Signal Transmitter Localization using Low-cost SDR receivers

Yago Lizarribar, *Master in Telematics Engineering, Universidad Carlos III de Madrid*

***Abstract*—The popularity of wireless communications is growing every year, and with it the possibility of threats in the electromagnetic spectrum increases. One of the key aspects to stop these attacks is to be able to localize where the threat is coming from. Current commercial products make use of expensive hardware and GPS-based synchronization techniques to perform geolocation. In this thesis we propose a network architecture based on low cost, GPS-free Software-Defined Radio (SDR) receivers that localizes a signal transmitter in a collaborative manner. We perform evaluations on different components of the architecture such as receiver imperfection correction, reference signals and multilateration approaches. Our results indicate that signal transmitter localization is feasible with the architecture proposed even using low-cost radio receivers.**

***Index Terms*—Outdoor Localization, TDOA, SDR, Sensor networks.**

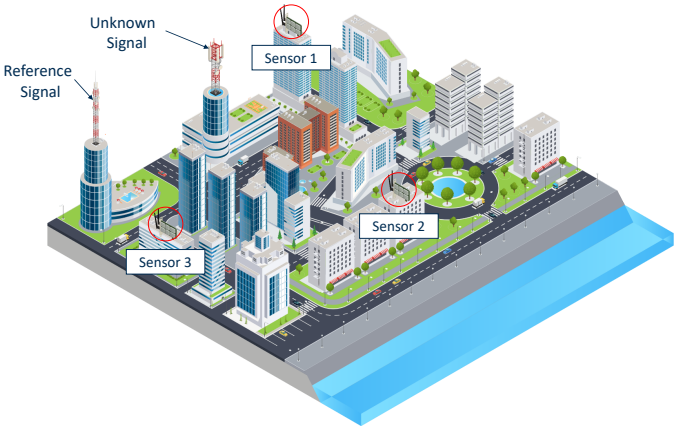


Fig. 1. Unknown transmitter can be localized collaboratively by using a network of sensors that make use of a reference signal transmitter to synchronize among them

I. Introduction

With the advent of wireless devices and technologies, wireless communication infrastructure is growing at a frantic pace [1]. Furthermore, the arrival of 5G networks will accelerate this growth. Wireless communication is possible due to a valuable and scarce resource, the Electromagnetic (EM) spectrum. Being a core pillar of today's society, from individuals to nations and companies, monitoring and detecting threats in the EM spectrum is of fundamental importance.

Attacks on the spectrum can be done with relatively affordable equipment and they can target any radio frequency, thus they can wreak havoc in highly populated areas. These attacks may come from other nations, criminal organizations or even individuals and they can take diverse forms: false telephony towers (that can be easily built with current technology) that may interfere with mobile devices, fake transmitters that emit deceptive signals, unauthorized transmissions to deactivate meteorological radars or GPS spoofing devices [2] [3].

Even though the EM spectrum is heavily regulated and the threats mentioned above have the potential to disrupt communications at a large scale, little is known yet of how spectrum is used and where are those anomalies. Current equipment used to detect these illegal transmissions is expensive, bulky and heavy and it is therefore not suitable to deal with the types of attacks above mentioned. Besides, detecting threats in a fast and accurate manner requires large financial and labor investments. In the past years, there have been attempts at monitoring the radio frequency spectrum, like Google Spectrum [4] or Microsoft Spectrum Observatory [5]. There exist companies like CRFS[1] who provide commercial spectrum monitoring solutions with custom hardware and thus higher costs.

In recent years, the cost of SDR has decreased drastically, and devices like the RTL-SDR, whose cost is around 15-20 $ have proliferated. The popularity of these low-cost receivers has led to the emergence of new spectrum monitoring solutions [6], that can outperform more advanced systems at a fraction of their cost. One of the most notable examples is Electrosense [7] [8], a crowd sourced EM spectrum monitoring initiative, that is based on low cost hardware like the aforementioned RTL-SDR and the Raspberry Pi[2] boards.

When it comes to detecting threats in the wireless channels, there are 2 main branches: Anomaly Detection and Transmitter Geolocation. The former is concerned with tagging anomalous or suspicious transmissions in the spectrum, whereas the latter looks at positioning the source of that anomalous signal. In this thesis, we will focus on the Transmitter Geolocation problem.

Literature regarding indoor localization is extense, and an ample variety of methods are employed such as Received Signal Strength (RSS), Angle of Arrival (AOA) or Time of Arrival (TOA). Transmitter localization on the other hand, has not been as widely studied, and implementations are often in controlled scenarios, that fail when deployed in more realistic environments. Previous research has focused on power based methods, which have lower costs and are easier to implement at the expense of lower accuracy. Methods based on Time Difference of Arrival (TDOA) can, in theory, reach higher accuracy but very few studies have attempted to deploy

[1] https://www.crfs.com/

[2] https://www.raspberrypi.org/

systems with these characteristics using low cost spectrum sensors.

One of the reasons is that for a TDOA based system to work, accurate synchronization among sensors is needed. This is achieved by the use of GPS receivers which increases the overall cost of the platform. An example of such case is KiwiSDR[3] an amateur radio initiative using custom SDR boards with GPS receivers and versatile web interface, which allows users to perform TDOA localization with different sensors.

However, low cost sensors like the RTL-SDR cannot take advantage of GPS synchronization. To counter attack this limitation we propose to use a known Reference Signal (RS) to synchronize the network of sensors and then make feasible the localization of the Unknown Signal (US), as Figure 1 shows.

The purpose for this thesis is to explore TDOA based methods for transmitter localization using the low cost RTL-SDR receiver and propose an architecture that can be implemented in a crowd sourced system like [8]. The main contributions are:

1) An architecture for transmitter localization based on TDOA, that is GPS free, signal agnostic and embeddable in a network of low-cost devices like the combination of Raspberry Pi and RTL-SDR. This architecture takes into account the SDR imperfections and corrects them. To solve the problem of time synchronization we will use a reference signal.
2) An evaluation on different common signals, such as Long-Term Evolution (LTE), Digital Video Broadcasting-Terrestrial (DVB-T) or Global System for Mobile Communications (GSM), and their potential to be used as reference for the localization system.
3) An evaluation on different multilateration algorithms and their performance using a well known platform.

The remainder of this thesis is organized as follows: Section II dives into the State of the Art of localization methods and Section III shows the most notable outdoor localization approaches that can be found in the literature. Section IV describes the main challenges of localization systems based on low cost hardware and Section V proposes an architecture for such systems. In Section VI we provide evaluation on the different stages of our architecture and last Section VII will show some conclusions and future lines of work.

II. Localization Methods

When tackling the problem of localizing unknown transmitters we can distinguish 2 big categories: range-based and range-free methods. Range is generally tackled from 4 different perspectives as it was summarized in [9]:

- **RSS**: The core of this method is to measure the power received at different sensors, and by means of a path loss model, estimating the distance to the transmitter. Sensors need to be calibrated and a Path-Loss model is needed to estimate the decay in power.
- **AOA**: Based on the use of directionally sensitive antennas, it requires specialized hardware.
- **TOA**: In this technique, the travelling time from the emitter to the different receivers is measured, and then position is obtained by performing a multilateration. For this approach to work, an accurate time synchronization among sensors and transmitters is needed.
- **TDOA**: It is similar to the TOA technique, but measuring the differences in arrival times of the signal to all the sensors. Since only sensors need to be synchronized in this scenario, it fits the case of unknown transmitter localization.

Table I provides a summary of the features of each of the methods, based on the analysis done in [10]. In the following sections, we will describe each of these methods in more detail.

1) RSS based methods: In the past years, a great deal of effort has been focused on RSS based systems, as it is stated in [11] and [12]. The situations in which these methods have been applied range from static transmitter localization with static sensors [13] [14] or mobile robots [15], distance sensing for collision avoidance between robots [16] or even wildlife tracking [17].

The main appeal of RSS based methods is the low hardware requirements [10], which makes an attractive case for deployments on large sensor networks. However, these techniques have their drawbacks such as: (i) Sensors need to be calibrated. (ii) A Path-Loss model is needed to convert from power measurements to distances. (iii) If there are multiple transmitters, an algorithm is needed to discriminate the contributions of each one, and they often have high time complexities [13]. (iv) It requires either knowing the transmitter power or a dense sensor network.

2) AOA based methods: AOA is an interesting approach since it only requires a pair of sensors and the angle of incidence is computed locally, so the amount of information sent to the backend is less than with other methods [18]. It also has a notable accuracy both in theory [19] and in practice [20].

One of the main issues of AOA based methods is that in order to compute the angle of incidence, an special array of

[3]http://www.kiwisdr.com/

Name	Principle of Operation	Special Hardware	Attenuation	Cost
RSS	Signal strength measurement	Not necessary	High	Low
AOA	Angle of signal arrival	Required	Medium	High
TOA	Time of signal arrival	Required	Low	Medium
TDOA	Time difference in signal propagation	Required	Low	Medium

TABLE I
RANGE BASED LOCALIZATION METHODS AND THEIR MOST PROMINENT FEATURES

antennas is needed, and that poses a notable obstacle when it comes to deploying a dense AOA system.

3) TOA based methods: TOA has been widely employed as a localization method in indoor Wireless Sensor Networks, since it provides accurate results and does not need to make assumptions on the signal [21]. The general idea is to calculate the time at which the signal arrives at each sensor (hence the name TOA), compute the circumferences and calculate the intersection by means of trilateration.

However, for the algorithm to work, it is necessary to know the signal's transmission time, thus making it unsuitable for the case of identifying and localizing rogue/unknown transmitters. Furthermore, transmitter and receivers need to be accurately synchronized, thus GPS based systems are often used.

4) TDOA based methods: TDOA algorithms have been employed for a long time for various types of localization tasks. In theory, it has slightly less accuracy to TOA based localization approaches [21] but with the advantage of only needing synchronization between receivers, and it is therefore an interesting approach for the task of localizing unknown transmitters. After TDOA values are computed for each pair of receivers, the position of the transmitter is estimated by means of multilaterating, in a similar manner as in TOA.

In recent years, there has been a part of research shifting towards hybrid solutions, as it is mentioned in [22] [18] and [19], where combining TDOA with other methods, authors achieved higher accuracies than using those methods isolated.

III. Related Work

In recent years, SDR receivers have sprouted, going from the low-cost RTL-SDR to more high-end ones like the Universal Software Radio Peripheral (USRP), and many others in between.

This increase in availability has led to a rise in the amount of research in the localization field using these devices. A great deal of articles related with localization are based on theoretical conjectures and simulations, but rarely are these assumptions tested on scenarios other than toy ones and areas greater than 100×100 m^2.

Being RSS based research the most popular in the past years, it comes as no surprise that SDR receivers are widely employed in these papers. Such examples appear in [12], where they make use of USRP, Pluto and other high-end receivers. The aforementioned receivers offer a great deal of features like GPS disciplined oscillators and higher bandwidths, although their cost is higher. In [17] we see an example of RSS based localization using the RTL-SDR although the covered distance is no more than 100 m, thus it cannot be considered a large scale deployment.

Another example of RSS based system is the one proposed in [23]. However, this system again was validated in simulations and on a relatively small area, and no realistic environment was considered in that work.

With regards to TDOA localization using SDR receivers, we have several examples in [24], [25] and [26]. However, as it happens with RSS based methods, the proposed systems are either focused on one specific technology, use high end hardware or remain conceptual and do not pursue any larger scale implementation.

Recalling KiwiSDR, the sensors employed in this project are custom boards that are GPS synchronized but their cost is relatively higher (around 200$) than the Raspberry Pi and RTL-SDR combination ($<$100$). Another drawback is that they can only measure frequencies from 10 kHz to 30 MHz, relatively "far" from technologies like LTE, DVB-T or GSM.

Another interesting case is the example of Panoradio[4], an attempt to use the low cost RTL-SDR devices to perform geolocation with notable results. To overcome the need of accurate time synchronization, they made use of Reference Transmitters with known locations. In their case, they deployed 3 sensors equipped with RTL-SDR in the city of Kaiserslautern. They were able to localize a number of transmitters at different frequencies and achieved notable results. One of the main drawbacks is that experiments were only done with a single set of sensors, and they do not perform any signal processing, thus they could achieve a precision of around 150 m.

IV. Challenges & Assumptions

The architecture we propose in this thesis must overcome the weak points we have observed in the related work section. The main requirements for our system are: (i) The system needs to be as low cost as possible, thus we will make use of the RTL-SDR. (ii) It must make use of commodity antennas and hardware, such as those sensors belonging to the Electrosense network. (iii) We want the system to be GPS-free, to reduce cost and because sensors may be placed in locations without GPS fix, so positioning and synchronization must be obtained with other methods or provided by other means.

In the following lines, we will describe what are the main challenges to achieve such a system and what assumptions we will make in this thesis.

A. SDR Imperfections

SDR receivers are susceptible to manufacturing errors, operating temperature and local oscillator variations. These imperfections may lead to small errors when an SDR is commanded to obtain data at the desired center and sampling frequencies. This error is called *Local Oscillator Offset* and it is usually measured in Parts Per Million (PPM). In this thesis we will assume that both errors in center and sampling frequencies are similar:

$$\frac{\Delta f_s}{f_s} \approx \frac{\Delta f_c}{f_c} \approx \varphi \tag{1}$$

where f_s, f_c and φ are the sampling frequency, center frequency and PPM respectively.

In Table II we show 4 popular SDR receivers and their reported PPM limits as per the information shown in their datasheets. Most high end receivers and the RTL-SDR incorporate devices such as Temperature Compensated Crystal Oscillator (TCXO) or GPS Disciplined Oscillator (GPSDO), that

[4]http://www.panoradio-sdr.de/

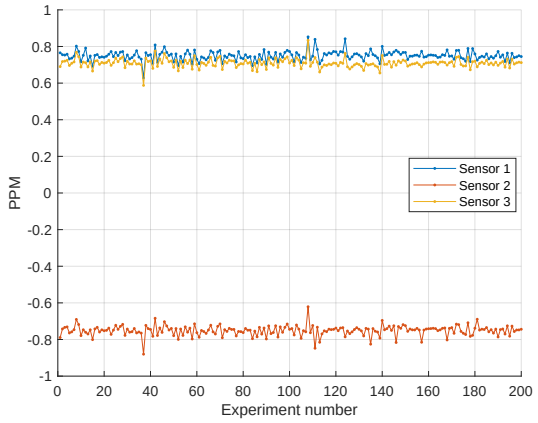


Fig. 2. PPM values over time

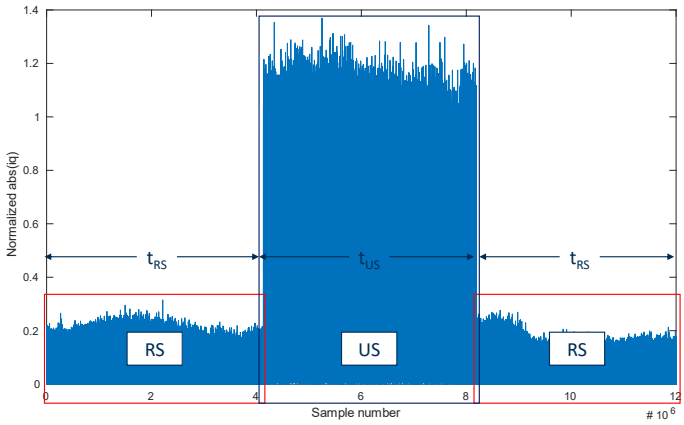


Fig. 3. Data obtained with the asynchronous library of RTL-SDR

areallow the SDR frontends to capture data at the requested parameters more precisely.

There are several tools to estimate the Local Oscillator offset. Two of those tools are *kalibrate-rtl*[5] and *LTESS-Track* [27]. Both approaches exploit the features of specific technologies to estimate the error in the oscillator (GSM for the former and LTE for the latter). In this work, we utilized *LTESS-Track* as our offset estimator, due to the accuracy and the low amount of time it requires to estimate the oscillator errors of our receivers.

Fortunately, for the RTL-SDR v3, which are the receivers we have analyzed in this thesis, the PPM values remain relatively stable. To show this, we took 200 LTE traces with 3 different sensors and we plot the PPM values over time. Figure 2 shows that throughout the experiments, PPM values for all 3 sensors remain constant. Given this fact, we could conclude that we do not need to continuously estimate the oscillator offset but we can use the same values for longer periods of time.

B. Sensor positioning

Another important aspect for a localization system is that sensors need to be accurately geolocated. Many commodity hardware devices we use in daily life, such as smartphones, already come with GPS receivers that can provide positioning accuracies of a couple meters.

However, for our system, we will not be able to rely on these GPS receivers to obtain the position of the sensors. In a crowdsourced initiative like Electrosense, users are in charge of the location of the sensor and we will rely on that. However, users may decide to obfuscate the actual location which may have an impact on the overall localization accuracy.

There has been work done however, to opportunistically exploit time information in aerial signals as it can be seen in [28]. By intelligently exploiting these signals, the authors were able to localize sensors in an indoor environment. This work is out of the scope of this thesis, but it would be an interesting approach to pursue.

[5]https://github.com/steve-m/kalibrate-rtl

SDR	Cost	Max. Bandwidth	PPM	Special HW
RTL-SDR v2	15-20$	2.4 MHz	±80	-
RTL-SDR v3	15-20$	2.4 MHz	±1	TCXO
USRP B210	1300$	60 MHz	±0.075	GPSDO
HackRF One	300$	20 MHz	±0.5	TCXO

TABLE II
POPULAR SDR RECEIVERS AND THEIR MAXIMUM PPM ERRORS

C. Signal Synchronization

We have already mentioned that in either TOA or TDOA based systems, sensors need to be synchronized. One of the most standardized manners to synchronize the clocks is to use the Network Time Protocol (NTP) which can achieve accuracies of a few milliseconds over Internet. However, for a localization system this is not enough. Given that the speed of light is $3 \cdot 10^8$ m/s, a synchronization error of 1 ms would yield:

$$10^{-3} \times 3 \cdot 10^8 = 300\ km \tag{2}$$

which is an error unacceptable for any localization system.

Some sensor networks utilize GPS as their source of synchronization, which can provide accuracies at the nanosecond scale. However, as we mentioned earlier, our system does not have any GPS source. To solve this we propose an approach that utilizes NTP to coordinate the sensing processes of different sensors, and the captures data from a Reference Transmitter (RS) and the Unknown Transmitter (US) we want to locate. By knowing the location of the RS, we can align the data incoming from 2 different sensors and obtain the real difference of arrival of the unknown signal.

To achieve this, we make use of *librtlsdr-2freq*[6], a reimplementation of the *rtl-sdr* library that allows for a center frequency change while measuring. The reasons to use this library is that by continously sampling data, we avoid losing samples in between sensing processes and thus we can obtain the actual synchronization delay between a pair of sensors.

In Figure 3 we show a trace of data obtained with the mentioned library. In this example we collected Digital Audio

[6]https://github.com/DC9ST/librtlsdr-2freq

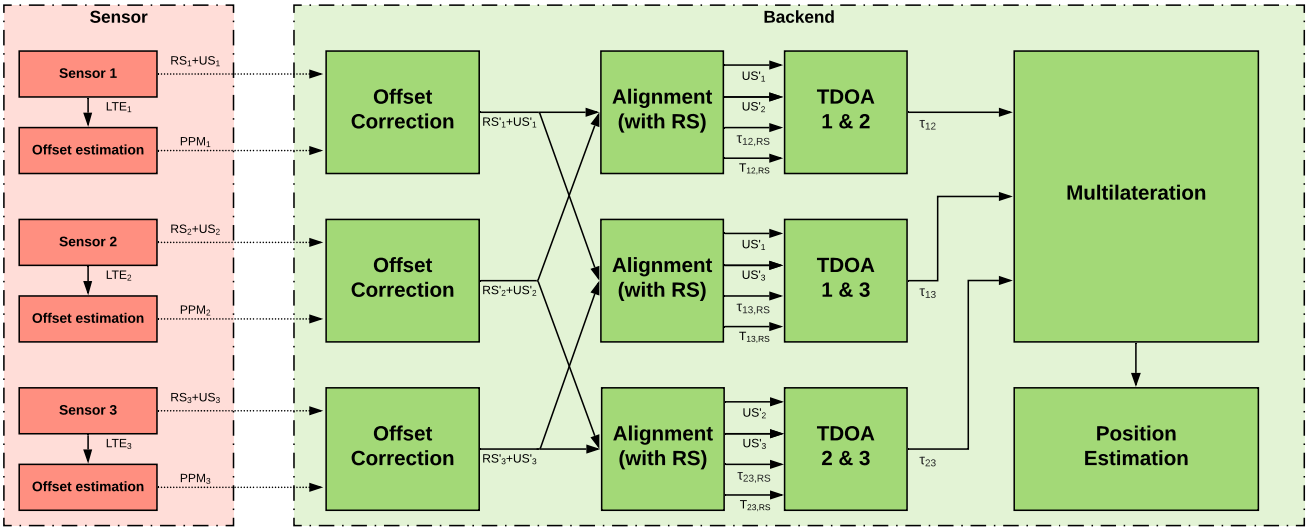


Fig. 4. General architecture of the TDOA localization module

Broadcasting (DAB) as reference (at 195 MHz) and tuned to LTE as our unknown signal. We collected a total of 12 million samples, 4 million per chunk. It can be seen that no samples are lost although we cannot assure that the converter switched frequencies exactly at 4 million samples. To avoid errors, we can set a guard interval and only use the "cropped" signals to calculate the difference of arrivals.

D. Other challenges

The challenges we mentioned above are by no means the only ones that appear in a system like this. Another issue that arises in a system like the one we are proposing in this thesis is the network usage. Sensors send In-Phase and Quadrature (IQ) samples with the RS and US at rates of around 2 Msps. At those rates, considering each sample coming from an RTL-SDR device, for example, contains 8 bits for each of the complex values, one second of data can account for:

$$8 \times 2 \times 2 \cdot 10^6 = 4MB/s \tag{3}$$

which is a considerable rate for most home networks.

One last issue concerns the bandwidth of the RTL-SDR. The maximum achievable bandwidth is of 2.4 MHz, but many of today's technologies have larger bandwidths, such as LTE or DVB-T to name a few. In [29] they managed to employ a number of RTL-SDR and reconstruct signals with larger bandwidth but this would require more advanced synchronization mechanisms between the sensors.

We will not analyze these issues since they would be out of the scope of this thesis but these are interesting challenges that will be explored in future works.

V. Architecture

Before describing the overall architecture of the system as shown in Figure 4, we will use these lines to provide a brief introduction to the Electrosense infrastructure [7] [8]. In essence, the sensors consist of a standard dipole antenna, an RTL-SDR and a Raspberry Pi (Figure 5), which the user connects at their home or office via ethernet or wifi.

Sensors use an open source tool, *es_sensor*[7] to scan the spectrum and can output values in 2 forms: Power Spectral Density (PSD) or IQ. Sensors then send this scans to the backend, that processes it and stores it in a database so that users can see the spectrum in different parts of the world, or decode different signals.

To coordinate the measurement campaigns, sensors communicate with the infrastructure via the Message Queuing Telemetry Transport (MQTT) protocol. In these campaigns, the type of measurement, the frequency ranges, duration and several other parameters are specified and sent to the corresponding sensors. Sensors are synchronized via NTP.

The localization architecture proposed in this thesis takes into account the design of Electrosense, and divides into 2 parts, **Sensor** and **Backend**. The process starts when the user wants to localize a transmitter in a given area, and selects a number of sensors (for an accurate location, at least 3 sensors are required). In the sensor side, the selected sensors would receive the measurement campaign with the start time and the

[7]https://github.com/electrosense/es-sensor

Fig. 5. Electrosense sensor kit. Source: Jetvision

two frequencies required (one as a reference and the other the target signal's). At the same time, sensors can use LTE signals and the *LTESS-Track* tool to compute their Local Oscillator Offset. After the given measurement time, these sensors send a trace of IQ samples as the one seen in Figure 3 and their corresponding Local Oscillator Offsets, measured in PPM.

For the backend, it is subdivided into several steps: *Frequency Offset Correction*, *Synchronization and TDOA Estimation* and *Multilateration*. We will use the following lines to explain them in more detail.

A. Frequency Offset Correction

Even though the TCXO of the latest versions of the RTL-SDR is guaranteed to be ± 1 PPM, it can still have a great influence in the measurement results as we will see in later sections. Using the *LTESS-Track* tool we can estimate the frequency offset, φ and correct both the center and the sampling frequency offsets.

To correct the center frequency, f_c, we must note that the actual measured center frequency, f_c', is:

$$f_c' = (1 + \varphi) \cdot f_c \tag{4}$$

therefore we must perform a frequency shift of the samples of the original signal denoted as $s'[nT_s']$. Each sample is received at time $n \cdot T_s'$, where T_s' is the observed sampling period, which can be related to the desired sampling period, T_s, as $T_s' = T_s/(1 + \varphi)$.

With all these we can generate the new signal, shifted to the desired center frequency:

$$s''[nT_s'] = s'[nT_s'] \cdot e^{-j2\pi \cdot nT_s' \cdot \varphi f_c} \tag{5}$$

Now that the center frequency is corrected, we can proceed to correct the sampling frequency. For that we can resample the data at the correct rate by performing an interpolation between samples. Thus the obtention of the corrected signal, $s[nT_s]$ can be regarded as changing from sampling rate $nT_s/(1 + \varphi)$ to nT_s:

$$s[nT_s] = I(s''[nT_s'], \frac{nT_s}{1 + \varphi}, nT_s) \tag{6}$$

B. Synchronization and TDOA Estimation

After the signals have been corrected, the next step is to estimate the TDOA values per each pair of sensors. To obtain the delay between 2 signals, we generally use first the cross-correlation function, which given 2 signals s_i and s_j is defined as follows:

$$(s_i \star s_j)[\tau] = \sum_{n=0}^{N-1} \overline{s_i(nT_s)} s_j((n + \tau)T_s) \tag{7}$$

After performing the cross-correlation, to obtain the delay τ_{ij} between both signals we take the maximum value of the resulting vector:

$$\tau_{ij} = \arg\max_{\tau} \left[(s_i \star s_j)[\tau]\right] \tag{8}$$

In this thesis, we have evaluated 3 different possibilities to perform the cross-correlation: using the raw IQ samples (`iq`), using the absolute value (`abs`) or the phase difference (`dphase`). If each IQ sample is represented by α and β respectively, the representation of each of the methods is the following:

$$s_{i,iq}[nT_s] = \alpha[nT_s] + j\beta[nT_s]$$

$$s_{i,abs}[nT_s] = \sqrt{\alpha[nT_s]^2 + \beta[nT_s]^2} \tag{9}$$

$$s_{i,dphase}[nT_s] = \angle s_{i,iq}[nT_s] - \angle s_{i,iq}[(n-1)T_s]$$

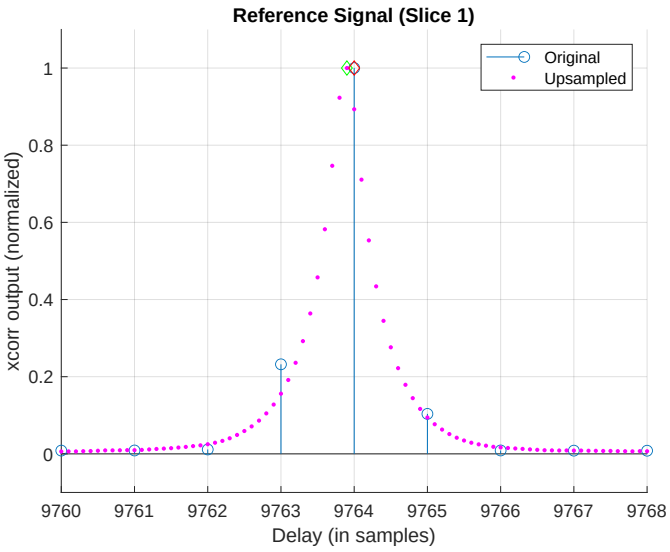


Fig. 6. Cross-correlation done on regular vs upsampled signals (factor 10)

To obtain the TDOA with the target transmitter, we first cross-correlate the first chunk of the signal, where we have RS (Figure 4) and obtain the delay $\tau_{ij,RS}$. By knowing the position of the reference transmitter, $\mathbf{p_{RS}}$ and the positions of the sensors $\mathbf{p_i}$ and $\mathbf{p_j}$ we can compute the expected difference of arrival, $T_{ij,RS}$:

$$T_{ij,RS} = ||\mathbf{p_{RS}} - \mathbf{p_i}|| - ||\mathbf{p_{RS}} - \mathbf{p_j}|| \tag{10}$$

With these values and the output from the cross-correlation of the chunks containing US, we can obtain the TDOA value for the unknown transmitter, τ_{ij} between sensors i and j:

$$\tau_{ij} = \tau_{ij,US} - (\tau_{ij,RS} - T_{ij,RS}) \tag{11}$$

One issue to consider is the fact of how much can affect an error in 1 sample when computing the TDOA. For the RTL-SDR, given it can sample at around 2 Msps, that would mean that a difference of 1 sample can result in an error of:

$$\frac{3 \cdot 10^8}{2 \cdot 10^6} = 150\ m \tag{12}$$

which can be interpreted as the approach having an uncertainty of 150 m. One way of mitigating the uncertainty region, is to upsample the signals by a reasonable factor and compute the cross-correlation on the upsampled versions. Figure 6 illustrates how by upsampling the signals by a factor of 10 in this case, we can estimate the position of the peak at a sub-sample level, and thus have more precise TDOA estimation.

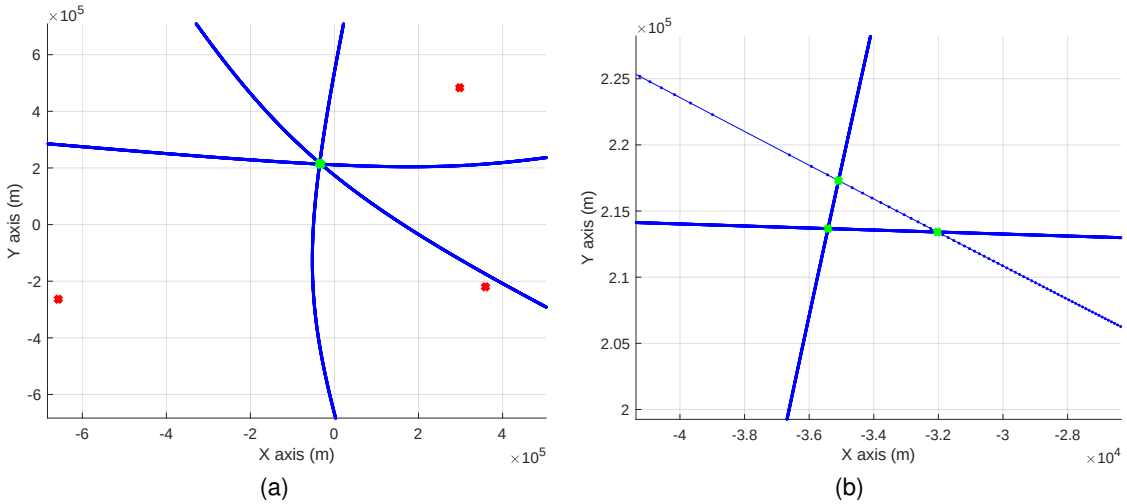


Fig. 7. How Fang's Algorithm works with real data from KiwiSDR. On the left the 3 sensors (red) and the associated hyperbolas. On the right the region defined by the intersections (green) after zooming

C. Multilateration

In the last component of the architecture the actual transmitter positioning is performed. For this thesis we have considered 3 approaches: **Linear Search**, **Fang's Algorithm** and **Least Squares**.

The Linear Search approach consists of discretizing the space into a grid of size M×N and computing the expected TDOA value per each of the squares in the grid and per pair of sensors ($\tau_{mn,ij}$). Then, the Mean Squared Error (MSE) is computed compared to the measured TDOA values:

$$e_{mn} = \sum_{i=0}^{N-2} \sum_{j>i}^{N-1} (\tau_{mn,ij} - \tau_{ij})^2 \tag{13}$$

then, the square with the lowest e_{mn} is assumed to be the transmitter location. The advantage of this method is that it is simple to implement and can provide good results. However, to achieve better accuracy finer discretization is needed and it is not difficult to see that if $M \sim N$, then the time complexity is of $\mathcal{O}(N^2)$, or in other words, grows cuadratically with the number of elements the region is divided into.

Fang's Algorithm is a method to obtain the intersection of 2 hyperbolas [30], for cases of 3 sensors, having 1 as a reference. There exist several variants depending on the geometry (e.g. planar, spherical) but for this thesis we have only implemented the planar case. Having 3 sensors, placed in $\mathbf{x_1} = [0,0]^T$, $\mathbf{x_2} = [b,0]^T$ and $\mathbf{x_3} = [x_3, y_3]^T$, and an unknown transmitter $\mathbf{x} = [x,y]^T$, the equations that result are:

$$\tau_{12} = \sqrt{x^2+y^2} - \sqrt{(x-b)^2+y^2} \tag{14}$$

$$\tau_{13} = \sqrt{x^2+y^2} - \sqrt{(x-x_3)^2+(y-y_3)^2} \tag{15}$$

The advantage of Fang's algorithm is that its computation time is lower than with the Linear Search method. The main issue with this algorithm is selecting the reference sensor, since with each sensor, due to the fact that TDOA measurements are not perfect, we will obtain different intersection points. To illustrate this, we took real data from 3 sensors from KiwiSDR and computed the intersection points using all 3 sensors as reference. As it can be seen in Figure 7, those intersections rarely land on the same point and thus further techniques need to be applied to determine the correct transmitter location.

The last of the methods implemented was the Least Squares approach. In this case we need 4 or more sensors, having one of them acting as a reference again. The TDOA between the reference sensor 1 and a sensor j:

$$\tau_{1j} = d_1 - d_j = d_1 - \sqrt{(x_j - x_t)^2 + (y_j - y_t)^2} \tag{16}$$

Operating on this equation we can end up with:

$$\begin{aligned} 2 \cdot (\tau_{1j} \cdot d_1 + (x_1 - x_j) \cdot x_t + (y_1 - y_j) \cdot y_t) \\ = \tau_{1j}^2 + x_1^2 + y_1^2 - x_j^2 - y_j^2 \end{aligned} \tag{17}$$

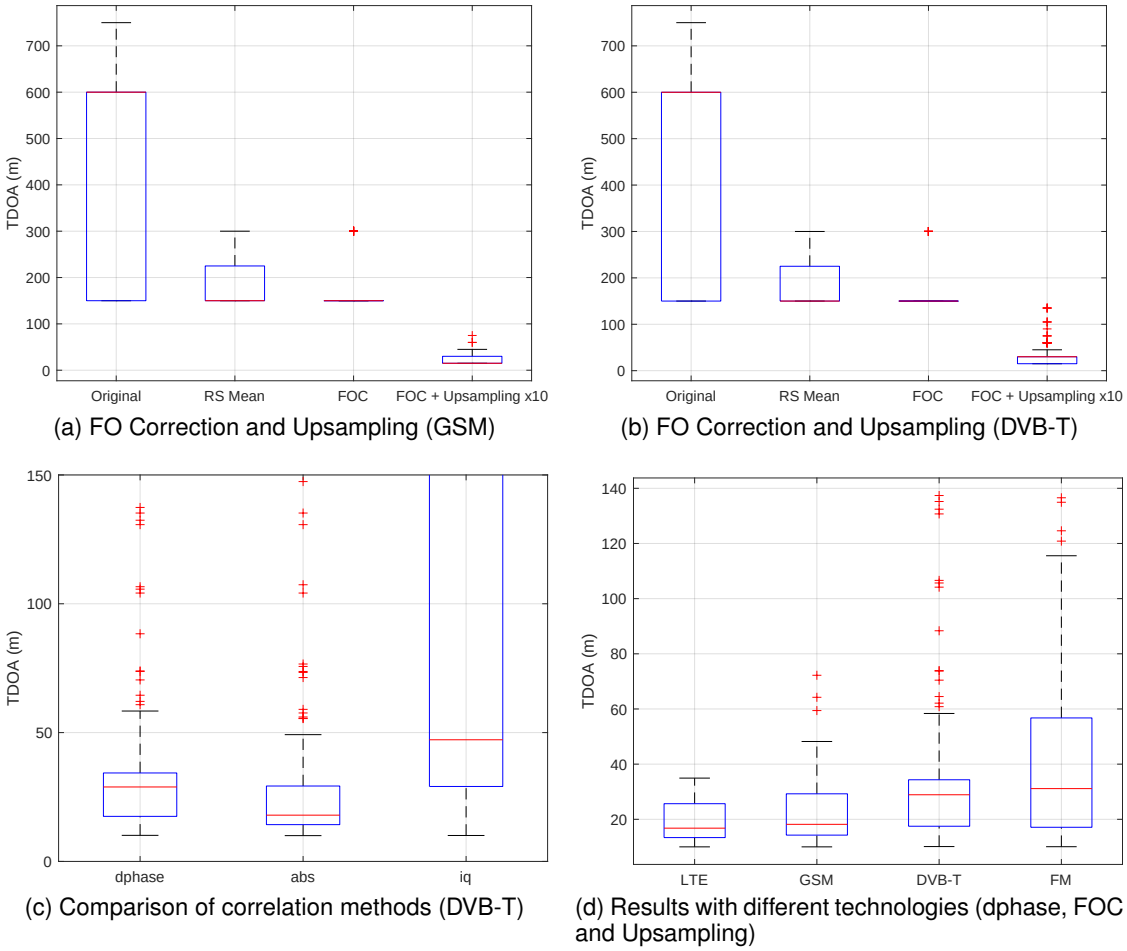
we can treat d_1, the distance from the reference sensor to the transmitter, as another variable and then we will have defined an over determined system with the form $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$. We can then solve it with the Least Squares approach:

$$\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \tag{18}$$

As with Fang's Algorithm, this method is more computationally efficient than Linear Search, but a reference sensor needs to be decided. Another issue with this approach is that linearizing d_1 can lead to erroneous results, thus further checks are required.

VI. Evaluation

To perform the evaluation of the system, we did 3 sets of experiments. The first were the feasibility studies, and we looked at how frequency offset correction, upsampling and different signals could affect the performance of the system. In the second set, we evaluated the real life applicability of different signals in our localization architecture. Last, we observed the performance of the different multilateration algorithms using real data from KiwiSDR.



(a) FO Correction and Upsampling (GSM)

(b) FO Correction and Upsampling (DVB-T)

(c) Comparison of correlation methods (DVB-T)

(d) Results with different technologies (dphase, FOC and Upsampling)

Fig. 8. Feasibility study results

A. *Feasibility Studies*

For these part of the evaluation, we used 3 RTL-SDR and collected samples in 5 different technologies: LTE, GSM, DVB-T, and FM radio. For each, receiver we collected 200 traces (as the ones in Figure 3). For each of the traces, we also collected 2 seconds of LTE traces to calculate the oscillator offsets. The experiment setup is summarized in Table III.

To collect data, we connected 3 RTL-SDR devices to an active splitter, and this to an antenna and an external power supply (Figure 9). The 3 SDR devices when then connected to the same master computer and the data collection processes where synchronized with `tmux`[8].

Measuring with this setup we can guarantee that the signal received is the same for all 3 sensors and the differences are due to their internal clock offsets. After data was collected, we analyzed all the traces with MATLAB and show the different results in Figure 8. Since the antenna is the same for all sensors, the observed differences should be 0. Any value different than that might be caused due to the offset between sensors not being properly corrected or that the signal is too noisy to be suitable to be employed as a reference signal.

1) Effects of Offset correction and Upsampling: Figures 8a and 8b show the results of the feasibility studies using GSM and DVB-T respectively in four scenarios: when no offset correction is made (Original), when we use the offset in both chunks that contain RS and average the delay (RS Mean), when we correct the offset with *LTESS-Track* (FOC) and when we correct offset and upsamle the signals (FOC + Upsampling).

In this plots is already visible that even though the Local Oscillator errors are relatively small (in our experiments we never saw values above 1 PPM, which matches the specification for the RTL-SDR), they can have a great influence in the TDOA values, and thus these errors must be corrected.

One approach to correct these errors is to assume the offset between 2 sensors is linear, which is a reasonable assumption since their oscillator errors are stable as we saw in Figure 2. By averaging the offset between both RS chunks and substracting it from the delay obtained with the US chunk, we can improve

[8]https://github.com/tmux/tmux/wiki

RS	f_{RS} (MHz)	US	f_{US} (MHz)	Bandwidth (MHz)
LTE	806	LTE	806	2
GSM	938.8	LTE	806	0.2
DVB-T	562	LTE	806	2
FM	98.8	LTE	806	0.2

TABLE III
EXPERIMENT SETUP

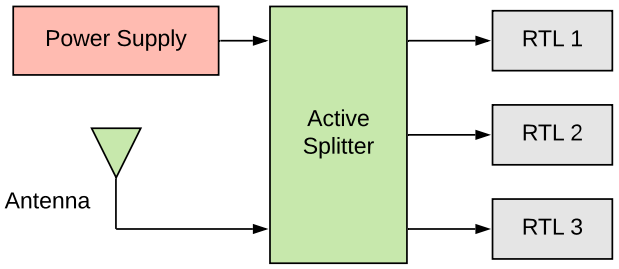
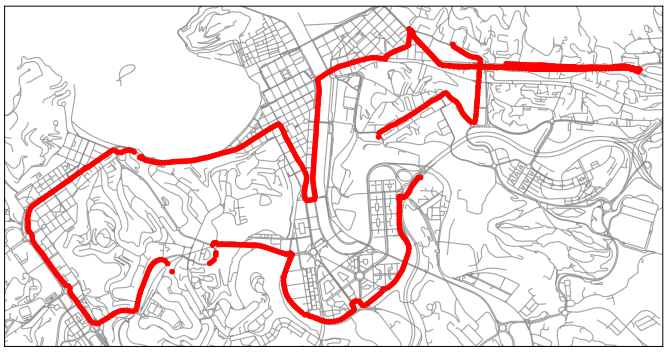
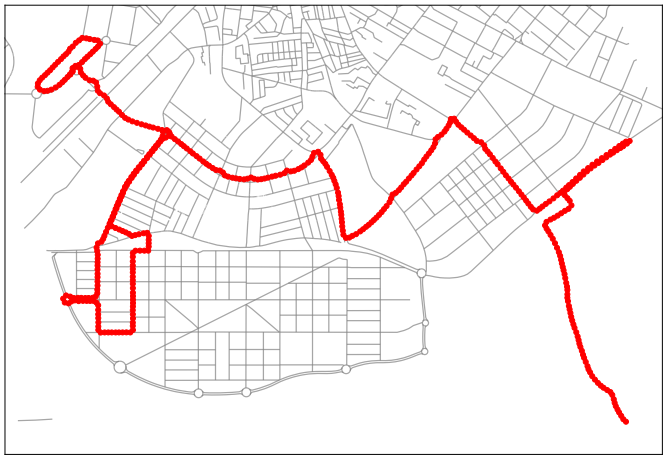


Fig. 9. Measurement setup for feasibility studies



(a) San Sebastian (Basque Country, Spain)



(b) Alcorcon (Community of Madrid, Spain)

Fig. 10. Trips in the selected cities to collect data on LTE cells

the results notably in all cases.

If instead we perform the offset correction, we can reduce the delays obtained in almost all experiments to 0. However, at a sampling rate of 2 Msps we would still have an uncertainty region, that for this experiments we considered it to be of 150 m. To observe were the actual peaks of the cross-correlation were happening, we upsampled the signals with different factors (in these figures we used a factor of 10). When upsampling, we observe that more than half of the results still yield a 0 delay in samples (which can be translated into an uncertainty region of 15 m).

2) Different Correlation methods: Figure 8c compares the 3 different correlation methods explained in Section V: `dphase`, `abs` and `iq`. In this plot we are showing the results using a DAB signal, but results were very similar with other technologies. From this figure we can clearly see that `abs` and `dphase` have similar performances (although the former slightly better), whereas `iq` method performed poorly on all the analyzed technologies.

3) Comparison of technologies: In the last plot, Figure 8d, we compare the performance of all the technologies analyzed. From this results we can see that LTE offers the best performance and FM signals the worst. One of the reasons for this difference in performance is that the former is a digital signal where as FM radio signals are analog. The fact of having discrete energy levels might probably help when correlating the signals.

Another aspect to notice is that the lower the frequency, the worse the results get (comparing LTE and DVB-T). This might be due the fact that DVB-T signals usually have greater coverage than LTE and thus the likelihood of multipath or fading happening is greater with DVB-T signals.

Last, one of the reasons that could explain why GSM performs slightly worse is due to its lower bandwidth. GSM signals have a bandwidth of 200 kHz, which is 10 times less information than what we can obtain with the RTL-SDR, thus there are less features on the signal for the cross-correlation to rely on, and more errors can occur.

B. Signal Evaluation

For a localization system to work, it is necessary that all the sensors involved in the task see the same signal. Lower frequency signals like DVB-T or DAB have greater coverage, thus it is more likely that sensors in the same area see the same signal. And as it can be seen from the previous studies, they offer similar performance to LTE since they are both digital as well.

One of the drawbacks of DAB, at least in Spain, is that currently there are only 2 cities with DAB transmitters: Madrid and Barcelona. Therefore its usage would be somewhat limited. Other countries, on the contrary, have denser deployments and utilizing this technology might be feasible.

LTE signals are interesting as well, since they performed the best in the previous set of experiments, and in addition to that, they can be used to correct the frequency offset of the sensors. However, the coverage of an LTE cell is much smaller thus dense deployments would be needed. To prove this intuition, we developed an Android app that outputs the current LTE Cell ID the phone is connected to as well as the GPS coordinates of the measurement. 2 trips were taken in the cities of San Sebastian and Alcorcon, with an area covered of 4.4 and 6 km^2 respectively, as shown in Figure 10.

To gain insigh on the results of these experiments, we aggregated the measurements by Cell ID and computed the maximum distance that could be observed between 2 measurements that shared the same Cell ID. The resulting cumulative distribution function is shown in Figure 11. For the case of San Sebastian, up to 64 different Cell IDS could be observed, whereas in the case of Alcorcon up to 20. The average

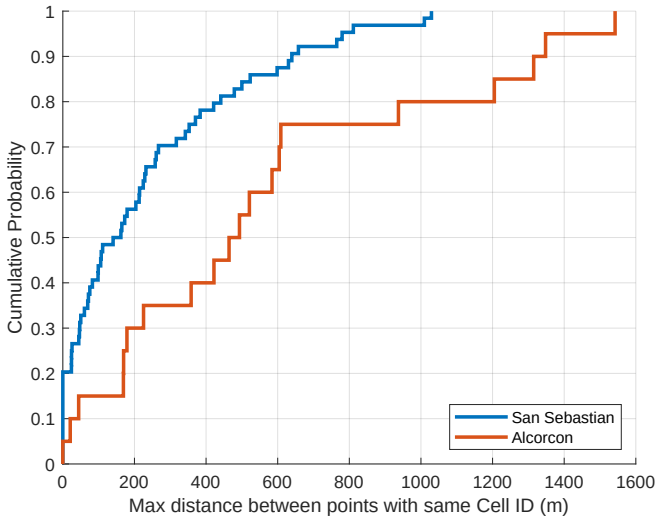


Fig. 11. ECDF for detected cells

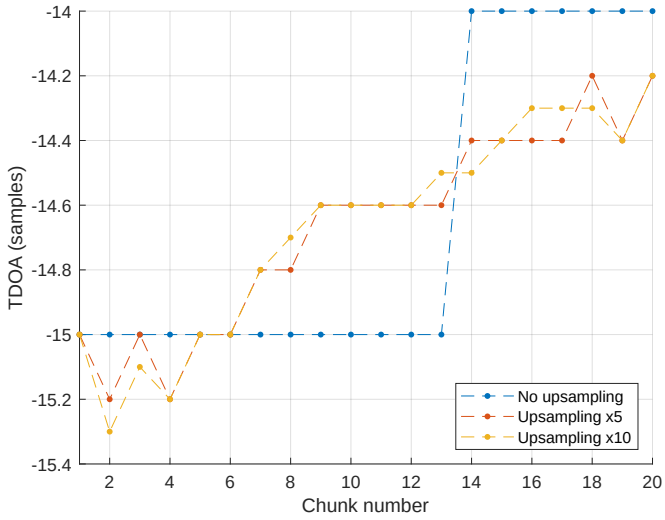


Fig. 13. Drift in delay caused by offsets in KiwiSDR sensors

maximum distance between measurements with same Cell ID was 236.6 m and 560.6 m respectively, but more than 50% of the cells fall in the range of less than 200 m for the case of San Sebastian and 500 m for Alcorcon.

What these results prove is that the intuition of the lower coverage of LTE was indeed correct, and the main conclusion that can be extracted is that to completely rely on LTE as a reference signal, an extremely dense deployment of sensors would be needed.

It is also important to note, that these results do not compromise the use of LTE signals to estimate the Local Oscillator Offset of the SDR receivers. For this case, each sensor can use any signal to estimate its own internal parameters and no interaction with other sensors is needed, thus the methods remains perfectly valid even though sensors use different LTE signals to calculate their clock offsets.

C. Multilateration Strategies

The last set of evaluations concerned the comparison of the different multilateration techniques aforementioned. And before real life deployments, we explored the use of real data like the one KiwiSDR provides. This platform offers a localization extension and allows users to download data for their own experiments. The sensors from KiwiSDR scan the spectrum from 0 to 30 MHz, thus they capture High Frequency (HF) bands. At these frequencies, transmitters can be seen from distances that range the hundreds of kilometers.

The data that comes from KiwiSDR is sampled at a frequency of 12 kHz and is GPS timestamped during 30 seconds, therefore no synchronization signal is needed. Data for each

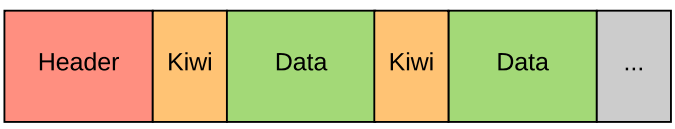


Fig. 12. KiwiSDR data structure

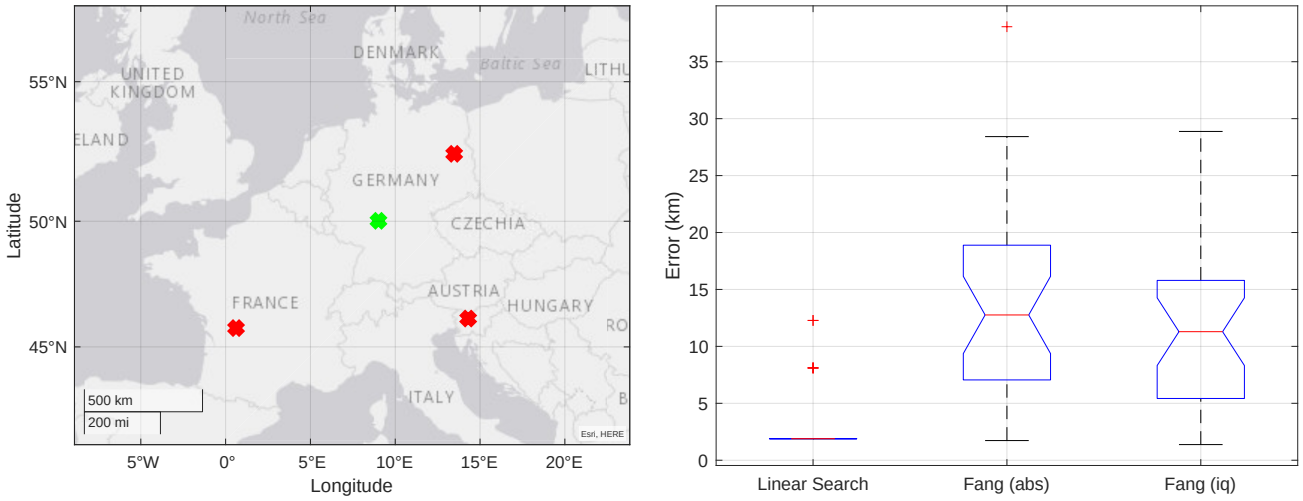
sensor is provided in `WAV` format and it consists of: 1) A general header with the location of the sensor and several parameters of the file. 2) Data chunks with 512 IQ samples. Preceeding each chunk there is a small header that contains the GPS timestamp of when the first sample of the chunk was received. Figure 12 shows a visual representation of the data structure.

Even though sensors are GPS synchronized, between sensors there is still a drift that affects the TDOA estimation between sensors. As an example, we divided the data incoming from 2 sensors tuned to the same transmitter into 20 different chunks and performed the cross-correlation between them. Figure 13 shows how the delay increases in a linear manner and also how upsampling the signal is able to capture this effect better. To correct this offset we added the average delay in each of the chunks to the delay obtained by cross-correlating in that chunk.

To compare the multilateration strategies we selected 3 sensors in Europe and a transmitter in Frankfurt, DCF77, a longwave time signal operating at 77.5 kHZ, with a bandwidth of 1 kHz and Amplitude Modulated. The location of the sensors and transmitter is shown in Figure 14a. We collected 30 traces per sensor and compared the Linear Search and the Fang's Algorithm results, summarizing the results in Figure 14b.

For Linear Search, most results are at 1 km from the actual position of the transmitter, independent of the correlation type. For Fang's Algorithm approach, the mean for `abs` correlation is 14.7 km and for `iq` correlation is 11.56 km. We also performed the tests with `dphase` but results were inaccurate since the signal is Amplitude Modulated and the phase did not carry information, therefore we have not included the results in this comparison. Even though the errors are in the order of km, since the distances involved are much larger, they are acceptable errors.

One of the reasons for Linear Search outperforming Fang's algorithm approach is that for the latter we did a conversion from geodetic coordinates to a planar space, whereas when



(a) Location of sensors (red) and transmitter (green)

(b) Comparison of strategies

Fig. 14. Selected KiwiSDR sensors and obtained results on localization

doing the Linear Search we can compute the distances on the ellipsoid directly.

VII. Conclusions and Future Work

In this thesis we have presented an architecture for a GPS-free transmitter localization system using low cost SDR receivers and offered some preliminary evaluations on the different aspects of it. These results indicate that a system of low-cost sensors may be feasible and capable of obtaining notable results in more real scenarios. One of the first conclusions is that correcting the sampling and frequency offsets improves notably the performance, as well as upsampling the signals. We observed as well that digital signals perform better than analog ones. It is also worth mentioning that the sphericity of the Earth might impact the accuracy of the localization, thus algorithms should be able to account for this.

Future works will look for real deployments utilizing the Electrosense hardware, testing the architecture as a whole. We will also look into modifications of the algorithms to account for Earth's geometry as well as newer and more accurate algorithms such as Gauss-Newton approaches [31].

Acknowledgment

This work was developed at IMDEA Networks Institute. I would like to show my gratitude to Domenico Giustiniano and Roberto Calvo-Palomino for their invaluable help and insights during the development of this thesis.

References

[1] J. Clement, "Mobile internet usage worldwide-Statistics & Facts," *Luettavissa*, vol. 16, p. 2020, 2019.

[2] W. A. Radasky, "Electromagnetic warfare is here," *IEEE Spectrum*, vol. 25, no. August, 2014.

[3] A. J. Kerns, D. P. Shepard, J. A. Bhatti, and T. E. Humphreys, "Unmanned aircraft capture and control via GPS spoofing," *Journal of Field Robotics*, vol. 31, no. 4, pp. 617–636, 2014.

[4] Google, "Google Spectrum Database." [Online]. Available: https://www.google.com/get/spectrumdatabase/

[5] Microsoft, "Microsoft Spectrum Observatory." [Online]. Available: http://spectrum-observatory.cloudapp.net/

[6] R. Calvo-Palomino, D. Giustiniano, V. Lenders, and A. Fakhreddine, "Crowdsourcing spectrum data decoding," *Proceedings - IEEE INFOCOM*, 2017.

[7] S. Rajendran, R. Calvo-Palomino, M. Fuchs, B. Van Den Bergh, H. Cordobes, D. Giustiniano, S. Pollin, and V. Lenders, "Electrosense: Open and Big Spectrum Data," *IEEE Communications Magazine*, vol. 56, no. 1, pp. 210–217, 2018.

[8] R. Calvo-Palomino, H. Cordobés, M. Engel, M. Fuchs, P. Jain, M. Liechti, S. Rajendran, M. Schäfer, B. V. den Bergh, S. Pollin, D. Giustiniano, and V. Lenders, "Electrosense+: Crowdsourcing radio spectrum decoding using IoT receivers," *Computer Networks*, vol. 174, 2020.

[9] F. Gustafsson and F. Gunnarsson, "Mobile positioning using wireless networks: Possibilities and fundamental limitations based on available wireless network measurements," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 41–53, 2005.

[10] N. A. Azmi, S. Samsul, Y. Yamada, M. F. Mohd Yakub, M. I. Mohd Ismail, and R. A. Dziyauddin, "A Survey of Localization using RSSI and TDoA Techniques in Wireless Sensor Network: System Architecture," *2018 2nd International Conference on Telematics and Future Generation Networks, TAFGEN 2018*, pp. 131–136, 2018.

[11] A. Zanella, "Best practice in RSS measurements and ranging," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 4, pp. 2662–2686, 2016.

[12] H. Nurminen, M. Dashti, and R. Piché, "A survey on wireless transmitter localization using signal strength measurements," *Wireless Communications and Mobile Computing*, vol. 2017, 2017.

[13] M. Khaledi, M. Khaledi, S. Sarkar, S. Kasera, N. Patwari, K. Derr, and S. Ramirez, "Simultaneous power-based localization of transmitters for crowdsourced spectrum monitoring," *Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM*, vol. Part F1312, pp. 235–247, 2017.

[14] X. Huang, K. Yan, H.-c. Wu, and Y. Wu, "Unmanned Aerial Vehicle Hub Detection Using Software-Defined Radio," in *2019 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*. IEEE, 2019, pp. 1–6.

[15] C. Y. Kim, D. Song, Y. Xu, J. Yi, and X. Wu, "Cooperative search of multiple unknown transient radio sources using multiple paired mobile robots," *IEEE Transactions on Robotics*, vol. 30, no. 5, pp. 1161–1173, 2014.

[16] C. Perkins, L. Lei, M. Kuhlman, T. H. Lee, G. Gateau, S. Bergbreiter, and P. Abshire, "Distance sensing for mini-robots: RSSI vs. TDOA," *Proceedings - IEEE International Symposium on Circuits and Systems*, no. May, pp. 1984–1987, 2011.

[17] M. Hartmann, T. Nowak, L. Patino-Studencki, J. Robert, A. Heuberger, and J. Thielecke, "A low-cost RSSI-based localization system for wildlife tracking," *IOP Conference Series: Materials Science and Engineering*, vol. 120, no. 1, 2016.

[18] Y. Wang and K. C. Ho, "Unified Near-Field and Far-Field Localization for AOA and Hybrid AOA-TDOA Positionings," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1242–1254, feb 2018.

[19] J. Yin, Q. Wan, S. Yang, and K. C. Ho, "A simple and accurate TDOA-AOA localization method using two stations," *IEEE Signal Processing Letters*, vol. 23, no. 1, pp. 144–148, 2016.

[20] D. Kim, K. E. Lee, and J. Kang, "Performance Analysis of AOA-based Localization with Software Defined Radio," *IGNSS Symposium 2015*, 2015.

[21] R. Kaune, "Accuracy studies for TDOA and TOA localization," *15th International Conference on Information Fusion, FUSION 2012*, pp. 408–415, 2012.

[22] Y. Wang and Y. Wu, "An efficient semidefinite relaxation algorithm for moving source localization using TDOA and FDOA measurements," *IEEE Communications Letters*, vol. 21, no. 1, pp. 80–83, 2017.

[23] A. Bhattacharya, C. Zhan, H. Gupta, S. R. Das, and P. M. Djuric, "Selection of sensors for efficient transmitter localization," in *IEEE INFOCOM*, 2020.

[24] J. Wei and C. Yu, "Performance analysis of TDoA based localization using SDRs," *2013 3rd Australian Control Conference, AUCC 2013*, pp. 91–92, 2013.

[25] F. Šturc, T. Morong, P. Kovář, and P. Puričer, "High performance SDR for monitoring system for GNSS jamming localization," *Proceedings - 2019 International Conference on Wireless Networks and Mobile Communications, WINCOM 2019*, pp. 1–5, 2019.

[26] R. Schreiber and J. Bajer, "Time difference measurement algorithm for TDOA positioning system using RTL-SDR," *ICMT 2017 - 6th International Conference on Military Technologies*, pp. 608–612, 2017.

[27] R. Calvo-Palomino, F. Ricciato, D. Giustiniano, and V. Lenders, "LTESS-track: A precise and fast frequency offset estimation for low-cost SDR platforms," *WiNTECH 2017 - Proceedings of the 11th Workshop on Wireless Network Testbeds, Experimental Evaluation and Characterization, co-located with MobiCom 2017*, pp. 51–58, 2017.

[28] M. Eichelberger, S. Tanner, K. Luchsinger, and R. Wattenhofer, "Indoor localization with aircraft signals," *SenSys 2017 - Proceedings of the 15th ACM Conference on Embedded Networked Sensor Systems*, vol. 2017-Janua, 2017.

[29] R. Calvo-Palomino, H. Cordobés, F. Ricciato, D. Giustiniano, and V. Lenders, "Collaborative wideband signal decoding using non-coherent receivers," *IPSN 2019 - Proceedings of the 2019 Information Processing in Sensor Networks*, pp. 37–48, 2019.

[30] B. T. Fang and Others, "Simple solutions for hyperbolic and related position fixes," *IEEE transactions on aerospace and electronic systems*, vol. 26, no. 5, pp. 748–753, 1990.

[31] N. Sirola, "Closed-form algorithms in mobile positioning: Myths and misconceptions," *Proceedings of the 2010 7th Workshop on Positioning, Navigation and Communication, WPNC'10*, pp. 38–44, 2010.