

Decimeter-Level Indoor Localization Using WiFi Round-Trip Phase and Factor Graph Optimization

Fangzhan Shi*, Wenda Li[†], Chong Tang*, Yuan Fang*, Paul Brennan[‡], Kevin Chetty*

*Department of Security and Crime Science, University College London, UK

[†]Department of Biomedical Engineering, University of Dundee, UK

[‡]Department of Electronic and Electrical Engineering, University College London, UK

fangzhan.shi.17@ucl.ac.uk, wli001@dundee.ac.uk, {chong.tang.18, yuan.fang.20, p.brennan, k.chetty}@ucl.ac.uk

Abstract—Indoor localization using WiFi signals has been studied since the emergence of WiFi communication. This paper presents a novel training-free approach to indoor localization using a customized WiFi protocol for data collection and a factor graph-based back-end for localization. The protocol measures the round-trip phase, which is very sensitive to small changes in displacement. This is because the sub-wavelength displacements introduce significant phase changes in WiFi signal. However, the phase cannot provide absolute range information due to angle wrap. Consequently, it can only be used for relative distance (displacement) measurement. By tracking the round-trip phase over time and unwrapping it, a relative distance measurement can be realized and achieve a mean absolute error (MAE) of 0.06m. For 2-D localization, factor graph optimization is applied to the round-trip phase measurements between the STA (station) and four APs (access points). Experiments show the proposed concept can offer a decimeter-level (0.26m MAE and 0.24m 50% CDF) performance for real-world indoor localization.

Index Terms—WiFi Localization, Round-trip Phase, Factor Graph Optimization

I. INTRODUCTION

THE emergence of location-based services (LBS) has profoundly shifted many facets of modern life, from navigating unfamiliar environments to logistics management. Global Positioning System (GPS) and Real-time Kinematic (RTK) have provided decimeter-level localization for outdoor applications, supporting various sectors and applications. However, the usefulness of GPS and RTK is greatly limited in the indoor environment due to signal attenuation and multipath effects. This has spurred the search for alternative solutions for indoor localization. Indoor spaces such as airports, shopping malls, hospitals and offices present unique challenges to navigation, asset tracking and space utilisation. An inaccurate or insufficient indoor localization can lead to sub-optimal user experiences and inefficiencies in operations (e.g. finding a specific parking space in a large underground parking area). Accurate indoor localization could significantly mitigate these issues, offering more user-friendly services. Likewise, accurate indoor localization extends to fields like indoor transportation (e.g. autonomous robots and vehicles), where positioning accuracy is the key factor for the task. Moreover, in emergencies, pinpointing an individual's exact location within a structure can be a life-saving tool for first responders.

Research in indoor positioning and localization using WiFi has attracted significant attention in recent years owing to its ubiquitous and pervasive nature. Existing WiFi-based localization systems are typically based on RSSI (received signal strength indicator), FTM (fine time measurement) / RTT (round-trip time) in 802.11mc [1] and CFR (channel frequency response). RSSI shows the signal power at the receiver, and curves can be fitted to present the relationship between range and RSSI [2] while RTT provides a meter-level accuracy in ranging [3] by measuring the time of flight directly. Then the localization can be accomplished by trilateration/multilateration using the range measurements between the user and several known devices. AoA (angle-of-arrival) can be estimated from CFR since it contains the phase of each subcarrier on each antenna in MIMO and then used for localization via triangulation/multi-angulation [4], [5] In real world-deployment, limitations apply to these methods. RSSI could vary up to 10dB error [2] for a given distance, leading to a significant error in ranging. For RTT, the accuracy is strictly limited by bandwidth and environment. According to evaluations in [3], the RTT ranging results could vary a few meters given a 10cm actual displacement in a typical indoor environment, which is insufficient for accurate localization. Using wider bandwidth could improve the performance, but it may introduce other issues like a high price, incompatibility with old devices, excessive power consumption and noncompliance with local radio regulations. For AoA, the accuracy depends on the number of antennas, and it also faces price, compatibility and power control issues if more antennas are adopted for better performance. Apart from that, it is not available on legacy or low-cost single-antenna WiFi devices. In addition to localization using range or angle estimation, radio fingerprinting also shows potential in indoor localization. All of RSSI, RTT and CFR can be used for radio fingerprinting, and since it does not utilize the range and angle directly, the limitation of theoretical accuracy discussed above can be bypassed. It generally consists of a training phase and a positioning phase. A database would be built based on the radio measurements at different locations in the training phase and used as a reference for localization during the positioning phase. For example, [6] divides the test area using 400 0.5cm×0.5cm squares that form a grid, then measures the CFR of WiFi packets from multiple WiFi APs on 3 Rx antennas at each square to build a fingerprint database.

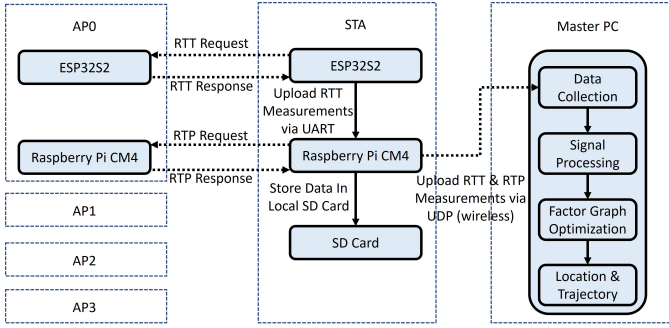


Fig. 1: System block diagram of proposed indoor localization system

For localization, the position is determined by the best match of the query CFR within the database. It demonstrates a localization accuracy of 1–2 cm, even in non-line-of-sight scenarios. This kind of fingerprinting-based method shows excellent performance in localization but may be impractical in real-world deployment. The 0.5cm-spacing grid in [6] requires a special mechanical structure to move the WiFi receiver to a specific location to collect data, and there will be 40,000 sample points per 1 m^2 , indicating a huge effort and cost to collect the data. To make matters worse, the database requires an update after the environment changes (e.g. several WiFi APs are moved), leading to high maintenance costs.

In this work, we present a novel method for indoor localization by tracking the WiFi round-trip phase (RTP) over time. It is achieved by using a customized WiFi protocol similar to Chronos [7] and using a factor graph optimization-based back-end for localization. Since measuring the phase can only provide relative distance due to angle wrap, a new factor graph model is developed to calculate the 2D location based on the RTP measurements between the STA and 4 APs. A testbed is built using commercial WiFi devices, and Fig. 1 illustrates the system. RTP is implemented on Raspberry Pi onboard WiFi chip where customized firmware runs the protocol. We also collect RTT data using ESP32 to compare the performance between RTT and RTP. Finally, the Raspberry Pi relays the collected data via UDP to a master PC that optimises the factor graph.

Contributions. By tracking WiFi RTP for localization, we have accomplished the following objectives for the indoor localization system:

- 1) Decimeter-level. Our localization system could reach decimeter-level accuracy such that it could support various applications, from indoor navigation to rescue missions in low-visibility scenarios like a fire.
- 2) Training-free. Our localization system only needs to know the location of the AP, and there is no requirement to collect the wireless fingerprints for training, reducing the complexity of real-world deployment. It makes the installation process more practical and easier for maintenance in the long run.
- 3) Compatible. Broadband and multi-antenna techniques for an improved position using WiFi signal have been standardized in IEEE 802.11az [8], but it may take a long

while to deploy them in the real world since new radio hardware is required to support it. On the contrary, we use 20MHz bandwidth and only one antenna to make our system compatible with various WiFi devices. This feature is not only required by legacy devices but also popular IoT devices such as ESP32 and Raspberry Pi with limited WiFi capabilities due to cost, size and power consumption considerations. Additionally, our system could be fused into future systems with range and angle measurements since it is relatively self-contained and can work independently.

- 4) Software only. Our method does not rely on modifying the radio hardware, which could reduce the risk of damage and make the deployment user-friendly.

The following paper is organized as follows: the related work is presented in Section II and we describe the principle of RTP and compare it with RTT in Section III. Our implementation and testbed are illustrated in Section IV. The experiment results for ranging are presented in Section V with the corresponding signal processing. Then the factor graph model for localization is detailed in Section VI. Finally, we conclude in Section VII and discuss future work.

II. RELATED WORK

In this section, we provide a review of the relevant literature and existing studies (Table I presents a selection of representative works in the field), offering a background of WiFi-based indoor localization. Following this, we discuss the baseline that underpins our work.

A. Localization Using WiFi RSSI

In [2] the authors fit a path loss model that presents the relationship between the RSSI and the distance. During localization, the distance is calculated from RSSI and then the trilateration is used for localization. RADAR [9] is one the first WiFi RSSI fingerprint-based localization work. 3 APs are set up, and the RSSI of the beacon signals from the APs are collected as signatures at the different sample locations. After the initialization of the signature database, K nearest neighbour (KNN) is used to search the location from the database given an RSSI measurement of the 3 APs at an unknown location. Horus [13] improves it by using a probabilistic model and builds a more extensive database for better fingerprinting. GIFT [14] uses RSSI gradient over the 2-D space as a map of signatures. Since this type of signature is more complicated, a particle filter is used to acquire the location. [15] treats RSSI signatures as a high-dimensional image and trains a CNN (convolution neural network) to learn the feature. However, the generalization of the aforementioned fingerprinting/learning approaches is not validated, and the model may be retrained if the environment changes.

B. Localization Using WiFi RTT

[10] calibrates RTT measurement to remove the bias and then uses the trilateration method to locate the STA. [12] uses a Kalman filter as a sensor fusion back-end that takes RTT

TABLE I: Typical Indoor Localization Work Using WiFi Signal

System	Device	Number Of AP	Measurement	Key Techniques	Localization Method	Accuracy (m)	Training Required
RADAR [9]	Lucent WiFi	3	RSSI	Fingerprint	KNN	2.9(50% CDF)	Yes
System in [10]	Smartphone	4	RTT	Sensor modelling	Trilateration	1.0(50% CDF)	No
DeepNar [11]	Smartphone	7	RTT	Fingerprint	Deep learning	0.62(50% CDF)	No
System in [12]	Smartphone	4	RSSI & RTT	Sensor modelling	Kalman filter & trilateration	0.35(50% CDF)	No
SpotFi [4]	Intel 5300	5	CSI & RSSI	AoA (using MIMO) & Sensor modelling	Quadratic optimization	0.6(50% CDF)	No
System in [6]	Not Specified	1	CSI	Fingerprint (frequency hopping over a 321MHz band, using MIMO)	Similarity matching	0.01(75% correlation with correct fingerprints)	Yes
Chronos [7]	Intel 5300	1	CSI (RTP)	Tracking RTP over frequency (frequency hopping over all 2GHz and 5GHz WiFi band, using MIMO)	Quadratic optimization	0.65(50% CDF)	No
Our system	Raspberry Pi	4	CSI (RTP)	Tracking RTP over time (no frequency hopping, no MIMO)	Factor graph optimization	0.24(50% CDF) 0.26 (mean error)	No

and RSSI as input. Before the Kalman filtering, the RTT is calibrated and the RSSI is converted to distance using a path loss model. The final step is to use trilateration for the location of the STA. Instead of trilateration, [16] builds a signature database using RTT and RSSI at different locations. A hand-crafted metric is used to evaluate the similarity between the signatures and map a given signature to a location. [17] presents a particle filter-based RTT localization using the Gaussian mixture model. DeepNar [11] develops an end-to-end neural network that maps the RTT measurement from multiple APs to the probabilities at different locations and then uses the probability as the weight to get the final estimation.

C. Localization Using WiFi CFR

PinLoc [18] builds a CFR database by collecting CSI data at different locations during initialization. Then a vote-based classification algorithm is developed to map the CSI as a signature to the localization. FILA [19] adopts a probabilistic model based on the correlation between the observed CSI and the fingerprint. CRISLoc [20] builds a reconstructable fingerprint database using transfer learning that can discard outdated signatures automatically. SpotFi [4] collected the CSI from a 3-antenna commercial WiFi card to calculate the AoA of the signal from the STA to the AP. MUSIC algorithm is used to improve the angular resolution and RSSI is also collected for coarse ranging. When multiple APs are utilized, the STA can be located using the bearing and range. Phaser [5] combines two 3-antenna WiFi cards to improve the AoA performance. For synchronization, one antenna is shared by both cards as a common reference such that there are 5 working antennas on the AP. DeepFi [21] trained a neural network similar to DeepNar [11] that maps the CSI data to location probability. To improve performance, [6] hops over different WiFi channels to form a wide-band sensing and shows a centimetre-level accuracy. It collects the CSI for a

321MHz band by frequency hopping and builds a fingerprint database with a centimetre-level localization resolution. It offers better performance than our system but is challenging to deploy in the real world. Occupying the 321MHz spectrum is not very practical for a WiFi system and collecting the WiFi fingerprint at the centimetre level is a costly and resource-intensive task.

We would like to highlight the Chronos project [7], a pioneering initiative that proposed the measurement of Wi-Fi RTP for localization and motivated our research. Chronos developed a custom round-trip Wi-Fi communication protocol for measuring Wi-Fi RTP across all available Wi-Fi channels in the 2.4GHz and 5GHz bands. This acted as a wide-band ranging solution using frequency hopping, allowing for tracking of RTP over frequency. For localization, Chronos adopts MIMO to measure the distance between various antenna pairs between the STA and AP, and framed the positioning problem as a quadratic optimization. However, their technique comes with some complications. Its use of all Wi-Fi channels can potentially lead to traffic collisions, and a full measurement duration of approximately 86ms doesn't cater well to robotic vehicles as they can move by a decimeter within this measurement window, thus reducing accuracy. Additionally, it hinges on the MIMO capability of radio hardware, which might not always be accessible. Despite these constraints, Chronos introduced the innovative method for Wi-Fi localization, which is to utilize the round-trip phase. It inspired us that accurate relative distance measurement can be acquired by tracking the round-trip phase over time and then processed for precise localization. This is primarily due to the phase's high sensitivity to displacement. Consequently, this approach can eliminate the need for frequency hopping and MIMO, substantially enhancing compatibility. However, it's crucial to note that since the measurement involves relative distance rather than absolute range information, the localization backend needs to

be re-engineered to accommodate this type of measurement.

D. Baseline Methodology

This study uses WiFi RTT-based trilateration/multilateration [10] localization as the baseline. The first reason is that our method shares very similar WiFi hardware specifications as RTT. Both ours and RTT can operate using a single antenna and a bandwidth of 20MHz, enhancing compatibility for real-world implementations. A crucial factor is the wide applicability of these specifications in current digital wireless communication protocols, meaning they can be generalized across a variety of systems (802.11n, 802.11a, 802.11ac, LTE and 5G). Given that these specifications are fundamental and compatible with more affordable hardware options, they have the capacity to utilize existing, cost-effective legacy hardware. This makes the implementation of the localization system not only straightforward but also budget-friendly. The second point of consideration is that RTT trilateration/multilateration localization relies on a geometric model, rather than utilizing a trained model derived from a fingerprint database. This characteristic aligns with our approach, which eliminates the need for training. Except for the method mentioned in [10], related works typically require multiple antennas and/or training. Consequently, it would not be equitable to use them as a benchmark for comparison.

III. MEASURING ROUND-TRIP PHASE

In this section, we present the method to measure the round-trip phase for relative distance measurement using WiFi signals. WiFi round-trip time is shown as the background first, and then we provide the details of the round-trip phase.

A. Background: Round-Trip Time

Round-trip communication can remove asynchronization by exchanging information. Fig. 2a shows the simplified workflow of RTT [1]. The STA sends a request and then the AP starts sending the response. The AP sends a response at $t1$ and arrives at STA at $t2$. Then, the AP sends an acknowledgement to the STA at $t3$ and arrives at the AP at $t4$. Finally, the AP sends $t1$ and $t4$ back to STA. Note that the time used above is the measured local time instead of the true time. Considering the offset between the true time and local time, $t1 = t1' + \Delta t_{STA}$, $t2 = t2' + \Delta t_{AP}$, $t3 = t3' + \Delta t_{AP}$, $t4 = t4' + \Delta t_{STA}$ where $t1'$, $t2'$, $t3'$, $t4'$ are the true timestamps and Δt_{STA} and Δt_{AP} are the time offset. The offset would be cancelled in round-trip time estimation as $t4 - t3 + t2 - t1 = t4' + \Delta t_{STA} - (t3' + \Delta t_{AP}) + t2' + \Delta t_{AP} - (t1' + \Delta t_{STA}) = t4' - t3' + t2' - t1'$. Consequently, the distance between the STA and AP would be $(t4 - t3 + t2 - t1)/2 \times c$ where c is the speed of light.

The RTT/FTM would be performed multiple times in a burst, and the average value is used as the final result. However, the bandwidth and multi-path effect still limit the RTT ranging accuracy in the real world. Radio-based ranging is strictly dependent on bandwidth [22], [23] and it is verified by [3] that 80MHz bandwidth outperforms 20MHz significantly.

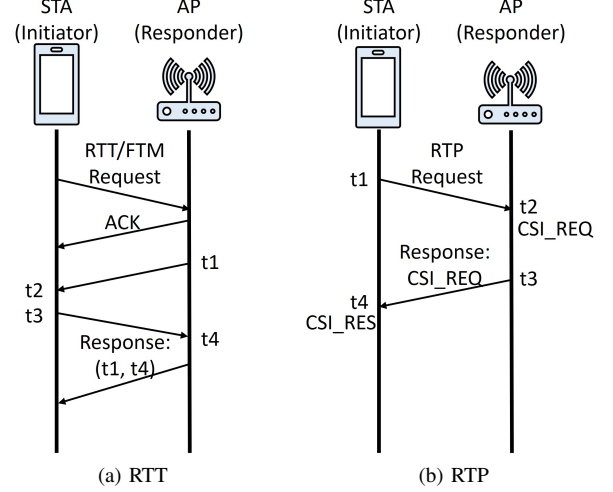


Fig. 2: RTT and RTP workflow

Additionally, [3] finds that the RTT ranging is not stable in the indoor area, varying up to 5m throughout the experiment with 20/40MHz bandwidth. [24] evaluates low-cost ESP32S2 chips for RTT ranging and finds that the error is up to 20m sometimes, and 75% of the errors are below 5m when using 20MHz bandwidth. The evaluations above indicate that RTT is an effective way for ranging, but the error is very large when using a relatively narrow bandwidth (20MHz) in an indoor scenario.

B. Details Of Round-Trip Phase

Instead of measuring time, we measure the RTP using a customized protocol. WiFi RTP was first proposed by Chronos [7] that the RTP is measured in all available WiFi channels in the 2GHz band and 5GHz band to simulate a wide-band ranging. Rather than tracking phase over RF frequency as Chronos [7], our work focuses on tracking phase over time for relative distance measurement and Fig. 2b also shows the workflow of RTP. The protocol shares a very similar structure to RTT, but the payload in the protocol is CSI from which the carrier phase can be inferred. CSI describes the wireless channel frequency response using complex numbers so that amplitude response and phase response can be presented simultaneously [25]. It is calculated from the fixed preambles during channel estimation, which the WiFi baseband processor operates. In the scenario that there is no asynchronization between the transmitter and receiver, the CSI would be

$$H(i, t) = \sum_{k=1}^N a_k(i, t) e^{-j2\pi(f_c + f_i)\tau_k(t)} \quad (1)$$

where t is the time, i is the subcarrier index, f_c is the RF centre frequency, f_i is the baseband frequency of subcarrier i , a_k is the attenuation and initial phase factor of path k presented in complex number, τ_k is the propagation delay of path k and N is the number of paths.

Considering inevitable asynchronization in real-world devices, the measured CSI is shown in equation (2) [25], [26].

$$\hat{H}(i, t) = s(t)H(i, t) \exp(j(\beta(t)i + \omega_{cfo}t + \theta(t) + \theta_{nl}(t, i))) \quad (2)$$

where $\hat{H}$ is the measured CSI, s is the scaling factor, j is the imaginary unit, β is the linear phase error factor, ω_{cfo} is the carrier frequency offset, θ is the phase offset error, θ_{nl} is the non-linear phase error.

Note that $\beta(t)$ can be interpreted as a time delay term in the receiver baseband. Thus, it adds a linear phase error with respect to subcarrier frequency similar to the propagation delay. In the real-world scenario, the true propagation delay is unknown, and $\beta(t)$ is time-varying and depends on propagation delay, signal detection delay and clock asynchronization, making it difficult to estimate. This effect would add unknown phase errors to every subcarrier, leading to non-negligible uncertainty in phase measurements. The only exception is the subcarrier 0, because the linear phase error is $\beta(t)i$, which would be 0 given $i = 0$. However, subcarrier 0 is not used in actual communications. The major reason is that the baseband frequency of subcarrier 0 is 0 Hz which would be mixed with inevitable DC offset/bias in radio hardware. Fortunately, the phase of subcarrier 0 can be interpolated from adjacent subcarriers even if the exact value of $\beta(t)$ is not given. For subcarrier i and subcarrier $-i$ ($i \neq 0$), the corresponding phase error would be $\beta(t)i$ and $-\beta(t)i$, whose sum is 0. It indicates that the sum of phases of subcarrier i and $-i$ has cancelled the effect from $\beta(t)$. Consequently, by using a few pairs of symmetrical subcarriers and finding the mean phase, subcarrier 0 can be estimated. In our implementation, only subcarrier $-5, -4, -3, -2, -1, 1, 2, 3, 4, 5$ are used, and since they are very close to subcarrier 0, the non-linear effect term $\theta_{nl}(t, i)$ is ignored. As a result, the CSI would be the following equation for subcarrier 0.

$$\hat{H}(0, t) = s(t)H(0, t) \exp(j(\omega_{cfo}t + \theta(t))) \quad (3)$$

To remove the phase offset from ω_{cfo} , round-trip communication takes effect. We can model that $\omega_{cfo}^{AP} = -\omega_{cfo}^{STA}$ since the carrier frequency offset would be opposite for the STA-To-AP packet and AP-To-STA packet. As a result, the product of those two CSI at subcarrier 0 would be

$$\hat{H}^{RT}(t2, t4) \quad (4)$$

$$= \hat{H}^{AP}(0, t2) \hat{H}^{STA}(0, t4) \quad (5)$$

$$= s^{RT}(t2, t4) \quad (6)$$

$$\times H^2(0, t4) \quad (7)$$

$$\times \exp(j\omega_{cfo}(t4 - t2)) \quad (8)$$

$$\times \exp(j(\theta^{RT}(t2, t4))) \quad (9)$$

where

$\hat{H}^{RT}$ is round-trip CSI (RTCSI),

$\hat{H}^{AP}$ is the measured CSI on the AP (CSI of the request, CSI_REQ),

$\hat{H}^{STA}$ is the measured CSI on the STA (CSI of the response, CSI_RES),

s^{RT} is the round-trip scaling factor,

θ^{RT} is the round-trip phase offset, which is caused by regular CSI phase offset $\theta(t)$.

s^{RT} and θ^{RT} come from the frequency response of the radio hardware and can be treated as constant over time. Also, ω_{cfo} and $(t4 - t2)$ are modelled stable over time, as a result, equation (8) is also a constant phase offset. In fact, $(t4 - t2)$ would change with $(t4 - t3)$, the propagation time, which would vary up to 0.1us, assuming the system works within a range of 30m. This timing error is applied to the carrier frequency offset that is generally tens of kilohertz (e.g. the frequency offset would be 58.25KHz given 5.825GHz carrier frequency and 10PPM frequency offset). In this case, the phase error introduced by unstable $(t4 - t3)$ is minimal (e.g. $0.1us \times 58.25KHz \times 2\pi = 0.037rad$) and can be neglected. Consequently, as long as $(t3 - t2)$, the baseband processing time, is stable, $(t4 - t2)$ can be assumed to be a constant value. $H^2(0, t4)$ is the true round-trip channel response and it is also the same as $H^2(0, t2)$ assuming that the channel response does not change between a $t2$ and $t4$ since the gap is very short (a few milliseconds). By tracking $\hat{H}^{RT}$ over time, we can monitor how $H^2(0, t)$ changes over time. Note that subcarrier 0 is not used in WiFi protocol, so it is interpolated from the adjacent subcarriers.

There are a few limitations in mapping the round-trip phase (phase of RTCSI, $\hat{H}^{RT}$) to the true phase. The phase would wrap by every 2π making it impossible to measure the absolute phase. As a result, it can only provide relative phase measurement, which is the displacement between two locations. Note that the derivative of phase w.r.t. time is frequency. If the STA is moving, it would be the instantaneous Doppler frequency shift. To track the phase over time, the minimum sampling rate must be higher than twice the instantaneous round-trip Doppler frequency shift, indicating the RTP measurement has to be performed frequently. Finally, equation (10) can be used to convert the unwrapped phase to an estimated distance, which is the true distance with an unknown offset.

$$d_{rtp}(t4) = -\frac{1}{2} \frac{\angle \hat{H}^{RT}(t2, t4)}{2\pi} \times \lambda = d_{true}(t4) + d_{offset} \quad (10)$$

where λ is the wavelength, $\angle \hat{H}^{RT}$ is the unwrapped angle of RTCSI, d_{true} is the true distance and d_{offset} is the unknown offset.

IV. IMPLEMENTATION AND TESTBED

Due to hardware limitations, RTT and RTP are implemented on different radio devices.

A. ESP32-Based RTT

ESP32 is a microcontroller unit (MCU) with an integrated WiFi function. It has been adopted for numerous internet of things (IoT) applications and ESP32S2 has a built-in function for WiFi RTT measurement. By calling the API from the manufacturer, RTT can be utilized. Additionally, there is a UART port on the development board so it can communicate with other devices.

B. Raspberry Pi-based RTP

Raspberry Pi is a popular microcomputer, and there is much research focusing on its on-board WiFi chip. Nexmon [27] provides a way to patch the WiFi firmware and the Nexmon CSI extractor [28] offers us access to the CSI of the WiFi packet. Using Nexmon and Nexmon CSI extractor, a customized wireless communication protocol is developed on Raspberry Pi CM4 to measure the WiFi RTP as described in Section.III.

C. Ground Truth

To collect the ground truth for localization, IOS ARKIT is used. It runs on an iPhone and is designed for AR applications by APPLE. Since motion tracking is crucial in AR applications, IOS ARKIT has a built-in real-time localization system that uses the sensors on the phone, such as IMU, camera and laser radar. By logging the localization result reported by IOS ARKIT, an indoor localization ground truth can be acquired. Two major strategies are used in our ARKIT-based ground truth system for optimal performance. The first is to add visual landmarks to the environment by attaching texture-rich posters to the wall and the floor such that the camera can capture more reference points for the internal localization system. The second is to scan the whole room multiple times allowing localization system build a map first, then start the tracking for experiments. By checking the localization result with the BOSCH GLM50C laser distance meter, the ARKIT system running on an iPhone 12 with IOS 13.1 can provide a centimeter-level localization.

D. System Integration

By attaching the ESP32 to the Raspberry Pi, RTT and RTP can work simultaneously (in different WiFi channels ¹). The ESP32 connects to the Raspberry Pi via the USB cable then the Raspberry Pi process and saves the uploaded RTT measurements from ESP32 using UART. For RTP measurements, the Raspberry Pi uses the onboard WiFi chip with the customized protocol and logs data. For the STA, an additional iPhone is mounted on top of the Raspberry Pi chassis to collect the ground truth. An NTP (network time protocol) server is set up in the local area network to synchronize the Raspberry Pi and the iPhone such that ground truth can share the same timestamp as Raspberry Pi. Fig. 3 shows the whole device.

V. EVALUATION FOR RTP-BASED RANGING

A. Multi-path Effect On RTP

A simulation is performed to evaluate the RTP performance in a strong multi-path scenario. Since the reflected signal can be modelled as the signal transmitted from a mirror of the original signal source with respect to the reflection interface, a few signal source mirrors are placed in the simulation as Fig. 4. Additionally, the trajectory of the STA is planned as a circle whose radius is 3m, and the linear velocity is 0.94m/s. The

¹Due to hardware limitations, ESP32 RTT can only work in 2.4G WiFi band and Raspberry Pi RTP can only work in 5G WiFi band.

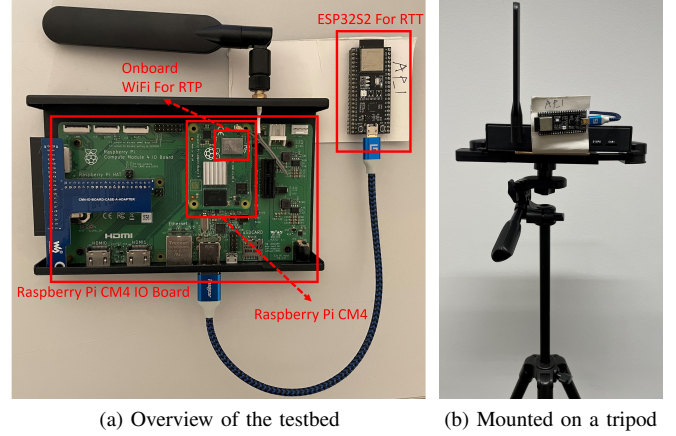


Fig. 3: The experimental testbed consisting of a Raspberry Pi and ESP32

propagation model is shown in equation (11), which consists of a power loss according to [29] and a phase shift due to the time delay caused by propagation. The final CSI collected would be $H_{sim} = \sum_i H_{sim_i}$ and then the RTCSI can be calculated as H_{sim}^2 . The phase of the RTCSI, H_{sim}^2 , is the RTP and we also plot the amplitude of the RTCSI in this simulation.

$$H_{sim_i} = \alpha_i \frac{1}{d_i} \exp(-2\pi j \frac{d_i}{\lambda}) \quad (11)$$

where α_i is the amplitude factor and d_i is distance between the STA and the signal source i . Three simulation results are shown with parameters in Table II. During the simulation, signal source 1 is the original signal source while others are the mirror signal sources due to reflections. The first simulation only uses signal source 1 and signal source 2 with amplitude factors of 1.0 and 0.5, respectively. It simulates that the line-of-sight path conveys higher energy than the reflection path, which is an acceptable assumption for indoor wireless communication. Fig. 5a shows the phase and amplitude of RTCSI and both fluctuate around the true values, which are the RTCSI using source 1 only. The fluctuation originates in the interference between the signals from the line-of-sight-path and reflection path. For the amplitude, since there are two coherent signal sources, the amplitude of the combined signal would change with location due to the constructive and destructive interference patterns. The signal amplitude would be higher at a constructive interference point, and vice versa. For the phase, as the line-of-sight path convey the major proportion of the energy, the phase deviation due to the multi-path is smaller and the line-of-sight-path signal takes the dominant role. As a result, the phase fluctuates about the true RTP. Fig. 5b shows the phase and amplitude when the reflection path takes the major proportion of the energy and it is clear that there are obvious errors. It simulates the case that the line-of-sight path is blocked and the other paths take the major effect. In this scenario, the RTCSI cannot work as expected, and it is a common problem in wireless localization. For example, the WiFi RTT ranging values would be larger [24] and the GPS also suffers from errors if urban infrastructures such as high-

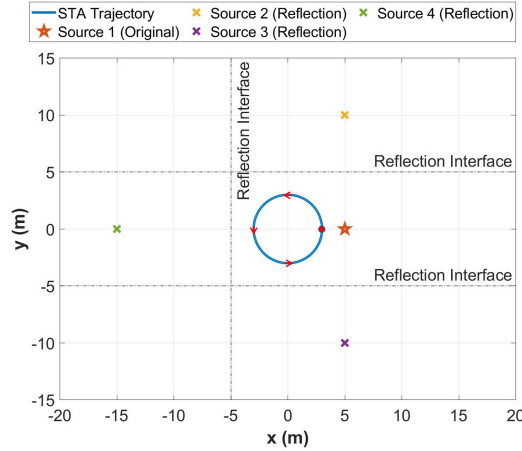


Fig. 4: Map for simulation. One original signal source (source 1) and 3 reflection signal sources (mirrors).

rise buildings block the direct signal from the satellites to the receiver [30]. The third simulation uses all 4 signal sources and Fig. 5c presents the result that the RTP shows a small error and the measurement is still valid. These simulations indicate that the RTP only works when the line-of-sight path conveys the major proportion of the power, or it will track a different path then lead to errors.

TABLE II: Parameters for RTCSI multipath simulation

Parameter	α_1	α_2	α_3	α_4
Simulation 1	1.0	0.5	0.0	0.0
Simulation 2	0.5	1.0	0.0	0.0
Simulation 3	1.0	0.5	0.5	0.5

B. Real-world Experiment Setup

The experiment is performed in a 6m x 6m meeting room as illustrated in Fig. 6. One STA and one AP (only AP0) are used. The STA leaves and approaches the AP periodically while collecting RTP, RTT and the ground truth. The Raspberry Pi measures RTP at 400Hz and the ESP32 measures RTT at 10Hz.

C. Signal Processing

1) *RTP*: Fig. 7 shows the procedure to get the RTP. The raw round-trip CSI is first calculated as equation (5) based on the CSI for the request and the corresponding response. A low-pass filter is then applied to the round-trip CSI to remove high-frequency noise. Then the phase of the round-trip CSI is extracted, followed by the next step to unwrap the phase over time. Finally, the phase is sub-sampled to 10Hz, which is the global sampling rate in the whole system.

2) *RTT*: Fig. 8 shows the procedure to get the RTT. The raw RTT reported from ESP32 is filtered by a moving mean filter and then sub-sampled to 10Hz for the sampling rate.

D. Line-of-sight Channel

One AP is installed on a tripod and one STA is mounted on a robot vehicle. The STA and the AP maintain a line-of-sight wireless channel during the experiment. The STA collects the RTP ranging data and RTP ranging data simultaneously while the robot is moving and the result is shown in Fig. 9a with the ground truth. Note that since the RTP cannot provide absolute range information, it is shifted vertically to align with the ground truth at the beginning. The RTP ranging curve exhibits a close correlation with the ground truth data, while the RTT ranging shows evident errors. Fig. 9b and Fig. 9c show the distributions of the errors and the cumulative distribution functions (CDF) of the absolute error, respectively. It is clear that the RTP ranging outperforms the RTT ranging that the MAE are 0.06m and 1.08m, respectively. This experiment demonstrates that our RTP system can reach centimeter-level ranging in the line-of-sight channel and agree with the predicted simulation results that this system can work in the real world as long as there exists a line-of-sight signal. Note that even if RTT is performed in a line-of-sight channel, the ranging result is generally above the ground truth. The error may come from hardware calibration and the inevitable multi-path in the real world.

E. Non-line-of-sight Channel

The line-of-sight channel cannot always be guaranteed in the real world and in this work we define the non-line-of-sight channel as the line-of-sight channel being blocked by radio-reflective material such as a metallic object resulting in multipath becoming dominant. We set up the experiment environment as in the previous section but an aluminium sheet is placed in front of the antenna on the AP to block the line-of-sight signal. Fig. 9d shows the ranging result and Fig. 9e and Fig. 9f show the errors. It presents that the RTT ranging is very likely to be higher than the ground truth but RTP ranging is always lower than the ground truth. These effects come from multi-path and signal reflection. Fig. 10 illustrates it with a simplified case that the line-of-sight path is blocked and the signal has to be reflected by the wall. The wall acts as a mirror and AP' can be treated as the actual location of the AP. For RTT ranging, it would be larger since the actual path would always be longer. When performing RTP ranging, the relative distance from STA(t_0) to STA(t_1) w.r.t. the AP' changes 0.8m meters while the ground truth is 1m. Consequently, the RTP measurement is lower than the ground truth since it only measures the changes in the propagation distance.

VI. EVALUATION FOR RTP-BASED LOCALIZATION

A. Setup

4 APs are set and mounted on tripods and the location is shown in Fig. 6. The STA is installed on the robot vehicle and it moves in the free space around the centre of the room. 100 RTP measurements and 10 RTT measurements are collected per second for each link between the STA and APs. The same signal processing method is applied as Section.V-C. Due to

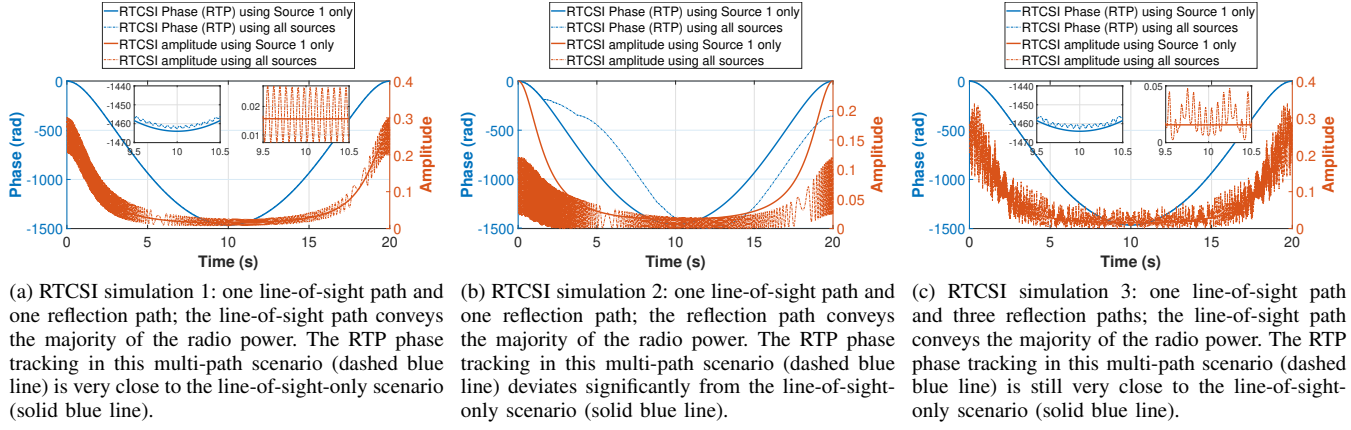


Fig. 5: RTCSI simulation in multi-path scenario

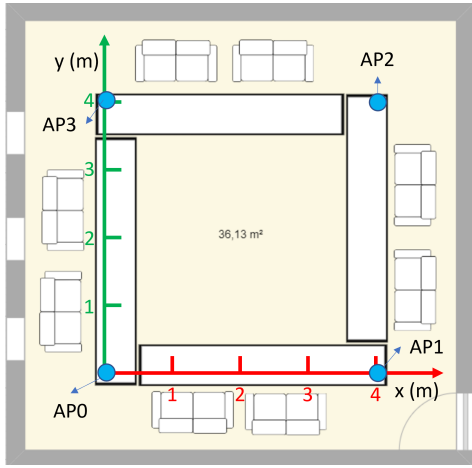


Fig. 6: Floor plan for the experiment. The meeting room is 6m x 6m and the experiment area is the 4m x 4m square in the centre, which is enclosed by four APs.

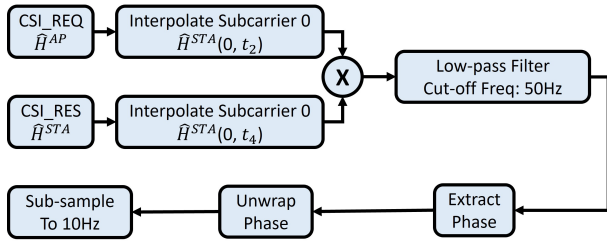


Fig. 7: RTP signal processing procedure

the limited computational resources on the Raspberry Pi, the measurements are forwarded to a master PC that runs the localization back-end.

B. RTP Modelling

According to observations in Section.V, the multipath effect has to be taken into consideration and the measurement model for RTP ranging becomes

$$d_{rtp}(i, t) = \alpha_{rtp}(i, t)dist(i, t) + \beta_{rtp}(i, t) + n_{rtp}(i, t) \quad (12)$$

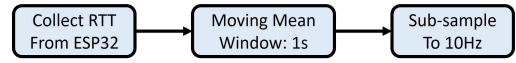


Fig. 8: RTT signal processing procedure

where

d_{rtp} is the RTP ranging result,

i is the index of AP,

t is the time,

$dist$ is the true distance from the STA to the AP,

α_{rtp} is the scaling factor,

β_{rtp} is the offset factor,

n_{rtp} is the additive white (zero-mean) Gaussian noise with variance σ_{rtp}^2 .

and

$$dist(i, t) = \left\| \begin{bmatrix} x_{sta}(t) \\ y_{sta}(t) \\ z_{sta}(t) \end{bmatrix} - \begin{bmatrix} x_{AP_i} \\ y_{AP_i} \\ z_{AP_i} \end{bmatrix} \right\| \quad (13)$$

where

$\| \cdot \|$ is the vector norm,

x_{sta} is the x-axis coordinate of the STA (similar for y_{sta} and z_{sta}),

x_{AP_i} is the x-axis coordinate of the AP i (similar for y_{AP_i} and z_{AP_i}).

It is a first-order approximation and the offset factor consists of not only the multipath effect but also the unknown initial distance since the RTP is based on relative distance. This prototype works in a 2-D space, so the location is modelled as $(x_{sta}(t), y_{sta}(t))$ and z_{sta} is set to a fixed value. Since the WiFi round-trip communication can be completed within a few milliseconds and considering a typical indoor application that the robot vehicle cannot move very fast, a round (4 measurements between the STA to each AP) of RTP measurements can be treated as occurring at the same time t . In this way, the status of the STA at time t can be presented as the following equation, and it has 10 degrees of freedom.

$$x(t) = \{x_{sta}(t), y_{sta}(t), \alpha_{rtp}(0, t), \beta_{rtp}(0, t), \alpha_{rtp}(1, t), \beta_{rtp}(1, t), \alpha_{rtp}(2, t), \beta_{rtp}(2, t), \alpha_{rtp}(3, t), \beta_{rtp}(3, t)\} \quad (14)$$

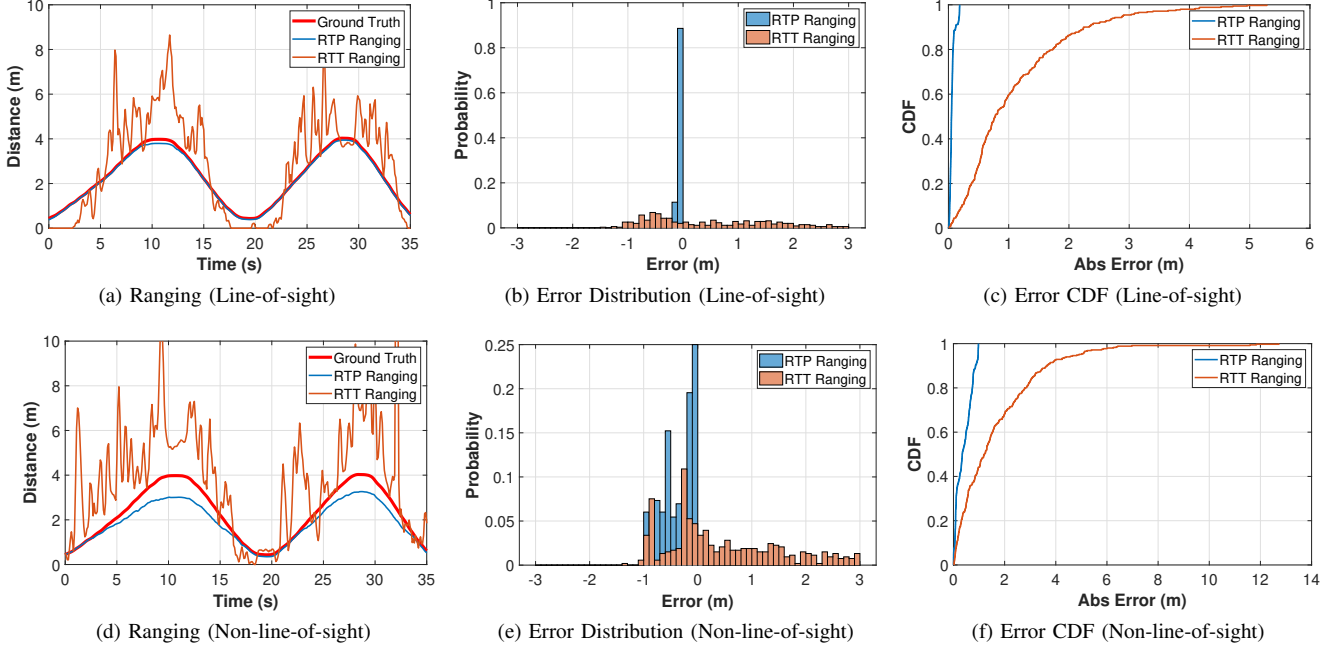


Fig. 9: Ranging evaluation for RTT and RTP

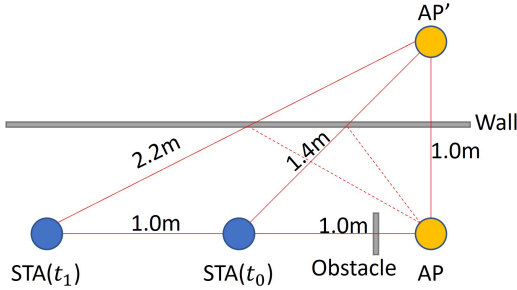


Fig. 10: Non-line-of-sight scenario example

where x is the STA status.

However, at time t , only 4 measurements are available, which means it is underdetermined. Two constraints over time are applied to reduce the underdetermination and improve the localization.

The first is that the scaling factor and the offset factor are constant within a short time window T_w . That is

$$\alpha_{rtp}(i, t) = \alpha_{rtp}(i, t_0), \forall t \in [t_0, t_0 + T_w] \quad (15)$$

$$\beta_{rtp}(i, t) = \beta_{rtp}(i, t_0), \forall t \in [t_0, t_0 + T_w] \quad (16)$$

The second is that the scaling factor and the offset factor change slowly over time and it is modelled as a random walk. The derivative of the values w.r.t time is assumed to follow zero-mean Gaussian distribution. Since the values have been windowed, the following equations show the actual model used in our system.

$$(\alpha_{rtp}(i, t + T_w) - \alpha_{rtp}(i, t)) \sim \mathcal{N}(0, \sigma_{rtp_\alpha}^2). \quad (17)$$

$$(\beta_{rtp}(i, t + T_w) - \beta_{rtp}(i, t)) \sim \mathcal{N}(0, \sigma_{rtp_\beta}^2). \quad (18)$$

C. Localization Backend: Factor Graph Optimization

As illustrated in Fig. 9c, RTP-based ranging considerably surpasses RTT, thus indicating a substantial potential for localization. However, due to the inherent periodicity of the radio wave, it is not feasible to estimate absolute range from RTP. The phase can only delineate the differential distances between the STA and AP at two distinct points in time during tracking, necessitating long-term memory for effective localization. To provide a clear and simple example, consider a scenario comprising two distinct APs situated at different locations, with one STA orbiting the first AP. The RTP tracking between the STA and the first AP would remain stable since the distance remains unchanged during the orbit. Conversely, the phase tracking for the second AP would exhibit a cyclical change. While these patterns are insightful, they can only be processed as a temporal sequence over an extended period. A traditional method such as the Kalman filter [31] cannot satisfy this requirement as it loses past measurements, with the memory retained solely in the covariance matrix. Similar problems exist in SLAM (simultaneous localization and mapping) in robotics, and research shows factor graph optimization is an effective approach [32]. It can model a variety of estimation problems using optimization. Compared to classic localization methods like the Kalman filter, the factor graph can keep all previous measurements for optimization instead of discarding previous measurements after updating the system status. Additionally, the factor graph also supports a more flexible way of organizing customized measurements and status. The STA status, $x(t)$, should be estimated for localization in this study and the whole system can be modelled as a conditional probability which is expressed in equation (19).

$$\begin{aligned}
P(X|Z) = & \\
& P(x(t_0), x(t_0 + \Delta T), x(t_0 + 2\Delta T), \dots \\
& x(t_0 + T_w), x(t_0 + T_w + \Delta T), x(t_0 + T_w + 2\Delta T), \dots \\
& |z_{rtp}(t_0), z_{rtp}(t_0 + \Delta T), z_{rtp}(t_0 + 2\Delta T), \dots \\
& z_{rtp}(t_0 + T_w), z_{rtp}(t_0 + T_w + \Delta T), z_{rtp}(t_0 + T_w + 2\Delta T), \dots \\
& z_{rtp\alpha}(t_0, t_0 + T_w), \dots \\
& z_{rtp\beta}(t_0, t_0 + T_w), \dots) \quad (19)
\end{aligned}$$

P is the probability,

X is the statuses,

Z is measurements and constraints (conditions),

ΔT is the sampling period of the measurement,

$z_{rtp}(t_0)$ the RTP ranging result from the AP at time t_0 modelled with scaling and offset factor (equation (12)),

$z_{rtp\alpha}(t_0, t_0 + T_w)$ is the RTP scaling factor constraint. The difference between α_{rtp} at two consecutive time windows, starting from t_0 and $t_0 + T_w$ respectively, follows a Gaussian distribution (equation (17)),

$z_{rtp\beta}(t_0, t_0 + T_w)$ is the RTP offset factor constraint. The difference between β_{rtp} at two consecutive time windows, starting from t_0 and $t_0 + T_w$ respectively, follows a Gaussian distribution (equation (18)).

The optimal X can be estimated using maximum posteriori probability estimation as equation (20). After applying Bayes' Rule, it would be equation (21).

$$\hat{X} = \arg \max_X P(X|Z) \quad (20)$$

$$= \arg \max_X \frac{P(Z|X)P(X)}{P(Z)} = \arg \max_X P(Z|X)P(X) \quad (21)$$

$P(X)$ is the priori and $P(Z|X)$ comes from the measurement model. Take equation (12) for example,

$$\begin{aligned}
P(z_{rtp}(i, t)|x(i, t)) &= \frac{1}{\sigma_{rtp^i} \sqrt{2\pi}} \\
&\times \exp\left(-\frac{1}{2} \left(\frac{d_{rtp}(i, t) - (\alpha_{rtp}(i, t) \text{dist}(i, t) + \beta_{rtp}(i, t))}{2\sigma_{rtp^i}} \right)^2 \right) \quad (22)
\end{aligned}$$

For equation (17) and equation (18), the probability would be equation (23) and equation (24), respectively.

$$\begin{aligned}
&P(z_{rtp\alpha}(i, t)|x(i, t), x(i, t + T_w)) \\
&= \frac{1}{\sigma_{rtp\alpha^i} \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\alpha_{rtp}(i, t) - \alpha_{rtp}(i, t + T_w)}{2\sigma_{rtp\alpha^i}} \right)^2 \right) \quad (23)
\end{aligned}$$

$$\begin{aligned}
&P(z_{rtp\beta}(i, t)|x(i, t), x(i, t + T_w)) \\
&= \frac{1}{\sigma_{rtp\beta^i} \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\beta_{rtp}(i, t) - \beta_{rtp}(i, t + T_w)}{2\sigma_{rtp\beta^i}} \right)^2 \right) \quad (24)
\end{aligned}$$

This process can be visually represented as a factor graph, as shown in Fig. 11. At each moment t , a state $x(t)$ is incorporated into the graph. During this phase, the location and parameters (α_{rtp} and β_{rtp}) within $x(t)$ are yet to be

determined. As such, they do not form part of the localization system's input. Upon obtaining the distance estimates from the most recent RTP tracking, these measurements are gathered and employed as system inputs to establish the RTP measurement factor within the graph, depicted as blue lines in Fig. 11. Throughout each time window, T_w , the scaling factor α_{rtp} and offset β_{rtp} are held constant to ensure their consistency. At the start of each time window, T_w , the scaling α_{rtp} and offset β_{rtp} constraint factors are introduced into the graph to fulfill the constraint of low variance over time. These are illustrated by the yellow lines in Fig. 11. After constructing the graph, the unknown status can be estimated by finding the optimal X that maximizes the probability as equation 21. We use GTSAM [32], a factor graph-based sensor fusion library, to run the optimization with our customized RTP factor. Internally, GTSAM uses a loss function $L_Z(X) = -\log P(X|Z)$ such that

$$\hat{X} = \arg \max_X P(X|Z) = \arg \min_X L_Z(X) \quad (25)$$

where $\hat{X}$ is the final estimated value. The logarithm can simplify the exponential calculation in the probability model and it becomes a non-linear optimization problem to find the optimal X to minimize the loss function $L_Z(X)$. Finally, Levenberg-Marquardt method [33], [34] is used to find the optimal value.

D. RTT Modelling

Similar to RTP, RTT ranging is modelled as follows.

$$d_{rtt}(i, t) = \alpha_{rtt}(i, t) \text{dist}(i, t) + \beta_{rtt}(i, t) + n_{rtt}(i, t) \quad (26)$$

where

d_{rtt} is the measured RTT ranging,

α_{rtt} is the scaling factor,

β_{rtt} is the offset factor,

n_{rtt} is the additive white Gaussian noise with variance $\sigma_{rtt^i}^2$.

Also, similar factor graph optimization can be applied as RTT. During factor graph construction, only the RTT measurements are used as input, and the unknown localization, together with the unknown parameters, are presented in the status. Then, the same optimization is performed as RTP via GTSAM to search for the optimal status to get the location. Note that there is only one key difference between RTT and RTP, the priori probability of the offset factor. For RTP, the offset factor also works for the initial phase of the relative distance measurement, and it means the RTP offset factor may vary a lot since it is related to the true distance. As a result, there is no specific priori for the RTP offset factor, which corresponds to a very large variance. However, RTT ranging can provide absolute distance measurement, so adding a low-variance priori for the RTT offset factor makes sense. The RTT offset factor is assumed to be drawn from of Gaussian distribution estimated from calibration, that is

$$\beta_{rtt}(i, t) \sim \mathcal{N}(u_{rtt\beta^i}, \sigma_{rtt\beta^i}^2) \quad (27)$$

It essentially formulates the same localization system as [10] that the scaling and offset values are used prior knowledge and the localization is reached by minimizing a cost function.

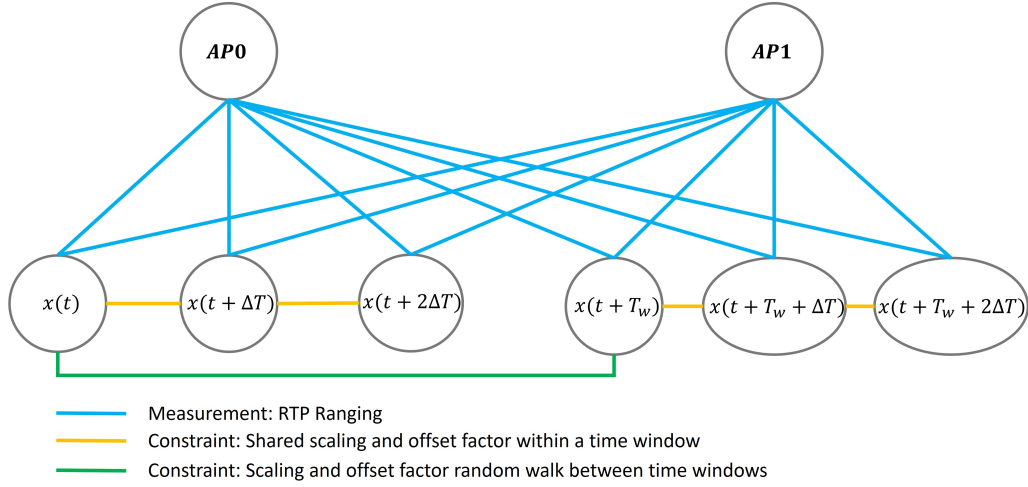


Fig. 11: Factor graph interpretation of localization model (only 2 APs are shown but 4 APs are used in the actual model). The blue lines symbolize the RTP measurements, serving as the input to the system. The $x(t)$ signifies the unknown state, comprising the location, scaling, and offset factors, which will be estimated from factor graph optimization after graph construction. The yellow and green lines denote the slow-changing constraints and are added automatically.

E. Performance

Since the non-linear optimization requires initial values, the centre of the room (2, 2) is set as the initial value for all the locations on the trajectory. The initial values for the scaling and offset factors are set to 1.0 and 0.0 respectively. Levenberg–Marquardt algorithm with GTSAM default parameters is used as the solver for the optimization. T_w is set to 1.0s and the variance of the distributions in the model is set empirically as Table III.

TABLE III: Variance Of The Distributions In Factor Graph Optimization

Name	Value	Name	Value
$\sigma_{rtp^i}^2$	0.005	$\sigma_{rtt^i}^2$	1.000
$\sigma_{rtp_\alpha^i}^2$	0.010	$\sigma_{rtt_\alpha^i}^2$	0.010
$\sigma_{rtp_\beta^i}^2$	0.050	$\sigma_{rtt_\beta^i}^2$	0.050

Factor graph optimization is applied to RTT ranging and RTP ranging individually. Fig. 12 shows the trajectories and Table IV shows the errors. It is clear that the RTP-based localization trajectories are very close to the ground truth, while the RTT-based trajectories seem chaotic. The average MAE for RTT and RTP are 1.83m and 0.26m, respectively. This evaluation indicates that our method using RTP is significantly better than the RTT. Note that in this evaluation, we plot the trajectory instead of discrete points. Even if the error for RTT localization is at meter level, which is not too bad, the trajectory looks disorderly since it is a time sequence-related visualization and the deviations will introduce significant distortion to the trajectory. The performance of our RTT localization method agrees with several RTT-only (fingerprint-free) localization research [10], [17] that the localization result is typically accurate around 1-2 meters. The RTT ranging error arises from the limitation in bandwidth and is the fundamental constraint. On the contrary, the proposed new method works

in the time domain by tracking RTP over time. In this way, it can bypass the issues caused by the limited bandwidth but requires frequent measurement. Fortunately, our prototype and experiments show that frequent RTP measurement is feasible and practical in the real world. Moreover, it can be implemented on commercial WiFi devices through a straightforward firmware patch. It offers an evident improvement that the RTP localization error is around one-tenth of the error when compared to RTT localization. It validates the performance of our system that realizes decimeter-level localization using existing WiFi infrastructure without data collection followed by training.

For time complexity, our method runs up to $2.2 \times$ real-time (raw data capture speed) on an Intel Core i7-7700 CPU. Even if this prototype is based on offline processing, the speed is fast enough for real-time processing, which could be implemented in the future.

Additional experiments are performed using non-line-of-sight channels that all of the APs are blocked by aluminum sheets as Section.V-E. Fig. 13 and Table V show the trajectory and error respectively. The trajectory of RTT-based localization is still random while the trajectory of RTP-based one tries to keep the general shape and the trend but with significant distortions and offsets. For the average localization error, RTT localization error increases by around 25% compared with line-of-sight channels but the value is almost doubled for RTP localization.

TABLE IV: Localization errors using line-of-sight channel

Trajectory	1	2	3	4	Average
RTT MAE (m)	2.19	1.47	1.84	1.82	1.83
RTT 50%CDF (m)	2.11	1.37	1.77	1.50	1.69
RTP MAE (m)	0.19	0.27	0.26	0.31	0.26
RTP 50%CDF (m)	0.18	0.23	0.24	0.32	0.24

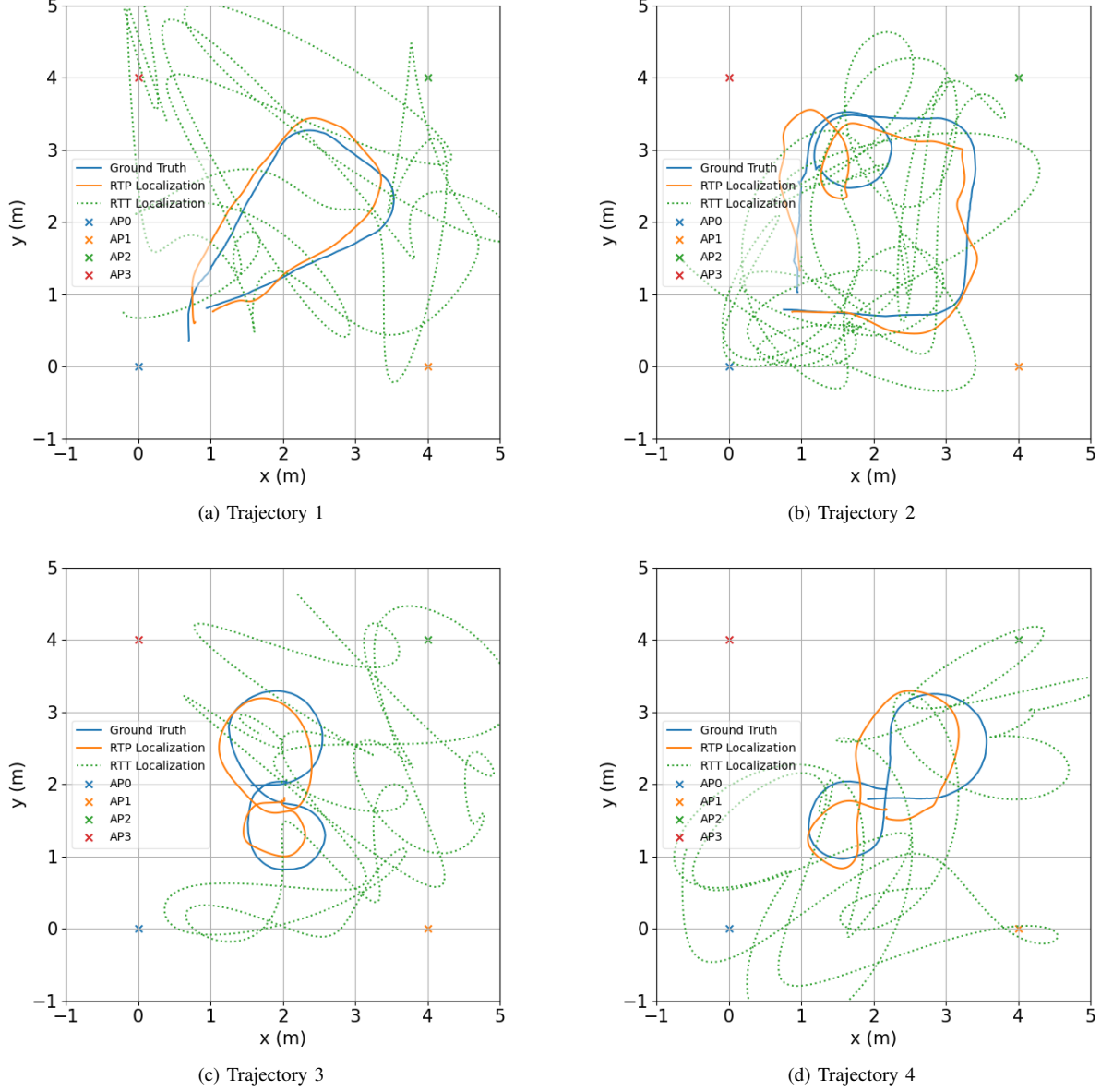


Fig. 12: Trajectories in the experiment using line-of-sight channel. Localization based on RTP demonstrates decimeter-level precision, in contrast to the meter-level precision achieved by RTT. The trajectories derived from RTT localization exhibit considerable disarray due to the substantial inaccuracies introduced by meter-level errors. For visualization, we plot the trajectories by sequentially connecting the waypoints, thereby amplifying the visibility of the errors for every waypoint. Given that the inaccuracies associated with RTT localization tend to be random, linking waypoints influenced by meter-level random errors produces chaotic patterns.

TABLE V: Localization errors using non-line-of-sight channel

Trajectory	1	2	3	4	Average
RTT MAE (m)	2.15	2.23	2.47	2.28	2.28
RTT 50%CDF (m)	2.00	1.96	2.33	2.36	2.16
RTP MAE (m)	0.52	0.45	0.84	0.42	0.56
RTP 50%CDF (m)	0.51	0.36	0.84	0.39	0.53

VII. CONCLUSION AND FUTURE WORK

This paper presents an indoor localization system based on the WiFi round-trip phase that uses a customized protocol and compares it with the WiFi round-trip time. Factor graph

optimization is adopted as the localization back-end and experimentally evaluated in real-world indoor scenarios. The prototype validates the idea of using WiFi carrier phase for localization and experiments show that the round-trip phase-based localization method offers decimeter-level accuracy. This method provides a practical solution for accurate indoor localization using WiFi signals because it is training-free and implemented on low-cost commercial devices. There are a few limitations in our system. The optimal performance can only be reached given good signal quality (e.g. line-of-sight) and it is a common issue with non-fingerprint-based localization.

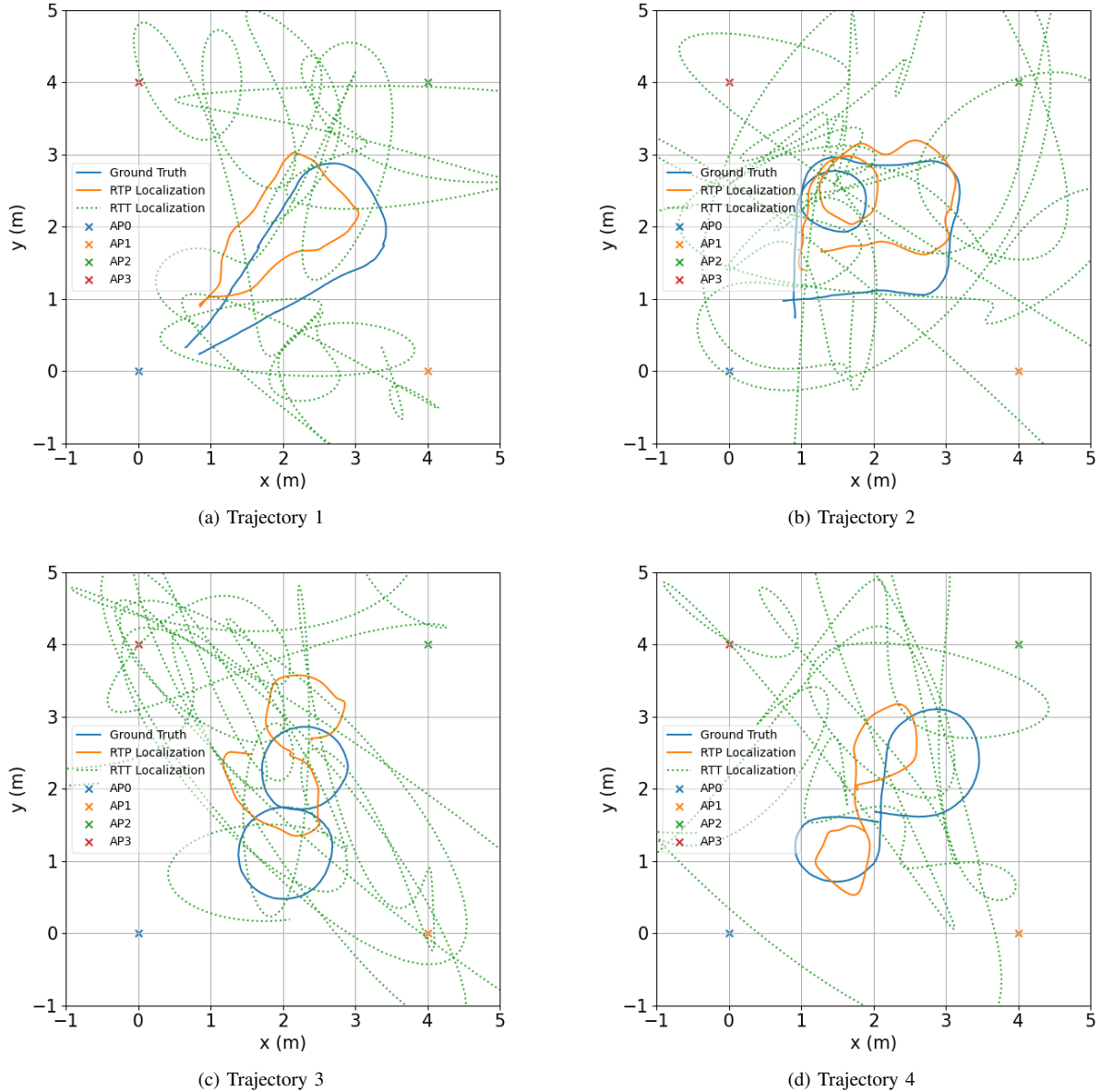


Fig. 13: Trajectories in the experiment using non-line-of-sight channel

Also, the sampling rate for RTP is limited in real-world scenarios and unwrapping the phase over time naively may not work if the fast motions of the robotic vehicle generate Doppler shifts greater than a given threshold. Fortunately, a similar problem has been solved in the GPS system (RTK) for carrier phase tracking and it can be solved by using an integer ambiguity model in phase tracking as [35]. Our proof-of-concept prototype is validated in a relatively small area and experiments in a larger area will be done in the future. It will involve utilizing professional motion capture or camera tracking devices in the filmmaking industry to acquire accurate ground truth because the IOS ARKIT cannot guarantee precise localization in a large area according to our observations. Since round-trip time and round-trip phase share very similar patterns in WiFi communication protocol, a new future WiFi protocol could be proposed that integrates both methodologies

for improved localization performance. Furthermore, similar to the time difference of arrival (TDoA), the round-trip phase system can be modified into a passive mode where the STA can measure the phase difference of arrival (PDoA) by listening to the communication between the APs for localization. In this way, the APs can serve an unlimited number of STAs simultaneously. Additionally, RTP can also be used in sensor fusion with other sensors like IMU for indoor localization. Last but not least, this round-trip phase-based method can also be applied to other wireless localization systems like UWB and 5G.

REFERENCES

- [1] "IEEE SA - IEEE 802.11-2016." [Online]. Available: <https://standards.ieee.org/ieee/802.11/5536/>
- [2] S. Sadowski and P. Spachos, "RSSI-Based Indoor Localization with the Internet of Things," *IEEE Access*, vol. 6, pp. 30 149–30 161, 6 2018.

- [3] M. Ibrahim, H. Liu, M. Jawahar, V. Nguyen, M. Gruteser, R. Howard, B. Yu, and F. Bai, "Verification: Accuracy evaluation of WiFi fine time measurements on an open platform," *Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICom*, pp. 417–427, 10 2018. [Online]. Available: <https://dl.acm.org/doi/10.1145/3241539.3241555>
- [4] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "SpotFi: Decimeter Level Localization Using WiFi," *Computer Communication Review*, vol. 45, no. 4, pp. 269–282, 8 2015.
- [5] J. Gjengset, J. Xiong, G. McPhillips, and K. 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*, ser. MobiCom '14. New York, NY, USA: Association for Computing Machinery, 2014, pp. 153–164. [Online]. Available: <https://doi.org/10.1145/2639108.2639139>
- [6] C. Chen, Y. Chen, Y. Han, H. Q. Lai, F. Zhang, and K. J. Liu, "Achieving Centimeter-Accuracy Indoor Localization on WiFi Platforms: A Multi-Antenna Approach," *IEEE Internet of Things Journal*, vol. 4, no. 1, pp. 122–134, 2 2017.
- [7] D. Vasisht, S. Kumar, and D. Katabi, "Decimeter-level localization with a single WiFi access point," in *Proceedings of the 13th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2016*, 2016.
- [8] "IEEE SA - IEEE 802.11az-2022." [Online]. Available: <https://standards.ieee.org/ieee/802.11az/7226/>
- [9] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proceedings - IEEE INFOCOM*, vol. 2, 2000.
- [10] C. Ma, B. Wu, S. Poslad, and D. R. Selviah, "Wi-Fi RTT Ranging Performance Characterization and Positioning System Design," *IEEE Transactions on Mobile Computing*, vol. 21, no. 2, pp. 740–756, 2 2022.
- [11] O. Hashem, K. A. Harras, and M. Youssef, "DeepNar: Robust Time-based Sub-meter Indoor Localization using Deep Learning," *Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks workshops*, 6 2020.
- [12] G. Guo, R. Chen, F. Ye, X. Peng, Z. Liu, and Y. Pan, "Indoor Smartphone Localization: A Hybrid WiFi RTT-RSS Ranging Approach," *IEEE Access*, vol. 7, pp. 176767–176781, 2019.
- [13] M. Youssef and A. Agrawala, "The Horus WLAN location determination system," in *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services, MobiSys 2005*, 2005.
- [14] Y. Shu, Y. Huang, J. Zhang, P. Coe, P. Cheng, J. Chen, and K. G. Shin, "Gradient-Based Fingerprinting for Indoor Localization and Tracking," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2424–2433, 4 2016.
- [15] W. F. Hassen and J. Mezghani, "CNN based approach for Indoor Positioning Services using RSSI Fingerprinting Technique," *2022 International Wireless Communications and Mobile Computing, IWCMC 2022*, pp. 778–783, 2022.
- [16] O. Hashem, K. A. Harras, and M. Youssef, "Accurate indoor positioning using IEEE 802.11mc round trip time," *Pervasive and Mobile Computing*, vol. 75, 8 2021.
- [17] C. Gentner, M. Ulmschneider, I. Kuehner, and A. Dammann, "WiFi-RTT Indoor Positioning," in *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. IEEE, 4 2020, pp. 1029–1035. [Online]. Available: <https://ieeexplore.ieee.org/document/9110232/>
- [18] S. Sen, B. Radunović, R. R. Choudhury, and T. Minka, "You are facing the Mona Lisa: Spot localization using PHY layer information," *MobiSys'12 - Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services*, pp. 183–196, 2012.
- [19] K. Wu, J. Xiao, Y. Yi, D. Chen, X. Luo, and L. M. Ni, "CSI-based indoor localization," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 7, pp. 1300–1309, 2013.
- [20] Z. Gao, Y. Gao, S. Wang, D. Li, and Y. Xu, "CRISLoc: Reconstructable CSI Fingerprinting for Indoor Smartphone Localization," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3422–3437, 3 2021.
- [21] X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-Based Fingerprinting for Indoor Localization: A Deep Learning Approach," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 763–776, 1 2017.
- [22] A. Goldsmith, *Wireless communications*. Cambridge University Press, 1 2005.
- [23] R. S. Campos and L. Lovisolo, *RF Positioning: Fundamentals, Applications, and Tools*. Artech House, 2015.
- [24] V. Barral, O. Campos, T. Dominguez-Bolano, C. J. Escudero, and J. A. Garcia-Naya, "Fine Time Measurement for the Internet of Things: A Practical Approach Using ESP32," *IEEE Internet of Things Journal*, 10 2022.
- [25] M. Speth, S. A. Fechtel, G. Fock, and H. Meyr, "Optimum receiver design for wireless broad-band systems using OFDM-part I," *IEEE Transactions on Communications*, vol. 47, no. 11, 1999.
- [26] Travis F. Collins, Robin Getz, Alexander M. Wyglinski, and Di Pu, *Software-Defined Radio for Engineers*. Artech House, 2018.
- [27] M. Schulz, D. Wegemer, and M. Hollick, "Nexmon: The C-based Firmware Patching Framework," 2017. [Online]. Available: <https://nexmon.org>
- [28] F. Gringoli, M. Schulz, J. Link, and M. Hollick, "Free Your CSI: A Channel State Information Extraction Platform For Modern Wi-Fi Chipsets," in *Proceedings of the 13th International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization*, ser. WiNTECH '19, 2019, pp. 21–28. [Online]. Available: <https://doi.org/10.1145/3349623.3355477>
- [29] S. Sen, J. Lee, K. H. Kim, and P. Congdon, "Avoiding multipath to revive inbuilding WiFi localization," *MobiSys 2013 - Proceedings of the 11th Annual International Conference on Mobile Systems, Applications, and Services*, vol. 13, pp. 249–262, 2013. [Online]. Available: <https://dl.acm.org/doi/10.1145/2462456.2464463>
- [30] S. Miura, L. T. Hsu, F. Chen, and S. Kamijo, "GPS Error Correction With Pseudorange Evaluation Using Three-Dimensional Maps," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 6, pp. 3104–3115, 12 2015.
- [31] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Fluids Engineering, Transactions of the ASME*, vol. 82, no. 1, 1960.
- [32] F. Dellaert, R. Roberts, V. Agrawal, A. Cunningham, C. Beall, D.-N. Ta, F. Jiang, lucacarlone, nikai, J. L. Blanco-Claraco, S. Williams, ydjian, J. Lambert, A. Melim, Z. Lv, A. Krishnan, J. Dong, G. Chen, K. Chande, balderdash-devil, DiffDecisionTrees, S. An, mpaluri, E. P. Mendes, M. Bosse, A. Patel, A. Baid, P. Furgale, matthewbroadwaynavenio, and roderick-koehle, "borglab/gtsam," 5 2022. [Online]. Available: <https://doi.org/10.5281/zenodo.5794541>
- [33] K. Levenberg and F. Arsenal, "A method for the solution of certain non-linear problems in least squares," *Quarterly of Applied Mathematics*, vol. 2, no. 2, pp. 164–168, 1944. [Online]. Available: <https://www.ams.org/qam/1944-02-02/S0033-569X-1944-10666-0/>
- [34] D. W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 6 1963.
- [35] Z. Zhang, S. Kang, and X. Zhang, "Indoor Carrier Phase Positioning Technology Based on OFDM System," *Sensors 2021, Vol. 21, Page 6731*, vol. 21, no. 20, p. 6731, 10 2021. [Online]. Available: [https://www.mdpi.com/1424-8220/21/20/6731](https://www.mdpi.com/1424-8220/21/20/6731/htmlhttps://www.mdpi.com/1424-8220/21/20/6731)