated by DFS. In addition, DynamicMusic fails to pinpoint the target when he is on the LoS path between two receivers. In such ill-conditioned case, AoAs at

two receivers coincide with each other and their point intersection spreads into a long line segment, leading to high localization error.

Second, the average performance of Widar2.0 is slightly worse than but still comparable to that of IndoTrack, which has average localization error of 0.48 m. However, IndoTrack has a much longer error tail than Widar2.0. The reason is that IndoTrack only leverages DFS for direct tracking, yet uses AoA as the confidential indicator of the estimated trace. As a result, while IndoTrack benefits from movement continuity at the start of tracking, its performance gradually degrades with the accumulation of DFS errors. In contrast, Widar2.0 uses both AoA and ToF calibrated by DFS for tracking, and thus avoids error accumulation. Section 5.3 further analyses the impact of walking distance.

Benefits of individual modules. This part studies the impacts of the proposed path matching (PM) and range refinement (RR) processes on system performance. Figure 15 shows the effects of the two steps. On one hand, without the PM process, the average localization error increases to 0.84 m. It demonstrates that the PM process is more robust to estimation noises and able to match more correct parameters of the signal of interest, in comparison with the Hungarian algorithm. On the other hand, the tracking accuracy is improved by 13 cm with the RR process. Specifically, the RR process smooths ToF-based range with DFS, and thus benefits from continuity of target movement.

5.3 Parameter Study

Impact of walking direction. To evaluate how walking direction impacts the system performance, we ask volunteers to walk along lines with various directions and track their locations with single link, as well as two orthogonal links. Figure 16 shows the distribution of localization errors with walking directions. Specifically, with single link, localization error statistically increases as targets tend to walk in parallel with the link, as DFS of reflection signals observed by the receiver becomes smaller and performance of the calibration process degrades.

As comparison, the performance improvement with additional link is from two folds. First, the level of localization error across all walking directions is reduced, as the system is more robust to noises with measurements from multiple receivers. Second, the error distribution along walking directions is more uniformed, since with two links can signals with significant DFS always be captured, leading to more accurate estimations of signal parameters.

Impact of walking distance. We further explore the impact of walking distance. Specifically, we ask volunteers to keep walking for
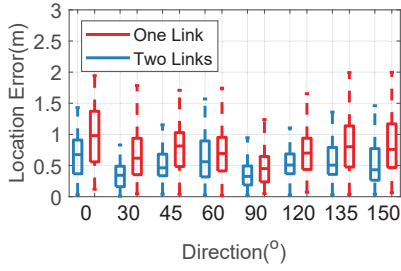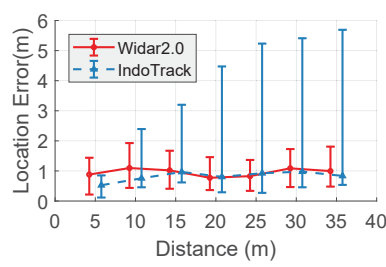
Figure 16: Impact of walking direction.

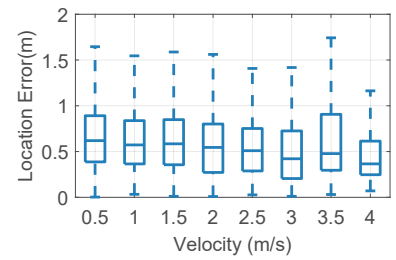**Figure 17: Impact of walking distance.**
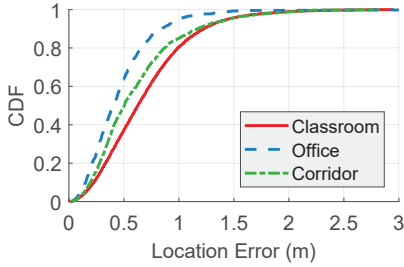

Figure 18: Impact of velocity.

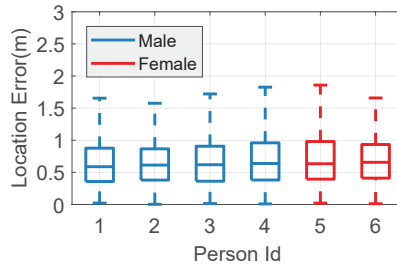**Figure 19: Impact of environment.**

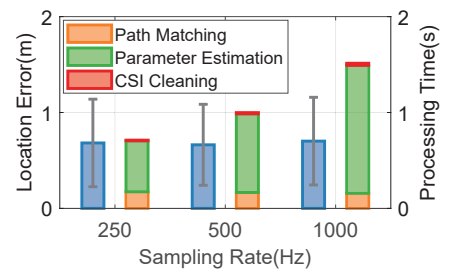
Figure 20: Impact of different persons.

Figure 21: Impact of packet rate.

a long distance of about 40 m in the monitoring area, and evaluate two approaches, Widar2.0 and IndoTrack with collected traces. Figure 17 shows the average localization errors of the two approaches. The error bar indicates the 10-percentile and 90-percentile error boundaries. As shown, while IndoTrack has small localization errors when the walking distance is short, its location results happens to drift away as the distance increases. The reason is that IndoTrack only uses DFS to estimate target velocity and further update target location, which suffers from accumulation of DFS errors. In contrast, Widar2.0 uses absolute ToF to avoid error accumulation and achieves consistent localization accuracy, and is thus feasible for continuous tracking.

Impact of walking velocity. During experiments, volunteers are allowed to walk freely at various speeds. To evaluate the impact of walking velocity, we compute groundtruth velocity from ground truth location captured by the digital camera. Figure 18 shows the distribution of location error with walking velocity. Typically, human walking velocity is no more than 4 m/s. As shown, the location error is stably around 0.6 m, yet gradually reduces as the walking velocity increases. The observation is intuitive, since large walking velocity leads to large DFS at receivers, which may facilitate the calibration process to extract reflection signals.

Impact of environment. In order to demonstrate the influence of environment diversity, we conduct experiments in three different scenarios: classroom, office and corridor, as in Figure 11. Note that two receivers are used in the classroom, while one receiver is used in the office and corridor. Figure 19 shows the localization errors of Widar2.0 at different spots. As shown, Widar2.0 achieves low average localization errors of 0.63 m, 0.4 m and 0.51 m respectively. Since Widar2.0 only relies on the reflection signal with non-zero DFS and the dominant LoS signal, it is robust to multipath effect in indoor environments, and is applicable to various indoor environments.

However, the system performance still varies among different environments, which is due to two factors. First, the system performance is inversely related to the size of tracking area. Specifically, the tracking areas in the classroom, office and corridor are about 30 m^2, 10 m^2 and 20 m^2 respectively. With larger tracking area, the signal reflected by targets tends to be weaker and more vulnerable to noises, leading to degradation of system performance.

Second, distance between the transmitter and receiver (i.e. link length) also impacts the system performance. Given the walking direction and velocity of the target, the larger the link length is, the smaller the DFS is observed at the receiver, which degrades the system performance. Specifically, the link length in classroom (6 m) is much larger than that in the office (2.5 m) and corridor (2 m), explaining the deterioration of system performance in the large classroom.

Impact of human diversity. To determine whether Widar2.0 consistently works for different users, 6 volunteers, with different genders, heights and body shapes, are recruited to participate the experiment. Before the experiment, volunteers are only explained with basic experimental settings, such as tracking area and typical walking traces. They are not specially trained and walk according to their own habits. Figure 20 plots the localization errors of Widar2.0 with different targets. As shown, Widar2.0 achieves consistent localization accuracy across all targets, without knowing any body features of targets.

Impact of packet rate. To find out the minimum packet rate required for Widar2.0 to correctly work, we initialize the packet rate at the transmitter as 1000 Hz, and gradually discard CSI collections to achieve packet rates of 500 Hz and 250 Hz. Figure 21 shows the localization errors with different packet rates (blue bars), which remains almost unchanged as the packet rate exponentially decreases from 1000 Hz to 250 Hz. It demonstrates that Widar2.0

works with moderate packet rate, and is practical with real Wi-Fi transmissions. However, further reducing the packet rate may lead to aliasing of DFS estimation. Specifically, as the human walking velocity is no more than 5 m/s, the range of corresponding DFS is within $\pm f_{D_{\max}} = \pm 100$ Hz. Suppose packets are evenly transmitted with interval of Δ_t, in order to uniquely determine DFS, the atomic DFS-induced phase, $2\pi f_D \Delta_t$ should be within the range of 2π. Thus, the maximum packet interval required by Widar2.0 is $\Delta_{t_{\max}} = \frac{1}{2f_{D_{\max}}} = \frac{1}{200}$ s, which corresponds to a minimum packet rate of about 200 Hz.

We further evaluate the per-second computation cost of steps in Widar2.0, as shown in Figure 21 (stacked bars). As steps of range refinement and localization cost less than 1 ms, we only consider processing time of the rest steps, i.e. CSI cleaning, parameter estimation and path matching. As shown, the major time cost comes from iterative optimization in the step of parameter estimation. By decreasing the transmission rate, the computation cost is reduced. Given that a transmission rate of 250 Hz is sufficient, the per-second processing time is 0.7 s, enabling real-time tracking with Widar2.0.

6 DISCUSSION

Multiple person tracking. We conduct preliminary experiments where two persons walks in the monitoring area. The results show that while Widar2.0 accurately estimates DFS and recognizes two persons, it fails to yield accurate AoA and ToF, and further track multiple persons. The reasons are two fold. First, the NIC has only 3 antennas, limiting the resolution of AoA for separating two reflection paths. Second, the channel bandwidth is only 20 MHz, resulting in small phase change by ToF and erroneous ToF estimation. In revision, we plan to combine multiple Wi-Fi NICs on one receiver [9], and splice multiple channels [37] for fine-grained AoA and ToF estimation for multiple person tracking.

Tracking in NLoS condition. Two types of LoS conditions should be satisfied to enable track with Widar2.0. First, the LoS path between any transceiver and the person should exist, since Widar2.0 requires the ToF of signal directly reflected by the person for localization (Section 4.3). Second, the LoS path between two transceivers should exist, since Widar2.0 can only estimate the difference of ToFs between the reflection signal and the strongest LoS signal of the link (Section 3.3). In cases where the LoS conditions are not fulfilled, ToF estimation becomes erroneous and cannot be used. However, as there are multiple links in typical indoor environments, it may group tracking results of these links and filter out outliers with NLoS conditions [40].

Device deployment. Displacement of devices acts as a main factor limiting tracking range of Widar2.0. On one hand, reflection signal is much weaker than LoS signal, due to longer propagation distance and additional reflection loss, and is hard to be captured in CSI. Thus, it is likely to displace devices around the height of human body to increase reflective surface. On the other hand, Widar2.0 recognizes reflection signal with its non-zero DFS, which is equivalent to the change rate of reflection path length [21]. However, increase in the distance between transceivers may reduce the change rate, and make it more difficult to distinguish reflection signal from static signals. So the link length should be controlled within a moderate range, which is 6 m verified through experiments.

7 RELATED WORK

Our work is broadly related to research in wireless sensing, which studies RF channel characteristics and derives both syntactic (e.g. location, velocity) and semantic (e.g. human activity) environmental contexts.

Device-based Localization. Device-based localization has been an area of active research in the last decade. Generally, it requires objects to carry devices that transmits RF signals, and calculates signal parameters, e.g. AoA [9, 12, 23, 24, 39], ToF [29, 37, 39, 40], using fine-grained channel state information [43]. SpotFi [13] applies JADE [28] algorithm to jointly estimate AoA and relative ToF of dominant incident signals. A pioneer work Splicer [37] splices the CSI measurements from multiple Wi-Fi channels, which greatly extends the bandwidth available and yields sub-nanosecond TDoA for accurate localization. Chronos [29] further calculates accurate sub-nanosecond ToF by leveraging phase differences between subcarriers spanning multiple Wi-Fi channels. WiCapture [14] achieves centimeter-level tracking accuracy by modelling CSI phase changes caused by transmitter motion. Since Wi-Fi transceivers are not synchronized, these works require irregular communication steps [29, 37] to splice multiple channel for accurate ToF estimation. In contrast, Widar2.0 embraces the multipath characteristics of Wi-Fi channel and calculates ToF of reflection path with only normal packet transmission.

Device-free Tracking. Various hardware [2–4, 11] are manufactured to capture extremely weak human reflections. WiTrack [2, 3] develops FMCW radar to accurately estimate ToFs of reflections in frequency domain. WiDeo [11] uses full-duplex Wi-Fi that enables self-interference cancellation, and jointly estimates ToFs and AoAs to localize all reflectors. xDTrack [38] applies SAGE algorithm to Wi-Fi signals to jointly estimate multi-dimensional signal parameters on SDR platforms. While xDTrack inspires Widar2.0, our work advances in modeling discrete CSI measurements, tackling with unknown CSI phase noises and supporting COTS Wi-Fi devices. Furthermore, Widar2.0 leverages multipath effects to enable estimation of absolute ToF.

Since dedicated hardware are difficult to be generalized, researches are shifted to ubiquitous COTS RF devices, such as RFID [26, 27, 33, 42], millimetre wave [35, 45] and Wi-Fi [21, 32]. Comparing with RFID and millimetre wave, Wi-Fi is much more ubiquitous in daily life, but it suffers from unknown phase noises [34] and cannot directly use signal phase for passive tracking. Instead, LiFS [32] employs radio tomography imaging approach for localization, which requires dense deployment of devices. WiDar [21] derives DFSs from CSI amplitude, and accordingly calculates the target's velocity and location. However, due to lack of phase information, WiDar cannot independently calculate locations and suffers from accumulative error. DynamicMusic [16] estimates AoA of reflection from CSI by MUSIC algorithm. IndoTrack [17] further incorporates AoA with DFS for successive tracking. Differently, Widar2.0 employs a maximum likelihood algorithm for joint signal parameter estimations and enables single-link tracking, while all existing works require at least two links.

Wi-Fi based Gesture and Activity Recognition. Wi-Fi-based activity recognition attracts considerable research interests recently.

Many innovative applications have been designed, including activity and gesture recognition [20, 30, 34], respiration detection [31] and direction estimation [22, 36], etc. Most of these works employ learning techniques to model relations between human activity and CSI variations. CARM [34] uses power distribution of DFS components as learning features of HMM model. Given that CSI variations are determined by not only DFS but also ToF and AoA, Widar2.0 can fertilize these works, as uncorrelated parameters ToF and AoA can be first estimated and decoupled from learning features, leading to position and orientation agnostic activity recognition.

8 CONCLUSION

In this paper, we present Widar2.0, the first passive tracking system that only requires one single WiFi link and achieves sub-meter level tracking accuracy, without support of any additional infrastructure or sensors. We implement and evaluate Widar2.0 on COTS Wi-Fi devices. The results show that Widar2.0 achieves a median location accuracy of 0.75 m in a 6 m × 5 m area, comparable to state-of-the-arts approaches based on multiple links. Widar2.0 opens up passive tracking to new applications where few devices are available or accessible, e.g., homes and mobile environments. Future work extends to multiple target tracking and through-wall monitoring.

ACKNOWLEDGEMENT

REFERENCES

[1] 2018. Microsoft Kinect. https://developer.microsoft.com/en-us/windows/kinect. (2018).
[2] Fadel Adib, Zachary Kabelac, and Dina Katabi. 2015. Multi-person localization via rf body reflections. In *Procs. of USENIX NSDI*.
[3] Fadel Adib, Zach Kabelac, Dina Katabi, and Robert C Miller. 2014. 3d tracking via body radio reflections. In *Procs. of USENIX NSDI*.
[4] Fadel Adib and Dina Katabi. 2013. See through walls with wifi!. In *Procs. of ACM SIGCOMM*.
[5] Maurizio Bocca, Ossi Kaltiokallio, Neal Patwari, and Suresh Venkatasubramanian. 2013. Multiple Target Tracking with RF Sensor Networks. *IEEE TMC* 13, 8 (2013).
[6] Liqiong Chang and Xiaojiang Chen. 2016. FitLoc: Fine-grained and Low-cost Device-free Localization for Multiple Targets over Various Areas. In *Procs. of IEEE INFOCOM*.
[7] A. P. Dempster, N. M. Laird, and D. B. Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society (B)* 39, 1 (1977).
[8] Jeffrey A. Fessler and Alfred O. Hero. 1994. Space-Alternating Generalized Expectation-Maximization Algorithm. *IEEE Trans. on Signal Processing* 42 (1994).
[9] Jon Gjengset, Jie Xiong, Graeme McPhillips, and Kyle Jamieson. 2014. Phaser: enabling phased array signal processing on commodity WiFi access points. In *Procs. of ACM MobiCom*.
[10] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2011. Predictable 802.11 packet delivery from wireless channel measurements. *Procs. of ACM SIGCOMM* (2011).
[11] Kiran Joshi, Dinesh Bharadia, Manikanta Kotaru, and Sachin Katti. 2015. Wideo: Fine-grained device-free motion tracing using rf backscatter. In *Procs. of USENIX NSDI*.
[12] Kiran Joshi, Steven Hong, and Sachin Katti. 2013. Pinpoint: Localizing interfering radios. In *Procs. of USENIX NSDI*.
[13] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. 2015. Spotfi: Decimeter level localization using wifi. In *Procs. of ACM SIGCOMM*.
[14] Manikanta Kotaru and Sachin Katti. 2017. Position Tracking for Virtual Reality Using Commodity WiFi. *CoRR* abs/1703.03468 (2017).
[15] H. W. Kuhn and Bryn Yaw. 1955. The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly* (1955).
[16] Xiang Li, Shengjie Li, Daqing Zhang, Jie Xiong, Yasha Wang, and Hong Mei. 2016. Dynamic-music: accurate device-free indoor localization. In *Procs. of ACM UbiComp*.
[17] 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. *Procs. of ACM IMWUT* (2017).
[18] J. Löfberg. 2004. YALMIP : A Toolbox for Modeling and Optimization in MATLAB. In *In Procs of the CACSD Conference*.
[19] H. Vincent Poor. 1994. *An Introduction to Signal Detection and Estimation (2Nd Ed.)*. Springer-Verlag.
[20] Qifan Pu, Sidhant Gupta, Shyamnath Gollakota, and Shwetak Patel. 2013. Whole-home gesture recognition using wireless signals. In *Procs. of ACM MobiCom*.
[21] Kun Qian, Chenshu Wu, Zheng Yang, Yunhao Liu, and Kyle Jamieson. 2017. Widar: Decimeter-Level Passive Tracking via Velocity Monitoring with Commodity Wi-Fi. In *Procs. of ACM MobiHoc*.
[22] Kun Qian, Chenshu Wu, Zimu Zhou, Yue Zheng, Zheng Yang, and Yunhao Liu. 2017. Inferring Motion Direction Using Commodity Wi-Fi for Interactive Exergames. In *Procs. of ACM CHI*.
[23] Souvik Sen, Dongho Kim, Stephane Laroche, Kyu-Han Kim, and Jeongkeun Lee. 2015. Bringing cupid indoor positioning system to practice. In *Procs. of ACM WWW*.
[24] Souvik Sen, Jeongkeun Lee, Kyu-Han Kim, and Paul Congdon. 2013. Avoiding multipath to revive inbuilding WiFi localization. In *Procs. of ACM MobiSys*.
[25] Souvik Sen, Božidar Radunovic, Romit Roy Choudhury, and Tom Minka. 2012. You are facing the Mona Lisa: spot localization using PHY layer information. In *Procs. of ACM MobiSys*.
[26] Longfei Shangguan, Zheng Yang, Alex X Liu, Zimu Zhou, and Yunhao Liu. 2017. STPP: Spatial-temporal phase profiling-based method for relative RFID tag localization. *IEEE/ACM Transactions on Networking (ToN)* 25, 1 (2017), 596–609.
[27] Longfei Shangguan, Zimu Zhou, and Kyle Jamieson. 2017. Enabling Gesture-based Interactions with Objects. In *Procs. of ACM MobiSys*.
[28] M.C. Vanderveen, B.C. Ng, C.B. Papadias, and A. Paulraj. 1997. Joint Angle and Delay Estimation (JADE) for Signals in Multipath Environments. In *Procs. of Conference on Signal, Systems and Computers*.
[29] Deepak Vasisht, Swarun Kumar, and Dina Katabi. 2016. Decimeter-level Localization with a Single WiFi Access Point. In *Procs. of Usenix NSDI*.
[30] Aditya Virmani and Muhammad Shahzad. 2017. Position and Orientation Agnostic Gesture Recognition Using WiFi. In *Procs. of ACM MobiSys*.
[31] Hao Wang, Daqing Zhang, Junyi Ma, Yasha Wang, Yuxiang Wang, Dan Wu, Tao Gu, and Bing Xie. 2016. Human respiration detection with commodity wifi devices: do user location and body orientation matter?. In *Procs. of ACM UbiComp*.
[32] 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 *Procs. of ACM MobiCom*.
[33] Ju Wang, Jie Xiong, Hongbo Jiang, Xiaojiang Chen, and Dingyi Fang. 2016. D-watch: Embracing" bad" multipaths for device-free localization with COTS RFID devices. In *Procs. of ACM CoNEXT*.
[34] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2015. Understanding and modeling of wifi signal based human activity recognition. In *Procs. of ACM MobiCom*.
[35] Teng Wei, Anfu Zhou, and Xinyu Zhang. 2017. Facilitating Robust 60 GHz Network Deployment By Sensing Ambient Reflectors. In *Procs. of USENIX NSDI*.
[36] Dan Wu, Daqing Zhang, Chenren Xu, Yasha Wang, and Hao Wang. 2016. WiDir: walking direction estimation using wireless signals. In *Procs. of ACM UbiComp*.
[37] Yaxiong Xie, Zhenjiang Li, and Mo Li. 2015. Precise Power Delay Profiling with Commodity WiFi. In *Procs. of ACM MobiCom*.
[38] Yaxiong Xie, Jie Xiong, Mo Li, and Kyle Jamieson. 2016. xD-track: Leveraging Multi-dimensional Information for Passive Wi-fi Tracking. In *Procs. of ACM HotWireless*.
[39] Jie Xiong and Kyle Jamieson. 2013. ArrayTrack: a fine-grained indoor location system. In *Procs. of USENIX NSDI*.
[40] Jie Xiong, Karthikeyan Sundaresan, and Kyle Jamieson. 2015. ToneTrack: Leveraging frequency-agile radios for time-based indoor wireless localization. In *Procs. of ACM MobiCom*.
[41] Chenren Xu, Bernhard Firner, Yanyong Zhang, and Richard E. Howard. 2016. The Case for Efficient and Robust RF-Based Device-Free Localization. *IEEE TMC* 15, 9 (2016).
[42] Lei Yang, Qiongzheng Lin, Xiangyang Li, Tianci Liu, and Yunhao Liu. 2015. See through walls with cots rfid system!. In *Procs. of ACM MobiCom*.
[43] Zheng Yang, Zimu Zhou, and Yunhao Liu. 2013. From RSSI to CSI: Indoor localization via channel response. *ACM Computing Surveys (CSUR)* 46, 2 (2013), 25.
[44] Chi Zhang and Xinyu Zhang. 2017. Pulsar: Towards Ubiquitous Visible Light Localization. In *Procs. of ACM MobiCom*.
[45] Yanzi Zhu, Yuanshun Yao, Ben Y. Zhao, and Haitao Zheng. 2017. Object Recognition and Navigation Using a Single Networking Device. In *Procs. of ACM MobiSys*.