# BLoc: CSI-based Accurate Localization for BLE Tags

Roshan Ayyalasomayajula
University of California, San Diego
sayyalas@ucsd.edu

Deepak Vasisht
Massachusetts Institute of Technology
deepakv@mit.edu

Dinesh Bharadia
University of California, San Diego
dineshb@ucsd.edu

## ABSTRACT

Bluetooth Low Energy (BLE) tags have become very prevalent over the last decade for tracking applications in homes as well as businesses. These tags are used to track objects, navigate people, and deliver contextual advertisements. However, in spite of the wide interest in tracking BLE tags, the primary methods of tracking them are based on signal strength (RSSI) measurements. Past work has shown that such methods are inaccurate, and prone to multipath and dynamic environments. As a result, localization using Wi-Fi has moved to Channel State Information (CSI, includes both signal strength and signal phase) based localization methods. In this paper, we seek to investigate what are the challenges that prevent BLE from adopting CSI based localization methods. We identify fundamental differences at the PHY layer between BLE and Wi-Fi, that make it challenging to extend CSI based localization to BLE. We present our system, BLoc, that incorporates novel, BLE-compatible algorithms to overcome these challenges and enable an accurate, multipath-resistant localization system. Our empirical evaluation shows that BLoc can achieve a localization accuracy of 86 cm with BLE tags, a 3X improvement over a state-of-the-art baseline.

## CCS CONCEPTS

• **Networks → Location based services**;

## KEYWORDS

RF-based Indoor Positioning, Bluetooth Low Energy, Indoor Localization

**ACM Reference Format:**
Roshan Ayyalasomayajula, Deepak Vasisht, and Dinesh Bharadia. 2018. BLoc: CSI-based Accurate Localization for BLE Tags . In *The 14th International Conference on emerging Networking EXperiments and Technologies (CoNEXT '18), December 4–7, 2018, Heraklion, Greece.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3281411.3281428

## 1 INTRODUCTION

The vision of the Internet of Everything has been to communicate with and to locate everyday objects around us. The Bluetooth Low Energy (BLE) protocol has been a massive boost towards this vision, primarily because of two reasons: (a) BLE devices can communicate with off-the-shelf cellphones and access points with a range of 10 m , (b) BLE can enable communication at low power budgets. Today, BLE devices are used to track pets [9, 33, 34], find lost objects like keys[9, 33, 34], enable navigation in shopping malls [10], and automate operation in factory floors [32]. In addition to promising startups [9, 10, 33, 34], top technological companies like Google, Apple, etc. have invested heavily in this domain through iBeacons [4], Project Eddystone[13], etc. Powered by this push, it is estimated that the market for BLE beacons will touch 58 billion by 2025 [15].

Localization is a key primitive enabled by the BLE tags and is crucial for BLE adoption in several applications ranging from asset tracking [9, 10, 33, 34] to contextual advertisements [3, 16, 24, 29]. Today, localization algorithms for BLE primarily rely on measuring the strength of the received signal and using it as a proxy for location [6, 40]. However, relying on signal strength alone is problematic for three reasons: (a) The accuracy achieved by modeling signal strength is often low (several m), (b) RSSI estimates are not robust to multipath which commonly exists in real-world environments, and, (c) RSSI modeling is susceptible to large errors when the environment is dynamic [21, 42, 47].

In contrast to BLE localization systems, Wi-Fi localization has moved towards CSI (channel state information) based algorithms, that utilize both the signal strength and the signal phase to estimate the angle or distance between the transmitter and the receiver [21, 23, 31, 35, 38, 42]. Such systems have achieved high accuracy (around 1 m median error), have incorporated algorithms to weed out multipath and have consequentially been robust to dynamic nature of environments. Enabling similar accuracy and robustness for localization of BLE devices can significantly enhance the utility of existing BLE tags as well as deliver new applications. For example, one can predict whether you left the keys in the cupboard or on the table, rather than just telling you that the keys are at home. Alternatively, one could use them to accurately track pet motion [28]. Similarly, higher accuracy and robustness in industrial localization can automate processing pipelines.

In this paper, we seek to answer a simple, yet fundamental, question – what does it take for us to enable CSI based localization for BLE devices? Our investigation reveals that there are three key roadblocks that prevent CSI based localization for BLE devices:

- **Phase Measurement:** In Wi-Fi, the wireless channel is measured by sending known symbols simultaneously at multiple fixed narrowband frequencies (OFDM subcarriers) and measuring the corresponding received signal. In contrast, BLE uses Gaussian Frequency Shift Keying modulation and as a result, the information is encoded as the frequency of the signal. Specifically, in BLE, the bits 0 and 1 correspond to two different frequencies (say $f_0$ and $f_1$) within a 2 MHz band. When the transmitter wants to transmit bit 0, the transmission frequency is shifted to $f_0$ in the frequency domain. Similarly, for bit 1, the transmission frequency is moved to $f_1$. Furthermore, a Gaussian filter is applied


*CoNEXT '18, December 4–7, 2018, Heraklion, Greece*

ACM ISBN 978-1-4503-6080-7/18/12…$15.00
https://doi.org/10.1145/3281411.3281428

to the bits to ensure smooth frequency switching. Thus, the signal has continuously varying frequency, making it challenging to measure the wireless channel in the first place.
- **Bandwidth:** State-of-the-art CSI based algorithms for Wi-Fi based positioning rely on wireless channels across wideband frequencies available in Wi-Fi to identify the direct path and weed out the multipath [21, 23, 38]. The smallest band available on Wi-Fi is 20 MHz wide and the largest band is 160 MHz wide. In contrast, BLE channels are 2 MHz wide, at least one order of magnitude narrower than Wi-Fi. Since larger bandwidths help localization systems to separate multipath that is close to each other, the bandwidth limitation hampers BLE's ability to deal with multipath.
- **Multipath Resolution:** BLE signals reflect off various reflectors (walls, screens, furniture, etc) in the environment. Some of these reflections might actually be stronger than the line-of-sight path because of obstructions. Then, how can we isolate the direct path and ignore reflections so as to avoid large errors in localization?

We present, BLoc, a system to enable sub-meter localization for BLE devices using a CSI-based localization method. From the perspective of a deployer, the system appears similar to a standard BLE system. BLE anchors deployed in the environment measure signals from the target device and use the signal measurements to compute a location estimate. However, unlike current systems, BLoc uses both the signal strength and signal phase to enable high accuracy, reliable localization. BLoc's solution consists of the following key components:

- **Measuring CSI:** As mentioned before, the frequency of transmission in BLE is constantly varying, thereby making it hard to measure CSI. Our insight to solve this problem is fairly simple. We design bluetooth localization packets that have long sequences of 0s, followed by long sequences of 1s. By sending a long sequence of 0s, we can force the transmission to converge to $f_0$ and measure channel at $f_0$. Similarly, by sending a long sequence of 1s, we force the transmission to converge to $f_1$. The CSI measured at each of these frequencies can then be processed in subsequent steps to get a location estimate.
- **Stitching Frequency Bands:** To obtain the channel information across a wide band, we leverage the insight that BLE devices hop across 37 frequency bands spanning a 80 MHz band to avoid collisions with Wi-Fi transmissions. Our idea is to combine channel information across these different channels and emulate the presence of a large frequency bandwidth. However, when the transmitter and the receiver switch frequency bands, they also incur a random phase offset per frequency. This random phase offset completely jeopardizes the phase of the wireless channels across the different frequencies. Thus, we need to compensate for this phase offset if we want to utilize the phase of the signal for localization. We show in section 5.1, that by using a novel collaborative approach across multiple anchors, we can eliminate this phase offset per channel.
- **Multipath Elimination:** The typical algorithm to eliminate multipath and isolate the direct path from the transmitter to the receiver is to select the shortest path, since the direct path travels the least distance. The use of a wideband frequency enables us to measure the length of the paths and this length can then be used to find the direct path. However, when we use multiple anchor points collaboratively to eliminate phase offsets, we need to re-visit this notion, primarily because the distance between the target and an anchor point is now measured relative to other anchor points. We show that it is possible to analytically identify the direct path in this scenario, in section 5.4. Furthermore, we observe that multipath peaks tend to be spread out in the spatial domain, because real-life reflectors are imperfect (and act as scatterers as well). We use this observation to further enhance BLoc's multipath resolution capability.

We have built BLoc on a software radio platform, while retaining compatibility with the BLE protocol. Our experiments reveal the following:

- BLoc achieves a localization accuracy of $86cm$ in a multipath-rich environment. In contrast, an angle-of-arrival baseline gets an accuracy of just $2.42m$.
- BLoc's bandwidth enhancement from 2 MHz to 80 MHz band reduces the median error from $1.6m$ to $86cm$.
- BLoc's novel multipath rejection algorithm improves the localization accuracy by a factor of 2X, thereby proving the utility of CSI measurements in combating multipath effect.

To summarize, in designing BLoc we make the following key contributions.

- We present the first CSI-based BLE localization system.
- We build a new algorithm to correct phase offsets across multiple frequency channels, without relying on channel measurements at the target device.
- We show that BLoc can analytically identify the direct path even when the distance information available to us is measured relative to other anchor points.

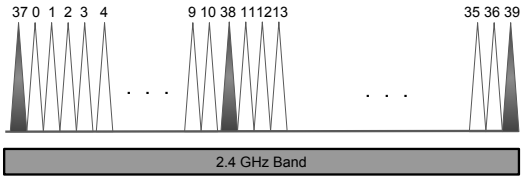
The rest of the paper is organized as follows, in section 2 we briefly describe the BLE protocol and how the localization based on CSI estimates work to familiarize the topic to the readers. Further, in section 3 we describe the deployment of the BLocsystem does not need major firmware and any hardware changes. In the following sections, we describe the design of BLoc's components. We start by describing how the CSI can be measured for BLE tags in section 4. Then, we discuss how we can use the CSI across multiple channel hops to localize the tag in section 5. The compatibility with BLE protocol is discussed in section 6. the experimental setup is described in section 7. Finally, we conclude with experimental evaluation of BLoc in section 8. We conclude the paper with past work in section 9.

## 2 PRIMER

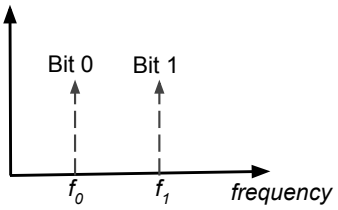
In this section, we include some information that will be useful in explaining BLoc's algorithms and system design.

### 2.1 Bluetooth Low Energy Protocol

Bluetooth Low Energy (BLE) is specifically designed for low power devices to communicate information over the ISM band (2.4 GHz - 2.48 GHz). While the exact specifics of the protocol are very detailed and quite complex, we try to abstract out the details necessary for understanding BLoc's design. We refer the reader to [12] for a detailed description of the BLE protocol.




(a) BLE Frequency Bands




(b) Frequency Shift Keying

**Figure 1—BLE Primer:** (a) BLE uses 40 frequency bands, 2 MHz wide each, spread over the 2.4 GHz ISM band. Of the 40 bands, 3 are designated advertisement bands (shaded) and the other 37 are data communication bands. (b) BLE uses Gaussian Frequency Shift Keying modulation where each bit corresponds to a different frequency.

The spectrum available for BLE is 80 MHz. This bandwidth is divided into 40 frequency bands, each of which is 2 MHz wide (see Fig. 1). BLE operates in two modes: broadcast mode and connected mode. The broadcast mode is used by devices to advertise their presence on the medium and contains information about the device, its type, manufacturer, etc. The broadcast mode operates on 3 of the 40 available bands for bluetooth. Upon receiving one of the advertised beacons on the broadcast bands, another device (master) can initiate a connection to this device (slave). During the connection establishment process, the master and the slave devices agree on several connection parameters, one of which is the frequency hop distance ($f_{hop}$).

Once the connection is established, the master and slave hop through the 37 non-broadcast bands, jumping by $f_{hop}$ bands every time a packet is exchanged between the master and the slave. Thus, if the first transmission happens at channel 10, and $f_{hop} = 3$, then the next transmission will be at channel 13. If $f_{cur}$ and $f_{next}$ denote the bands used for current and next transmissions respectively, then $f_{next} = f_{cur} + f_{hop} \mod 37$. Since the total number of bands is prime (37), the transmissions will hop through all available bands before repeating a band that has been used before. In section 5, we will show how BLoc uses this property to its advantage to distinguish between direct and reflected paths.

## 2.2 Localization

The key premise of wireless localization is to measure properties of wireless channels (like amplitude, phase, etc.) from a transmitter to one or more receivers and convert the measured properties to a location estimate for the transmitter. For simplicity, let us consider a signal with wavelength, $\lambda$, travelling in free space from a transmitter to a receiver separated by distance $d$. Then, the wireless channel, $h$, measured at the receiver can be modeled as[1]:

$$h = \frac{A}{d} e^{-\iota \frac{2\pi d}{\lambda}} \quad (1)$$

[1]We assume that the transmitter ans receiver do not have any hardware imperfections.

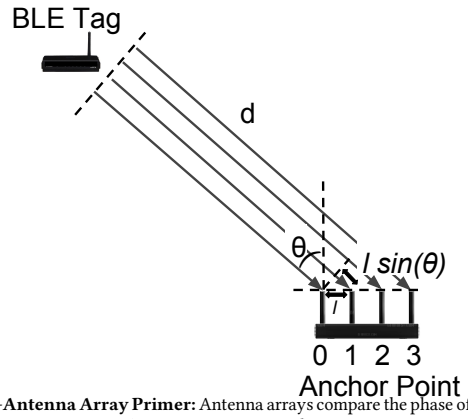



**Figure 2—Antenna Array Primer:** Antenna arrays compare the phase of the received signal at multiple antennas (separated by distance $l$) to identify the angle of the arrival of the signal.

Here, $A$ is the attenuation constant and $\iota = \sqrt{-1}$. When the signal is travelling along multiple paths, the wireless channel obtained is just the sum of the channel along each of these paths:

$$h = \sum_{i}^{N} \frac{A_i}{d_i} e^{-\iota \frac{2\pi d_i}{\lambda}} \quad (2)$$

where $N$ is the number of paths and $d_i$ is the length of the $i-th$ path.

**RSSI-based Localization:** In RSSI-based localization, the system measures $|h|$, i.e. the absolute value of the wireless channel. As can be seen in Eq. 1, for a single path, the absolute value of the channel directly depends on the distance. However, when there are multiple paths (Eq. 2), the absolute value of the channel is determined by how the channels along different paths combine. They can combine in-phase (making the channel amplitude high) or out-of-phase (completely canceling each other out), creating huge variations in amplitude for small changes in positions. Thus, RSSI based systems tend to suffer in the presence of multipath. Some RSSI-based systems try fingerprinting-based approaches wherein the RSSI at different locations is manually measured and these measurements are used to guide the predictions. However, any changes in the multipath characteristics of the environment (like moving furniture around) will require the system to do fingerprinting again, involving massive efforts.

**Measuring Angles:** When both the signal strength and the phase of the channel are available, a set of channel measurements at multiple antennas can be used to measure the angle of arrival of the signal. Consider a linear antenna array with $N$ antennas, as shown in Fig. 2. The target is placed at an angle $\theta$ with respect to the antenna array. Assume that antenna $i$ measures channel $h_i$. Observe that the distance from the target to antenna $i$ is larger than the distance from the target to antenna 0 by $il\sin\theta$, where $l$ is the separation between adjacent antennas. Thus, the channel to antenna $i$ incurs an additional phase of $\frac{-2\pi il\sin\theta}{\lambda}$ based on Eq. 1, i.e. $h_i = h_0 e^{-\iota \frac{2\pi il\sin\theta}{\lambda}}$. This transform can be reversed to identify the likelihood of transmission from each direction.

$$Pa(\theta) = |\sum_{i=1}^{N} h_i e^{\iota \frac{2\pi il\sin\theta}{\lambda}}| \quad (3)$$

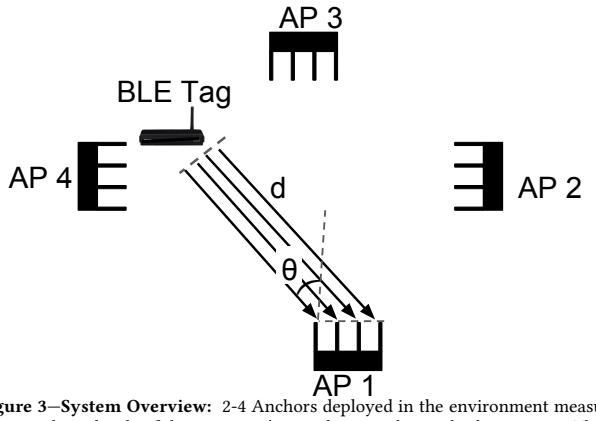



**Figure 3—System Overview:** 2-4 Anchors deployed in the environment measure phase and amplitude of the target tag's signal received at multiple antennas. These channels are then processed to give information about distance and angle, which is further mapped over 2D space to identify the location of the target devices.

where $Pa(\theta)$ gives the likelihood of the signal received from direction $\theta$. Even if there are multiple paths (direct and reflected), the signal along different directions can be separated out based on the direction that the signals arrive from. For a detailed description, see [21, 42].

**Measuring Distances:** Measurement of distance between a transmitter and a receiver is done by using multiple frequencies. Assume that we measure the channels at $K$ different frequencies such that $f_i = f_0 + i\delta f$. Then, using Eq. 1, the channel measured on frequency $f_i$, $h_{fi} = h_0 e^{-\iota \frac{2\pi i(f_0+\delta f d)}{c}}$, where $c$ is the speed of light. As you can see, the phase of the channels is a linear function of distance. Thus by comparing phases of the channels measured at multiple frequencies, the distance between a transmitter and a receiver can be estimated. Similar to Eq. 3, given channel measurements at different frequencies, we can compute the likelihood of the signal coming from each distance, $Pt(d)$ as (for details, see [35]):

$$Pt(d) = |\sum_{i=1}^{K} h_i e^{\iota \frac{2\pi i \delta f d}{c}}| \quad (4)$$

Finally, both multiple antennas and multiple frequencies can be combined to obtain a 2-d function, $P(\theta, d)$, that can compute the likelihood the signal coming from direction $\theta$ and distance $d$:

$$P(\theta, d) = |\sum_{i=1}^{K} h_{ik} e^{\iota \frac{2\pi i l \sin\theta f_0}{c}} e^{\iota \frac{2\pi k \delta f d}{c}}| \quad (5)$$

where $h_{ik}$ is the wireless channel measured on antenna $i$ and frequency $k$. Note that, in the joint case, $d$ is the distance measured from antenna 0.

## 3 SYSTEM OVERVIEW

In designing BLoc, we strive to achieve the following 3 objectives:

- **BLE Compatibility:**BLoc should not make any changes to the BLE protocol.
- **No change to user devices:** We intend BLoc to work with off-the-shelf target devices. BLE tags have seen wide spread deployments as trackers for pets, objects, factory instruments, etc. Therefore, BLoc should be backwards compatible and be able to work with off-the-shelf target devices. Thus, we cannot require BLE tags to have more capabilities (for example, the ability to measure and report phase) than today's deployed tags.
- **Sub-meter Accuracy:** We aim to achieve sub-meter localization accuracy for BLE targets. High localization accuracy for BLE can enable several novel applications. For instance, one can identify the exact location of lost objects, not just that they are at the shopping mall. It can also enable tracking of objects on factory floors, tracking of people in shops (down to the aisle and shelf) for local advertisements, etc.

An overview of BLoc's deployment setting is shown in Fig. 3. As shown in the figure, BLoc's anchor points, similar to BLE beacons, are deployed in the environment. The BLE tag connects to one of these anchor points (we call the connected anchor point the master) while the other anchor points passively listen for communication between the tag and the anchor. By listening to this conversation, they measure the CSI (both amplitude and phase) for both the transmissions (from tag to the anchor and anchor to the tag). Then, all the anchor points communicate to a central server to estimate the location of the tag.

## 4 MEASURING PHASE INFORMATION FOR BLE

As mentioned before, the first step for localization of BLE tags is to extract the phase information of the signal. The phase information can then be used for identifying the angle of arrival of the signal (using multiple antennas) and/or the distance between the tag and the anchor (using multiple frequency channels). Since the measurement of phase happens at the PHY layer and is intricately related to the PHY protocol, we start by describing the relevant details of the PHY protocol.

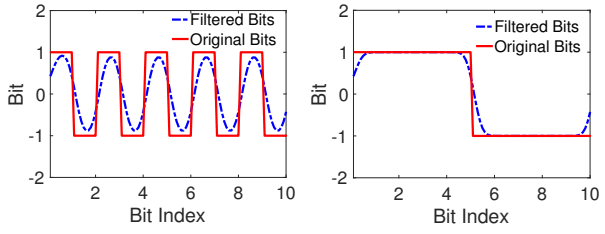



**Figure 4—GFSK**: (a) Gaussian filter applied to random BLE data leads to smooth changes in bits, but consequentially means that the frequency of the transmission is never static. (b) BLoc batches together long sequence of bits to ensure that a stable frequency is reached to obtain meaningful phase and amplitude measurements.

BLE uses Gaussian frequency shift keying (GFSK) modulation for transmitting data. In traditional frequency shift keying (FSK) protocols, each symbol corresponds to a frequency. For example, when using just two symbols (say, bit 1 and bit 0), there are two frequencies (say, $f_1$ and $f_0$) that correspond to the symbols. Thus, if the transmitter wants to transmit bit 1, it transmits at frequency $f_1$ and if the transmitter wants to transmit bit 0, it transmits at frequency $f_0$. However, to avoid frequent jumps in frequency (and out-of-band noise), BLE uses a Gaussian Filter on the bits. As a result, the bits are smoothed versions of 0 and 1 and hence, the frequency transitions are continuous. This issue is highlighted in Fig. 4(a), wherein the bit variation becomes continuous as a result

of the Gaussian filter. This continuous variation in bits (and hence frequency) implies that the frequency of transmission is never stable and prevents accurate channel measurements.

To overcome this problem, we leverage a simple technique. We construct BLE data packets with long sequences of bit 0 followed by long sequences of bit 1. Because we send long sequences of bit 0, the frequency value settles at $f_0$ and we can then measure the wireless channel at $f_0$, $h_0$. The wireless channel can simply be measured by taking the ratio of the received symbol to the transmitted symbol. If the transmitted symbol is $x_0$ and it is received as $y_0$ at the receiver, the channel $h_0$ at frequency $f_0$ can be measured as: $h_0 = \frac{y_0}{x_0}$. Similarly, we can measure the channel $h_1$ at frequency $f_1$ by transmitting a sequence of one-bits. We demonstrate this graphically in Fig. 4(b). As shown in the figure, sequences of 5 zero-bits followed by 5 one-bits lead to almost constant frequency for significant chunks of time. These stable transmissions can then be used to measure the phase of the channel at these frequencies.

## 5 FROM CSI TO LOCATION

Now that we have obtained the complex-valued wireless channel for a BLE band, we present BLoc's algorithm to infer the location of the device. Before we dig into the details, let us establish some standard notation. We denote the complex-valued channel, measured at anchor $i$, antenna $j$ and frequency band $f$ by $h_{ij}^f$. Note that, BLE allows multiple frequency bands. As mentioned before, we can measure two channel values for each band. We combine the two values into a single value per band by averaging the channel amplitude and channel phase separately and combining them into a single channel value. This channel value is assumed to be the wireless channel at the center frequency of the band. Furthermore, we denote the amplitude and phase of $h_{ij}^f$ by $|h_{ij}^f|$ and $\angle h_{ij}^f$ respectively. Finally, we have a total of $I$ anchors, $\mathcal{J}$ antennas per anchor and $f \in \{f_k | k = 1...K\}$.

### 5.1 Combining Frequency Bands

The key benefit of using channel phase for localization is the additional ability to resolve multipath. By using channel phase, one can identify the delay of individual paths (both direct and reflected) and hence, pick out the shortest paths to be the direct paths. However, the ability to separate out different paths depends on the available bandwidth. If the closest separation between paths is $\delta d$, then the frequency bandwidth, $BW$ required to identify them is given by:

$$BW \geq \frac{c}{\delta d} \tag{6}$$

where $c$ is the speed of light. Thus, for BLE's effective bandwidth of 1 MHz[2], the corresponding distance is $300m$. This is larger than all distances in indoor settings and hence, any paths that are closer to each other than $300m$ cannot be distinguished from each other. Clearly, this bandwidth is not sufficient to resolve indoor multipath.

To solve this problem, we make the observation that while BLE has effective bandwidth of $1MHz$, it has frequency hopping built in. It changes the frequency of transmission after every packet. Specifically, if the current frequency band is $k_1$ (with frequency $f_{k1}$, then for the next packet, the transmission will happen in frequency band, $k_1 + k_{hop} \mod 37$, where $k_{hop}$ is the parameter specific to a connection and 37 is the total number of BLE frequency bands possible. Since 37 is a prime number, $k_1 + k_{hop}$ is guaranteed to go through all frequency bands for all values of $k_{hop}$. Thus, if we can measure and combine the wireless channel information across multiple frequency bands, we can span a total of $80MHz$ (the total span of the BLE frequency bands, see Fig. 1(a)).

So, to achieve a large bandwidth and improve the multipath resolution, BLoc measures channel data on all BLE frequency bands. Thus, we can measure channels on multiple antennas per anchor and on different frequency bands. Can we just use these measured wireless channels and plug them in Eq. 5 to obtain the likelihood of the signal coming from different angles and distances? Unfortunately, this is not as straightforward. The prime reason for this is the lack of phase synchronization between the transmitter and receiver. Every BLE device has a local oscillator responsible for generating the signals. This oscillator is used to tune the system to different frequencies. However, every time this oscillator is used to tune the frequency, it incurs a random phase offset. Thus, if the transmitter has a phase offset, $\phi_T$[3] and the receiver has a phase offset, $\phi_R$, the channel measured at the receiver is given by $\hat{h}_{ij}^f = h_{ij}^f e^{\phi_T - \phi_R}$. This phase offset $(\phi_T - \phi_R)$ is random and changes per frequency switch. Thus, the phase of the channel measurements is completely garbled and hence, cannot be used for localization directly.

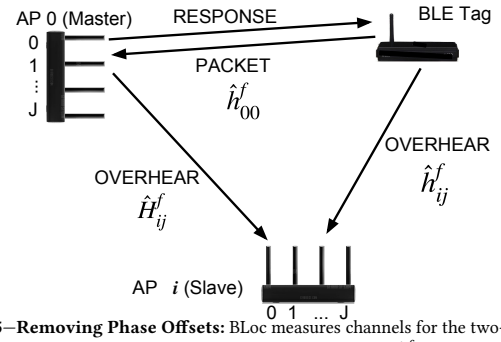



**Figure 5—Removing Phase Offsets:** BLoc measures channels for the two-way communication at the slave anchor's (AP $i$'s) first antenna (j=0), $\hat{h}_{ij}^f$, from the BLE tag to the master anchor(AP0), $\hat{H}_{ij}^f$ and the master anchor to the target, $\hat{h}_{00}^f$. It then combines these measurements as shown in Eq. 10 to remove phase offsets.

### 5.2 Combating Phase Offsets

How can we retrive the underlying physical channel, $h_{ij}^f$ from the measured wireless channel with garbled phase values, $\hat{h}_{ij}^f$? To solve this problem, we once again rely on the BLE protocol. In the BLE protocol, during one communication period, the master and the slave talk to each other, i.e. there is a two-way exchange of packets. In BLoc, we designate one of the anchors as a master and it exchanges packets with the target BLE tag. All other anchors measure CSI for both sides of the conversation, i.e., they measure the channel from the target tag to themselves and from the master anchor to themselves. To understand how this helps, let us assume that the

[2]While each band in BLE is $2MHz$, the separation between the two data bits is just $1MHz$.

[3]Since all antennas on an anchor are driven by the same oscillator, the phase offset only varies across anchors and not within one anchor.

master anchor is anchor 0, i.e. $i = 0$ for the master anchor. Then, let us assume that anchor $i$ measured channels $\hat{h}^f_{ij}$ from the target to itself and measured channels, $\hat{H}^f_{ij}$ from antenna 0 on the master AP to itself. Then, we claim that: $\hat{h}^f_{ij}\hat{H}^f_{i0*}\hat{h}^{f*}_{00}$ is independent of any random phase offsets, where $(.)^*$ denotes the complex conjugate operation. We present a proof of this claim below. For this proof, we denote the phase offset of the target tag by $\phi_T$ and the phase offset of the $i-th$ anchor by $\phi_{Ri}$. Further, let us assume that $H^f_{ij}$ is the true physical wireless channel from antenna 0 on anchor 0 to antenna $j$ on anchor $i$, measured at frequency $f$.

$$\hat{h}^f_{ij} = h^f_{ij} e^{\phi_T - \phi_{Ri}} \tag{7}$$

$$\hat{h}^f_{00} = h^f_{00} e^{\phi_T - \phi_{R0}} \tag{8}$$

$$\hat{H}^f_{i0} = H^f_{i0} e^{\phi_{R0} - \phi_{Ri}} \tag{9}$$

$$\implies \hat{h}^f_{ij}\hat{H}^{f*}_{i0}\hat{h}^{f*}_{00} = h^f_{ij}H^{f*}_{i0}h^{f*}_{00} \tag{10}$$

Eq. 10 shows that the quantity $\hat{h}^f_{ij}\hat{H}^{f*}_{ij}\hat{h}^{f*}_{00}$ is independent of phase offsets and just depends on the underlying physical wireless channels. In equation 10 there are three terms of channel:

- from the tag to the anchor $i$ (slave anchor), measured using the overhear of the packet transmission from the tag at frequency $f$, $\hat{h}^f_{ij}$
- from the anchor 0 (master anchor) to the anchor i (slave anchor), measured using the overhear of the packet response from the Master anchor to the tag at frequency $f$, $\hat{H}^f_{i0}$
- from tag to the master anchor at frequency $f$, $\hat{h}^f_{00}$.

And as discussed in 3, the central server has access to all of these channel estimates and can use them to correct for phase offsets. For ease of exposition, let us denote, $\alpha^f_{ij} = \hat{h}^f_{ij}\hat{H}^{f*}_{i0}\hat{h}^{f*}_{00}$. We call $\alpha^f_{ij}$ to be corrected channels for the rest of the discussion. The corrected channels are free from any random phase distortions caused due to frequency switching. We have been able to achieve this desirable property by relying on the observation that both the anchor and the tag transmit data during a communication period.

## 5.3 Estimating Location Probabilities

Even though the corrected channels are free from phase distortions per frequency, do they retain the information about the underlying geometric world that can enable us to do localization? To answer that question, let us re-write the corrected channels, $\alpha^f_{ij}$ in their expanded geometric form. In the equations below, we use the term $d^{lm}_{ij}$ to denote the distance from antenna $j$ on anchor $i$ to antenna $m$ on anchor $l$. Furthermore, $d^{ij}_T$ represents distance from the tag to antenna $j$ on anchor $i$.

$$\alpha^f_{ij} = \hat{h}^f_{ij}\hat{H}^{f*}_{i0}\hat{h}^{f*}_{00} \tag{11}$$

$$= h^f_{ij}H^{f*}_{i0}h^{f*}_{00} \tag{12}$$

$$= e^{-\iota d^{ij}_T \frac{2\pi f}{c}} e^{\iota d^{i0}_{00} \frac{2\pi f}{c}} e^{\iota d^{00}_T \frac{2\pi f}{c}} \tag{13}$$

$$= e^{-\iota \frac{2\pi f}{c}(d^{ij}_T - d^{i0}_{00} - d^{00}_T)} \tag{14}$$

There are two important aspects to note about Eq. 14:

- **Effect on Angle Measurements:** As we explained in Eq. 3, the relative phase measured by the different antennas on a single anchor point determines the angle-of-arrival of the incoming signal (i.e., the direction of the signal from the target tag). Thus, if we add the same constant phase to the channel measurements on all the antennas of a single anchor point, the angle distribution does not change. Because we multiply the channels on all antennas on an anchor point by the same conjugate channel values, we add the same phase measurements to all antennas. Thus, the angle distribution of the received signal can be computed using the corrected channels, $\alpha$, directly. Thus, we can just replace the true channels($h$) by corrected channels ($\alpha$) in Eq. 3 and then, write the angular distribution of received signal at anchor $i$ as:

$$Pa_i(\theta) = \left| \sum_{j=1}^{\mathcal{J}} \alpha^f_{ij} e^{\iota \frac{2\pi j l \sin\theta}{\lambda}} \right| \tag{15}$$

  This distribution gives us information about the angle of arrival of the signal (and its multipath reflections). A sample of this distribution mapped over the 2D space is shown in Fig. 6(a).
- **Effect on Distance Measurements:** First of all, note that $d^{i0}_{00}$ is a fixed-distance known a priori because the distance between anchor $i$ and anchor 0 can be measured one-time during deployment. Thus, Eq. 14 shows that the corrected channel contains information about relative distances, i.e., distances measured with respect to anchor 0, antenna 0 (i.e. $d^{ij}_T - d^{00}_T$). As we saw in Eq. 4, we can use channel measurements at multiple frequencies to extract this distance information out. All we need to do is to replace distance in Eq. 4 by the relative distance and the channels by corrected channels, $\alpha$. Thus, we can denote the likelihood of the signal received at antenna $j$ on anchor $i$ coming from a relative distance $d^{ij}_T - d^{00}_T$ as,

$$Pt_{ij}(d^{ij}_T - d^{00}_T) = \left| \sum_{k=1}^{K} \alpha^{f_k}_{ij} e^{\iota \frac{2\pi f_k}{c}(d^{ij}_T - d^{00}_T - d^{i0}_{00})} \right| \tag{16}$$

  A sample of this distribution mapped over the 2D space is shown in Fig. 6(b). Note that, because we measure relative distances as opposed to absolute distances, the shape of the high probability region looks like a hyperbola.

We can combine Eq. 15 and Eq. 16 to write a joint distribution over distance and angles with respect to each anchor. Let us denote this joint distribution as $P_i$ for the $i-th$ anchor.

$$P_i(d^{i0}_T - d^T_{00}, \theta) = \left| \sum_{j=1}^{\mathcal{J}} \sum_{k=1}^{K} \alpha^{f_k}_{ij} e^{\iota \frac{2\pi f_k}{c}(d^{ij}_T - d^{00}_T - d^{i0}_{00})} e^{\iota \frac{2\pi j l \sin\theta f_k}{c}} \right| \tag{17}$$

Eq. 17 can further be mapped onto the 2-D cartesian coordinates by a simple change of coordinates. Thus, Eq. 17 gives us a likelihood of the signal originating from every point in space. A sample of this likelihood is plotted in Fig. 6(c). We get this likelihood for each anchor. We simply add the likelihood obtained from each anchor together to get the joint likelihood over the X-Y space. Now that we have obtained this distribution over space, we need to pick the position of the target, given this distribution. Before we do that, let us recall the steps we have done so far:

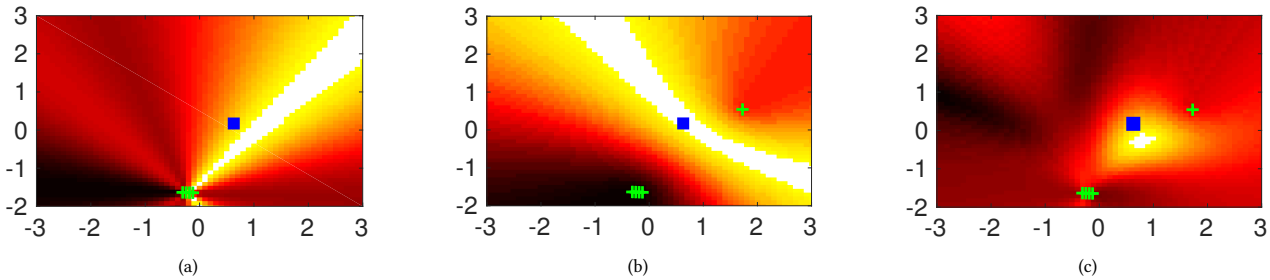



**Figure 6—CSI to Location:** (a) The likelihood distribution of a signal source obtained using multiple antennas on a single anchor point using Eq. 15, (b) The likelihood distribution of a signal source obtained using multiple frequency bands using Eq. 16. The shape of the high likelihood region is hyperbolic because the distances measured are relative. (c) The combined likelihood distribution obtained using Eq. 17. Blue square marks the actual location of the source.

- We measure wireless channels on each anchor for two packets, one from the target tag to the master anchor and the other from the master anchor to the tag.
- We measure the wireless channels above for all BLE frequency bands.
- We use the measured channels to obtain a spatial distribution of the likelihood of the tag presence in a given location. We do this by computing $P_i$ using Eq. 17.
- We add the likelihood values obtained by each AP to obtain a joint likelihood distribution. An example of this likelihood distribution mapped over space is shown in Fig. 6(c).

## 5.4 Resolving Multipath

So far, we have used the measured wireless channels on the anchors to calculate the likelihood of the tag being present at a given location in space. Given this likelihood distribution, how do we ascertain where the tag really is? A naive way to solve this problem would be to pick the point with the highest likelihood. However, given multipath effects and obstructions in the environment, the direct path may not always be the strongest. In a lot of the cases, the reflections of the tag might overwhelm the direct path. How can we reason about the correct position of the tag in that case?

There are two approaches in existing works to resolve the issue of multipath. The first approach is to isolate the direct path among all the existing paths and thus filtering out the reflections [21, 30, 42]. The second approach is to utilize reflected components in combination with the direct path to enable better localization [31]. But, the reflected components have a diffused distribution over time in their angle of arrival and time of flight estimates [21]. Thus, we use the former method to resolve multipath. So, how can we identify the direct path and filter out the reflections?

First, observe that in Eq. 16, even though we are measuring relative distances, the shortest path continues to be the shortest path even in the case of multipath. This is because the reference distance (in this case, $d_T^{00} + d_{00}$) is being subtracted from all the paths[4]. Thus, we can rely on the fact that direct paths have shortest distance as compared to reflected paths. Therefore, for each peak in the likelihood profile, we evaluate if it has the shortest distance for each anchor point.

[4]While we discuss this observation in the context of a single path from the tag to the master anchor point, it continues to hold even if multiple paths exist from the tag to the master anchor points

Second, we observe that multipath reflections are bound to be spread out in space as opposed to direct paths which are more peaky. The reason behind this intuition is that the direct paths are being directly transmitted by an antenna, whereas the reflection is happening off surfaces which are non ideal reflectors. Since, they are non-ideal reflectors, they can scatter some parts of the incident signal. Furthermore, different anchors see reflections from different parts of the reflector, making the likelihood distribution more spread out. To quantify this intuition, we compute the spatial entropy of the likelihood distribution around all peaks, i.e., for each peak in the likelihood distribution, we compute the entropy of the likelihood distribution in its immediate neighborhood. If the likelihood distribution is almost flat, the entropy will be low and hence, the path is more likely a reflected path.

We do a weighted combination of these two factors to determine the position of the tag. For each peak in the likelihood distribution over space, we define a score, $s_x$ given by:

$$s_x = p_x e^{b\mathcal{H} - a\sum_i d_i} \tag{18}$$

Here, $x$ is the location of the peak, $p_x$ is the joint likelihood value of the peak, $d_i$ is the distance measured from anchor $i$ corresponding to this peak and $H$ is the entropy in the neighborhood of the peak. $a$ and $b$ are weights for the two components of the score function. Once we have computed the score function for all peaks, we can just pick the peak with the highest score to be the direct path.

# 6 DISCUSSION

To conclude the discussion of the techniques behind BLoc, a few points are worth mentioning:

- The system requires no changes to BLE tags. The tags talks to a master anchor using the BLE protocol.
- BLoc complies with BLE protocol, to the best of our knowledge. The frequency hopping used for BLoc is in-build in BLE. We leverage the existing hopping to our advantage and used the increased frequency bandwidth to mitigate the multipath effect.
- BLoc requires software/firmware changes to the BLE anchors to report signal phase. Currently, no off-the-shelf BLE devices provide access to signal phase in the application layer. Thus, to deploy BLoc, today, one would have to use anchors which are customized. However, the operation of the anchors is compliant with BLE protocol.

Further, it would seem that the long streams of 0/1's employed for the BLoc's channel estimation would effect the usual BLE communication. But as we know, BLE hops through all channels 40 times every second. Thus, even if one complete hop is used for localization, the other hops can be used to communicate data as usual. This location frequency suffices for typical indoor navigation applications. Thus, for the CSI estimate to be possible, we would just need $8\mu sec$ for each 0 and 1 , similar to the way [17] achieve tones for Bluetooth, which should not effect the throughput of the usual BLE communication.

# 7 IMPLEMENTATION

**Hardware Setup:** We implement BLoc on USRP (Universal Software Radio Peripheral) platform ([11]). We use USRP N210s to create four 4-antenna BLE anchor points. All antennas on one anchor point are driven by the same clock to ensure time and frequency synchronization within one anchor point. The target device is also run using a USRP N210, but it has only one antenna, as is common with BLE tags.

**Software Setup:** We implement the BLE PHY layer on the USRP platform in C as a patch to the UHD (USRP Hardware Driver) code. The USRP software radios are connected to PC's over ethernet and the data sent by them is processed on a central server. The server collects the complex-valued signals and processes the signals to estimate the location of a client in MATLAB.

**Ground Truth Estimates:** We conduct our experiments in a $5m$ by $6m$ room equipped with the VICON motion capture system [1]. The VICON motion capture system relies on an array of infrared cameras to deliver mm-level accuracy in tracking objects equipped with visible infrared markers. We equip our target device with 4 infrared markers to track it accurately using the VICON motion capture system. The VICON estimates are used as ground truth for our evaluation. We *do not* use the output of the VICON system at all in our algorithmic implementation. Finally, note that the VICON room is a shared space and is full of metallic objects, like robotic equipment, large metal cupboards, etc. As a result, the room is rich in multipath and presents a challenging localization environment.

**Experiment Procedure:** We manually move the tag to randomly sampled locations in the VICON room. The anchor points are present on the 4 edges of the VICON room, in the centre of each edge. When the tag is moved to a location, its ground truth location is estimated using the VICON system. The channel measurements are performed on each antenna of every anchor point. Once the measurements are done, the tag is moved to a different location. Overall, we measure the ground truth of channels in 1700 different locations, which serves as the dataset for our evaluation. The 1700 points cover the entire space. The average separation between two nearest neighbors is 10 cm.

**Compared Schemes:** We compare BLoc with an angle-of-arrival (AoA) baseline localization system. Many of the state-of-the-art Loclaization systems [21, 42], build on AoA. Thus, we take AoA-combining as a baseline comparison. We further implement both BLoc and the baseline similar to [21] using the same number of antennas and the same set of channel measurements to compare with the state-of-the-art wireless localization.

Based on emperical observations, for all the results reported in section 8 we use $a = 0.1$ and $b = 0.05$, and use a circular neighborhood window of window size $7 \times 7$ in the BLoc algorithm for entropy calculations.

# 8 EXPERIMENTAL EVALUATION

We present our evaluation of BLoc below.

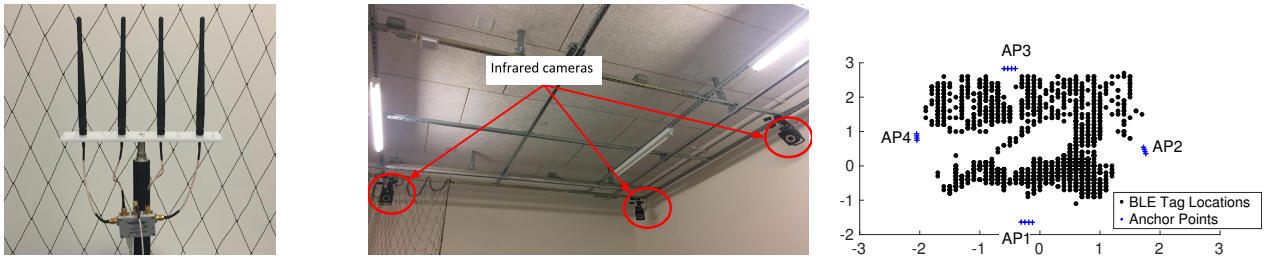
## 8.1 Microbenchmarks

Before we delve deeper and analyze the overall localization accuracy of BLoc, we discuss some microbenchmarks to illustrate some of its important aspects.

First, as we discussed before, measuring consistent CSI measurements for BLE is non-trivial, since the frequency continuously varies with time (within one packet) depending on the sequence of data bits. Thus, we design special packets with long sequences of 0 bits followed by long sequences of 1 bits. Does this sequence design lead us to measure consistent phase across time? To check this intuition, we plot the CSI measured by BLoc for 10 consecutive measurements on 4 different frequency channels in Fig. 8(a). Note that, BLE has 40 different frequency channels but we choose 4 bands for illustration. As can be seen in the figure, the phase of the channel remains consistent across measurements, revealing the stability of the CSI measurements.

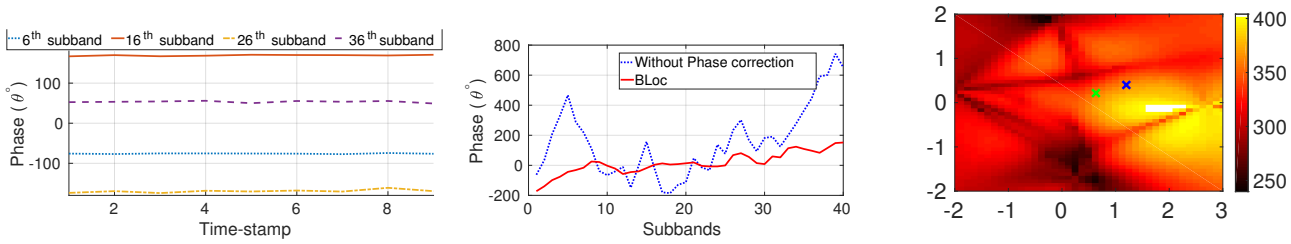
Further, we want to see if we can combine the signal across multiple anchor points, to avoid channel-dependent random phase offsets. To check this intuition, we place the target and two APs in line of sight in a relatively multipath free environment. In this case, the expectation is that the phase across multiple channels varies linearly with frequency. However, random phase offsets will make this phase offset vary randomly across frequency. Thus, we compare the phase of the CSI measurements in two cases: (a) when BLoc's phase offset cancellation is applied (red curve in Fig. 8(b)), (b) when there is no offset cancellation (blue curve in Fig. 8(b). As can be seen in this figure, the blue curve varies randomly with frequency, whereas the red curve shows linear behavior across frequency. This graph shows that the random phase offsets incurred due to channel switching can be cancelled by use of BLoc's offset cancellation scheme.

Finally, we plot a sample localization profile plotted over space in Fig. 8(c). The x-y axis in this plot correspond to the spatial X-Y axis. The colormap defines the probability of presence of the device at that location (white is the highest probabilit). The real location of the target is marked by a green x, while the prediction by BLoc is in blue x. There are two interesting aspects of this figure. First, there are multiple locations that are possible for the device due to the multipath present in the environment. This highlights the requirement for BLoc's multipath cancellation algorithm. Second, the multipath peaks are more spread out than the direct path. This is because the reflectors in the environment are not ideal reflectors and so, there is not one single point on them that reflects to all the anchor points and all the antennas. This leads them to be spread out. This observation validates our insight mentioned before. A detailed evaluation of BLoc's multipath cancellation algorithm is given in section 8.7. We can further observe that BLoc has predicted



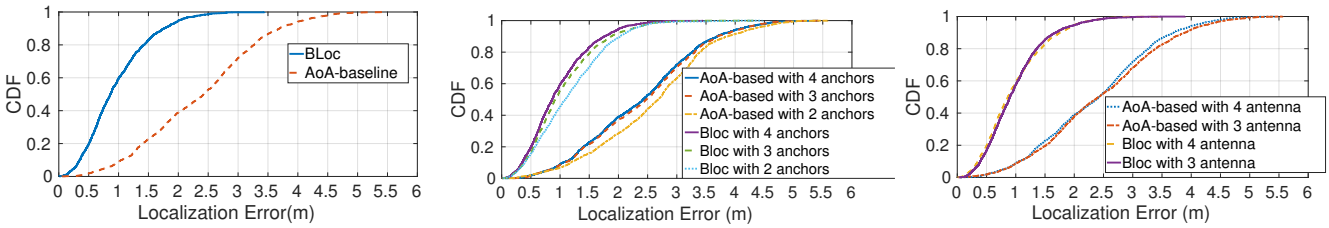

(a) Anchor Point (b) Testbed (c) Point Distribution

**Figure 7—Implementation:** (a) A shot of one of the 4 antenna Anchor points of the setup, (b) The infrared camera to measure the ground truth for calculating localization errors, (c) The top-view of the setup, we have 4 Anchors with 4 antenna each (in blue +) and the 1700 ground truth positoins of the BLE tag.




(a) CSI Measurement (b) Combining Across Anchor Points (c) Multipath Profiles

**Figure 8—Microbenchmark**: (a) CSI measurements taken for {6,16,26,36} subbands for 9 instances are constant over time. (b) Phase of without BLoc (Combining over APs) and BLoc (Combining across) for a given client location across subbands, (c) Sample combined Multipath profile in X-Y coordinates, combined across subbands and multiple antennas on an AP and across APs




(a) Localization Accuracy (b) Effect of number of Anchors (c) Effect of number of antennas

**Figure 9—Localization Accuracy**: (a) CDFs of Localization error *inmeters* for BLoc and AoA-baseline, (b) CDF plots of Localization errors (in meters) for {2,3,4} anchor points for both BLoc and AoA-combining baseline, (c) CDF plots of localization errors (in meters) for {3,4} antennas on each anchor point for both BLoc and AoA-combining baseline.

the right peak, in that the predicted location and the actual location belong to the same Maxima's neighborhood.

## 8.2 Localization Accuracy

To measure the localization accuracy of the system, we deploy BLoc in the environment shown in Fig. 7. The environment is equipped with 4 BLoc anchor points and has 1 target which is moved to 1700 different positions to evaluate. We measure the location of the device using two different schemes: using BLoc and using least ToF based AoA localization systems [21, 42], which is the state-of-the-art in localization. We report the distance between the estimated position of the target and the actual position of the target. The cdf of the localization errors are plotted in Fig. 9(a).

As shown in the figure, BLoc achieves a median error of 86 cm, whereas the AoA-combining based system achieves a median error of 242 cm. The 90th percentile of the localization error is 170 cm and 340 cm for BLoc and the baseline respectively. Thus, BLoc clearly outperforms traditional AoA-combining based methods in our evaluation. The primary reason behind the better performance of BLoc is its ability to deal with multipath effects. It has been well studied that for a given number of antennas per AP, there is only a fixed number of multipath that one can resolve.

## 8.3 Effect of Number of Anchor Points

Further, we want to analyze the effect when the number of anchor points is changed. To evaluate this behavior, we compute the localization error for 3 anchors and 4 anchors, for both the baseline and BLoc. For the 3 anchor scenario, we take all possible subsets of the 4 deployed anchors and report the average of those errors for each data point. The cdf of the localization errors for 3 and 4 anchors for both AoA baseline and BLoc are plotted in Fig. 9(b).

As expected, the results for 3 anchors is slightly worse than the 4 anchors for both the schemes. The median error for BLoc goes up to 91.5cm from 86cm and the 90th percentile goes up from 170 cm to 175cm when one goes to 4 anchors to 3 anchors. On the other hand, for AoA-combining based baseline, the median error goes up from 242 cm to 247 cm, and the 90th percentile from 340 cm to 350 cm. However, in spite of this reduction, BLoc continues

to achieve sub-meter median accuracy even with 3 anchors. We can see from these results that just using AoA-combining based localization would degrade if one does not have sufficient anchor points who have (Line-of-Sight) LOS. While BLoc uses both AoA and distance information simultaneously to localize, so even if one has fewer anchors one can get better performance. We can further see from Fig. 9(b) that for 2 anchors there is a significant increase in median and 90th percentile errors for both BLoc and AoA-combining baseline.

## 8.4 Effect of Number of Antennas

Prior work has shown that increasing the number of antennas improves the resolution of the antenna array and hence improves localization accuracy. We want to verify what effect this reduction in the number of antennas has on BLoc and the AoA-combining baseline. To understand this effect, we compare the localization errors achieved with 3 antennas and 4 antennas for both the schemes. In this experiment, we use all 4 anchors. The cdf of the localization errors for 3-antenna anchors and 4-antenna anchors are shown in Fig. 9(c).

As shown, the median 3-antenna localization error for BLoc is 90 cm (90th percentile is 171 cm). For AoA-combined baseline, the median 3-antenna localization error is 241 cm (90th percentile 320 cm). This shows that the reduction in the number of antennas causes a minimal effect on the localization accuracy. This is mainly because BLoc relies on two different components for its multipath resolution: number of antennas and the frequency bandwidth. A reduction in the number of antennas is compensated by the frequency bandwidth and hence, does not effect the accuracy very much.

## 8.5 Bandwidth Variation

One of the goals of BLoc is to increase the frequency bandwidth used for BLE localization, so as to achieve higher accuracy. We presented an algorithm to achieve this goal in section 5.1. Now, we intend to empirically investigate the effect of the enhanced bandwidth on localization accuracy. To this effect, we measure the localization errors observed when the bandwidth is 2 MHz (just 1 BLE channel), 20 MHz, 40 MHz and 80 MHz. We plot the median errors as a function of frequency in Fig. 10. The error bars are standard deviation.

As can be seen in the figure, the localization error decreases as the available bandwidth increases. The median errors for the 4 frequencies are 86 cm, 110 cm, 134 cm, 160 cm respectively. Observe that for a bandwidth of just 2 MHz, which is equivalent to just 1 BLE channel, the localization error is really high (almost 2 times that of 80 MHz). This shows the importance of combining information across multiple BLE channels. Without BLoc's combination across frequencies, the ability of BLoc to resolve multipath is greatly limited and as a result, the localization error increases significantly.

## 8.6 Interference Avoidance

BLE co-exists with Wi-Fi in 2.4GHz frequency band and is prone to interference from Wi-Fi. As a result, BLE can sometimes blacklist certain channels that won't be used for BLE transmissions. How does BLoc cope with missing CSI information on these BLE

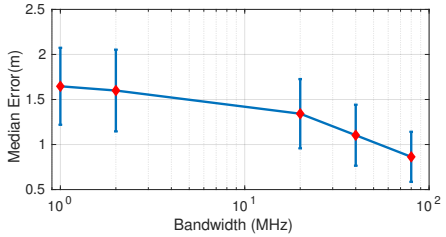



**Figure 10—Effect of Bandwidth:** Shows how the median localization error for BLoc varies with increasing the possible bandwidth of hops for BLE

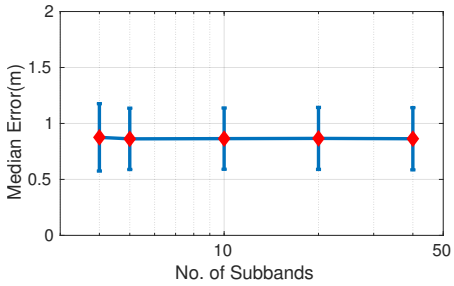



**Figure 11—Interference Avoidance:** Shows how the median localization error for BLoc varies with limiting the number of possible available subcarriers wihtout contention over 80MHz bandwidth for BLE

channels? Note that this issue is different from reduced bandwidth because the bandwidth is not being reduced in this case. Its just that there are gaps in available frequency bands. To evaluate this effect, we subsampled the available BLE channels by a factor of 2 and by a factor of 4 and compute BLoc's localization accuracy using subsampled data. The median errors are plotted in Fig. 11.

As can be seen in the figure, subsampling the available channels has almost no effect on the localization accuracy. This observation is also backed by theoretical understanding. The span of frequencies available determines the resolution of the system (i.e., more bandwidth implies higher accuracy). However, the gaps in the frequency bandwidth determine the aliasing, i.e., if we have 4 MHz gaps in adjacent frequency bands, then the system will be unable to differentiate between distances separated by 75m (speed of light/4 MHz). Even for gaps as large as 20 MHz (one Wi-Fi channel), the aliasing distance is 15 m. Since most indoor environments are less than 15 m large, such aliasing does not effect the accuracy of the system. The slight reduction in accuracy due to sub-sampling is thus attributable to lower SNR caused by sub-sampling.

## 8.7 Multipath Rejection

Further, we evaluate the effectiveness of the multipath rejection algorithm proposed in section 5.4. We turn-off the multipath rejection algorithm from the pipeline. We replace the multipath rejection with a naive baseline that just picks the shortest distance path as the direct path. The cdf of the localization error is plotted in Fig. 12. Note that, this measurement uses 4 anchors with 4 antennas each. Furthermore, it uses all the 40 available BLE channels. As can be seen in the figure, the median error increases from 86 cm to 195 cm (a factor of 2X), while the 90th percentile increases from 178 cm to 331 cm. This clearly shows that the multipath rejection algorithm is crucial to the accuracy of BLoc. This multipath rejection is enabled

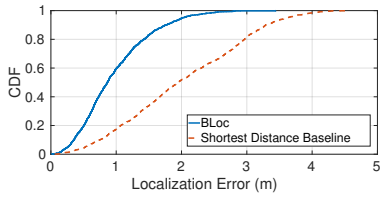



**Figure 12—Effect of Multipath Rejection:** BLoc's novel multipath rejection scheme improves the accuracy by a factor of 2.

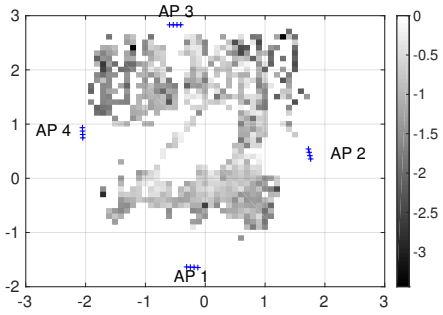



**Figure 13—Correlation of accuracy with Location:** BLoc's accuracy variation with the location of the BLE tag across the environmental setup.

by the high bandwidth obtained by channel combination (described in section 5.1) and phase offset removal (section 5.2).

## 8.8 Location Dependency

Finally, we looks at how the variation in the location of the BLE tag within the environment effects the accuracy of the BLoc's RMSE. In figure 13 we plot the RMSE values at different locations of the BLE tag within the environment. We can observe that the errors is particularly high in the corner locations of the setup, which can be attributed to the closely spaced values of the sinusoid at near 90 deg angles. Apart from that, we see that there is no consistent pattern observed showing that the accuracy of the BLocis not dependant on the location of the BLE tag.

# 9 RELATED WORK

## 9.1 Context for BLE Localization

Academia and industry have worked on wireless localization since the inception of wireless communication, for various wireless communication protocols (WiFi, satellite signals, LTE, passive RFID, active RFID, Bluetooth) [5, 8, 18, 21, 23, 25, 35–38, 41, 42, 45]. The primary motivation has been that several devices use one or more of the wireless protocols for communication and reusing the communication protocol for localization makes it easy to add on.

Wi-Fi based localization has been a long studied topic and has seen great advances in recent years. Wi-Fi based systems started with using RSSI [5, 8], but have moved to CSI-based localization in recent years [21, 23, 35, 42] due to improved accuracy of CSI based systems, their ability to combat multipath and no requirement for fingerprinting. Such CSI-based systems have been able to locate off-the-shelf devices with sub-meter accuracy. But, WiFi communication and thereof localization require high power. With the advent of Internet of Everything, low power communication is needed to enable long lasting battery powered devices. Therefore, these devices cannot use WiFi for communication. To mitigate the high power challenge, Bluetooth (BLE), passive RFID, active RFID, backscatter communication with WiFi have been proposed [17, 19, 20, 48].

In the low power localization domain, passive RFID's provide zero power, short distance communication/identification protocol. They are deployed in large scale factories where cheap inventory management is required. RFID localization has seen several innovations in using phase measurements for localization [25, 26, 38, 39, 46]. However, because of their zero power nature, RFIDs are low range and have known to be unreliable. Furthermore, RFIDs require dedicated RFID readers for localization and scanning.

Another popular communication paradigm recently evolved is using low power backscatter radios to backscatter ambient signals which can be decoded by existing infrastructure like WiFi, therefore providing low power Internet connectivity without needing any new infrastructure [17, 19, 20, 48]. We refer to this as backscatter on WiFi, which provides similar attributes of low power and short range communication. Recently, [22] has shown high accuracy localization for such communication systems. However, like passive RFIDs, backscatter on WiFi suffers from the low range of communication.

In contrast to passive RFID's and backscatter on WiFi, active RFIDs (100m) or Bluetooth (20-30m) provides medium range communication as they have active radios on them, while maintaining low power. Active RFID's typically are deployed outdoors; for example, highway toll booth use active RFID for communication, thereby provided an opportunity to localize and count cars using their EZ-pass, again reusing existing communication infrastructure for localization [2].

In indoor environments, BLE tags are the methods of choice [9, 10, 33, 34]. They provide sufficiently long range indoors, are resistant to frequency selective fading and have low-power operation. BLE tags are readable by off-the-shelf smartphones and access points, because of their co-existence in the 2.4 GHz Wi-Fi band. BLE tags are, therefore, getting very popular for tracking operations in homes, factory floors, etc. Google's vision for physical web is based on extensive deployment of BLE beacons [13, 14]. It is in this context that localization for BLE devices becomes crucial. The goal of BLoc is to improve localization accuracy for the BLE tags that increasingly form a part of our daily lives. In summary, each communication protocol has different applications and different deployment scenarios, therefore localization of each protocol is important. BLoc is geared to advance the Bluetooth localization towards the indoor application scenario.

## 9.2 BLE Localization

Past work on Bluetooth localization has significantly relied on using RSSI as the input[7, 40]. Similar to RSSI for Wi-Fi localization, this work either relies on extensive fingerprinting or is inaccurate. Its also prone to multipath effects and changes in the environment. The most recent work on Bluetooth localization in [7] provides median localization accuracy of 1.2 meter using RSSI. However, it requires finger printing of the environment, therefore it needs to

be trained for every new environment and retrained every time the environment changes.

Similar to the transition of localization algorithms for Wi-Fi, BLoc shifts this paradigm of bluetooth localization to use channel state information to perform localization using geometry (triangulation and trilateration), therefore requiring no training for every new environment. BLoc builds a novel algorithm to recover channel state information from bluetooth transmission. However, Bluetooth localization accuracy suffers due to low bandwidth. BLoc also presents a novel technique to provide bandwidth expansion by combining channel state information across multiple band, while requiring no change to the Bluetooth module. By combining multiple bands, Bluetooth achieves comparable resolution in time or distance measurement. BLoc open the doors to apply wideband channel state information based localization algorithms which were developed for WiFi localization to the Bluetooth communication protocol.

Finally, we do note that BLoc requires deployment of new anchors in the environment which can measure CSI, whereas the RSSI based system could use smartphones to gain RSSI information. But, we believe this is an essential first step towards enabling zero startup cost Bluetooth localization which requires no training and is based on channel state information.

### 9.3 RF-based Localization

In terms of the algorithms presented in the paper, three systems are close to our work. First, [21] uses channel state information available for 40 MHz Wi-Fi bands on multiple access points to measure both distance and angle to a target device. We presented a simplified version of [21] in background section (section 2). [35] uses stitching of multiple Wi-Fi channels to get a wide bandwidth, but requires CSI measurement on the target itself, which is infeasible for BLE tags. [27, 43, 44] use relative channels, but require synchronization across multiple access points.

In contrast to these systems, BLE does not provide wide bandwidth or access to CSI. We present novel algorithms to stitch multiple channels, measure CSI and cancel phase offsets using anchor collaboration. Furthermore, our approach to solving multipath using relative distances and spatial entropy differs from these systems.

Finally, we believe the techniques to use CSI across multiple anchor points to cancel phase offsets and the technique to use spatial entropy for multipath cancellation are applicable beyond BLE and can benefit general localization systems.

## 10 CONCLUSION

We present, BLoc, a CSI-based localization system for BLE tags. BLoc includes novel algorithms to compute CSI for BLE packets, to increase bandwidth of BLE signals by combining the frequency hops and to isolate the direct path from multipath reflections. By doing so, BLoc achieves sub-meter localization accuracy in a real world environment. We believe BLoc will open new avenues for localization of tens of millions of already deployed BLE tags. Furthermore, we hope that BLoc will serve as a tool for the research community to test out CSI-based localization algorithms for BLE devices.

**Acknowledgements**– We thank anonymous reviewers and our shepherd, Kate Lin, for their insightful comments and feedback. We thank Austin Duffield for help with initial implementations and Hariharan Rahul for his help with running the experiments.

## REFERENCES

[1] VICON T-Series. www.vicon.com/products/documents/Tseries.pdf.
[2] O. Abari, D. Vasisht, D. Katabi, and A. Chandrakasan. Caraoke: An e-toll transponder network for smart cities. In *ACM SIGCOMM*, 2015.
[3] I. Afyouni, C. Ray, and C. Claramunt. Spatial Models for Context-Aware Indoor Navigation Systems: A Survey. JOSIS, 2012.
[4] Apple. iBeacon. https://developer.apple.com/ibeacon/.
[5] V. Bahl and V. Padmanabhan. RADAR: An In-Building RF-based User Location and Tracking System. INFOCOM, 2000.
[6] L. Chang, X. Chen, Y. Wang, D. Fang, J. Wang, T. Xing, and Z. Tang. Fitloc: Fine-grained and low-cost device-free localization for multiple targets over various areas. *IEEE/ACM Transactions on Networking (TON)*, 2017.
[7] D. Chen, K. G. Shin, Y. Jiang, and K.-H. Kim. Locating and tracking ble beacons with smartphones. In *CoNEXT*, 2017.
[8] K. Chintalapudi, A. Padmanabha Iyer, and V. N. Padmanabhan. Indoor Localization Without the Pain. MobiCom, 2010.
[9] Chipolo. Chipolo Classic. https://chipolo.net/.
[10] Estimote. Estimote. https://estimote.com/.
[11] Ettus Research. Universal Software Radio Peripheral N210. https://www.ettus.com/.
[12] C. Gomez, J. Oller, and J. Paradells. Overview and evaluation of bluetooth low energy: An emerging low-power wireless technology. *IEEE Sensors*, 2012.
[13] Google. Project Eddystone. https://developers.google.com/beacons/.
[14] Google. Web Bluetooth. https://webbluetoothcg.github.io/web-bluetooth/.
[15] Grand View Research Inc. Bluetooth Beacon Market Worth $58.7 Billion By 2025, 2017. https://www.grandviewresearch.com/press-release/global-bluetooth-beacons-market.
[16] A. Henkin, Y. Shaham, and I. Brickner. Contextual advertising techniques for implemented at mobile devices, July 18 2017. US Patent 9,710,818.
[17] V. Iyer, V. Talla, B. Kellogg, S. Gollakota, and J. Smith. Inter-technology backscatter: Towards internet connectivity for implanted devices. In *SIGCOMM*, 2016.
[18] K. Joshi, S. Hong, and S. Katti. PinPoint: Localizing Interfering Radios. NSDI, 2013.
[19] B. Kellogg, A. Parks, S. Gollakota, J. R. Smith, and D. Wetherall. Wi-fi backscatter: Internet connectivity for rf-powered devices. In *ACM SIGCOMM Computer Communication Review*, 2014.
[20] B. Kellogg, V. Talla, S. Gollakota, and J. R. Smith. Passive wi-fi: Bringing low power to wi-fi transmissions. In *NSDI*, 2016.
[21] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti. SpotFi: Decimeter Level Localization Using Wi-Fi. SIGCOMM, 2015.
[22] M. Kotaru, P. Zhang, and S. Katti. Localizing low-power backscatter tags using commodity wifi. In *CoNEXT*, 2017.
[23] S. Kumar, S. Gil, D. Katabi, and D. Rus. Accurate Indoor Localization with Zero Start-up Cost. MobiCom, 2014.
[24] J. K. Y. Lau, J. P. Bruno, et al. Location and contextual-based mobile application promotion and delivery, 2018. US Patent 9,936,333.
[25] Y. Ma, N. Selby, and F. Adib. Drone relays for battery-free networks. In *SIGCOMM*.
[26] Y. Ma, N. Selby, and F. Adib. Minding the billions: Ultra-wideband localization for deployed rfid tags. In *MobiCom*, 2017.
[27] D. Musicki and W. Koch. Geolocation using tdoa and fdoa measurements. In *Information Fusion, 2008 11th International Conference On*, pages 1–8. IEEE, 2008.
[28] Q. Pu, S. Gupta, S. Gollakota, and S. Patel. Whole-home Gesture Recognition Using Wireless Signals. MobiCom, 2013.
[29] M. Ros, J. Boom, G. d. Hosson, and M. D'Souza. Indoor Localisation Using a Context-Aware Dynamic Position Tracking Model. International Journal of Navigation and Observation, 2012.
[30] S. Sen, J. Lee, K.-H. Kim, and P. Congdon. Avoiding Multipath to Revive Inbuilding Wi-Fi Localization. MobiSys, 2013.
[31] E. Soltanaghaei, A. Kalyanaraman, and K. Whitehouse. Multipath triangulation: Decimeter-level wifi localization and orientation with a single unaided receiver. In *MobiSyS*, 2018.
[32] The Lake Companies Inc. Beacon-Trak. http://www.lakeco.com/lake-products/beacon-trak/.
[33] Tile. Tile Mate. https://www.thetileapp.com/.
[34] TrackR. TrackR Pixel. https://secure.thetrackr.com/.
[35] D. Vasisht, S. Kumar, and D. Katabi. Decimeter-Level Localization with a Single Wi-Fi Access Point. NSDI, 2016.
[36] J. Wang, F. Adib, R. Knepper, D. Katabi, and D. Rus. RF-compass: Robot Object Manipulation Using RFIDs. MobiCom, 2013.
[37] J. Wang, H. Jiang, J. Xiong, K. Jamieson, X. Chen, D. Fang, and B. Xie. LiFS: Low Human-effort, Device-free Localization with Fine-grained Subcarrier Information. MobiCom, 2016.

[38] J. Wang and D. Katabi. Dude, Where's My Card?: RFID Positioning That Works with Multipath and Non-line of Sight. SIGCOMM, 2013.
[39] J. Wang, D. Vasisht, and D. Katabi. Rf-idraw: Virtual touch screen in the air using rf signals. *ACM SIGCOMM*, 2014.
[40] Y. Wang, Q. Ye, J. Cheng, and L. Wang. Rssi-based bluetooth indoor localization. In *2015 11th International Conference on Mobile Ad-hoc and Sensor Networks (MSN)*, 2015.
[41] Y. Xie, J. Xiong, M. Li, and K. Jamieson. xd-track: leveraging multi-dimensional information for passive wi-fi tracking. In *HotWireless*, pages 39–43. ACM, 2016.
[42] J. Xiong and K. Jamieson. ArrayTrack: A Fine-grained Indoor Location System. NSDI, 2013.
[43] J. Xiong, K. Jamieson, and K. Sundaresan. Synchronicity: Pushing the envelope of fine-grained localization with distributed mimo. In *HotWireless*, 2014.
[44] J. Xiong, K. Sundaresan, and K. Jamieson. ToneTrack: Leveraging Frequency-Agile Radios for Time-Based Indoor Wireless Localization. MobiCom , 2015.
[45] C. Xu, B. Firner, Y. Zhang, R. Howard, J. Li, and X. Lin. Improving RF-based Device-free Passive Localization in Cluttered Indoor Environments Through Probabilistic Classification Methods. IPSN, 2012.
[46] L. Yang, Y. Chen, X.-Y. Li, C. Xiao, M. Li, and Y. Liu. Tagoram: Real-time tracking of mobile rfid tags to high precision using cots devices. MobiCom, 2014.
[47] Z. Yang, Z. Zhou, and Y. Liu. From rssi to csi: Indoor localization via channel response. *ACM Comput. Surv.*, 2013.
[48] P. Zhang, D. Bharadia, K. Joshi, and S. Katti. Hitchhike: Practical backscatter using commodity wifi. In *SenSys*, 2016.