Location and Identification of People in the Home Environment

Tomás da Silva Marques

Thesis to obtain the Master of Science Degree in

Information Systems and Computer Engineering

Supervisor: Prof. Renato Jorge Caleira Nunes

Examination Committee

Chairperson: Prof. João António Madeiras Pereira
Supervisor: Prof. Renato Jorge Caleira Nunes
Member of the Committee: Prof. Alberto Manuel Ramos da Cunha

September 2020

Agradecimentos

Em primeiro lugar, quero agradecer a toda a minha família, especialmente à minha Mãe e à minha Avó, por me terem dado esta oportunidade. Por todo o apoio, suporte e paciência nestes últimos 8 anos, e por todo amor que me deram nestes primeiros 26. Obrigado.

Quero agradecer ao meu orientador, Renato Nunes, pela sua ajuda valiosa neste trabalho. Obrigado.

Quero também agradecer ao Pedro, à Nena, ao Janeca, à Joana ao Pedro, e a todos os outros amigos da terrinha que trouxe comigo para Lisboa. Ao Manel, ao Antunes, ao Fábio, ao Xu, ao Loureiro, ao Lord, ao Alcino, ao Palhas, ao Titó, à Nádia, ao Tínoní, ao Leonoldo, a toda a CPLEIC, às amigas de Medicina, e à Bukkatuna. Por todas as missões, aventuras e momentos inesquecíveis que partilhamos juntos. Pela vossa amizade e companhia. Obrigado.

Finalmente, quero agradecer à Inês. Por todo o apoio e motivação que me dás. Pelas aventuras que temos juntos e por me amares, como eu te amo a ti. Obrigado.

Abstract

The growing interest and development of smart home technology with location aware applications has lead to a surge in the developing of Indoor Positioning Systems. Existing commercial Positioning Systems, however are quite expensive and complex. In this work we introduce and survey the different existing technologies and techniques used for indoor positioning. Furthermore, we present and detail the development of own, very low cost indoor positioning system, using a radio network of fixed and mobile nodes, with room level precision. Finally, we describe the tests performed to validate the system and evaluation metrics.

Keywords

Indoor Positioning System; Context Aware Applications; Sensor Network; Assisted Living.

Resumo

O crescente interesse e desenvolvimento em tecnologias para casas inteligentes e aplicações cientes de contexto levou a uma onda de criação de Sistemas de Localização Interior. No entanto, as soluções comerciais existentes são caras e complexas. Neste trabalho pesquisamos diferentes técnicas de localização e tecnologias existentes de localização interior. Em seguida apresentamos em detalhe o desemvolvimento do nosso Sistema de Localização Interior de muito baixo custo, consistindo numa rede de rádio com nós fixos e nós móveis, com precisão ao nível da divisão. Finalmente, descrevemos os testes usados para validar o sistema e as métricas de avaliação.

Palavras Chave

Sistema de Localização Interior; Aplicações Cientes de Contexto; Redes de Sensores; Vida Assistida.

Contents

List of Figures

List of Tables

Acronyms

AGPS Assisted Global Positioning System

AOA Angle of Arrival

BLE Bluetooth Low Energy

CSI Channel State Information

GNSS Global Navigation Satellite System

GPS Global Positioning System

IPS Indoor Positioning System

KNN K Nearest Neighbours

RF Radio Frequency

RFID Radio Frequency Identification

RSS Received Signal Strength

RSSI Received Signal Strength Indicator

RTOF Roundtrip Time of Flight

SPI Serial Peripheral Interface

TDOA Time Difference of Arrival

TOA Time of Arrival

TOF Time of Flight

UWB Ultra Wideband

WLAN Wireless Local Area Network

WPS WiFi Positioning System

1

Introduction

Contents

1.1 Context and Motivation

With the development of home technology, the location and tracking of people is an increasingly useful information for a number of context aware applications in smart buildings, ranging from safety, security, resource efficiency, visitor navigation, health care and asset tracking. [1]

An Indoor Positioning System (IPS), is a system designed for determining the position of objects or people in an indoor physical environment. When used in real time, an IPSs could provide location information of users, for assistance in emergencies. Smart homes could make use of user location information provided by IPSs for optimizing resource efficiency by directing resources to target locations. They could also be used to detect anomalous behaviour in users with risk conditions.

Most commercially available IPSs, are designed for use in the industry, health-care and commerce. For asset tracking in factories and warehouses, for patient and equipment tracking in hospitals and clinics, and for client navigation and browsing behaviour analysis in shops and malls.

There is currently no standard for Indoor Positioning Systems. They differ on several characteristics, based on types of information provided, system topology, technologies used and location estimation methods implemented.

Active and Passive Systems. Positioning systems can be classified into two categories: Active or passive systems. [2] Active systems require the user to actively take part in the process of positioning (e.g., by carrying a fixed device). Oppositely, passive positioning, functions completely without the need for the users to interact with the system.

Despite the apparent inconvenience to the users, active systems actually have the advantage of solving identification trivially, by having the interaction between user and system include identity information. On the other hand, passive systems either don't care for user identification or use another method for deriving this information. Passive identification systems also exist [3], so a combined system could be designed to function entirely without user intervention. However, this might require extra infrastructure, which in turn means a higher system cost, or, if it can be designed to use the same infrastructure it will lead to a higher computational complexity, which in turn might affect the capacity of the system to operate in real time.

Different location-based-services (LBS) require different types of location information. In [4] the authors categorize location information in regards to two proprieties.

Physical vs Symbolic Location. Physical Position identifies a point on a map (2D, 3D). Symbolic Location identifies a location in a human language way, (e.g., ”in the basement”, ”next to the window”). One type can be derived from the other depending on the system resolution, for example, a system with an accuracy of 3 meters can translate the physical position determined to the room if there are no walls within that range.

Absolute vs Relative Location. Absolute location uses a shared reference grid for all located objects. Different systems report the same absolute location for the same located target. Relative location depends on the reference frame. Each device reports a location relative to a reference point, usually the devices own position. If we know the absolute location of the reference points, then we can transform relative location to absolute location.

Finally, another way IPSs can be categorized is regarding the system topology. There are three different classifications for active positioning systems, in which users carry a fixed device, regarding their topology. [5] That is, where measurements are collected and where the algorithm calculations are made.

Self-Positioning. In self-positioning systems, the mobile unit measures the signals received from the fixed transmitters in known locations, and computes its own location. In a system like this, user side applications can use this computed position information to make location aware decisions. Also, this is the ideal topology for user navigation, since we want to display the location information to the user.

Remote Positioning. In remote positioning systems, the mobile unit transmits a signal that is received and measured by the fixed devices with known locations. The measurements are then combined in a "central site". This topology might be better suited for a smart home environment, where a potentially large number of context aware applications could co-exist.

Indirect Positioning. A system where the mobile unit sends the result to the remote side is an indirect positioning system, specifically an indirect remote positioning system. Conversely, a system in which the remote side computes the location of the user and sends the result to the mobile unit, is an indirect self-positioning system.

Multipath Effect. One of the biggest difficulties positioning systems based on radio signals run into is the multipath effect. Multipath effect consists of refracted and reflected radio signals arriving at the receiver by distinct paths. This causes different readings and measurements obtained from the same source which results in poor positioning accuracy.

1.2 Objectives

Several commercially available indoor positioning systems have been designed for a large number of different contexts, like industry, healthcare and commerce. In spite of this, we found none that was as affordable for use at home as we recognize they could be.

The objective of this work is to design an open, low cost, Indoor Positioning System, that can locate a user in real time. At least room-level accuracy is required, a finer granularity may be achievable but is not the focus of this work. Designing for low-cost takes precedence over a high accuracy. In this sense, a balance between cost and accuracy is our goal.

The proposed system should provide location to other applications or notify them when users enter areas or locations defined by them. Additionally, the system should be able to be used for analyzing and detecting anomalous or unusual behaviour of users that have a need to be under constant or frequent watch, such as, elderly people. For instance, detecting that a user has not moved in a long time, or that the stove as been left turned on, with no one nearby for a while.

Our system should also have some degree of scalability, that is, it should function well for small or big homes. It could also be interesting for the system to be installed in a nursing home or small hospital.

Finally, the system should be relatively easy to install.

1.3 Organization of the Document

In this work we first start by surveying technologies and techniques used for indoor positioning, as well as other related existing systems in chapter 2. In chapter 3 we produce a general description of our solution, following the objectives set above and taking into account the survey presented. In chapter 4 we give all details pertaining to the implementation of the solution described. In chapter 5 we describe the validation tests and present the relevant metrics and the observations. Finally in chapter 6 we leave notes on possible future work, and give some concluding remarks.

2

Related Work

Contents

Here we present some positioning and location sensing algorithms, techniques and technologies, that are relevant for designing a solution to our problem, as well as some previous works on the subject of Indoor Positioning and available solutions.

2.1 Positioning Algorithms and Techniques

In this section we introduce the main position estimation techniques and algorithms. There are 3 main categories of position estimation techniques, Triangulation, Scene Analysis and Proximity based techniques. [6]

2.1.1 Triangulation

The Triangulation method is based on using the geometric properties of triangles to estimate the subject location. Triangulation can further be categorized in two sub-categories, Lateration and Angulation.

2.1.1.A Lateration

Lateration computes the target position by measuring the distance from the tag location to the fixed receivers, anchors (Fig. 2.1). The distance measurement can be determined indirectly through the Time of Flight (TOF) or attenuation of signals. Three different TOF measurements can be used to compute distance: Time of Arrival (TOA), Time Difference of Arrival (TDOA) and Roundtrip Time of Flight (RTOF).

In TOA, the target mobile tag transmits a signal to the fixed devices. At arrival, the fixed devices calculate the time of flight as the difference between the time of arrival and the time of departure and calculate the distance to the target using the time of flight and the signal propagation speed. As such, at least three fixed devices are needed to compute the position and all devices need to be synchronized with a time source.

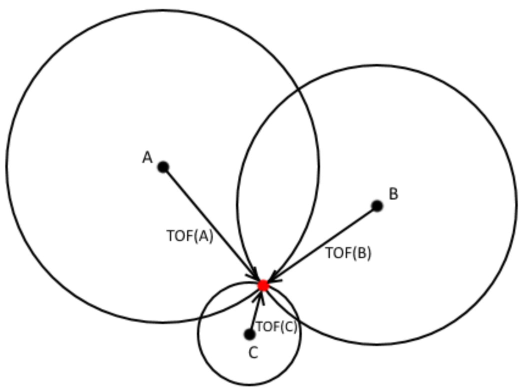

Figure 2.1: Time of Flight based trilateration. TOF can be determined through TOA or RTOF

In TDOA, the fixed devices, with well known positions, receive a transmission from the mobile unit with an unknown starting time, and use the differences between the times of arrival to compute the mobile units location (Fig. 2.2). With this method only the receivers need to be synchronized.

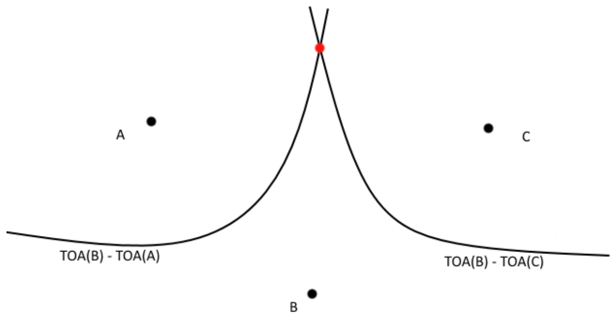

Figure 2.2: TDOA based trilateration.

RTOF is the time of flight from the transmitter to the receiver and back. This way, only the time at the transmitter is necessary, which solves the problem of time synchronization between devices. Despite this, ranging to several fixed devices at the same time may cause latency that will result in inaccurate time measurements.

The intensity of a signal decreases with distance. Computing distance with signal attenuation usually makes use of the Received Signal Strength Indicator (RSSI). Refraction and reflection of signal, as well as the multipath effect, affect the attenuation of signals non linearly.

2.1.1.B Angulation

Instead of distance between sender and receiver, Angulation uses Angle of Arrival (AOA) measurements to determine the position of the target. At least the angles at two receivers and the distance between them are necessary to compute the target position (Fig. 2.3).

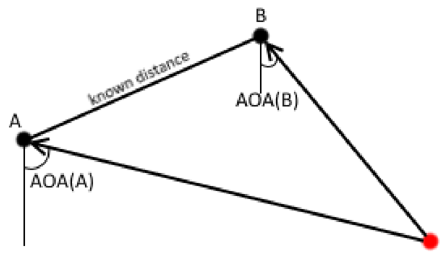

Figure 2.3: AOA based angulation. Only two receivers and the distance between them is required.

2.1.2 Scene Analysis

The Scene Analysis method consists of using observed or measured characteristics to infer the location. In static scene analysis observed characteristics are looked up in a predefined dataset, that maps them to the target location. This is also called Location Fingerprinting. Scene analysis can be made using images, captured by a camera, or using any other measurable propriety, like RSSI. Received Signal Strength (RSS) based fingerprinting is commonly used in scene analysis. [7]

Location Fingerprinting consists of two phases: the offline phase, and the runtime phase. In the offline phase, a location is surveyed, the coordinates and respective RSS values from fixed measuring units are collected to build the dataset. During the runtime phase, an algorithm matches the currently measured RSSI to a location using the dataset as a dictionary. Several algorithms or approaches can be used for the pattern matching phase. (e.g., probabilistic approaches, K Nearest Neighbours (KNN), neural networks)

2.1.3 Proximity

Providing relative symbolic location, Proximity location sensing methods usually depend on a grid of sensors, with shorter range than those used in other location sensing methods, with well known positions. The range of these sensors needs to be adapted to the corresponding location, so that when a target is within its range, it can be considered to be in that same location.

This technique can be implemented with physical contact detectors (e.g., pressure sensors, touch sensors), Infrared Radiation, passive or active Radio Frequency Identification (RFID), and Bluetooth. The accuracy of proximity location sensing methods depends mostly on the density of sensors deployed.

2.2 Relevant Technologies and Solutions

Next, we present different technologies that we considered to solve our problem, and some proposed or commercially available systems that use those technologies.

The different technologies are based on analyzing different kinds of signals and measurements. Here we talk about technologies using radio frequency, sound and light signals, as well as inertia measurements.

2.2.1 Global Positioning System

Global Positioning System (GPS), or Global Navigation Satellite System (GNSS), has long since been the technology most used for outdoor positioning. However, due to satellite signal attenuation caused by walls and buildings it becomes unsuitable for indoor localization. [8]

Besides the expected increase in receiver sensitivity, some effort as been put in trying to make GPS a viable solution for indoor positioning.

2.2.1.A Assisted GPS

Assisted Global Positioning System (AGPS) uses a mobile phone data link to download satellite information that would usually be received from the GNSS satellites directly, making it easier for the AGPS receiver to lock on to available satellites. As such AGPS needs a mobile phone with internet connection, which may not be possible in some scenarios. Also, even assisted by Differential-GPS techniques, AGPS may not be able to acquire the signal for the required 4 different satellites.

2.2.1.B Pseudolites

Pseudo satellites, or pseudolites, are ground based transceivers that generate a GPS-like signal from a known position. This signal is a bit different from normal GPS signals. Because of this, a slight modification to the receivers software is necessary [9]. In [10] the authors design a pseudolites-based positioning system that doesn't require any modification to standard GPS receivers.

In conclusion, GPS is not suitable for most indoor environments unless mobile phones are used as receivers and an internet connection is always present, or without an expensive [11] pseudolite system.

2.2.2 RFID

Radio Frequency Identification, or RFID, is a way to retrieve information through electromagnetic transmission to a Radio Frequency (RF) compatible circuit. A RFID system is made up of readers, tags and the communication used among them. [12] RFID tags can be either passive or active.

Passive RFID

Passive tags reflect and modulate the signal transmitted to them using no battery, therefore making them much cheaper. However, their read ranges are very short. Also, Passive RFID readers can have a relatively high cost. [7]

Active RFID

Active tags have a built-in battery and radio to broadcast their identity periodically or in reply to an interrogation. Because of this, active tags have a much longer range than passive tags.

SpotOn

SpotOn [13] is a positioning system based on active RFID tags designed by the researchers. Fixed readers then take RSSI from nearby tags and a central server aggregates the readings and triangulates the position of the tags. Each tag was estimated to have a 30-40$ cost after revision and manufacturing for quantity.

LANDMARC

LANDMARC [12], also uses active RFID tags and fixed readers to compute the tag's position based on RSSI analysis. Additionally, they use extra tags with fixed positions that server as reference points to help with location calibration. LANDMARC also uses the KNN method to determine the position of the tags.

Lastly, the authors in [14] implemented a high performance positioning system using a Kalman-filter to remove the RSSI drift and Heron-bilateration to reduce the execution time of the position determining algorithm and the distance measurement error, and have reported sub-meter level accuracy.

2.2.3 Ultra Wideband

Different from other radio technologies, Ultra Wideband (UWB), concurrently transmits signals over multiple bands of frequencies. UWB works by transmitting ultrashort pulses with a low duty cycle [7]. Because of this, UWB signals have a high penetration capability, tags use less power and can be utilized near other radio signals without suffering or causing interference. Moreover, it is easy to determine which signals are correct and which resulted from multipath as well as determining an accurate TOA [15] [16].

The authors in [17] designed a positioning system based on a UWB radar built using off the shelf components and an android device. Able to achieve sub-meter level accuracy and maximum error, their system offers a cost of 120€ per host per tag.

2.2.4 Bluetooth

Bluetooth is a very ubiquitous technology, available in most smartphone devices. Bluetooth is another one of the most popular technologies used for indoor positioning, particularly through the use of Bluetooth Low Energy (BLE) beacons. BLE beacons are essentially used as tags, have a very low power consumption and are relatively low cost. The main drawbacks of using beacons tags in an IPS are the high signal attenuation and low range [18]. Bluetooth based IPSs can function using proximity or lateration techniques.

2.2.5 Wireless Local Area Network

Another popular way to build IPSs is using the Wi-Fi in a Wireless Local Area Network (WLAN). Relative to other technologies, this method has a higher energy cost and lower positioning accuracy, with a mean error of over 2 meters. However, since it is an infrastructure already present in most indoor environments, it can become a very cheap alternative. Most WiFi Positioning System (WPS) determine positions based on proximity detection via RSSI analysis or through fingerprinting techniques. [19]

In [20] the authors present a methodology for indoor localization based on fingerprinting using RSSI. They report results with an accuracy of 2m.

Recently, researchers have started using channel response, instead of RSSI, to determine location in Wi-Fi networks [21]. Channel response is the power feature of the PHY layer and it is able to discriminate multipath characteristics, allowing for more accurate location sensing. It used to be only obtainable by professional equipment, however, as it becomes available in off-the-shelf equipment in the format of Channel State Information (CSI), some works have reported sub-meter level accuracy [2] [22] [23].

2.2.6 ZigBee

Is a wireless communication standard designed for low power consumption, low data rate and low cost, with easy access to RSSI, aimed to be used for remote home automation applications. For these reasons, several IPSs have been proposed based on ZigBee networks. [24]

A Zigbee network works as sensor network and is composed of three kinds of components, coordinator, router, and end devices. The coordinator component is responsible for initiating the formation of the network. Every Zigbee network needs one coordinator. In [25], the authors use the end devices as mobile tags, routers as fixed receiver beacons.

2.2.7 Dead Reckoning

Dead Reckoning is and inertial positioning system and makes use of sensors on the user calculate relative position. It needs no other sensors to be installed on the building, however, with no reference points accuracy decreases quickly with time. [1]

2.2.8 Ultrasound

An ultrasound beacon periodically emits an ultrasonic pulse and and a radio message simultaneously. Since the radio message travels much faster than the ultrasound pulse, TDOA can be used to determine distance between the beacon and the receiver [8]. Ultrasound technology presents presents some downsides. The speed of sound varies with temperature. Ultrasound is sensitive to interference and multipath signals and is unable to penetrate walls [8] [26].

2.2.9 Infrared Radiation

Like Bluetooth, Infrared is available in many devices. Most Infrared Radiation devices are based on Line-of-Sight mode. It is small, lightweight, short-range narrow-transmission-angle beam fit for selective reception and aiming. As disadvantages, Infrared Radiation suffers interference from fluorescent light and sunlight and systems based on it use expensive hardware.

2.2.10 Computer Vision

Vision based tracking has been the focus of many and extensive researches [27] [28]. Positioning systems built on computer vision are usually passive systems that use scene analysis and also use computer vision for identification of the subjects. Despite this, these systems are dependent on light conditions and a large number of cameras is necessary to obtain complete line of sight coverage, which makes them expensive [29].

2.2.11 Visible Light Communication

Visible Light Communication is the use of visible light pulses or flickers to transmit data [30]. This is possible as light sources have the ability to turn on and off in very short intervals, such that, while humans cannot perceive them, light intensity sensors or cameras can register these flicks and provide relative location information accordingly.

An advantage of these systems is that they may be able to utilize light infrastructure already present in the indoor environment. On the other hand, if compatible infrastructure is not already present, it can be expensive.

2.2.12 Magnetic Positioning

Indoor Magnetic Positioning is based on using the magnetic field disturbances, created by the materials inside of buildings with magnetic proprieties, for fingerprinting. Aside from the sensor carried by the user no extra infrastructure is necessary. In some locations, the magnetic field disturbances can be quite complex, which increases positioning accuracy, however, it also makes site surveying for building the fingerprint database much more difficult [31].

2.2.13 Technologies conclusion

Many different technologies exist with potential for building solutions for user location in indoor environments. Though some technologies might be better suited to the problem than others, they all have their advantages and drawbacks. Table 2.1 displays the principle radio based technologies and their advantages and drawbacks.

Technology	Accuracy	Cost	Strengths	Drawbacks
GPS	Sub-meter using pseudolites	High	Already ubiquitous for outdoor pos.	Not suited for indoor environments without expensive infrastructure.
RFID	1m - 5m ([14] reported $<$1m)	Low	Very low cost passive tags	Low accuracy
UWB	Sub-meter	High	Very accurate and precise	Expensive tags
Bluetooth	2m - 3m	Low	Could reuse existing infrastructure	Low range
WiFi	2m	Medium	Low-cost if reuses common building infrastructure	Low accuracy Fingerprinting methods vulnerable to access point changes
ZigBee	Sub-meter	Medium	Can implement wireless sensor network techniques	Medium cost

Table 2.1: Comparing the main radio-based technologies.

2.3 Commercial Solutions

In this chapter we write about existing commercial solutions, the technologies and techniques they use and other details.

Many different commercial solutions exist. Most solutions we found online are designed specifically for one or more industries, the most common being warehouse and stock management, mining, factories and healthcare.

The most common technology used seems to be Ultra Wideband. Systems using UWB consist of a number of fixed anchors and mobile tags. At least three anchors are needed, four for a 3D space, and more depending on the area to cover. A mobile tag is designated for each asset to track. The system then uses TOA measurements to triangulate each target.

Another common solution is to use BLE or Active RFID in the same way as UWB systems, with RSSI to laterate the mobile tags locations.

Hybrid systems are also commonly used. Tracktio[1] employs a hybrid GPS and BLE system to seamlessly locate assets both indoors and outdoors.

Sewio[2] and Pozyx[3] were the only ones we found with pricing information available to the public. They both sell their Ultra Wideband real time location systems kit, with five and six anchors, respectively, and a mix of six tags for different use cases. The Sewio kit costs 2850$ and the Pozyx kit costs 4490$. Additionally, Pozyx also sell a developer version of their tags/anchors for 135$.

2.4 Hardware Options

Because designing for low cost is one of our main objectives, choosing the hardware to use in our location system will be one of the most important decisions. In fact, our accuracy requirements are low enough (room level) that cost should be the driving factor in choosing the equipment.

In this section, we go over some of the hardware options for building our IPS.

2.4.1 Arduino Nano

The Arduino Nano is a small development board based on the ATmega328 microcontroller. It has 32 KB of program space, 22 digital pins, 8 analog pins and low power consumption. The Nano can be powered with a 6-20V unregulated external power supply, or a 5V regulated external power supply.

The biggest advantage of the Arduino is the support and code available, through the Arduino development environment as well as available libraries for it. It costs around 2.5€.

2.4.2 nRF24L01+

The nRF24L01+ is an ultra low power radio transceiver module. It is operable and configurable through Serial Peripheral Interface (SPI). It has four power amplifier levels, and a one bit signal strength check.

It has support libraries, for development, requiring only a microcontroller unit, like the arduino nano. The nRF24L01+ costs less than 1€.

2.4.3 ESP8266

The ESP8266 is a Wi-Fi transciever module with integrated TCP/IP protocol. It has a built in microcontroller with 32 KB for instruction space, and 16 general purpose IO pins, but it can also be controlled by an external microcontroller. It costs around 2€.

[1] tracktio.com
[2] sewio.net
[3] pozyx.io

While the ESP8266 does not include a USB adapter, necessary for development, there are development boards that already include both. These development boards cost around 4€.

2.4.4 XBee

XBee is a family of Zig-Bee radios. It has mesh networking capabilities and RSSI access. The XBee has a big drawback however, it costs over 20€.

2.4.5 HC-05 Bluetooth Module

The HC-05 is a bluetooth module. It can communicate with a microcontroller through serial communication, has low power consumption and access to RSSI.

Cost: 3€

3

Solution Description

Contents

In this chapter we explain our choices of hardware, broadly describe our system, its architecture, main components and functionalities.

3.1 Choice of Hardware

As explained in the previous section, the main factor in choosing the hardware is cost. Taking this into consideration, the two obvious options are either a combination of Arduino Nano with the nRF24L01+ transceiver module or the ESP8266 Wi-Fi transceive. Because the Arduino has more support than the ESP we ended up choosing to build our system with Arduino Nano microcontroller board and the nRF24L01+ radio module.

With the simple RSSI capabilities of the nRF24L01+ the most suitable location sensing technique we can implement is proximity based. Each Arduino Nano and nRF24 will form a pair and each pair will the basis for the proximity nodes that will be spread throughout the home.

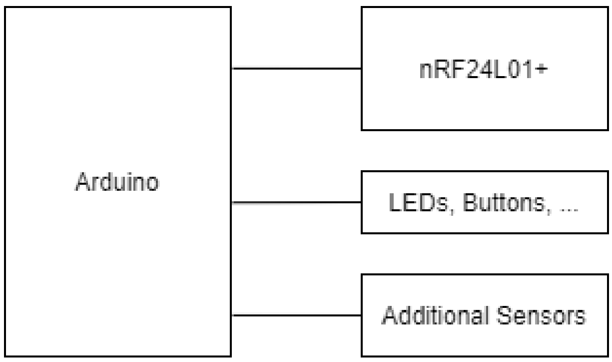

Figure 3.1: Diagram of the node.

Additionally, nodes will have some extra LEDs, buttons and may have extra sensors added later.

The power source will depend on the type and position of node. Mobile nodes will have a battery, while fixed nodes may either connect directly to a power socket or also have a battery.

3.2 The System

Our IPS consists of a node network with fixed anchor nodes and mobile nodes. Each user will carry with them a mobile node, which will periodically broadcast a message to all fixed nodes in range. Fixed radios are arranged in a tree network, and upon receiving a broadcast from a mobile node, will direct the message up through the network to the root node, also called the master node or central node. The master node in turn is feeding the messages to the server application, running in a computer.

The fixed radios are positioned one or more per each room in the house. When a fixed node receives

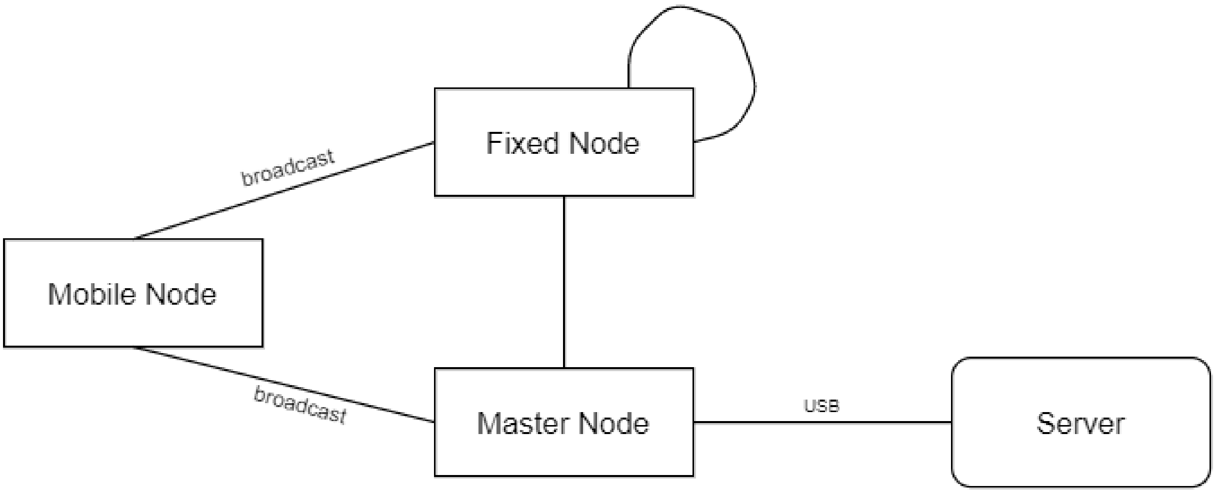

Figure 3.2: General model of the system.

a mobile broadcast, the system positions the respective user in the same room as the fixed node. When more than one fixed node receives the same broadcast a simple RSSI measurement is used. Further more, as users are not, usually, just standing about in a room, but rather doing something in a specific location, fixed nodes may be positioned for smaller, sub areas within rooms, e.g. in the desk or in the couch.

So, our system can be characterized as an active, remote positioning system, reporting absolute symbolic location. Active, because the users are required to carry a mobile node with them. Remote, as the signals are processed in a ”central site”. Absolute symbolic, considering we use the same reference grid for all locations and describe them in a human language way.

In regards to the techniques used, our solution uses both proximity sensing combined with RSSI measurements to determine the locations of the users.

Server Application

This application reads and processes all the messages from the network in order to maintain and display a representation of the network and the house plan. Through this server application, users may also interact with the nodes in the network, changing their configuration with a series of commands. The server is also responsible for computing the location of each mobile node. To do this, the server implements some disambiguation logic as well as a feature to automatically find an appropriate power level for each mobile node.

The computer running the application could be a desktop, laptop or a single board computer like a Raspberry Pi.

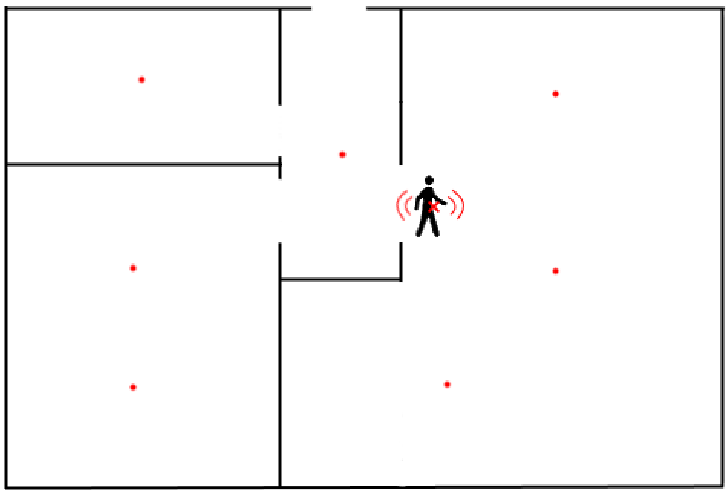

Figure 3.3: Possible deployment of the fixed nodes.

4

Implementation

Contents

In this chapter we describe in detail all implementation specifics, from th node hardware to radio and server functionality.

4.1 nRF24L01+

The nRF24L01+ is a low power, low cost, radio transceiver module that operates in the ISM band, between the 2.4 GHz and 2.525 GHz frequencies. The module allows for multiple modes of operation, implements automatic message acknowledgement and message re-transmission. It interfaces with a microcontroller unit through SPI.

In the following sections we specify the most important characteristics of the node.

4.1.1 General Characteristics

The nRF24L01+ has two different data FIFOs, one for storing transmitting payloads, the TX FIFO, and one for storing received payloads, the RX FIFO. They both have 3 levels, each level with 32 bytes of space. During Receive Mode, the TX FIFO is used to store ACK payloads for up to three different devices.

When operating in Receive Mode the nRF24L01+ works as a Multiceiver (**Multi**ple transmitters single re**ceiver**). This is a feature that allows a single module to have six different data pipes, each corresponding to a different address. This means that a nRF24L01+ operating in Receive mode can receive data from up to six other modules in Transmit mode sharing the same frequency channel.

Each data pipe has a 5 byte address. Data pipe 0's address is unique among the six data pipes. Data pipes 1 through 5 share the 4 most significant bytes. All six pipes are required to have a different least significant byte.

As previously mentioned, the radio module supports an air data rate of 250Kbps, 1Mbps or 2Mbps. At 250Kbps and 1Mbps we can have non overlapping channels with 1MHz spacing between each channel. This way, between 2.4GHz and 2.525GHz, we can have up to 126 different channels with no overlaps. At 2Mbps, 2MHz channel spacing is required for no channels to overlap. Receiver sensitivity is also affected by by the air data rate. With -82dBm receiver sensitivity at 2Mbps, -85dBm at 1Mbps and -94dBm at 250Kbps. This means that using a 250Kbps should result in a significant increase in the range of communication.

In terms of power amplification, the device has 4 options for transmitting packets. At maximum power it will transmit at 0dBm, at high -6dBm, at low -12dBm and at minimum it will transmit at -18dBm. These options consume 11.3mA, 9mA, 7.5mA and 7mA of current, respectively.

The NRF24L01 has 5 different main modes of operation, Receive mode, Transmit mode, Standby-I, Standby-II and Power Down mode.

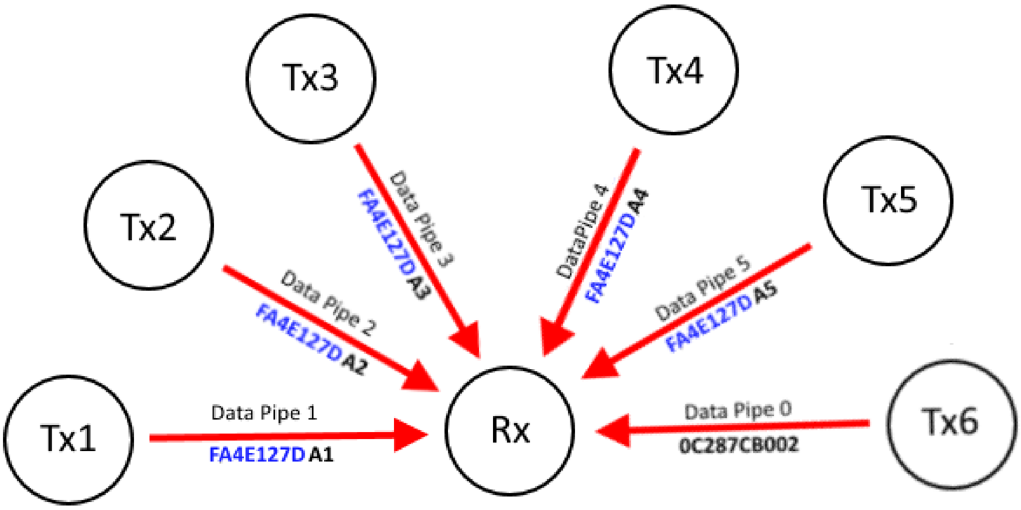

Figure 4.1: Depiction of a nRF24L01+ multiceiver. Note the shared bytes of the addresses for pipes 1 - 5.

4.1.2 Power Down Mode

In Power Down mode the SPI is still active and can be used to read and write to the data registers. In this state the module uses 900nA supply current. Transitioning from the Power Down mode will put the radio in Standby-I mode and this transition will take 1.5ms.

4.1.3 Standby Modes

There are 2 standby modes, Standby-I and Standby-II.

In Standby-I the module uses $26\mu A$ and takes $130\mu s$ to transition to either the transmit or receive mode. This state is used for its low current consumption and start times. Standby-II mode is used when the TX FIFO is emptied during Transmit mode, or when the radio tries to transit from Standby-I to Transmit Mode with the TX FIFO empty. Standby-II takes more supply current, $320\mu A$, compared to Standby-I.

4.1.4 Transmit Mode

Transmit mode is the operational mode used to send packets. In this state the radio sends, and empties, 1 or more levels of the TX FIFO, as well as handles any acknowledgements and retransmissions that may need to be done. In Transmit mode the radio consumes between 7.0mA and 11.3mA, depending on the output power.

Output Power	Power Consumption
-18 dBm	7.0 mA
-12 dBm	7.5 mA
-6 dBm	9 mA
0 dBm	11.3 mA

Table 4.1: Power consumption in Transmit Mode

4.1.5 Receive Mode

Receive mode is used to receive packets. When a received packet is validated (matching address, valid CRC) it is put in the RX FIFO. If there are no slots available in the RX FIFO, the received packet is discarded. Power supply on Receive mode depends on the air data rate, and it varies from 12.6mA with 250Kbps to 13.5mA with 2Mbps. Higher air data rate consumes slightly more power, but sending is done considerably faster, resulting in a lower average power consumption as well as reduced chance of collisions.

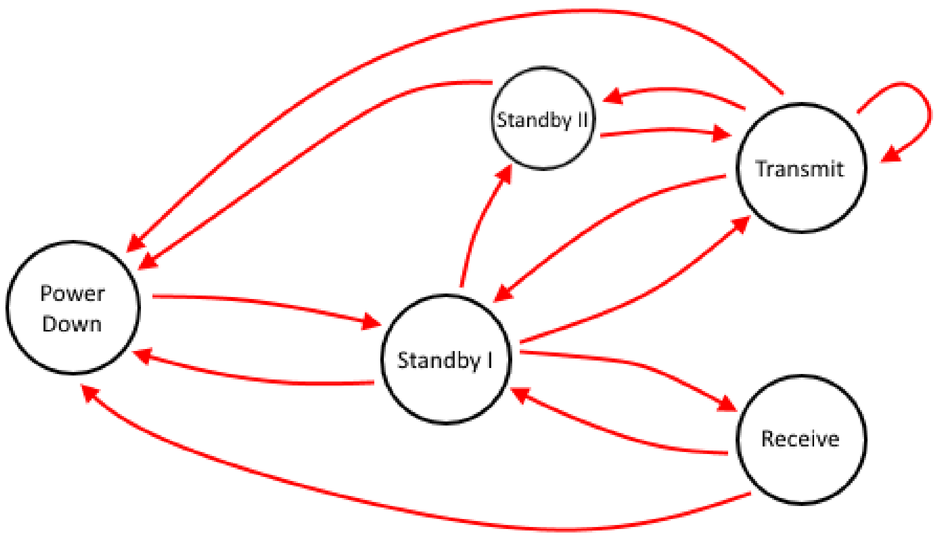

Figure 4.2: State transitions diagram for the nRF24L01+.

4.1.6 RSSI in the nRF24L01+

The device features a very basic mechanism to differentiate between messages through signal strength. This mechanism is called Received Power Detector (RPD) and it turns a bit in a register ON when a signal is present in the current channel with power above -64dBm.

4.1.7 Enhanced ShockBurst

The device's built-in data link layer, Enhanced ShockBurst, features automatic packet assembly, automatic acknowledgements and packet retransmissions.

The Enhanced ShockBurst packet format is made up of 5 fields, a preamble, the receiver address, a control field, the payload and the CRC.

The preamble consists of a 1 byte sequence used to synchronize the incoming bit stream with the receiver's demodulator. The sequence is either 10101010 or 01010101, depending on the first bit of the receiver's address.

The address field can be either 3, 4 or 5 bytes long and ensures the packet is received by the correct device.

The packet control field is made up of 3 fields, totaling 9 bits. The payload length field, 6 bits, determines the length of the payload field. The packet ID, 2 bits, prevents the receiver from sending a packet to the controller unit more than once. The NO_ACK flag tells the receiver if the received packet requires an automatic acknowledgement or not.

The payload can have between 0 and 32 bytes, defined by the user. We use a static payload size of 13 bytes.

Finally, the CRC is the mechanism used to detect errors in the packet. Can be 1 or 2 bytes long. It is a polynomial applied to the address, packet control and payload fields.

4.1.8 Control Interface

To control the nRF24L01+ a microcontroller interfaces with six digital signals available. CE (Chip Enable), CSN (Chip Select Not), SCK (Serial Clock), MOSI (Master Output Slave Input), MISO (Master Input Slave Output) and IRQ. The CSN, SCK, MISO and MOSI signals are the 4 pins used for the SPI. The IRQ is a maskable interrupt pin. The device's register map is easily configurable through the built-in 8 bit SPI commands.

4.1.9 Conclusion/ Implementation Choices

With all the capabilities and characteristics of the nRF24L01+ radio module mentioned in this section, some key aspects were taken into special consideration.

Since the vast majority of the time, of any nodule in the network, will be spent in Receive mode, minimizing power consumption in this mode is very important. Higher air data rates mean a lower transmission time, which, in turn, translates to less power consumed, on average, by the transmitter as well as a decreased chance for packets to collide. On the other hand, increasing receiver sensitivity is essential for accurate positioning.

Since our system should be adaptable to different locations, adjusting both the Air Data Rate and the Output Power in runtime is fundamental to assure ideal functioning of the system.

Finally, to control the radio device we are using the RF24 driver library[1], by TMRh20. This is an existing library designed to use the radio in accordance with the specifications from the manufacturer as laid out in this section.

4.2 Node Board

Each node consists of a microcontroller and nRF24L01+ connected in a circuit board with some additional LEDs and switches. This section details the Arduino microcontroller, the power supply and the schematic.

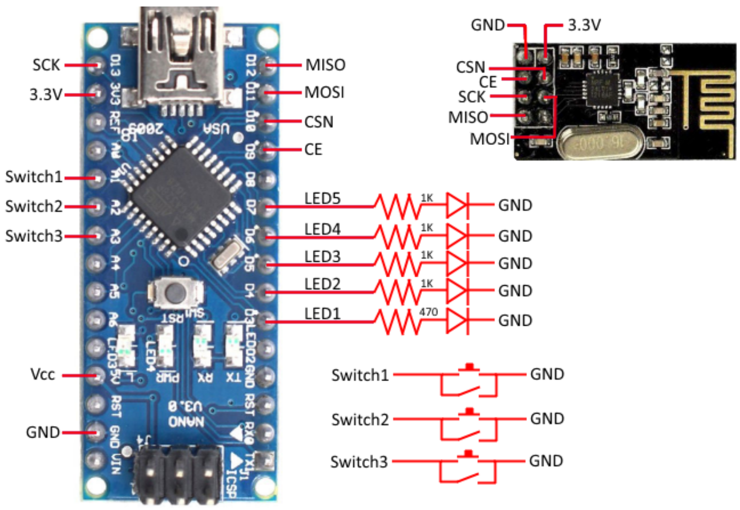

Figure 4.3: Schematic of the node's circuit.

The board features five different colored LEDs. These may be used to signal different internal states or actions in the node or network. Furthermore, three buttons and switches can be used to change the operation of the node. The functions of these LEDs and switches will be detailed in section further ahead.

Arduino

Together with the nRF24L01+ radio modules we are using Arduino Nano as a microcontroller. The Arduino Nano is an electronic platform based on the ATmega328p 8 bit AVR microcontroller. It has analog and digital pins to connect other electronic components. It also has 1KB of EEPROM, 2KB SRAM, 32KB of flash memory, 2KB of which are occupied by the bootloader and a 64 byte serial buffer.

[1] http://tmrh20.github.io/RF24

Arduinos are programmed in the Arduino language, which is based on the C and C++ programming languages.

Power can be supplied to the Arduino in several different ways. With an USB cable and ac adapter we can draw current directly from a power outlet. Another way is to connect the Arduino to a computer through USB, this is what happens with the master node. Both of these options are fairly immobile, suitable for the fixed nodes. Otherwise, power can be supplied with a powerbank or a nine volt battery. These last two are ideal for the mobile nodes.

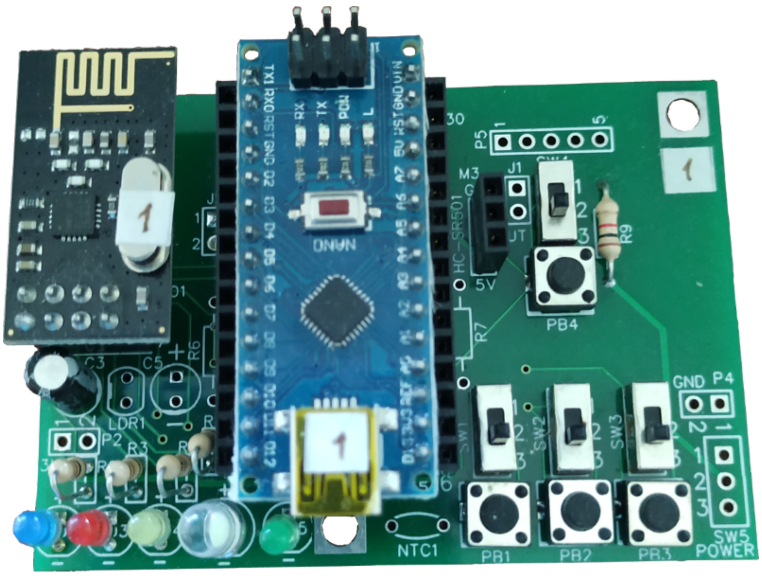

Figure 4.4: The node, with the components soldered and an Arduino Nano and nRF24L01+ radio module mounted.

4.3 Serial Communication Protocols

In this section we present the two serial protocols used.

4.3.1 SPI

SPI is a synchronous data bus, that uses different lines for communicating. One of these lines is for the clock signal, which tells the receiving devices when to sample the data bits. The protocol functions in a master/slave setup, with the master device providing the clock signal. Besides a line for the clock, SPI uses a line for the master to send data to the slaves, MOSI (Master Out Slave In), a line for the slaves to send information to the master, MISO (Master In Slave Out), and finally a line for each slave, for selecting which slave is to read or send data, SS (Slave Select).

This is the communication protocol used between the Arduino and the radio module.

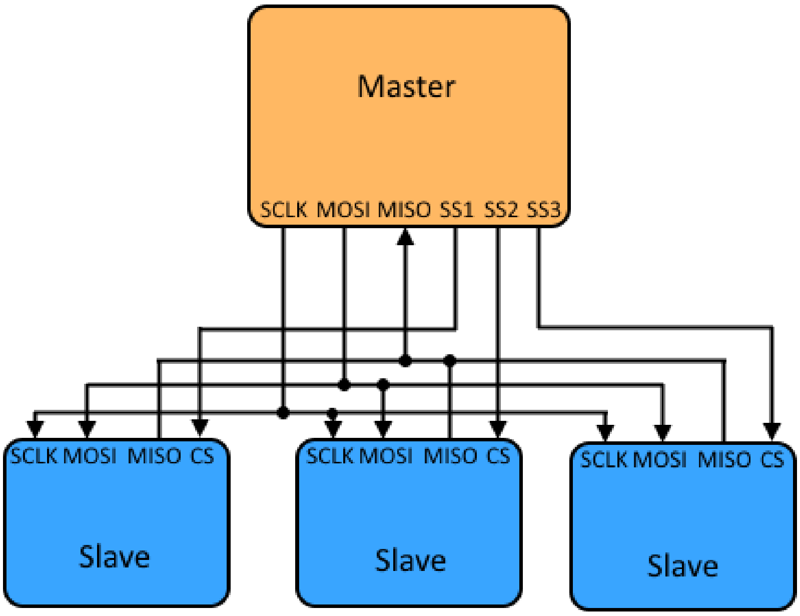

Figure 4.5: Common SPI configuration. (Image source: wikipedia)

4.3.2 Asynchronous Serial

Unlike synchronous protocols, in asynchronous serial communication the devices communicate with no clock signal telling the devices when to sample the data. Instead the devices use only 2 lines, one for transmitting and another for receiving. To make sure the devices are sampling the data at the right time both devices must agree on a number of rules that confirm that the data is transferred with no errors.

The devices must agree on a baud-rate, which determines how fast the data is transmitted and read, expressed in bits per second (bps).

When sending data, the data is divided into frames which contain a prearranged number of start bits, data bits, parity bits and stop bits in that order. The start and stop bits are used for synchronization, there is always one start bit and there can be one or two stop bits. Finally, there can be one parity bit, optionally, which is used for error checking.

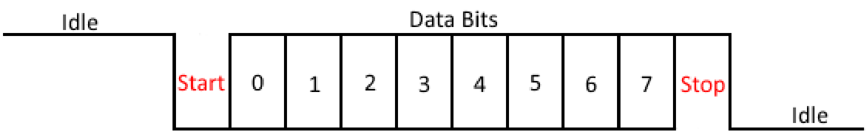

Figure 4.6: 8-N-1 configuration example. With 8 data bits, 0 parity bits and 1 stop bit.

This is the communication protocol used between the Arduino and the server.

4.4 Network Layer

The network layer we are using is implemented by the RF24 Network Layer library[2], by TMRh20, which runs on top of the RF24 driver library. This network layer library implements network ACKs, automatic routing, multicasting and an intuitive addressing format. In addition to these features, we also implemented some basic fault detection capabilities. The next few sections detail these features.

4.4.1 Multicasting

Multicasting allows one node to broadcast a message to several nodes at once. This works simply by having the receivers all listening to one address, common between them. When multicasting, acknowledgements are turned off. This is how the mobile nodes broadcast their location to fixed nodes, and how fixed nodes send commands to mobile nodes. Importantly, duplicate multicasts are not sent.

4.4.2 Topology and Addressing Format

Motivated by the radio's multiceiver capabilities, where a receiver can listen up to six other addresses, the network takes the shape of a tree. This way, each node can be listening to one parent and up to 5 children. Applying this to our intended solution, since fixed nodes need to also receive messages from the mobile nodes, we reserve one of the children's spot for the multicasting address. So, each fixed node, is listening to one parent fixed node, up to 4 children fixed nodes and the multicasting address for the mobile nodes.

Each node, has a two byte address which corresponds to a child index, '1' through '4', followed by its parent's address. So, starting with the central node, with address '0', and its children addresses '1', '2', '3' and '4', the addresses for node '1''s children would be '11', '21', '31' and '41'. The children of node '41' are '141', '241', '341' and '441'. This means that the address represents a node's position in the network and therefore, the only information a node needs to route a message is its own address and the message's destination address. Mobile nodes take the addresses with '5' as the least significant digit, which correspond to the unused branch of the network tree.

Also, since only numbers 0 through 5 are used to represent addresses, they can be stored in octal format, using only 3 bits per digit. This also allows for easy masking and manipulation of addresses. Taken this format, with 2 bytes an address can have up to 5 octal digits. So while in one hand, we may have up to 1365 fixed nodes, including the master node, on the other hand, no node can be more than 4 hops away from the master node. This means that, in large houses, with horizontal layouts, the fixed node network coverage could be limited by the distance to the central node. However, in these cases more than one network could be used.

[2] http://tmrh20.github.io/RF24Network

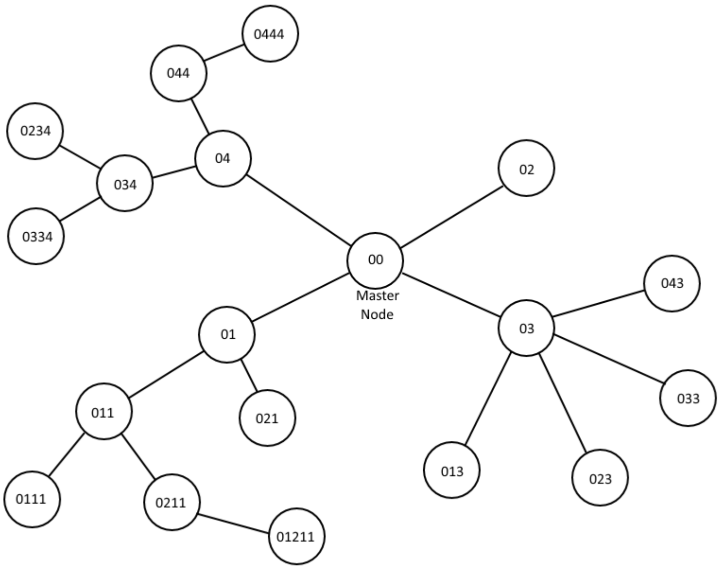

Figure 4.7: An example of a possible topology for the fixed node network.

4.4.3 Network Messages

Messages in the network consist of a header and a payload. The header is 8 bytes long and contains the destination address, the origin address, a message id, a message type, and another field reserved for use by the library. The payloads can be dynamic and be to up to 24 bytes long.

Message	Type	Payload Fields (Size in bytes)			Directionality
Location Broadcast	49	Counter (4)			Mobile Node → Fixed Nodes
Fixed Location	50	Counter (4)	Mobile Address (2)	Signal Strength (1)	Fixed Node → Master Node
Life Signal	51	Empty			Fixed Node → Fixed Node
Children List	52	Children Alive List (2)			Fixed Node → Master Node
Sync Mobile Nodes	53	Empty			Master Node → Mobile Nodes
Reset Children List	54	Empty			Master Node → Fixed Node
Request Children List	55	Empty			Master Node → Fixed Node
Change PA Level	56	Mobile Address (2)		PA Level (1)	Master Node → Fixed Node Fixed Node → Mobile Node
Change Data Rate	57	Data Rate (1)			Master Node → Fixed Node Fixed Node → Mobile Node
Change Location Rate	58	Location Rate (1)			Master Node → Fixed Node Fixed Node → Mobile Node
Change Chanel	59	Channel (1)			Master Node → Fixed Node Fixed Node → Mobile Node

Table 4.2: Payload structure for each message type in the network.

The message type field is used to identify the kind of message, so that the receiver knows the payload structure. Additionally, message types 65 through 127 will receive a network ACK and message types 128 through 255 are reserved for library use.

4.5 Nodes Common Functionality

This section describes functionality common between node types.

Nodes periodically blink a green LED, to indicate they are functional. When receiving a message from the master node, nodes will blink a yellow LED. When a fixed node receives a broadcast from a mobile node it will blink the red LED. Finally when receiving a life signal from a child node it blinks a blue LED. Additionally, the switch1 will power down the radio and light a red LED.

4.5.1 Configuration Commands

There are three operation parameters that are configurable from the server, these are, power amplifier level, air data rate and frequency channel.

To change the PA level the server will send a message to a fixed node with the new PA level and a mobile address. If the mobile address is 0 the command is addressed to the fixed node and it will change its PA level to the new one, otherwise the command is addressed to a mobile node broadcasting from nearby and the fixed node will broadcast the command. A mobile node with address matching the mobile address field in the message will change its PA level.

To change the air data rate or the frequency channel the server will send a message to all fixed nodes with the new parameter. Upon receiving the message, every fixed node will broadcast it to potential mobile nodes in the area before changing the corresponding parameter. Because, changing either the data rate or the frequency channel will affect the communication between nodes, the server must send the command to the nodes in order, by furthest away from the central node first.

4.6 Mobile Node

This section details the operation and functionality in the mobile nodes.

Periodically, mobile nodes will multicast a message to a shared address between all fixed nodes and the master node. This message contains an updated broadcast counter, which serves two purposes. The first is to identify the order of departure from the mobile node, so the server can tell which message is the most recent. The other is to avoid sending duplicate broadcasts, as they will be ignored. The broadcasting period is defined by the Location Rate, which is the parameter that determines how frequently location broadcasts are made.

Besides the configuration commands shared with other node types, mobile nodes have one other configurable parameter, the location rate. To change the location rate the server will send a message to all fixed nodes, just like when changing the air data rate or the frequency channel, the only difference being that the fixed nodes don't change anything, just broadcast the message to nearby mobile nodes.

4.6.1 Synchronization

As several Mobile Nodes can be present in the same location, each of them multicasting to all nearby Fixed Nodes, collisions become more frequent. Furthermore, since all mobile nodes are multicasting with the same period, if a collision happens then it is very likely to happen again in the future. Since multicasts cannot be acknowledged, the nodes will never know when collisions happen, therefore message retransmissions are not a viable solution. To help mitigate this problem we implemented a simple mechanism to synchronize the nodes and allot different timing windows to each one. This works by having the master node send a message to all mobile nodes. The mobile nodes will all receive the message at around the same time. Upon receiving the message the mobile nodes will delay the next multicast by an amount determined by each own unique address.

For this mechanism to work, all mobile nodes must be in direct range of the master node, as the message is a multicast. If the message were sent through the fixed node network to the locations of each mobile node, the message would not reach all mobile nodes at around the same time because the routes the message would take to each location in the network can be very different.

4.6.2 Addressing

Addressing in mobile nodes is similar to addressing in fixed nodes, with a few extra rules.

As mentioned above, mobile nodes take the addresses from the unused branch of the fixed node network tree. This means that all mobile nodes addresses have a '5' as the least significant digit. Besides this, each digit after the least significant will be multiplied by a fraction of the broadcast period. The more significant the digit, the smaller the fraction. After the least significant digit, the second least significant will be multiplied by 200ms, assuming a 1000 milliseconds multicast period, the third digit will be multiplied by 40ms and the forth digit will multiplied by 8. In order to avoid different addresses mapping to the same delay, as we can see in figure 4.8, digits after the least significant digit cannot be '5'. The only exception to this rule is node '55', this is because both address '55' and '05' map to the same delay, this way, we opted to keep the node '55' in favor of '05' because otherwise there would be no address mapping to delays between 0 ms and 200 ms.

4.7 Fixed Node

The fixed nodes act as the communication infrastructure in our system. Their job in the network is to route messages from the mobile nodes to the server and also from the server to the mobile nodes.

Whenever a fixed node picks up a broadcast from a mobile node, it makes a new message with the counter from the broadcast message, the mobile address from the header, the signal strength from the

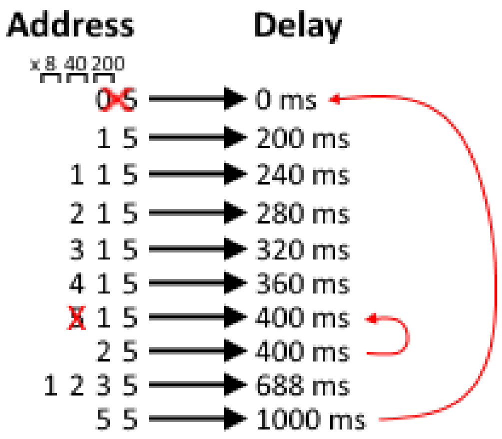

Figure 4.8: Addressing in mobile nodes and sync delays.

RPD flag and sends it to the master node through the fixed network.

Moreover, the fixed nodes have a fault detection mechanism that allows the server and the users to know which fixed nodes may have issues that prevent communication with the rest of the network, like having no power or operating outside range of communication or improper configuration parameters.

In the next two subsections we detail the fault detection mechanism and commands specific to fixed node.

4.7.1 Fault Detection

Every five seconds, each fixed node sends a ping to their parent node. After sending this life signal, the fixed node updates its list of children alive based on the pings received. Finally, nodes send their list of children to the master node. This allows the server to know if and which nodes and sections in the network may be unreachable.

To avoid congestion in the master node, fixed nodes only send their children list if they have changed from the last one sent.

4.7.2 Commands

Besides the configuration commands there are 2 other commands a fixed node can receive from the server.

'Reset Children List' makes the node empty its own list of children nodes. This will force a change in the list if the node actually has any children nodes, thereby triggering the node to re-send the list to the master node. This command is useful when loading a previously used network configuration and or when a node that was previously unreachable becomes available again.

’Request Children List’ directly asks the node to send its list of active children to the master node. While ’Reset Children List’ command is targeted at all fixed nodes, ’Request Children List’ is usually targeted at only one fixed node at a time.

4.8 Master Node

The master node is the bridge between the server and the wireless network. Its job is to send the location messages from mobile nodes and children node lists from fixed nodes to the server, via USB, as well as distributing and routing commands from the server to the network.

Serial Communication

As the master node receives messages from the fixed node network and from the mobile nodes, it encodes these messages as byte arrays and sends them through the serial interface to the server.

Furthermore, every program cycle, the node will check for available commands from the server. The master node does this by first reading a single byte. This byte identifies the command and how many arguments it has. Following the command identifier, the master node will read the available corresponding arguments.

When a command is addressed to several nodes, as is the case of some of the configuration commands, the master node will read one address at a time and send the command to the corresponding fixed node. In order to not block, the master node will only read one address per cycle and send the respective command. Because of reasons explained earlier, in section 4.5.1, the list of addresses will start with farthest nodes first, this way the last address is 00 and it marks the end of the list and the end of the command arguments.

Command	**Arguments**
Reset Children Lists	List of all fixed addresses (2 bytes each), ending with 00
Sync Mobile Nodes	None
Request Children List	Fixed address
Change Data Rate	List of all fixed addresses, ending with 00
Change Channel	List of all fixed addresses, ending with 00
Change Location Rate	List of all fixed addresses, ending with 00
Change PA Level	Mobile address (2 bytes) + New PA level (1 byte)

Table 4.3: List of commands and respective arguments read from the server.

Because the server reads and writes bytes in big endian format, when the master node reads a set of bytes from the serial port the bytes are ordered in reverse, with the most significant bytes first in the sequence, so when reading arguments with more than one byte, the byte order must be reversed.

Similarly, when encoding a packet as a byte array to be sent to the server each field must be reversed, having most significant bytes towards the start of the array.

4.9 Server's Application

The server runs an application that controls the network and interfaces with the user. The server essentially combines the information received from mobile and fixed nodes to maintain a representation of the network and derive each mobile node's positioning. It was made with Java and other platform independent tools, designed to be platform independent itself, so it should be able to compile and run on Windows, Mac OS or Linux desktops, laptops or single board computers, like a Raspberry Pi.

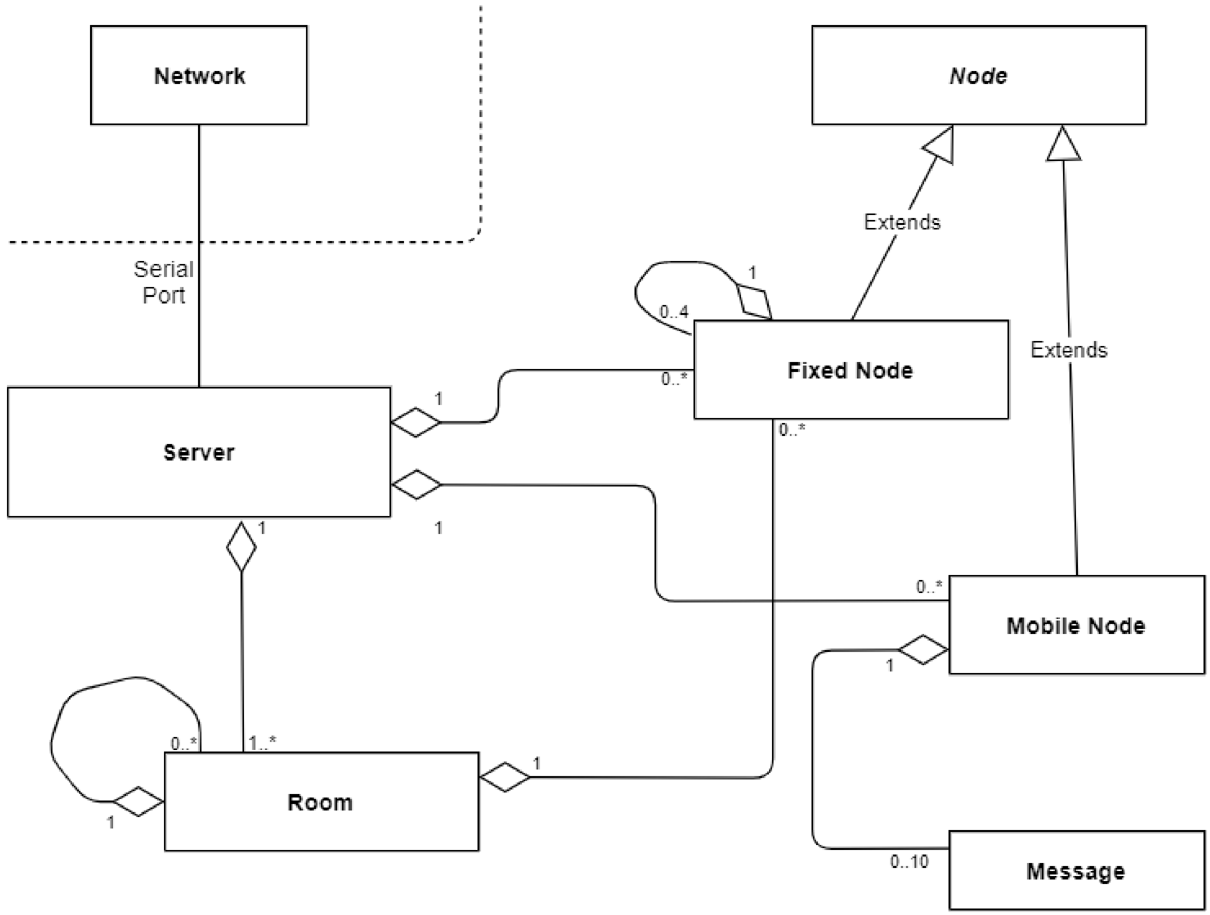

Figure 4.9: Model of the Server's Application.

This section will describe in detail how the network is represented in the server, the communication between the application and the master node, and other functionalities.

4.9.1 Message

The message class main purpose in the server is for deriving the mobile nodes' location.

Messages in the server have 5 fields: Mobile node address, fixed node address, counter, time of arrival and signal strength. These are the same fields as a location message from a fixed node, plus a time of arrival attributed as they reach the server. This time of arrival will serve as context for when displaying the history of locations of a mobile node.

A message is also used in the server to decode the bytes from the serial port.

4.9.2 Fixed Nodes

In the server, each fixed node class, besides the address and power level, is associated with a location, maintain list of fixed children nodes and a field detailing if it is currently reachable.

Every time a fixed node is registered in the server, the server finds the new node's parent node and adds the new node to its children node list. Additionally, the server also looks for children nodes that may already have been registered, and adds them to the new node's list of children. This means the fixed nodes can be registered in the server in any order. When created in the server, fixed nodes start out as unreachable until its parent reports it as alive in a children list.

4.9.3 Mobile Nodes

Each mobile node class keeps its address, power level and an ID of the person holding the node. Mobile nodes also keep a list of the last 20 messages received in the server corresponding to its own location broadcasts.

When a packet arrives in the server from a fixed node receiving a location broadcast from a mobile node, the server adds it to the list of messages of the corresponding mobile node. The list is then sorted in descending order by the message's Counter first, Signal Strength second and finally in ascending order by Time of Arrival.

This way, the last messages to leave the mobile node are at the beginning of the list, with ties broken by the "Strong Signal" field and further ties broken by the Time of Arrival; e.g., message with counter: 1000, strong signal: false, ToA: 00:00:10 $>$ message with counter: 999, S.S.: true, ToA:00:00:09 $>$ message with counter: 999, S.S.: false, ToA: 00:00:08 $>$ message with counter: 999, S.S.: false, ToA: 00:00:09 (use a table for this example?). Finally, if the list contains more than the maximum, messages are removed from the end of the list; i.e. the older ones. This list is key in determining the location of each mobile node.

When inspecting a mobile node in the interface, the full list of messages can be displayed as a history of previous broadcasts.

4.9.4 Floor Plan

The floor plan is represented as a graph, implemented as an adjacency list, where each room has a list of connected rooms. Additionally, each room has a name and a list of fixed nodes located in the room. Each fixed node in a room may be associated with a sub-location within a room or with the whole room.

The purpose of keeping the floor plan in the system is to help solve ambiguities in location sensing. For example, if the system is locating a user by two different locations but both locations are in the same room, then the user must be in that room.

4.9.5 Positioning

Mobile node location is determined from the list of messages kept by each mobile node. Location is determined directly from the first message if there are no ties at the beginning of the list and if the message was received less than two seconds ago. If there is a tie, we first check if the tied locations are in the same room. If they are not in the same room we then count the number of messages from the tied locations present the message list. This should help as the closest fixed node should also receive messages more consistently. If the tie persists, we count the number of messages from the tied rooms. If there is still a tie then, the system will indicate all tied locations.

4.9.6 Auto-Ranging

Auto-Ranging is what we call a feature that automatically helps adjust the power amplifier in the mobile nodes, to find a setting that doesn't cause an ambiguous positionings.

Whenever a mobile node has had an ambiguous location for more than two straight seconds, the system will increase its PA level, or set to minimum if its already at max level.

4.9.7 Serial Communication

To access and interface with the serial ports in the system we use jSerialComm, a lightweight and efficient library that runs on different architectures. This library allows the application to display all available serial ports, along with a description of the connected devices. This is used at the beginning of execution to have the user choose the correct serial port shared between the server device and the master node. The communication protocol used is asynchronous and the library also allows for configuration of the serial port, changing the baud rate, data bits, parity bits and stop bits. The port configuration is 57600/8-N-1.

The server can receive three kinds of messages from the network. A location message, an active children list message or an error message. The server differentiates between them by checking the message fields, if the mobile address field is 0 and the signal strength is false its a message from the master node about an error, otherwise if the counter is 0 then its a message from a fixed node with the active children list.

Another important point to consider is the compatibility of types between the server, in Java, and the network, in Arduino. Since Java does not have unsigned types, when reading the message counter from the serial port the server must cast this variable to a wider type and if the result is negative apply a bit mask to it. This is done in order to not read negative counters. The server does not need to take the same precautions with the fixed and mobile addresses because Java interprets octal values as unsigned. Furthermore, the highest address admissible is 55555, 23405 in decimal, which means all addresses are in the positive range of a short type(-32,768 to 32,767).

Furthermore, a final aspect to take into consideration. When sending a list of addresses to the master node, we may only ever send 31 at a time. This is because the arduino's serial buffer only has 64 bytes, so we may send one byte to identify the command, plus 31 two byte addresses. If we need to send more than 31 addresses we must stagger the sending.

4.9.8 Saving and Loading a Network

The application allows the user to save current network configuration to a file, to be loaded at a later time. The saved configurations include each registered node's address and power output, fixed nodes' locations, mobile nodes' holder IDs, the network's data rate, location rate and frequency channel. This does not include fixed nodes' reachability, mobile nodes' location or past broadcasts. This feature is located in the "Network Configurations" tab of the user interface.

4.9.9 User Interface

The user interface was built with Swing. Swing is a lightweight and platform independent Java toolkit for designing graphical user interfaces.

The application's graphical interface consists of 2 different windows. The first window is presented at the start of execution to display the serial ports available in the system and to have the user select the one corresponding to the master node.

The second window correlates to the application's main operations. This window contains four tabs. The first tab displays the mobile nodes in the network, their addresses, the holder's identification, and their location. The second tab displays the fixed nodes, their addresses, locations and whether their are reachable or not. In both of these two tabs the user may edit each node's information, add new nodes or delete existing nodes. When doing this dialog windows are deployed to enter new values, change configurations and report warnings. To display the nodes' two tables are used, one for mobile nodes and

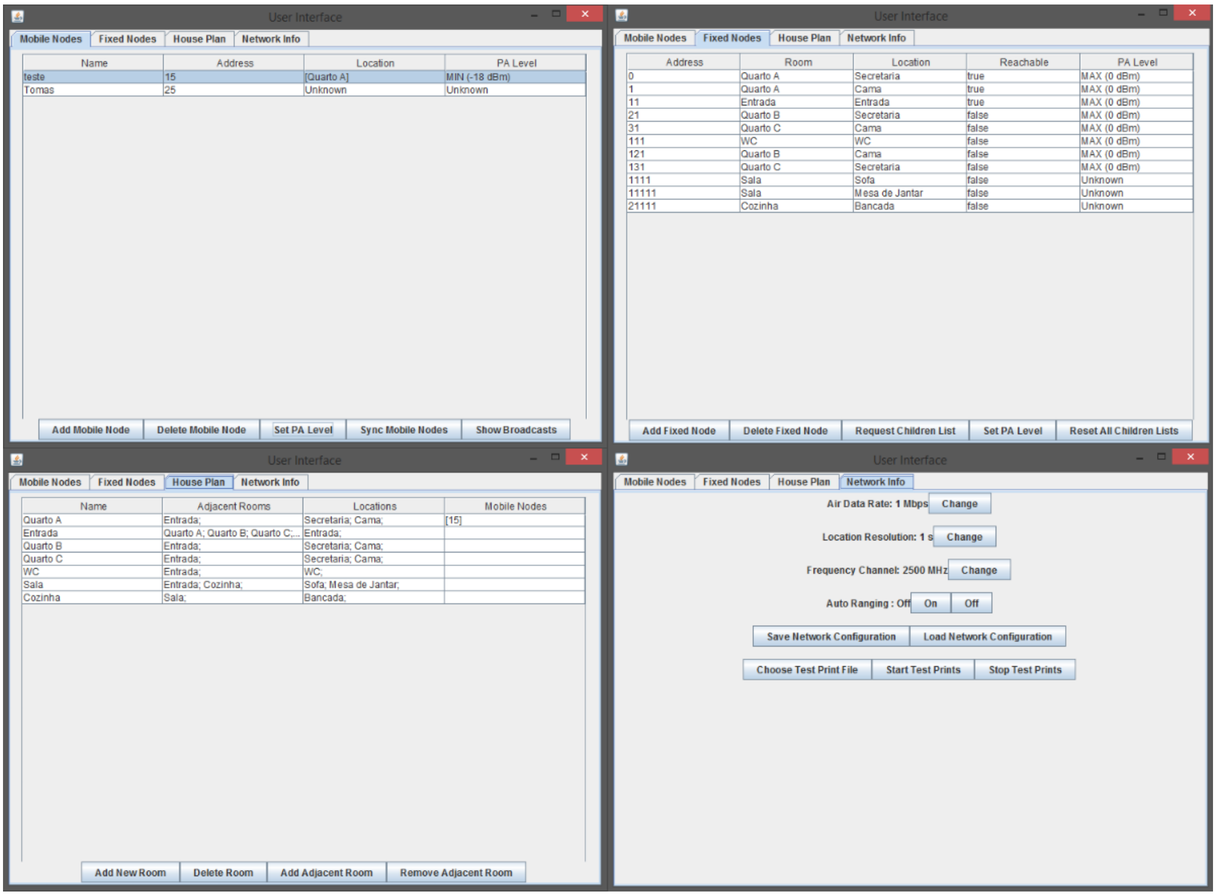

Figure 4.10: The four main screens in the interface.

another for fixed nodes. These tables are updated every second and the interface is redrawn.

The third tab display the floor plan, the fixed nodes in each room and the adjacent rooms. Another table is used for displaying the floor plan.

Lastly, in the fourth tab, network properties are displayed and may be changed. The functionality of saving and loading the network configuration to and from a file is also available in this tab.

5

Testing and Evaluation

Contents

In this chapter we first describe the tests carried out to evaluate the system's capabilities and location sensing accuracy. Next we write about how the system preformed in the tests and in other relevant metrics.

5.1 Tests

Several different tests were done, one focusing on testing the range of communication, another on normal system operation and finally a third test, with increased node density.

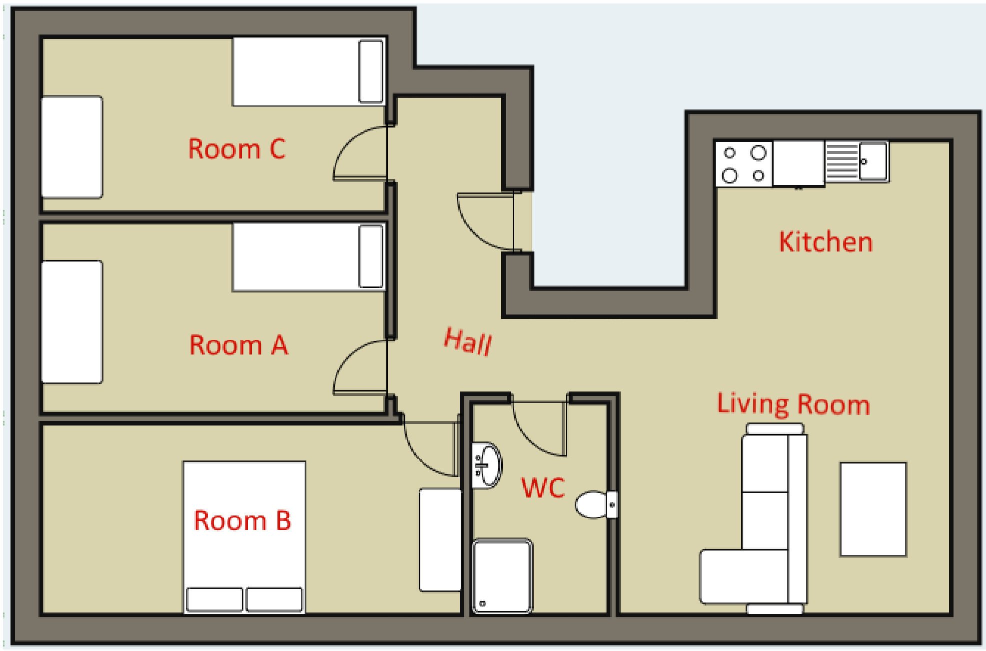

Figure 5.1: Floor plan used in the testing.

All tests were carried out in the same house, according to the room disposition in figure 5.1.

5.1.1 General Testing

This test consists of moving both a mobile node and a fixed node around another fixed node. By moving the onboard antenna we intend to study message reception and signal strength variation between mobile to fixed and fixed to fixed communications. For every position two sets of one hundred messages are broadcast from the mobile node to the fixed node. Additionally, in between each set, the mobile node is rotated ninety degrees. Finally, we also vary the the data rate to observe the receiver sensitivity.

Observations

From this test we can see that for fixed nodes to communicate consistently between adjacent rooms the power amplifier should be be kept at MAX level. In the case of fixed nodes communicating in the same room, power amplifier can be lowered to HIGH or even to LOW level.

Another important observation to note is that a strong signal is almost never received from the other side of a wall. This may vary with wall thickness, the walls in our test environment were approximately 20 cm thick.

Finally, regarding the data rate variation, we didn't find the difference in the reception to be very meaningful. Furthermore, when nodes get reset, either by ran out of battery or by manually resetting a malfunctioning node, data rate gets reset to a predefined value, which may not coincide with data rate being used by the network This results in the node not being able to communicate with the rest of the network. Because of these reasons we decided to not have the data rate be configurable.

5.1.2 One Node per Room

This test consists of analysing regular system operation, with one fixed node per room.

To position fixed nodes for this test we considered the center of each room. The mobile nodes testing locations were chosen based on the consideration that people spend most time doing things in specific locations, opposed to commuting between locations. Therefore we chose to focus on testing positions that relate to where time is most likely to be spent within the home environment, like sleeping (bed), working (desk), cooking (kitchen counter), watching TV (sofa).

In figure 5.2 we can see fixed node positions in red with node addresses and in blue the mobile positions being tested. In table 5.1 we see the location accuracy in function of the mobile node PA level.

Node	PA Level			
Position	Min	Low	High	Max
A	99%	100%	100%	100%
B	24%	93%	21%	13%
C	46%	47%	100%	88%
D	0%	100%	100%	100%
E	35%	0%	85%	91%
F	100%	0%	0%	0%
G	2%	90%	99%	76%

Table 5.1: Location accuracy.

Looking at the table, we can see that increasing the power level not always helps, in truth, it may make it so fixed nodes farther away detect the mobile node with the same signal strength as fixed nodes closer. We can see this having a big impact in locating positions F and and B.

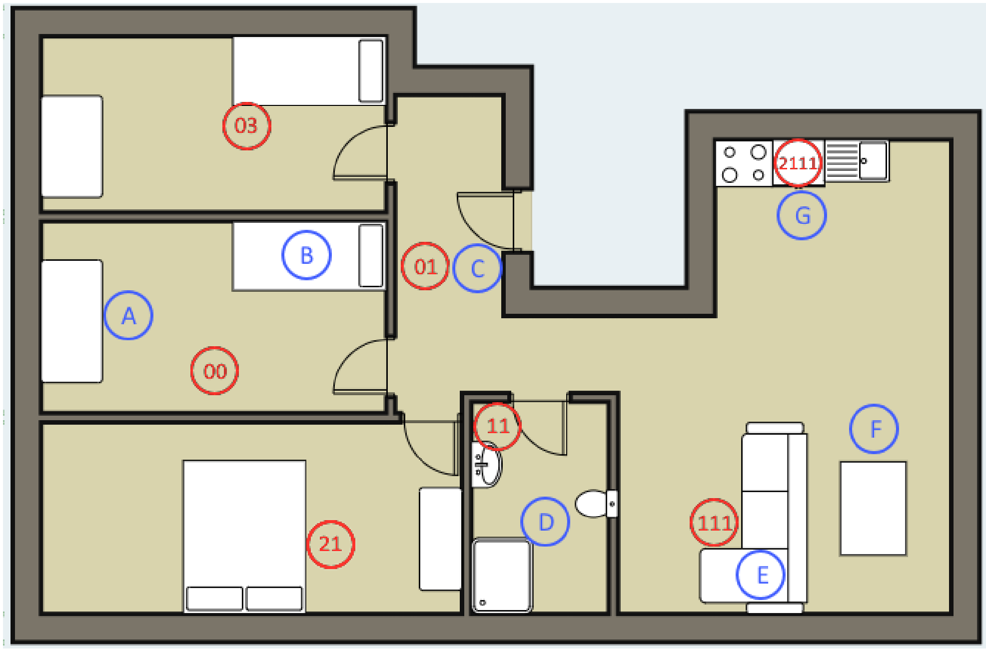

Figure 5.2: One fixed node per room.

Otherwise, the most important thing to note from the results from this test is that all tested locations seem to have a broadcast PA level for which the system accurately determines the correct room.

Also, using the number of messages received from a fixed node to solve ambiguous locations doesn't always work well. Because the system keeps a fixed number of location messages, 20 as mentioned in chapter 4, if a mobile node is in range of 3 fixed nodes during 7 consecutive broadcasts, that is a total of 21 location messages. In this case the system will untie in favor of the 2 fixed nodes closest to the central node, as they are ordered by time of arrival.

5.1.3 Increased Fixed Node Density

Instead of using one node in each room, this test consists of using one node for each commonly used space in the home. We do this for the same reasons we chose the test locations, in the previous test, as well as to test the effect of having more fixed nodes in the positioning accuracy, message throughput as well as system robustness. See figure 5.3.

When preforming the test we noticed that multiple fixed nodes were receiving the mobile broadcasts but somewhere along the way to the central node the majority of the messages were lost. This was not caused by any fixed node in specific, as after we switched fixed nodes around the problem persisted. We believe this problem is being caused by message collisions.

Overall, having more fixed nodes does not seem to be worth the cost increase, as the increase in message density causes more harm to the network than any benefit it may bring.

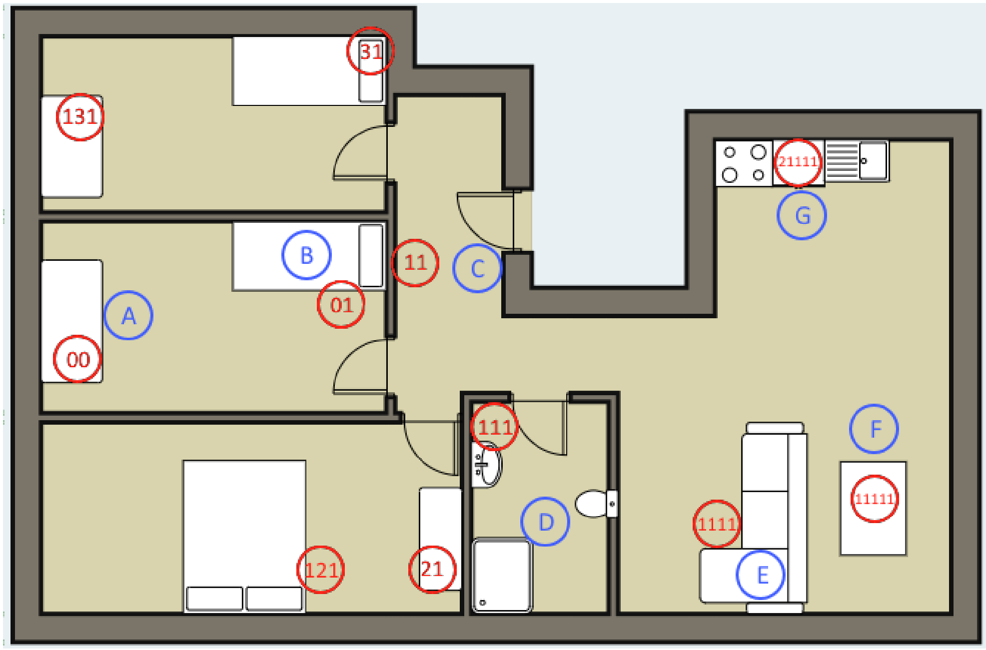

Figure 5.3: Two fixed nodes per room.

5.2 Evaluation

Location Accuracy

We calculate accuracy as the percentage of unambiguous correct locations sensed by the system. That is, the number of instances the system calculates only the correct room.

As noted in section 5.1.2, in all tested positions there is at least one PA level for which the system senses the correct room with 90% accuracy. This means that overall accuracy of the system is ensured by the Auto-Ranging function.

Cost

Cost is one of the most important measures in our work. In regards to cost, each node cost around 7€ to make. Manufacturing in scale this cost would be significantly reduced. In the main test described in this chapter we cover a small apartment of 6 rooms using 7 fixed nodes and 1 mobile node all while achieving our accuracy goals. This cost in total around 50€, which is a fraction of the cost of other systems mentioned in chapter 2.

Installation Complexity

Installation is relatively simple. As the network communication is wireless, positioning the fixed nodes is very easy. Besides this, the system requires planning the network topology and manually attributing addresses to fixed and mobile nodes. It may require a rehearsal to confirm correct functioning.

Scalability

While the system is limited in the number of hops it can have from the centre to the leaf nodes, several instances can run parallel to each other, in different frequencies. This allows for a wider coverage of bigger homes, per example, in a multi storey house we could have one instance of the system per floor, each with its own node network.

6

Conclusion

Contents

In this final chapter we go over the work done in this dissertation and leave some suggestions for possible future additions to the system.

6.1 Conclusions

In this thesis, we have discussed how emergent smart home technology is increasing the demand for context aware applications and how location systems can fit that role, focusing on indoor positioning systems.

To start, we survey commonly used techniques and technologies used to build positioning systems. We also looked at existing commercial solutions and concluded they mostly focus on achieving fine grain accuracy and there isn't an affordable option available, especially for small scale applications, like small to medium size homes.

Following that, we then describe our own solution, designed to be low cost and capable of identifying the room users are located in. Driven by cost, we write about the hardware chosen to implement our system, which forms a network of nodes spread throughout the home environment, with both fixed and mobile nodes, using radio to communicate and RSSI analysis to position mobile nodes..

Additionally, to implement our system we had to apply and build upon the existing network layer features provided by the RF24Network library, adding fixed node fault detection, mobile node synchronization and a suite of commands for node operation.

Furthermore, a platform independent server application was built to compute and consolidate location information, interface with users and manage the network. In order to achieve our positioning accuracy objectives, the server application includes some added positioning logic to help solve ambiguous locations, as well as automatic power adjusting in the broadcasting mobile nodes

Finally, we carried out some tests to validate the system, where we show the system can accurately identify the room a user is in, with a total system cost being a fraction of the one made available by commercial solutions and even other academic solutions mentioned in this dissertation.

6.2 Future Work

Although the developed system meets its objectives, we leave some improvements and additional features that could be added to improve it.

- Integrate a platform for the server application to make location information available to other context aware applications. This feature could include remote access to server information and functionality over the internet. This would be useful for taking care of people who require frequent monitoring like elderly people.

- Implement some pattern analysis to identify dangerous behaviours, like not moving or not using the bathroom for a long time, and possibly notify the user or the user's carers.
- Add an accelerometer or other inertial sensors to the node to help with positioning. Simply knowing the user's speed may be enough to make a tracking algorithm effective at telling where a user can go from any position, based on the floor plan. Additionally, these sensors could make the system capable of detecting falls.
- Integrate more sensors with the node, like a heart rate monitor or other medical sensors to keep track of the well being of the user.

Bibliography

[1] R. Harle, “A survey of indoor inertial positioning systems for pedestrians,” *IEEE Communications Surveys Tutorials*, vol. 15, no. 3, pp. 1281–1293, 2013.

[2] H. Yu, G. Chen, S. Zhao, and C. Chang, “A passive localization scheme based on channel state information in an indoor environment,” in *2017 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA)*, 2017, pp. 576–580.

[3] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, “Wifi-id: Human identification using wifi signal,” in *2016 International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2016, pp. 75–82.

[4] J. Hightower and G. Borriello, “A survey and taxonomy of location systems for ubiquitous computing,” vol. 34, 09 2001.

[5] C. Drane, M. Macnaughtan, and C. Scott, “Positioning gsm telephones,” *IEEE Communications Magazine*, vol. 36, no. 4, pp. 46–54, 1998.

[6] J. Hightower and G. Borriello, “Location sensing techniques,” 09 2001, ftp://ftp.cs.washington.edu/tr/2001/07/UW-CSE-01-07-01.pdf, last accessed on 14/09/2020.

[7] H. Liu, H. Darabi, P. Banerjee, and J. Liu, “Survey of wireless indoor positioning techniques and systems,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1067–1080, 2007.

[8] R. Mautz, “Overview of current indoor positioning systems,” *Geodezija ir Kartografija*, vol. 35, no. 1, pp. 18–22, 2009. [Online]. Available: https://www.tandfonline.com/doi/abs/10.3846/1392-1541.2009.35.18-22

[9] R. Eriksson and V. Badea, “Indoor navigation with pseudolites (fake gps sat.),” 2005, https://www.diva-portal.org/smash/get/diva2:20395/FULLTEXT01.pdf, last accessed on 14/09/2020.

[10] C. Kim, H. So, T. Lee, and C. Kee, “A pseudolite-based positioning system for legacy gnss receivers,” Mar 2014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4029644/

[11] M. Zhenhua, Z. Xingqun, and X. Hongliang, “Design and implementation of base band and if of low-cost gps l1 pseudolite,” *Computer Measurement Control*, pp. 151–153, 2010.

[12] L. M. Ni, Yunhao Liu, Yiu Cho Lau, and A. P. Patil, “Landmarc: indoor location sensing using active rfid,” in *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, 2003. (PerCom 2003).*, 2003, pp. 407–415.

[13] J. Hightower, G. Borriello, and R. Want, “Spoton: An indoor 3d location sensing technology based on rf signal strength,” 03 2000, http://www.hightowerweb.org/pubs/hightower2000indoor/hightower2000indoor.pdf, last accessed on 14/09/2020.

[14] C. Huang, L. Lee, C. C. Ho, L. Wu, and Z. Lai, “Real-time rfid indoor positioning system based on kalman-filter drift removal and heron-bilateration location estimation,” *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 3, pp. 728–739, 2015.

[15] S. Gezici and H. V. Poor, “Position estimation via ultra-wide-band signals,” *Proceedings of the IEEE*, vol. 97, no. 2, pp. 386–403, 2009.

[16] S. Gezici, Zhi Tian, G. B. Giannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, and Z. Sahinoglu, “Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks,” *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 70–84, 2005.

[17] M. Gunia, F. Protze, N. Joram, and F. Ellinger, “Setting up an ultra-wideband positioning system using off-the-shelf components,” in *2016 13th Workshop on Positioning, Navigation and Communications (WPNC)*, 2016, pp. 1–6.

[18] A. A. Kalbandhe and S. C. Patil, “Indoor positioning system using bluetooth low energy,” in *2016 International Conference on Computing, Analytics and Security Trends (CAST)*, 2016, pp. 451–455.

[19] G. Jekabsons, V. Kairish, and V. Zuravlyov, “An analysis of wi-fi based indoor positioning accuracy,” *J. Riga Technical University*, vol. 44, pp. 131–137, 01 2011.

[20] H. Chabbar and M. Chami, “Indoor localization using wi-fi method based on fingerprinting technique,” in *2017 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS)*, 2017, pp. 1–5.

[21] Z. Yang, Z. Zhou, and Y. Liu, “From rssi to csi: Indoor localization via channel response,” *ACM Comput. Surv.*, vol. 46, no. 2, Dec. 2013. [Online]. Available: https://doi.org/10.1145/2543581.2543592

[22] K. Wu, Jiang Xiao, Youwen Yi, Min Gao, and L. M. Ni, “Fila: Fine-grained indoor localization,” in *2012 Proceedings IEEE INFOCOM*, 2012, pp. 2210–2218.

[23] J. Xiao, K. Wu, Y. Yi, and L. M. Ni, "Fifs: Fine-grained indoor fingerprinting system," in *2012 21st International Conference on Computer Communications and Networks (ICCCN)*, 2012, pp. 1–7.

[24] R. Brena, J. García-Vázquez, C. Galván Tejada, D. Munoz, C. Vargas-Rosales, J. Fangmeyer Jr, and A. Palma, "Evolution of indoor positioning technologies: A survey," *Journal of Sensors*, vol. 2017, 03 2017.

[25] T. Alhmiedat and S.-H. Yang, "A zigbee-based mobile tracking system through wireless sensor networks," *Int. J. Adv. Mechatron. Syst.*, vol. 1, 01 2008.

[26] Z. Farid, R. Nordin, and M. Ismail, "Recent advances in wireless indoor localization techniques and system," *Journal Comp. Netw. and Communic.*, vol. 2013, pp. 185 138:1–185 138:12, 2013.

[27] T. B. Moeslund and E. Granum, "A survey of computer vision-based human motion capture," *Computer Vision and Image Understanding*, vol. 81, no. 3, pp. 231 – 268, 2001. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S107731420090897X

[28] T. Moeslund, A. Hilton, and V. Krüger, "A survey of advances in vision-based human motion capture and analysis," *Computer Vision and Image Understanding*, vol. 104, pp. 90–126, 11 2006.

[29] H. Koyuncu and S.-H. Yang, "A survey of indoor positioning and object locating systems," 2010, http://www.academia.edu/download/33814983/A_Survey_of_Indoor_Positioning_and_Object_Locating_Systems.pdf, last accessed on 14/09/2020.

[30] T. Komine and M. Nakagawa, "Fundamental analysis for visible-light communication system using led lights," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 1, pp. 100–107, 2004.

[31] B. Li, T. Gallagher, A. G. Dempster, and C. Rizos, "How feasible is the use of magnetic field alone for indoor positioning?" in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2012, pp. 1–9.