SaTrack: LoS/NLoS State-Aware WiFi Indoor Tracking System

Yinnan Zhou*†, Hao Zhou*†, Luowei Li*†, Siming Chen*†, Zhi Liu‡ and Xiang-Yang Li*†
*LINKE Lab, School of Computer Science and Technology, University of Science and Technology of China
†CAS Key Laboratory of Wireless-Optical Communications, University of Science and Technology of China
‡Department of Computer and Network Engineering, The University of Electro-Communications, Tokyo 182-8585, Japan
Email: yinnanzhou@mail.ustc.edu.cn, kitewind@ustc.edu.cn,
{liluowei, chensiming}@mail.ustc.edu.cn, liu@ieee.org, xiangyangli@ustc.edu.cn

***Abstract*—WiFi-based technology is appealing for indoor localization due to the widely deployed infrastructures. Recently, path separation solutions have been proposed to address the multipath effects and achieve decimeter-level localization accuracy in line-of-sight (LoS) scenarios. However, these solutions experience serious performance degradation in non-line-of-sight (NLoS) scenarios, and couldn't be used for mobile device tracking where continuous LoS/NLoS switching happens. In this paper, we propose SaTrack, a LoS/NLoS state-aware mobile device tracking system. SaTrack identifies LoS/NLoS states based on the diversity of the strongest estimated paths when using different reference antennas. With the observation of spatial aggregation and temporal continuity for the Tx-Rx direct path, SaTrack chooses the direct path through two-step clustering, i.e., clustering in the spatial domain to form candidates and clustering again in the temporal domain to select the winner. Extensive experiments are conducted to evaluate the effectiveness of SaTrack. In a typical indoor environment with abundant multipath, SaTrack achieves 0.64m and 1.27m for the median and 90th percentile tracking errors, outperforming the state-of-the-art (SOTA) solutions.**

I. Introduction

Indoor localization technologies are widely used in various applications such as navigation in shopping malls or underground parking lots. These technologies include cameras [1], inertial sensors [2], ultrasound [3], Bluetooth [4], RFID [5, 6], infrare [7], UWB [8, 9], and WiFi [10, 11, 12]. Researchers have focused on WiFi-based technologies due to the low cost of deployment and versatility in deployment scenarios [13, 14], making them the most promising option.

WiFi-based localization requires obtaining fine-grained path information since WiFi signals are multipath superimposed signals in typical indoor environments. Such a target could be fulfilled through path separation or subspace solutions where the former has better performance. For example, the Expectation Maximization (EM)-based path separation solutions achieve decimeter-level localization accuracy [15, 16].

The general path separation solutions consist of two main tasks, i.e., path estimation and path selection [11, 17]. For the path estimation task, signal decomposition methods are commonly used to extract possible signal transmission paths from the superimposed WiFi signals. Here, the estimated paths are characterized by various parameters such as Angle-of-Arrival (AoA), Time-of-Flight (ToF), Doppler Shift, and Complex Attenuation [18]. For the path selection task, the direct path between the transmitter (Tx) and receiver (Rx) is chosen from the estimated paths. Subsequently, the device can be localized based on the parameters of the Tx-Rx direct path.

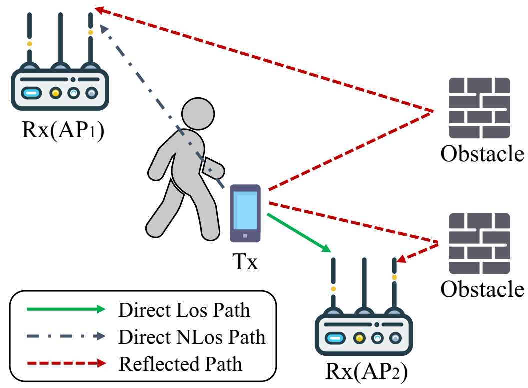


Fig. 1. A typical indoor mobile device tracking scenario.

However, the distinctions between Line-of-Sight (LoS) and Non-Line-of-Sight (NLoS) conditions pose great challenges to the path separation task [16]. The current path estimation process is usually performed without the awareness of the LoS/NLoS states. Moreover, for path selection, the inherent characteristics of the direct path may not hold in NLoS scenarios. For example, the assumption that the direct path corresponds to the strongest power is valid only in LoS conditions [16]. Therefore, current solutions suffer from obvious performance degradation in NLoS scenarios.

As we know, LoS/NLoS state switching is common for mobile device tracking applications [19]. Fig. 1 depicts a typical scenario where one person walks around with a handheld mobile device acting as Tx, and multiple access points (APs) acting as Rx to receive signals and conduct localization. We can see frequent LoS/NLoS state switching of the Tx-Rx direct path due to human body obstruction, which impacts path separation and leads to unstable tracking results. Therefore, we need a LoS/NLoS state-aware solution that utilizes the state information to ensure reliable path separation and achieves accurate indoor mobile device tracking. To implement such a WiFi tracking system, we need to address the following challenges:

- Real-time LoS/NLoS state identification. On one hand, the signal decomposition methods employed in the existing solutions fail to comprehend the physical differences between the paths. On the other hand, although statistical characteristics can be used to identify LoS and NLoS states after receiving a large number of packets [20], the time needed will contradict the purpose of mobile device tracking. Thus, it is non-trivial to implement a real-time state identification method within the framework compatible with existing path separation solutions.

- Accurate direct path selection in NLoS state. Effective path selection metrics are still lacking in NLoS scenarios even if NLoS states could be identified. It is difficult to accurately select a direct path due to power degradation caused by obstruction, resulting in the direct path exhibiting similar characteristics to other reflected paths.

In this paper, we successfully address these challenges and propose a LoS/NLoS **S**tate-**A**ware **Track**ing System, denoted as SaTrack. Our contributions are summarized as follows:

1) To the best of our knowledge, SaTrack is the first LoS/NLoS-aware path separation solution, which is robust to LoS/NLoS state switching and therefore suitable for indoor mobile device tracking.
2) We analyze the decomposition results of the EM-based path separation algorithm. We solve the LoS/NLoS identification problem based on the observation of the strongest path diversity, which means that the strongest estimated paths in the NLoS state (i.e., reflected paths in most cases) exhibit obvious diversity when using different reference antennas. We further utilize the observation of spatial aggregation and temporal continuity to solve the direct path selection problem, i.e., using Density-Based Spatial Clustering of Applications with Noise (DBSCAN)-based algorithm to highlight the direct path candidates in the spatial domain, and cluster again in the temporal domain to select the winner.
3) We implement a prototype of SaTrack in multi-AP scenarios, and conduct extensive experiments to evaluate the effectiveness. The results demonstrate that SaTrack achieves the median and 90th percentile tracking error of 0.64m and 1.27m in typical indoor environments, outperforming the state-of-the-art (SOTA) solutions.

II. PRELIMINARY

In this section, we provide a preliminary exploration of WiFi-based indoor tracking.

A. Wireless Model

For the indoor environment, the received wireless signals at Rx can be regarded as superimposed signals transmitted from Tx along different paths. The characteristic of the l-th path can be expressed through multi-dimensional parameters, including Time of Flight (τ_l), Angle of Arrival (φ_l), Doppler Shift (γ_l) and Complex Attenuation (α_l) [17].

The WiFi Network Interface Card (NIC) measures channels discretely in time (packet), frequency (subcarrier), and space (antenna) [10]. Assuming that the Channel State Information (CSI) measurement is a superimposed signal of L path signals, we can denote the discrete CSI measurement at index $\langle i,j,k\rangle$ (i.e., the i-th packet, j-th subcarrier, and k-th antenna) as

$$\begin{cases} H(i,j,k) = \sum_{l=1}^{L} H_l(i,j,k) + N(i,j,k), \\ H_l(i,j,k) = \alpha_l e^{-j2\pi f_l(i,j,k)}, \\ f_l(i,j,k) \approx f_c\tau_l + \Delta f_j \tau_l + f_c \Delta s_k \cdot \varphi_l - \gamma_l \Delta t_i, \end{cases} \tag{1}$$

where $H_l(i,j,k)$ is the signal of the l-th path, $N(i,j,k)$ is the complex Gaussian white noise, $f_l(i,j,k)$ is the signal phase (divided by 2π) of the l-path in $H(i,j,k)$ taking $H(0,0,0)$ as reference, f_c is the carrier frequency of the channel, Δf_j, Δs_k and Δt_i are differences of frequency, spatial position and time between $H(i,j,k)$ and $H(0,0,0)$.

B. Framework of Existing Path Separation Solutions

Current SOTA path separation solutions usually contain three parts, i.e., CSI pre-processing, EM-based path estimation, and path selection.

1) CSI Pre-processing: Raw CSI measurements could not be directly used to estimate parameters due to the existence of hardware defects caused by phase noises [21, 17]. Specifically, the erroneous version of signal measurement is

$$\widetilde{H}(i,j,k) = H(i,j,k) e^{2\pi(\Delta f_i \delta_{t_i} + \Delta t_i \delta_f) + \zeta}, \tag{2}$$

where δ_{t_i} and δ_f are the Time Offset (TO) and Carrier Frequency Offset (CFO) between Rx-Tx, respectively. ζ is the initial phase of the receiver antenna.

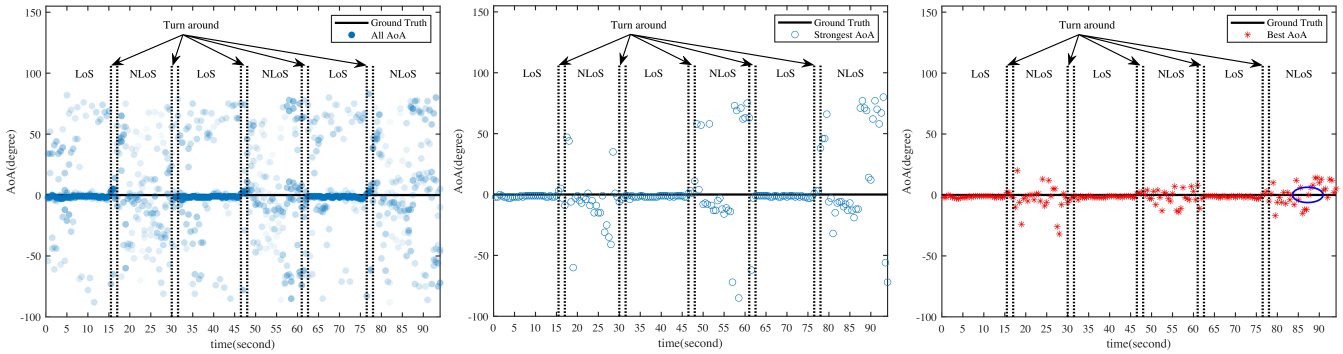
Conjugate multiplication is a typical method to remove phase noises [12, 21], the principle of which is to compensate $\widetilde{H}(i,j,k)$ based on the phase noises of $\widetilde{H}^*(i,j,k_0)$. This is because the CSI noises, such as TO and CFO, only vary in the time and frequency domain instead of the spatial domain. We can select a reference antenna (e.g., k_0-th antenna) and calculate the conjugate multiplication $\hat{H}(i,j,k)$ as follows, where TO and CFO are successfully eliminated.

$$\begin{aligned} \hat{H}(i,j,k) &= \widetilde{H}(i,j,k) \times \widetilde{H}^*(i,j,k_0) \\ &= H(i,j,k) e^{2\pi(\Delta f_i \delta_{t_i} + \Delta t_i \delta_f) + \zeta} \times H^*(i,j,k_0) \\ &\quad e^{-2\pi(\Delta f_i \delta_{t_i} + \Delta t_i \delta_f) - \zeta} \\ &= H(i,j,k) \times H^*(i,j,k_0). \end{aligned} \tag{3}$$

2) EM-based path estimation: Current SOTA separation solutions, such as Widar2.0 [17] and mD-Track [18], generally use an EM-based path estimation algorithm to separate multipath superimposed signals.

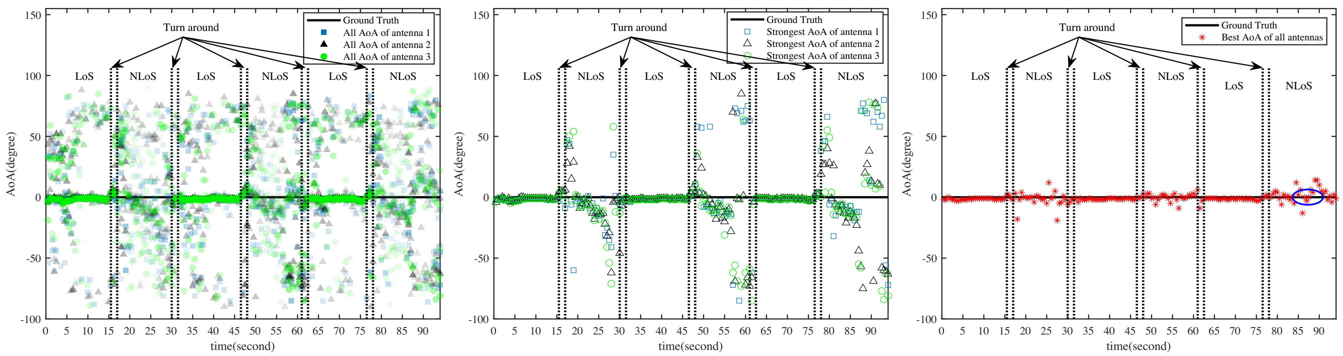
The input of the EM-based path estimation algorithm is the multipath CSI measurement after pre-processing, i.e., $\hat{H}(i,j,k)$. The output is the maximum likelihood estimate (MLE) parameters for multiple paths.

Specifically, the EM-based algorithm assumes that there are L possible paths in the space, normally set L=5 since there are usually 5 major signaling paths in indoor environments [11, 22]. The algorithm initializes the parameters of all paths to zeros, and then iteratively performs the Expectation step and the Maximization step to separate the path signals and update the path parameters until convergence [18, 23]. The



(a) All AoA estimated results. (b) The strongest estimated paths. (c) The best estimated paths.

Fig. 2. AoAs estimated results with one reference antenna.



(a) All AoA estimated results of all antennas. (b) The strongest estimated paths of all antennas. (c) The best estimated paths of all antennas.

Fig. 3. Influence of different reference antennas.

Expectation step reconstructs the signal for each path based on the estimated parameters from the last iteration and compares them with the measured signal to obtain the "expected" signal for each path in the current iteration. The Maximization step updates the parameters for each path to well fit with its "expected" signal.

3) Path selection: The goal of path selection is to select the direct path from the estimated results. Current solutions generally use the Complex Attenuation as the metric [17, 24], because it indicates the power strength of the signal and the signal power of the direct path tends to be the strongest in LoS scenarios. However, this assumption couldn't be assured in NLoS scenarios since the direct path is obstructed.

C. Path Decomposition Results

In this subsection, we analyze the results of current path separation solutions to demonstrate their drawbacks and seek guidance for performance enhancement in NLoS scenarios.

1) Performance degradation in NLoS scenarios: We conduct a preliminary experiment for mobile device tracking. A volunteer walks in the room with the Tx in hand, then moves toward the Rx, turns around, and returns to the starting point. Such process is repeated three times. For detailed deployments and configurations of the Tx/Rx, please refer to Sect. IV-A.

Fig. 2(a) shows the AoA estimation results of the EM-based path separation algorithms, where the color indicates the complex attenuation, i.e., the stronger path appears darker.

Firstly, we witness obvious LoS/NLoS state switching during the experiment due to the human body obstruction. This phenomenon shows the general existence of NLoS scenarios, which couldn't be ignored for mobile device tracking.

Secondly, we notice that the strongest estimated paths keep consistency with the ground truth in LoS scenarios, but are biased in NLoS scenarios as shown in Fig. 2(b). The reason lies in the fact that the obstructed direct path might have much lower power than the reflected paths. Such observations imply that we need additional metrics for direct path selection in NLoS beyond traditional strongest attenuation metrics.

Last, we examine the best estimated paths, i.e., those closest to the ground truth, as shown in Fig. 2(a). They are close to the ground truth, demonstrating the ability of the EM-based algorithm to extract the direct path in most cases, even in NLoS scenarios. Meanwhile, the gap might be large in certain cases, e.g., the elliptical region in Fig. 2(c), which introduces additional difficulty to our device tracking system.

2) The influence of reference antennas: As introduced in the previous subsection, current EM-based path separation solutions choose one reference antenna to perform conjugate multiplications for phase noise elimination. Here, we evaluate the influence of reference antenna choices.

Fig. 3(a) shows the path estimation results using different reference antennas under the same setup as in Fig. 2. After careful inspection, we make the following observation:

The strongest estimated paths: We focus on the distribution of the strongest estimated paths as in Fig. 3(b). In LoS states, we notice clear consistency, i.e., all the strongest paths from different reference antennas keep close to the ground truth, while in NLoS states, there exists obvious diversity.

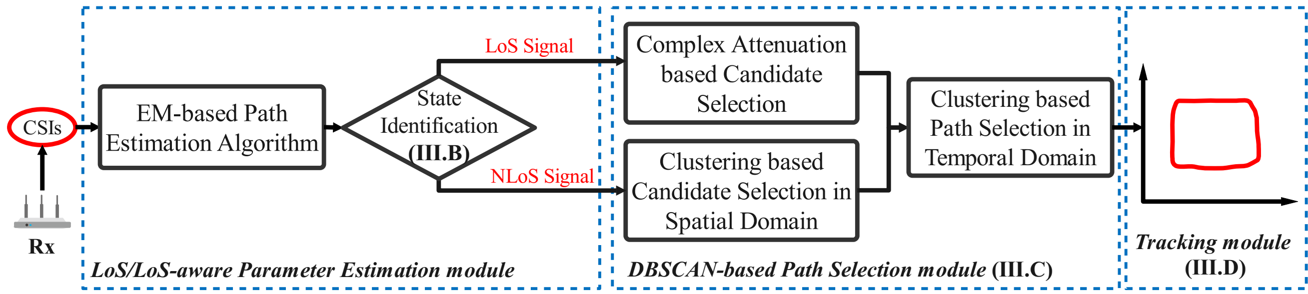


Fig. 4. System overview of SaTrack.

The observation can be explained as follows. Let's take a close look at Eq. (3) in the conjugate multiplication operation, i.e., $\hat{H}(i,j,k) = H(i,j,k) \times H^*(i,j,k_0)$. $\hat{H}(i,j,k)$ is dominated by the signals from strong paths, i.e., $\hat{H}(i,j,k) \approx \sum_{l_1,l_2 \in \mathcal{L}^{strong}} H_{l_1}(i,j,k) \times H^*_{l_2}(i,j,k_0)$, where $\mathcal{L}^{strong}$ denotes the set of strong paths. In LoS scenarios, the direct path is usually the strongest, with one order of magnitude higher than the reflected paths [18], i.e., $\mathcal{L}^{strong} = \{l^{direct}\}$. It means $\hat{H}(i,j,k)$ is dominated by the direct path itself with whatever k_0. Thus, the estimations for the strongest estimated path are consistent across different reference antennas. While in NLoS scenarios, the reflected paths might be stronger than the obstructed direct path. Due to the presence of multiple reflected paths (i.e., $|\mathcal{L}^{strong}| > 1$), the result of $\hat{H}(i,j,k)$ are far more complex than that in LoS scenarios, and the differences of the chosen reference antennas cannot be ignored.

In summary, the obstruction status of the direct path determines the distribution of the strongest estimated paths. Therefore, we can identify the LoS/NLoS state by examining the distribution.

The best estimated paths: We now focus on the best estimated paths in NLoS scenarios as shown in Fig. 3(c), i.e., whether they are close enough to the ground truth, and do they have some special properties helping for identification from the other estimated paths.

Firstly, we witness that the best estimated paths show spatial aggregation near the ground truth since the EM-based algorithm could extract the correct direct path in most cases. Furthermore, by comparing the elliptical region with Fig. 2(c), the gap is narrowed to the ground truth because the special cases in Fig. 2(c) can be compensated through multiple executions of EM-based algorithm with different reference antennas.

Secondly, we also notice the temporal continuity of the best estimated paths, i.e., there is no sudden jump along the time. The reason lies in the fact that the tracking mobile device moves continuously in the temporal domain.

As for now, these observations could help us perform direct path selection, i.e., we can use spatial aggregation to choose possible candidates, and further use temporal continuity to refine the path selection results.

III. System Design

In this section, we give a detailed explanation of our implementation of the SaTrack system.

A. System Overview

As shown in Fig. 4, SaTrack consists of three main modules, namely path estimation, path selection and tracking.

The path estimation module includes path decomposition and LoS/NLoS state identification. The former is achieved through a refined EM-based algorithm. The specific details are omited here, as the algorithm has been explained in the previous section, with the modification of using all three antennas as reference antennas. The latter is then accomplished based on the properties of the strongest path distribution.

In the path selection module, SaTrack processes the path estimation results separately according to their LoS/NLoS states. SaTrack employs a traditional method based on Complex Attenuation for LoS signals and applies the spatial aggregation distribution to cluster the results of NLoS signals. Subsequently, temporal continuity is utilized to further cluster the processed results to select the direct path.

In the tracking module, SaTrack is deployed in multi-AP scenarios. Based on the estimated AoA parameters, we utilize the least squares method to obtain the target location.

The technical details are introduced in the following parts.

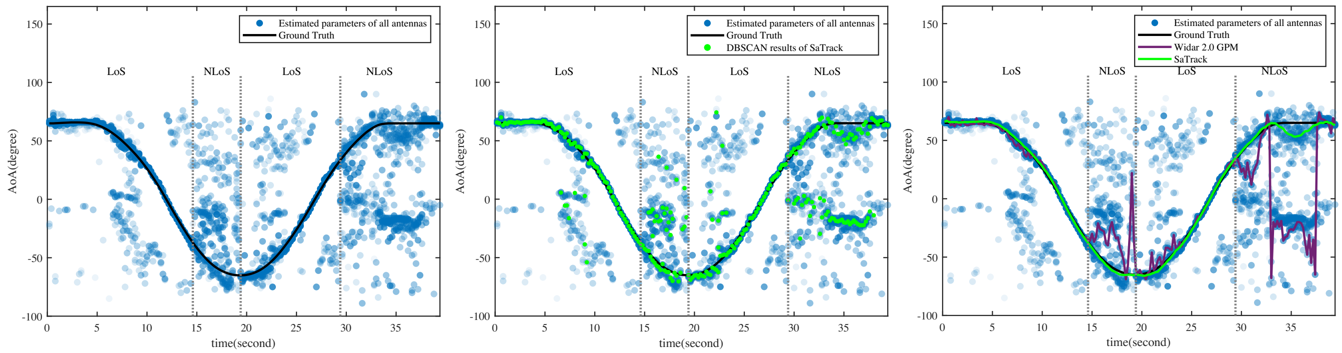
B. State Identification

Based on the preliminary experiment results in Sect. II-C, we can identify the LoS/NLoS state based on the distribution of the strongest estimated paths as they show obvious diversity in NLoS scenarios.

Specifically, our approach consists of three steps. First, we use each antenna as the reference antenna and obtain their path estimation results through an EM-based algorithm. Then, we select the strongest estimated path for each reference antenna individually. At last, we compute the variance of the AoA for these strongest paths and compare it with a pre-defined threshold to identify the LoS/NLoS states.

C. Path Selection

In the path selection module, the inputs include the multidimensional parameters of multiple estimated paths and the state identification results. For LoS state, we use the traditional selection method with Complex Attenuation as the metric, i.e., select the strongest path as the direct path. For NLoS state, we utilize spatial clustering to cluster the parameter estimation results and obtain more accurate direct path candidates. Then we cluster all LoS & NLoS estimation outcomes by employing temporal continuity, ultimately determining the optimal direct path. Such a process includes the following two steps:



(a) Estimation results.
(b) Clustering results in the spatial domain.
(c) Clustering results in the temporal domain.

Fig. 5. Examples for direct path selection.

1) Spatial domain clustering: With the spatial aggregation property of the best estimated path, we can cluster all the estimated paths to obtain reliable direct path candidates.

Here, we employ the DBSCAN method to cluster the AoAs of the estimated paths. Specifically, the key parameters of DBSCAN include Epsilon (i.e., neighborhood radius) and minPts (i.e., the smallest number of elements of the cluster). To ensure the distance between the clustered results, we set Epsilon as AoA resolution times a given constant, e.g., 5 in our prototype. Meanwhile, we set minPts equal to the number of Rx antennas M, since any physical path should have M corresponding estimations after the EM-based algorithm is executed with M reference antennas separately.

Fig. 5(a) plots a segment of AoA estimation results. As shown in Fig. 5(b), the original scattered estimations (blue) form several well-separated candidates (green) after clustering, which is beneficial for subsequent direct path selection.

2) Temporal domain selection: Based on the observation of temporal continuity, we further cluster the candidates in the temporal domain to obtain the final direct path.

In our prototype, we execute EM-based path estimation periodically (e.g., once every 0.1 seconds), yielding discrete estimates in the time domain. We use a sliding window with a duration of 1s (i.e., 10 time slices) and perform DBSCAN-based clustering within the window. We set the same Epsilon parameter as that for clustering in the spatial domain, and set minPts as half of the time slice amount in the sliding window.

This method allows us to utilize the reliable estimated results in LoS states to refine the ones in NLoS states. Fig. 5(c) compares the path selection of SaTrack with the graph-based path matching (GPM) algorithm in Widar2.0 [17]. We can see serious performance degradation of GPM in NLoS scenarios since it is LoS/NLoS-unaware, while results from SaTrack are more reliable and stable.

D. Tracking

So far, we have obtained the LoS/NLoS states and the parameters of the Tx-Rx direct path. This information can support mobile device localization. In this section, we introduce the localization method in multi-AP scenarios.

SaTrack utilizes AoA-based localization since it is a widely used and accurate localization technique. Specifically, we denote (x, y) as the coordinate of the target, and (x_i, y_i) as that for i-th AP whose estimated AoA is φ_i. Then we have the $tan\,\varphi_i = \frac{y-y_i}{x-x_i}$ as defined by geometry. By combining K APs, we obtain the linear equation group as

$$\begin{bmatrix} y_1 - x_1\,tan\,\varphi_1 \\ y_2 - x_2\,tan\,\varphi_2 \\ \vdots \\ y_K - x_K\,tan\,\varphi_K \end{bmatrix} = \begin{bmatrix} -tan\,\varphi_1 & 1 \\ -tan\,\varphi_2 & 1 \\ \vdots & \vdots \\ -tan\,\varphi_K & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \quad (4)$$

When $K = 2$, we can easily solve Eq. (4) and obtain a unique solution. When with $K \geq 3$, we couldn't directly solve the equation group due to the possible estimation errors. Thus, we should use the least squares method [25] to solve Eq. (4) and obtain the coordinates of the target device.

In this paper, we focus on AoA-based localization in multi-AP deployment since the estimated AoA is more accurate than other path parameters, such as RSSI and ToF. Moreover, although single-AP implementation is beyond the scope of this paper, we still emphasize that LoS/NLoS identification is also effective for such cases, e.g., we can perform compensation for direct path parameters (e.g., RSSI) in NLoS scenarios.

IV. Performance Evaluation

A. Experiment Setup

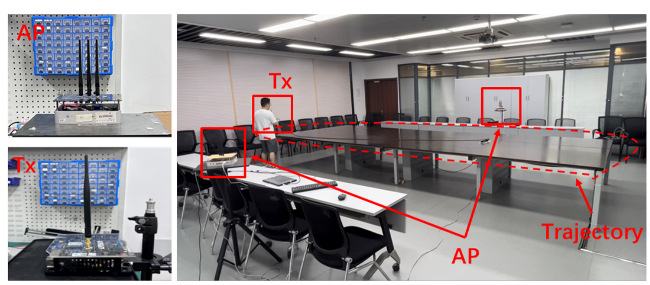
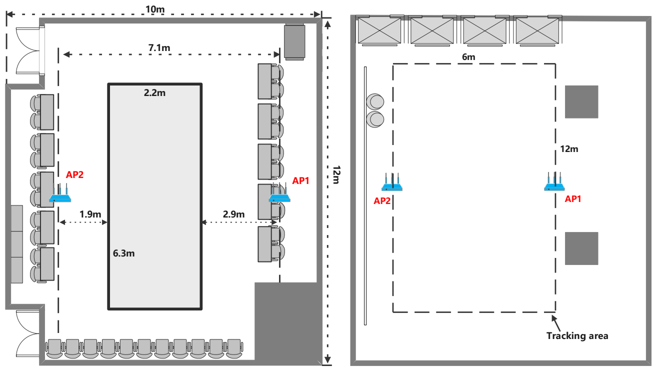


Fig. 6. A typical indoor mobile device tracking scenario.

1) Implementation: We implement SaTrack using Commercial Off-The-Shelf (COTS) WiFi devices with Intel 5300 WiFi cards. The transmitter has 1 antenna and broadcasts packets into the air. Each receiver has 3 antennas, which form a uniform linear array. Each pair of antennas is placed with a spacing of 2.6 cm (nearly $1/2$ wavelength of 5 GHz band), which is shown in Fig. 6. The Linux 802.11n CSI Tool [26] is installed in APs to collect CSI measurements. APs are set to work with monitor mode, on channel 165 at 5.825 GHz. The transmission rate of packets is set to 500 Hz. The desktop uses an Intel i5-11500 2.70Hz CPU and processes CSI data using MATLAB.

2) Deployment: To fully evaluate the performance of SaTrack, we conduct experiments in two different environments, i.e., a complex environment as a meeting room with tables and chairs, and a simple environment for an open public area. Figs. 7(a) and 7(b) illustrate the device deployment and tracking areas in these two environments, respectively. The volunteer walks along different shapes of trajectories, such as that shown in Fig. 8. To save system deployment costs, we evaluate SaTrack only with dual-APs set up in this paper.



(a) Meeting room deployments. (b) Public area deployments.

Fig. 7. Deployments in different environments.

3) Comparison baselines: We choose SpotFi [11] and Widar2.0 [17] with three antennas as baselines for comparison because they are the SOTA localization systems that can be deployed without any hardware or firmware modification to the AP [27]. Specifically, Widar2.0 uses a graph-based path matching method (GPM) to improve path selection accuracy. Note that since Widar2.0 is oriented towards passive localization, we compare only the EM-based path estimation and the GPM components of Widar2.0.

Here, we don't compare SaTrack with other SOTA solutions since they couldn't be deployed on typical commercial WiFi devices directly. For example, M^3 [15] employs frequency-hopping which interferes with communication, and NLoc [23] requires multiple antennas on both the Rx and the APs.

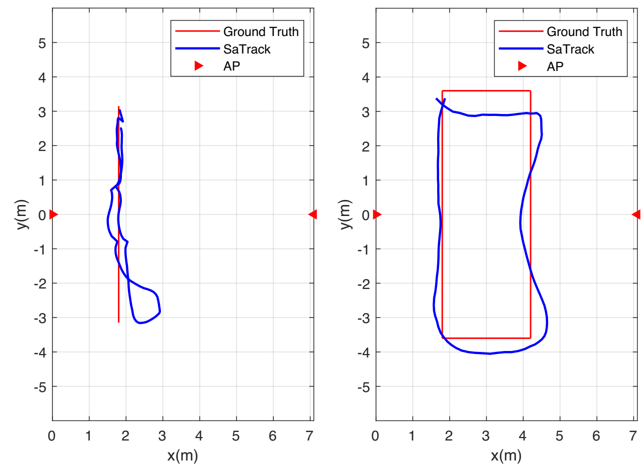
B. Overall Performance

We evaluate the performance of SaTrack in different environments. Following is a summary of the main results, with the experiment details given in the later subsections.

- SaTrack performs well during challenging movements with LoS/NLoS state switching. SaTrack achieves median tracking errors of 0.64m for typical indoor environments, outperforming SpotFi and Widar2.0, which achieve 0.92m and 0.85m, respectively. At the 90th percentile, SaTrack's tracking error is 1.27m compared to 1.39m and 2.45m from SpotFi and Widar2.0.
- SaTrack's accuracy stems from the state-identification-based path estimation and DBSCAN-based path selection. It has a median AoA estimation error of only 6.29 degrees, outperforming the 7.85 and 7.26 degrees of SpotFi and Widar2.0, respectively. Moreover, SaTrack achieves 12.30 degrees in the 90th percentile estimation error, with improvements of 7.87 and 22.23 degrees over SpotFi and Widar2.0, respectively.

C. Accuracy of Tracking

1) Evaluation in meeting room: The proposed SaTrack system is designed to support complex indoor environments. Therefore, we first evaluate it in a typical indoor environment with rich multipath, i.e., a meeting room of size 10m×12m with various furniture.



(a) Linear (b) Rectangular

Fig. 8. Trajectory localization results.

We conduct evaluations for dual-AP tracking. Fig. 8 plots examples of tracking results along with AP deployment. As shown in the figures, our system achieves decimeter-level localization results, close to the ground truth. Fig. 9 gives CDF curves for different tracking schemes. As shown in the figures, SaTrack achieves the median and 90th percentile tracking errors of 0.64m and 1.27m, outperforming SpotFi whose median and 90th percentile localization errors are 0.92m and 1.39m, respectively. SaTrack also outperforms Widar2.0's median and 90th percentile errors of 0.85m and 2.45m, respectively.

2) Impact of environments: In order to evaluate the robustness of SaTrack under different environments, we further conduct experiments in an open public area with 6m×12m, as shown in Fig. 7(b). As compared with the meeting room in Fig. 7(a), the chosen public area has less furniture which means fewer multipath effects.

Fig. 10 plots the CDFs for different solutions under two environments. As expected, both algorithms achieve better results with less than 0.5m median tracking error in the open public area since it has fewer reflected paths. Furthermore, SaTrack exhibits robust performance in different environments, i.e., the tracking error difference under two environments is 0.26m for median values and 0.41m for 90th percentile values. Meanwhile, Widar2.0 shows a larger gap, i.e., 0.67m differences for the 90th percentile tracking errors. This again arises from the performance degradation of EM-

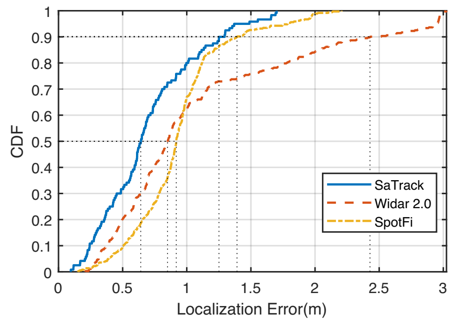


Fig. 9. Overall performance.

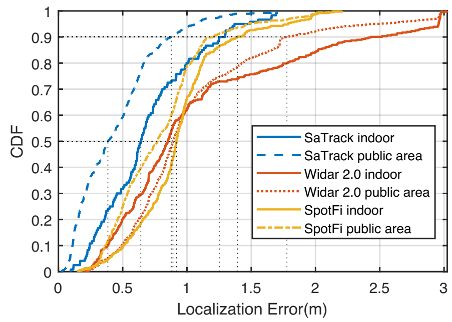


Fig. 10. Impact of environment.

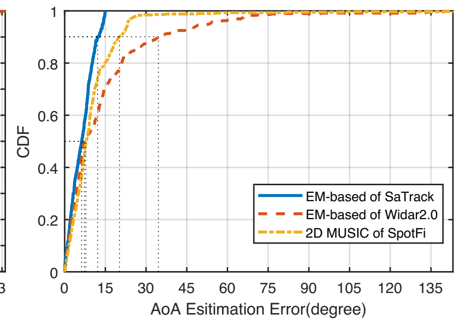


Fig. 11. AoA estimation error.

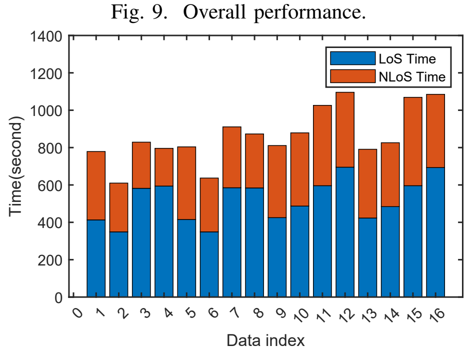


Fig. 12. Statistical of LoS/NLoS signals.

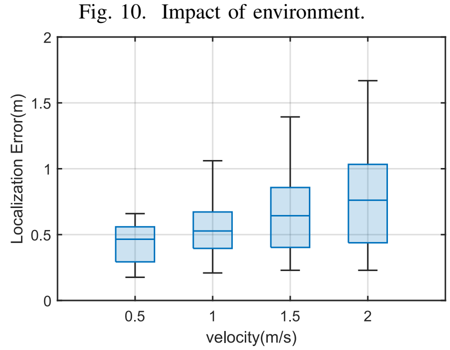


Fig. 13. Impact of walking velocity.

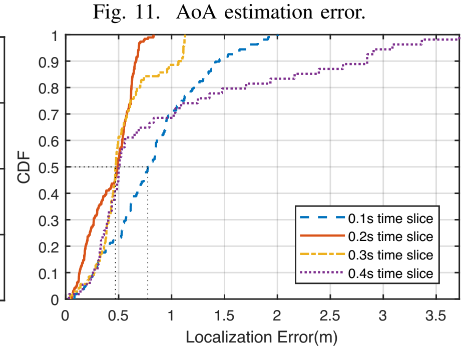


Fig. 14. Impact of time slice length.

based algorithms in NLoS states, especially in challenging environments with rich multipath and various occlusions.

D. Accuracy of AoA Estimation

Here, we evaluate the path estimation accuracy of the proposed LoS/NLoS-aware solution. We record the moving trajectory with ground markers and calculate the ground truth. Thus, the estimation error can be obtained by comparing the ground truth and the estimated AoAs.

We compare SaTrack with SpotFi and Widar2.0 again. Fig. 11 plots the CDFs for AoA estimation error of the three algorithms. SaTrack has a median estimation error of 6.29 degrees, slightly outperforming SpotFi (7.85 degrees) and Widar2.0 (7.26 degrees). Meanwhile, SaTrack achieves 12.3 degrees for the 90th percentile estimation error, which obviously surpasses 20.17 degrees for SpotFi and 34.53 degrees for Widar2.0. Such phenomenon comes from the fact that SaTrack provides significant performance improvement for NLoS scenarios, as compared with other EM-based solutions.

E. Evaluation of state identification

1) Ratio of LoS/NLoS states: In order to demonstrate the existence of switching between LoS and NLoS states during a typical movement, we collect 8270 seconds of CSI data for 16 sets of movements in total, including linear or rectangular trajectories. The statistical results in Fig. 12 show that the Tx-Rx direct path keeps in LoS state for 5552 seconds (67.13%), and in NLoS state for the rest of time (32.87%). The experiment results again demonstrate that the NLoS state is non-negligible for indoor localization, and it is important to improve the localization accuracy in NLoS states.

2) Accuracy of State Identification: We conduct an analysis of the state recognition accuracy for the aforementioned 8270 seconds of data, with 91.89% accuracy for LoS signals, and 94.52% for NLoS signals. The reason for the lower performance of the former is that volunteer movements cause large fluctuations in the signal, which may occasionally be misclassified as NLoS signals. However, such state identification errors won't severely affect the overall localization accuracy due to the presence of path-following selection operations.

F. Parameter Study

1) Impact of Walking Velocity: To evaluate the impact of walking velocity, we record the ground truth on the ground with a marker and conduct four sets of comparison experiments with different walking velocities. Fig. 13 shows the distribution of tracking errors of volunteers. As shown in the figure, the tracking error increases progressively with increasing walking velocity, which compounds our intuition that faster velocity leads to greater fluctuations in the signal and estimated results. Meanwhile, the median tracking error of SaTrack is still around 0.7m, which indicates its robustness.

2) Impact of time slice length: SaTrack executes an EM-based algorithm periodically for path estimation, i.e., once for 0.1 time slice which means 50 CSI packets are used since our sampling rate is 500 Hz. Fig. 14 plots the tracking performance with different lengths of the time slice. Generally, longer time slices can reduce noise and improve performance, with median tracking errors of 0.78m, 0.49m, and 0.47m for 0.1s, 0.2s, and 0.3s slices, respectively. However, the results with 0.4s time slice in Fig. 14 indicate that an excessively long time slice might also cause performance losses. This is because the dynamic signals are relatively stable only within the coherence time [23]. If the length of the time slice is larger than the coherence time, the unstable signals will cause larger estimation errors for the EM-based algorithm.

3) Performance in long-term NLoS scenarios: The main concept behind SaTrack is to use accurate results in LoS states as a reference to improve performance in NLoS states. To evaluate its effectiveness in long-term NLoS scenarios, we conduct an experiment. The experiment involves a 5-second LoS phase followed by a 15-second NLoS phase, during which we measure the average localization error per second. As depicted in Fig. 15, during the initial 5 seconds in LoS states, the error remains relatively low (less than 1 meter). However, the error gradually increased upon transitioning to NLoS states, reaching approximately 2 meters by the 17th second and peaking at 3 meters by the 20th second.

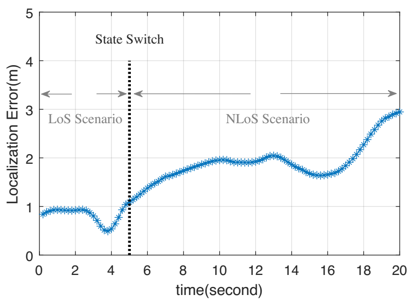


Fig. 15. Impact of prolonged NLoS state.

The results indicate that SaTrack's performance tends to worsen during prolonged NLoS states, but remains stable for around 10 seconds. Meanwhile, since extremely long-term NLoS scenarios are rare due to frequent switching between LoS and NLoS states for a typical human movement, our approach can ensure reliable performance in practical settings.

G. Computational complexity

It's worth noting that SaTrack has a longer computation time because it performs path estimation multiple times with different reference antennas. To evaluate its impact on real-time tracking, we experiment with the computational complexity.

We measure the computation overhead with different sampling frequencies and parameter L values (i.e., the number of possible paths in the environment). Table I compares the computational time of three algorithms for processing 1-second CSI data. SpotFi has the longest processing time due to its subspace-based matrix processing strategy. SaTrack's computation time is about three times that of Widar2.0, as our prototype has three antennas for Rx and performs path estimation once for each antenna acting as the reference one. However, we still find that SaTrack only requires 0.8886 seconds to process 1-second CSI data in a typical setup (with a sampling rate of 500 Hz and L = 5), demonstrating that SaTrack is capable of supporting real-time device tracking.

TABLE I
TIME CONSUMPTION OF 3 ALGORITHMS FOR 1-SECOND CSI DATA

	Solution	$L=2$	$L=5$	$L=8$	$L=10$
200 Hz	SaTrack	0.0532(s)	0.2406	0.5841	0.9812
	Widar2.0	0.0173	0.0792	0.1887	0.3161
	SpotFi	5.5014	12.6499	21.8051	27.2685
500 Hz	SaTrack	0.2305	0.8886	2.2135	3.7589
	Widar2.0	0.0725	0.2773	0.7136	1.2328
	SpotFi	12.7869	34.1197	54.5731	68.2186

Although SaTrack takes three times longer than Widar2.0 to compute, there are ways to reduce this duration. Our experiments show that accurate parameter estimation can be achieved even with L = 2 in LoS scenarios. Therefore, we start by conducting parameter estimation and LoS/NLoS state identification with L = 2. If the result indicates a LoS state, we can accept the path estimation parameters. Otherwise, we need to rerun the algorithm with L = 5 to obtain accurate estimation results in NLoS scenarios. With this strategy, the computational time is $t_{L=2} + \alpha_{NLoS} \times t_{L=5}$, where t represents computation time and α_{NLoS} represents the proportion of NLoS state. As shown in Fig. 12 and Table I, the expected computation time is approximately $0.2305 + 32.87\% \times 0.8886 = 0.5226$ seconds in a typical setup, which is obviously shorter than 0.8886 seconds when L is always set to 5.

V. Related Work

Currently, WiFi-based indoor localization techniques can be mainly categorized into the following two types:

1) Active localization: SpotFi [11] achieved a median localization error of 0.4m in LoS scenarios with 5 APs, which increased to 1.6m in NLoS scenarios, though. UbiLocate [16] applied an EM-based algorithm to extract wireless signal parameters and achieved a medium localization error of 0.3m in LoS scenarios with 4 APs, reduced to 1.1m in NLoS scenarios with 5 APs. However, UbiLocate relied on Multiple-Input Multiple-Output (MIMO) antenna systems and modification of device firmware to obtain fine time measurements (FTM), which is challenging to deploy on most WiFi devices. M^3 [15] used frequency hopping to increase bandwidth and improve ToF resolution, but it may affect wireless communication. NLoc [23] used the EM-based algorithm to extract fine-grained reflected path parameters and then calculated target positions using geometric principles, achieving decimeter-level localization accuracy with 4 to 5 APs. Furthermore, most of the aforementioned works primarily focus on the localization of static devices [28, 29, 30, 31] and pay less attention to tracking mobile devices, which is important in indoor localization.

2) Passive tracking: In recent years, some passive tracking technologies [17, 18, 32, 33] have been proposed to localize the target without carrying or connecting WiFi devices. They achieved impressive results by exploring various solutions to utilize the multi-dimensional path parameters extracted through EM-based algorithms.

However, passive tracking has evident limitations, such as limited tracking range and inability to support multi-user and complex environments, which limit its applications. Nevertheless, the methods for path parameter extraction employed in passive tracking still offer valuable insights for our research.

VI. Conclusion

In this paper, we proposed SaTrack, a LoS/NLoS state-aware WiFi indoor tracking system. SaTrack identified LoS/NLoS states by inspecting the distribution of the strongest estimated paths from different reference antennas. For NLoS states, SaTrack utilized spatial aggregation and temporal continuity for direct path selection, i.e., clustering in the spatial domain to form direct path candidates, and clustering in the temporal domain to select the final victor. Based on the selected direct path, SaTrack fulfilled mobile device tracking under the multi-AP deployment.

We prototyped SaTrack in COTS WiFi devices and conducted experiments in two different environments. The experiment results suggested that SaTrack achieves significant improvement in tracking accuracy as compared with other baselines, which also demonstrated the importance of LoS/NLoS state identification for indoor tracking system works.

VII. Acknowledgment

The research is partially supported by the Key-Area Research and Development Program of Guangdong Province with Grant No. 2020B0909050001, China National Natural Science Foundation with No. 62172379, Key Research Program of Frontier Sciences, CAS, Grant No.ZDBS-LY-JSC001. Hao Zhou and Xiangyang Li are the corresponding authors.

References

[1] Harlan Hile and Gaetano Borriello. Positioning and orientation in indoor environments using camera phones. *IEEE Computer Graphics and Applications*, 28(4):32–39, 2008.

[2] Antonio R Jimenez, Fernando Seco, Carlos Prieto, and Jorge Guevara. A comparison of pedestrian dead-reckoning algorithms using a low-cost mems imu. In *2009 IEEE International Symposium on Intelligent Signal Processing*, pages 37–42. IEEE, 2009.

[3] Sangki Yun, Yi-Chao Chen, Huihuang Zheng, Lili Qiu, and Wenguang Mao. Strata: Fine-grained acoustic-based device-free tracking. In *Proceedings of the 15th annual international conference on mobile systems, applications, and services*, pages 15–28, 2017.

[4] Loizos Kanaris, Akis Kokkinis, Antonio Liotta, and Stavros Stavrou. Fusing bluetooth beacon data with wi-fi radiomaps for improved indoor localization. *Sensors*, 17(4):812, 2017.

[5] Longfei Shangguan, Zimu Zhou, and Kyle Jamieson. Enabling gesture-based interactions with objects. In *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*, pages 239–251, 2017.

[6] Jue Wang and Dina Katabi. Dude, where's my card? rfid positioning that works with multipath and non-line of sight. In *Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM*, pages 51–62, 2013.

[7] Yasin Yelkovan, Hilal Güneren, Ahmet Akgöz, Furkan Eybek, Ahmet Türk, İsa Göl, and Ömer Çetin. Infrared beacon based sub-meter indoor localization. In *2014 22nd Signal Processing and Communications Applications Conference (SIU)*, pages 1427–1430. IEEE, 2014.

[8] Sinan Gezici, Zhi Tian, Georgios B Giannakis, Hisashi Kobayashi, Andreas F Molisch, H Vincent Poor, and Zafer Sahinoglu. Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks. *IEEE signal processing magazine*, 22(4):70–84, 2005.

[9] Fadel Adib, Zachary Kabelac, and Dina Katabi. Multi-person localization via rf body reflections. In *12th USENIX Symposium on Networked Systems Design and Implementation (NSDI 15)*, pages 279–292, 2015.

[10] Robert K Harle and Andy Hopper. Deploying and evaluating a location-aware system. In *Proceedings of the 3rd international conference on Mobile systems, applications, and services*, pages 219–232, 2005.

[11] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. Spotfi: Decimeter level localization using wifi. In *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, pages 269–282, 2015.

[12] Kun Qian, Chenshu Wu, Zimu Zhou, Yue Zheng, Zheng Yang, and Yunhao Liu. Inferring motion direction using commodity wi-fi for interactive exergames. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pages 1961–1972, 2017.

[13] Wei Gong and Jiangchuan Liu. Sifi: Pushing the limit of time-based wifi localization using a single commodity access point. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(1):1–21, 2018.

[14] Jon Gjengset, Jie Xiong, Graeme McPhillips, and Kyle Jamieson. Phaser: Enabling phased array signal processing on commodity wifi access points. In *Proceedings of the 20th annual international conference on Mobile computing and networking*, pages 153–164, 2014.

[15] Zhe Chen, Guorong Zhu, Sulei Wang, Yuedong Xu, Jie Xiong, Jin Zhao, Jun Luo, and Xin Wang. M3: Multipath assisted wi-fi localization with a single access point. *IEEE Transactions on Mobile Computing*, 20(2):588–602, 2019.

[16] Alejandro Blanco Pizarro, Joan Palacios Beltrán, Marco Cominelli, Francesco Gringoli, and Joerg Widmer. Accurate ubiquitous localization with off-the-shelf ieee 802.11 ac devices. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*, pages 241–254, 2021.

[17] Kun Qian, Chenshu Wu, Yi Zhang, Guidong Zhang, Zheng Yang, and Yunhao Liu. Widar2. 0: Passive human tracking with a single wi-fi link. In *Proceedings of the 16th annual international conference on mobile systems, applications, and services*, pages 350–361, 2018.

[18] Yaxiong Xie, Jie Xiong, Mo Li, and Kyle Jamieson. md-track: Leveraging multi-dimensionality for passive indoor wi-fi tracking. In *The 25th Annual International Conference on Mobile Computing and Networking*, pages 1–16, 2019.

[19] Swarun Kumar, Stephanie Gil, Dina Katabi, and Daniela Rus. Accurate indoor localization with zero start-up cost. In *Proceedings of the 20th annual international conference on Mobile computing and networking*, pages 483–494, 2014.

[20] Chenshu Wu, Zheng Yang, Zimu Zhou, Kun Qian, Yunhao Liu, and Mingyan Liu. Phaseu: Real-time los identification with wifi. In *2015 IEEE conference on computer communications (INFOCOM)*, pages 2038–2046. IEEE, 2015.

[21] Xiang Li, Daqing Zhang, Qin Lv, Jie Xiong, Shengjie Li, Yue Zhang, and Hong Mei. Indotrack: Device-free indoor human tracking with commodity wi-fi. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):1–22, 2017.

[22] Jie Xiong and Kyle Jamieson. Arraytrack: A fine-grained indoor location system. In *10th USENIX Symposium on Networked Systems Design and Implementation (NSDI 13)*, pages 71–84, 2013.

[23] Xianan Zhang, Lieke Chen, Mingjie Feng, and Tao Jiang. Toward reliable non-line-of-sight localization using multipath reflections. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(1):1–25, 2022.

[24] Harold W Kuhn. The hungarian method for the assignment problem. *Naval research logistics quarterly*, 2(1-2):83–97, 1955.

[25] Åke Björck. Least squares methods. *Handbook of numerical analysis*, 1:465–652, 1990.

[26] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. Tool release: Gathering 802.11 n traces with channel state information. *ACM SIGCOMM computer communication review*, 41(1):53–53, 2011.

[27] Kiran Joshi, Dinesh Bharadia, Manikanta Kotaru, and Sachin Katti. Wideo: Fine-grained device-free motion tracing using rf backscatter. In *12th USENIX Symposium on Networked Systems Design and Implementation (NSDI 15)*, pages 189–204, 2015.

[28] Jie Xiong, Karthikeyan Sundaresan, and Kyle Jamieson. Tonetrack: Leveraging frequency-agile radios for time-based indoor wireless localization. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, pages 537–549, 2015.

[29] Deepak Vasisht, Swarun Kumar, and Dina Katabi. Decimeter-level localization with a single wifi access point. In *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*, pages 165–178, 2016.

[30] Manikanta Kotaru and Sachin Katti. Position tracking for virtual reality using commodity wifi. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 68–78, 2017.

[31] Yaxiong Xie, Zhenjiang Li, and Mo Li. Precise power delay profiling with commodity wifi. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, pages 53–64, 2015.

[32] Xiang Li, Shengjie Li, Daqing Zhang, Jie Xiong, Yasha Wang, and Hong Mei. Dynamic-music: Accurate device-free indoor localization. In *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing*, pages 196–207, 2016.

[33] Ju Wang, Hongbo Jiang, Jie Xiong, Kyle Jamieson, Xiaojiang Chen, Dingyi Fang, and Binbin Xie. Lifs: Low human-effort, device-free localization with fine-grained subcarrier information. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, pages 243–256, 2016.