A Hybrid Approach to GPS-Free Geolocation over LoRa

JONAS DANEBJER
VALTHOR HALLDÓRSSON
MASTER´S THESIS
DEPARTMENT OF ELECTRICAL AND INFORMATION TECHNOLOGY
FACULTY OF ENGINEERING | LTH | LUND UNIVERSITY

A Hybrid Approach to GPS-Free Geolocation over LoRa

Jonas Danebjer
kin13jda@student.lu.se
Valthor Halldórsson
fys11sha@student.lu.se

Department of Electrical and Information Technology
Lund University

Supervisor: Flavius Gruian

Examiner: Erik Larsson

September 23, 2018

Printed in Sweden
Tryckeriet i E-huset, Lund

Abstract

The last few years have seen a rapid expansion within the area of the Internet of Things (IoT). For many IoT applications, such as asset tracking and distributed sensor networks, the problem of localization (i.e. determining the physical location of nodes in the network) is fundamental. For such applications, constraints on size, cost, and power consumption often prohibit the use of GPS. This thesis designs, implements, and evaluates a GPS-free localization system using the Long Range Radio (LoRa) technology, built using low cost off-the-shelf hardware. The system is based on combining Time of Flight (ToF) and Received Signal Strength Indicator (RSSI) distance measurements using a weighted multilateration estimator. By using an accelerometer that detects significant movement, subsequent measurements made in the same place can be leveraged to provide increased accuracy. The evaluation shows that the hybrid system outperforms both ToF and RSSI used on their own, achieving a mean stationary position error of 272 meters, which is comparable to other LoRa localization techniques.

Acknowledgements

We would like to thank Flavius Gruian for keeping us on track throughout the thesis, for his continuous and thorough feedback, and for generously lending us his personal LoRa equipment.

This thesis was conducted at Sigma Connectivity, who not only provided the necessary equipment and working environment, but also shared their invaluable expertise. We would especially like to thank our supervisor Arvid Elmér, for his guidance and support throughout the entire process, as well as Sadik Masar, whose expertise and willingness to help saved us many hours of radio debugging.

Jonas Danebjer & Valthor Halldorsson

GPS-free geolocation for low cost IoT devices

Jonas Danebjer, Valthor Halldorsson

Geolocation is a problem that for a long time has been solved entirely with the use of GPS. However, with the emergence of low cost Internet of Things (IoT) devices, the evolving demands placed on these devices require new solutions to old problems.

Introduction

With the emergence of the Internet of Things, a growing number of low cost devices intended to operate on their own for extended periods of time, often far away from WiFi access points. Long Range Radio (LoRa) is a radio technology that allows such devices to communicate across large distances (up to 20km under good conditions), using comparatively low power. The ability to geolocate these devices is often of interest in many applications such as for tracking moving objects or keeping track of sensor locations in large sensor networks. This is traditionally achieved using GPS modules on the devices, but adding such a module can be costly. Furthermore, a GPS module can significantly increase the overall power consumption, making some applications unfeasible. If geolocation can be done without requiring a GPS module, that would open up the possibility of low cost solutions to such problems.

GPS free geolocation

In this thesis, we present a system for localization using only a LoRa module. It uses both *time* and *signal strength* measurements to estimate distances between a sender and receiver, and combines them to produce an estimated position through a technique known as multilateration. By keeping track of whether the device has moved, the system can safely combine measurements that were taken in the same place, increasing the overall accuracy of the position estimation. As a result, it achieves an accuracy comparable with that of other LoRa geolocation systems, using only cheap hardware and simple setup procedures.

Conclusion

It is possible to achieve a reasonable position estimation using only the LoRa radio, but the use of cheap hardware has some clear drawbacks. Radio waves propagate at the speed of light, and for geolocation the travel time between two devices is very short. Since cheap clock hardware measures time in relatively large intervals (often in whole microseconds), this can cause large rounding errors when measuring the travel time between sender and receiver. With light traveling roughly 300m in a single microsecond, this rounding error alone can cause a large error in the estimated distance.

While taking multiple measurements allows the system to achieve better accuracy than a single measurement, the overall position accuracy remains in the area of a few hundred meters on average. With this in mind, use cases are those where only a rough location estimation is required, or where using a GPS simply is not an option due to battery or cost constraints.

Table of Contents

List of Figures

List of Tables

Acronyms

AM Amplitude modulation. 5

CSS Chirp spread spectrum. 5

ETSI European Telecommunications Standards Institute. 5, 6, 19

EU European Union. 5

FSK Frequency-shift keying. 5

Galileo European Global Navigation System. 11

GLONASS Globalnaya Navigatsionnaya Sputnikovaya Sistema. 11

GPS Global Positioning System. 3, 5, 11, 13, 14, 34, 35, 37

IoT Internet of Things. i, 3

LoRa Long Range Radio. i, v, 3, 4, 5, 6, 14, 34

NFER Near Field Electromagnetic Ranging. 8

RS Ranging Signal Message. 9, 14, 17, 34, 43

RSR Ranging Signal Response Message. 9, 14, 34, 43

RSSI Received Signal Strength Indicator. i, 8, 9, 13, 21, 34

SF Spreading Factor. 6

TDoA Time Difference on Arrival. 3, 4, 8

TM Timing Message. 9, 14, 19, 34, 43

ToA Time on Air. 6, 14

ToF Time of Flight. i, 3, 8, 9, 14, 15, 21, 22, 26, 34, 37

Zigbee Zigbee. 3

Chapter 1

Introduction

With the expansion of IoT markets, devices operating within previously uncharted territories are incrementally appearing. Using emerging radio technologies such as LoRa, Zigbee, and Bluetooth Mesh, many of these devices operate in a resource-constrained context, thus making the development of more resource efficient technologies more and more pressing.

LoRa is a radio technology that provides long range and low power consumption networking, at the cost of low data bandwidth. Because of these properties, LoRa is a good fit for technologies which require long range communications, but impose hard constraints on device sizes, power consumption, and cost.

One common feature of LoRa based applications is a reliance on some form of geolocation. Mature technologies such as Global Positioning System (GPS) that provide high accuracy exist, but the inclusion of a GPS module on a device incurs some costs for the manufacturer. For a device operating in a low energy context, the power requirements of such a module may necessitate larger and more expensive batteries, or require the device to be charged more often [1]. Further, the price of the GPS module itself and the costs of fitting it onto a board can be prohibitively expensive. Due to this, geolocation solutions that do not require the use of GPS (GPS-free) are of interest even in cases where their accuracy does not match that of GPS.

Implementing such a system is a challenging problem. Since power and cost constraints are a primary motivator for using LoRa localization, nodes cannot be required to possess expensive specialized hardware. Many LoRa applications also use low power microcontrollers that can not natively perform high precision time measurements, limiting the accuracy of time-measurement based localization techniques. Existing LoRa localization techniques are promising but do not meet the accuracy demands of many common applications.

This thesis presents a system that combines two existing distance estimation techniques into a hybrid system for an overall increase in accuracy. Further, it compares the accuracy of the hybrid system with the two techniques in isolation, and discusses its advantages and disadvantages.

1.1 Previous work

Many systems for outdoor positioning over radio technologies other than LoRa have been proposed, based on techniques including RSSI [2, 3, 4, 5, 6, 7, 8], ToF [9, 10, 11, 12], or Time Difference on Arrival (TDoA) [13, 14]. Work has also been done to combine multiple ranging techniques into hybrid systems [15, 16].

However, localization systems employing LoRa have received much attention in the past year. Recent research has in large part focused on TDoA as a positioning technique. Fargas et. al. [1] use TDoA, achieving a static error of around 100 meters using 4 gateways. In [17], TDoA localization was used to track moving nodes in a larger public LoRaWAN network, resulting in a median per-measurement accuracy of 200 meters. The accuracy of the position estimation was then further improved by accounting for the movement trajectory of the node.

The use of RSSI for localization in LoRa has also been explored. Aernouts et. al. [18] collected a large dataset of RSSI measurements over a period of several months, applying RSSI fingerprinting to achieve a mean per-measurement accuracy of 398 meters. Lam et. al. [19] also propose RSSI localization algorithms, based on detecting noisy gateways and excluding their measurements from the localization process.

1.2 Research questions

The purpose of this thesis is to examine the feasibility of a hybrid ToF and RSSI localization system over LoRa. The following research questions were investigated:

- What level of position accuracy can be achieved using only a LoRa radio and an accelerometer?
- Is a clock resolution of 1μs sufficient for ToF distance estimation to yield accuracy comparable to that of RSSI based distance estimations?
- Can ToF and RSSI based distance estimations be combined to provide an overall increase in position estimation accuracy?

1.3 Contributions

This thesis makes the following contributions:

- Evaluates the accuracy of ToF and RSSI as methods for estimating distances between nodes and gateways over LoRa.
- Proposes a hybrid method for combining ToF and RSSI distance estimations to a stationary node for increased overall accuracy.
- Evaluates the accuracy of the hybrid positioning method and demonstrates that accuracy can be increased by combining ToF and RSSI.

Chapter 2
Theory

To better understand the limitations of wireless communication and what effects this has on the domain of localization, a basic understanding of the phenomena affecting radio waves is helpful. This chapter presents an overview of factors that affect localization and their underlying causes, as well as an introduction to the techniques used for localization. It also touches on configurable parameters within LoRa that influences radio wave propagation.

2.1 Long Range Radio

The localization system presented in this thesis uses the LoRa radio technology, a radio modulation technique created by Semtech Corporation. It is intended to operate in low energy contexts and at long range, at the cost of low bandwidth [20]. In addition to data transfer, LoRa provides the basis for GPS-free geolocation functionality. The technology operates in the unlicensed spectrum which comes with both advantages and disadvantages. Anyone may put up a device that operates in the unlicensed spectrum as long as they follow the regional data restrictions. LoRa operates at 868MHz in Europe and is therefore limited to the duty cycle set by the European Telecommunications Standards Institute (ETSI) [21]. An important property of LoRa is that it uses a so called Chirp spread spectrum (CSS) modulation technique, which make radio waves sent over LoRa more resistant to narrow band interference compared to alternative modulation techniques such as Frequency-shift keying (FSK) or Amplitude modulation (AM). Narrow band interference usually spans a small range of frequencies, and CSS encodes the data of a wave over a wider range of frequencies making it more resistant to such interference.

There are a multitude of configurable and parameters for signals sent over LoRa, as well as certain limitations imposed either by the laws of physics or local regulations. Factors that are relevant for the proposed localization system are explained in detail below.

2.1.1 Duty Cycle

A duty cycle is the percentage of a time period a device is allowed to transmit information. This is regulated by telecommunications authorities, which for the

European Union (EU) is the European Telecommunications Standards Institute (ETSI). ETSI allows a maximum duty cycle of 1% for end devices (nodes). This limits how much data can be transmitted over a given amount of time, which directly influences the precision of the localization system. Since duty cycles revolve around a one hour period, a device transmitting over LoRa in the EU may only transmit for a total of 360 seconds every hour per channel.

2.1.2 Spreading Factor

The Spreading Factor (SF) is a configurable parameter that adjusts the transmission time of a signal transmitted over LoRa. It spreads the signal over the time, hence *spreading* factor. This value can be set to any value between 7 and 12 in steps of 1, where each step doubles (approximately) the time it takes to transmit the signal. A high spreading factor allows for a longer reach, at the cost of a lower data rate and increased transmission time. The opposite is true for a lower SF. Consequently, correctly configuring the SF for the intended purpose can be of great importance.

2.1.3 Time on Air

The Time on Air (ToA) is the time elapsed from the transmission of the first bit of information to the last bit [22]. The total ToA is the preamble transmission time plus the payload transmission time

$$T_{on\_air} = T_{payload} + T_{preamble} \tag{2.1}$$

These can be calculated using the following set of formulas:

$$T_{payload} = PL_{sym} \cdot T_{sym}$$

$$T_{preamble} = (P_{sym} + 4.25)T_{sym}$$

$$T_{sym} = \frac{2^{SF}}{BW}$$

$$PL_{sym} = 8 + max\left(ceil\left(\frac{8PL - 4SF + 28 + 16 - 20H}{4(SF - 2DE)}\right)(CR + 4), 0\right)$$

where SF is the spreading factor, BW is the bandwidth (125kHz for the EU), P_{sym} is the amount of preamble symbols (normally 8), PL is the packet payload size (bytes), H is the header option (0 if enabled, 1 otherwise), DE is data optimization (1 of enable, 0 otherwise) and CR is the coding rate used (between 1 and 4). Excluding the spreading factor, the values of these constants were set to the defaults provided by the firmware used.

Solving equation 2.1 with P_{sym} = 8, PL = 10, CR = 1,H = 1, SF = 12 ,BW = 125 and DR = 0 the total airtime is roughly 990ms. Adhering to duty cycle restrictions, this means that each device has to wait 99 seconds before sending another packet. Since a cycle is based around a one hour period, it is therefore possible to send multiple packets within a shorter time frame and still stay within the permitted airtime, although each node is required to wait for a longer period

afterwards to comply with the ETSI standard. From this follows a limitation in the number of measurements that can performed each day, impacts the overall accuracy of the hybrid localization system.

2.2 Path Loss Models

Electromagnetic fields propagating through free space dissipate with distance. This dissipation incurs a reduction in measured signal strength that is proportional to the inverse of the squared distance [23]. The total reduction in energy over a given distance defined as the path loss (PL). Some phenomena that cause this reduction of measured signal strengths are multipath fading, reflection, absorption, obstruction and interference. An estimation of how far a radio wave has propagated given a signal strength therefore has to incorporate the influence of these phenomena. Due to the inherent complexity in accounting for all the factors that affect signal strength, empirically derived models are most often used. These models are explicitly formulated to fit a certain type of environment, which also makes it impossible to choose a single model that fits all environments.

2.2.1 Log-distance path loss

A simple way of estimating the signal strength as a function of distance is by use of the Log-distance path loss model. It is defined as

$$PL = PL_{d_0} + 10\gamma \log \frac{d}{d_0} \tag{2.2}$$

where γ is the path loss exponent, d_0 is the reference distance (m), PL_{d_0} the path loss measured at the reference distance, and d is the distance of interest. Due to absorption in the material of an antenna and impedance in the circuitry, some of the energy diverted to the transmitter will not be radiated away in the form of radio waves. A path loss PL_{d_0} is therefore measured at the reference distance d_0 to incorporate the actual transmitted signal strength into the model.

When assuming the log distance path loss model, a path loss exponent is to be determined. This exponent can be derived from signal strength measurements in the local environment at points where the distance between them is known and chosen to best fit these measurements. While there are some empirically derived values for the path loss exponent suiting common environments, each area might introduce new factors that were previously unaccounted for. Table 2.1 gives a rough understanding of what the path loss exponent can be for various transmission frequencies and environments.

Table 2.1: Empirically measured path loss exponents for different environment.

Environment	Path loss Exponent	Transmission Frequency
Vacuum, Infinite free space	2	Any
Retail Stores	2.2	914 (MHz)
Grocery Store	1.8	914 (MHz)
Indoor Street	3	914 (MHz)
Urban area	2.7-3.5	Cellular Radio
Obstructed in building	4-6	Cellular Radio

2.2.2 Hata

To better incorporate the surround environment, more complex path loss models are often adopted. Some of these models are the Stanford University Interim model (SUI), Okamura's model and Hata model. The work of this thesis tested the Hata model for small cities [24]. It is formulated as follows:

$$L_H = 69.55 + 26.16 \log_{10} f - 13.82 \log_{10} h_t - a \cdot h_r + \big(44.9 - 6.55 \log(h_t)\big) \log(d) \quad (2.3)$$

where L_H is the path loss in urban environments (dB), f the transmission frequency (MHz), h_t the height of the transmitter (m), h_r the height of the receiver (m), a the antenna correction factor, d the distance between transmitter and receiver (km).

The following correction factor is suitable for small to medium sized cities:

$$a = (1.1 \log_{10} f - 0.7) h_t - (1.56 \log_{10} f - 0.8)$$

This model can be modified to better fit suburban environments with the following formulation:

$$L_{SU} = L_U - 2(\log_{10} \frac{f}{28})^2 - 5.4$$

where L_{SU} is the suburban path loss (dB), L_U the path loss from the urban Hata model (dB) and f the transmission frequency (MHz).

Alternatively, to fit rural environments, the Hata model can be modified as follows:

$$L_R = L_U - 4.78(\log_{10} f)^2 + 18.33 \log_{10} f - 40.94$$

where L_R is path loss for rural areas (dB), L_U the path loss from the urban Hata model (dB) and f the transmission frequency (MHz).

2.3 Distance Estimation

There are multiple ways to estimate distances between nodes in a radio network, including Time of Flight (ToF), Time Difference on Arrival (TDoA), RSSI and

Near Field Electromagnetic Ranging (NFER) [9]. This section focuses on ToF and RSSI, since these are the techniques employed in this thesis.

2.3.1 Time of Flight

The time of flight is the time it takes for a signal to travel from one point to another. In the context of this thesis these points are represented by a node and a gateway. Having measured the time of flight and by knowing the position of either device, the distance Δd between them can be calculated using

$$\Delta d = c \cdot t$$

where c is the speed of light and t the measured time of flight. This can be computed using two devices with synchronized clocks by applying following equation

$$\Delta d = c(T_2 - T_1)$$

where T_1 is the time at which the signal left the transmitting node, and T_2 is the time at which the signal was received by the other node. In order to avoid the need for clock synchronization, the two way ranging protocol shown in figure 2.1 can instead be used to calculate the ToF. It works as follows:

- Node A sends a ranging signal and stores the time of departure for that signal as T_{rsA}
- Node B stores the time of reception for the ranging signal as T_{rsB}
- Node B sends a ranging signal response and stores the time of departure as T_{rspB}, and calculates $T_{processing} = T_{rspB} - T_{rsB}$
- Node A stores the time of reception for the ranging signal response as T_{rspA} and calculates $T_{roundtrip} = T_{rspA} - T_{rsA}$
- Node A sends a timing message containing $T_{roundtrip}$ to Node B
- Node B receives $T_{roundtrip}$ and can now calculate the time of flight as

$$ToF = \frac{T_{roundtrip} - T_{processing}}{2}$$

The accuracy of the ToF depends on the accuracy of each time measurement taken during the protocol run. These in turn depend on the precision of the hardware timer used. Factors such as interference and reflection may also impose further variation on the measured timestamp values.

2.3.2 Received Signal Strength Indicator

Another way of estimating the distance between a transmitter and receiver is by use of a path loss model that predicts the loss of signal strength over a distance. By measuring the strength of a signal and comparing with the model, the estimated distance to the sender can be calculated. The estimated distance then depends directly on the chosen path loss model, which is an approximation that itself is

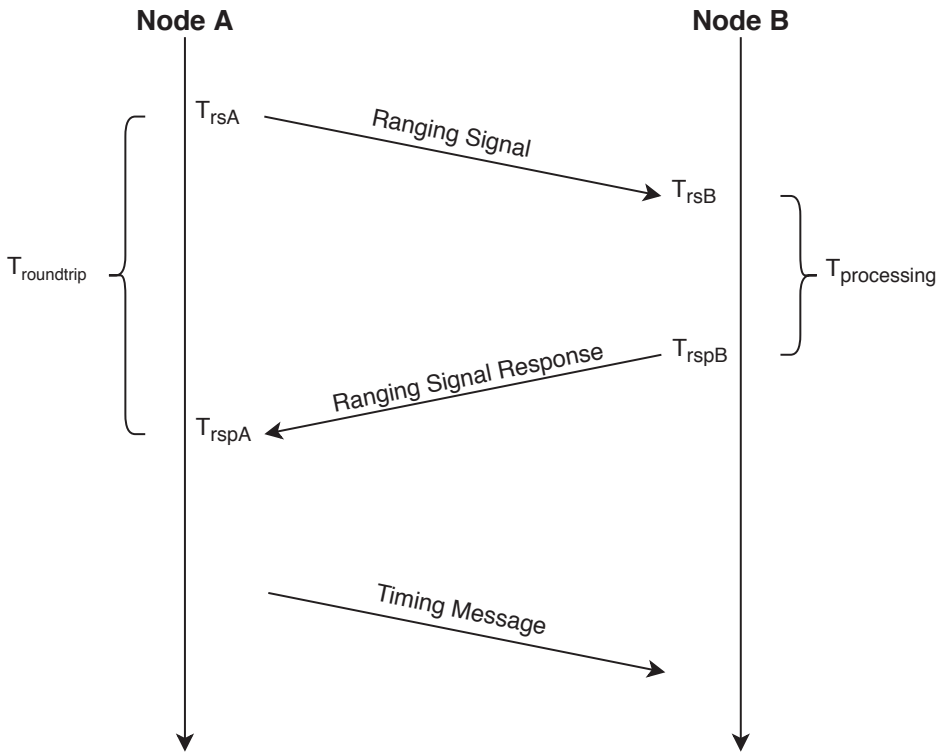

Figure 2.1: A visual overview of the two way ToF protocol, with time flowing from top to bottom. The three different protocol messages Ranging Signal Message (RS), Ranging Signal Response Message (RSR) and Timing Message (TM) can be seen in sequence.

subject to noise due from interference and obstruction. If, for example, the receiver is placed behind a house where the signals are obstructed and consequently the signal strength reduced, then the distance predicted by the model will be much larger than in reality.

To estimate the distance using the Hata model, one can solve equation 2.3 for d, yielding

$$d = 10^{(L_H - 69.55 - 26.16 \log_{10} f + 13.82 \log_{10} h_t + a \cdot h_r)/(44.9 - 6.55 \log(h_t))} \tag{2.4}$$

This can then be directly used to estimate the distance for a measured signal strength. To get the distance estimation for the log distance path loss model, one can solve for d in equation 2.2, yielding

$$d = d_0 \cdot 10^{(PL - PL_{d_0})/(10\gamma)} \tag{2.5}$$

2.4 Position Estimation

The position p of an object can be represented using a vector of coordinates $(x_1, x_2, ..x_n)$ in an n-dimensional coordinate system. In a geographical coordinate system $n = 2$ if elevation is omitted, which is commonly represented using latitude and longitude (lon, lat). By use of a set of reference lines known as the equatorial plane and the prime meridian, latitude is defined as the angular divergence from the equatorial plane and longitude is defined as the angular divergence from the prime meridian. Estimation of (lon, lat) for an object can be done through several different techniques, and the choice of which to use can depend on factors such as what type of data is measurable and the accuracy of those measurements.

The three most common types of position estimation techniques are triangulation, trilateration and multilateration [1]. To find (lon, lat), triangulation requires the angle between the transmitter and two receivers, as well as the geographical position of the two receivers. Estimation of the position then becomes a problem of solving trigonometric equations.

Trilateration on the other hand is achieved with distance measurements, in contrast to angular measurements, but requires three such measurements to determine (lon, lat). This is the techniques used by global navigation satellite systems such as Global Positioning System (GPS), Globalnaya Navigatsionnaya Sputnikovaya Sistema (GLONASS) and European Global Navigation System (Galileo) [25][26][27], which can achieve an average precision of down to 4.9m for smartphones [28]. More generally, for $n \geq 3$ distinct distance measurements this process is called multilateration.

2.4.1 Multilateration

In a 2D coordinate system the distance r from point p specifies a circle of possible positions as can be seen in figure 2.2.

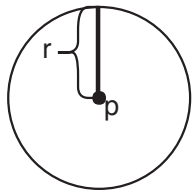

Figure 2.2: Possible positions given a single distance. With the exact known distance r from point p, the circle represent all possible positions at this distance in a 2D environment.

In practice, distance measurements are subject to noise which gives rise to an area of uncertainty at the boundary of each circle. When three or more such approximate distances are known, an area of confidence can be determined if these boundaries overlap (see figure 2.3). The position is then guaranteed to be within this area if the uncertainty of the error is known.

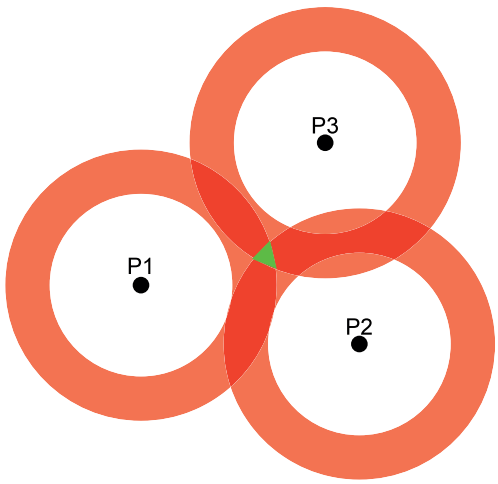

Figure 2.3: Illustration of multilateration. The borders of each circle represents the uncertainty of a distance estimation originating from points P1, P2, and P3. The highlighted intersection between all borders is the area of confidence.

This becomes an over determined system if the number of measurements m $m > d+1$, where d is the dimensionality of the position (2 in the case of geographical coordinates without altitude). This thesis uses the iterative Gauss-Newton algorithm [29] to solve this system. It solves the problem by iteratively finding the value of the variables that minimizes the sum of the squares of the difference between the variables and the measurements.

Consider three circles whose center lies a $C1 = (x_1, y_1), C2 = (x_2, y_2), C3 = (x_3, y_3)$ and with radii $R1$, $R2$, $R3$, that correspond to distance estimations. Finding the position $P = (x, y)$ that satisfies the distance estimation represented by $P1$ is then done by solving

$$\sqrt{(x_1 - x)^2 + (y_1 - y)^2} = R1$$

in which the solution is found with any point that lies on the circle. This can be rewritten as

$$r1(x, y) = \sqrt{(x_1 - x)^2 + (y_1 - y)^2} - R1$$

where r1(x,y) is the residual error of the solution. In the case where the point lies on the circle, $r1(x, y) = 0$.

The iterative Gauss-Newton algorithm finds the point (x, y) that minimizes multiple such residual errors [30]:

$$\begin{aligned} r1(x, y) &= \sqrt{(x_1 - x)^2 + (y_1 - y)^2} - R1 \\ r2(x, y) &= \sqrt{(x_2 - x)^2 + (y_2 - y)^2} - R2 \\ r3(x, y) &= \sqrt{(x_3 - x)^2 + (y_3 - y)^2} - R3 \end{aligned}$$

By applying Gauss-Newton to a collection of distance measurements with known origin, an estimated position within the area of confidence can be calculated. While Gauss-Newton finds an approximate solution given already error prone distance measurements, improvement on the position estimation can be made by improving the accuracy of the individual distance estimations. However, the distance estimations from each gateway might not overlap (in terms of the borders shown in figure 2.3), in which case the Gauss-Newton iteration will still terminate with a position that minimizes the residual errors.

2.4.2 Fingerprinting

By mapping an area and measuring relevant sensor data at various locations, a measurement signature each location can be stored in a database. Upon taking a new measurement, the signature of that measurement can be compared with those stored in the database, and the closest match can be found. This technique is known as fingerprinting and is often used with RSSI measurements [2]. The duty cycle restrictions for LoRa transmission make the data gathering process for fingerprinting time consuming. Due to this, the method was not employed in the implementation.

2.5 Power Consumption

Power consumption is of special relevance when comparing the systems overall consumption with that of GPS. The following subsections give the theoretical

consumption as defined by the manufacturer's datasheet. Note that power consumption is commonly specified using only the current consumption, under the assumption that a fixed voltage is used. For all energy calculations described in this thesis, a constant voltage of 3.3V is assumed.

2.5.1 Accelerometer

The power consumed by the LIS2HH12 accelerometer used in this study depends on the sampling frequency used, as can be seen in table 2.2.

Table 2.2: The power consumption of a LIS2HH12 accelerometer for various sampling frequencies [31] at a voltage of 3.3V.

Sampling frequency (Hz)	Power consumption (µA)
Idle	5
10	50
50	110
100	180
200	180
400	180
800	180

2.5.2 LoRa vs GPS

The average power consumption of the LoRa radio when used with the EU configuration is 38.9mA when transmitting and 14.2mA when receiving (both at 3.3V) [32]. In standby mode, when neither transmitting nor receiving, the radio module consumes 2.8mA. Using the LoRa configuration from table 3.1, the ToA of a RS is approximately 827ms. Combined with the TM of 10 bytes, this sums to a total of $992ms + 827ms = 1819ms$ of transmission time for the node during a single ToF iteration. The node also receives a single RSR, which has the same size as the RS (approximately 827ms).

In total, the theoretical energy usage of a single ToF run for a node sums up to $1.819s \cdot 38.9mA + 0.827s \cdot 14.2mA \approx 82.5mAs$.

The power consumption and operational details of GPS modules vary between different products. Using the GPS module found on the LoPy4 as a reference, it takes approximately 35s for the module to lock on to a first GPS position after starting from a powered down state [33]. During this cold start, the module draws a current of 29mA, adding up to $35s \cdot 29mA = 1015mAs$. Once a fix has been found, the GPS module drops its consumption to 19mA while continuously tracking.

Chapter 3
Implementation

This chapter describes the design and motivations behind the localization system implemented in this thesis.

3.1 Hardware

The localization system was implemented using LoPy4 (see figure 3.1) devices mounted on PyTrack (figure 3.2) development boards. The LoPy4 and PyTrack were chosen for a number of reasons relating to practicality:

- The LoPy4 runs programs written in the Python programming language, allowing rapid prototyping and development.
- The LoPy4 is capable of acting both as a device and as a gateway, allowing code reuse between both sides of the system.
- The PyTrack development board contains both an accelerometer and a GPS module. The accelerometer was used to enhance position accuracy, while the GPS module provided a ground truth position to be used for evaluation.

The board includes a Semtech LoRa transciever SX1276 radio, to which an SMA Tilt Swivel 1/2 Wave Whip Dipole antenna was externally attached. The underlying chipset for the development board is an Espressif ESP32.

3.1.1 Firmware

The LoPy4 firmware associates a microsecond precision timestamp to each *received packet*. This timestamp is then made available to the python application by means of an API call. In contrast, *sent packets* are not timestamped at the firmware level.

The ToF method requires timestamps for both sent and received packets, which requires modification of the original firmware. It is possible to manually timestamp in the python code before and after calling the firmware's *send packet* method from the application level, but there is an unpredictable time delay between the time when the send call returns and the radio finishes transmitting the message. While this delay is short, even microseconds of added variance in execution times can cause problems for ToF time measurements.

To address this, the LoPy4 devices were flashed with a modified micropython-sigfox-pycom firmware (version v1.17.3.b1), following the official guide presented at

Figure 3.1: The LoPy4 development board [34].

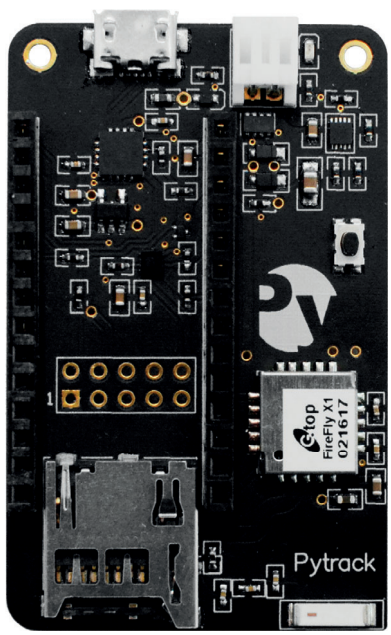

Figure 3.2: The Pytrack sensor shield [35].

`https://github.com/pycom/pycom-micropython-sigfox`. This implementation has been modified to register a timestamp after each successful packet transmission. This time critical source code modification was made to the `sx1276.c` source file, and can be seen in appendix A.

3.2 Localization system design

The localization system consists of the LoPy4 devices and a backend server. A total of four devices were used, allowing for three gateways and one roaming node. This is the minimum number of devices required to acquire a position estimation.

The LoPy4 devices were configured to use raw LoRa rather than transmitting over LoRaWAN. While LoRaWAN offers some desirable features, such as automatic CRC checks and adaptive data rates (i.e. automatically lowering spreading factors when lower spreading factors are sufficient), LoRaWAN device classes impose limitations regarding the duration of reception windows after transmissions. Since the ranging protocol is sensitive to time factors, raw LoRa was chosen to simplify the implementation.

3.2.1 Ranging protocol

When the node is powered up it attempts to initiate the ranging protocol by and broadcasting a Ranging Signal Message (see figure 2.1) to a gateway. Because the gateway must respond immediately after receiving the Ranging Signal Message, there is not enough time for the gateways to coordinate in order to dynamically select only one gateway to respond. For that reason, the Ranging Signal Message is explicitly addressed to a single gateway. The node iterates through a list of known gateway IDs, attempting to reach each gateway in order. If the gateway does not respond within a set timeout period, the node moves on to the next one.

The list of known gateways IDs is set manually when the node is flashed. This list could instead be maintained dynamically by the node, but the details of such an implementation may depend on the specific use case. Since data bandwidth is limited, IDs must be kept small. One implication of this is that for a large network with many gateways, IDs should preferably be dynamically allocated, ensuring gateways and nodes are locally unique.

Figure 3.3 illustrates the ranging process. When a gateway receives a Ranging Signal Message addressed to itself it immediately responds, continuing the ranging protocol. Once finished, it sends the resulting ToF measurement to the backend, where it is stored in an in-memory database and analyzed. When a gateway receives a node transmission addressed to a gateway other than itself, it records the transmission's RSSI and forwards it to the backend. For the system to be able to correlate the various incoming RSSI and ToF measurements belonging to a node, the node must prefix each message with its own unique node ID. As with the gateway IDs, dynamically allocating unique IDs presents a challenge which is left unaddressed in this thesis. Allocating locally unique node IDs is especially difficult in large networks, since node movement could require the node ID to be reissued to avoid conflicts.

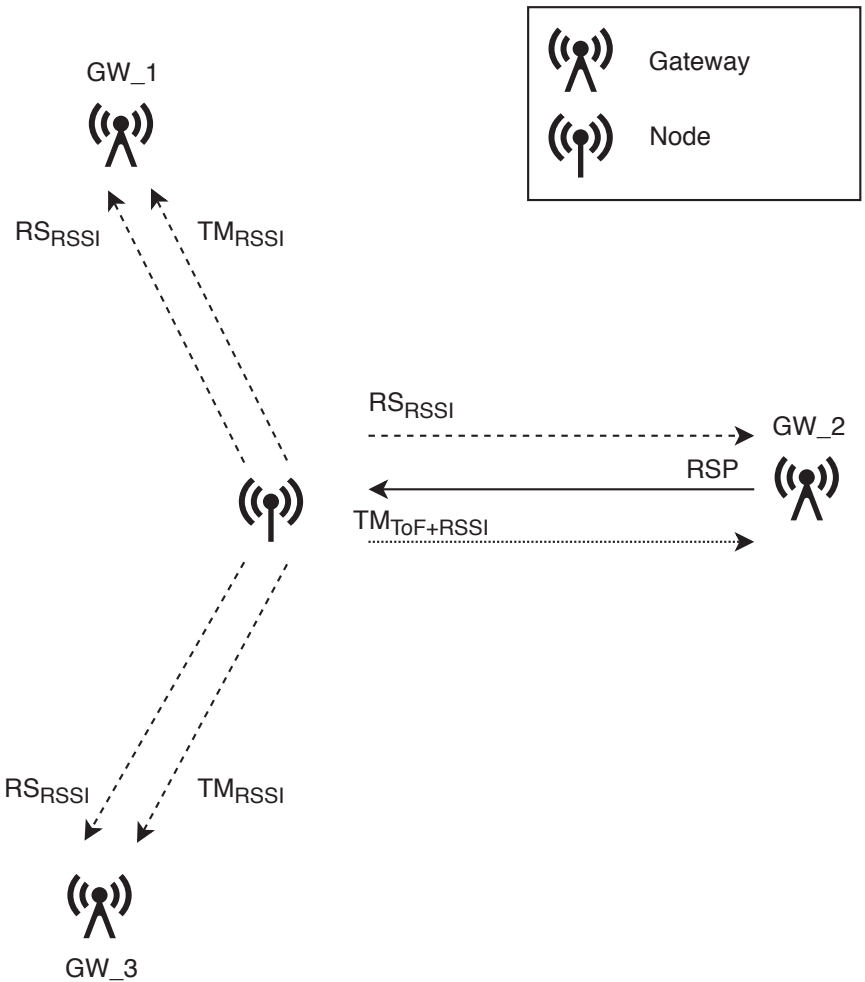

Figure 3.3: Signal transmissions of a full ranging protocol iteration. All gateways within reach measure the signal strength of signals sent by the node. The gateway involved in the current run of the ranging protocol measures both ToF time and the signal strengths.

Before attempting to start the ranging protocol with a gateway, the node generates a unique session counter that is associated with the current ranging protocol iteration. This session counter is not strictly necessary for the localization to function, but it allows for the gateways' measurements to be retroactively paired with the ground truth GPS locations taken by the node.

3.2.2 Packet structure

Since ETSI limits the transmission time of any single device operating in the 868MHz range to a maximum transmission rate of 1% per hour, minimization of packet size becomes an important factor to take into consideration. The ranging protocol implementation tries to minimize the packet size while still making the data collection and data analysis steps as effortless as possible.

The packet sizes (see appendix B) entail a total time on air of approximately 2 seconds per ranging protocol run per device using spreading factor 12. Consequently, the ranging protocol was only allowed to run once every 200 seconds, adding up to a total of 17 iterations an hour.

3.2.3 Movement detection

For the localization system to perform multilateration on a node over multiple measurements, it must know that the node has not moved in between ranging protocol iterations with the three gateways.

To achieve this, the node keeps track of when it is moved using the accelerometer on the PyTrack. The accelerometer is polled at regular intervals, and if the measured acceleration exceeds a threshold, it is registered as moving. Once the next ranging protocol finishes, the level of movement detected (encoded as the relative share of accelerometer polls above the threshold since the last ranging protocol) is sent sent to the gateway in the Timing Message.

The threshold was set conservatively to ensure that the movement detector would not yield false negatives.

3.2.4 LoRa parameters

Table 3.1 shows the parameters used for the LoRa radio. Notably, the spreading factor of 12 was chosen to maximize the range of the signal, at the cost of increased transmission times (and thereby slower ranging protocol iterations). The coding rate of 1 (sometimes described as 4/5) was used to minimize the redundancy (forward error correction) in the message, speeding up transmission.

Table 3.1: Configuration parameters used for the LoRa radios.

Parameter	Value
Frequency	868 (MHz)
Spreading factor	12
Bandwidth	125 (kHz)
Preamble size	8
Data optimization	off
Coding rate	1
Explicit header	off

3.3 Backend

The backend server is responsible for collecting the measurements made by the gateways and using the measurements to perform distance and position estimations.

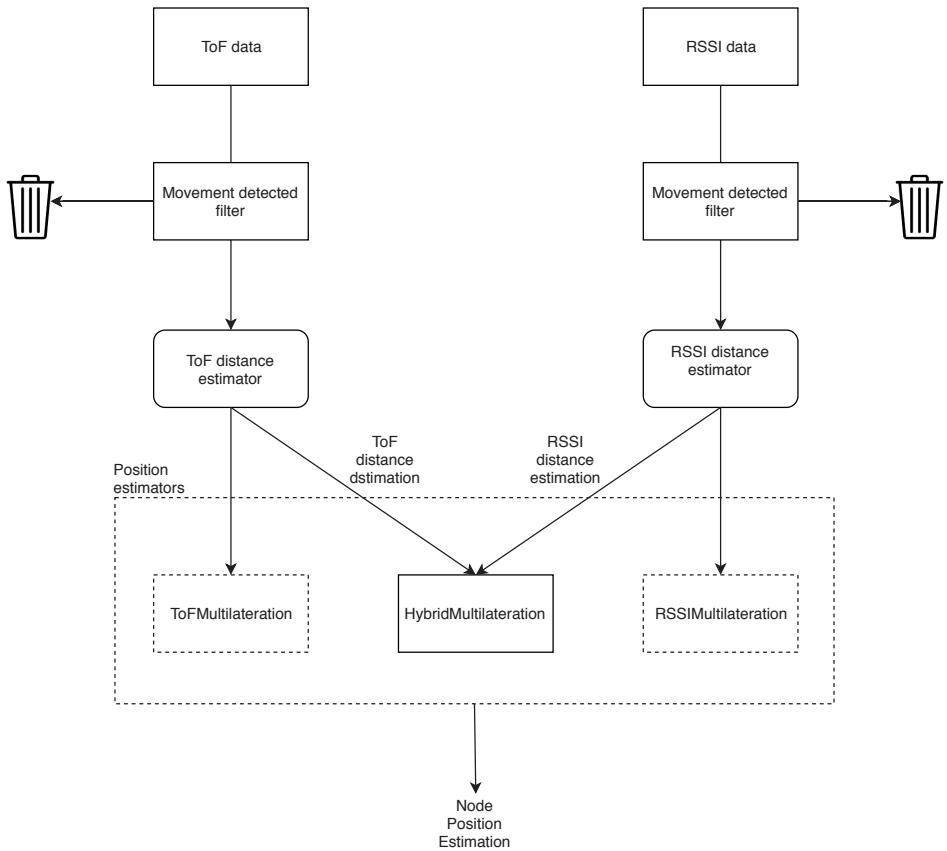

Figure 3.4: An overview of the data flow on the backend server.

When incoming measurements arrive at the backend, they are grouped into *movement chunks* - a series of measurements taken over a time period where no movement was detected. Position estimation is then seen as a set of functions taking a movement chunk as input and yielding a position estimate as an output. Each function first converts the set of measurements in the movement chunk to gateway distance estimates before applying multilateration to find a position. This process is illustrated in figure 3.4.

When testing the ranging protocol, there was a need to monitor whether packets sent by the roaming node were being reliably received by the various gateways, and vice versa. Since the sender does not receive an acknowledgement, this is not directly supported by the ranging protocol. Instead, each device was configured to stream its logs to the backend. These logs were then made viewable in a web application, allowing the status of remote gateways to be monitored while measurements were being taken in the field. This information was then used to find node locations where all gateways were reachable.

Devices were connected to the internet over WiFi. In cases where the devices did not have access to a WiFi network (both the node and some gateway locations, such as many of the rural gateway locations), mobile phones were used as ad-hoc access points and left with the device.

3.3.1 Ranging protocol parameters

In order to perform distance estimation with ToF or RSSI, some parameters had to be empirically determined. In the case of ToF, the measured time t_{ToF} includes the time required to transmit the entire message (which is large in comparison to the time it takes for the signal to propagate to the receiver) as well as other static delays. The base time offset t_0 was determined by minimizing the squared error of distance predictions.

Similarly, for the log distance RSSI path loss model the parameter γ was determined by using a least square error fit over the measured RSSI data. A separate γ value was calculated for each of the datasets, and both of the exponents were evaluated for position estimation.

To provide the RSSI models with a reference strength for the LoPy4 node, measurements were taken at a true distance of 1km, and the mean RSSI value was used as a reference strength.

Table 3.2 shows the results of reference measurements as well as model parameters derived from the measurement data.

Table 3.2: Model parameter values for all distance estimators.

Model parameter	Value
Suburban RSSI reference strength (1km)	-118.49 dBm
Rural RSSI reference strength (1km)	-115.92 dBm
Best fit ToF base time offset	1.01891501 s
Best fit log distance path loss exponent (rural)	2.55
Best fit log distance path loss exponent (suburban)	2.04

3.3.2 ToF distance estimation

Given a measured ToF time t_{tof}, the estimated distance $\hat{d}$ is calculated as:

$$\hat{d} = (t_{tof} - t_0)c$$

Where c is the speed of light and t_0 is the base time offset extracted from the empirical measurements.

3.3.3 RSSI distance estimation

Given an RSSI measurement r, the distance estimation $\hat{d}$ is calculated using two different path loss models. Previous work has concluded that it is not always apparent which choice of path loss models is suitable [36]. The authors selected the Hata model due to its previous success within the working frequency range of LoRa, as well as the commonly used log distance path loss model that allows for a configurable path loss exponent.

3.3.4 Position estimation

The position estimators can be divided into three groups: pure ToF estimators, pure RSSI estimators, and hybrid estimators. Each estimator can be seen as a function from a single movement chunk to an estimated position. Each estimator also takes some number of parameters specific to the techniques it applies.

The position estimators used were the following:

ToFMultilateration(t_0) Pure time of flight position estimator using the base time offset t_0 (to be subtracted from the ToF time). Uses multilateration to estimate a position given the distances.

RSSIMultilateration(M) Pure RSSI position estimator using the path loss model M. Uses multilateration to estimate a position given the distances.

HybridMultilateration(t_0, M, w_{ToF}) Hybrid position estimator that first reduces the measurements in a movement chunk to one ToF time and one RSSI value per gateway by the *pruned mean function*. Then calculates distance estimates from the result using the base time offset t_0 and the RSSI path loss model M as described above. Finally, weighted multilateration is used to estimate a position given the distances. In order to control the relative importance given to the various distance estimators during error minimization, the error contribution from all ToF distances is multiplied by a factor w_{ToF}, and the error contribution from RSSI distances is multiplied by $1 - w_{ToF}$. Setting w_{ToF} to 1 will cause the optimization process to discard the contribution of the RSSI measurements entirely, while a setting of 0 will discard all of the ToF measurements.

Since extreme outliers appear in the measured data, all estimators rely on a *pruned mean function* that reduces the set of all measurements to a single value per gateway. This function works by trimming all values outside of a set threshold from the median (20 microseconds for ToF, and 25 dBm for RSSI), and taking

the mean of the remaining values. These limits were set to include all common variation observed, excluding only values that were abnormally high.

Chapter 4
Evaluation

This chapter describes how the localization system was evaluated, presents the results of the empirical measurements, and finally discusses the implications of those results.

4.1 Methodology

In order to assess the accuracy of the hybrid position estimator, a series of empirical tests were performed. This section describes how these tests were conducted, and how their results were evaluated.

4.1.1 Data gathering

Measurements were carried out in two different environments. One being a rural setting with gateways spread 2-3 kilometers apart and the other a suburban environment with gateways spread 1-2 kilometers apart.

Over multiple days, three gateways were placed at high points in the surrounding environment, and a roaming node was moved in the area between the gateways. The node was left in place for extended periods of time without movement such that multiple measurement could be taken in the same place. The node's movement detector was used to tag the logged data with information about whether the node had moved since the last measurement.

Both nodes and gateways stored all measurement data that was generated while running the ranging protocol to an SD card in the form of JSON logs (see listings 4.1 and 4.2). The logs were also sent over the network to the backend, allowing them to be remotely monitored while in the field. In order to be able to correlate measurements taken by the different devices (for example to identify which RSSI measurements belong to which ranging protocol iteration) the session counter field of the transmitted packet was used. To ensure that the session counter remained unique, the node retrieved the highest recorded session counter from the network server upon start up.

In order to provide the measurements with an associated ground truth, the node's GPS module was used. Before each ranging protocol iteration, a fresh GPS position was collected and then logged along with the measurement. The true distances were produced by applying the Haversine distance formula to the

gateway's location and the node's measured GPS location. In order to ensure that the node had not moved since the GPS location was taken, the ranging protocol was run only once directly after the node received each new GPS location.

Listing 4.1: ToF data bundle for a single ranging protocol iteration.

```
{
  "node_gps": [55.718868, 13.228985],
  "gateway_gps": [55.716397, 13.232338],
  "gateway_id": 1,
  "tof_time": 1.018919,
  "session_counter": 0,
  "motion_detected_ratio": 0.4980392,
}
```

Listing 4.2: Processed RSSI measurement chunks. The json object displayed below is the data structure after chunking the RSSI measurements into bundles where no movement has occured. This is the training data and includes the true position of the node.

```
{
  "node_gps": [55.718868, 13.228985],
  "rssis":  [
  {"gateway_id": 1, "gateway_gps": [55.716397, 13.232338],
     "rssi": -101, "gateway_gps_age": 4.324, "session_id":
      0},
  {"gateway_id": 2, "gateway_gps": [55.715852, 13.230311],
     "rssi": -103, "gateway_gps_age": 23.24092, "
     session_id": 0},
  {"gateway_id": 3, "gateway_gps": [55.715852, 13.240311],
     "rssi": -93, "gateway_gps_age": 223.541, "session_id"
     : 0}
           ]
}
```

4.1.2 Data pruning

Since many measurements included extreme levels of noise (often by orders of magnitude), some pruning of the data was required before analysis. ToF times corresponding to distances larger than 20 kilometers were removed. In addition, faulty values of the true distances occurring due to GPS issues were excluded.

This pruning was not applied when evaluating position estimators. Rather, the position estimators rely on the *pruned mean function* to filter out outliers in a movement chunk.

4.1.3 Data analysis

Before attempting to combine ToF and RSSI measurements into a hybrid system, the behavior of each distance estimation method was investigated on its own. Analysis was conducted to examine how distance estimation errors were distributed, whether errors depended on the true distance between the node and the gateway, and which RSSI path loss models best described the measurements.

Because the ToF protocol requires multiple iterations of measurements to reach usable levels of precision, emphasis was put on analysing the *stationary error* (where the node makes multiple measurements without movement) and not the *per-measurement* error (the error for a single measurement).

In order to apply the hybrid multilateration algorithm, relative weights had to be given to the two estimation methods. Exactly how the weights should be chosen for optimal results is a research question on its own. For the sake of simplicity, it was assumed that the weights should be equal. To understand the impact of this decision, the relationship between the positioning error and the chosen weights was investigated, but the results were not used to improve position accuracy due to the risk of overfitting the test data.

4.2 Results

This section presents the results obtained from empirically evaluating the localization system as well as its distance estimation methods.

4.2.1 ToF distance estimation

As shown in figure 4.1, there is a linear relationship between the times measured by the ranging protocol and the true distance between the node and gateway. The time of flight increase at a rough rate of $1/c$ seconds per meter of added distance. The figure also shows that there is considerable variance in the measurements. The discrete steps on the time axis are due to the microsecond precision of the timestamps. Figure 4.2 shows that the time errors correspond to distance errors of up to 2000 meters. The distance errors can roughly be described as being normally distributed with an expected value of 22 nanometers and a standard deviation of 766 meters. The expected value is thus close to zero.

Figure 4.3 shows the error as a function of true distance between node and gateway. Because this graph shows the distances between the points and the line in figure 4.1, the discrete measurement levels here appear rotated along the slope of the fitted line in the original graph. The two lines show the best (lowest square error) fit to either the whole dataset, or only to measurements taken at distances larger than 300 meters. The difference in slope illustrates that there is a positive error bias at shorter ranges, but that for longer distances the error stabilizes around a zero mean. This indicates that ToF distance measurements are stable for use at longer distances.

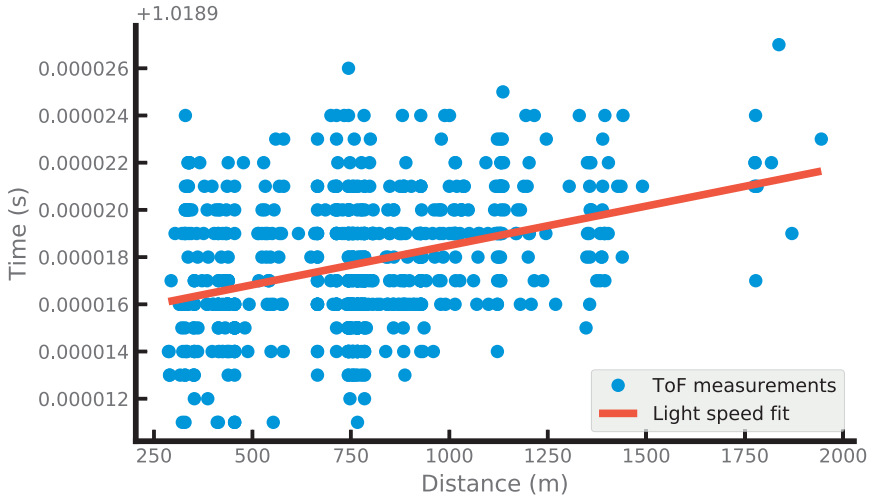

Figure 4.1: ToF time measurements as a function of true distance (as measured by GPS). The fitted line has a fixed slope of $1/c$ and a y-intercept of 1.0189150. The discrete spacing of the measurement arise from the microsecond resolution of the hardware timer.

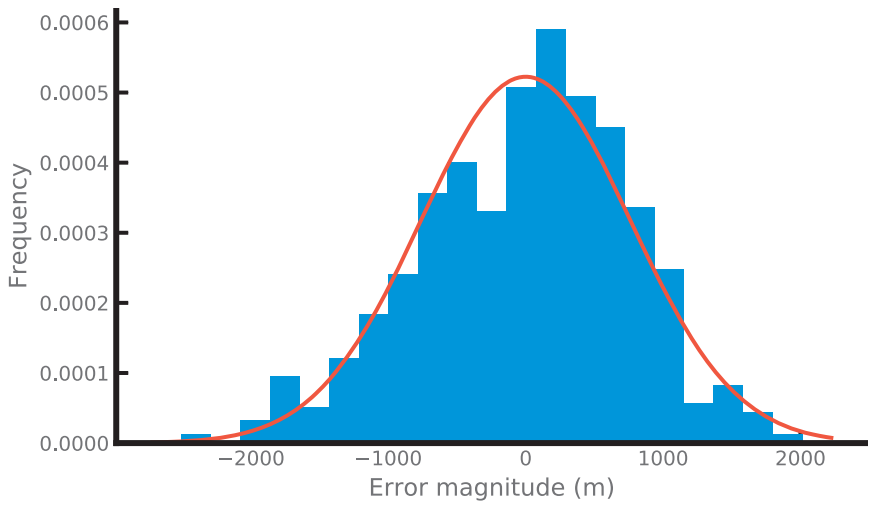

Figure 4.2: ToF distance estimate error magnitude frequencies. Overlaid with a $\mathcal{N}(\mu = 3.82^{-08}, \sigma = 767.39)$ gaussian for comparison.

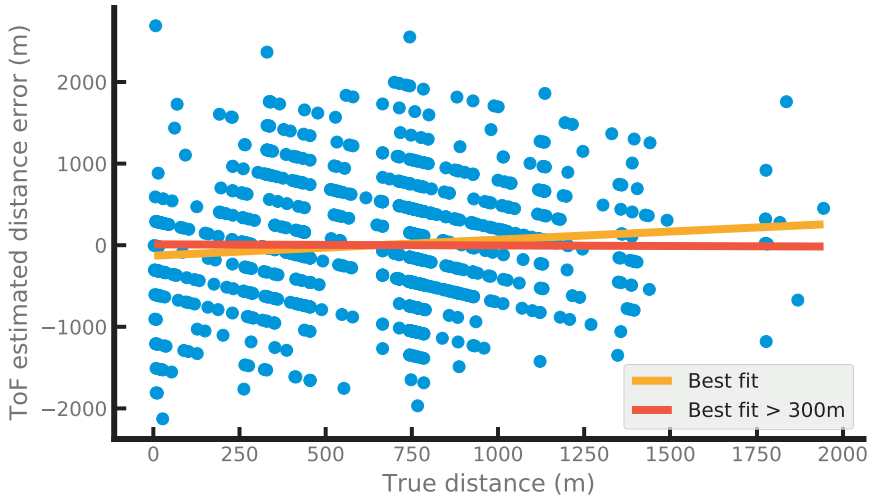

Figure 4.3: Error of ToF distance estimate as a function of true distance (as measured by GPS).

4.2.2 RSSI distance estimation

Figures 4.4 and 4.5 show RSSI measurements in rural and suburban environments respectively. One notable difference between the graphs is that the rural data shows a number of measurements in the 250-1100 meter range where a direct line of sight path existed between the node and a gateway, yielding high signal strength measurements.

Table 4.1 and 4.2 show the mean error of the distance estimations for all measured signal strengths in the respective environments. The hata model outperformed the two log models approximately by a factor of 3 in the suburban area, and by a factor of 2 in the rural area. As can be seen in figure 4.6, the average error for the hata model at most distances was less than that of either log model. Estimation errors were averaged in bins of 50m with respect to the true distance. The first datapoint represents the average error between 0-50 meters, the second 50-100 meters, and so on.

Table 4.1: Mean per-measurement error for distance estimations using each RSSI path loss model in a suburban area

Model	Error
Hata	326.08
Log exp=2.04	943.88
Log exp=2.55	961.15

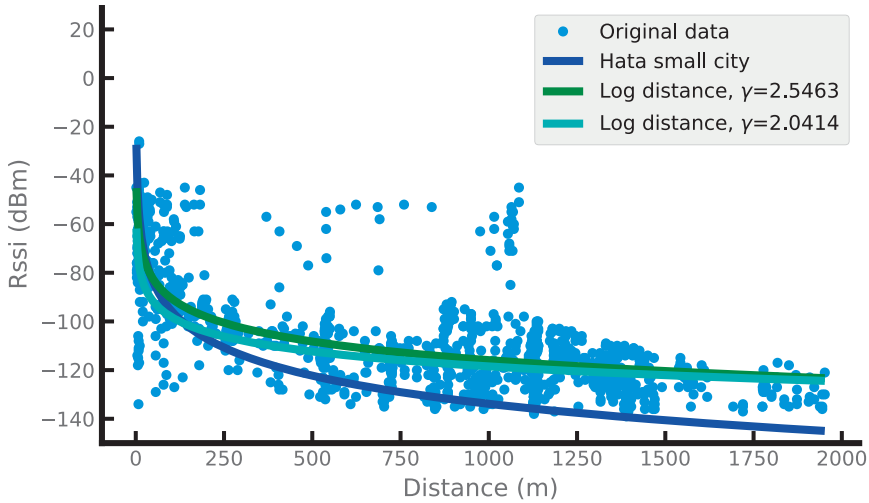

Figure 4.4: The three different path loss models and their approximation of the signal strength with respect to distance. The original data is measured in a rural environment.

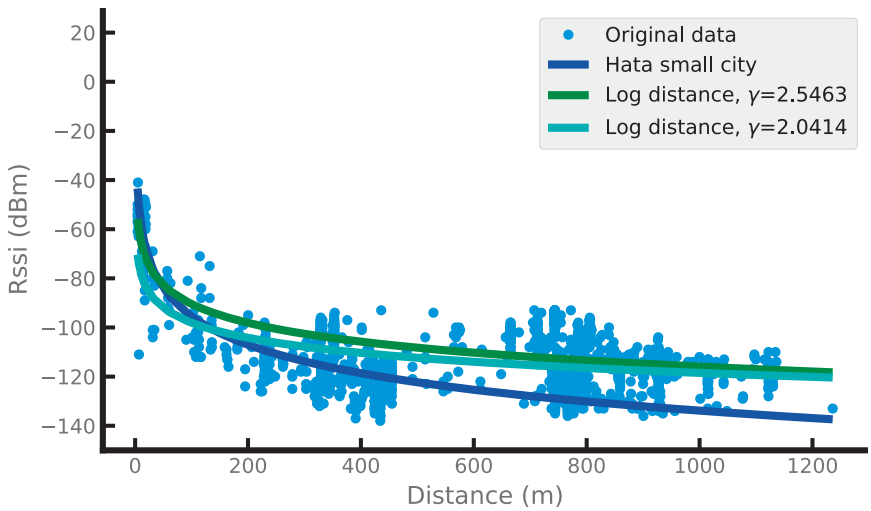

Figure 4.5: The three different path loss models and their approximation of the signal strength with respect to distance. The original data is measured in a suburban environment. Note that the range of the x-axis differs from that of figure 4.4.

Table 4.2: Mean per-measurement error for distance estimations using each RSSI path loss model in a rural area

Model	Error
Hata	475.63
Log exp=2.55	894.70
Log exp=2.04	981.61

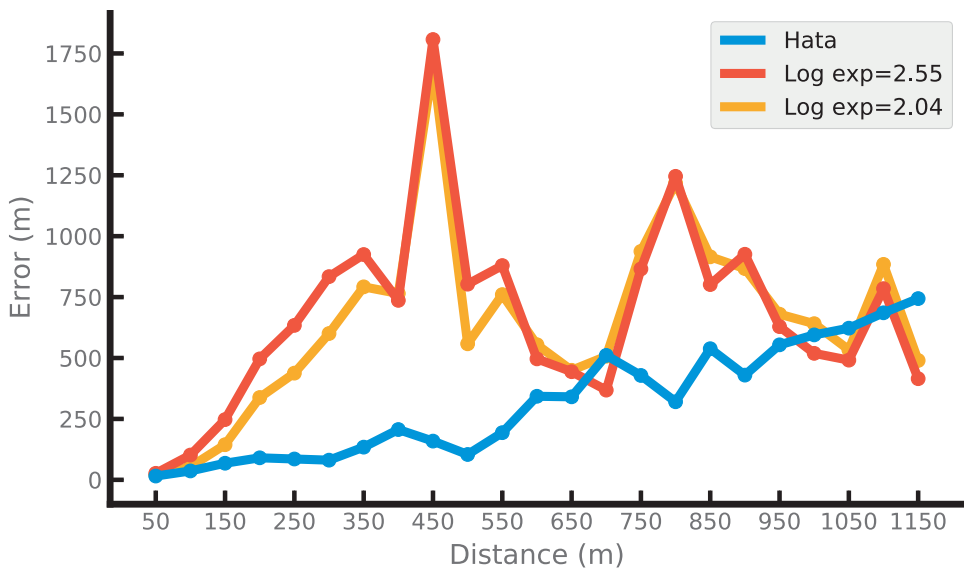

Figure 4.6: RSSI distance estimation error averaged at intervals of 50m for each of the three path loss models used.

4.2.3 Position estimation

The accuracy of the different position estimators can be seen in table 4.3, where the best performing estimator was the hybrid estimator with a mean error of 272m. Figure 4.7 shows effect on mean position error for different configurations of W_{ToF}. It can be seen that for the measured data, the optimal W_{ToF} is approximately 0.4.

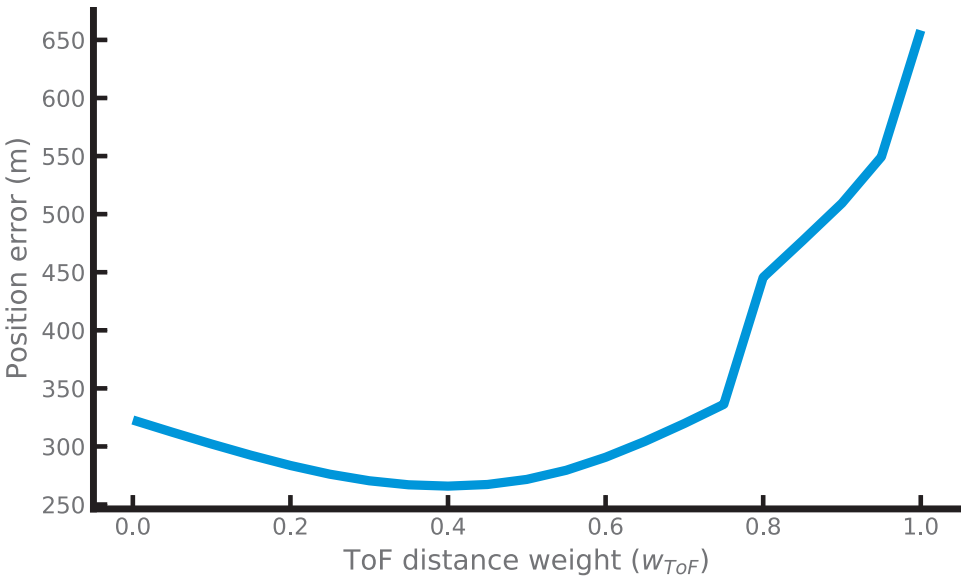

Figure 4.7: Position error as a function of the ToF weighting used in multilateration for the hybrid system.

Table 4.3: Mean static error for position estimations using each of the proposed multilateration estimators.

Position estimator	RSSI model	Error
HybridMultilateration	Hata	271.64
RSSIMultilateration	Hata	322.77
ToFMultilateration	None	658.24
HybridMultilateration	Log exp=2.04	715.71
HybridMultilateration	Log exp=2.55	787.45
RSSIMultilateration	Log exp=2.04	1037.19
RSSIMultilateration	Log exp=2.55	1203.67

4.3 Discussion

This section discusses the performance of the system, its limitations, optional configurations and effects that follow from these.

4.3.1 Localization accuracy

The results of the position estimation show that the best performing estimator is HybridMultilateration using the Hata RSSI path loss model, with its mean error of 272 meters, followed by pure Hata RSSIMultilateration with a mean error of 323 meters. Since the node remained stationary for an average of 37 protocol iterations per location, this error is best understood as a stationary error.

In comparison to previous work, the error is larger than the 100 meter median stationary error achieved by Fargas et. al. [1], but uses fewer than their 168-1728 measurements per location. Since the error arising from the ToF measurements is roughly normally distributed with an expected value close to zero the position error can be expected to decrease further as more ToF measurements are made. Fargas et. al. also employ one additional gateway over our three. To what extent this influences the accuracy of the position estimate is unclear, but when considering Podovijn et al., who use a considerably larger number of gateways (albeit spread over a larger area) and achieve a 200m median error with only a single measurement, it is likely that TDoA accuracy scales favorably with added gateways over the minimum 3. In addition, Lam et al. show that when RSSI measurements are noisy, accuracy can be improved considerably by excluding noisy gateways. These results indicate that our error could be reduced by increasing the number of gateways, but this needs to be investigated further.

All of the positioning methods used in this study are based on first estimating distances to all gateways and then estimating a position based on those distances. However, both the ToF and RSSI techniques can be used in combination with fingerprinting to directly estimate a position. Aernouts et al. [18] achieved a mean per-measurement error of 398m using fingerprint positioning of RSSI measurements. This per measurement error is comparable to our stationary error, but requires the area being tested to undergo a mapping phase where measurements with known locations are taken throughout the area. In contrast, our method can in principle be applied to any area, although for optimal results the area should be classified as rural, suburban or urban so that a suitable RSSI model can be selected.

One advantage of our system is that it does not rely on expensive gateway hardware. Some of the previous work mentioned in this study relies on gateways capable of nanosecond precision time stamps and high precision time synchronization. It is unclear how using cheaper hardware might have affected the results of these studies. Time synchronization can be achieved using GPS modules (as in Fargas), but this requires gateways both to possess a GPS module and be placed in locations where the GPS signal can be locked on to, preventing some indoor placements.

One notable disadvantage of ToF as a ranging method is that ranging is done at one gateway at a time, rather than to all gateways simultaneously as in TDoA.

Coupled with the limitations on transmission frequency in higher spreading factors, this can be prohibitive. This drawback means that ToF positioning could potentially scale poorly with the number of reachable gateways, which depending of the application may be more or less impactful. However, since RSSI measurements are made by all gateways in range of a node's transmission it may be sufficient to do a fixed number of ToF iterations, relying instead on the larger amount of RSSI measurements to provide accuracy scaling as the number of reachable gateways increase.

Looking at the relationship between RSSI distance estimation errors and the true distance shown in figure 4.6, it is clear that relying on RSSI path loss models for accurate ranging is difficult. Even though the Hata model yielded a mean per-measurement distance error of only 326m (see table 4.1), it loses accuracy as distances increase. This is in contrast to the behavior of ToF shown in figure 4.3, which has a large per-measurement error, but where the error stable at longer distances. This difference in behaviour allows for the two methods to complement each other.

With the LoRa spreading factor of 12 used during measurements, transmission times were 827ms for the Ranging Signal Message and Ranging Signal Response Message. The transmission time for the Timing Message was 992ms. With the 1% duty cycle limitation, this limits the ranging protocol to a frequency of at most once every 100 seconds. By reducing the spreading factor, the ranging protocol could be run at a faster rate and using less energy (up to a factor of 32 using SF7) but this would also limit the range of the transmission and consequently the number of gateways reached. Podevijn et al. [17] reason that an optimal spreading factor exists and can be calculated as a function of the probability that a transmission using a given spreading factor will be received by three or more gateways. In short, the more densely the gateways are placed, the lower the optimal spreading factor.

The estimated energy usage for a ranging protocol run is 82.5mAs, see section 2.5) with spreading factor 12, permitting a total of 12 protocol runs before consuming more than that of the reference GPS. With the accelerometer sampling at a frequency of 50Hz it consumes an extra 0.11mA, having little overall impact on the total power consumption. Using an average of 37 protocol iterations per location, the average energy usage per position estimation was 3011mAs. This is roughly three times as much as that of a cold started GPS, for an average accuracy that is roughly 50 times less accurate. By reducing the spreading factor to 7, the break even point with GPS is instead at 384 protocol iterations. Since the ToF errors are normally distributed, the added samples are likely to greatly increase the overall accuracy. Thus, if power consumption is the primary concern and high levels of accuracy are desired, multiple protocol iterations with a high spreading factor are not an effective solution, but a lower spreading factor might be sufficient.

Additionally, unlike GPS the proposed system requires only a LoRa radio and an accelerometer to run the ranging protocol. The system might therefore still be a viable alternative for low cost devices where the price of a GPS module is too high.

There exists a multivariate trade off between maximum reach, spreading factor, power consumption, accuracy and transmission time. The local environment also puts limitations on the maximum reachable distance, which might make the use

of an RSSI path loss model alone a better choice if power consumption is a critical variable.

It should be noted that all of the position estimations were conducted using measurements from a single suburban area, and it is unclear how representative these measurements are. The size of the dataset is also limited, having only measured at four distinct locations. Further testing is required to fully understand the true accuracy of the positioning system.

4.3.2 Possible Applications

For the position estimation system to function effectively in any given scenario, nodes must be able to consistently reach three or more gateways. Otherwise, the system is not able to triangulate the node's position. With this in mind, the most promising applications are likely to be in areas such as cities or suburbs. City applications allow for the possibility of using gateways which you do not directly control (assuming they implement the gateway side of the ranging protocol), reducing the cost of the application. The system could also be feasible for use in rural areas provided the user ensures there are enough gateways.

Suitable applications are likely to include cases where limitations of power and cost are more important than high levels of accuracy. One such application is that of wildlife tracking, where battery replacement of a module mounted on an animal might be an invasive action. An alternative that limits the total number of battery replacements could prove a better alternative even if the accuracy is lacking. An obvious downside for this application is that it requires multiple gateways to be spread out in the nearby area, which GPS does not. Placing these gateways in the surrounding landscape might itself be intrusive.

More generally, many forms of asset tracking such as locating cargo containers in large container ports may not require high accuracy, but can take advantage of the reduced hardware cost. Applications where object positions are being tracked centrally also require the position to be communicated to the central server somehow. If GPS is used for positioning, some form of radio technology like LoRa is likely required to transmit the coordinates anyway. In contrast, LoRa-based localization techniques solve both the localization and communication problems using the same hardware.

Chapter 5
Conclusion

The goal of this thesis was to explore the viability of a hybrid ToF and RSSI localization system using low-cost LoRa devices. The evaluation of the localization system demonstrates that the hybrid localization system performs better than a pure version of either, achieving a mean stationary position error of 272 meters after an average of 37 protocol iterations per location.

The empirical results indicate that distance estimation errors for ToF are normally distributed with an expected value close to zero, and that while a 1 μs clock resolution limits the accuracy of single measurements, averaging multiple ToF measurements provides a reliable distance estimation. In addition, ToF error does not increase with distance. The evaluation also shows that RSSI distance estimation is dependent on the environment and the path loss model used. The best performing path loss model overall was the small city Hata model.

Depending on the spreading factor there is a break even point for the number of ranging protocol iterations, where the system's energy consumption exceeds that of a single GPS lock. At spreading factor 12, this break even point is 12 iterations, and for spreading factor 7 it is 384.

It is clear that there is a trade off between location accuracy, price, and power consumption. Since the core purpose of this thesis is partly to explore a hybrid solution as a means to provide localization in a constrained context, efforts were made using a minimal viable set of hardware to combine techniques and have them augment each other. The result is a system that does not require any top down control such as time synchronization, works on gateways and nodes with limited processing power and clock resolution, and requires only a LoRa radio and accelerometer (for nodes).

The implementation presented in this study could be further improved upon in different ways depending on which aspect is prioritized. If accuracy is the prime objective, adding TDoA measurements to the hybrid system could be effective, in exchange for requiring time synchronization between gateways. If added accuracy may not come at the cost of an increased prize or power consumption, environment mapping opens up the possibility for accurate RSSI path loss models as well as the use of fingerprinting for accurate positioning.

Improvements could also be made by applying more complex techniques to combining the results of the different ranging methods. Machine learning techniques could potentially make use of more complex patterns in the measurement data, allowing for increased accuracy. Filtering out noisy gateways, e.g. using k

means clustering (see [19]), may also lead to better results.

Finally, implementing the system on top of the LoRaWAN stack would allow for it to be used in public LoRaWAN networks such as The Things Network. With its focus on low cost devices and decentralization the system could be a means to providing IoT solutions with an easy answer to localization problems.

Appendices

Appendix A

Firmware Modifications

```
IRAM_ATTR void SX1276OnDio0Irq( void )
{
            /* other code ... */
            TimerStop( &TxTimeoutTimer );
            // Transmit Done interrupt
            switch( SX1276.Settings.Modem )
            {
            case MODEM_LORA:
                // Hardware timer timestamp call
                SX1276.Settings.LoRaPacketHandler.TimeStamp
                    = mp_hal_ticks_us_non_blocking();
                // Clear Irq
                SX1276Write( REG_LR_IRQFLAGS,
                    RFLR_IRQFLAGS_TXDONE );
            /* other code ... */
    }
```

Listing A.1: Timestamp after a successful packet transmission using a call to the hardware timer.

Appendix B

Packet structure

Every packet sent included a one byte header (see figure B.1) field, a two byte session counter field and an accompanying four byte CRC field used to validate packet correctness.

A total of three packets were sent during each iteration of the protocol. The two type of signal messages (figure B.2) used, Ranging Signal Message and Ranging Signal Response Message, were identical except for the message type field of the header.

At the end of each protocol run the Timing Message (figure B.3) sent to the gateway included the same data as the signal messages (except for a different message type) as well as the node's processing time and the rate of movement detected by the motion detector.

Figure B.1: Header Structure

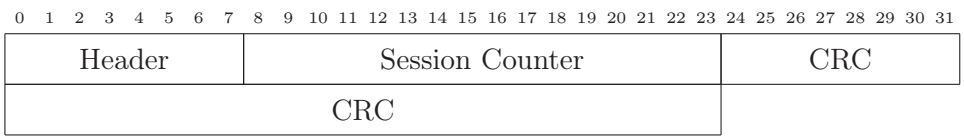

Figure B.2: Signal Message Packet Structure

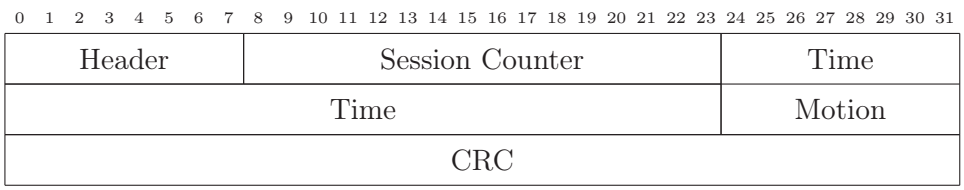

Figure B.3: Timing Message Packet Structure

Appendix C

Contributions

Overall, the work done over the course of this thesis has been evenly distributed between the authors. During implementation of the hybrid positioning system, most of the work was done through pair programming. Similarly, the field work (empirical tests, gathering data, calibrations) was done together. During the writing of this paper no explicit areas of responsibility were assigned, but the first drafts of each section were distributed as follows: Valthor wrote the introduction, methodology, implementation, and discussion. Jonas wrote the theory, results, abstract, and designed all of the figures and illustrations. The remaining parts were written together.

Bibliography

[1] B. C. Fargas and M. N. Petersen. "GPS-free geolocation using LoRa in low-power WANs". In: *2017 Global Internet of Things Summit (GIoTS)*. June 2017, pp. 1–6. DOI: 10.1109/GIOTS.2017.8016251.

[2] A. Zhang, Y. Yuan, Q. Wu, S. Zhu, and J. Deng. "Wireless Localization Based on RSSI Fingerprint Feature Vector". In: *International Journal of Distributed Sensor Networks* 11.11 (2015), p. 528747. DOI: 10.1155/2015/528747.

[3] J. Vander Stoep. "Design and implementation of reliable localization algorithms using received signal strength". PhD thesis. Citeseer, 2009.

[4] O. Oguejiofor, V. Okorogu, A. Adewale, and B. Osuesu. "Outdoor localization system using RSSI measurement of wireless sensor network". In: *International Journal of Innovative Technology and Exploring Engineering* 2.2 (2013), pp. 1–6.

[5] W.-H. Kuo, Y.-S. Chen, G.-T. Jen, and T.-W. Lu. "An intelligent positioning approach: RSSI-based indoor and outdoor localization scheme in Zigbee networks". In: *Machine Learning and Cybernetics (ICMLC), 2010 International Conference on*. Vol. 6. IEEE. 2010, pp. 2754–2759.

[6] M. L. Sichitiu, V. Ramadurai, and P. Peddabachagari. "Simple Algorithm for Outdoor Localization of Wireless Sensor Networks with Inaccurate Range Measurements." In: *International Conference on Wireless Networks*. Vol. 2003. 2003.

[7] J. Graefenstein and M. E. Bouzouraa. "Robust method for outdoor localization of a mobile robot using received signal strength in low power wireless networks". In: *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*. IEEE. 2008, pp. 33–38.

[8] C. Alippi and G. Vanini. "A RSSI-based and calibrated centralized localization technique for Wireless Sensor Networks". In: *Pervasive Computing and Communications Workshops, 2006. PerCom Workshops 2006. Fourth Annual IEEE International Conference on*. IEEE. 2006, 5–pp.

[9] B. Thorbjornsen, N. M. White, A. D. Brown, and J. S. Reeve. "Radio frequency (RF) time-of-flight ranging for wireless sensor networks". In: *Measurement Science and Technology* 21.3, 035202 (Mar. 2010), p. 035202. DOI: 10.1088/0957-0233/21/3/035202.

[10] S. Lanzisera, D. T. Lin, and K. S. Pister. "RF time of flight ranging for wireless sensor network localization". In: *Intelligent Solutions in Embedded Systems, 2006 International Workshop on*. IEEE. 2006, pp. 1–12.

[11] S. Lanzisera, D. Zats, and K. S. Pister. "Radio frequency time-of-flight distance measurement for low-cost wireless sensor localization". In: *IEEE Sensors Journal* 11.3 (2011), pp. 837–845.

[12] A. V. Nieuwenhuyse, J. Wyffels, J.-P. Goemaere, L. D. Strycker, and B. Nauwelaers. "Time of Arrival Based on Chirp Pulses as a means to Perform Localization in IEEE 802.15.4a Wireless Sensor Networks". In: *Advances in Electrical and Computer Engineering* 10.2 (2010), pp. 65–70. DOI: 10.4316/aece.2010.02011.

[13] L. Yang and K. Ho. "An approximately efficient TDOA localization algorithm in closed-form for locating multiple disjoint sources with erroneous sensor positions". In: *IEEE Transactions on Signal Processing* 57.12 (2009), pp. 4598–4615.

[14] H. Jamali-Rad and G. Leus. "Sparsity-aware multi-source TDOA localization". In: *IEEE Transactions on Signal Processing* 61.19 (2013), pp. 4874–4887.

[15] M. Laaraiedh, S. Avrillon, and B. Uguen. "Hybrid data fusion techniques for localization in UWB networks". In: *Positioning, Navigation and Communication, 2009. WPNC 2009. 6th Workshop on*. IEEE. 2009, pp. 51–57.

[16] M. Laaraiedh, L. Yu, S. Avrillon, and B. Uguen. "Comparison of hybrid localization schemes using RSSI, TOA, and TDOA". In: *Wireless Conference 2011-Sustainable Wireless Technologies (European Wireless), 11th European*. VDE. 2011, pp. 1–5.

[17] N. Podevijn, D. Plets, J. Trogh, L. Martens, P. Suanet, K. Hendrikse, and W. Joseph. "TDoA-Based Outdoor Positioning with Tracking Algorithm in a Public LoRa Network". In: *Wireless Communications and Mobile Computing* 2018 (2018).

[18] M. Aernouts, R. Berkvens, K. Van Vlaenderen, and M. Weyn. "Sigfox and LoRaWAN Datasets for Fingerprint Localization in Large Urban and Rural Areas". In: *Data* 3.2 (2018), p. 13.

[19] K. H. Lam, C. C. Cheung, and W. C. Lee. "LoRa-based localization systems for noisy outdoor environment". In: *2017 IEEE 13th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. Oct. 2017, pp. 278–284. DOI: 10.1109/WiMOB.2017.8115843.

[20] Semtech Corporation. *What is LoRa?* https://www.semtech.com/technology/lora/what-is-lora. Accessed: 2018-02-23.

[21] European Telecommunications Standards Institute. *Electromagnetic compatibility and Radio spectrum Matters (ERM); Short Range Devices (SRD)*. Standard 2.4.1. Valbonne, FR, Jan. 2012.

[22] Semtech Corporation. *LoRa Modem Design Guide*. English. Version Revision 1. Semtech Corporation. 9 pp. November 16, 2011.

[23] Semtech Corporation. *Calculating Radiated Power and Field Strength for Conducted Power Measurements*. https://www.semtech.com/uploads/documents/semtech_acs_rad_pwr_field_strength.pdf. Accessed: 2018-02-23.

[24] P. K. Sharma and S. R.K. "Comparative Analysis of Propagation Path loss Models with Field Measured Data". In: 2 (June 2010).

[25] United States Government. *2008-SPS-performance-standard*. https://www.gps.gov/technical/ps/2008-SPS-performance-standard.pdf. Accessed: 2018-02-23.

[26] N. Ivanov and V. Salischev. "The GLONASS System – An Overview". In: *Journal of Navigation* 45.2 (1992), pp. 175–182. DOI: 10.1017/S0373463300010675.

[27] European Global Navigation Satellite Systems Agency. *Galileo is the European global satellite-based navigation system*. https://www.gsa.europa.eu/european-gnss/galileo/galileo-european-global-satellite-based-navigation-system. Accessed: 2018-02-23.

[28] S. von Watzdorf and F. Michahelles. "Accuracy of Positioning Data on Smartphones". In: *Proceedings of the 3rd International Workshop on Location and the Web*. LocWeb '10. Tokyo, Japan: ACM, 2010, 2:1–2:4. ISBN: 978-1-4503-0412-2. DOI: 10.1145/1899662.1899664. URL: http://doi.acm.org/10.1145/1899662.1899664.

[29] N. Sirola. "Closed-form algorithms in mobile positioning: Myths and misconceptions". In: *2010 7th Workshop on Positioning, Navigation and Communication* (2010), pp. 38–44.

[30] T. Sauer. *Numerical Analysis*. 2nd. USA: Addison-Wesley Publishing Company, 2011. ISBN: 0321783670, 9780321783677.

[31] STMicroelectronics. *LIS2HH12: MEMS digital output motion sensor ultra-low-power high-performance 3-axis "pico" accelerometer*. English. Version Revision 1. AN4662. STMicroelectronics. 43 pp.

[32] Libelium Comunicaciones Distribuidas S.L. *Waspmote LoRaWAN Networking Guide*. Accessed: 2018-06-04.

[33] Quectel Wireless Solutions. *Extremely Compact GNSS Module with Ultra Low Power Consumption*. Accessed: 2018-06-04.

[34] Pycom Ltd. *LoPy4*. Accessed: 2018-06-04. 2018. URL: https://pycom.io/product/lopy4/.

[35] Pycom Ltd. *LoPy4*. Accessed: 2018-06-04. 2018. URL: https://docs.pycom.io/chapter/datasheets/boards/pytrack.html.

[36] H. Linka, M. Rademacher, O. G. Aliu, and K. Jonas. "Path Loss Models for Low-Power Wide-Area Networks: Experimental Results using LoRa". In: *VDE ITG-Fachbericht Mobilkommunikation*. May 2018.

Printed by Tryckeriet i E-huset, Lund 2018

LUND
UNIVERSITY

Series of Master's theses
Department of Electrical and Information Technology
LU/LTH-EIT 2018-670
http://www.eit.lth.se