OwLL: Accurate LoRa Localization using the TV Whitespaces

Atul Bansal
Carnegie Mellon University
atulb@andrew.cmu.edu

Akshay Gadre
Carnegie Mellon University
agadre@andrew.cmu.edu

Vaibhav Singh
Carnegie Mellon University
vaibhav3@andrew.cmu.edu

Anthony Rowe
Carnegie Mellon University
agr@ece.cmu.edu

Bob Iannucci
Carnegie Mellon University
bob@sv.cmu.edu

Swarun Kumar
Carnegie Mellon University
swarun@cmu.edu

ABSTRACT

LoRa is a popular Low-Power Wide-Area Networking (LPWAN) technology that allows devices powered by a ten year AA battery to connect to radio infrastructure miles away. One of the most promising features of LoRa is the ability to track the location of radios from a distance, enabling applications ranging from inventory tracking, smart infrastructure monitoring and structural health sensing. Yet, state-of-the-art LoRa localization systems experience errors of several tens or even hundreds of meters in location tracking, owing to the narrow bandwidth and limited battery life of LoRa devices.

This paper presents OwLL, a LoRa localization system that limits location error to few meters with commodity LoRa clients in a wide-area network. Our key innovation is the development of a distributed base station network made of individually low-cost components that together span a wide bandwidth that encompasses the TV whitespaces and offers high aperture, crucial to localization accuracy. We demonstrate how this network can aggregate signal measurements made across multiple different narrowband channels of a LoRa client to triangulate it at fine accuracy. We implement and evaluate OwLL on a testbed spanning a large university campus centered in a major U.S. city and demonstrate a 9 m (across line-of-sight and non-line-of-sight) median error in localization.

CCS CONCEPTS

• **Networks** → **Wireless access points, base stations and infrastructure**; *Location based services*; • **Computer systems organization** → *Sensor networks.*

KEYWORDS

LPWAN, Localization, Sensor Networks, TV whitespaces

1 INTRODUCTION

Low Power Wide Area Networks (LP-WANs) are touted to provide long range ubiquitous connectivity for everyday devices with a AA battery. One such LP-WAN technology is LoRa that aims to enable IoT devices to send data at a few kbps to a base station several kilometers away while maintaining a 10 year battery life. An important vision with which LoRa was designed was to provide accurate asset tracking solutions with high accuracy in urban environments enabling many applications such as wide-area inventory tracking, infrastructure monitoring, and structural health sensing. Yet, there remains a gap in the literature for a LoRa localization system which can provide enough accuracy to realize this vision.

While there have been several attempts in both academia and industry to build LoRa localization solutions [17, 27, 34], they have achieved accuracy in tens or even hundreds of meters, depending on topography and testbed size. A key culprit behind this low accuracy is LoRa's narrow bandwidth–125 KHz–which only allows for range resolution of more than a kilometer (with the traditional c/B resolution being 2.4 km). Indeed, this low bandwidth rules out most types of localization approaches which have worked well for other technologies – time [45], time-difference of arrival [32, 49] and phase-based location tracking systems [21, 22]. In addition, traditional RSSI or fingerprinting systems [32] prove challenging to deploy at high accuracy in the wide-area context due to channel dynamics, multipath and the large area of coverage.

IPSN '21, May 18–21, 2021, Nashville, TN, USA

ACM ISBN 978-1-4503-8098-0/21/05.
https://doi.org/10.1145/3412382.3458263

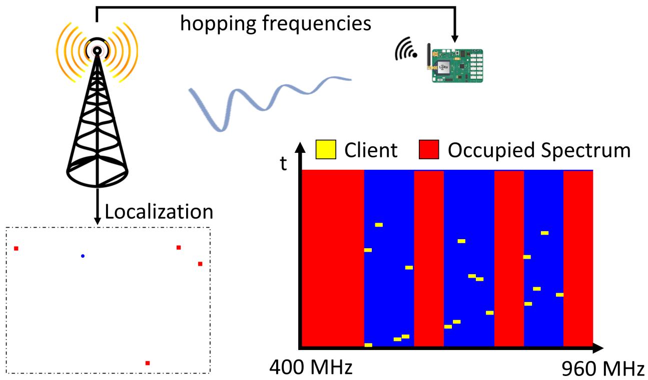


Figure 1: OwLL enables accurate outdoor LoRa localization using TV whitespace bands. We use a TDOA based approach where a LoRa client hops multiple frequencies spanning a wide bandwidth at TV whitespaces and ISM bands

This paper presents OwLL[1] – a LoRa localization system which can provide high localization accuracy (~9 meters) over long range using off-the-shelf commodity LoRa radios. Our approach circumvents the low-bandwidth of LoRa by stitching together channel responses at multiple TV whitespace and ISM frequency bands. We show how we can piece together samples of wireless spectrum available in these frequencies to improve the effective range resolution of LoRa and achieve resilience against multipath. We deploy a network of software radio base stations that operate on the TV whitespaces based on FCC experimental licenses available to us. Future commodity LoRa base stations which support ISM and TV whitespace bands would not require these experimental licenses. Our system deployment achieves 9 meters median accuracy across both line-of-sight and non-line-of-sight contexts which is a considerable improvement over state-of-the-art LoRa localization approaches (~ 50m)[27].

OwLL's fundamental approach is to give the illusion that the infrastructure is a network of wideband and perfectly phase synchronized base stations whose antennas are geo-distributed to collectively span a wide area. This effective geo-distributed antenna array can now measure signals received from any low-power LoRa client to measure phase difference across its antennas, which loosely can be translated to time-difference-of-arrival when measured across multiple frequencies spanning the wide bandwidth. The time-difference-of-arrival can then be measured across the wideband geo-distributed antenna array to trilaterate the LoRa device in the cloud. In order to realize such a system, we have to overcome two critical challenges: (1) We need to create the illusion of the presence of wideband radios using low-cost client and base station hardware; (2) We need to develop a distributed phase synchronized array of base stations to mitigate the impact of signal multipath, especially blockages in city environments.

Emulating Wide Bandwidth in LP-WAN: A OwLL client performs frequency hopping on new TV whitespaces so that the channels can be stitched together at the base station. However, designing base stations that listen across wide bandwidths of the order of a GHz would significantly increase cost. A more practical solution is to make narrowband base stations hop along with the client to frequencies on which the client transmits. This fits in very naturally with the design of LoRa as the base stations can feedback this frequency information to the client during acknowledgement packets. While the notion of stitching many narrow band frequencies to emulate wide bandwidth has been explored in other contexts (e.g. WiFi [45] and RF backscatter [30]), the LP-WAN context brings unique challenges. Specifically, LoRa packets are extremely narrowband (125 kHz), long (up to seconds each) and energy consuming, meaning that the act of hopping through all frequencies in the whitespaces is both time and energy-draining.

To tackle this problem OwLL performs an extensive outdoor experimental study and discovers that sampling LoRa packets at small number of carefully chosen frequencies results in highly accurate localization. We believe this stems from the fact that even while multipath outdoors is rich, it tends to have a sparse number of dominant paths that can be discovered from sparse sampling of the spectrum [13]. We then develop a search space exploration system running in the cloud that iteratively instructs clients to hop frequency bands and actively learns which frequency bands to next explore and maximize localization accuracy. We demonstrate that our approach achieves significant savings in time and energy, with minimal impact on localization accuracy. We further show how our system is designed to be compatible with FCC regulations on access to TV whitespaces and respects the presence of incumbents.

Emulating Distributed Arrays in LP-WAN: Our approach leverages the aperture provided by wide bandwidth to accurately localize clients outdoors. A key challenge is signal multipath which could often cause the line-of-sight path to disappear. While some multipath resilience is offered by the wide bandwidth of our system, dealing with blocked line-of-sight paths in urban contexts is a major challenge.

Our observation is that while the line-of-sight paths could often be blocked, they are unlikely to completely disappear across a large number of antennas, each spanning a wide aperture in terms of bandwidth at sub-GHz frequencies. In addition, certain large, dominant reflectors in the environment (e.g. buildings) are likely to affect multiple clients in a

[1]Outdoor whitespace-band LoRa Localization

predictable manner that allows them to be weeded out of consideration. We present a particle filter inspired optimization problem that explores consistency in the wireless channels measured from diverse locations to weed out reflected signal paths and identify the direct path to at least some locations in the environment. We further tackle various challenges pertaining to efficient phase synchronization of LoRa base stations to emulate distributed arrays.

Limitations: (1) OwLL targets localization of quasi-static devices and is not designed for moving objects owing to the need for clients to frequency hop and the slow data rates of LoRa. This should not pose a problem for most LoRa deployments that are quasi-static (e.g. sensor networks). (2) The evaluation of this paper is restricted to localization in 2-D space and primarily outdoors, where social distancing is easier. This is due to COVID-19 related university campus access restrictions that limited access to multiple buildings and high vantage points. (3) Like most wireless localization systems, our system struggles under extreme occlusion, e.g. deep indoors or underground, where all line-of-sight paths are completely obscured across frequencies.

Implementation and Evaluation: We implement OwLL on USRP N210 base stations and off-the-shelf Semtech SX1262 radios as LP-WAN clients. These clients enable us to hop across a wide spectrum ranging from 500 MHz to 960 MHz. We hold experimental FCC licenses to operate on the TV whitespace bands across all our hardware. We evaluate the system on a testbed spanning 66,000 sq.m. at Carnegie Mellon University campus in Pittsburgh. We also evaluate our system in the presence of occlusions due to buildings, trees and topography that obscure line-of-sight. Our results reveal the following:

- A median error in time-difference of arrival of 3.6 and 14.8 meters respectively in line-of-sight and non-line-of-sight environments
- A median error in localization of 3.9 and 15.7 meters respectively in line-of-sight and non-line-of-sight environments
- Time and energy savings of 67% compared to exhaustive sampling of available frequencies.

Contributions : This paper's contributions are as follows:

- The development of the first accurate localization system in whitespace frequencies that operates on commodity LoRa devices and achieves an accuracy of 9 meters.
- An approach that treats diverse base stations as a wideband distributed array, with minimal overhead in terms of battery life of client devices.
- A detailed experimental evaluation on a large neighborhood of a city, demonstrating high performance.

2 RELATED WORK

Low-Power Wide-Area Networks: Recent years have witnessed accelerated growth in the field of Low Power Wide Area Networks (LP-WANs) [14] encompassing various implementations and protocols, both proprietary and open source, supported by various hardware platforms. While cellular providers use standards like LTE-M [16] and NB-IoT [35] for low-power IoT communication in licensed spectrum, companies like Semtech [1, 40] and SigFox [37] focus more on utilizing unlicensed spectrum including the available channels in the TV whitespaces [12], thereby improving the spectral availability. Hence, even though the former offers better data rates, MAC sophistication, and better features for routing, firmware broadcast, etc., the latter offers a cheaper and a battery efficient solution for larger city scale deployments.

Localization: As the number of deployed IoT clients has increased, there has been a rising interest in leveraging LPWAN technologies for real time tracking and tagging of objects. As localization accuracy is a pre-requisite for these applications, there have been extensive research done in the field of localization – both indoor and outdoor systems, satisfying different constraints on power, range, cost, multi-path resilience, etc. Based on these constraints, a simple classification on RF-localization systems can be done. This is captured in Table. 1. A key observation across technologies is that there remains a gap for a solution that is simultaneously accurate, long-range (i.e., operates over hundreds of meters) and low-power (i.e., supports a 10-year battery life).

Long range systems with high accuracy: Till date, global positioning systems (GPS) dominate in the category of existing long range localization systems in terms of accuracy. However, a key disadvantage of GPS is its power consumption. Military grade GPS systems are expensive with high power consumption – restricting its reach across smaller devices. Also, off-the-shelf GPS chips which consume about the same power as Semtech LoRa devices would incur extra energy and hardware costs if deployed on an already energy constrained IoT device leading to significant drop in battery life. Cost-effective positioning systems involving Bluetooth [7, 19, 26] and WiFi [22, 45, 47] provide low power localization with centimeter level accuracy, but are constrained in their range. Systems for sensor positioning [15, 38], wildlife tracking [5, 29] and LP-WAN based systems [17, 24, 25, 27], operate over long distances at low-power outdoors, yet are highly susceptible to multipath effects, thereby losing accuracy in urban environments. Most LoRa based localization approaches have leveraged extensive RSSI + Fingerprinting Datasets [3, 9, 11] with some of them being restricted to theoretical evaluation [31, 42, 43] or are evaluated indoors[18]

Papers	Features			
	Class	Range	Low Power	Accuracy
[22, 45, 47]	WiFi	12-15m	-	< 50 cm
[7, 19, 26]	Bluetooth	<10m	✓	≈ 100 cm
[23, 41]	Cellular	35-60m	X	≈ 85 cm
[4, 15]	Sensor-based	≈ 50m	✓	≈ 3-5 m
[29, 46]	Wildlife Tracking	500m+	✓	15-24m
[28, 30, 48]	Back-scatter	tens of m	✓	< 50 cm
[2, 33]	Prior LP-WAN	500 m+	✓	100s of m
OwLL	LP-WANs	500 m+	✓	≈ 9 m

Table 1: Comparison of Related Work

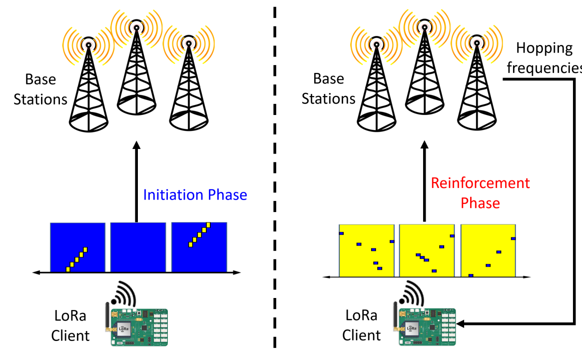


Figure 2: OwLL iteratively requests client transmissions at specific frequencies to localize accurately at low power.

or outdoors in very controlled setting[36]. Using only timestamping or RSSI based approaches [2, 33], demonstrate relatively high errors in estimating the location of the client.

Indeed, there remains a gap for accurate RF-based positioning that spans outdoor spaces that is suited to low-power wide-area networks.

3 OVERVIEW OF OWLL

OwLL aims to achieve a few meters of localization accuracy using commodity LoRa radios. Our choice of system design is motivated by two important constraints: (1) First, we seek to be compatible with commodity LoRa clients. This will allow our system to be deployed at scale on existing LoRa deployments; (2) Second, we do not seek to build complex multi-antenna base stations, since these antennas would be bulky and expensive to deploy at sub-GHz frequencies. In other words, we seek to preserve the simplicity of LoRa infrastructure that has been key to its adoption.

Our constraint on the simplicity of base stations leads us to a natural design point – what if we can fuse multiple individual base stations together to operate as one virtual multi-antenna array. Specifically, we aim to develop algorithms that synchronize distributed base stations in time, frequency and phase. We further allow these base stations to hop frequency bands without losing synchronization – effectively behaving as a wide band distributed array. We then fuse signal measurements received on multiple frequency bands from a commodity LoRa client to triangulate its location. Our system will span frequencies starting from the TV whitespaces through the unlicensed 915 MHz ISM band. The resulting array has two desirable properties that are key to achieving high localization accuracy: (1) It emulates wide bandwidth that is key to accurate time-of-flight and therefore range estimates of the LoRa clients; (2) It achieves wide spatial aperture owing to the significant spatial separation between base stations, allowing for effective triangulation of LoRa clients.

The rest of this paper describes two key components of OwLL's design that relate to the above two properties: how do we achieve high bandwidth and coordinate distributed base stations?

Emulating Wide Bandwidth: We first describe our approach to emulate wide bandwidths by frequency hopping over available spectrum in the TV whitespace bands and 915 MHz ISM band. While there has been much literature on frequency stitching in Wi-Fi, LoRa packets often last much longer (~seconds) and are energy constrained, meaning that we would need to minimize this hopping. We also consider the policy implications as well as the choice of frequencies in the TV whitespaces needed to perform effective frequency hopping to optimize for multipath resilience. Sec. 4 describes our solution to these challenges.

Emulating Distributed Arrays: Next, we describe our solution to deal with signal multipath, a common problem in localization systems in general. We focus especially on blockages of the direct path in urban settings – demonstrating how the distributed location of base stations, as well as an effective modeling of multipath due to dominant reflectors, can help mitigate this problem to some degree. We also consider challenges pertaining to phase synchronizing these base stations. Sec. 5 describes our approach.

4 EMULATING WIDE BANDWIDTH

The core idea of OwLL is to combine wide swaths of bandwidths available at the TV whitespaces to improve localization accuracy of LoRa. We use phase information from the I/Q samples received at 2 different base stations across all frequencies to estimate the Time Difference of Arrival (TDOA) between them, which in turn provides the location of the client if extended to multiple such pairs. While the notion of stitching together frequencies has been explored in prior work in domains such as Wi-Fi [45], UWB [20] and RFIDs [28], the LP-WAN context brings forth unique challenges and policy implications. We discuss these challenges and our solutions to overcome them in this section.

4.1 Challenges unique to LP-WAN

The core challenge that makes frequency stitching challenging to implement on LP-WANs is that it can be quite expensive in terms of both time and energy. This is in contrast to Wi-Fi where hopping across all frequency bands can be efficiently performed with tiny packets over a few milliseconds.

LP-WAN Packets are Long: First, LP-WAN packets are much longer – often lasting seconds. This is not accidental – it is designed as such to provide low data rate connectivity at extended distances over which LP-WAN packets need to be decoded from extremely battery-starved clients. This means that simply transmitting short packets on any given frequency band is latency-intensive.

LP-WAN Clients are Narrowband: Another challenge that complicates the aforementioned problem is the fact that LP-WAN clients are narrowband. This once again is by design, due to the battery-constraints of Low-power clients. For instance, LoRa's bandwidth is 125 kHz – a factor of 160× smaller than that of Wi-Fi. A natural consequence of this low bandwidth is the huge number of frequencies that need to be hopped to span even a modest bandwidth (∼160 packets for 20 MHz). This scales to an inordinately large number of packets required to span the 400 MHz of frequency bands available in the TV whitespaces.

LP-WAN Clients are Energy-Starved: A third challenge that makes hopping difficult is that LP-WAN clients are energy-starved. Prior work has shown that data transmission, particularly transmission time is a critical power bottleneck in LoRa sensor networks [10]. This means that the simple act of hopping between frequency bands and transmitting short packets, when repeated over many frequencies can quickly lead to energy drain.

4.2 Can we Minimize Frequency Hopping?

The above challenges lead us to a simple conclusion: to perform localization in a both time and energy-efficient manner, OwLL must minimize the number of frequencies that need to be hopped. Thus, it is critical for OwLL to carefully select frequencies that provide the most bang for the buck – i.e. **maximize localization accuracy with the minimum number of frequencies hopped**. However, which subset of frequencies need to be hopped to achieve this may vary depending on the location of an LP-WAN radio and its environment. This leads us to a natural question: "how do we choose which frequencies to hop to if we do not already know where the radio is located in the first place?".

A Strawman Solution: Prior to answering this question, let us design a näive version of the frequency-hopping system that relies on all frequencies available to OwLL, and therefore can maximize localization accuracy. For simplicity, we consider the case of estimating the Time-Difference of Arrival (TDOA) between the signals received at two different LP-WAN base stations from an LP-WAN client. We note that should this time-difference be accurately calculated relative to four or more base stations, the location of the client can be accurately triangulated. We also assume in this section that the clocks of the two LP-WAN base stations are perfectly synchronized in time, frequency and phase – we explain how this can be achieved later in Sec. 5.1.

Mathematically, let us estimate the time-of-flight between two base stations that estimate wireless channels of $h_{1,i}$ and $h_{2,i}$ (where $i = 1, \ldots, N$) respectively on a set of N frequencies among $\{f_1, \ldots, f_n\}$ that are uniformly spaced. We then leverage the Bartlett algorithm [8], which effectively has the structure of the Discrete Fourier transform to compute $P(\tau)$, the power of the signal component received corresponding to a time-difference of arrival of τ:

$$P(\tau) = \sum_{i=1}^{n} h_{1,i} h_{2,i}^{*} e^{-2\pi f_i \tau}$$

We note that the above $P(\tau)$ function could peak at multiple time-difference of arrival values due to signal multipath – a topic we discuss in Sec. 5.2. Clearly however, the above formulation does assume that frequencies are uniformly separated and finely sampled. If they are not, this would cause artifacts such as spurious peaks that would appear in the $P(\tau)$ profile due to aliasing, when frequencies are not sufficiently separated. It should be noted that to maximize accuracy, one should choose as many frequencies that are available – both finely sampled and measured across a wide bandwidth. Unfortunately, this remains impractical in the real world for two reasons: (1) First, hopping all frequencies across 400 MHz

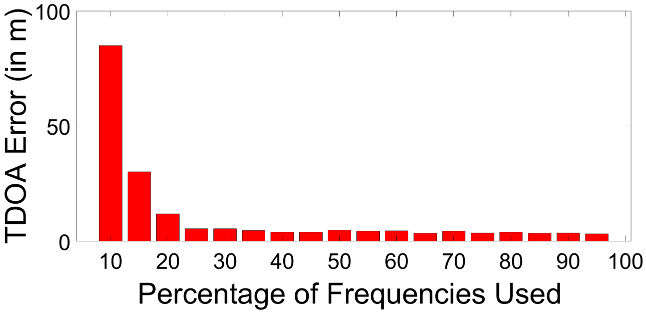


Figure 3: Carefully hopping over a small fraction (20%) of all frequencies leads to dramatic performance improvements.

bands would consume 3%[2] of client battery life and require several minutes to hours making it latency-impractical. (2) FCC rules only allow LoRa clients to utilize **unutilized** TV whitespace bands. This means across the wide bandwidth, OwLL client can only hop in a few frequency bands that change every day and must be constantly monitored.

Does Sub-sampling Frequencies Work? A trivial optimization on the above approach would then be to choose a subset of frequencies to hop from the expansive list of frequencies that may be available in the TV whitespaces. Of course, in this case, we should not directly use the Discrete Fourier Transform based Bartlett algorithm which would cause unnecessary artifacts in $P(\tau)$ due to the non-uniformity of frequencies sampled. Instead, we would use the Non-Uniform Discrete Fourier Transform (NDFT), an approach that extends the Fourier Transform to non-uniform frequencies and has been used in other localization contexts (e.g. Wi-Fi [45]). The flavor of the NDFT algorithm we use that is LoRa compatible is summarized in Algm. 1. Inherently, NDFT removes artifacts by making an assumption that the number of dominant multipath peaks in $P(\tau)$ is sparse. Fortunately, prior studies specific to LoRa have shown that this is indeed the case [13] in urban outdoor testbeds, which means NDFT could indeed hold promise.

We now empirically evaluate based on field measurements in a large outdoor testbed (details of the testbed are described in Sec. 7) consisting of two LP-WAN base stations and a client moved across line-of-sight and non-line-of-sight locations.

We then choose a large number of random sub-samples of these frequencies containing a different number of them ranging from 10% through 90%. Note that the fraction is only in relation to available bands, and not all bands in the TV whitespaces (typically, about 30 MHz of total bandwidth inter-spersed between 500 to 800 MHz is available in our city).

[2]Available Battery Energy: 2900 mAh; 125 kHz channels in 400 MHz: 3200; Energy of a typical LoRa packet: 100 mAs; Battery spent: 3.06%

Algorithm 1: Algorithm to estimate TDOA

Input : $\tilde{\mathbf{h}}$: Measured relative channels at selected frequencies f_i
F: Non-Uniform DFT matrix such that $F_{i,k} = e^{-j2\pi f_i \tau_k}$ for some set of τ_k
α_{min} and α_{max}: range of sparsity parameter
α ϵ: $Convergence parameter$

Initialize $\gamma = \frac{1}{\|F\|_2^2}$;
while $\alpha < \alpha_{max}$ **do**
  Initialize $\mathbf{p_0}$ to a random value, $converged = false$, $t = 0$;
  while $\|\mathbf{p_{t+1}} - \mathbf{p_t}\|_2 > \epsilon$ **do**
    $\mathbf{p_{t+1}} = \mathbf{p_t} - \gamma F^*(F\mathbf{p_t} - \tilde{\mathbf{h}})$ with values $< \gamma\alpha$ zeroed;
    $t = t + 1$;
  $numberpeaks$ = Number of peaks in $\mathbf{p_{t+1}}$;
  if $numberpeaks \neq 1$ **then**
    Increase α;
  else
    return $\tau_{TDOA} = \underset{t}{argmax}\ \mathbf{p_{t+1}}$;

In each instance, we report the max accuracy across these sub-samples. Fig. 3 plots the accuracy achieved in estimating time-difference of arrival vs. the fraction of frequencies chosen (%) among the entire range of available frequencies.

Our experiments reveal two surprising results:

- *Sampling a small number of frequencies is sufficient*: First, we observe that sampling as few as 20% of all available frequencies yields localization error that is equivalent to the 75%ile error achieved across these locations when all frequencies are used. This shows that one can gain remarkable benefits in hopping latency and energy overhead should we somehow know which frequencies to hop in advance.
- *Optimal Frequencies change over time*: Second, we see that these so-called optimal frequencies do not remain static over time or across clients. We observe that changes in the environment or client locations can substantially change which frequencies are sampled. This tells us that a static approach (~ say uniformly sub-sampled) to choose these frequencies would not suffice.

Observation #1: Hopping a small fraction of available frequencies suffices for accurate localization with minimal loss, yet these *optimal* frequencies vary dynamically.

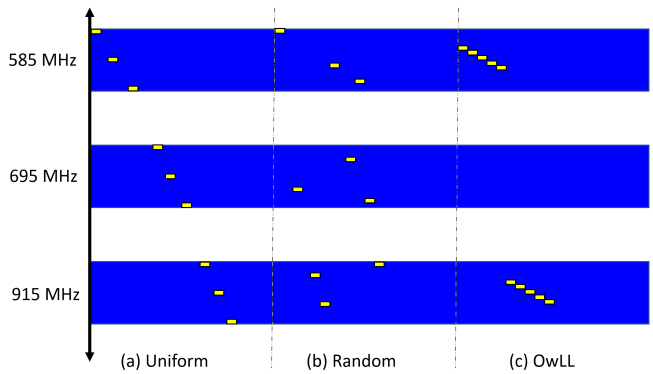


Figure 4: OwLL's approach to select initial frequencies.

4.3 OwLL's Frequency Hopping Design

Based on our observations in the previous section, our approach to localize a client in a time and energy-efficient manner hinges on selecting this small set of optimal frequencies instead of exhaustively sampling all possible frequencies. Our approach must naturally be dynamic – i.e. learn progressively which frequencies to hop based on prior measurements. There is a natural exploration vs. exploitation trade-off here however. To achieve better accuracy, we need to explore as many frequencies as possible. However, at some point, we do also need to terminate our algorithm to report time-difference of arrival (and in turn remain energy efficient). To counter this problem, we design an iterative maximum-likelihood algorithm which tries to progressively glean on an increasingly clear view of multipath propagation through the space to guide its choice of frequencies in an environment-invariant manner.

An Iterative Approach to Model Multipath: OwLL's algorithm effectively seeks to build a model of signal multipath based on measurements seen so far to then guide which measurements to take next. (see Fig. 2). Initially, OwLL begins with a small number (10% in our implementation) of initial frequencies among our available frequencies (see Observations). We then run the NDFT algorithm (Algm. 1) to obtain a sparse set of candidate TDOA values and their corresponding signal amplitudes and phases. Now, we identify the fact that given the amplitudes and phases of these TDOA peaks, there exists a frequency (f_{opt}) which, given this prior, would amplify the peaks the most. The best case scenario would be to transmit a packet at that f_{opt}, and only pick out the peaks that get amplified (the ones that did not are likely spurious). Unfortunately, this frequency f_{opt} is often not one of the frequencies that we are allowed to transmit on (e.g. due to incumbents or licensed users).

Thus, to select a frequency from within our spectrum constraints we define a goodness metric which measures how much a given frequency theoretically should amplify our prior belief of the TDOA. In other words, should this frequency be added to our pool of frequencies, it would maximize the amplitudes of TDOAs seen so far. We iterate over all available frequencies and select a few (5% in our implementation) that maximize goodness. We then explore those frequencies and validate whether the TDOAs discovered are truly amplified, new peaks emerge or spurious peaks are dampened. If the former occurs (to within a threshold), we terminate – otherwise, we repeat the algorithm with newly discovered peaks.

Mathematical Details: More formally, let us assume that the result of the NDFT algorithm based on the channels from the frequencies sampled so far h_{prev} are a set of time of flights τ and the corresponding amplitudes and phases $\mathbf{p}$. Should our NDFT algorithm be a good approximation of our observed channels, the wireless channel at any new frequency f should be given by:

$$h_{pred,f} = \sum_{\alpha \in \mathbf{p}, \tau \in \tau} \alpha e^{-2\pi f \tau}$$

We then define the goodness metric as the frequencies that would truly maximize the powers at the peaks previously observed, i.e.

$$\text{goodness}(f) = \sum_{\alpha \in \mathbf{p}_{f,new}} ||\alpha||^2,$$
$$\text{where } \mathbf{p_{f,new}} \leftarrow NDFT(h_{prev}, h_{pred,f})$$

We then pick the set of frequencies that maximize goodness(f) and stop when the goodness measured from observed channels closely aligns (or exceeds) that of the predicted channels. Algm. 2 summarizes our approach in greater detail.

Algorithm 2: Identify next frequency to be queried

$\vec{\mathbf{h}}_{prev}$,$\vec{\mathbf{p}}_{prev}$: Current channels and current estimate of TDOA
f_{opt} : Next Frequency to be queried
$\forall f \in$ Unoccupied Frequency Bands
  $h_{pred,f} = \sum_{\alpha \in \mathbf{p}, \tau \in \tau} \alpha e^{-2\pi f \tau}$;
  goodness$(f) = \sum_{\alpha \in \mathbf{p}_{f,new}} ||\alpha||^2$
 where $\mathbf{p_{f,new}} \leftarrow NDFT(h_{prev}, h_{pred,f})$;
return $f_{opt} = \underset{f}{arg\,max}$ goodness(f);

Compatibility with LoRa: We note that our system can be designed without modifications to the LoRa protocol on the client's end. OwLL can provide the clients an initial set of frequencies to hop with further suggestions on which frequencies to hop piggy-backed on acknowledgment packets. Note that all of this computation is performed at the resource-rich cloud connected to the base stations and not on the low-power client devices themselves.

Observations: A few points are worth noting:

- *Choice of Initial Frequencies:* A natural question one might ask is if the system could be misled by a poor choice of initial frequencies that mislead us with erroneous peaks. Further, the choice of these initial frequencies cannot be highly environment specific. To circumvent this, we require a minimum of extra 10% frequencies to be hopped to ensure that our data is not entirely dominated by our initial choices. Furthermore, we provide a very intuitive approach to choose initial frequencies which would be true for any environment. We ensure that this choice of frequencies allows us to fully span the desired range of TDOAs based on testbed geometry. However, choosing frequencies that are individually well separated could lead to aliasing, where a measured channel leads to multiple TDOA peaks due to 2π wraparound. This problem is particularly acute outdoors where the range of possible TDOA values can be large. However, this phenomenon can be easily avoided by ensuring the set of frequencies also include a few that are closely separated. How close these frequencies should be is easily calculated using the maximum TDOA value possible in the testbed geometry. We compare in Fig. 4 the different approaches one can take to decide initial frequencies, including our own which maximizes bandwidth while minimizing aliasing amongst the various TDOA values.
- *Energy and Time Overhead:* Our results in Sec. 8.3 show that in practice, we need to hop around 40 frequencies, requiring 20% ($\sim$ 20.97s) of time required by exhaustive sampling to achieve roughly 75th percentile error bound of our TDOA estimation, effectively increasing energy efficiency by 66%. We note that the time overhead is practical for most quasi-static LoRa deployments such as sensor networks, given the slow data transmission rate of the LoRa protocol. Our system is therefore limited to quasi-static LoRa localization.

4.4 TV Whitespace Policy Considerations

OwLL leverages 400 MHz of available bandwidth in the TV whitespaces and ISM bands between 500 to 960 MHz. Thus, it is critical to understand the implication of TV whitespace policy on OwLL.

Hardware Compatibility: OwLL exploits readily available LoRa hardware that is increasingly allowing for support on the TV whitespace bands. We use the SemTech SX-1262 client devices which operate in the TV whitespaces and have native support for a wide frequency range from 150 MHz to 960 MHz. While effectively they traverse upto 800 MHz of spectrum, FCC rules in US enable us to only leverage the vacant TV whitespace spectrum spanning from 500 to 960 MHz. Our base stations are USRP N210 software radios to allow for maximum flexibility in I/Q channel processing and we hold FCC experimental licenses to operate these USRPs in the TV whitespaces in the U.S. city where our network operates. We believe that with recent SemTech hardware allowing for I/Q channel measurement (e.g. SX1257), our approach can be extended to future commodity LoRa base station deployments with no license requirements whatsoever.

Policy and Impact on System: We note that recent policy changes to the TV whitespaces [39] allow dedicated powered-radios (e.g. base stations) to periodically (every day) check the shared whitespace database on behalf of the clients to check for and avoid incumbents. This eliminates the burden from low-power clients to coordinate with the database. It further aligns well with our design where base stations guide clients on which frequencies to hop. We note that the presence of whitespace incumbents may necessitate OwLL base stations to identify vacant frequency bands from the TV whitespaces every 12 hours and only run its optimization algorithm (Algm. 2) over available frequencies in the whitespaces.

5 EMULATING DISTRIBUTED ARRAYS

In this section, we describe the various challenges in building a localization system that uses distributed LP-WAN base stations that serve as a wideband and multi-antenna array. We discuss challenges such as synchronization and signal multipath.

5.1 Ensuring Phase Synchronization

Our discussion so far has assumed that base stations are synchronized in time, frequency and phase – however this is challenging to implement with geo-distributed base stations, some potentially indoors. To this end, we build on Chime [13], a prior approach that has demonstrated synchronization across base stations for a different problem – radio configuration. At a high level, this approach uses a known reference transmitter (typically a base station) that transmits a long LoRa packet concurrently at a frequency adjacent to the LoRa client. The algorithm then interpolates the wireless channel from both the reference and client across two different base stations at the common guard band between these frequencies. It then shows that if the location of the reference transmitter is known, a combination of four wireless channels (reference and client to each base station) can be used to infer time-difference of arrival to the desired client. We refer the reader to Chime [13] for more details on

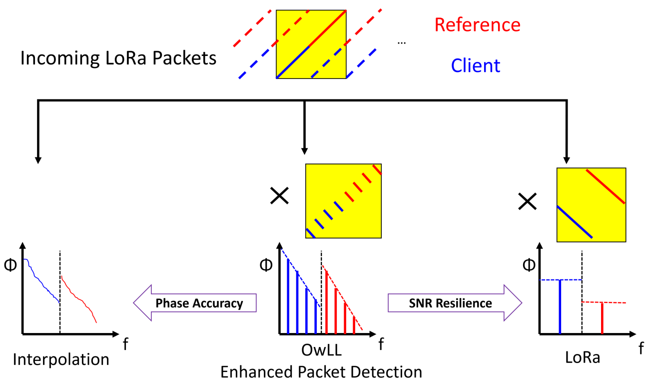


Figure 5: OwLL Enhanced Phase Estimation balances between SNR resilience and phase accuracy outdoors

the algorithm. While Chime was intended for radio configuration, we highlight two important modifications that are needed to extend it to OwLL's localization context:

Interpolation at the Guard Band: In performing interpolation at the guard band, OwLL must ensure that the phases of the wireless channel obtained are extremely stable for a static client in a quasi-static environment. Thus, OwLL designs the reference base station to sense and align its chirps at sample accuracy ($\sim 10\ \mu s$) by listening for the client's preamble and turning around to transmit on the adjacent band. Further, we note that we must only interpolate phases at specific points in the guard band when two chirps – one each of the reference and the client "meet" as highlighted in Fig. 5. This is important because chirps are sparse signals transmitting at different frequencies over time and the highlighted spot is one of the few where both transmitters truly agree in time and frequency (modulo time-of-flight, that we seek to measure).

One important consideration when performing this interpolation is the case of low-signal to noise ratio (SNR). In noisy settings, interpolation of the chirp signals (especially in phase) tends to loose robustness. To mitigate this, one approach could be to rely on correlation rather than interpolation. Specifically, one could correlate each received preamble symbol with the corresponding ideal up-chirp in both bands to compute phase values. Unlike interpolation of phase, correlation is known to be highly noise resilient. However, correlation would not allow for the ability to find the phase value at the guard band, meaning that the result would still be error-prone to timing offsets due to the frequency separation between the bands. OwLL therefore adopts a hybrid approach between these two extremes of interpolation and correlation. Rather than correlating with an ideal preamble up-chirp that spans all frequencies, we splice this chirp into smaller chunks that span smaller chunks of frequencies. We use the computed correlations to obtain multiple noise-resilient phase values. This allows us to then perform interpolation across these phase values at the guard band. Fig. 5 illustrates our enhanced phase estimation approach. Our results in Sec. 8.1 demonstrate the utility of this solution in achieving phase stability and noise resilience simultaneously.

Designing the Reference: We note that unlike Chime, the reference must hop alongside OwLL on an adjacent frequency based on the pre-determined schedule provided by the base stations. Should the adjacent frequency be unavailable (e.g. due to the presence of incumbents), the reference could transmit a non-interfering simultaneous LoRa packet on the same frequency using a different spreading factor, but at the expense of some resilience owing to the near-far effect. It is also prudent to choose the base station closest to the client as reference to ensure minimal detection delays as well as low TDOA values which are more noise resilient. Thus, before localizing any client, we run a quick calibration step where client transmits a packet to enable us to determine which base station to designate as the reference base station (coarsely, based on similarity of RSSI values). We elaborate in Sec. 5.3.

5.2 Impact of Multipath

Next, we make some experimental observations regarding signal multipath – a key challenge in any localization system. We specifically ask: "How frequently is the direct path between the clients and base station completely blocked in the urban LP-WAN context?". Given that our system has multiple base station locations and spans multiple frequency bands, we want to characterize the importance of both of these factors in truly discovering the TDOA corresponding to the direct path even in completely occluded settings.

Fig. 6 describes the accuracy of localization measured across the number of base station pairs based on our wide-area testbed described in Sec. 7. This result demonstrates that as we use more number of base stations pairs from the raw samples, we get better and better accuracy improving from 10.5 m for 3 pairs to 6.2 m for 6 pairs for the chosen clients.

We make the following broad conclusions based on these observations that guide our localization algorithm that follows:

- Adding more base stations to take multiple views of the object increases the likelihood that we check consistency across more base stations. Given that base stations are highly distributed, the likelihood that the line-of-sight across all of them is severely attenuated becomes low.

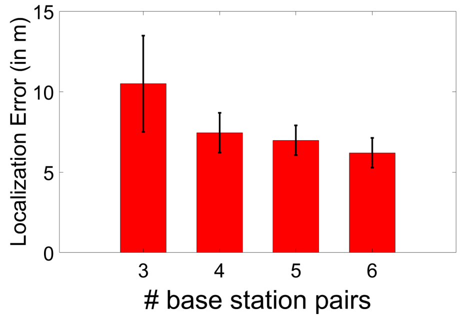


Figure 6: OwLL location accuracy across base station pairs

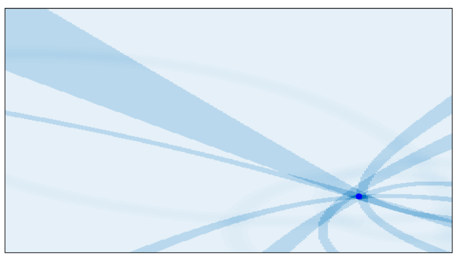


Figure 7: OwLL uses particle filtering atop the measured TDOA hyperbolas to estimate the location of the client

- A similar argument can be made about more diverse frequencies over the wide TV whitespaces bands, where frequency selective fading allows the direct path to survive in some frequency bands even while being attenuated at others.
- We note that the low frequencies of the LoRa band and TV whitespaces have an improved ability to penetrate through obstacles compared to higher frequencies. This in part contributes to some of our observations above.

Observation #2: Even in occluded settings, combining wireless channels from multiple base stations and frequencies can improve detection of the line-of-sight path.

5.3 Localization Design

Motivated by the above observations, we now seek to combine the diverse TDOA measurements made across base stations to obtain the true location of the device. While we could naïvely do so by trilateration across base stations, this misses two key additional advantages of our system: (1) First, the synchronization approach in Sec. 5.1 ensures that all base stations are phase-sychronized, which means that we could effectively treat the entire system as one holistic distributed array; (2) Second, trilateration may miss opportunities from considering obvious geometric constraints of the layout of our testbed. (3) Third, choosing a relatively proximate reference (based on RSSI) helps constrain the range of possible locations of the client even further (see Sec. 5.1). (4) Finally, OwLL can account for sources of known multipath from large, dominant reflectors that have previously been observed for other clients in the environment, if any.

A Particle Filter Based Search: Our approach to identify the true location of the device uses a particle filter [44] to weave in the geometric constraints of the system along with prior TDOA estimates as well as phase estimates from distributed base stations. We point to Fig. 7 that visualizes our solution. Our approach initially chooses a set of randomly chosen candidate locations in the space of interest uniformly sampled while respecting geometric constraints of the testbed. For each sample, we compute a goodness metric that is the sum of two components: (1) How well does this estimate fit the observed TDOA peak values (inverse of the mean squared error) across pairs of base stations? (2) How well does this location agree with the observed wireless channels across base stations and frequencies (we define this quantity in Algm. 3 below)? (3) Are the TDOA values consistent with the location of the chosen reference base station based on RSSI? (i.e., locations extremely far from the reference are unlikely) (4) We further add a penalty discounting consistent sources of multipath peaks resulting from large reflectors discovered in localizing prior clients (or reference base stations).

Based on these four observations, we assign probabilities to each region of space surrounding where the particles were originally obtained. We now re-sample the space while respecting geometric constraints, with the sampling biased by this new probability distribution. We repeat this process until the probability of a specific location within the geometric constraints surpasses a set threshold.

Mathematical Details: Algm. 3 describes the design of our particle filter and the goodness metric used to assign probabilities to the different particles.

Observations: A few points are worth noting:

- *Geometry of Base Stations:* Our approach is agnostic to the arrangement of base stations, however the precise layout can impact accuracy just as in any antenna array. In this paper, we assume that the relative geometry of the base station is given (e.g. based on logistic constraints) and leave to the precise optimization of this geometry to improve localization accuracy to future work.

Algorithm 3: Localization algorithm

Input : Base Station locations $\mathbf{BS_{i,j}}$ and Reference locations $\mathbf{R_{i,j}}$, $(\tilde{\mathbf{h}}, \mathbf{f})$ for all possible Base station pairs (BS^m, BS^n) from a client
Set of Geometric constraints - $(\mathbf{X}, \mathbf{Y})_{allowed}$, TDOA margin *mar*, Probability distribution variance σ, *Penalty*(x,y): penalizes dominant reflectors discovered when localizing other clients

Output : Estimated Client location (x_{opt}, y_{opt})

Initial Calibration:- Reference Base Stations **R** receive a single packet from client C to select $R_{optimal}$;

$R_{mask}(x,y) = 1\ \forall(x,y)|(||(x,y) - R_{optimal}||_2 < ||(x,y) - R||_2 \forall R \neq R_{optimal})$

$(\mathbf{X}, \mathbf{Y})_{allowed} = (\mathbf{X}, \mathbf{Y})_{allowed}. * R_{mask}$;

$P(x,y) = 0\ \forall(x,y) \in (\mathbf{X}, \mathbf{Y})_{allowed}$;

for *each* (BS^m, BS^n) *pair* **do**

  $tdoaMeasured^{m,n} = \underset{t}{argmax}\ InverseNDFT(\tilde{\mathbf{h}}, \mathbf{f})$

$tdoaActual^{m,n} = tdoaMeasured^{m,n} + tdoaRef^{m,n}_{R_{optimal}}$;

return ParticleFilter($(\mathbf{X}, \mathbf{Y})_{allowed}$, GetProbability);

`/* Run Particle filter [44] sampling allowed space using GetProbability(.) */`

Function `GetProbability`(x,y):

  for *each* (BS^m, BS^n) *pair* **do**

    if $z = |||(x,y) - BS^m||_2 - ||(x,y) - BS^n||_2 - tdoaActual^{m,n}| < mar/Penalty(x,y)$ **then**

      $P(x,y) = P(x,y) + e^{-z/2\sigma^2}$

- *Temporal Constraints:* We note that our algorithm does not implement additional temporal filters (e.g. Kalman Filtering), however these can readily be incorporated to improve accuracy (at the cost of energy).

6 DISCUSSIONS AND LIMITATIONS

We highlight a few important limitations of OwLL:

Mobility: Our system is not designed for moving objects owing to the need for clients to frequency hop. This should not pose a problem for most LoRa deployments are quasi-static (e.g. building sensor networks). We believe it may be possible to extend OwLL to aggregate measurements across frequency hops for a continuously moving object to trace out its trajectory, if the object moves relatively slowly. We leave this for future work.

Extending to 3-D: The evaluation of this paper is restricted to localization in 2-D space owing to COVID-19 related restrictions that reduced our access to multiple buildings and different floors. We had to place our base stations on temporary mobile platforms in open spaces outdoors, where social distancing was easier to ensure, both in line-of-sight and occluded settings. We expect new challenges to arise for extending it to 3-D localization which remains an exciting direction for future work.

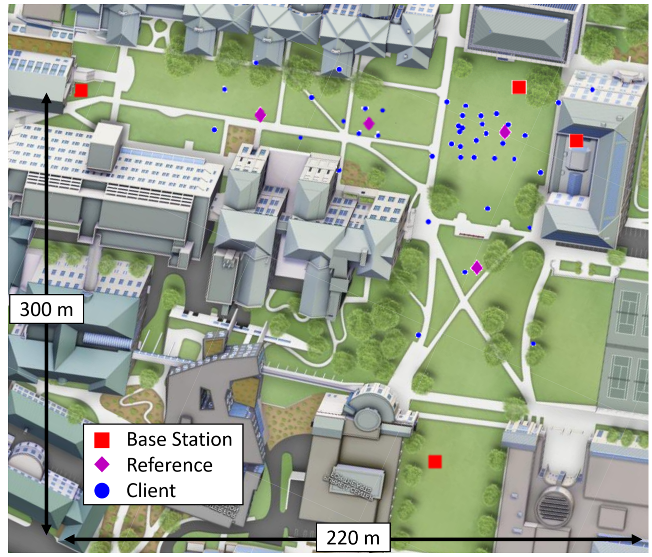


Figure 8: OwLL's campus-scale deployment.

Extreme Multipath/Occlusion: While our system does consider the impact of obstructions, including buildings, it assumes that the line-of-sight path is present (i.e. above noise) at least in a subset of frequencies and base stations, even if intermingled with other paths. However, under extreme occlusion (e.g. devices in the basement, or deep inside buildings), this assumption may not hold. One may be able to detect these extreme occlusion cases by carefully observing the phase trends across frequencies, but this has not been evaluated in this paper. However, performing localization even in these occluded scenarios continues to remain an open problem for future work.

7 IMPLEMENTATION AND EVALUATION

We implement OwLL using off-the-shelf Semtech SX1262 radios as LP-WAN clients, and NI USRP N210s as software-radio base stations. These clients enable us to hop across a wide spectrum ranging from 500 MHz to 960 MHz. We hold experimental FCC licenses to operate on the TV whitespace bands across all our hardware and explicitly avoid bands with incumbent transmitters.

An important factor to take into consideration is the choice of antenna. Most off-the-shelf antennas either have narrowband gains in signal or provide a low gain or directionality

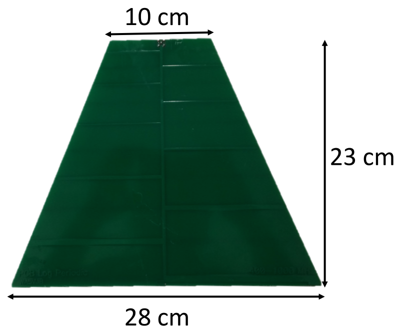


Figure 9: Log Periodic Ettus LP0410 antenna used in OwLL's evaluation.

across such a wide bandwidth. This is due to the large difference in wavelength (31cm - 75cm) across the spectrum we are aiming to hop. While there exist highly specialized whitespace antennae providing high gains across this spectrum (and there remain interesting open questions in this regard), we used multiple Ettus LP0410 log-periodic directional antennas (shown in Fig. 9) connected using SMA splitters to emulate a high gain omnidirectional whitespace antenna. With scale, omnidirectional white space antennas will become cheaper and ubiquitously deployed commercially

We receive packets from the USRP base stations using UHD/C++. Our base station software is implemented in python for live processing of packets as they are collected and a hybrid python/MATLAB script in the cloud to localize the client. Note that this live processing of packets provides upto 25× compression by only transferring the I/Q data when the packet was transmitted.

We evaluate our system on a campus-scale testbed of 66,000 sq. m. with 4 base stations and client moved across 50 locations. Our client locations face various outdoor wireless obstructions such as hills, trees, buildings, etc. Chosen client locations ensure diversity in terms of being in LOS or NLOS to every base station and deployed across a varied topography. To enable our solution, we also place the reference at 4 locations to ensure coverage across all areas.

Ground Truth: An important problem with measuring accurate localization errors in outdoor environments is measuring the locations with enough accuracy. In fact, a typical GPS receiver has a 3-5 m median variance in location tracking. Instead, we use a Nikon Forestry Pro II to accurately survey our testbed with a central reference point. We use these measured distances between points to formulate a L2-minimization to estimate the location of all clients, base stations and references. The average error in our estimates falls below 10 cm between any two locations surveyed.

Base Line: LoRa localization has been explored with various side channels using satellite images, GPS locations along with other sensors[27]. However, our system localizes off-the-shelf LoRa clients, without using other side-channel information, and thus we compare OwLL to a LoRa localization system purely based on RSSI. For every client, we calculate the distances from the client to all the base stations using relative RSSI value taken with respect to the RSSI of the signal from the reference base station. We find the location of the client using simple trilateration. This acts as baseline system to compare our results against.

8 EXPERIMENTAL RESULTS

We first evaluate the primitives that work together to enable OwLL's time difference of arrival estimates (Sec. 8.1-8.3). We then evaluate OwLL's capability to accurately estimate the time difference of arrival and location across LOS and NLOS scenarios across our campus-scale testbed. (Sec. 8.4-8.5)

8.1 Phase Stability and SNR resilience

In this experiment, we attempt to ascertain the SNR resilience and phase stability of OwLL's approach over prior work.

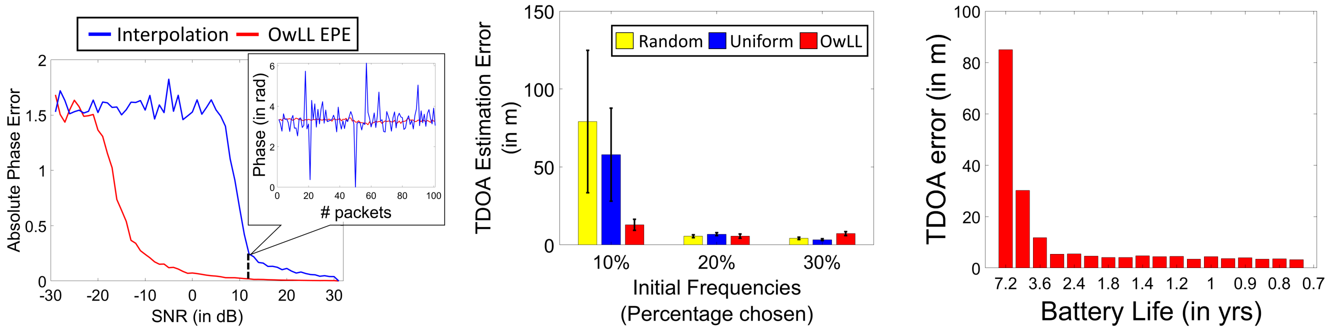
Method: We transmit 100 packets from transmitter and reference at known locations at a single frequency (915 MHz). We compute the phase after synchronization (described in Sec. 5.1) from their signals using two approaches: (i) Interpolation approach (Fig. 5 left) and (ii) OwLL Enhanced Phase Estimation (Fig. 5 center). We then find the error in measured phase from the actual phase to demonstrate SNR resilience and phase stability.

Results: Fig.10a shows that OwLL's enhanced phase estimation is $\sim$ 25 dB more resilient when compared to direct interpolation of phase across frequencies. Due to the inherent error in interpolating the phase across frequencies from a large number of points, we see that even at high SNRs, enhanced phase estimation outperforms the interpolation method. The smaller graph on the top right depicts the phase over 100 packets for both approaches.

8.2 Initial Frequency Selection

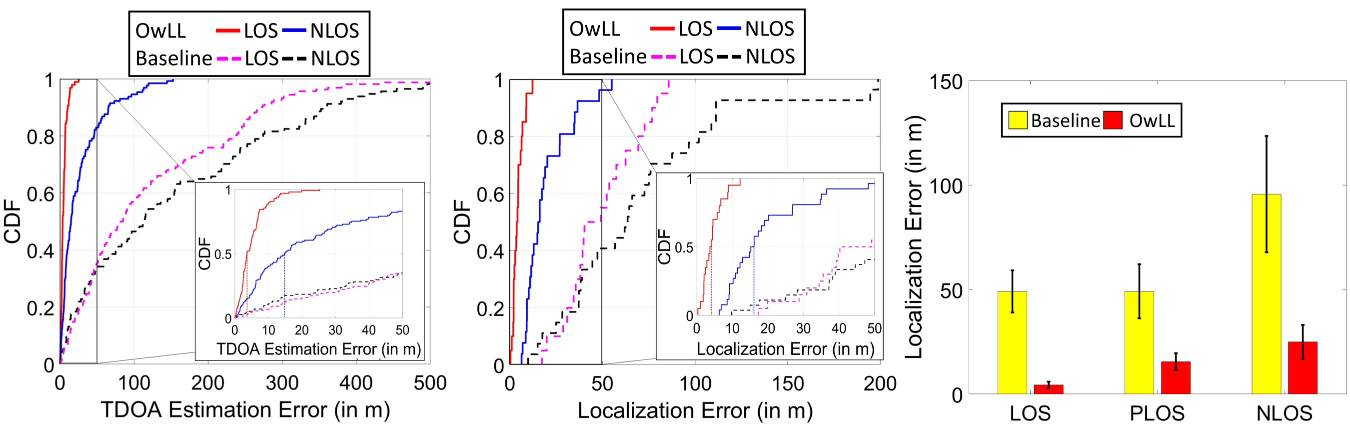
In this section, we compare OwLL's approach of identifying the initial frequencies vs. other baseline approaches.

Method: We chose 10 locations with less than median localization error to identify what choice of frequencies contain the most amount of information in identifying the accurate time-difference-of-arrival of client across base stations. As proposed in Sec. 4.3, we typically choose a few initial frequencies to create a rough estimate of TDOA which we reinforce



(a) (b) (c)

Figure 10: OwLL Design Considerations: (a) Stability and SNR resilience of OwLL Enhanced Phase Estimation (b) Impact of choice of initial frequencies; (c) Accuracy-Power Tradeoff



(a) (b) (c)

Figure 11: Localization Results: (a) CDF of time-difference-of-arrival estimation error (b) CDF of location estimation error (c) Impact of Multipath on location estimation

using Algm. 2. However, it is critical to understand that a bad choice in these initial frequencies can adversely change the TDOA as well as localization accuracy. Based on our observations in Sec. 4.3, we devise a custom initial frequency selection policy that chooses two chunks of frequencies across the first and last available band for transmission. We also choose the frequencies near the center frequencies of receivers as typical software radio filters provide the highest SNR gain at these frequencies. We compare OwLL's initial frequency selection policy with two baselines: (i) choosing the frequencies *randomly* and (ii) choosing the frequencies *uniformly*.

Results: Fig. 10b shows significant benefits of OwLL when the number of initial frequencies chosen is small. This is especially important as we would like to minimize the number of frequencies queried to save power on these power-starved LoRa clients. However, as we choose more frequencies we see that the baseline policies do approach similar errors to that of OwLL. Note, however, in really large TDOA scenarios OwLL's policy will be superior to others due to its ability to disambiguate large range of TDOA values as well as providing reasonably large bandwidth.

8.3 Accuracy-Power Tradeoff

In this result, we address the key drawback of hopping frequencies – the additional power consumption incurred.

Method: We attempt to identify how much power overhead is incurred due to hopping multiple frequencies to estimate the location of the client. We assume that the client needs to be localized twice everyday. We assume the typical battery energy of a AA battery of 2900 mAh and a typical LoRa client consumes 100 mAs of energy for every packet (based on battery life model used in [13]). This means that a typical LoRa client can send approximately 104400 packets in its whole lifetime. We then study how OwLL's localization accuracy changes as we transmit more and more packets.

Results: As shown in Fig. 10c, we see that after transmitting around 80 packets per localization query, we are able to achieve the 75th percentile localization error for the chosen points enabling us a 1.8 year battery life assuming the client is getting localized twice a day. Further to achieve the 50th percentile localization error, we need to transmit 40 more packets per localization query leading to 1 year battery life. Note that here we have assumed that the client transmits *only* to get localized. Any other transmissions by the client to communicate sensor data will reduce the battery life further. However, we would also like to note that it is quite atypical to localize a quasi-static client at the frequency we have considered and the behavior will likely be more energy-efficient in real world (~ localizing once a month) removing localization as the energy bottleneck.

8.4 Time Difference of Arrival Accuracy

In this result, we identify the estimation error in the time-difference-of-arrival across the client and reference at the two base stations.

Method: We evaluate our system across 50 locations on our campus scale testbed - 20 line-of-sight (LOS) and 30 non-line-of-sight (NLOS) points. Note that some of the NLOS points may be in line-of-sight of one or two base stations (we differentiate between the accuracy of these points from those completely non-line-of-sight in Sec. 8.5). We implement our TDOA estimation algorithm as defined in Algm. 1 using three available TV whitespace + ISM bands from 580 MHz to 920 MHz. (the available bands change over time but the edge bands were vacant across all experiments).

Results: Fig. 11a depicts a 3.6m median error for OwLL in estimating the TDOA of LOS clients across 4 base stations. This median error becomes 14.8m for the NLOS locations. We surmise that across some of these NLOS pairs, the path corresponding to the LOS TDOA is severely attenuated thus denying our algorithm the ability to extract it from the phase of the signals. Note that we still significantly outperform the baseline system of localizing using RSSI, which provides a poor median accuracy of 90m across all scenarios. Note that this is not surprising as it is quite close to observations in prior work [27].

8.5 Localization Error

Next, we evaluate the localization error using OwLL.

Method: We implement OwLL across the same 50 locations as the previous experiment and attempt to localize the clients based on their estimated TDOA distances as shown in Algm. 3.

Overall Results: Our results in Fig. 11b demonstrate a median accuracy of 3.9 m for the LOS locations. This accuracy drops to 15.7m for localizing the NLOS points. This reduction in accuracy directly corresponds to the reduced accuracy of estimating the TDOA across base stations for the client location. We note that despite the increased error, our system significantly outperforms the state-of-the-art baseline under identical settings. We also note the longer tail for NLOS settings due to certain locations that experience significant occlusions compared to others. We believe new techniques are required to inherently understand the multipath around a client and reinforce the severely attenuated LOS TDOA paths to improve the accuracy. Note that while recent work has shown promise using deep learning for addressing this problem in the WiFi context [6], the same for LoRa remains a direction for future work. The accuracy of the RSSI-based baseline is 53.6 m across all points.

Impact of Multipath: It is important to note in the above evaluation that some of the NLOS points were partially in line-of-sight (PLOS) of some of the base stations. Thus, it is critical to differentiate the accuracy in such scenarios from locations which are completely NLOS of the base stations. Fig.11c shows the localization errors across the three scenarios using OwLL. As we can clearly see the mean errors for PLOS and NLOS scenarios are 15.5 m and 24.9 m respectively. However, the median errors for both scenarios are 14.8 m and 16.1 m. This means much of the error in purely NLOS scenarios is dominated by a few locations with poor localization accuracy (highlighting the importance of both metrics). The mean localization error for LOS scenarios is 4.4 m.

9 CONCLUSION AND FUTURE WORK

This paper presents OwLL, a LoRa localization system that achieves few meters accurate localization of commodity LoRa clients. We design a distributed base station network made of individually low-cost components that together spans a wide-bandwidth and offers high aperture. We show how this network can aggregate signal measurements made across multiple different narrowband channels of a LoRa client to triangulate it at fine accuracy. We further optimize our system to demonstrate accurate localization with minimal energy overhead from the clients. We deploy OwLL on a testbed spanning 66,000 sq.m. centered in a major U.S. city and show a 9 m median error in location estimation. While the evaluation in this paper is restricted to 2-D, we believe 3-D is possible with careful placement of the base stations and leave this for future work. We further believe that handling moving objects and addressing significant attenuation of the direct path for devices deep inside buildings or underground remain important problems for future work.

ACKNOWLEDGMENTS

We would like to thank NSF (grants 1942902, 1718435, 1837607, 2030154, and 2007786), IoT@Cylab, CyLab Presidential Fellowship, the CMU Scott Institute and the Kavcic-Moura grant for their support.

REFERENCES

[1] LoRaWAN – What is it? A Technical Overview of LoRa and LoRaWAN. Tech. Rep. November, LoRa Alliance, 2015.

[2] Aernouts, M., BniLam, N., Berkvens, R., and Weyn, M. TDAoA: a Combination of TDoA and AoA Localization with LoRaWAN. *Internet of Things* (2020).

[3] Aernouts, M., et al. Sigfox and LoRaWAN datasets for fingerprint localization in large urban and rural areas. *Data* (2018).

[4] Alippi, C., and Vanini, G. A RSSI-based and calibrated centralized localization technique for Wireless Sensor Networks. In *IEEE PerCom Workshops 2006* (2006).

[5] Andrade, T., Alves, L. N., Cupido, L., and Santos, D. A reverse GPS architecture for tracking and location of small objects. In *IEEE ICL-GNSS* (2011).

[6] Ayyalasomayajula, R., et al. Deep learning based wireless localization for indoor navigation. In *ACM MobiCom* (2020).

[7] Bargh, M. S., and de Groote, R. Indoor localization based on response rate of bluetooth inquiries. In *ACM Workshop on Mobile entity localization and tracking in GPS-less environments* (2008).

[8] Capon, J. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE 57*, 8 (1969), 1408–1418.

[9] Choi, W., et al. Low-Power LoRa signal-based outdoor positioning using fingerprint Algorithm. *ISPRS International Journal of Geo-Information* (2018).

[10] Dongare, A., et al. Charm: exploiting geographical diversity through coherent combining in low-power wide-area networks. In *ACM/IEEE IPSN* (2018).

[11] Du, H., et al. A hybrid outdoor localization scheme with high-position accuracy and low-power consumption. *EURASIP Journal on Wireless Communications and Networking* (2018).

[12] FCC. Second Report and Order and Memorandum Opinion. *Tech. Rep.* (2010).

[13] Gadre, A., et al. Frequency configuration for low-power wide-area networks in a heartbeat. In *USENIX NSDI* (2020).

[14] Gadre, A., Yi, F., Rowe, A., Iannucci, B., and Kumar, S. Quick (and dirty) aggregate queries on low-power wans. In *2020 19th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)* (2020), IEEE, pp. 277–288.

[15] Gholami, M. R., et al. RSS-based sensor localization in the presence of unknown channel parameters. *IEEE Transactions on Signal Processing* (2013).

[16] GSM Association. 3GPP Low Power Wide Area Technologies (White Paper).

[17] Gu, C., Jiang, L., and Tan, R. Lora-based localization: Opportunities and challenges. In *ACM EWSN* (2019).

[18] Islam, B., Islam, M. T., Kaur, J., and Nirjon, S. Lorain: Making a case for lora in indoor localization. In *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)* (2019), IEEE, pp. 423–426.

[19] Kalliola, K., et al. System and Methods for Direction Finding using a Handheld Device, US Patent 8115680.

[20] Kempke, B., Pannuto, P., and Dutta, P. Harmonia: Wideband spreading for accurate indoor RF localization. *ACM SIGMOBILE Mobile Computing and Communications Review 18*, 3 (2015), 19–25.

[21] Kotaru, M., Joshi, K., Bharadia, D., and Katti, S. Spotfi: Decimeter level localization using wifi. In *ACM SIGCOMM* (2015).

[22] Kumar, S., et al. Accurate Indoor Localization with Zero Start-up Cost. In *ACM MobiCom* (2014).

[23] Kumar, S., et al. LTE radio analytics made easy and accessible. In *ACM SIGCOMM Computer Communication Review* (2014).

[24] Lam, K.-H., et al. New rssi-based lora localization algorithms for very noisy outdoor environment. In *IEEE COMPSAC* (2018).

[25] Lam, K.-H., et al. RSSI-Based LoRa Localization Systems for Large-Scale Indoor and Outdoor Environments. *IEEE Transactions on Vehicular Technology* (2019).

[26] Lazik, P., et al. Alps: A bluetooth and ultrasound platform for mapping and localization. In *ACM SenSys* (2015).

[27] Lin, Y., et al. SateLoc: A Virtual Fingerprinting Approach to Outdoor LoRa Localization using Satellite Images. In *ACM/IEEE IPSN* (2020).

[28] Ma, Y., et al. Minding the billions: Ultra-wideband localization for deployed rfid tags. In *ACM MobiCom* (2017).

[29] MacCurdy, R., et al. Automatic Animal Tracking Using Matched Filters and Time Difference of Arrival. *JCM* (2009), 487–495.

[30] Nandakumar, R., Iyer, V., and Gollakota, S. 3D localization for sub-centimeter sized devices. In *ACM SenSys* (2018).

[31] Niu, R., Vempaty, A., and Varshney, P. K. Received-signal-strength-based localization in wireless sensor networks. *Proceedings of the IEEE 106*, 7 (2018), 1166–1182.

[32] Plets, D., et al. Experimental performance evaluation of outdoor tdoa and rss positioning in a public lora network. In *IEEE IPIN* (2018).

[33] Podevijn, N., et al. TDoA-based outdoor positioning with tracking algorithm in a public LoRa network. *Wireless Communications and Mobile Computing* (2018).

[34] Podevijn, N., et al. Experimental tdoa localisation in real public lora networks. In *IEEE IPIN* (2019).

[35] Ratasuk, R., et al. NB-IoT system for M2M communication. In *IEEE WCNC* (2016).

[36] Sanchez, C., et al. Performance Evaluation of RSSI-based Positioning System with Low-cost LoRa Devices. In *ACM PE-WASUN* (2019).

[37] Sanchez-Iborra, R., and Cano, M.-D. State of the Art in LP-WAN Solutions for Industrial IoT Services. *Sensors* (2016).

[38] Sichitiu, M. L., and Ramadurai, V. Localization of wireless sensor networks with a mobile beacon. In *IEEE Conference on Mobile Ad-hoc and Sensor Systems* (2004).

[39] Singh, A., et al. TV white spaces exploration for cognitive radio: taxonomy and research issues. *Telecommunication Systems* (2020).

[40] Sornin, N., et al. LoRaWAN Specification.

[41] Sun, G., et al. Signal processing techniques in network-aided positioning: a survey of state-of-the-art positioning designs. *IEEE Signal Processing Magazine* (2005).

[42] Tomic, S., and Beko, M. A bisection-based approach for exact target localization in NLOS environments. *Signal Processing* (2018).

[43] Tomic, S., et al. On target localization using combined RSS and AoA measurements. *Sensors* (2018).

[44] Van Der Merwe, R., Doucet, A., De Freitas, N., and Wan, E. A. The unscented particle filter. In *Advances in neural information processing systems* (2001).

[45] Vasisht, D., et al. Decimeter-level localization with a single wifi access point. In *USENIX NSDI* (2016).

[46] Weiser, A. W., et al. Characterizing the Accuracy of a Self-synchronized reverse-GPS Wildlife Localization System. In *ACM/IEEE IPSN* (2016).

[47] Xiong, J., and Jamieson, K. Arraytrack: A fine-grained indoor location system. In *USENIX NSDI* (2013).

[48] Yang, L., et al. Tagoram: Real-time tracking of mobile rfid tags to high precision using cots devices. In *ACM MobiCom* (2014).

[49] Zhang, C., et al. Accurate UWB indoor localization system utilizing time difference of arrival approach. In *IEEE radio and wireless symposium* (2006).