SiFi: Pushing the Limit of Time-Based WiFi Localization Using A Single Commodity Access Point

WEI GONG and JIANGCHUAN LIU, Simon Fraser University

There has been a booming interest in developing WiFi localization using multi-antenna (MIMO) access points (APs). Recent advances have demonstrated promising results that break the meter-accuracy barrier using commodity APs. Yet these state-of-the-art solutions require either multiple APs that are not necessarily available in practice, or multiple-channel measurements that disrupt normal data communication. In this paper, we present SiFi, a single AP-based indoor localization system that for the first time achieves sub-meter accuracy with a single channel only. The SiFi design is based on a key observation: with MIMO, the multiple (typically three) antennas of an AP are frequency-locked; although the accurate Time-of-Arrival (ToA) estimation on commodity APs is fundamentally limited by the imperfect time and frequency synchronization between the transmitter and receiver, there should be only one value for the ToA distortion that can cause three direct-path ToAs of the antennas to intersect at a single point, i.e., the position of the target. We develop the theoretical foundations of SiFi and demonstrate its realworld implementation with off-the-shelf WiFi cards. Our implementation introduces no hardware modification and is fully compatible with concurrent data transmission. It achieves a median accuracy of 0.93 m, which significantly outperforms the best known single AP single channel solution.

CCS Concepts: •**Networks** → **Location based services; Mobile networks;**

Additional Key Words and Phrases: WiFi OFDM, MIMO, Single Access Point, Indoor Localization

ACM Reference format:
Wei Gong and Jiangchuan Liu. 2017. SiFi: Pushing the Limit of Time-Based WiFi Localization Using A Single Commodity Access Point. *PACM Interact. Mob. Wearable Ubiquitous Technol.* 0, 0, Article 0 (2017), 21 pages.
DOI: 0000001.0000001

In the past few years, we have witnessed a growing interest in developing ubiquitous indoor localization systems using WiFi infrastructure, which have broken the meter accuracy barrier using off-the-shelf devices [21, 33, 37]. They however have yet to enable ready-to-use indoor navigation service as what GPS does for outdoors. Taking the characteristics of current WiFi infrastructure into consideration, an ideal WiFi-based localization system should meet the following requirements:

Universal: It should use standard WiFi interfaces on both the AP and target devices, without introducing specialized hardware modification or extra hardware (e.g., camera, accelerometer, gyroscope, and etc) that are not readily available on all sizes of WiFi-devices.

Compatible: As data communication is the essential task of WiFi, the localization sub-system should be compatible with data transmission, without blocking it.

This work was supported by a Canada Technology Demonstration Program (TDP) grant, a Canada NSERC Discovery Grant, and an NSERC E.W.R. Steacie Memorial Fellowship, and NSFC under Grant No. 61472268.

Author's addresses: W. Gong and J. Liu, School of Computing Science, Simon Fraser University.

DOI: 0000001.0000001

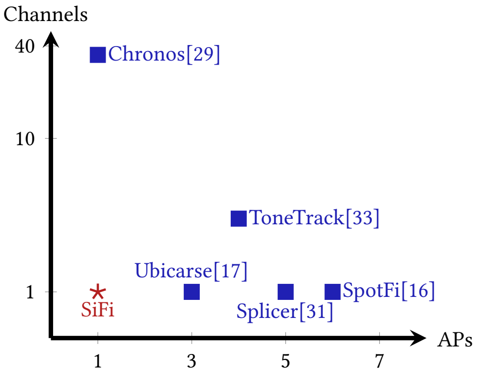


Fig. 1. We compare SiFi with the state-of-the-art WiFi-based localization systems that can achieve sub-meter accuracy. While all of them can deliver desirable accuracy, they require either multiple APs or multiple channel measurements. In contrast, SiFi requires only a single AP with a single channel. Note that here we only include systems that can achieve sub-meter median accuracy. Thus, some schemes are not involved. For example, CUPID achieves 2.7 m median accuracy using 3 APs.

Accurate: The pursuit of accuracy is boundless. Experiences have suggested that one meter or below is considered practically useful, which is also the design goal of state-of-the-art localization systems [21, 33, 37].

The above requirements seem simple but are hard to meet at the same time. The fingerprint-based schemes require densely deployed APs to deliver desirable accuracy [39, 41]. The single AP-based approaches for smartphones suffer from meter level accuracy and require extra inertial sensors [25, 29]. Recently a number of Angle-of-Arrival (AoA)-based techniques are proposed with high accuracy [14, 21, 36]. They however need at least three APs working together, which are not necessarily available in practice [29]. The recent Time-of-Arrival (ToA)-based advances on a single AP can deliver decimeter accuracy by creating a virtual wide band (or even ultra wide band) [33, 35], but their channel-hopping mechanisms disrupt data communication. We compare our scheme with state-of-the-art WiFi localization systems in terms of the number of channels and APs as in Figure 1. In short, indoor localization with a single commodity AP that enables concurrent data transmission (one single channel) remains a challenging goal.

With this goal in mind, we first carefully examine the physical layer design of WiFi to characterize the root causes of inaccuracy in channel measurement, i.e., Channel State Information (CSI). We find that the residual errors from the Symbol Time Offset (STO) [1] and Sampling Frequency Offset (SFO) are two major sources. While both the SFO and STO contribute delay distortions to all paths, they vary vastly in time scales. Specifically, SFOs may keep stable on the order of minutes, whereas STOs are different from packet to packet. As such, the ToA-based WiFi localization faces three critical challenges:

Dynamic time distortion: ToA estimates, even in channel coherence time, are highly dynamic due to STOs that are different across packets. This complicates super-resolution-based ToA estimations, as they need multiple measurements from the same distribution [31].

Direct path identification: Differentiating the direct path from reflection paths is vital for ToA-based localization since only the direct path captures the true ToA. Unfortunately, the prior insight for direct path identification becomes invalid due to fast-changing STOs, which contradicts the assumption that the variation of direct path is less than that of reflection paths across packets. Furthermore, inevitable spurious estimates arise due to the inaccurate estimate of path cardinality.

[1] It is sometimes called packet detection delay.

Tangled time delay: As both the STO and SFO add delay distortions to all paths, they are tightly coupled with the true ToA as will be shown in section 2.1. The prior solution for correcting tangled time delays rely on either pilot subcarriers in the data symbol [1] or additional measurements from other channels [35, 37], both of which become inapplicable if only the CSI on a single channel is available.

In this paper, we show that it is possible to address the above three challenges using only a single off-the-shelf AP on a single channel. In a nutshell, our method is based on an important observation: *there is only one value for the delay distortion that should cause all three direct-path ToAs of the receiving antennas to intersect at a single point where the signals physically come from.* Based on this insight, we develop a set of key techniques that deliver accurate localizations. First, we build a super-resolution algorithm based on Hankel matrix decomposition that can work with a single packet, avoiding the dynamic delay distortion. Then, leveraging the insight that STOs have a Gaussian distribution, we design a clustering scheme using ToA spreads across packets to remove STOs and identify the direct path at the same time. This clustering algorithm also evaluates the estimated direct path and outputs a likelihood score for it. By utilizing an advantage of the MIMO design that three antennas on an AP are frequency-locked, we make a key observation that although the delay distortion due to SFO is unknown, there is only one true value for the three direct-path ToAs to intersect. We accordingly model our localization process into a weighted iterative least square problem that estimates the unknown time delay distortion and location at the same time.

To demonstrate the feasibility of our design, we build SiFi, a single AP-based localization system for indoor environments, using Intel 5300 commodity WiFi cards. We evaluate it under the same settings and environments with two state-of-the-art single AP-based systems, Splicer [35] and CUPID [29]. Our experiments show that on a 40 MHz channel, SiFi achieves a median localization accuracy of 0.93 m, significantly outperforming Splicer (2.96 m) and CUPID (5.11 m). We also show that SiFi works robustly in challenging Non-Line of Sight (NLoS) scenarios, where Splicer and CUPID fail to provide desirable results.

Contributions: To our knowledge, SiFi is the first single AP-based localization system that achieves sub-meter accuracy using a single channel. SiFi does not require any hardware modification for both the AP and target devices, nor does it affect regular data transmission. Due to its simplicity, a range of indoor localization applications shall greatly benefit from SiFi.

1 RELATED WORK

The research of WiF-based indoor localization has a long history and thus numerous systems have emerged, we only survey methods that are closely related to ours here. For more complete surveys, please refer to [40]. Broadly speaking, there are three different types of approaches: ToA, AoA, and fingerprinting.

ToA: Early ToA-based methods use RSSI and the prorogation model to deduce the range between the transmitter and receiver [5, 9]. These methods are fundamentally limited by the RSSI that is an indirect measurement. CUPID [29] combines the range that is estimated by the energy of direct path, the angle that is based on AoA-MUSIC, and human movements together to realize single AP-based localizations. SAIL [25] further improves CUPID by coupling a built-in 88 MHz clock of an Atheros WiFi card and CSI to measure the ToA, leading to a median localization accuracy of 2.3 m. Other schemes that rely on maintaining highly accurate time synchronization between access points [26, 27] or between the transmitter and receiver [4], are quite hard to implement on commodity APs. Recently several advances use the physical layer information to accurately deduce the ToA/TDoA, breaking the meter accuracy barrier [33, 35, 37]. Although the localization accuracy has been greatly improved, they require information from other channels, resulting in data communication disruptions. However, SiFi is quite different from those methods, since it can deliver sub-meter accuracy with a single AP while keeping concurrent data transmission unaffected.

AoA: Due to the MIMO design of commodity APs, many AoA-based techniques are proposed using antenna-array [16]. ArrayTrack [36] pioneers MIMO-based WiFi localization by using WARP and USRP software radios. Later, Ubicarse [22] and LTEye [23] use synthetic aperture radar to improve accuracy. More recently, SpotFi [21] and Phaser [14] successfully apply phased array on commodity APs. Nevertheless, almost all AoA-based methods need at least three APs working together due to the triangulation requirement. While SiFi also makes use of the MIMO design, it only requires a single AP. Actually, SiFi can be an excellent complement to state-of-the-art AoA-based systems if more APs are available.

Fingerprinting: Fingerprinting based methods assume that every distinct location should have a distinct radio frequency fingerprint. While early fingerprinting based methods [41] always require the manual site survey, recent crowdsourcing-based schemes [39] receive lots of attentions. Although decent accuracy is not a problem for fingerprint based methods, e.g., a median accuracy of 0.6 m is achieved in [41], the major problem is its slow adaption to environment changes. For example, the replacement of AP or the movement of large indoor objects might require a sweeping new site survey/crownsourcing. Another problem is the performance would have degraded significantly when only a limited number of APs are available. Unlike those approaches, SiFi is a universal, easy-to-maintain, and cost-effective solution.

There are also a number of sensor-based localization schemes on mobile devices, e.g., acoustic sensor [8, 11], RFID [13, 15]. Yet, those schemes are not as ubiquitous as WiFi-based localization systems, which can deal with all sizes of devices, from desktop, laptop to tablet, cell phone, and even tiny-size tags [2].

2 SINGLE AP LOCALIZATION

2.1 CSI Primer

In wireless communication, multipath is the phenomenon that a signal reaches a receiving antenna by two or more paths. The mathematical model of multipath propagation can be presented using the channel impulse response function, i.e., $h(\tau) = \sum_{k=1}^{K} a_k \delta(\tau - \tau_k)$, where h is the impulse response of channel, τ is the time, K is the number of received impulses (paths), δ is the Dirac delta function, τ_k and a_k are the time delay and complex amplitude of k-th path. We call the estimates $(\hat{\tau}_1, \hat{\tau}_2, ..., \hat{\tau}_K)$ an estimated ToA spread and one of the estimates, e.g., $\hat{\tau}_i$, that is associated with the direct path, the true ToA. By Fourier transform, an impulse response can be equivalently converted to a channel frequency response, i.e.,

$$H(f) = \sum_{k=1}^{K} a_k \mathrm{e}^{-j2\pi\tau_k f}. \tag{1}$$

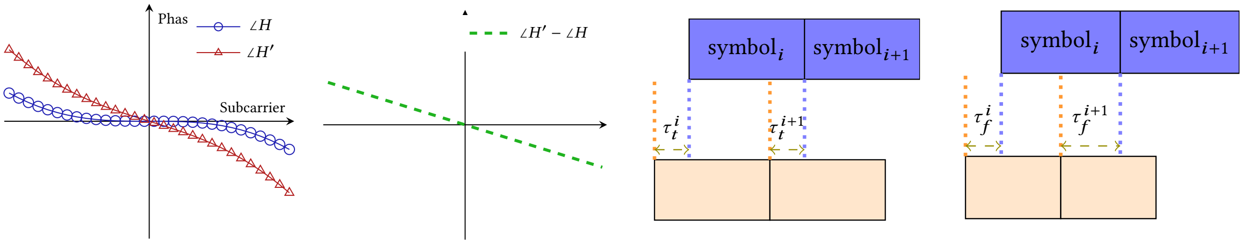
In a WiFi OFDM receiver, the frequency response is discretized by subcarriers, i.e., $H[f_i] = \sum_{k=1}^{K} a_k \mathrm{e}^{-j2\pi\tau_k f_i}$, where f_i is the frequency of i-th subcarrier.

The main difficulty of solving Equation 1 is in the nonlinear dependence of frequency responses on unknown delays. Fortunately, a bunch of well-researched super-resolution algorithms can be used to tackle this [32]. MUSIC [28] is probably the most prominent representative due to its robustness and effectiveness in many areas.

Unfortunately, the CSI in practice always contains errors, of which the STO and SFO are two major sources [2]. The STO stems from the residual error of symbol synchronization module in the receiver after the detection of frame start. According to the standard [1], the requirement of symbol synchronization is *one sample resolution*. To put this in context, as the subcarrier spacing is 312.5 KHz in WiFi, a useful symbol time is 3.2 µs (excluding long/short guard interval). For a 20 MHz channel, there are 64 samples (including null subcarriers) per symbol, then the time of one sample is 3200/64 = 50 ns, which is corresponding to a distance of about 15 meters. Differently, the SFO [3] comes from that the sampling frequency of DAC at the transmitter, f_t, and the sampling frequency of

[2] In this paper, we do not discuss other errors that are intractable by software methods, e.g., thermal noise, quantization error.

[3] It is also called sampling clock error sometimes.



(a) STO&SFO phase distortion **(b)** phase difference between the true and distorted channel **(c)** STO symbol distortion **(d)** SFO symbol distortion

Fig. 2. (a) The channel distorted by STO&SFO, H', has an added phase across subcarriers to the true channel H. (b) The phase differences between H and H' are linear with subcarriers. (c) The STO manifests in a constant offset to all symbols. (d) The SFO manifests in accumulated offsets across symbols.

ADC at the receiver, f_r, are not in sync. So, the fractional SFO of sampling frequency correction in the receiver is defined as $\zeta = \frac{f_t}{f_r} - 1$.

The common effect of those two errors in frequency domain is the phase rotation of CSI [4] across subcarriers as shown in Figure 2a. The magnitude of this added phase varies linearly across subcarriers as shown in Figure 2b. In time domain, these two errors result in delay distortions to all paths for a single packet, i.e.,

$$\tau'_k = \tau_k + \tau_f + \tau_t, k \in [1, K], \tag{2}$$

where τ_k is the true time delay of k-th path in theory, τ'_k is the tangled time delay of k-th path in practice, τ_f and τ_t are the delay distortions by the SFO and STO, respectively.

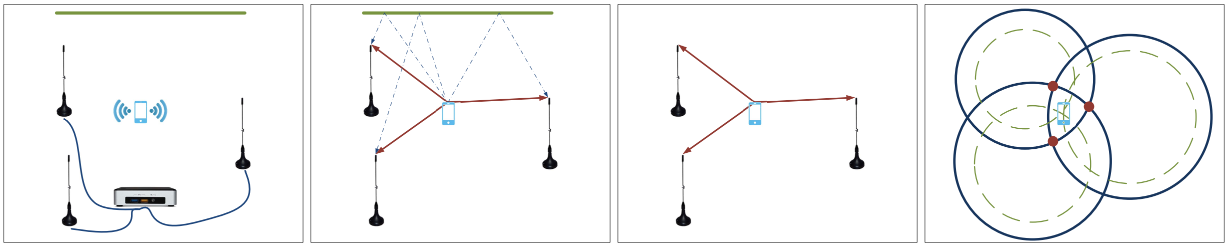
Nevertheless, the effects of STO and SFO on time-domain symbols are quite different. The STO leads to a constant symbol offset for all symbols as shown in Figure 2c, whereas the SFO results in accumulated symbol offsets across symbols as shown in Figure 2d. So an initial small fractional STO, ζ, (e.g., the first few symbols of a packet) can lead to a significantly larger offset in Long Training Fields that are directly used to estimate the channel. More importantly, the STO and SFO vary differently in time scales. While the STO varies from packet to packet, the SFO may keep stable on the order of minutes [20, 35].

2.2 Localization framework

We present the framework of SiFi and how it tackles challenges of the ToA estimation as follows.

The first step is to acquire the CSI from three antennas of an AP when the target is sending out packets, as shown in Figure 3a. Currently, WiFi cards of almost all the major WiFi manufacturers, e.g., Atheros, Intel, support exporting the CSI for each packet. Then in the second step, we apply a super-resolution algorithm to resolve time delays of all the paths as shown in Figure 3b. Our super-resolution algorithm is built upon several key techniques that are embedded in existing spectral analysis [32]. We will detail it in section 3. The third step is to identify the direct path of each antenna as shown in Figure 3c. By leveraging the observation that STOs follow a Gaussian distribution [30, 35], we design a clustering algorithm to identify the direct path and remove the STO at the same time, which are detailed in section 4. The final step is to estimate the location by untangling the true ToA

[4] Here, we do not discuss phase distortions that have little impact on the ToA estimation. For example, the carrier frequency offset manifests in a constant added phase for all subcarries [14, 35].



(a) CSI acquisition **(b)** multipath super-resolution **(c)** direct path identification **(d)** location estimation

Fig. 3. The architecture of SiFi. (a) It first collects CSI from the three antennas of the AP; (b) It uses a super-resolution algorithm to estimate the ToA spread; (c) It applies a clustering algorithm to identify the direct path and remove the STO at the same time; (d) It estimates the location by using a weighted iterative least square model to combat SFO.

from the delay distortion due to the SFO, as shown in Figure 3d. We model the unknown delay estimation into a weighted iterative least square problem, which is detailed in section 5.

3 ESTIMATING TOA SPREAD

Based on Equation 1, ideally the CSI of m equally spaced subcarriers $(f_{i_1}, f_{i_2}, ..., f_{i_m})$ in a WiFi channel with K paths can be written in vector form as

$$\begin{aligned} \mathbf{H} &= \mathrm{S}\alpha, \\ \mathrm{S} &\equiv [s(\tau_1), s(\tau_2), ..., s(\tau_K)], \\ s(\tau) &\equiv [\mathrm{e}^{-j2\pi f_{i_1}\tau}, \mathrm{e}^{-j2\pi f_{i_2}\tau}, ..., \mathrm{e}^{-j2\pi f_{i_m}\tau}]^T, \\ \alpha &\equiv [a_1, a_2, ..., a_K]^T, \\ \mathbf{H} &\equiv [H[f_{i1}], H[f_{i2}], ..., H[f_{im}]]^T \equiv [H_1, H_2, ..., H_m]^T, \end{aligned} \tag{3}$$

where the channel matrix $\mathbf{H}$ is of size $m \times 1$, the steering matrix S is of size $m \times K$, the amplitude matrix α is of size $K \times 1$, and $s(\tau)$ is the steering vector of size $m \times 1$. We use H_k to denote $H[f_{ik}]$ for brevity.

As each subcarrier is considered to be a narrowband flat-fading channel, the measured channel is usually modeled as $\mathbf{H_n} = \mathbf{H} + \mathbf{n}$, where $\mathbf{n}$ is the circular symmetric complex normal noise [5], of which the mean value is zero and the noise covariance matrix is known, and $\mathbf{H_n}$ is the measured CSI.

If multiple measurements can be collected from the same distribution, the traditional MUSIC algorithm is able to estimate the covariance matrix of $\mathbf{H}$ and then separates the signal subspace from the noise subspace, which is the key of super-resolution. However, the delay distortion brought by STO changes the parameter of the distribution, τ_i, across packets. So it calls for a single packet based super-resolution method. Fortunately, we observe that the number of available subcarriers is much larger than the number of paths. For example, with Intel 5300 WiFi cards, the CSI is of length 30 for an antenna in a packet, whereas the number of dominant paths

[5] Note the delay distortions due to the STO and SFO are coupled with the delay of each path, e.g., τ_i, which are different from the channel noise modeled here.

for indoor environments is around 5 [36]. Therefore, we can form a Hankel matrix as follows

$$\mathcal{H} = \text{Hankel}(\mathbf{H}) = \begin{pmatrix} H_1 & H_2 & \cdots & H_{m-l} \\ H_2 & H_3 & \cdots & H_{m-l+1} \\ \vdots & \vdots & \vdots & \vdots \\ H_l & H_{l+1} & \cdots & H_m \end{pmatrix}, \tag{4}$$

where l is an integer parameter that satisfies $l \geq K$ and $m - l \geq K$ [6]. Actually, the Hankel data matrix is widely used in many super-resolution algorithms [24] and can even be dated back to 1795 [10]. Then, we apply Singular Value Decomposition on the $\mathcal{H}$, i.e.,

$$\mathcal{H} = \mathbf{U}\Sigma\mathbf{V}^*, \tag{5}$$

where $\mathbf{U}$ is an $l \times l$ unitary matrix, Σ is an $l \times (m - l)$ diagonal matrix, and $\mathbf{V}^*$ is an $(m - l) \times (m - l)$ unitary matrix. * denotes conjugate transpose. Here we are interested in $\mathbf{U}$ and Σ. In particular, Σ is in form of $\text{diag}(\beta_1, \beta_2, ..., \beta_K, 0, ..., 0)$ with singular values $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_K > 0$. Based on these singular values, $\mathbf{U}$ can be separated into $\mathbf{U_s}$ of size $l \times K$ and $\mathbf{U_n}$ of size $l \times (l - K)$ that are corresponding to non-zero singular values and diagonal elements of zeros in Σ, respectively, i.e., $\mathbf{U_s}$ denotes the signal space and $\mathbf{U_n}$ denotes the noise space.

Note that the singular-value decomposition of $\text{Hankel}(\mathbf{H})$ is not possible in practice as the measured result is always $\mathbf{H_n}$. This is where MUSIC comes in. The core of MUSIC is the observation that the signal space should be orthogonal to the noise space. Therefore, by the singular-value decomposition of $\mathcal{H}_\mathbf{n} = \text{Hankel}(\mathbf{H_n})$, the noise space is obtained as $\mathcal{U}_\mathbf{n}$ by Equation 5 that should be orthogonal to the steering matrix, $\mathbf{S}^l$, that coincides with the signal space $\mathbf{U}^\mathrm{s}$, where $\mathbf{S}_l \equiv [s_l(\tau_1), s_l(\tau_2), ..., s_l(\tau_K)]$, $s_l(\tau) \equiv [\mathrm{e}^{-j2\pi f_{i_1}\tau}, \mathrm{e}^{-j2\pi f_{i_2}\tau}, ..., \mathrm{e}^{-j2\pi f_{i_l}\tau}]^T$. Hence, the ToA spread $(\tau_1, \tau_2, ..., \tau_K)$ can be identified as the peaks of the following orthogonal projection function

$$\mathcal{D}(\tau) = \frac{1}{\|\mathcal{U}_\mathbf{n}^* s_l(\tau)\|_2}, \tag{6}$$

where $\|\cdot\|_2$ denotes L^2 norm.

However, finding peaks of the above equation always requires either human interaction or a discretized search algorithm. To address this issue, we turn this peaks search into a model-based parameter estimation similar to [6], which directly results in numeric values to avoid discretization errors. The detail is included in Appendix A.

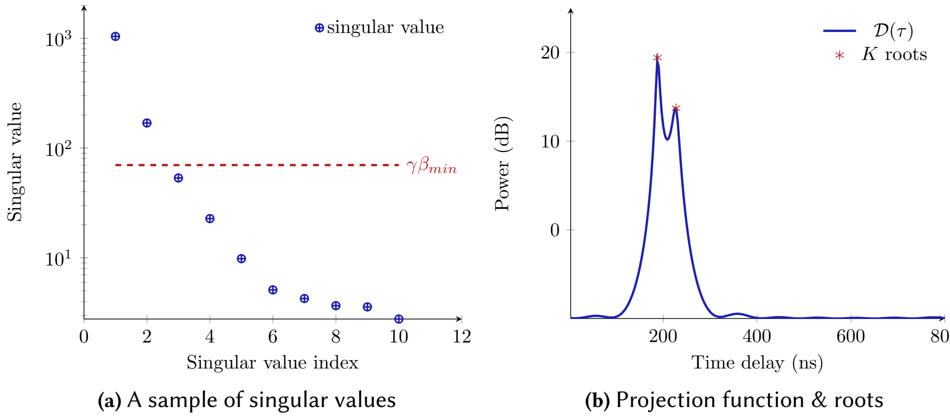
So far we have assumed that K is known. In practice, however, we need to estimate it, so we introduce a singular value-based threshold method. Remember, in noiseless case, Σ should include K positive singular values. But with the presence of noises, the zero entries in the diagonal of Σ might become positive, leading to more singular values than expected. Therefore, we set a threshold γ to estimate K. In particular, a singular value that is no less than $\gamma\beta_{min}$ is considered to be corresponding to the signal, where β_{min} is the minimal singular value. In our tests, we empirically set $\gamma = 8$ for moderate SNR scenarios and $\gamma = 25$ for high SNR scenarios. A sample of ToA spread estimation in practice is depicted in Figure 4.

4 FINDING THE DIRECT PATH

Now we need to pick up the direct path from the estimated ToA spread, $(\hat{\tau}_1, \hat{\tau}_2, ..., \hat{\tau}_K)$. We design a two-step procedure for this, identifying the paths from spurious peaks using a clustering algorithm and picking up the direct path by likelihood evaluation.

Identifying paths: We observe that although STOs change fast across packets, they follow a Gaussian distribution [30, 35]. So we can plot the ToF spread and corresponding amplitudes across packets and apply a fit clustering algorithm. The intuition is that as STOs add the same delay to all the paths, the delay of the same path across

[6] In [19], the suggested l is $m/2$ or $m/3$. We test both of them and find they provide quantitatively similar results in terms of accuracy at settings that are $m = 30$, $K = 5$.



(a) A sample of singular values

(b) Projection function & roots

Fig. 4. (a) A sample of singular values in practice when $l = m/3, m = 30, \gamma = 25$. It clearly shows that all singular values are non-zero due to noises. By using a threshold $\gamma\beta_{min}$, only two singular values that are above the threshold are selected, i.e., the estimated $\hat{K} = 2$; (b) The peaks of orthogonal projection function $\mathcal{D}(\tau)$ are correctly identified by the polynomial roots of Equation 19.

packets should also follow a Gaussian distribution. In particular, we apply the well-known Gaussian Mixture Model [7] clustering. For the most important input parameter, the number of clusters, η, we adopt a dynamic selection process. We vary η from 2 that is for simple LoS scenarios to 5 [7] that is for complex NLoS scenarios. The best η is chosen based on the intuition that real paths should be more tightly clustered than spurious peaks, i.e.,

$$\eta' = \underset{\eta \in [2,5]}{\operatorname{argmin}} \sum_{i=1}^{\eta} \mathbf{Var}[\text{cluster}_i]. \tag{7}$$

Note that during the iteration of clustering, the fitted covariance matrix may become ill-conditioned when η is much more than the ground truth or the data is highly correlated. The solution is to add a small positive number to the diagonal elements of the covariance matrix, resulting in a guaranteed positive-definite covariance matrix.

Based on clustered results, we further apply two filters to identify real paths. The first filter is the cardinality of a real-path cluster should be more than a percentage of the number of clustered packets. The justification of this filter is that the dominant path should exist in most of the packets. For instance, we can set this percentage as $\epsilon = 50\%$, which is very conservative. The second filter is that the variance of a real-path cluster should be less than ς. In practice, we set $\varsigma = 1.5 \cdot$ (one sample duration), as the standard [1] specifies the resolution of STO is one sample. While we recommend the threshold values for the above two filters in common indoor environments, they can be easily adjusted according to the needs of different scenarios.

Evaluating paths: After filtering the clustered results, we use the mean of each cluster as the ToA of each path ($\hat{\tau}_k$) and identify the direct path using likelihood evaluation. Our likelihood evaluation is inspired by the work [21]. In particular, we incorporate all the positive and negative factors of being a direct path into a likelihood function, i.e.,

$$\rho(k) = \mathrm{e}^{(w_\kappa \kappa_k - w_\sigma \sigma_k - w_\tau \bar{\tau}_k)}, \tag{8}$$

where $\rho(k)$ is the likelihood score of the k-th path, κ_k, σ_k, and $\bar{\tau}_k$ are the cardinality, variance, and mean delay of the cluster that is corresponding to the k-th path, respectively. The weights $w_\kappa, w_\sigma, w_\tau$ are constants

[7] Note that the maximum value for η is set at 5 because there are usually 5 dominant paths in indoor environments [21, 36].

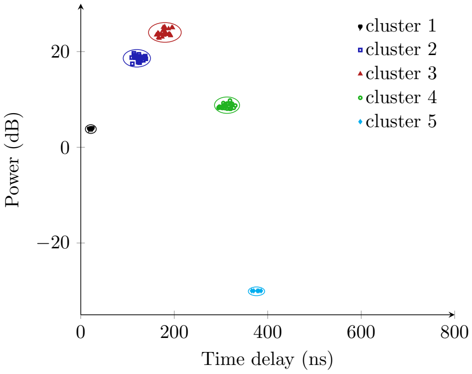


Fig. 5. A sample of clustered paths based on 30 packets in an NLoS scenario.

across clusters but different in scales. w_κ is on the order of 10 points in the cluster. w_σ, w_τ are on the order of nanoseconds. The intuition of the above likelihood function is the more points (larger κ_k) in the cluster, the higher possibility of being a direct path, while more dispersive (larger σ_k) and larger delay values (larger $\bar{\tau}_k$), the lower possibility of being a direct path as the direct path tends to be stable and always travels in the shortest time.

Now we can pick up the direct path that is with the highest likelihood score. A sample of clustering process in given in Figure 5. Five candidate paths (clusters) are clear to spot. First, the first and fifth clusters are removed by filtering due to their low cardinalities. Then for likelihood evaluation, as the variance and time delay of the second path are smaller than the third path, so $\rho(2) > \rho(3)$. For the fourth path, its likelihood score is also less than the second path due to the less number of points in the cluster and longer time delay. Therefore, the second path is identified as the direct path. Note that the salient feature of our direct path identification is to remove dynamic time distortions brought by STOs and identify the direct path at the same time. Also the associated likelihood score ρ is quite useful in localizations as we shall see in next section. We denote the output of this section as τ_D and ρ_D for a transmitter-receiver pair.

Although our clustering and likelihood evaluation are inspired by SpotFi [21], our solution is quite different from it. Specifically, our method differentiates itself in objectives and techniques. For example, our clustering is to remove the fast-changing STOs based on the observation that STOs follow a Gaussian distribution, whereas SpotFi uses a linear regression to remove STOs and then apply a clustering algorithm to help identify the correct AoA. Moreover, SpotFi uses a fixed number of clusters for clustering, which is 5 in [21]; in contrast, our clustering is more flexible and adaptable since we use a range of different number of clusters, $\eta \in [2, 5]$, to account for dynamic indoor scenarios, e.g., LoS and NLoS environments.

5 COMBATING SFO

Even the effect of STO could be alleviated by the above clustering process, the delay distortion incurred by SFO still exists in all direct paths of transmitter-receiver pairs according to Equation 2. There are some existing methods to correct this error. For example, in the WiFi standard [1], pilot subcarriers in data symbols are employed to correct this residual phase offset. Splicer [35] and ToneTrack [37] make use of CSI from other channels. Unfortunately, they are not applicable in our case since our goal is to do localization with the CSI of only one channel. Nevertheless, we observe that an opportunity arises in the MIMO design of commodity APs that all antennas on board are frequency-locked, which means the STO should be the same across all the transmitter-receiver pairs, i.e., $\tau_f^1 = \tau_f^2 = \tau_f^3$, where τ_f^i is the delay distortion on the i-th receiving antenna due to STO. Since these delay distortions are the same, we use τ_F to denote this. At the same time, according to the

Algorithm 1 The localization algorithm of SiFi

```
Input: collected CSI for each receiving antenna on the AP, the locations of three antennas.
Output: location of the target.
for each receiving antenna do
    for each packet do
        Construct the data matrix as in Equation 4;
        Use singular-value decomposition to obtain eigenvectors, $\mathcal{U}_{\mathrm{n}}$;
        Estimate $K$ as in Figure 4a;
        Obtain $K$ roots from Equation 19;
    end for
    Filter and cluster ToA spreads across packets;
    Identify the direct path using equation 8;
end for
Iteratively minimize Equation 13 until results converge.
```

observation that SFOs keep stable on the order of minutes [20, 35], we can treat τ_F as a constant across packets in a short-time interval. Based on the above observations, we can model the measured distance at the i-th antenna as follows,

$$\mathcal{R}_n^i = \mathcal{R}^i(x, y, \tau_F) + n_R^i, \tag{9}$$

$$\mathcal{R}^i(x, y, \tau_F) = \sqrt{(x^i - x)^2 + (y^i - y)^2} + c\tau_F, \tag{10}$$

where c is the speed of light, $\mathcal{R}_n^i$ and $\mathcal{R}^i$ are the measured distances with and without measurement noises, n_R^i, (x^i, y^i) is the position of i-th receiving antenna, and (x, y) is the unknown position of target. Note that $\mathcal{R}_n^i$ is computed using the estimated delay of the direct path from the last section, i.e., $\mathcal{R}_n^i = c\tau_D^i$.

As Equation 10 is non-linear, we apply Taylor series to linearize it. Specifically, we can expand it at the point (x_0, y_0, τ_{F0}) and omit the second and higher order terms, i.e.,

$$\mathcal{R}^i(x, y, \tau_F) = \mathcal{R}^i(x_0, y_0, \tau_{F0}) + \frac{\partial R^i}{\partial x}\Delta x + \frac{\partial R^i}{\partial y}\Delta y + \frac{\partial R^i}{\partial \tau_F}\Delta \tau_F, \tag{11}$$

, where $\Delta x \equiv x - x_0$, $\Delta y \equiv y - y_0$, $\Delta \tau_F \equiv \tau_F - \tau_{F0}$.

Therefore, by putting Equation 11 into Equation 9, we rewrite data from three antennas into a vector form as

$$\begin{pmatrix} \Delta R^1 \\ \Delta R^2 \\ \Delta R^3 \end{pmatrix} = \begin{pmatrix} \Delta\frac{\partial R^1}{\partial x} & \frac{\partial R^1}{\partial y} & \frac{\partial R^1}{\partial \tau_F} \\ \Delta\frac{\partial R^2}{\partial x} & \frac{\partial R^2}{\partial y} & \frac{\partial R^2}{\partial \tau_F} \\ \Delta\frac{\partial R^3}{\partial x} & \frac{\partial R^3}{\partial y} & \frac{\partial R^3}{\partial \tau_F} \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta \tau_F \end{pmatrix} + \begin{pmatrix} n_R^1 \\ n_R^2 \\ n_R^3 \end{pmatrix}. \tag{12}$$

Now it is clear that the above equation is in the standard form of least square problem. However, since the variances of n_R^1, n_R^2, n_R^3 are not equal, we need to give proper weights to each direct path and this is where likelihood scores come in. In other words, we need to estimate the location in a way that well explains the quality of each direct path. Mathematically, we need to find the location that minimizes the following objective function,

$$\sum_{i=1}^{3} \rho^i (\Delta R^i - \frac{\partial R^i}{\partial x}\Delta x + \frac{\partial R^i}{\partial y}\Delta y + \frac{\partial R^i}{\partial \tau_F}\Delta \tau_F)^2, \tag{13}$$

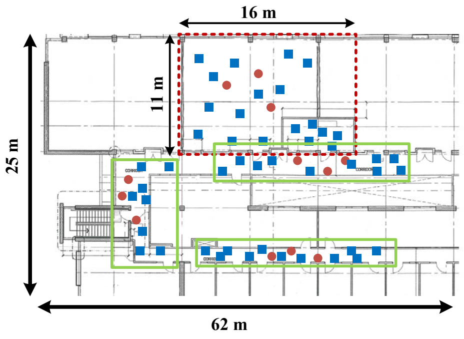


Fig. 6. Our testbed deployment includes target locations in blue squares and antenna positions in red circles. The red dashed box stands for a typical indoor office environment, while the other three green solid boxes denote corridor scenarios. In each test, only a single AP with three antennas is available.

where ρ^i is the likelihood score of a direct path from i-th antenna. Actually, minimizing the above object function is exactly the weighted version of Equation 12.

By solving the above equation, we can obtain a one-time estimate, $[\Delta x, \Delta y, \Delta \tau_F]^T$. Based on this, we replace the initial guess (x_0, y_0, τ_{F0}) with a new point $(x_0 + \Delta x, y_0 + \Delta y, \tau_{F0} + \Delta \tau_F)$ in equation 11 to start another around of weighted least square estimation until the solution converges below a threshold, e.g., the $(\Delta x, \Delta y)$ is at 1-meter level. For cases that it may not be able to reach the convergence threshold, we also set a maximum iteration times to prevent infinite loops. For the initial guess point, both a random point or a fixed point would work. We wrap up all the above localization algorithms briefly in Algorithm 1.

6 IMPLEMENTATION

We implement SiFi using off-the-shelf Intel 5300 WiFi cards, which can export CSI using Linux CSI Tool [17]. For each transmitter-receiver pair, the CSI tool outputs truncated CSI of length 30 (for both 20/40 MHz channels) for each packet. Each element of CSI is a complex number, of which the real and imaginary parts are quantized using 8 bits. We use a dell desktop that is with an Intel 5300 WiFi card installed using a Mini PCI-E to PCI-E adapter as the AP. Each of three antennas on the AP is connected to a 5-meter long antenna extension cable with a magnetic base [8]. For the target device, we use an Intel Mini PC, NUC D34010WYB, that is equipped with another Intel 5300 WiFi card. To support the mobility of this Intel NUC, we connect it to a portable charger, RAVPower Xtreme series power bank. All the experiments are done using the 802.11n protocol [1]. Our evaluations are mainly concerned with the localization error, which is the Euclidean distance between the estimated position and the real position.

Deployments: We deploy SiFi in different scenarios, including common offices, corridors, and high NLoS scenarios. We tested over 200 different locations, a part of which is shown in Figure 6. Note that although there are several deployment positions for the AP, only a single AP is available for each test. The only requirement for positions of antennas is to keep them non-collinear, which is basic for all ToA based solutions. We obtain the ground truth of locations by combing the architectural floor plan and a Bosch GLM35 laser distance finder that can achieve mm accuracy.

[8] For typical extension cords, the cable loss is 0.1 dB/ft (about 0.3 dB/meter). The common sizes of those cords are 3-meter, 5-meter, and 7-meter. As the maximal difference of such a signal loss is about 1.2 dB between 3-meter and 7-meter cords, we observe no significant difference in localization performances with cords of those sizes. In this paper, we report the results based on 5-meter extension cords.

CSI acquisition: First, the AP works in the monitor mode on a pre-selected channel. At each location, the target sends out 200 packets with 10 ms interval. Then, the AP shall collect CSI for each packet on all transmitter-receiver pairs. Later, it uses Algorithm 1 to estimate the target's location. Our algorithms are implemented using MATLAB. All competitions are fed by the same raw CSI data.

Note that we do not need the synchronization of CSI, which is usually implicitly or explicitly required in many other systems, e.g., timestamps in [21], and the wireless synchronization protocol in [37]. Because the synchronization across antennas on an AP is done by a frequency-locked loop on-chip. Due to the instability of firmware [14] on 2.4 GHz, both the AP and target operate on a 40 MHz channel of 5 GHz spectrum, unless otherwise stated. For example, one of the channels we tested is channel 100+, of which the frequency range is 5490-5530 MHz. Actually, the operation on 5 GHz frequencies cannot perform better than that on 2.4 GHz due to the worse penetration for indoor applications.

Competitions: We compare SiFi with two state-of-the-art single AP based methods, Splicer [35] and CUPID [29]. We choose these two methods because, among recently proposed ToA based designs, they can directly operate on a single commodity AP on a single channel without any hardware or driver modification. We did not compare SiFi with Chronus [33], as it requires a modified drive that can support fast channel hopping, which is not compatible with WiFi standards. While SAIL [25] is an improved version of CUPID, it relies on the timing reading from an internal clock on the chip, which is not universal for all commodity WiFi cards. Hence, we did not include it for comparison.

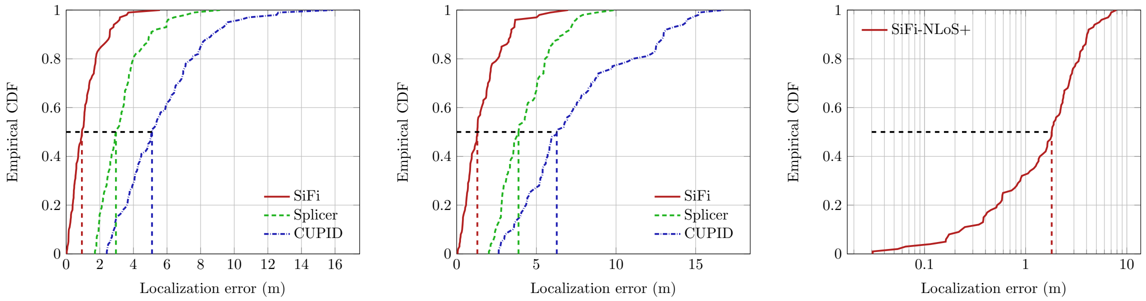
We try our best to faithfully implement CUPID as stated in the paper [29] and achieve comparable results. For a fair comparison, instead of the estimated distance of human movement using inertial sensor data, which is required in CUPID, we feed it the ground truth distance of such movement. For all the rest of settings that are used in our implementation of CUPID [29] and Splicer [35], we follow the original setup specified in [29, 35], such as 2.6 cm for the antenna spacing, which is the half of the wavelength of carrier frequency at 5G Hz. Such a setup is mainly for maximizing the AoA range to [0°, 180°]. For the rationale of those settings, please refer to [29, 35] for more details. While Splicer in the original paper [35] does have the ability to collect CSI from multiple channels, we only adopt its basic version that only has the access to CSI on a single channel. For other settings, Splicer is the same as CUPID stated above.

7 EXPERIMENTAL EVALUATION

7.1 Comparisons

7.1.1 Indoor office. We first examine the performance of SiFi and its competitions in indoor offices. The indoor office is one of the most representative scenarios for indoor localizations as it is quite multipath rich due to walls, tables, metallic objects, etc. For indoor offices, we test locations that are both in LoS and NLoS, such as, table corners, drawers, and places obstructed by humans. One of the tested offices is highlighted by the red dashed box in Figure 6.

We plot results in Figure 7a, which shows that SiFi achieves a median localization accuracy of 0.93 m compared to 2.96 m for Splicer and 5.11 m for CUPID. The 90th percentile tail errors are 2.59 m, 5.06 m, and 8.69 m for SiFi, Splicer, and CUPID, respectively. To put these numbers in context, the best AoA-based algorithm, SpotFi, achieves 0.4 m median accuracy using 6 APs with a 40 MHz channel, and the most advanced ToA-based scheme, Chronus, delivers 0.58 m median accuracy using a single AP but with 35 channels, which means 700 MHz bandwidth has been used. So we believe SiFi indeed has pushed the current limit for WiFi-based localization systems in the way that only a single AP with a single channel is involved. Another thing worth noting is that SiFi can locate stationary targets, which are quite useful for indoor applications, e.g., search for missing objects, whereas Splicer and CUPID cannot.



(a) Indoor offices

(b) Corridors

(c) Challenging NLoS scenarios

Fig. 7. (a) CDFs of the localization error of SiFi and other two competitions for indoor offices; (b) CDFs of the localization error of SiFi and other two competitions for corridors; (c) CDF of the localization error of SiFi in challenging NLoS scenarios, where the stable direct path is available for at most 2 antennas.

7.1.2 Corridor. Next, we conduct tests in corridors that are quite common for almost all indoor buildings. There are two major difficulties in locating objects in corridors. First, the number of APs that can be seen by the client is usually not too much, so robust single AP based solutions are always desired for this situation. Second, the interval of ToA spread becomes much smaller (e.g, 3-5 ns), worsening multipath channel distortions, especially for narrow corridors. For example, the width of the narrowest corridor in our test is only 1.2 m, which is way smaller than that of an office room.

The results for corridors are plotted in Figure 7b. We observe that in corridors, the median accuracy of SiFi is 1.29 m, whereas Splicer's accuracy degrades to 3.88 m and CUPID's accuracy worsens to 6.28 m. Note that these results of Splicer and CUPID are assisted by the ground truth distance of human movement we feed, because usually the deck-reckoning becomes more challenging for distance estimation in corridors [25]. The better performance of SiFi is due to two aspects. First, the super-resolution algorithm of SiFi is able to resolve multipath more accurately, compared to inverse Fourier transform used by CPUID and Splicer. Second, the direct path identification scheme of SiFi can efficiently differentiate the direct path from refection paths. In contrast, Splicer and CUPID rely on the human movement which is not that stable.

7.1.3 Challenging NLoS scenario. Furthermore, we evaluate SiFi under more stressful NLoS scenarios. Specifically, this test is conducted in locations where the stable direct path is only available for at most 2 antennas. These locations usually are severely interfered, e.g., 2-3 thick walls, metal poles, quite narrow corners, due to undesirable diffractions, refractions, and even absorptions. We denote such scenario as NLoS+. Much prior time-based work does not study such challenging scenarios [33, 35, 37].

We test SiFi and two competitions in NLoS+ locations and report results in Figure 7c. However, since antennas of Splicer and CUPID are quite close in space (2.6 cm separated), three antennas always experience the same serious interference, making them unable to provide meaningful results. So their results are not included. Seen from Figure 7c, SiFi experiences a degradation in accuracy as expected, but it still achieves a median accuracy of 1.81 m. The main reason for this is that SiFi uses a likelihood score to assess the quality of direct path, which rewards the stable path and gives less weight to the low quality path accordingly.

7.2 Investigating SiFi in detail

7.2.1 Traditional MUSIC vs. single packet MUSIC. To further investigate the effects of essential parts of SiFi, we first examine the importance of single packet based MUSIC. The result of traditional MUSIC [28] is shown

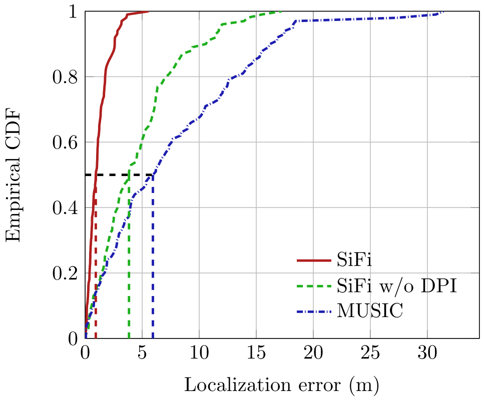


Fig. 8. We investigate the essential modules of SiFi by replacing them with counterparts. First, our super resolution module is substituted by the traditional MUSIC algorithm using measurements across packets. Second, our direct path identification (DPI) module is replaced by a simple strategy, which is to pick up the path that is with the smallest ToA.

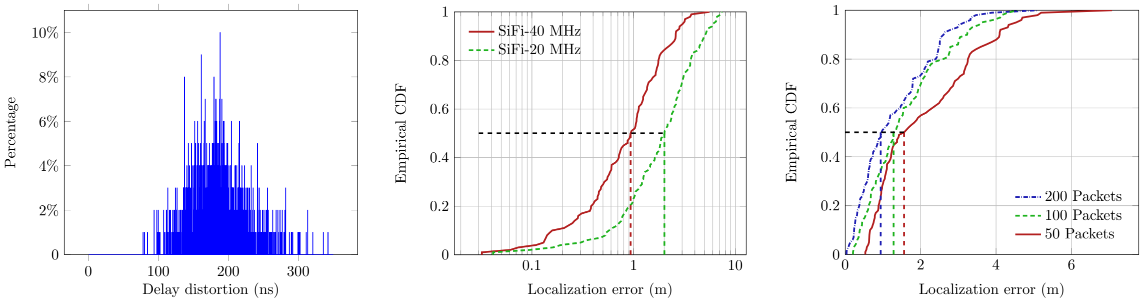
in Figure 8. Unsurprisingly, the performance of traditional MUSIC is quite poor. Specifically, it only achieves a median accuracy of 5.95 m and the 90-th percentile error is 16.2 m. The main reason for this is that the fast changing STOs make one of the prerequisites for MUSIC invalid, which assumes multiple measurements are sampled from the same distribution. In contrast, our customized MUSIC in SiFi is not affected due to its single packet based design.

7.2.2 Impact of direct path identification. Next, we examine the impact of direct path identification by replacing it with a simple strategy, which is to pick up the path that is with the smallest ToA as the direct path. Actually, this intuitive strategy is used in several other systems, such as LTEye [23].

Its result is shown in Figure 8. We observe that the SiFi without direct path identification suffers from poor accuracy. In particular, the median accuracy worsens from 0.93 m to 3.84 m. This is because the smallest ToA may contain spurious peaks produced by inaccuracy estimates of number of paths and highly dynamic STOs. Fortunately, SiFi takes good care of this by using filtering and dynamic clustering techniques together to remove spurious ToAs.

7.2.3 Delay distortions. Moreover, we examine delay distortions brought by STOs and SFOs, which are computed as the delay of the estimated direct path by our single packet MUSIC subtracts the accurate direct path ToA that is derived from the ground truth distance. To remove the effect of NLoS, we conduct this experiment in LoS locations. Hence, we use a very stable solution for direct path identification in LoS, which is to pick up the path that is with the smallest ToA and the highest power at the same time. Figure 9a plots the histogram of delay distortions we measure at 50 locations for 20 times with 5-minute interval. We have two observations from this figure. First, the delay distortion is much larger than the actual ToA. In particular, the mean of delay distortions is 186 ns, whereas common direct paths of 5-20 m take ToA from 16.7 to 66.7 ns. Second, the delay distortion is highly dynamic, of which the standard deviation is 40.2 ns. These two factors make the correction for inaccurate channel measurement necessary.

7.2.4 Impact of channel bandwidth. We also investigate how SiFi performs under a 20 MHz channel, which is also commonly used in nowadays' WiFi. Figure 9b plots the localization error of SiFi using a 20 MHz channel. Like all ToA based methods, the performance of SiFi degrades on a narrower channel. Yet, it still delivers a median localization accuracy of 2 m. Actually, we find that when the resolution of time becomes worse on a



(a) Distribution of delay distortions **(b)** Different channel bandwidth **(c)** Different number of packets

Fig. 9. (a) The distribution of delay distortions due to STOs and SFOs in 1000 measurements. (b) CDFs of the localization error of SiFi under different channel bandwidths. (c) CDFs of the localization error of SiFi under different number of packets.

narrower band, the iterative least square contributes the most to making the final estimate stable. Note that due to the simple single channel design of SiFi, it can be seamlessly adapted to an AP that is compatible with 802.11ac standard, which promises to deliver higher accuracy as the channel bandwidth increases to 80/160 MHz. We leave SiFi with the 802.11ac standard as one of future work.

7.2.5 Impact of number of packets. As stated in Section 4, multiple packets are quite helpful in clustering, as they can help differentiate the direct path and remove the instability incurred by STO. In the mean time, we also want to minimize the localization delay by using a small number of packets. So we investigate the impact of number of packets on SiFi by varying it from 50 to 100. Figure 9c plots the results, which demonstrate SiFi achieves a median localization accuracy of 1.56 m using 50 packets compared to 1.28 m using 100 packets. Even more, SiFi achieves 0.94 m median accuracy when 200 packets are available. There are two observations from these results. First, SiFi can adapt its accuracy to different number of packets. Second, the more accurate results by SiFi are achieved at the cost of delay compared to other schemes, e.g., Splicer, CUPID, that only involve a limited number of packets (e.g., 10-30).

7.2.6 Investigation in a larger area. To further study the performance of SiFi in a larger area, we conduct experiments in the West Mall Center of Simon Fraser University as shown in Figure 10. The size of this testbed is 4620 m^2. In this testbed, there are about 4 shops, 10 classrooms, and 23 tables in the lobby. During the day, all the areas are packed with students coming and going. During the night, shops are closed and fewer students are in classrooms and with tables, 2 persons/table and 7 persons/classroom on average. Same as the experiments done in Figure 6, multiple APs are deployed as a single AP cannot cover the whole area. But we ensure that each test client is covered only by a single AP. We have done two groups of experiments. In the first group, we intend to investigate the impact of a larger area. So we conduct tests both in a small area, which is 176 m^2 shown as a red box in Figure 6, and in a large area, which is 4620 m^2 shown in Figure 10. All experiments done in this group are controlled, which means no moving participants passing by. Results are shown in Figure 11a. It is as expected that the accuracies of both scenarios are quite similar. This is because a much bigger area does not bring much difficulty to the problem and each client is still covered a single AP. For the second group, we examine the performance of SiFi in the large area for uncontrolled (real) scenarios, where there are a number of irrelevant people moving around. We group the results into different time periods: morning, afternoon, and evening, which are shown in Figure 11b. We observe that those irrelevant participants indeed impact the localization accuracy because they may occlude objects, bring more dynamic reflections, and even create interference by using their

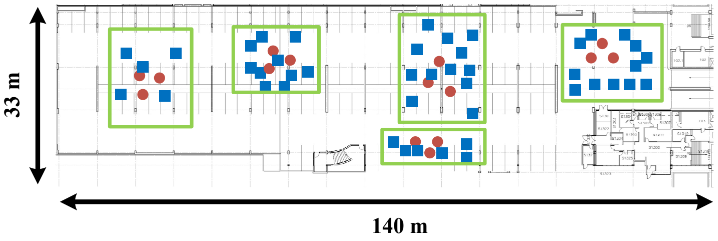
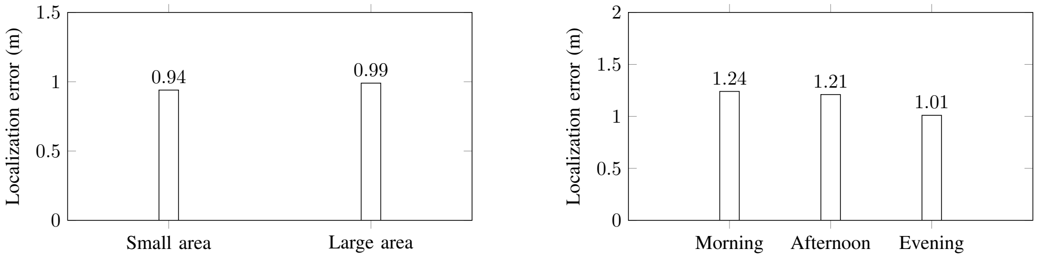


Fig. 10. Larger test area of 4620 m^2 in the West Mall Centre of Simon Fraser University. There are 20 positions for APs and 500 positions for client devices in the whole area. For brevity, part of locations are shown in the map where client positions are denoted as blue squares and antenna positions are denoted as red circles.



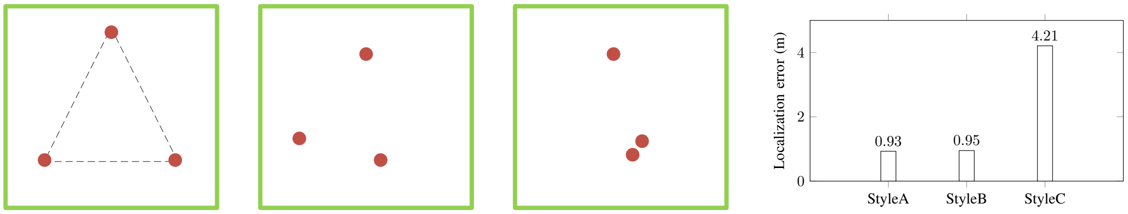
(a) Comparison of small and large areas for controlled experiments

(b) Performance of the large area for uncontrolled experiments

Fig. 11. Investigation of performance of SiFi in a large area of 4620 m^2 for controlled experiments (a) and uncontrolled experiments (b).

WiFi devices. Specifically, during morning and afternoon time, the median accuracies deteriorate to 1.24 m and 1.21 m, respectively. The result at evenings is better those at daytime, which is 1.01 m and is quite close to 0.99 m, the performance of controlled scenarios. This is because there are usually just a handful of people self studying out there during the night. Therefore, irrelevant people, whose locations are not of our interest, can still pose challenges for localization accuracy. Mechanisms that can remove the unwanted reflections from moving objects, like frequency modulated carrier wave [3], are worth future investigation.

7.2.7 Impact of antenna positions. Next, we investigate the impact of antenna positions. Apart from the general rule that three antennas cannot be collinear, we empirically find that if two antennas are too close, or clustered, the localization performance drops. Based on this, we perform experiments in three different groups. The first group, StyleA, includes equilateral triangles where the length of the leg is more than 3 m. Intuitively, such triangles cover the area evenly and are used as benchmarks. The second group, StyleB, is with random triangles but each leg is more than 2 m while random triangles in the third group (StyleC) have a leg less than 2 m. Three examples of antenna positions for each group are shown in Figure 12a, 12b, and 12c, respectively. For each group, we have 10 different schemes for positioning. Averaged results are depicted in Figure 12d. No surprisingly, the median localization accuracy of group StyleA is the best, which is 0.93 m and thus used as a benchmark. The performance of group StyleB, 0.95 m, almost catches up with that of StyleA. Group StyleC, however, suffers a lot from too close antennas and achieves only 4.21 m median accuracy. Our observations confirm that too close anchors can degrade the network localization performance because the intersection point of two close anchors is



(a) StyleA, equilateral triangles with legs more than 3 m

(b) StyleB, random triangles with each leg more than 2 m

(c) StyleC, random triangles with a leg less than 2 m

(d) Comparison of different antenna positions

Fig. 12. Impact of antenna positions. Experiments are done in three different styles of antenna positions. Three examples are shown in (a), (b), and (c). The corresponding localization performances are compared in (d).

Table 1. Comparison of state-of-the-art systems (sub-meter accuracy) with ours.

	Comm. compatibility	Median accuracy (m)	Commercial device support	# of APs	# of antennas per AP	# of channels	Testbed area (m^2)	Approach
ArrayTrack [36]	✓	0.23	×	≥3	16	1	N/A	AoA
Phaser [14]	✓	0.8	✓	≥3	5	1	190	AoA
SpotFi [21]	✓	0.4	✓	≥3	3	1	160	AoA
Ubicarse [22]	✓	0.39	✓	≥3	2	1	150	AoA
ToneTrack [37]	✓	0.9	×	≥3	1	3	500	ToA
Splicer [35]	×	0.95	✓	≥1	3	10	N/A	ToA&AoA
Chronus [33]	✓	0.65	×	≥1	3	35	400	ToA
SiFi	✓	0.93	✓	≥1	3	1	4620	ToA

easily influenced by the distance error, thus making the final 3-intersection highly unstable [18]. Therefore, in practice we should adopt StyleA or StyleB for better performance. Our current settings of 3 m for StyleA and 2 m for StyleB are empirically learned. Finding the optimal positions for antennas is a very important problem and we intend to investigate it using rigid graph [12] and network localizability theories [38] together in the future.

8 DISCUSSION & CONCLUSION

To better understand our system, we first compare it with other state-of-the-art systems in Table 1. Note that we only include systems that achieve sub-meter median accuracy here. For more comprehensive comparisons, please refer to surveys [34, 40]. The comparisons are done from eight aspects, namely, communication compatibility, median accuracy, commercial off-the-shelf device support, No. of APs, No. of antennas for each AP, No. of channels, testbed area size, and approach. Obviously, our system is not the most accurate one among those. Yet, it achieves a good trade-off between accuracy and resources for scenarios where a single AP or a very limited number of APs are available. Specifically, Arraytrack [36] achieves the best median accuracy, which is 0.23 m. It, however, does so at the cost of hardware resources. It requires a software-defined radio platform and 16 antennas for each AP. To put this in context, a typical commercial AP usually comes with 3 antennas, which is the standard setup for most other systems, like SpotFi, Chronus, Splicer, and ours. Phaser [14] tries to implement Arraytrack's idea with commercial devices by synchronizing two wireless cards as an AP. It achieves around 0.8 m median accuracy but requires hardware modification, like extra signal splitters, and additional alignment algorithms. Furthermore, it requires at least 3 APs to do localization because it adopts the AoA approach. Such a requirement also applies to the other two excellent AoA systems, SpotFi [21] and Ubicarse [22]. Therefore, those AoA systems would face difficulties when there are not enough APs. For example, it is common that there is only a single AP for small businesses, offices, and homes. Also, sometimes only a limited number of AP are seeable even in a larger space due to interferences, occlusions, and far distances. ToA based solutions promise to use less APs for localization. In particular, Splicer [35], a most recent improved version of CUPID, combines AoA and ToA approaches and achieves sub-meter median accuracy using 10 channels. Chronus [33] goes even further and achieves 0.65 m median accuracy using all the 35 WiFi channels and a single AP. They, however, require customized drivers to do channel hopping, which inevitably disrupts the ongoing communication. In contrast, our scheme, SiFi, leverages novel clustering based direct-path finding and builds a weighted least square method to estimate SFO, achieving sub-meter median accuracy. The primary reason of our system's better performance over previous single-AP solutions, like Splicer and CUPID, is that our new direct-path finding method and the corresponding weighted iterative algorithm, which helps us estimate SFO. Such a solution is novel and vastly different from previous solutions. For example, Chronus and Splicer use multiple channels to correct SFO, which introduces unavoidable communication disruption. Other benefits of our system include the adaptive clustering for direct-path identification and model-based parameter estimation that avoids discretization errors. By contrast, SpotFi's clustering is fixed and cannot handle dynamic paths. On the other hand, our system does have some limitations, which are as follows.

(1) Prolonged Processing Delay: Compared with other single-AP schemes [29, 35], SiFi introduces more delays. It comes from two main factors. First, SiFi uses more packets and an adaptive clustering algorithm to achieve more accurate results. Specifically, the computation complexity of SiFi is $O(N_s * N_p^3 * N_c)$ while those of CUPID and Splicer are $O(N_s)$ and $O(N_s * N_p * N_{ch})$, where N_s, N_p, N_c, N_{ch} are the numbers of subcarriers, packets, clusters, and channels, respectively. We can see that SiFi has the highest computation complexity and thus best accuracy. To some extent, the better localization accuracy comes at the expense of time-efficiency. Therefore, SiFi is not suitable for real-time applications, e.g., tracking moving devices. Promising improvements for this include coherent MUSIC that can process multi-packet with different time delays, and advanced clock calibration techniques that eliminate the need of clustering. Second, the scheduling time in the medium access layer is another contributing factor. Currently, SiFi localizes a single user once at a time and thus cannot handle multiple users at the same time. This problem applies to most single-AP solutions and needs more investigation. The potential solutions involve broadcasting mechanisms for multi-user localization on the client side and novel distributed localization architectures that can do load-balancing of localization requests on the AP side.

(2) Limited Coverage Area: Both Chronus and our system require cable extensions for antennas, which incurs some signal losses. A decent off-the-shelf antenna extension cable can achieve about 0.1 dB/ft loss at 5 GHz, which translates to about 3.2 dB loss if a 10-meter cable is required. Hence, SiFi's current design works the best with a small office, like 10 m × 10 m. For a larger area, we can deploy multiple single-AP systems like we have done in our evaluation or explore multi-AP localization solutions using both ToA and AoA where AoA ensures fair coverage and ToA provides accurate distance measurements.

(3) Completion Latency. As our system currently uses the Linux CSI Tool [17], CSIs can only be obtained after the packet is successfully and fully decoded, incurring unnecessary latency. In fact, the CSI acquisition can be done at the first few byte of the packet and it should not be related to successful decoding, i.e., CSIs may be acquired even when the packet's payload is corrupted. To do so, we intend to migrate our system into software-defined radio platforms where more controls and flexibility are provided.

Besides, there are some directions worth future investigations.

(1) Existence of the Direct Path: Almost most of the existing direct path solutions, including ours, assume there do exist the direct path. However, the signal along the direct path might be too weak to detect due to serious occlusion. This can result in large localization errors as one of the reflection paths is deemed as the direct path. We may leverage the geometry of the target and AP to eliminate such outliers when more antennas and APs are available.
(2) More Antennas/Bandwidth: SiFi could be improved by extending the current design to 3D localizations by leveraging more antennas of advanced off-the-shelf APs, e.g., NETGEAR R8500, ASUS RT-AC88U, and Netis WF2471, which are equipped with 4 antennas. Also, we plan to explore potentials of the 802.11ac design, which promises to boost accuracy as the bandwidth of data channel increases to 80/160 MHz.
(3) Data Fusion: Although SiFi is a pure WiFi-based system for now, it has lots of opportunities to deliver higher accuracy when inertial sensors become available to clients, such as accelerometer and gyroscope.

The main result of SiFi is that it can achieve sub-meter localization accuracy using a single AP with 3 antennas. Another benefit of SiFi is compatibility with ongoing communication. These pros make it suitable for a range of applications. For example, SiFi can be used to extend drones' localization capability when GPS is not available. SiFi can protect drones from crashing or maintain a safe distance from the target for indoor environments. Another realworld application is that small businesses that usually only has a single AP can use SiFi to provide WiFi connectivity to customers within the store, restricting connections outside the facility.

Overall, we believe SiFi pushes the limits of single-AP indoor localization using a single channel. It offers decent localization accuracy while keeping data communication unaffected. Its main insight is that only one value for the delay distortion that should cause all the direct-path ToAs of antennas to intersect at a single point due to the frequency-locked-antenna design of MIMO. We believe that SiFi can benefit current indoor localization and navigation services greatly in many ways due to its simple requirement.

A MODEL-BASED PARAMETER ESTIMATION

To find peaks of Equation 6, we translate it into polynomial root finding problems. Specifically, as subcarriers are equally spaced, if we use Δf to denote the frequency spacing of two consecutive subcarriers [9], we can rewrite the steering vector as

$$s_l(\tau) = \mathrm{e}^{-j2\pi f_{i_1}\tau}[1, z, ..., z^{l-1}]^T, \tag{14}$$

$$z \equiv \mathrm{e}^{-j2\pi\Delta f\tau} \tag{15}$$

At the same time, the denominator of projection function can be rewritten as

$$\mathcal{D}^{-1}(\tau) = s_l^*(\tau)\mathrm{C}s_l(\tau), \tag{16}$$

$$\mathrm{C} \equiv \mathcal{U}_{\mathbf{n}}\mathcal{U}_{\mathbf{n}}^*. \tag{17}$$

[9] For example, on a 40 MHz channel, Intel 5300 WiFi cards output the CSI of size 30 out of 114 subcarriers. The indexes are [-58,-54,-50,...,-2,2,...,50,54,58], which means those measured subcarriers are equally spaced by $\Delta f = 312.5$ KHz $\cdot\ 4 = 1.25$ MHz.

Hence, we can put Equation 14 and 16 together to obtain polynomials, i.e.,

$$\mathcal{D}^{-1}(\tau) = \mathrm{e}^{-j4\pi f_{i_1}\tau}\tilde{\mathcal{D}}^{-1}(\tau), \tag{18}$$

$$\tilde{\mathcal{D}}^{-1}(\tau) = \sum_{p=-l+1}^{l-1} c_p z^{-p}, \tag{19}$$

$$c_p \equiv \sum_{i-j=p} \mathrm{C}(i,j), \tag{20}$$

where i and j are the row and column index of matrix $\mathbf{C}$, respectively, c_p is the sum of entries of $\mathbf{C}$ along the q-th diagonal.

It is easy to see that the polynomial of Equation 19 is with $(2l-2)$ roots. Those roots come in pair as $(z, 1/z^*)$, that have the same phase but reciprocal amplitudes. Note that only the phase of root carries our interested parameter τ. So first we need to find the roots of Equation 19 in $(l-1)$ pairs and only keep $(l-1)$ roots that are within the unit circle. Then we pick up the K roots that are closest to the unit circle. Finally, we put the K roots into Equation 15, resulting in the estimated ToA spread $(\hat{\tau}_1, \hat{\tau}_2, ..., \hat{\tau}_K)$. Then corresponding amplitudes $(\hat{a}_1, \hat{a}_2, ..., \hat{a}_K)$ are derived by a simple linear regression.

REFERENCES

[1] 2009. IEEE 802.11n-2009 standard. (2009). http://standards.ieee.org/getieee802/download/802.11n-2009.pdf.

[2] 2016. Wi-Fi Tags. (2016). http://www.ekahau.com/real-time-location-system/technology/wi-fi-tags.

[3] Fadel Adib, Zachary Kabelac, Dina Katabi, and Robert C Miller. 2014. 3D Tracking via Body Radio Reflections.. In *Proc. of USENIX NSDI.*

[4] Marcello Ascione, Aniello Buonanno, Michele D'Urso, Leopoldo Angrisani, and Rosario Schiano Lo Moriello. 2013. A new measurement method based on music algorithm for through-the-wall detection of life signs. *IEEE Transactions on Instrumentation and Measurement* 62, 1 (2013), 13–26.

[5] Paramvir Bahl and Venkata N Padmanabhan. 2000. RADAR: An In-Building RF-based User Location and Tracking System. In *Proc. of IEEE INFOCOM.*

[6] Arthur J Barabell. 1983. Improving the resolution performance of eigenstructure-based direction-finding algorithms. In *Proc. of IEEE ICASSP.*

[7] Gilles Celeux and Gérard Govaert. 1995. Gaussian parsimonious clustering models. *Pattern recognition* 28, 5 (1995), 781–793.

[8] Yuchi Chen, Wei Gong, Jiangchuan Liu, and Yong Cui. 2018. I Can Hear More: Pushing the Limit of Ultrasound Sensing on Off-the-Shelf Mobile Devices. In *Proc. of IEEE INFOCOM.*

[9] Krishna Chintalapudi, Anand Padmanabha Iyer, and Venkata N Padmanabhan. 2010. Indoor Localization Without the Pain. In *Proc. of ACM MobiSys.*

[10] Baron Gaspard Riche de Prony. 1795. Essai éxperimental et analytique: sur les lois de la dilatabilité de fluides élastique et sur celles de la force expansive de la vapeur de lalkool,a différentes températures. *Journal de lécole polytechnique* 1, 22 (1795), 24–76.

[11] Haishi Du, Ping Li, Hao Zhou, Wei Gong, Gan Luo, and Panglong Yang. 2018. WordRecorder: Accurate Acoustic-based Handwriting Recognition Using Deep Learning. In *Proc. of IEEE INFOCOM.*

[12] Tolga Eren, OK Goldenberg, Walter Whiteley, Yang Richard Yang, A Stephen Morse, Brian DO Anderson, and Peter N Belhumeur. 2004. Rigidity, computation, and randomization in network localization. In *Proc. of IEEE INFOCOM.*

[13] Xiaoyi Fan, Wei Gong, and Jiangchuan Liu. 2017. i2tag: RFID mobility and activity identification through intelligent profiling. *ACM Transactions on Intelligent Systems and Technology (TIST)* 9, 1 (2017), 5.

[14] Jon Gjengset, Jie Xiong, Graeme McPhillips, and Kyle Jamieson. 2014. Phaser: Enabling Phased Array Signal Processing on Commodity WiFi Access Points. In *Proc. of ACM MobiCom.*

[15] W. Gong, S. Chen, J. Liu, and Z. Wang. 2018. MobiRate: Mobility-Aware Rate Adaptation Using PHY Information for Backscatter Networks. In *Proceedings of IEEE INFOCOM.*

[16] Wei Gong and Jiangchuan Liu. 2017. Robust Indoor Wireless Localization Using Sparse Recovery. In *Proc. of IEEE ICDCS.*

[17] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2011. Tool release: gathering 802.11 n traces with channel state information. *ACM SIGCOMM Computer Communication Review* 41, 1 (2011), 53–53.

[18] Guangjie Han, Huihui Xu, Trung Q Duong, Jinfang Jiang, and Takahiro Hara. 2013. Localization algorithms of wireless sensor networks: a survey. *Telecommunication Systems* (2013), 1–18.

[19] Yingbo Hua and Tapan K Sarkar. 1990. Method for Estimating Parameters of bxponentially Damped/Undamped Sinusoids in Noise. *IEEE Transactions on Acoustics, Speech and Signal Processing* 38, 5 (1990), 814–824.

[20] Suman Jana and Sneha Kumar Kasera. 2008. On Fast and Accurate Detection of Unauthorized Wireless Access Points Using Clock Skews. In *Proc. of ACM MobiCom.*

[21] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. 2015. Spotfi: Decimeter Level Localization Using WiFi. In *Proc. of ACM SIGCOMM.*

[22] Swarun Kumar, Stephanie Gil, Dina Katabi, and Daniela Rus. 2014. Accurate Indoor Localization With Zero Start-up Cost. In *Proc. of ACM MobiCom.*

[23] Swarun Kumar, Ezzeldin Hamed, Dina Katabi, and Li Erran Li. 2014. LTE Radio Analytics Made Easy and Accessible. In *Proc. of ACM SIGCOMM.*

[24] Wenjing Liao and Albert Fannjiang. 2016. MUSIC for Single-Snapshot Spectral Estimation: Stability and Super-resolution. *Applied and Computational Harmonic Analysis* 40, 1 (2016), 33–67.

[25] Alex T Mariakakis, Souvik Sen, Jeongkeun Lee, and Kyu-Han Kim. 2014. SAIL: Single Access Point-Based Indoor Localization. In *Proc. of ACM MobiSys.*

[26] Hariharan Rahul, Haitham Hassanieh, and Dina Katabi. 2011. SourceSync: A Distributed Wireless Architecture for Exploiting Sender Diversity. In *Proc. of ACM SIGCOMM.*

[27] Hariharan Rahul, Swarun Kumar, and Dina Katabi. 2012. MegaMIMO: Scaling Wireless Capacity with User Demands. In *Proc. of ACM SIGCOMM.*

[28] Ralph O Schmidt. 1986. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation* 34, 3 (1986), 276–280.

[29] Souvik Sen, Jeongkeun Lee, Kyu-Han Kim, and Paul Congdon. 2013. Avoiding Multipath to Revive Inbuilding WiFi Localization. In *Proc. of ACM MobiSys.*

[30] Michael Speth, Stefan A Fechtel, Gunnar Fock, and Heinrich Meyr. 1999. Optimum receiver design for wireless broad-band systems using OFDM. I. *IEEE Transactions on Communications* 47, 11 (1999), 1668–1677.

[31] Petre Stoica and Nehorai Arye. 1989. MUSIC, Maximum Likelihood, and Cramer-Rao Bound. *IEEE Transactions on Acoustics, Speech and Signal Processing* 37, 5 (1989), 720–741.

[32] Petre Stoica and Randolph L Moses. 1997. *Introduction to spectral analysis.* Vol. 1. Prentice hall Upper Saddle River.

[33] Deepak Vasisht, Swarun Kumar, and Dina Katabi. 2016. Decimeter-Level Localization with a Single WiFi Access Point. In *Proc. of USENIX NSDI.*

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

[35] Yaxiong Xie, Zhenjiang Li, and Mo Li. 2015. Precise Power Delay Profiling with Commodity WiFi. In *Proc. of ACM MobiCom.*

[36] Jie Xiong and Kyle Jamieson. 2013. ArrayTrack: A Fine-Grained Indoor Location System. In *Proc. of USENIX NSDI.*

[37] Jie Xiong, Karthikeyan Sundaresan, and Kyle Jamieson. 2015. ToneTrack: Leveraging Frequency-Agile Radios for Time-Based Indoor Wireless Localization. In *Proc. of ACM MobiCom.*

[38] Zheng Yang, Yunhao Liu, and X-Y Li. 2009. Beyond trilateration: On the localizability of wireless ad-hoc networks. In *Proc. of IEEE INFOCOM.*

[39] Zheng Yang, Chenshu Wu, and Yunhao Liu. 2012. Locating in Fingerprint Space: Wireless Indoor Localization with Little Human Intervention. In *Proc. of ACM MobiCom.*

[40] 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.

[41] Moustafa Youssef and Ashok Agrawala. 2005. The Horus WLAN Location Determination System. In *Proc. of USENIX MobiSys.*