2023-08-31

Ultra-Wideband Trained Artificial Neural Networks for Bluetooth Proximity Detection in Small Crowded Areas

Debnath, Satinath

UNIVERSITY OF CALGARY

Ultra-Wideband Trained Artificial Neural Networks for Bluetooth Proximity Detection in Small Crowded Areas

by

Satinath Debnath

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN GEOMATICS ENGINEERING

CALGARY, ALBERTA
AUGUST, 2023

Abstract

Estimating the distance between indoor users is increasingly important in unexpected ways. One specific example is the need for electronic contact tracing demonstrated during the recent global pandemic. Smartphones are now routinely equipped with Bluetooth Low Energy radios, among other sensors, and these can be used for proximity detection based on received signal strength that is subject to errors due to poor modelling of the indoor propagation environment. Some high-end smartphones have now also been equipped with ultra-wideband ranging radios that provide a much more precise range measurement.

This thesis demonstrates the concept of using a limited number of UWB-equipped smartphones to gather data to train Artificial Neural Networks (ANN) to improve short-range distance estimation among Bluetooth users. The trained RSSI to range model can be used for proximity determination by other Bluetooth users in small, crowded areas. Two ANN algorithms were trained using RSSI measurements from three BLE advertising channels and UWB range as ground truth and training data. The initial training and testing were conducted in a semi-empty office laboratory with 2130 observations. The RF model used 1917 samples (90% of data) for training and 213 samples (10%) for testing, while the CNN method used 1704 samples (80% of data) for training and 426 samples (20%) for evaluation.

The trained neural network models were tested in two other office environments under different user conditions. The results indicate that the ANN models can estimate proximity in a new environment without further training with a mean error of less than 1.2 metres, within a range of up to 6 metres at line-of-sight (LOS). In highly constrained non-line-of-sight (NLOS) areas in the first office room, the proposed models provided proximity accuracy better than 2.9 metres. Furthermore, during testing across two adjacent office environments, each containing a single BLE device with complex furniture arrangements, the ANN models showed the proximity between the BLE devices with an error of less than 2-3 metres.

Acknowledgements

My deepest gratitude and heartfelt thanks go to my supervisor, Dr. Kyle O' Keefe, for his continuous support, guidance, valuable discussions, and financial assistance throughout my graduate studies. I am extremely thankful for the source of my research funding: Natural Sciences and Engineering Research Council of Canada (NSERC). I want to thank my colleagues from the "PLAN" group: Naman Agarwal, Mike Blois, Rene Manzano Islas, Baoyu Liu, and Zelin Zhou. Their contribution to creating a pleasant and supportive environment for studies and research has been invaluable. Thank you, guys, for all your support and many trips to beautiful places surrounding Calgary, including Banff, Drumheller, Yoho, and many others. In addition, I extend my gratitude to other graduate students, faculty, and staff members of Geomatics Engineering for always being friendly and interactive.

Many have selflessly shaped my life through their work, love, affection, and blessings. My grandfather (Late Laxminarayan Debnath), Ms. Sadhana Bhowmik (my primary school teacher), and Mr. Ramakrishna Das (teacher and well-wisher) are some of them who have gone the extra mile during my formative years to help me way beyond the scope of a typical relationship. My heartfelt gratitude goes to Mr. Jayanta Kumar Ray, my guide and mentor, whose encouragement and inspiration led me to pursue graduate studies at Calgary. I am also profoundly thankful to Accord Software & Systems, Bangalore, India, for providing me with my initial professional platform and the opportunity to delve into the fascinating world of GPS. I sincerely appreciate Mr. Naveen GS, Mrs. Ashwitha Setiger, and Mr. Kiran S of Accord Software and System, whose insightful discussions significantly influenced my decision to embark on the M.Sc. program.

Finally, I am deeply grateful to my parents, brother, friends, and well-wishers for their unwavering support, encouragement, motivation, and driving force throughout my life. Last but not least, I thank Almighty for allowing me to travel to North America and complete my graduate studies at the esteemed University of Calgary.

Dedicated To

My parents, grandparents

My brother: Santanu

And

Almighty for what I am today

Table of Contents

List of Tables

List of Figures

List of Symbols

Symbol	Definition
B_{frac}	Fractional bandwidth
c	Speed of light=299792458 m/s
f	Signal frequency
f_h	Upper frequency
f_l	Lower frequency
G_r	Receiver antenna gain
G_t	Transmitter antenna gain
H	Measurement design matrix
J	Mean square error
K_k	Kalman gain
m	Number of training samples
n	Path loss exponent
P_k	State error covariance
P_r	Receiver power (mW)
P_t	Transmitted power (mW)
Q_k	System noise covariance
R_k	Measurement noise covariance
R_x	Receiver
T_x	Transmitter
w_0	Neural network bias
X_σ	Observation error
x_k	State estimate
y_k	Measurement residual
z_k	Observations
λ	Wavelength
$\phi(x)$	Activation function

List of Abbreviations

Abbreviation	Definition
ADS-TWR	Asymmetric Double-Sided Two-Way Ranging
AFH	Adaptive Frequency Hopping
AI	Artificial Intelligence
ANN	Artificial Neural Network
AOA	Angle of Arrival
AP	Access Point
AT&T	American Telephone and Telegraph Company
BLE	Bluetooth Low Energy
BPNN	Back Propagation Neural Network
CCIT	Calgary Centre for Innovative Technology
CDF	Cumulative density function
CSI	Channel State Information
CNN	Convolutional Neural Network
dB	Decibel
dBm	dB milliWatt
DEV	Development
1-D	One Dimensional
2-D	Two Dimensional
3-D	Three Dimensional
DL	Deep Learning
DT	Decision Tree
DNN	Deep Neural Network
DoD	Department of Defense
DS-TWR	Double-Sided Two-Way Ranging
FCC	Federal Communications Commission
FFSPM	Friis Free Space Propagation model

GFSK	Gaussian Frequency Shift Keying
GHz	Gigahertz
GNSS	Global Navigation Satellite Systems
GPS	Global Positioning System
GSM	Global System for Mobile Communication
IEEE	Institute of Electrical and Electronics Engineer
IMU	Inertial Measurement Unit
IoT	Internet of Things
IPS	Indoor Positioning Systems
ISM	Industrial, Scientific, and Medical
INS	Inertial Navigation Systems
KF	Kalman Filter
LED	Light Emitting Diode
LOS	Line-Of-Sight
LPS	Local Positioning System
LSTM	Long Short-Term Memory
MEMS	Micro-Electromechanical Systems
MHz	Megahertz
ML	Machine Learning
MSE	Mean Square Error
MLP	Multilayer Perceptron
mW	MilliWatt
NFC	Near Field Communication
NLOS	Non-Line-Of-Sight
NRF	Non-Radio Frequency
OWR	One Way Ranging
PSO	Particle Swarm Optimization
PC	Personal Computer
PCS	Personal Communication Service

PSD	Power Spectral Density
RBF	Radial Basis Function
ReLU	Rectified Linear Unit
RF	Random Forest
RFID	Radio Frequency Identification
RNN	Recurrent Neural Network
RP	Reference Point
RSSI	Received Signal Strength Indicator
RTLS	Real-Time Locating System
SDS-TWR	Symmetric Double-Sided Two-Way Ranging
SIG	Special Interest Group
SS-TWR	Single Sided Two-Way Ranging
TDD	Time Division Duplexing
TDoA	Time Difference of Interval
TH	Time-Hopping
TOF	Time of Flight
ToA	Time of Arrival
TV	Television
TWR	Two Way Ranging
UMTS	Universal Mobile Telecommunications System
US	United States
UWB	Ultra-Wideband
VLC	Visible Light Communication System

Chapter 1

Introduction

1.1 Background

The Global Positioning System (GPS) was successfully designed and developed by the US Department of Defense in the early eighties to serve the US military in outdoor environments with good navigation accuracy. However, GPS signals can be attenuated, reflected, or entirely blocked by physical materials such as concretes, steel and glass inside buildings leading to a weak or lost signal. Subsequently, researchers began designing alternative technologies, widely known as Indoor Positioning Systems (IPS), and locating objects in indoor environments. Over the last three decades, indoor positioning technologies such as Zigbee, ultrasound, visible light communication, WiFi, Bluetooth, Ultra-wideband, etc., have evolved in terms of cost, accuracy, and reliability. Today one of the fastest-growing and widely used technologies in wireless systems for precise indoor positioning is Bluetooth Low Energy (BLE). Bluetooth has emerged as the lowest power, low cost, low complexity, and low maintenance requirements system of all available indoor positioning systems. Despite this, the technology still faces limitations owing to its inherent signal transmission characteristics. BLE estimates the distance between two devices by measuring the signal strength using a radio propagation model, whose accuracy highly depends upon the model and the environment. Alternatively, the Ultra-wideband (UWB) technology, the contemporary positioning system, is an emerging precise indoor positioning method that uses the time of flight of radio waves to estimate the distance between two devices. Although the cost of UWB radios has decreased significantly in the past decade, they still need to be made available to mass users. This thesis assumes that in the near term, a small fraction of mobile phones (or expert users) will be equipped with UWB-ranging radios and will provide more accurate training data for general BLE users. That is, the UWB range information will be used to train the artificial neural network (ANN) whenever available, providing better range estimation and thus improving the accuracy of the BLE system.

1.2 Indoor Positioning

The term “Indoor positioning” refers to determining the position of an object in closed environments such as offices, buildings, malls, airports, hospitals, etc. Global Navigation Satellite Systems (GNSS) are known for determining the 3D position of an object in the open outdoor environment with a minimum of four satellites in view. The same satellite signals become weak after reaching the ground and cannot penetrate buildings to provide range measurements for locating the position of the object inside. Because the COVID-19 pandemic and post-pandemic era created a need for high-accuracy proximity detection, the demand for highly reliable indoor localization applications has skyrocketed. There are widely used outdoor systems such as GNSS. In-contrast, there are a broad range of Indoor Positioning Technologies that can be classified into two major categories: Radio Frequency Based Systems (RF) and Non-Radio Frequency Based Systems (NRF) (Kim Geok et al., 2021a). Radio Frequency (RF) technologies includes Bluetooth, WiFi, UWB, Zigbee, Cellular and Near field Communication (NFC) system, etc. Similarly, Non-Radio Frequency (NRF) technologies includes Inertial Navigation System (INSs), Visible Light Communication (VLC), Ultrasound and Vision Sensors, etc. Among them, the most popular and often used technologies for indoor positioning are Radio Frequency (RF) technologies, Inertial Navigation Systems (INSs), Vision sensors. However, inertial sensors suffer from time integrated drift errors from their sensor measurements, and frequent calibration or external update is required to control these errors (Aggarwal et al., 2008). Vision-based surveillance or monitoring technologies are vulnerable to illumination and require a large data set of features (Minh Dang et al., 2020).

1.3 Literature Review

With continuous advancements in science and technology, indoor positioning techniques experienced cost reduction, improved accuracy, decreased power consumption, and enhanced reliability. There are several indoor positioning technologies available today. The indoor positioning technologies are divided into two groups (i) Radio frequency-based technology and (ii) Non-radio frequency-based technology. First, the chapter explores common non-radio frequency-based positioning techniques, including the Ultrasonic positioning method, Inertial Navigation Systems, Computer vision techniques, visible light technology, etc. Then it provides the background of a few standard radio-based positioning systems, which include: WiFi, Bluetooth, UWB, and Zigbee. Next, the chapter discusses Artificial Neural Networks (ANN), a recent attractive methodology for radio-frequency-based indoor positioning. Each of the indoor positioning technology is briefly summarized in subsequent subsections to provide a comprehensive understanding.

1.3.1 Indoor Positioning Technologies

1.3.1.1 Non-radio Frequency-based Technology

Infrared Positioning Technology uses electromagnetic signals or radiation within the infrared band of the electromagnetic spectrum for positioning purposes. Infrared rays, which are invisible to the human eye and primarily emitted as heat, have traditionally been used for thermal imaging and scientific research. However, advancements in technology have enabled the use of infrared light for information transmission and positioning. Infrared positioning can generally be categorized into two modes: active and passive. Active infrared involves the artificial emission of infrared light from specialized sensors (IR sensors) to measure angles or distances. In contrast, passive infrared uses infrared light generated by sources like the human body or other animals, which is then captured by an infrared camera for angle or distance measurements. The measured distance or angle, combined with a positioning algorithm, can be employed for object localization.

Active Badge is a renowned indoor location sensing system initially developed by AT&T labs at the University of Cambridge (Want et al., 1992), using active infrared light. A combination of an active infrared sensor and passive landmarks has been proposed for localizing mobile robots in indoor environments, resulting in significantly improved accuracy (Oh et al., 2014). In another study, a local positioning system was investigated for smart devices emitting active infrared light captured by stationary mounted cameras. The research demonstrated high positioning accuracy of approximately 8 cm for short-range applications and around 16 cm for an area of approximately 100 m^2 (Aitenbichler and Muhlhauser, 2003). Recent work proposed using low-range infrared signals to achieve highly precise angle of arrival (AoA) estimates for localization without requiring line-of-sight propagation (Arbula and Ljubic, 2020). A passive indoor visible light localization system was introduced, employing deep learning techniques to localize specific objects of interest (OI). The research indicated that the position of the object influences the impulse response (IRs) among source and receiver pairs, enabling object localization without the need for a line-of-sight path measure (Majeed and Hranilovic, 2021). However, infrared positioning technology faces several limitations despite its simple mechanism, relatively low cost, and high accuracy. Infrared light is sensitive to obstacles and cannot penetrate them, resulting in a limited operational distance. Typical working coverage areas using infrared technology range between 1-5 metres (Mautz, 2012). Another study reported that the accuracy of infrared positioning methods using angle of arrival algorithms is limited to a few metres.

Moreover, infrared positioning technology has not made significant advancements compared to other technologies. With its limited transmission coverage and susceptibility to interference, it is not widely adopted for indoor localization purposes.

Ultrasound positioning-based technology uses sound frequencies higher than the audible range (above 20 KHz) for positioning purposes. Sound signals, which are pressure waves, travel through the air faster than electromagnetic signals. In an ultrasound system, the time taken for an ultrasound signal to travel from a transmitter to a receiver is measured. This time of flight (ToF) information is then used to calculate distance and determine the user's position through trilateration. However, accurate ToF estimation relies on temporal synchronization between the transmitter and receiver. Ultrasound positioning offers high accuracy, typically within a few centimetres. A robust and highly accurate ultrasound indoor positioning system is proposed in (Qi and Liu, 2017). The simulation results showed that when the ultrasound signal had a clear line of sight (LOS), the maximum positioning error for a moving robot was only 10.2 mm. One disadvantage of the ultrasound technique is calculating flight time for distance estimation. The velocity of sound waves in air is not constant and is significantly influenced by environmental factors such as humidity and temperature (Bohn, 1988). Humidity causes ultrasound signals to attenuate rapidly and travel shorter distances, while temperature directly affects the speed of sound. To address this issue, most ultrasound systems incorporate temperature compensation sensors. Additionally, ultrasound technology is highly susceptible to interference. Persistent noise sources can significantly degrade system performance, requiring the implementation of separate algorithms to filter out location estimates when non-persistent noise is present (Ijaz et al., 2013). Due to these aforementioned disadvantages, the ultrasound technique is not considered reliable for indoor positioning.

Visible Light Communication (VLC) is a wireless communication technology that uses visible light, which is perceptible to humans. In VLC positioning systems, light-emitting diodes (LEDs) are employed to transmit signals for determining the position of mobile devices. The mobile devices are equipped with built-in cameras or photodiodes to receive signals containing identification information. The user's location is determined based on the received signal strength from the light source. VLC-based Indoor Positioning Systems (IPS) offer several advantages. They do not generate radio frequency (RF) interference, making them suitable for radio-sensitive areas such as healthcare units. Additionally, VLC signals are not susceptible to issues like multipath or propagation loss, ensuring a more reliable and predictable positioning system. Moreover, VLC-based IPS are cost-effective, environmentally friendly, and possess other notable benefits. The accuracy of VLC-based positioning systems discussed in these papers demonstrated precision within a few centimetres (Kim et al., 2013; Wang et al., 2013).

Inertial Navigation Systems (INS) is a self-contained navigation technique that uses measurements from accelerometers and gyroscopes to track the position and orientation of an object relative to a known starting point, orientation, and velocity. It operates independently without relying on external signals or connectivity. INS is a non-radiating and non-jammable system, making it resistant to interference and capable of providing dead-reckoning navigation (Barshan and Durrant-Whyte, 1995). The system derives measurements, such as attitude, velocity, and direction, from the object's acceleration measurements, assuming a known starting point. An external GPS receiver or an operator can provide this initial starting point. The object's location is continuously updated based on the forces experienced by the accelerometers through mathematical calculations. INS uses a computer system to process sensor measurements, perform complex calculations, and generate location information. Inertial Navigation finds application in a wide range of areas, including aircraft navigation, tactical and strategic guided missiles, submarines, ships, and satellite launches for important scientific missions. Unlike GPS technology, inertial navigation is autonomous after initialization, does not rely on satellite connectivity, and provides more precise data. Recent Micro-Electro-Mechanical Systems (MEMS) advancements have resulted in smaller, lighter inertial navigation systems. Inertial navigation systems can be found in various indoor navigation applications. For instance, a study proposes an inertial navigation system with a wireless reference system to accurately estimate the long-term position of a moving cart (Coronel et al., 2008). Integration of Received Signal Strength Indication (RSSI) and inertial measurements is employed to locate and track pedestrians in indoor scenarios (Zhang et al., 2014). Another research presents an INS/WiFi-based hybrid smartphone indoor localization system (Chen et al., 2018). Additionally, an indoor laser-aided inertial navigation system that uses an Inertial Measurement Unit (IMU) and a 2D laser scanner is presented for assisting the visually impaired, among other applications.

1.3.1.2 Radio Frequency-based Technology

RF signals are electromagnetic or radio waves that lie in the frequency range between 20 KHz to 300 GHz. The primary objective of this technology was for the communication system as at these frequencies, energy from oscillating currents radiates off to space as radio waves. Most commonly known devices, such as transmitters, receivers, and televisions, use radio frequencies for communication. Radio communication systems can provide location information based on three key characteristics of radio signals: the power of the transmitted signal, the propagation time, and the direction of the transmitted signal (Kim Geok et al., 2021a). Based on the signal measurements, (Liu et al., 2007) classified RF-based indoor localization into three categories: (i) Received Signal Strength Indicator (RSSI), (ii) Time of Flight (TOF) and (iii) Angle of Arrival (AOA)-based methods.

ZigBee is an IEEE 802.15.4 standard designed to meet the requirements of low-cost implementation, low-power devices, and low data rates (20-250 Kbps) for short-range wireless communication. Its primary objective is device control and monitoring. The ZigBee system operates in three frequency bands: 868 MHz (Channel 0), 915 MHz (Channels 1-10), and 2.4 GHz (Channels 11-26). The availability of different channels within each frequency band enables efficient spectrum utilization. The standard incorporates dynamic channel selection functionality, allowing devices to scan and select the appropriate channel based on various factors such as beacon detection, receiver energy detection, link quality indication, and channel switching (Ergen, 2004). The ZigBee system follows three network architectures: star, tree, and mesh. Devices within the ZigBee network are classified into three categories: ZigBee Coordinator, ZigBee Router, and ZigBee End devices. ZigBee End devices are equipped with ZigBee radios to facilitate communication (Kim Geok et al., 2021b). ZigBee technology initially found applications in areas such as home automation (remote light and thermostat monitoring and control), urban traffic light control, medical care, agriculture, and more, thanks to its low energy consumption and enhanced security features (Wheeler, 2007).

A ZigBee positioning system consists of a network of sensors, including reference nodes with known physical locations and a target node, along with wireless network positioning algorithms. These algorithms use RSSI values and employ techniques like fingerprinting and propagation models similar to WiFi and Bluetooth positioning. In (Tadakamadla and Oelmann, 2006), a model is presented that uses ZigBee technology, RSSI, and the Euclidean distance method to monitor the presence and movement of vehicles. Another approach described in (Fang et al., 2012) proposes an enhanced ensemble method that combines positioning and fingerprinting algorithms to achieve more accurate location estimation within a ZigBee sensor network. However, some research suggests that ZigBee-based positioning may suffer from inaccuracies due to signal strength and interference (Chu et al., 2011). Additionally, the short-range and high latency limitations of 802.15.4 wireless technology pose challenges for real-time ZigBee positioning using RSSI measurements, primarily due to network interference (Jianyong et al., 2014). Therefore, further improvements are needed to address these drawbacks and make ZigBee technology suitable for indoor localization applications.

Radio Frequency Identification (RFID) is a radio technology used to track people or objects through RFID tags. RFID localization can be categorized into a reader (decoder or transceiver) and tag localization (Sanpechuda and Kovavisaruch, 2008). There are two main types of RFID systems: Active RFID and Passive RFID. Active RFID operates in the ultra high frequency and microwave range, requiring a power source to transmit identification and other data. They offer a longer range for localization and tracking. Passive RFID operates at a shorter range (1-2 m) and does not require battery power. They are smaller, cheaper, and lighter than active RFID (Zafari et al., 2019). Research has demonstrated that RFID can be used for absolute positioning. For example, a study proposed a localization technique for a mobile robot reader using RFID to reduce accumulated errors caused by the robot's movement mechanism (Lee and Lee, 2006). Another research focused on received signal intensity-based reader positioning, using machine learning techniques to overcome the geometric relationship between tags and readers (Yamano et al., 2004). Although RFID offers accurate positioning, high-speed functionality, low cost, and reliability, its major drawback is its limited communication range. This restricts its use to smaller areas and makes it less suitable for many positioning applications. RFID is commonly employed in inventory management, asset tracking, passport information management, shipping, and other similar applications. Compared to other wireless radio technologies, it is less commonly used for localization purposes.

WiFi Technology is a common wireless internet protocol (802.11) for indoor localization due to its widespread availability in mobile devices. The IEEE 802.11 standard is a family of network protocols evolved with different generations of WiFi, consisting of various sub-standards such as IEEE 802.11a/b/g/n/ac/ad/ax to improve the communication bandwidth or speed and IEEE 802.11e/i/v/s/p/ deals with the quality of service, security and encryption, management and configuration, mesh networking and access to the vehicular environments respectively. All these sub-standards operate in the 2.4 or 5 GHz Industrial, Scientific, and Medical (ISM) bands, with 2.4 GHz being the most crowded and 5 GHz less crowded frequency bands. The standards provide different ranges of coverage for both indoor (35-70 m) and outdoor (100-250 m) applications and data speed of 11,54 or 300 Mbps (Abdelrahman et al., 2015; Bhoyar et al., 2013). WiFi can serve as a positioning technology by using various algorithms and parameters, such as proximity, TOF, Channel State Information (CSI), and RSSI -based fingerprinting methods. In the context of WiFi-based indoor localization, fingerprinting and trilateration algorithms are commonly employed. However, it has been observed that the fingerprinting localization method demonstrates superior performance and effectiveness compared to other approaches. A time-based solution is proposed by (Makki et al., 2016) for determining the positioning of a transmitter using multiple and mutually synchronized 802.11g receivers. Although most of the researchers have focused mainly on using IEEE 802.11a/b/g standards for providing location estimation services (Maglogiannis and Hadjiefthymiades, 2007), other standards such as IEEE 802.11n/ac/v/ax can also be used for enhancing localization due to additional features associated with these standards.

The WiFi fingerprinting algorithm is notable for its ability to operate without the need for LOS or time synchronization between devices, as highlighted by (Varshney et al., 2016). It achieves an accuracy of approximately 20-40 m, which can be further enhanced through dense deployment of access points (APs) or by integrating with other technologies (Mainetti et al., 2014). Moreover, CSI which contain fine-grained feature of WiFi channels are also used for fingerprinting and attracted great attention in recent times (Wang et al., 2017).

The Bluetooth Technology (IEEE 802.15.1) is currently attracting much attention for the local position system (LPS), and it is managed by Bluetooth Special Interest Group (SIG). Although the IEEE originally standardized Bluetooth as IEEE 802.15.1, the organization no longer actively maintains this particular standard. The technology is designed for short-range wireless communication with a high data speed of up to 1 Mbps between digital devices such as printers, keyboards, personal computers, mouse. Bluetooth specification version 4.0 adopted two major technologies: classic Bluetooth and Bluetooth Low Energy (BLE). Classic Bluetooth carries legacy Bluetooth protocols, whereas BLE is intended to support low power consumption while maintaining similar cost and range. Currently, Bluetooth version 5.0 is also available with additional features and meets the requirements of Internet of Things (IoT) devices. However, this thesis work is carried out using existing hardware supporting BLE version 4.0. Both WiFi and Bluetooth system operates in the same license-free 2.4 GHz ISM radio frequency band. The sharing of the same frequency band by both the WiFi and Bluetooth technology makes the band very crowded and also makes both technologies vulnerable to interference. To mitigate this problem and to improve performance, Bluetooth uses the Frequency Hopping Spread Spectrum (FHSS) and Time Division Duplexing (TDD) access schemes. The Frequency hopping scheme allows classic Bluetooth to hop between 79 frequency channels in a pseudorandom pattern at 1600 times per second. Further, in the Bluetooth networks, all the devices communicating with each other are synchronized so that they can hop together between the channels (Tabassam et al., 2007). It is observed that WiFi suffers from interference due to Bluetooth hopping sequence from channels 1 to 79. In other words, the narrow bandwidth of the Bluetooth spectrum corrupts wider WiFi bandwidth, as shown in Figure 1.1. So, in order to avoid interference in a coexistence environment, Bluetooth now uses an enhanced technology called Adaptive Frequency Hopping (AFH) which can choose the best frequency for hopping (Pei et al., 2017). For a succinct explanation, AFH permits Bluetooth devices to measure the quality of wireless signal present in the environment and determine whether any channel has been corrupted by interference or due to an intentional attack so that Bluetooth adjusts its hopping pattern and avoids the corrupted channel. Thus, AFH provides extra robustness to the transmitted signals so that any unauthorized receiver cannot identify the channel on which communication is built between two synchronized Bluetooth devices. Because hopping last for only a fraction of a second and immediately changes to other frequency channels in a pseudorandom way.

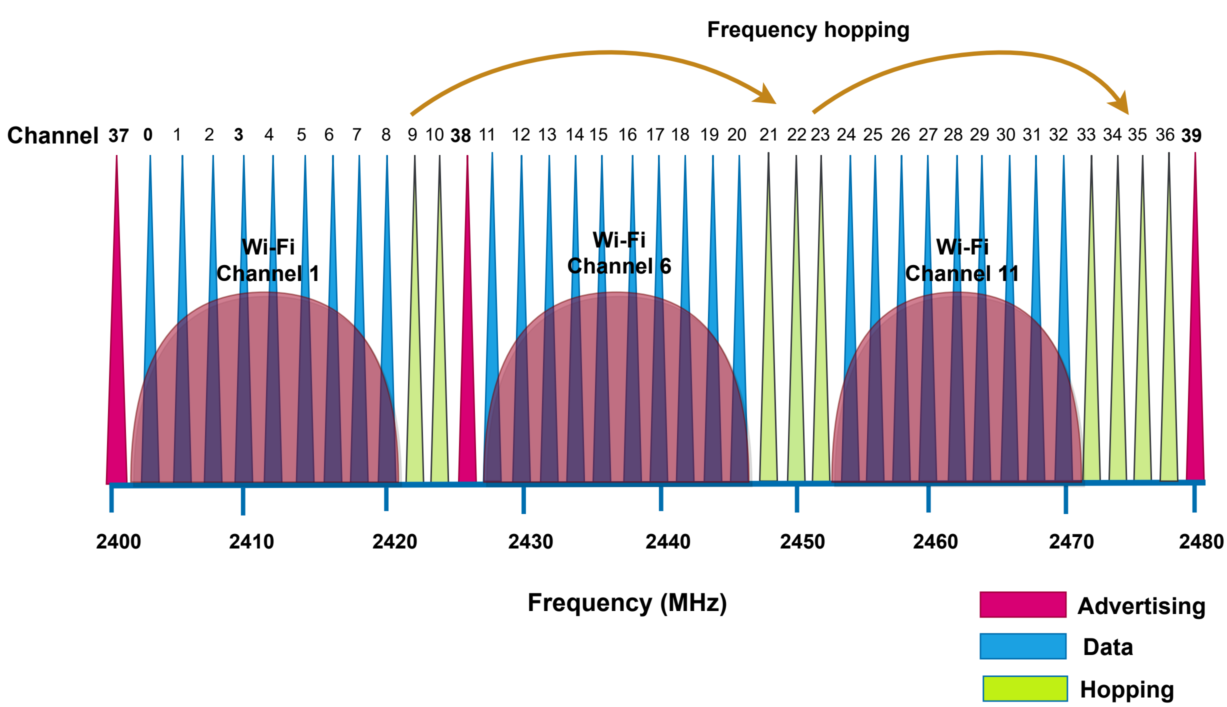


Figure 1.1: WiFi and classic Bluetooth channels in the 2.4 GHz ISM band.

As stated before, BLE 4.0 version is derived from classic Bluetooth technology to have low power consumption and cost that provides longer battery life and supports more applications. This resulted in BLE 4.0 having 40 channels with 2 MHz bandwidth from 79 channels with 1 MHz spacing as in the case of the classic one. The main objective behind this was to minimize the power consumption in BLE by reducing the number of channels to 40 while keeping the total bandwidth the same. More specifics about the Bluetooth technology are discussed in Chapter 2 and Chapter 3.

Ultra-Wideband (UWB) is an emerging wireless radio technology for indoor positioning that uses a wide spectrum of frequency bands to transmit data over short-distances with high bandwidth. Specifically, UWB signals occupy a frequency band that is at least 500 MHz wide or that has a fractional bandwidth greater than 20%. The use of a wide frequency band in UWB technology enables high-speed data transmission, while factors such as low duty cycle, short-range operation, and efficient circuit design contribute to its low power consumption (Fontana, 2004b). Moreover, this wide bandwidth enables UWB signals to be transmitted without interfering with other narrow-band transmissions, such as Television (TV), Global System for Mobile Communication (GSM), Universal Mobile Telecommunications System (UMTS), and GPS signals within the same frequency band (Rahayu et al., 2008b). The technology is designed to use short-duration pulses for short-range transmission mainly to protect against multipath interference and attenuation, enabling the UWB system to achieve centimetre to decimetre-level positioning accuracy (Silva et al., 2014).

Although WiFi and BLE can perform two-way flight ranging between fixed anchors and mobile users (tags), they are not as precise as UWB two-way ranging. This is because, in comparison to BLE and WiFi, UWB radios measure distance using signal travel time between anchor and tag. Distance is then computed by multiplying signal travel time between UWB radios with the speed of light. The distance estimation using very narrow time resolution UWB signals eliminates most of the time biases and provides precise range measurements (Dabove et al., 2018). Tag position is estimated by solving three unknown range measurements using trilateration.

Although still not widely adopted, UWB technology is known for its precision and suitability for indoor positioning. It can measure range in different two-way modes, eliminating the need for precise clock synchronization between UWB radios. Apple recently introduced the U1 chip, which supports ultra-wideband (UWB) ranging in their iPhone 11 series. This U1 chip adheres to the 802.15.4z standard and enables spatial awareness and precise location tracking capabilities (Coppens et al., 2022). The availability of UWB technology in mobile devices such as smartphones has sparked interest among researchers, who are exploring its potential and considering it an appealing and cost-effective option for indoor localization applications.

In recent years, UWB signals have been extensively explored across various technological domains. UWB technology has found diverse applications in communications, radar, and navigation. By leveraging broad RF bandwidth spectrum and UWB techniques, communication networks have successfully adopted UWB for seamless wireless connectivity (Rahayu et al., 2008a). UWB's unique characteristics, including very short-pulse signals, enable radar applications with remarkable features such as highly accurate range measurement, superior range resolution, improved target recognition, immunity to passive interference (rain, fog, clutter, aerosols, etc.), enhanced resilience to co-located radar transmissions, increased detection probability for specific target types, and the ability to detect stationary or slow-moving targets (Fontana, 2004a). Lastly, positioning and ranging represent the third and most prevalent application domain for UWB technology.

The wide bandwidth of UWB signals enables precise range measurements by accurately measuring the time delay between transmitted and received signals, achieving sub-centimetre to sub-millimetre positioning precision (Chong et al., 2007; Mahfouz et al., 2011). Over the past two decades, researchers have conducted extensive experiments on UWB technology for reliable indoor positioning and its applications. In one study, a short-range high-accuracy indoor localization system using UWB technology was proposed (Mahfouz et al., 2008). The research employed advanced receiver hardware design for sampling incoming UWB pulses and detecting the main LOS peak, demonstrating the feasibility of achieving millimetre-level accuracy in highly reflective environments. Zhoue et al (2011) introduced an asynchronous absolute range-based elliptical position system for indoor localization using UWB technology (Zhou et al., 2011). By leveraging the differential time of arrival (TOA) between the direct transmitter and the tag signal, the research showcased the elimination of synchronization requirements and the determination of the absolute range of the target.

Chiu and O'Keefe (2008) evaluated UWB radio range accuracy, revealing that bias and scale factor errors affected the calculation of radio accuracy for various UWB ranging radios (Chiu and O'Keefe, 2008). To determine the distance between two transceivers, Jiang and Leung (2007) implemented an asymmetric double-sided two-way ranging (TWR) method based on the TOF of UWB signals (Jiang and Leung, 2007). Indoor positioning is enhanced by combining measurements from inertial sensors (accelerometers and gyroscopes) with the time of arrival measurements from an ultra-wideband system (Kok et al., 2015). Furthermore, Banerjee et al. (2012) presented a research study on improving indoor positioning tracking in UWB radios through noise modeling and augmented distance measurements (Banerjee, 2012).

Recently, numerous studies have focused on artificial neural network techniques for indoor applications using UWB technologies. A CNN-based algorithm to monitor heartbeat and identify individuals using UWB radar signal simultaneously is shown by (Wu et al., 2019). The study found that heart rate can be estimated by measuring the interval between two adjacent heartbeat patterns. Additionally, researchers have explored the identification of UWB line of sight (LOS) and non-line of sight (NLOS) signals using convolutional neural networks and LSTM methods, followed by using ranging information to calculate position (Jiang et al., 2020). More insights into the UWB technology are presented in Section 2.2 of Chapter 2 and Section 3.4 of Chapter 3, respectively.

1.3.2 Artificial Neural Network (ANN) in Indoor Positioning

An artificial neural network, or just a neural network, thought as a computer system designed to function and make decisions like the human brain. The concept of artificial neural network models is mainly developed from how the brain's biological structure is organized. ANN consists of a network of artificial neurons specifically designed to match the characteristics of brain neurons. These neurons together are capable of executing specific tasks. Over the past decade and more, artificial neural network models are extensively used in many applications such as image classification, signature classification, speech and handwritten recognition,medical diagnosis,and many others. Recently, they have become very handy and attractive tool for positioning applications. ANN are well-suited for finding patterns in complex datasets, which makes them a valuable tool for indoor positioning applications (Rani, 2011). Unlike traditional path-loss propagation models that rely on simplified assumptions about indoor environments which limit their accuracy (Sarkar et al., 2003), ANNs can learn from specific data collected in those environments, leading to more accurate predictions of signal strength. Additionally, ANNs excel at capturing non-linear relationships between input and output variables, making them more effective at modeling complex phenomena compared to models based on simplified assumptions (Ostlin et al., 2010).

Several studies have demonstrated the effectiveness of ANN in indoor positioning applications using UWB and Bluetooth technologies. For example, a Convolutional Neural Network (CNN)-based indoor localization system using received signal strength (RSS) values from WiFi access points in a multi-building, multi-floor scenario (Ibrahim et al., 2018a). The CNN approach achieved remarkable results, boasting a 100% accuracy rate in predicting both building and floor, along with a 2D mean error in coordinates of 2.77 m. In another indoor classification application, CNN model based on RSSI fingerprints outperformed several other deep neural networks such as AlexNet, ResNet, ZFNet, Inception v3, and MobileNet v2 (Sinha and Hwang, 2019). The model achieved a test accuracy of 94.45% with average location error of 1.44 m. A weighted indoor positioning algorithm (WIP) using two neural models is proposed to predict indoor positions from five different UWB signal features (Lu et al., 2021). Another study presents a fingerprinting localization technique that combines feed-forward neural networks with UWB signals (Yu et al., 2011). Moreover, Cheng et al. (2020) proposed a real-time positioning system for smart grid application based on UWB and artificial intelligence techniques (Cheng et al., 2020). Additionally, researchers have improved BLE RSSI-distance models using backpropagation neural networks (BPNN) optimized through particle swarm optimization (PSO-BPNN) to minimize position errors (Li et al., 2018). Further examples of high-performance localization using neural networks can be found in the papers (Hoang et al., 2019; Song et al., 2019; Chen et al., 2020). For more detailed information on Artificial Neural Networks (ANN), refer to Section 2.3 of Chapter 2 and Section 4.3 of Chapter 4, which together provide a comprehensive overview of ANN principles and functioning. Section 4.3 in Chapter 4 explicitly discusses the application of ANN in proximity detection.

1.4 Motivation and Objectives

Accurate range measurement is essential for precise positioning in many applications. Any inaccuracies in range measurement lead to erroneous position solutions, rendering them unreliable for applications that require higher accuracy. Currently, in the literature, the commonly used technique for range estimation from Bluetooth Low Energy (BLE) Received Signal Strength Indicator (RSSI) involves employing a propagation model. However, due to the time-varying characteristics of Bluetooth RSSI and the complex nature of indoor environments, fitting the RSSI distance model with existing radio models poses challenges. Hence, achieving accurate range estimation continues to pose a significant challenge, mainly when applied to proximity detection between two BLE users within indoor environments.

The proposed research uses a limited number of accurate UWB sources as ground truth to train the BLE RSSI to range model, which BLE-only users can later use for determining proximity with reasonable accuracy. A few critical points need to be focused on to meet the primary objective of this research. The key points are as follows:

- Analyze the signal characteristics of the three BLE primary channels' RSSI and the UWB range in line of sight (LOS) from a fixed transmitter to ascertain their behavior.

- Investigate the benefits of using UWB range measurements as ground truth for deriving patterns from the non-linear relationship between BLE RSSI and range instead of using standard radio model in a test scenario.

- Collect BLE advertising channel RSSI values and UWB range data from a new environment as the training dataset for the ANN model.

- Evaluate the reliability of the trained model in determining the proximity of BLE users in novel environments where UWB is not accessible without requiring additional training datasets.

- Furthermore, the study explores the influence of non-line-of-sight (NLOS) and dynamic environments on the ANN model's feasibility to determine proximity and evaluate the variability of both BLE and UWB signals.

1.5 Thesis Outline

A brief overview of the current chapter and the remaining chapters is as follows:

Chapter One covers the problem definition to be investigated and researched during this project. It provides a brief description of the working principles of commonly used indoor positioning technologies, recent developments in AI models in this field, and a relevant literature review to show the current research topic into the right direction. The chapter also describes the motivation and objectives of this research.

Chapter Two provides an overview of the Bluetooth system, Ultra-Wideband system, and Artificial Neural Network (ANN) models primarily used in this work. It presents comprehensive signal characteristics and compares the Bluetooth and UWB systems. Furthermore, it describes the evolution and theory of artificial neural networks in positioning applications.

Chapter Three thoroughly analyzes the RSSI-based propagation model and its limitations in indoor environments. It emphasizes the significance of precise time-based UWB ranging methods and introduces the application of Kalman filtering for processing RSSI and range data. The chapter also showcases the hardware components used in this research. Furthermore, it presents specific results obtained from evaluating BLE RSSI and UWB in a test scenario, which serve as the basis for training the proposed proximity model.

Chapter Four describes the proposed core algorithm for providing BLE RSSI-based distance estimation using neural network models. It begins by explaining the experimental setup, training environment, data collection scheme, and data preprocessing. It then discusses the tuning parameters required to construct a generic model operating efficiently in natural environments.

Chapter Five provides the results of the proposed neural network models for estimating range and proximity using BLE only in real environments, incorporating different scenarios and user conditions. The results of the experiments are then analyzed and explained in detail.

Chapter Six summarizes the results obtained from this research and presents distinct conclusions. The chapter concludes by describing the challenges encountered with the proposed model and providing suggestions for future research work.

Chapter 2

Overview of Systems

This chapter provides an overview of BLE technology, UWB technology, and Artificial Neural Networks. It covers the historical background of BLE technology, its communication protocol, signal characteristics, and a common proximity detection application using this technology. The chapter also presents an introduction to UWB technology, discussing its signal characteristics and applications in indoor positioning. Furthermore, the chapter delves into the fundamental theory of artificial neural networks and explores popular neural network algorithms commonly employed in indoor positioning.

2.1 Bluetooth Technology

Bluetooth Low Energy (BLE) is a wireless radio communication technology and standard developed by the Bluetooth Special Interest Group (SIG) in 2010 to reduce power consumption (Abbas and Yoon, 2015) and promote the Internet of Things (IoT). The standard is designed to transmit data over 40 channels with 2 MHz spacing based on a frequency hopping scheme that avoids channel congestion. After adopting the standard BLE 4.0, BLE is immediately found to be an attractive technology for many indoor positioning applications (Faragher and Harle, 2014; Bai et al., 2020; Giuliano et al., 2020). Several companies, such as Inpixon, Esri, and Mapsted, offer BLE-based indoor positioning solutions commercially for different applications (Indoor Positioning Systems & Location Tracking by Inpixon, 2021; Indoor Positioning System for Indoor Tracking by Esri, 2021; Indoor Positioning & Navigation System-Mapstead, 2021).

As stated before, BLE uses 40 channels, each 2 MHz wide, and classifies the channels into two types: advertising channels (also known as primary channels) and data channels (also known as secondary channels). The primary channels: channel 37 (2402 MHz), channel 38 (2426 MHz), and channel 39 (2480 MHz), are mainly used for discovery purposes. The data channels, i.e., the other 37 channels (channels 0-36), occupy different band frequencies and are only used for bidirectional data transmission. BLE uses Gaussian Frequency Shift Keying (GFSK) modulation, and the transmit power is limited within -20 to +20 dBm (Nikoukar et al., 2018). Figure 2.1 depicts how the BLE channels are arranged in the frequency band, and Table 2.1 refers to the primary channels and their corresponding assigned frequencies. The first primary channel, i.e., channel 37, is centered at 2402 MHz, the second primary channel (channel 38) is centered at 2426 MHz, and the last channel (channel 39) is centered at 2480 MHz. It can be observed that the three primary frequency channels are not sequential, and the two advertising channels (37 and 39) occupied the lowest and highest center frequency channels.

Table 2.1: Primary channel assignments and their frequencies.

BLE Channel Number	Frequency Value (MHz)
37	2402
38	2426
39	2480

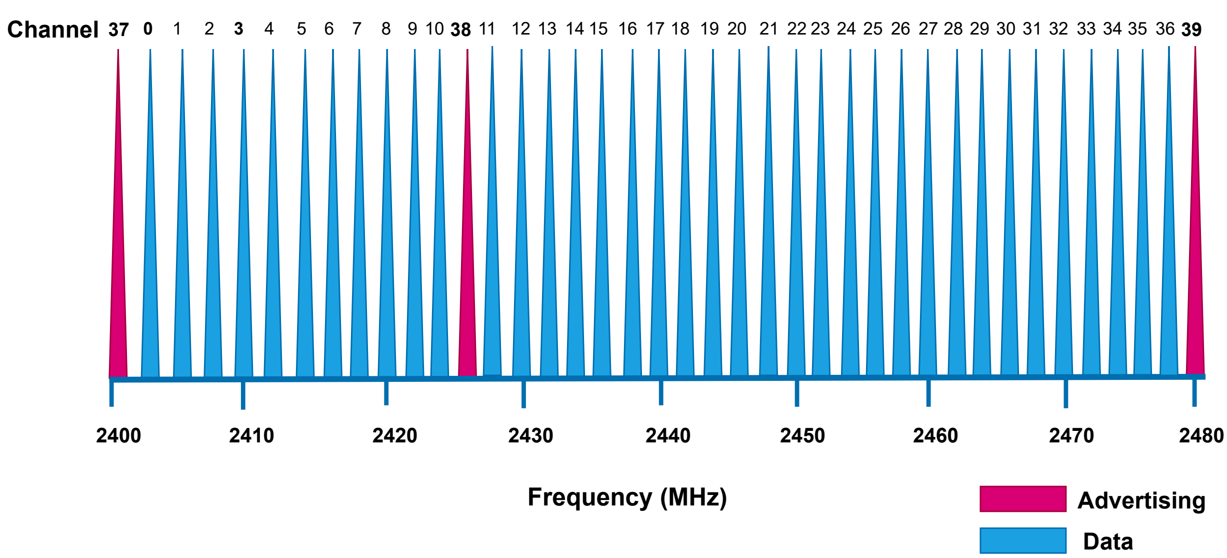


Figure 2.1: The distribution of BLE frequency channels, divided into 37 channels for data transmission (blue) and three channels for advertising (brown).

2.1.1 BLE Communication

Before multiple devices that are compatible with BLE technology can begin communicating, it is essential to establish a secure wireless connection for proximity detection. This mode of communication is commonly referred to as connectionless. BLE-compatible devices are categorized into two types: central or transmitters and peripheral or receivers. Central devices initiate the communication process, while peripheral devices either collect information or participate in the communication. In the context of indoor positioning, these devices are commonly referred to as anchors (transmitters) and tags (receivers), respectively.

A BLE transmitter broadcasts or advertises advertising packets across three channels sequentially, expecting nearby BLE receivers to receive these packets. However, the broadcaster does not know the number of advertising packets the receiver receives. BLE technology was primarily introduced for a wide range of IoT applications, and it has gained significant popularity for indoor positioning applications due to the availability of BLE sensors. It has been extensively used in various areas such as asset tracking, determining a user's location in airports, malls, buildings, museums, healthcare facilities, and proximity services.

Recently, with the emergence of transmissible diseases and the ongoing importance of controlling their spread, BLE gained a new and significant application in determining user proximity to other users in small crowded areas for the purposes of contact tracing. Due to its low cost, energy efficiency, and user-friendly nature, Bluetooth technology has seen an increase in users, making it an ideal choice for this critical application.

Figure 2.2 illustrates an application of BLE technology, where BLE users estimate their proximity to an infected or vulnerable patient in small crowded areas. In this scenario, all users have BLE, while some users may possess additional sensors such as GPS, UWB, or vision sensors. The figure showcases how BLE technology facilitates proximity estimation in line-of-sight conditions, with the assistance of a few expert users. This enables the identification of potential risks and promotes safety measures in such environments. Furthermore, by employing suitable algorithms or models, the application can be extended to support non-line-of-sight scenarios. The significance of expert users and mathematical modeling in enabling BLE users to compute proximity with other BLE users will be discussed in detail in Chapter 4.

The BLE receiver, also known as a scanner, periodically scans for available advertising packets from nearby transmitters. The transmitters broadcast advertising packets at regular intervals on the three primary channels to avoid packet loss, while the receiver collects these packets. However, the advertiser and scanner are not necessarily to be strictly synchronized, and there is uncertainty in the time it takes for a packet to be successfully received (Siva et al., 2019). Therefore, advertising interval, which determines the rate at which advertising packets are sent across the channels, is a crucial parameter for the advertiser.

Similarly, for the scanner, the scanning intervals and the scanning window time decide when the scanner will be activated and start scanning in each scanning time period. Optimizing the advertising interval of the advertisers can aid the scanner in quickly discovering advertisers and reducing energy consumption (Shan and Roh, 2018). It is important to note that scanning intervals and scanning window time have a significant impact on power consumption. The scanner needs to determine the duration it can remain turned on based on these factors. Additionally, the BLE transmitter broadcasts advertising packets at a high transmission power and rate, while the scanner scans one channel at a time and at a lower rate to conserve battery power (Faragher and Harle, 2015).

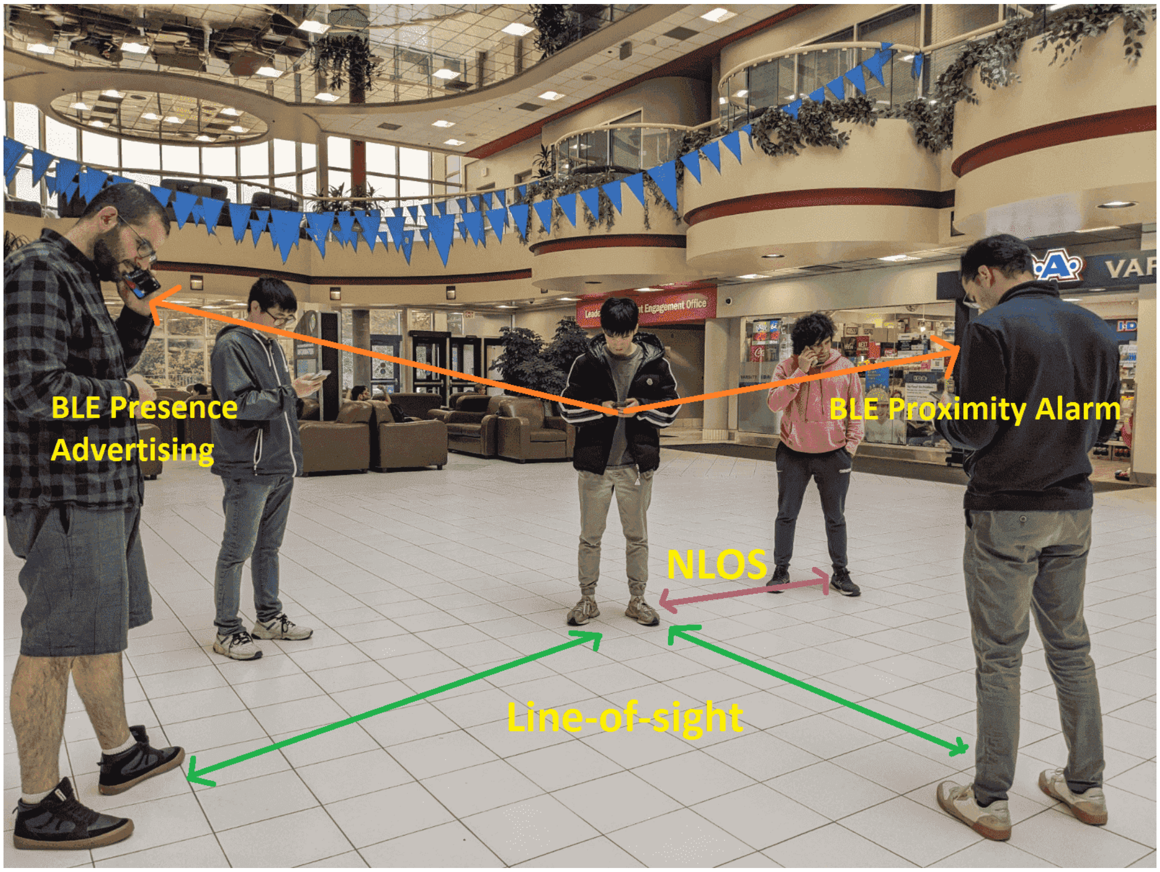


Figure 2.2: Illustrating a common application of a BLE system for detecting proximity in small crowded areas.

2.1.2 BLE Signal Measurements (RSSI)

The term RSSI stands for Received Signal Strength Indicator, which indicates the strength of a transmitted signal when it is received by a device. The RSSI value depends on factors such as the distance between the transmitter and receiver and the transmitted power level. A higher RSSI value indicates a stronger signal, and it is typically represented as a negative integer. Moreover, as the distance from the transmitter increases, the RSSI values tend to exhibit more randomness. The measured signal power at the receiver can be used to estimate the relative distance between the transmitter and receiver. Assuming a constant antenna gain and considering the transmit power (Pt) and received signal power (Pr), the path loss of the radio signal propagating through an environment can be defined as the ratio of transmit power to receiver power. The path loss is defined as

$$\text{Path Loss} = \frac{\text{Transmit Power}}{\text{Receiver Power}} \tag{2.1}$$

RSSI is a representation of the power of the received signal, Pr, as perceived by the user. It is typically measured in the logarithmic scale or in the unit of dBm (decibels relative to one milliwatt).

$$\text{RSSI} = 10\log\left(\frac{\text{Receiver Power}}{1\text{mW}}\right)\text{dBm} \tag{2.2}$$

RSSI-based positioning is the most commonly used technique in major indoor applications. The two most widely used RSSI-based indoor positioning techniques are Trilateration and Fingerprinting. The BLE uses RSSI information from at least three reference nodes and computes a user position using range or range-free measurements. Chapter 3 will provide further details on RSSI-based ranging using propagation models and discuss the positioning algorithms employed in trilateration and fingerprinting techniques.

2.1.3 BLE Primary Channels

In section 2.1.1, use of three BLE primary channels for advertising packets are discussed. When considering indoor positioning, the radio signals transmitted on these channels exhibit distinct propagation characteristics due to their specific frequency bands. Consequently, factors such as path loss, multipath effects, refraction, and fading can cause variations in signal power levels or received signal strength indicators (RSSI). Path loss and multipath effects arise from various elements, including signal obstructions, signal reflections off side walls, and the presence of nearby objects in the propagation medium. It is important to note that these effects can differ among the different BLE primary channels.

Figure 2.3 illustrates the signal characteristics of raw RSSI measurements obtained from individual BLE primary channels and the aggregated channel of a fixed transmitter. The initial 150 RSSI samples from each primary channel are used to compute separate mean and standard deviation. Additionally, 150 samples obtained in aggregate mode, where all three channels are combined, are used to calculate a single mean and standard deviation. The analysis reveals that the aggregated mode exhibits a larger standard deviation than the separate channels. The separate channels demonstrate stable RSSI values with minimal fluctuations, while the aggregated mode yields unreliable values with significant fluctuations. Based on these findings, this thesis employs separate channel measurements to ensure more reliable and stable RSSI values throughout the study.

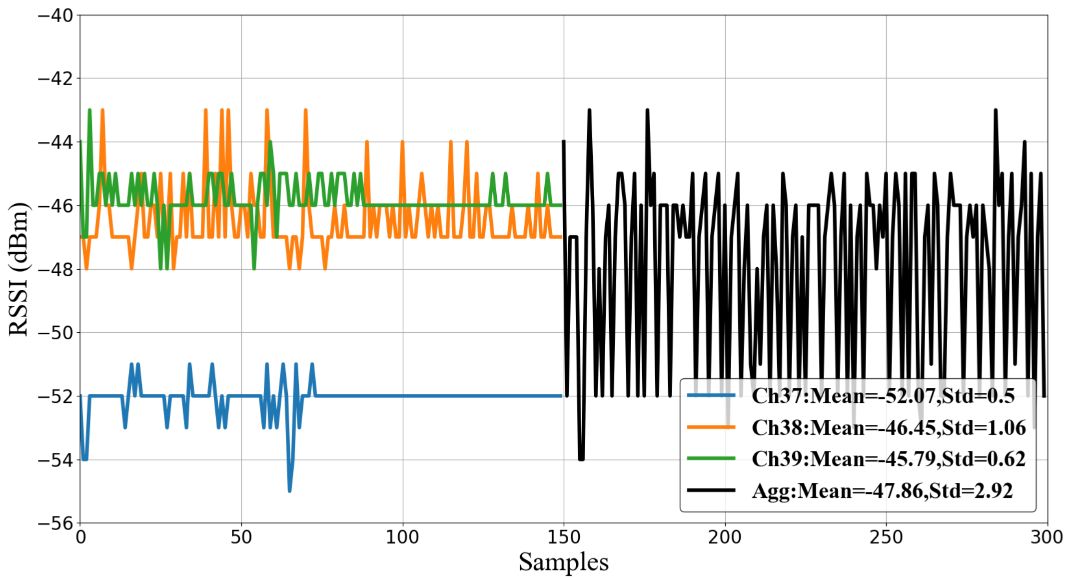


Figure 2.3: Comparison of Received Signal Strength Indicator (RSSI) samples received from the three advertising channels with the aggregate.

2.2 Ultra-Wideband (UWB) Technology

2.2.1 Background

The Ultra-Wideband (UWB) is a recently revolutionizing wireless technology for indoor positioning systems (IPS) (Gezici and Poor, 2009). The fast evolution and milestone of various UWB-based communication applications can be traced back to the late 1960s by the U.S. Department of Defense (Lakkundi, 2006). The technology, however, was restricted to only highly secured military and Department of Defense (DoD) applications until 1990. With the fast advancement of semiconductor technology and the availability of time-hopping (TH) impulse radio, the focus on the commercial use of UWB has started (Win and Scholtz, 2000). In February 2002, the Federal Communications Commission (FCC) approved making the frequency band between 3.1-10.6 GHz available for unlicensed civilian operation of UWB devices, under strict restrictions on the power emission (Win et al., 2009).

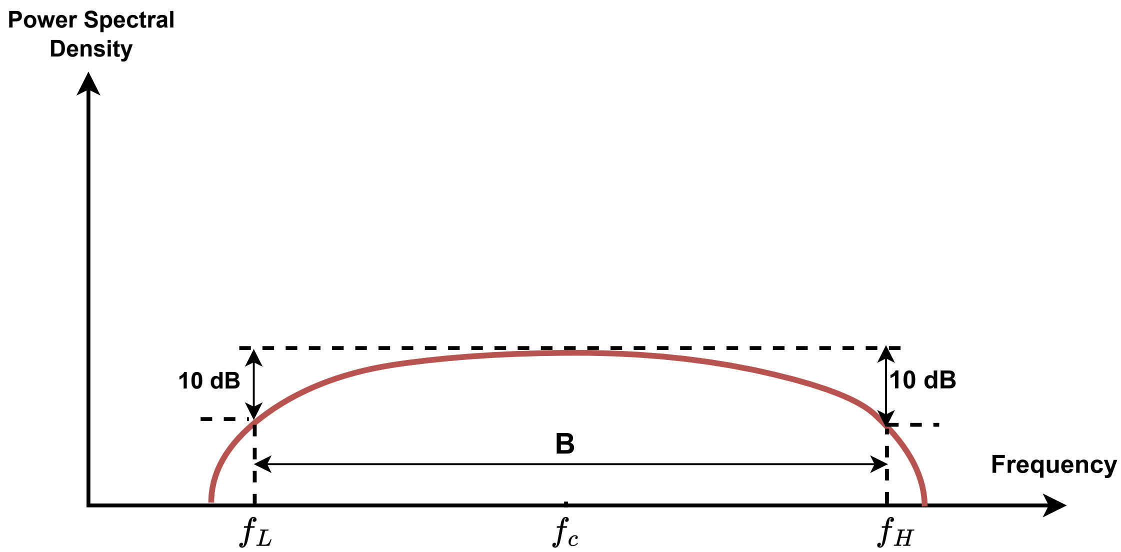


Figure 2.4: A UWB signal with an absolute bandwidth B of at least 500 MHz or a fractional bandwidth B_frac greater than 0.2 (Yang, 2007).

As per FCC, a UWB signal is a signal to have an absolute system bandwidth greater than 500 MHz or a fractional bandwidth (bandwidth divided by centre frequency of band) greater than twenty percent (Federal Communications Commission, 2002).

Figure 2.4 depicts a typical UWB signal with absolute bandwidth B is the difference between upper frequency f_h of -10 dB emission threshold and lower frequency f_l of -10 dB emission threshold. Mathematically, the bandwidth B can be expressed as below

$$\text{B} = f_h - f_l \tag{2.3}$$

The fractional bandwidth B_frac is calculated as

$$B_{frac} = \frac{B}{f_c} \tag{2.4}$$

where f_c is the center frequency and is defined as

$$f_c = \frac{f_h + f_l}{2} \tag{2.5}$$

From Figure 2.4, it is seen that f_c is the frequency which has the highest power spectral density.

Therefore, using Equation 2.3, 2.4 and 2.5, the fractional bandwidth B_frac can be deduced as

$$B_frac = \frac{2(f_h - f_l)}{f_h + f_l} \tag{2.6}$$

Since UWB has large bandwidth, it can interfere with other narrow-band communication systems operating in the same frequency band. Therefore, FCC has set a transmission power threshold for the UWB system to coexist smoothly without causing significant interference. According to FCC regulations, the power spectral density (PSD) should be low as -41.3 dBm/MHz over the 7.5 MHz available spectrum, and it must be even further lower for outside this band, depending on the applications.

The large bandwidth of the UWB system enables it for high data-rate communication. According to Shannon's -Hartley information theorem, in presence of noise and low signal power, the UWB system trade off a portion of its bandwidth for signal power (Chavez-Santiago and Balasingham, 2014). Shannon's theorem states the maximum capacity of information (bits per second) that can be transmitted to a receiver without any error in the presence of noise through a band-limited channel. The theorem expresses this in below mathematical equation

$$\mathrm{C} = \mathrm{B}\log_2(1 + \frac{\mathrm{S}}{\mathrm{N}}) \tag{2.7}$$

Where C is the channel capacity in bits/second (maximum rate of data), B is the bandwidth in Hz available for data transmission, S is the average signal power at receiver, and N refers to the total noise power over bandwidth B (Parker, 2017).

The equation 2.7 can be interpreted in following ways:

- A higher signal-to-noise ratio and a wider bandwidth result in a higher data transmission rate.
- The channel capacity increases more rapidly with an increase in bandwidth compared to the signal-to-noise ratio (SNR).
- The availability of a large bandwidth allows the signal power to be kept at a minimum level, resulting in increased battery life. It also helps minimize interference with other systems (Gezici and Poor, 2009).

- The equation indicates that the channel capacity increases linearly with the increase in bandwidth. On the other hand, the equivalent capacity increases exponentially with the increase in signal power.

- In the presence of noise and low signal power, the bandwidth of the UWB system needs to be traded off to balance the channel capacity. The trade-off improves the system's performance and ensures reliable communication.

Indeed, the equation confirms that UWB technology can achieve high transmission data rates while operating at very low power levels. Figure 2.5 provides a comparison of UWB power spectral density and frequency bands with various commonly used wireless signals.

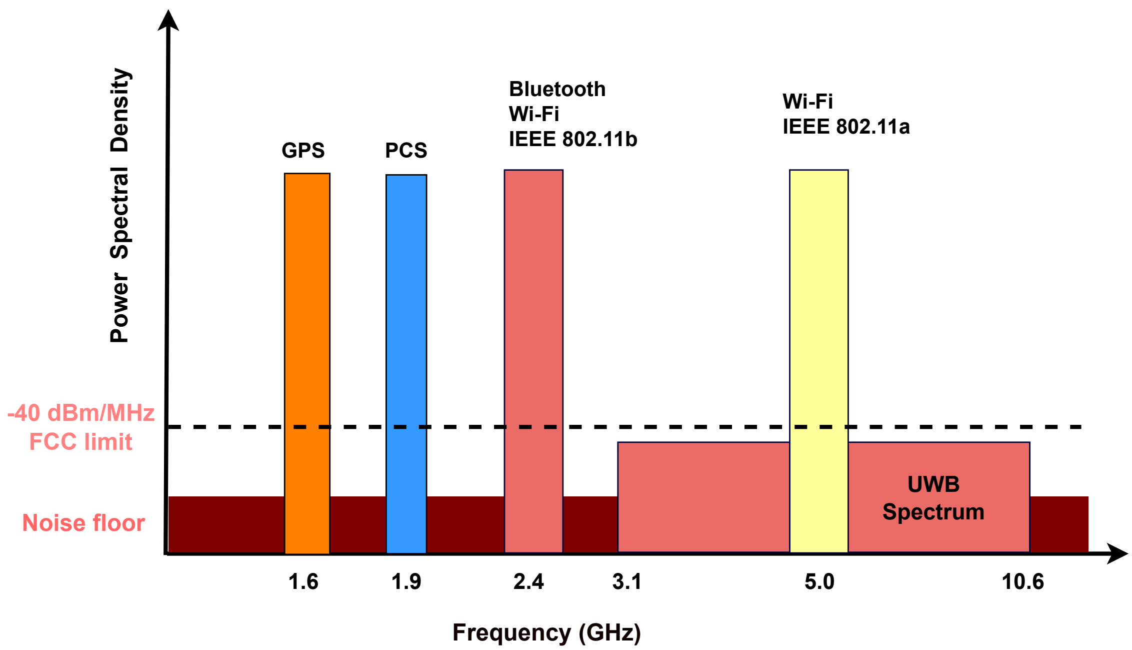


Figure 2.5: Comparison of power density and frequency of UWB signal with various wireless signal (Yadav and Malviya, 2020).

Figure 2.5 shows that the UWB signal occupies a larger bandwidth (3.1-10.6 GHz) compared to other signals. The dashed line represents the power limit for UWB signal transmission set by the FCC. Other indoor positioning systems, such as Bluetooth and WiFi, coexist in the radio frequency spectrum's lower band (2.4 GHz). The Global Positioning System (GPS), a well-known outdoor positioning system, operates in the lowest band of the spectrum (between 1-2 GHz).

Due to its large bandwidth, UWB technology offers several advantages over other indoor positioning techniques. The broad frequency spectrum of UWB enables the generation of short and high-resolution signals in the time domain, resulting in enhanced accuracy for distance estimation. These range estimation methods will be discussed in Chapter 3. The large bandwidth of UWB also enables the system to effectively resolve multipath and interference, ensuring robust performance even in non-line-of-sight (NLOS) scenarios and complex environments (García et al., 2015). In harsh environments, UWB systems exhibit remarkably high accuracy compared to other positioning techniques (Dabove et al., 2018).

.

2.2.2 UWB Signal Characteristics

The bandwidth and time duration of a signal are inversely proportional. Because of the large bandwidth, the UWB system has short-duration pulses (on the order of nanoseconds). The UWB system uses very short-duration pulses with low-duty cycles. In other words, the ratio of the instant of signal transmission to the average time between two consecutive signal transmissions is small.

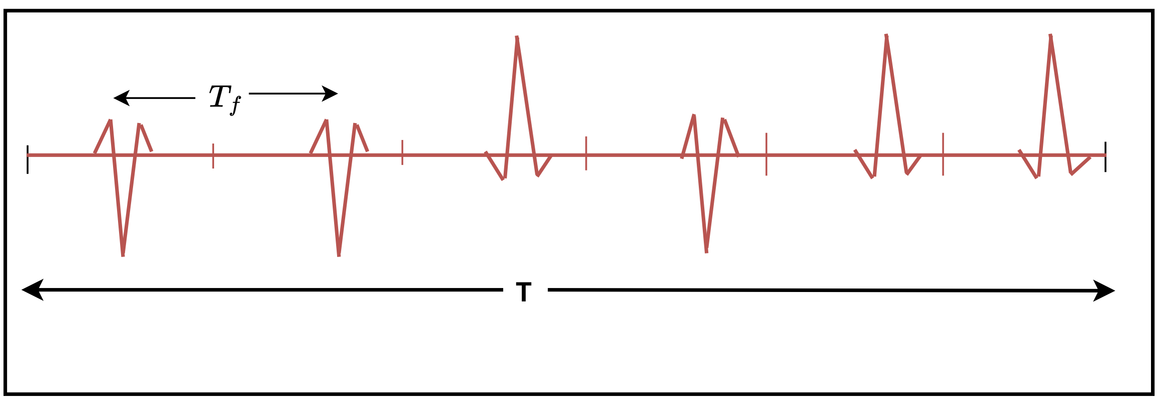


Figure 2.6: UWB signal pulses with low duty cycle.

Figure 2.6 shows an example of a UWB signal consisting of multiple short-duration pulses with a low-duty cycle. T represents the signal duration, and T_f represents the pulse repetition interval. The main advantage of UWB signals is their large bandwidth, which makes them suitable for various applications such as communication, navigation, and radar. In the context of indoor positioning systems, UWB signals offer several important features, which are summarized below:

- The high-time resolution UWB signal enables high short-range accuracy. This means UWB measure distance using time of flight (ToF) information of fine pulses. Moreover, they can be used in synchronous time-of-arrival or asynchronous systems known as two-way time-of-flight ranging (TWR).

- The low-frequency component of UWB signals enables them to penetrate various obstacles, including reinforced concrete building materials, concrete blocks, sheetrock, bricks, wood, plastic, tiles, fiberglass, and even ground or snow (Rovnakova et al., 2008).
- Theoretically, a large bandwidth helps mitigate multipath effects like higher pre-correlation bandwidth in GPS can reduce multipath errors significantly. Similarly, the larger bandwidth of the UWB system strengthens it against jamming and multipath interference.
- From Shannon's expression of channel capacity, the UWB system can have high-speed communication with low power due to its considerable bandwidth advantage.
- Because of the FCC regulations, UWB signal power spectral density is limited to the 7.5 GHz operational spectrum. This condition makes the UWB system avoid interference with other narrow-band communication systems, thus increasing reliability and security.
- The UWB pulses can be transmitted in baseband, i.e., without a carrier, which makes it feasible to implement with simple hardware.

Due to its numerous advantages, including high accuracy, affordable price, and low power consumption, the UWB system is an ideal choice for indoor positioning solutions. Its ability to provide precise position information makes it well-suited for various applications in indoor environments.

2.2.3 UWB Measurements and Applications

As described in the previous section, UWB uses high-resolution time signals between the transmitter and receiver. These high-resolution signals can be used for accurate distance estimation. To illustrate this, Figure 2.7 depicts a histogram plot of range measurements obtained from a UWB source positioned at a fixed location, specifically at a distance of 1 metre from the receiver.

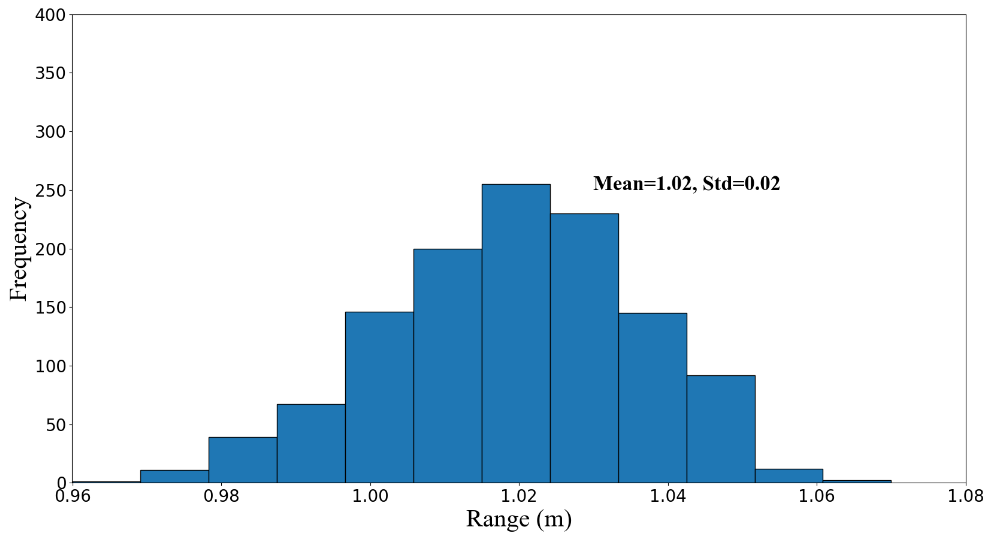


Figure 2.7: Measured UWB range from a fixed transmitter at 1 m distance.

Figure 2.7 demonstrates that the mean value of the range measurements (1.02 m) is very close to the reference ground truth separation (1 m). The small standard deviation of 0.02 m indicates that the UWB measurements exhibit low variability. These measurements were acquired under line-of-sight condition over several minutes. The majority of measurements cluster around the mean value, highlighting the consistency of the observations. Considering the numerous advantages and accurate range measurements offered by the UWB system, this project has chosen to use it as a reliable source of ground truth. Simultaneously, UWB range data is employed in training and calibrating BLE RSSI measurements to enhance BLE accuracy and reliability in estimating range. The Ultra-wideband technology until recently was only for specialized RTLS but now with things like air-tags it is becoming a mass market application. Some of the applications are listed below:

Inventory Management: UWB technology is highly effective in real-time tracking of shipments and goods, making it a popular choice for equipment location in various environments such as hospitals, shopping malls, and warehouses.

Search and Rescue Operations: UWB technology plays a crucial role in search and rescue operations due to its strong penetration capabilities. It is used for locating missing children and smart devices, such as car keys, and provides valuable assistance to emergency responders, miners, firefighters, and other professionals in critical situations.

Security Applications: UWB technology plays a crucial role in ensuring security by enabling the tracking and locating of military personnel in highly sensitive areas. It allows real-time monitoring of their positions during important missions, enhancing situational awareness and security measures.

Medical: UWB system can provide a real-time, seamless tracking option for medical staff or a wandering patient with centimetre accuracy. It enhances patient care, improves workflow efficiency, and provides an added layer of security for medical personnel and patients alike.

Smart Homes: UWB technology can provide high security and easy control of home appliances. For example, the digital front door lock can be controlled through a UWB chip, allowing it to unlock from a distance without tapping or touching it. Similarly, finding digital smart devices such as car keys, smartphones, and many others can be traced and located easily.

2.3 Artificial Neural Network (ANN)

2.3.1 Background

Artificial neural networks, also known as neural networks, have evolved based on the functioning of the human brain. These networks are constructed by connecting a series of neurons or nodes arranged in layers. To understand neural network algorithms for indoor positioning applications, it is first essential to explain a neural network's hierarchical structure. At a broader level, Artificial Intelligence (AI) is defined as the science and engineering behind creating intelligent machines, such as intelligent computer programs (McCarthy et al., 2007). Machine learning (ML) algorithms, a subset of AI, can be described as computational procedures that use input data to accomplish specific tasks without explicit programming. Artificial Neural Networks (ANN) constitute a subset of machine learning, which is, in turn, a subset of Artificial Intelligence. Neural networks are designed to mimic the performance of the human brain. Lastly, Deep Learning (DL), the final category in this hierarchy, is a subset of Artificial Neural Networks (see Figure 2.8).

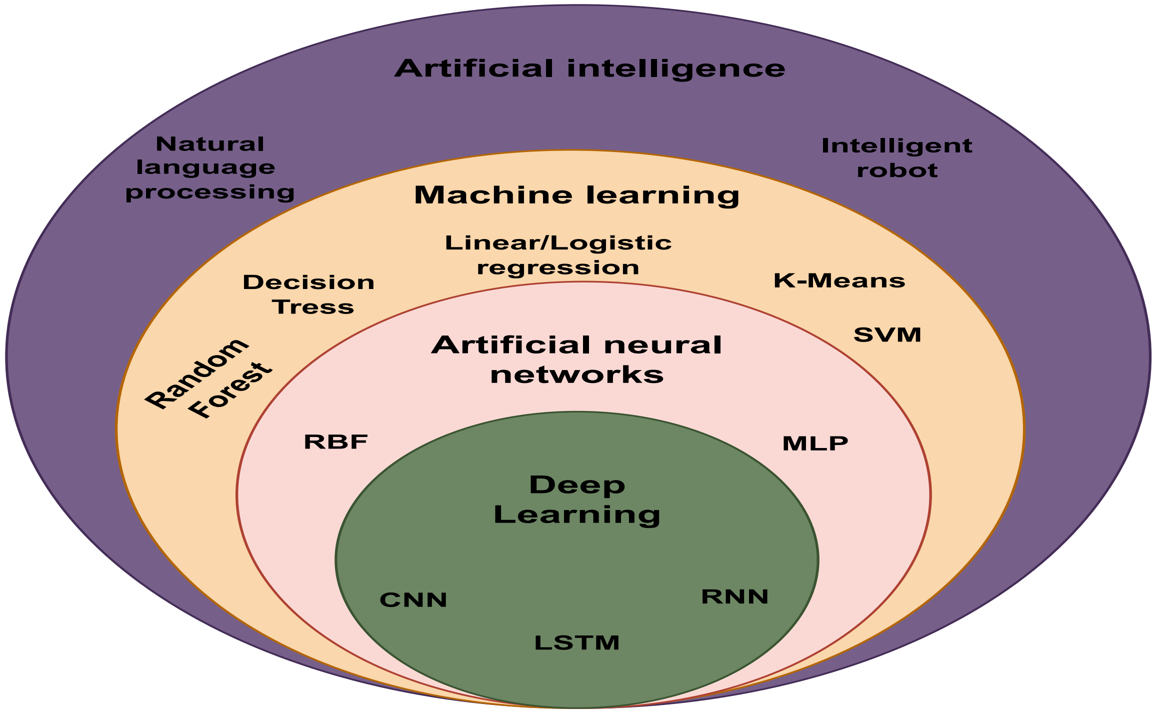


Figure 2.8: Relationship between artificial intelligence, machine learning, neural network, and deep learning (Li et al., 2021).

Deep Learning means deep artificial neural networks with more layers and neurons to train neural network models more efficiently. In other words, neural networks form the backbone for deep learning algorithms. In Deep Learning, the number of layers are more than three, including input and output layers.

With the top-level conceptual description of a basic artificial intelligence system, the underlying principles of artificial neural network become simple. An artificial neural network (ANN) structure consists of three main components: an input layer, one or more hidden layer, and an output layer. The input layer is the first layer that receives signals from outside, and the output layer is the last in the network that produces results for given inputs. In contrast, a hidden layer is a processing layer whose function is hidden from the outside. An architecture of deep neural network with one input layer, three hidden layers and one output layer is illustrated in Figure 2.9. Each neural network layer consists of neurons or nodes always connected to the neurons of at least another layer. Neural networks can be considered as layers of filters in which each filter learn a specific feature from the previous layer and passes its output to the next layer.

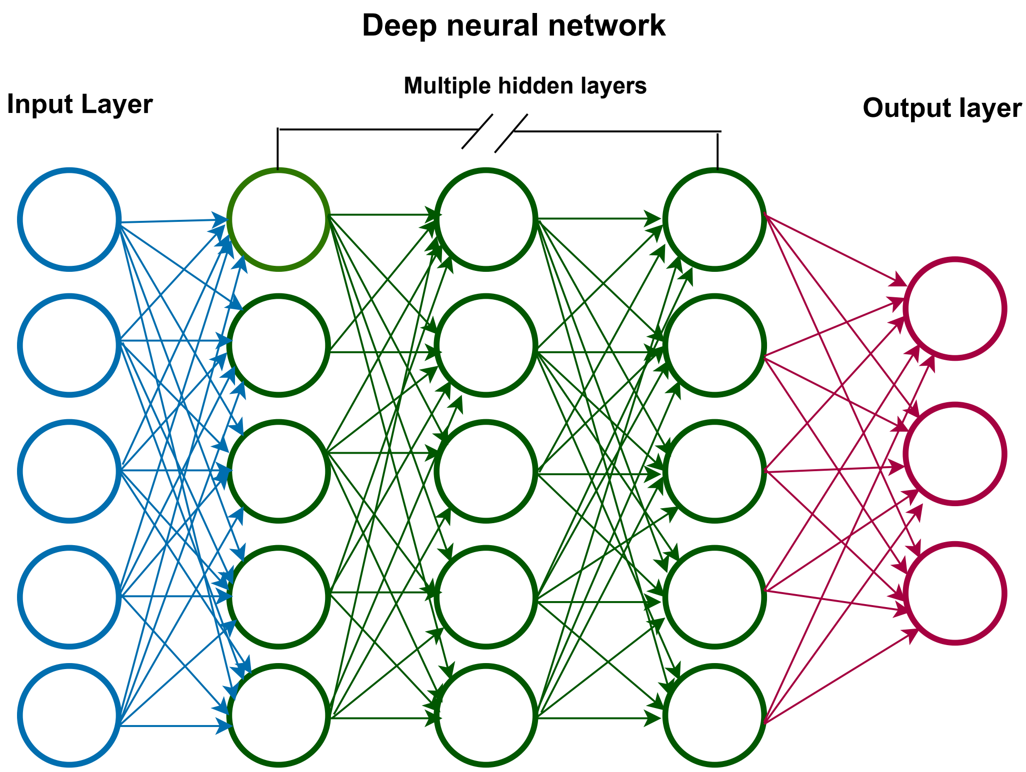


Figure 2.9: An example of deep neural network architecture with one input layer, three hidden layers and one output layer in sequence (Parmar, 2018).

2.3.2 The Artificial Neuron

A neuron (or perceptron), the fundamental component of a neural network, performs four functions to generate output for the next layer neuron. The following mathematical operation represents these functions:

$$\hat{y} = g\left(\sum_{i=0}^{n} x_i w_i + w_0\right) \tag{2.8}$$

where n is the number of input neurons with x_k input values and w_k is weight given to each input values, g(x) is the activation function, w_0 is the bias and $\hat{y}$ is the k^{th} neuron output. Figure 2.10 illustrates the structure of single perception model receiving n inputs and single bias. The weighted sum of inputs plus a bias term is pass through an activation function to produce output.

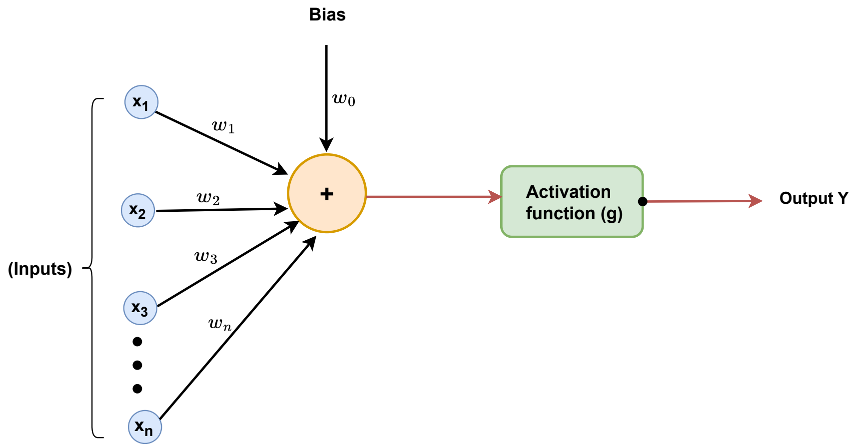


Figure 2.10: Structure of a single perceptron.

The bias is a constant parameter added to the weighted sum of inputs which enables the model to shift activation function towards positive and negative side. This helps the neural network model to fit the input data better and adjust the output.

The activation function (or transfer function) is a mathematical operation to limit the output of a neuron and activate it based on the threshold value of the function (i.e., the neuron is activated if the output value exceeds the threshold). There are two types of activation functions: linear and nonlinear. Linear activation functions, such as the identity function, lead to a neural network that behaves like a linear regression model, which lacks performance. Generally, a neural network model uses nonlinear activation functions to introduce non-linearity into the output of a neuron to estimate weights and biases of input data. The real-world data has nonlinear characteristics, which enables the neural network to learn well in capturing complex patterns and relationships in the data.

In the literature, several non-linear activation functions are commonly used depending upon classification or regression problems. Some of them are discussed here: (i) Logistic Sigmoid , (ii) Hyperbolic Tangent and (iii) Rectified Linear Unit (ReLU). The logistic sigmoid function (Eq. 2.9) maps the input to a value between 0 and 1, providing a smooth S-shape curve as shown in Figure 2.11.

$$\phi(x) = \frac{1}{1 + e^{-x}} \tag{2.9}$$

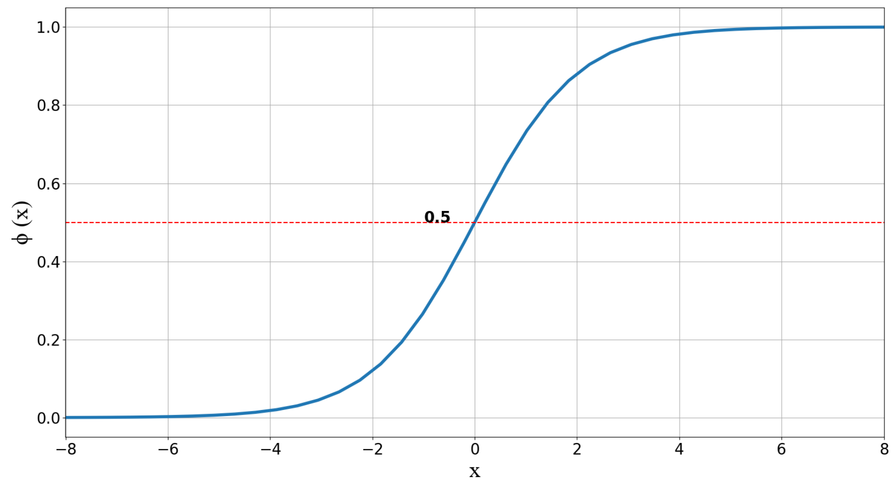


Figure 2.11: Sigmoid activation function.

Similarly, the hyperbolic tangent function (Eq. 2.10) also produces a smooth curve but maps the input to a value between -1 and 1, giving negative outputs as depicted in Figure 2.12.

$$\phi(x) = tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{2.10}$$

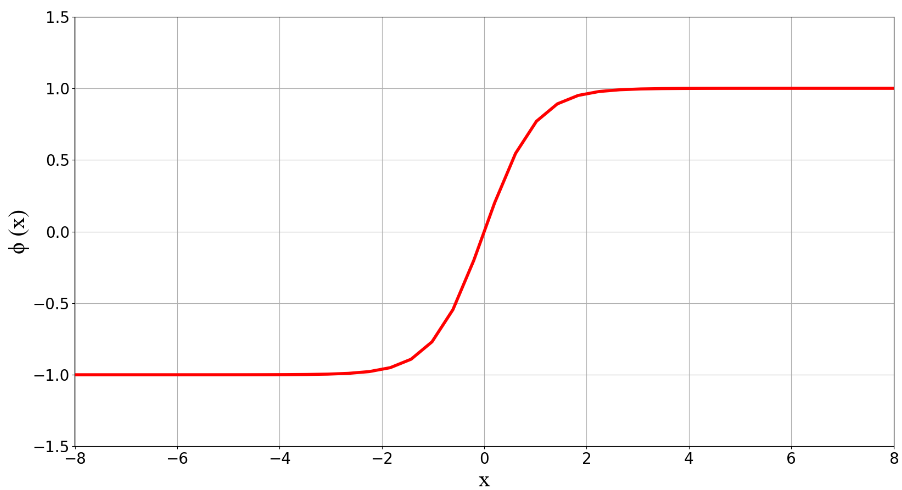


Figure 2.12: Hyperbolic activation function.

The rectified linear unit function (Eq. 2.11) returns the input as the output if it is positive, and 0 otherwise. as shown in Figure 2.13. It offers advantages over sigmoid and tanh functions by preventing saturation, allowing for efficient computation, inducing sparse activation, and addresses the vanishing gradient problem in deep networks. These properties make ReLU to facilitate better representation of learning and use in various deep learning architectures.

$$\phi(x) = f(x) = \max(0, x) \tag{2.11}$$

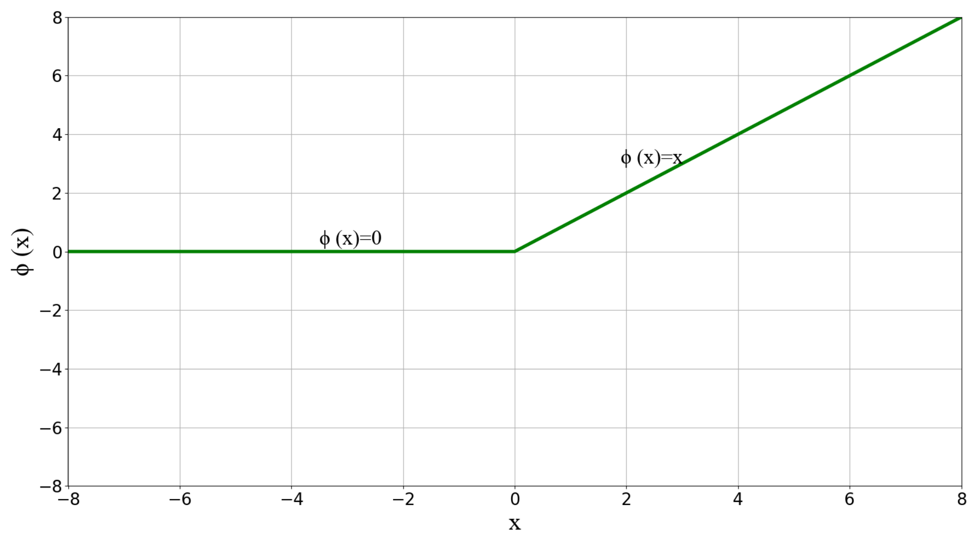


Figure 2.13: ReLU activation function.

2.3.3 Neural Network Learning Process

Artificial neural networks are a machine learning approach that requires training to effectively learn from the provided data, also known as training data, and perform well when presented with new unseen data. There are mainly three ways a neural network can learn and make decisions:

Supervised Learning: As the name suggests, the neural networks learn about the data under supervision. In supervised learning, the neural network is trained with labeled data which means there is already a known output for each input. Supervised learning can be used for both classification and regression problems.

Classificationis the process of predicting the class or category of input data. For instance, determining whether an email is spam is a binary classification problem. In binary classification, the goal is to classify instances into one of two classes, such as "spam" or "not spam.

Regression is the process of predicting continuous output value based on the input data. A typical example of a regression problem is predicting the prices of a house given the input features of the house, such as size, location, builder, and security. Some of the commonly used supervised learning algorithms are Linear regression, Logistic regression, Support Vector Machine, K Nearest Neighbors, Decision Trees, Random Forest etc.

Unsupervised Learning: In the unsupervised learning technique, the training data is unlabeled, meaning there is no known output for the input data. The model can learn and discover information by understanding the pattern and trend in the data. Examples of unsupervised learning algorithms are Principal Component Analysis (PCA), K-Means Clustering, Hierarchical Clustering, etc.

Reinforcement Learning: Reinforcement learning involves training a neural network model to follow a trial-and-error approach to get the desired solution. It uses an agent and environment to produce action and rewards. After accomplishing a task, the agent receives a reward. Common reinforcement learning algorithms include Q-Learning, Monte Carlo, Deep Q Network, and others.

2.3.4 Neural Network Architecture and Models

In a neural network, information can flow in two ways. These are described below:

Feed-Forward Networks: In this architecture, the data fed from input layer (left) travels in one direction towards the output layer (right). Feed-Forward networks have one input layer, one output layer and can have zero or many hidden layers. Feed-forward neural networks sometimes called as Multilayer Perceptron (MLP) model or just Artificial Neural Network. This architecture is widely used in classification and regression problems.

Feedback Networks: Feedback network architecture allows data to travel in both directions through the hidden layer loops in the network. The network have internal state (memory) to process and remember past sequence of data. Feedback networks are dynamic as internal state changes frequently and get complex. This architecture addresses mainly time-series and sequential problems. There are five common types of neural networks models which are applied in various application.

- Feed Forward Neural Network
- Multilayer Perceptron (MLP)
- Convolutional Neural Network (CNN)
- Radial Basis Functional (RBF)
- Recurrent Neural Network (RNN)

2.3.5 Neural Network Model and Machine Learning Algorithms

The Convolutional Neural Network (CNN), a deep learning model, and Random Forest (RF), a classical machine learning algorithm, have been selected in this work. CNN and RF have demonstrated remarkable power and effectiveness as signal processing tools in numerous indoor localization applications (Ibrahim et al., 2018b; Hsieh et al., 2019; Jedari et al., 2015; Bai et al., 2020). For the sake of simplicity, both CNN and RF is referred to as Artificial Neural Network (ANN) techniques in the following sections and chapters. The following sections exclusively focus on the theory behind these two models.

2.3.5.1 Random Forest (RF) Algorithm

Random Forest is a classical supervised machine learning algorithm. The forest it creates is an ensemble or group of many decision trees trained using the bagging method. A decision tree is a top-down flow-chart structure that displays all possible outputs for a given input data (Figure 2.14). Using a single decision tree constructed on the entire dataset with all features can result in inadequate predictions and drawbacks, including overfitting and reduced computational efficiency. On the other hand, random forest combines multiple decision trees to take advantage of their strengths. The bagging method in random forest involves randomly sampling subsets of the training dataset with replacement, fitting a model to each decision tree using the smaller datasets, and aggregating their predictions. This approach uses a collection of models for making predictions instead of relying on a single individual model. In other words, random forest constructs multiple decision trees and combines them to achieve more accurate predictions. The structure of the random forest is illustrated in figure 2.15.

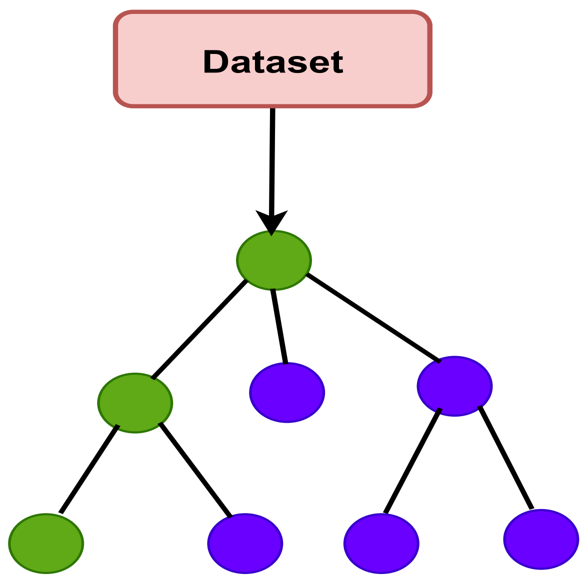


Figure 2.14: Illustration of the Decision Tree structure.

The advantages of the random forest lie in its versatility in handling classification and regression problems. In classification, the final output is determined through majority support or voting among the decision tree outputs. Conversely, for regression, the output of each decision tree is averaged to provide the final prediction.

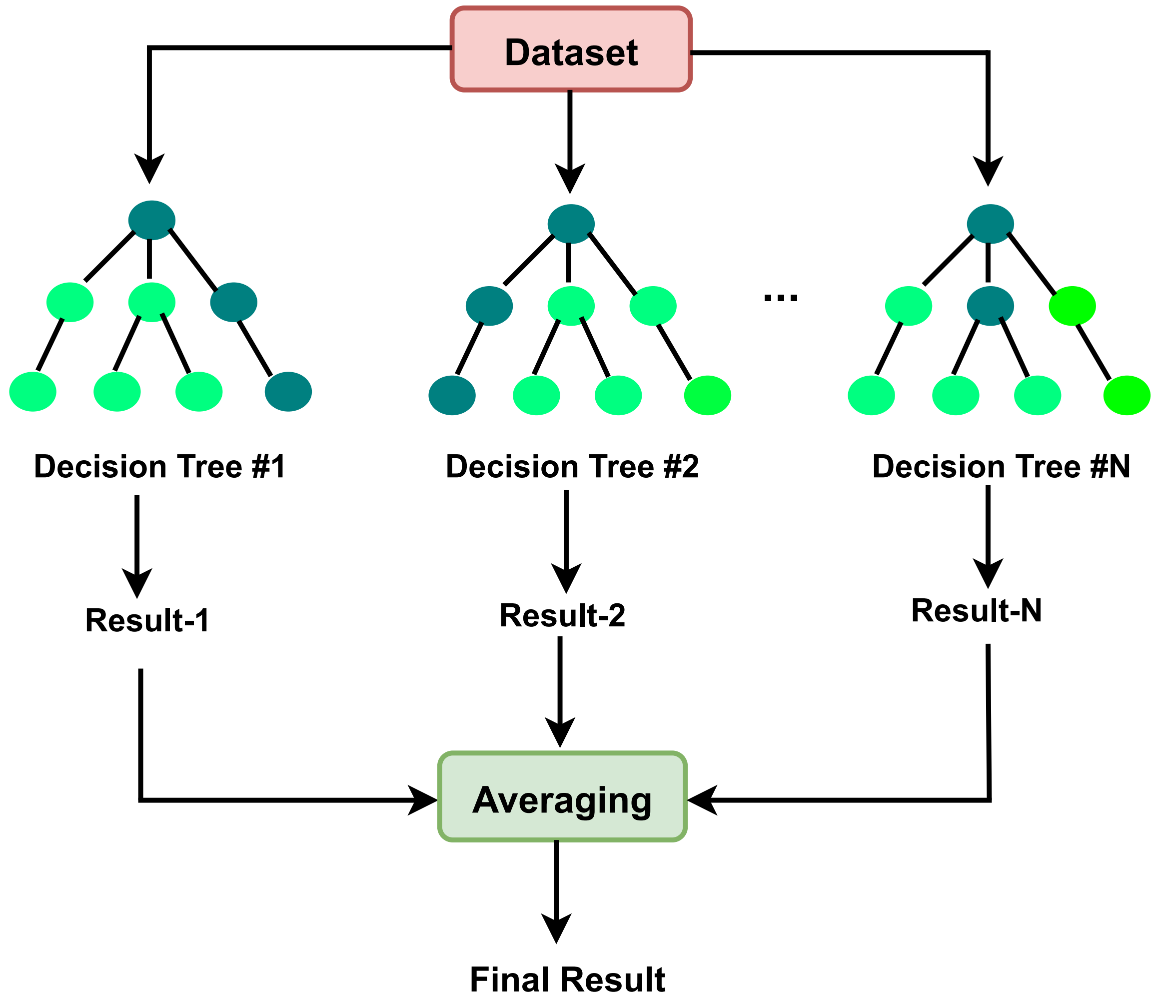


Figure 2.15: Architecture of the Random Forest algorithm for regression-based range estimation using BLE RSSI and UWB values (Brital, 2021).

Each decision tree in the random forest works on a random subset of data derived from the original training dataset. As a result, each tree is different, and not all attributes are available during the creation of an individual tree. Having separate random features for each decision tree introduces randomness to the model. The model searches for the best feature that can influence a maximum number of decision trees while splitting a node, thus creating diversity and building an efficient model. Each decision tree consists of three types of nodes: the root, leaf, and decision nodes. The root node, located at the top left with a random dataset, serves as the starting point. Leaf nodes are the final nodes without child nodes, where the final decisions are made. Decision nodes are where the splitting occurs based on certain conditions. An important parameter used in tree splitting is entropy or information gain. Entropy represents the uncertainty or randomness in the data, with higher values indicating greater randomness. Information gain, on the other hand, measures the difference in entropy before and after a split and is used to make decisions. The total entropy of the sub-branches must be lower than the entropy of the parent node for further splits to occur.

2.3.5.2 Convolutional Neural Network (CNN) Model

The Convolutional Neural Network (CNN) is a supervised deep neural network model primarily used for image recognition and computer vision applications. It is specifically designed to process data with a grid pattern. A typical CNN consists of three main layers: (i) a convolutional layer, (ii) a pooling layer, and (iii) a fully connected layer. The complete structure of the Convolutional Neural Network is depicted in Figure 2.17 (Balaji, 2020). The convolutional layer is a fundamental building block of CNNs and differs from the fully connected hidden layers in traditional artificial neural networks. In CNNs, the neurons in the convolutional layer are not connected to all output neurons. Instead, each neuron in a convolutional layer is associated with a specific region of the input data, allowing for local feature detection.

A convolutional layer in CNNs contains a set of filters called kernels or feature detectors. These filters perform convolution, which involves sliding the filter over the image data and computing the dot product at each position to generate a feature map. The filter size is typically smaller than the input data size, and multiple filtering operations are applied to cover the entire image data. To illustrate, consider an example of convolution with an input image size of 5x5 and a filter size 3x3, as shown in Figure 2.16.

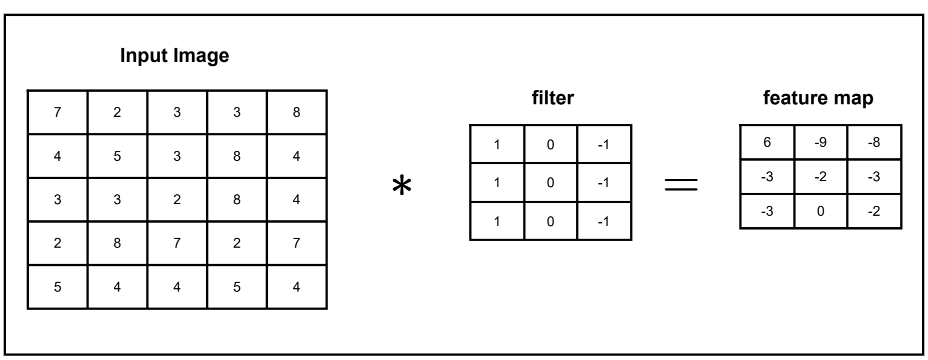


Figure 2.16: An example of convolution.

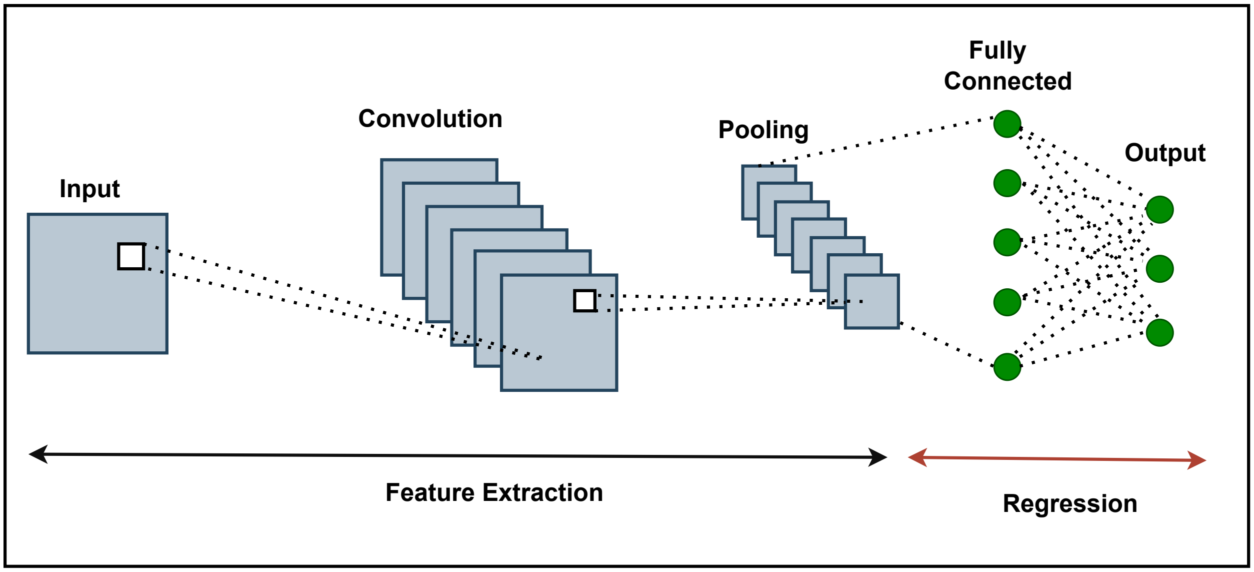


Figure 2.17: Architecture of the Convolutional Neural Network algorithm for the regression problem of range estimation using BLE RSSI and UWB values (Balaji, 2020).

After creating the feature map, each value within it undergoes a non-linear activation function, such as the Rectified Linear Unit (ReLU), commonly used in CNNs. ReLU is a lightweight and computationally efficient activation function compared to others. The advantage of ReLU is its ability to output zero for negative input values while maintaining linearity for positive inputs (see Figure 2.13). By only activating neurons for positive inputs and deactivating neurons for negative inputs, ReLU helps prevent overfitting on the training data. This selective activation contributes to the regularization of the network.

After the convolution and ReLU operations, pooling layers are typically applied in CNNs. The primary purpose of using pooling layers immediately after the convolutional layer is to reduce the dimensionality or size of the feature maps. This dimensionality reduction helps decrease the number of trainable parameters and improves computational efficiency without significant loss of extracted features. There are two commonly used pooling functions in CNNs: Max Pooling and Average Pooling. In Max Pooling, the operator selects the maximum value within the filter's coverage area on the feature map. That means the most prominent or essential features are retained from the previous feature map layer. On the other hand, Average Pooling selects the average or mean value of the feature map within the filter's coverage area. It is worth noting that the features obtained through Max Pooling tend to be superior to those obtained through Average Pooling, as they capture the most salient features (Li et al., 2019).

In a CNN, multiple layers of convolution and pooling can exist between the input image data and the fully connected layer at the output. These layers reduce the data size while retaining meaningful information by reducing noise at each layer. That helps address the problem of overfitting and improves computational speed. The last two stages of a CNN consist of flattening and fully connected layers. Since the pooled feature map at this stage is two-dimensional, flattening converts it into a single long vector or one-dimensional array. This allows it to be used as input for the subsequent layer. Unlike the neurons in the convolutional layers, the neurons in the fully connected layer are connected to all neurons in the preceding and succeeding layers. Finally, the fully connected layer is connected to the number of neurons required for the classification task. As a result, the CNN model can be broadly divided into two parts: (i) feature extraction, which is performed by the left portion of Figure 2.17, and (ii) classification or regression, which includes a fully connected layer at the output shown in the right portion of the network.

2.4 Summary

This chapter provides a comprehensive overview of BLE and UWB technologies, delving into various categories of neural network architectures and their learning processes. It focuses on two specific ANN algorithms (RF and CNN) from a signal-processing standpoint in the context of this coursework. In Chapter 3, a more detailed examination will be conducted on BLE RSSI-based and UWB range positioning techniques, accompanied by the presentation of the system model architecture for general BLE users. Additionally, in Chapter 4, the training procedures of the two neural networks: random forest and convolutional neural network (CNN) will be discussed.

.

Chapter 3

Bluetooth and Ultra-Wideband Ranging Methods & System Design

3.1 Background

Chapter 2 presents a high-level overview of Bluetooth, UWB, and ANN systems. It explores the signal characteristics of Bluetooth and UWB systems, highlighting the differences in signal components that affect range estimation methods. This chapter presents the standard approaches for range estimation using Bluetooth RSSI and various time-based UWB ranging techniques. Furthermore, it describes the hardware components used as both BLE and UWB sources and receivers. Then the chapter provides a test example scenario that demonstrates the advantages of using UWB as ground truth and training BLE using both BLE RSSI and UWB range measurements. The intricacies of this test example scenario are thoroughly explained. Finally, leveraging a comprehensive analysis of the results, the architecture for the BLE proximity detection model is designed.

Radio frequency technologies provide user localization information based on three distinct features of the radio signals: (i) the power level of the transmitted signal, (ii) the propagation time, and (iii) the direction of the transmitted signal. Among them, the Received Signal Strength Indicator (RSSI)-based method, which indicate the power level of the transmitter is more common in Bluetooth positioning. The BLE RSSI values are converted to distance using one of the standard propagation models (Section 3.2).

Two commonly employed indoor positioning algorithms, trilateration (range-based) and fingerprinting (range-free), are derived using RSSI measurements.

3.2 Path-loss propagation model

The characteristics of radio signals, also known as electromagnetic waves, cause them to spread and decrease power density as they propagate away from the transmitter. This decrease follows the inverse square law (Eq. 3.1). In real-world scenarios, radio signals spread out and weaken as they interact with and pass through objects. This reduction in signal power due to object interaction is called path loss. Path loss is defined as the ratio of the transmitted power to received power, typically expressed in decibels (dB) (Eq. 2.1). Various factors, including transmitter power level, transmitter and receiver antenna gains, signal frequency, and the distance between antennas, influence the received power level. In simpler terms, the received power level and path loss can be expressed as functions of distance, aiding in the estimation of the distance between the transmitter and the receiver. In an environment without obstacles, when a radio signal travels from a transmitter to a receiver over a line-of-sight distance d, the received power P_r is determined by the transmitted power P_t. The relation between P_r and P_t can be derived using Friis Free Space Propagation model (FFSPM) as shown below

$$P_r = P_t G_t G_r \left(\frac{\lambda}{4\pi d}\right)^2 \tag{3.1}$$

where

G_t and G_r are gain of transmitting and receiving antennas.

λ is the wavelength of signal.

d is the distance between transmitter and receiver.

G_t and G_r are dimensionless quantities

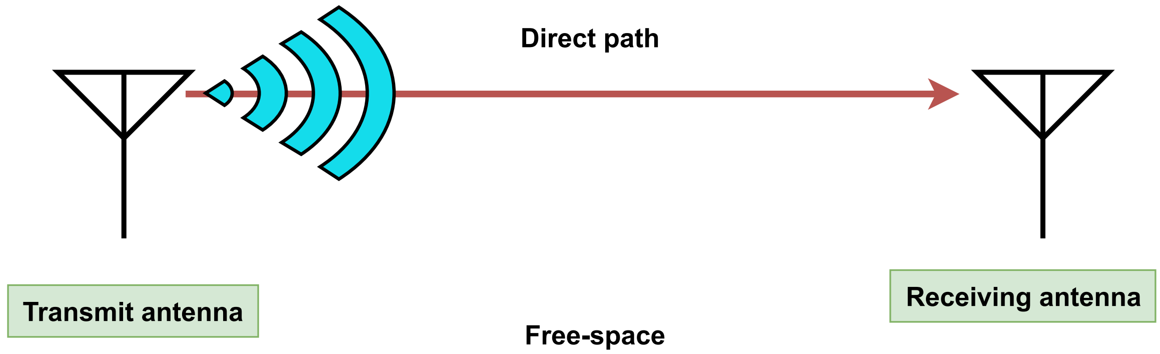


Figure 3.1: Free space propagation model of radio signal without obstruction.

Equation 3.1 can be rearranged to find distance between transmitter and receiver:

$$d = \frac{c}{4\pi f}\sqrt{\frac{P_t}{P_r}G_tG_r} \tag{3.2}$$

where wavelength λ is replaced with ratio of velocity of light c and signal frequency f.

As described earlier received power level depends on path loss which is the ratio of transmit power and received power. Therefore, rearranging Eq. 3.1 the loss can be shown as:

$$\text{Path loss} = \frac{P_t}{P_r} = \frac{1}{G_tG_r\left(\frac{\lambda}{4\pi d}\right)^2} \tag{3.3}$$

Applying logarithmic scale on both sides of Equation 3.3 gives

$$10log_{10}\left(\frac{P_t}{P_r}\right) = 10log_{10}\left(\frac{1}{G_tG_r\left(\frac{\lambda}{4\pi d}\right)^2}\right) \tag{3.4}$$

$$\text{Path loss (dB)} = 10log_{10}\left(\frac{P_t}{P_r}\right) = -10log_{10}[G_tG_r\left(\frac{\lambda}{4\pi d}\right)^2] \tag{3.5}$$

$$\text{Path loss (dB)} = -10log_{10}\left(G_tG_r\right) - 10log_{10}\left(\frac{\lambda}{4\pi d}\right)^2 \tag{3.6}$$

$$\text{Path loss (dB)} = -10log_{10}\left(G_tG_r\right) + 20log_{10}\left(\frac{4\pi d}{\lambda}\right) \tag{3.7}$$

Let assume that antenna gains are constant and equal to unity, then above equation reduces to Equation 3.8

$$\text{Path loss (dB)} = 20log_{10}\left(\frac{4\pi d}{\lambda}\right) \tag{3.8}$$

$$\text{Path loss (dB)} = 10log_{10}\left(\frac{4\pi d}{\lambda}\right)^2 \tag{3.9}$$

In real-world scenarios, the path loss experienced by radio signals can vary due to reflection, multipath fading, interference, and diffraction. Consequently, the power term of 2 in Equation 3.9 is replaced with a constant denoted as n. This constant, referred to as the path loss exponent or propagation constant, represents the extent of signal power loss along the propagation path in a specific environment. Thus, Equation 3.9 is modified to its new form as shown in equation 3.10.

$$\text{Path loss}(dB)_{indoor} = 10log_{10}\left(\frac{4\pi d}{\lambda}\right)^n \tag{3.10}$$

Applying log on both sides of above equation transformed to equation 3.11

$$\text{Path loss}(dB)_{indoor} = 10nlog_{10}\left(d\right) + 10nlog_{10}\left(\frac{4\pi f}{c}\right) \tag{3.11}$$

Both velocity of light and π are constant values,however, frequency of Bluetooth signal varies between 40 different channel frequencies (between 2.4-2.4835 GHz). Therefore, second part of the above equation is substituted with constant value denoted by A (also called as measured power).

$$\text{Path loss}(d)_{indoor} = 10nlog_{10}\left(d\right) + A + \chi_{\sigma} \tag{3.12}$$

where χ_{σ} is a zero mean Gaussian distributed random variable with standard deviation σ showing shadowing effect or observation error.

In indoor navigation and localization application, the received signal strength indicator (RSSI) at a distance d is as follows:

$$\text{RSSI(d)} = P_t - \text{Path loss}(d)_{indoor} \tag{3.13}$$

where P_t is the actual signal transmission power.

Also, A can be defined as the averaged received signal strength indicator (RSSI) values of receiver at a reference distance of 1 metre (d_0) from the transmitter.

$$\text{A} = \text{RSSI}(d_0) = P_t - \text{Path loss}(d0)_{indoor} \tag{3.14}$$

From Equations 3.12, 3.13, and 3.14, the simplified obtained model is known as standard log-distance path loss model.

$$\text{RSSI(d)} = \text{RSSI}(d_0) - 10nlog_{10}\,(d) + \chi_\sigma \tag{3.15}$$

The path loss exponent n has a value of 2 in free space. However, in environments with higher attenuation levels, the value of n tends to be greater, whereas in wave guides, it is typically less than 2. Estimating the path loss exponent n and the reference signal strength indicator $RSSI(d0)$ can be achieved by fitting a line to a set of measurements or using standard values as references. Since the standard log-distance path loss model depends on environment-specific factor n while additional effort is required to precisely determine reference $RSSI(d0)$ value, it is challenging to identify the correct RSSI-distance relationship using this model. This thesis proposes a new technique and overcomes the shortcoming of the standard log-distance model in range estimation on three BLE primary channels.

3.3 RSSI-based Positioning Techniques

Different types of positioning solutions can be achieved using the Received Signal Strength Indicator (RSSI) values obtained at the receiver. In indoor localization applications three different RSSI-based positioning methods are commonly observed: (i) Proximity, (ii) Fingerprinting, and (iii) Trilateration. Each of these positioning techniques will be briefly discussed in the following sub-sections.

3.3.1 Proximity

Proximity positioning is one of the most straightforward RSSI-based localization techniques, which determines if an access point (AP) such as WiFi can be connected to a mobile user so that the user position is located in the proximal region of the access point. The region can be considered a circular area, and a user is detected using threshold approaches as shown as example in Figure 3.2.

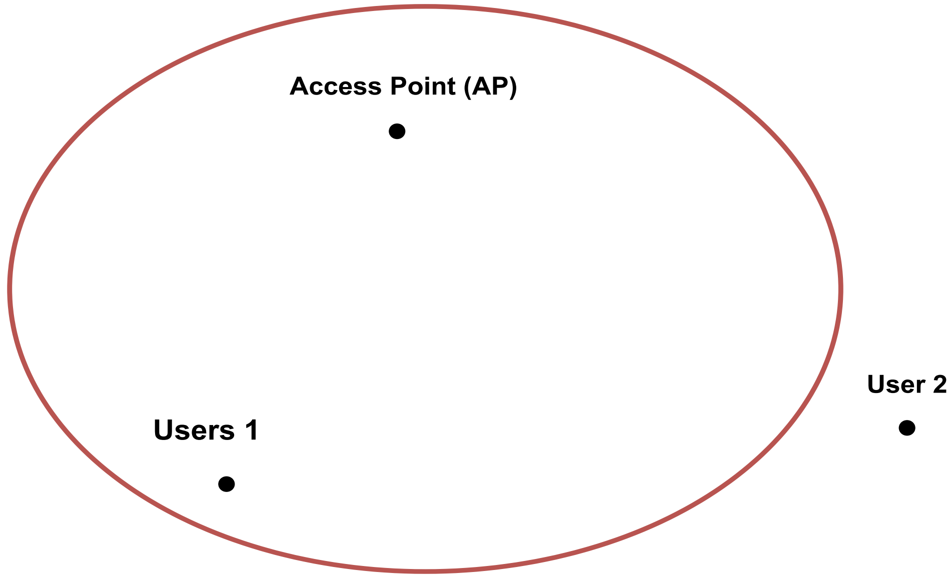


Figure 3.2: Proximity positioning technique.

The circular area specifies a proximity zone of an access point. The two mobile users, user one and user two, are monitored whether they are in the vicinity of the access point. It is sensed by the access point that user one is in its vicinity and user two is unidentified. That means the proximity localization technique can only provide a rough estimation of user position and not give the user absolute positioning information. In addition, user two were situated farther away from the access point than user one, suggesting that this technique is used for only short-range communication. An approach to improve the accuracy of the BLE proximity estimation using Bayesian filtering was proposed by (Mackey et al., 2020). The proposed method achieved proximity error of 0.27 m when BLE beacon was within 3 m from the receiver in two test environments. BLE RSSI-based proximity localization suffers from high variance; thus, a robust technique is necessary to improve the accuracy of proximity services in short-range applications.

3.3.2 Fingerprinting

The second category of RSSI-based positioning is the fingerprinting method, the most preferred and widely used technique in indoor positioning systems. There are two stages in fingerprinting positioning: (i) the Training (offline) phase and (ii) the Online positioning or testing phase. During the offline training phase, the interested site is divided into a specific number of rectangular cells. RSSI measurements are collected from the fixed reference nodes for each cell and stored in a database.

In the online phase, the RSSI values are again observed and mapped with the location information from the training phase stored database to compute the user's position.

It is important to note that the fingerprinting method for positioning does not rely on range calculation to determine the user's location. While RSSI-based fingerprinting has the potential to achieve high accuracy, the performance of this localization method depends not only on having much training data but also on the algorithm used to map RSSI measurements during the online phase with the offline radio map. Furthermore, constructing a robust radio database involves deploying a dense network of access points to cover the desired areas, which increases hardware costs and requires significant manual effort. This process can be time-consuming. The algorithm also faces challenges when there are environmental changes, such as in an office building, which may necessitate rebuilding the radio map. Additionally, NLOS multipath environments pose further difficulties, including extensive data analysis and reduced accuracy in determining the user's position.

Nevertheless, unlike the proximity sensing technique, the RSSI-based fingerprinting positioning method can provide absolute positioning information of a user. Since the main objective of the work is to obtain an accurate and reliable short-range distance estimation from BLE RSSI values and improve proximity sensing applications, fingerprinting positioning will not be used in this project.

3.3.3 RSSI-based Distance and Positioning

This is the last category of RSSI-based positioning where a path loss model is used first to estimate the range from RSSI values. Following the distance estimation, the trilateration (solved with least square or Kalman filtering) method is used to estimate the position of the user. Trilateration is the technique of estimating the location of a target by measuring distance between the target and at least three reference nodes whose coordinates are already known. The location of the target is the point obtained from the single intersection of all three different circular distance measurements. An example of trilateration to find the coordinates of a user using distance measurements from three anchors is illustrated in Figure 3.3. As observed, all the anchors have known locations $A_1 = (x_1, y_1, z_1)$, $A_2 = (x_2, y_2, z_2)$, $A_3 = (x_3, y_3, z_3)$. The target user measures distance from all the three anchors with respect to its unknown coordinates (x, y, z) as shown in equation 3.16.

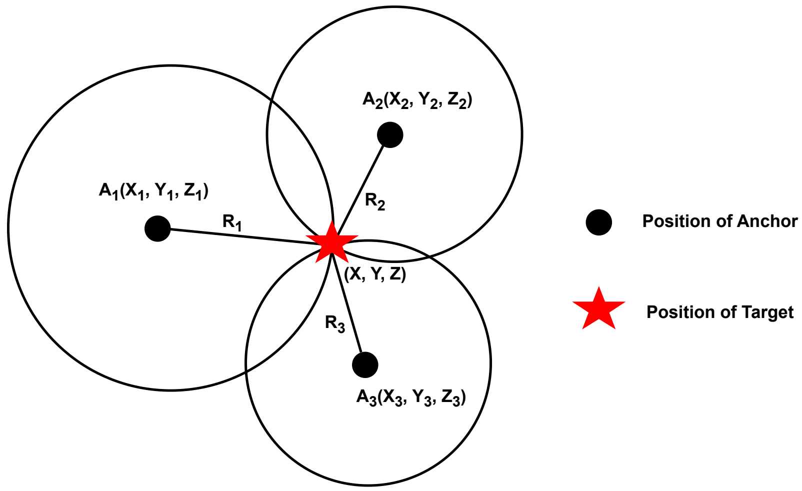


Figure 3.3: Example of an ideal trilateration.

$$\mathrm{R}_1 = \sqrt{(x_1 - x)^2 + (y_1 - y)^2 + (z_1 - z)^2} \tag{3.16}$$

$$\mathrm{R}_2 = \sqrt{(x_2 - x)^2 + (y_2 - y)^2 + (z_2 - z)^2} \tag{3.17}$$

$$\mathrm{R}_3 = \sqrt{(x_3 - x)^2 + (y_3 - y)^2 + (z_3 - z)^2} \tag{3.18}$$

Where (R_1, R_2, R_3) are known as measured ranges. The target user already have calculated ranges from RSSI values using path-loss propagation model technique. Therefore, with two sets of equations (measured range and calculated range), a target user can estimate 3D position solution (x, y, z) using non-linear least square method that minimizes the mean square error of the residuals.

An improvement in positioning accuracy using Bluetooth RSSI for indoor positioning showed by (Subhan et al., 2011). They used standard radio propagation model to estimate distance from RSSI values. The average distance error improved from 5.87 m to 2.67 m using gradient filtered measurements. Similarly, improvement in position solution was showed by applying Kalman filter on meta-heuristic RSSI methods (Amer and Noureldin, 2016). When compared with the fingerprinting approach, their proposed system reduced position error from 4.5 m to 2.8 m with 80% probability. In another study, performance of four wireless technologies (WiFi, BLE, Zigbee, and LoRaWAN) in indoor localization are compared using RSSI values and path-loss model (Sadowski and Spachos, 2018). The study showed the approximated position error using BLE up to 5 m distance is more than 1.1 m when tested in two environments. Furthermore, BLE positioning accuracy showed improvement in centimetre level (0.2-0.5 m) by pre-processing the RSSI measurements using moving average filter (Chai et al., 1604). Distance is then computed using Kalman filter and triangulation algorithm to determine position approximately within a 4 x 4 m room.

The objective of this project is to achieve accurate short-range distance estimation based on BLE RSSI values using an artificial neural network model, focusing on improving the accuracy of proximity detection between smartphone users. In this context, the project does not involve computing the absolute positioning of the user. Consequently, trilateration and least square estimation methods will not be employed in this thesis.

3.4 UWB time-of-flight (ToF) Ranging

BLE systems, which employ signal power (RSSI) to estimate range using propagation models, UWB (Ultra-Wideband) technology calculates the distance between a transmitter and a receiver based on signal travel time. As mentioned in section 2.1.1, a receiver is usually a tag node which depending upon the applications can either initiate or respond to the transmitter (or anchor node). From basic physics, distance between a tag and an anchor node can be calculated using speed and time information of the transmitted signal. The speed of radio wave signal is equal to speed of light in vacuum C. Once the time of flight is known, the distance between tag and anchor can be written as

$$\text{Distance} = ToF(\text{time of flight}) \times C(\text{velocity of light}) \tag{3.19}$$

However, as the transmitter and receiver clocks are not synchronized, the ToF information can not be used directly just from the difference between transmitter and receiver time stamps. This necessitates the calculation of ToF in different ways, which are described below.

3.4.1 One-way Ranging (OWR)

In this mode, both the transmitter and receiver clock has synchronized clocks and devices transmit ranging frame to the other device in one direction (Lee et al., 2009); therefore, ToF can be calculated precisely from the difference between receiver and transmitter time-stamps.

3.4.2 Two-Way Ranging (TWR)

The two-way ranging (TWR) measures the total time taken by a signal to travel from the sender to the receiver and back from the receiver to the sender. This is called round trip time (RTT), as the signal completes a round between transmitter and receiver. Since the clock of only the sender node is referenced to calculate the difference between transmitting and receiving time-stamp instants, the two-way ranging method does not need a synchronized clock of both initiator and responder devices.

3.4.2.1 Single-Sided Two-Way Ranging (SS-TWR)

As described before, a tag node can be an initiator which sends a poll message or packet to the responder (or anchor node). The responder device receives the message, process it, and sends the response back after a delay to the initiator device. The initiator device receives the message, too, after a delay which includes the processing time of the responder. In other words, the initiator device uses an exchange of information to compensate for the synchronized clocks and measure the accurate time of flight (Figure 3.4).

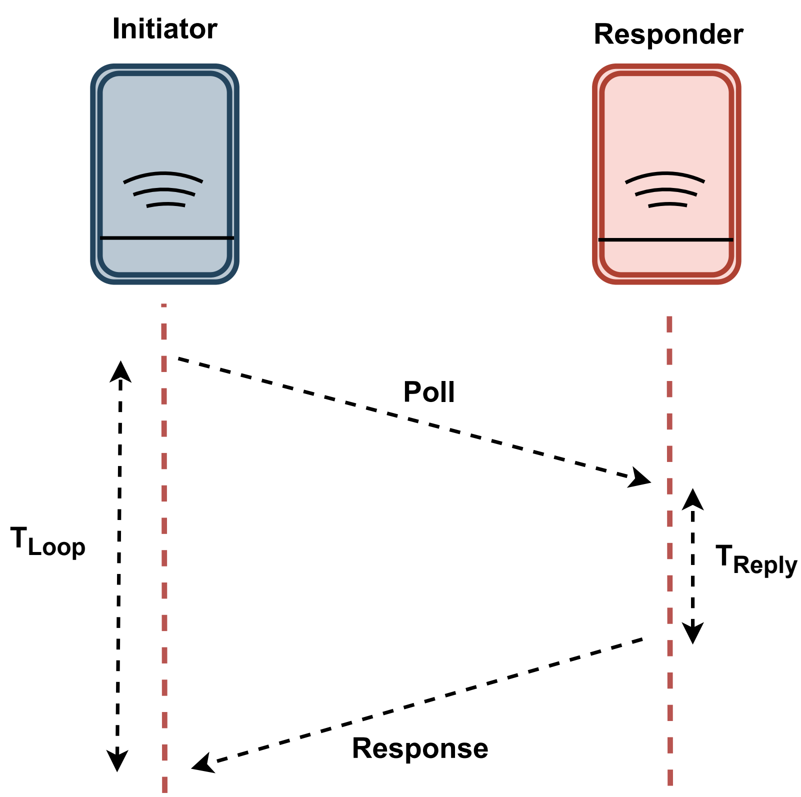


Figure 3.4: Single-sided two-way ranging principle for measuring distance.

The two important parameters the initiator device uses are : (i) T_{Loop} which is the time between when the pooling message was sent and the time response packet was received by the initiator device, and (ii) T_{Reply} is the responder processing time of the received signal which is sent in a response message to the initiator device. The initiator device uses these two parameters to calculate ToF as shown in Equation 3.20.

$$ToF = \frac{T_{Loop} - T_{Reply}}{2} \tag{3.20}$$

Once the ToF is calculated, the distance between the initiator and responder can be obtained as per Equation 3.19. However, to avoid any minor clock deviations in the Single-Sided Two-Way Ranging (SS-TWR), usually, a correction is applied to the relative clock drift from carrier frequency offset estimation (Dotlic et al., 2018).

The Double-Sided Two-Way Ranging (DS-TWR) involves two Single-Sided Two-Way Ranging (SS-TWR) measurements where an additional message is sent to correct relative clock drifts. The DS-TWR can be divided into (i) Symmetric Double-Sided Two-Way Ranging (SDS-TWR) and (ii) Asymmetric Double-Sided Two-Way Ranging (ADS-TWR). Each of these two types is explained below:

3.4.2.2 Symmetric Double-Sided Two-Way Ranging (SDS-TWR)

In DS-TWR, both the initiator and the responder exchange poll messages one more time compared to the traditional SS-TWR method. Figure 3.5 illustrates the message exchange in SDS-TWR. After receiving a reply from the responder, the initiator waits for a certain amount of time and then sends the final polling message. By exchanging the polling message twice, the timing error caused by clock deviations can be reduced, enabling the calculation of Time of Flight (ToF) as described in (Neirynck et al., 2016):

$$ToF = \frac{T_{Loop1} \times T_{Loop2} - T_{Reply1} \times T_{Reply2}}{T_{Loop1} + T_{Loop2} + T_{Reply1} + T_{Reply2}} \tag{3.21}$$

However, increasing the number of polling messages results in a longer overall wait time (or reply time) for both devices. This introduces a delay that is relatively longer than the actual Time of Flight (ToF). The Asymmetric Double-Sided Two-Way Ranging method (ADS-TWR) takes care of this excessive time delay.

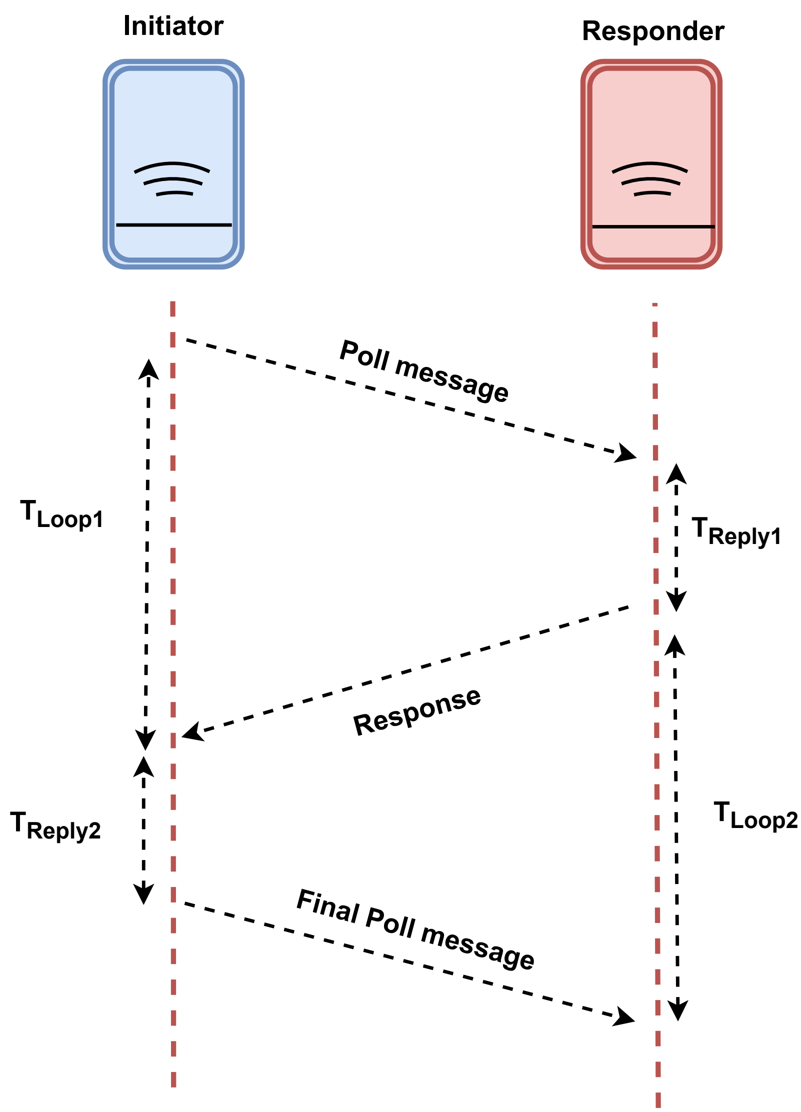


Figure 3.5: The principle of measuring distance through symmetric double-sided two-way ranging (DecaWave, 2015).

3.4.2.3 Asymmetric Double-Sided Two-Way Ranging (ADS-TWR)

An ADS-TWR (Asymmetric Double-Sided Two-Way Ranging) is similar to the SDS-TWR (Symmetric Double-Sided Two-Way Ranging) in the sense that both involve sending additional polling messages to reduce clock deviations in the initiator and responder devices. However, in ADS-TWR, the long wait time is shortened to facilitate quick exchange of polling messages. As depicted in Figure 3.6, once the initiator device receives a reply from the responder device, it immediately sends the final polling message instead of waiting (Jiang and Leung, 2007). By examining Figure 3.6 and

3.5, it is observe that the first round trip time T_{Loop1} is the same as the Equation 3.20. Similarly, the subsequent round trip time T_{Loop2} is also similar to T_{Loop1} but lacks the reply term. The round trip time for both the initiator and responder devices can be summarised as below:

$$T_{Loop1} = 2 \times ToF + T_{Reply1} \tag{3.22}$$

$$T_{Loop2} = 2 \times ToF \tag{3.23}$$

Using Equations 3.22 and 3.23, ToF between two UWB devices can be calculated as:

$$TOF = \left(\frac{T_{Loop1} + T_{Loop2} - T_{Reply1}}{4} \right) \tag{3.24}$$

It is important to note that the key factor influencing the ranging error due to clock deviation is a fraction of 1/4 instead of the standard TWR's 1/2. Consequently, the waiting time is significantly reduced in ADS-TWR compared to SDS-TWR, assuming the same clock offset values.

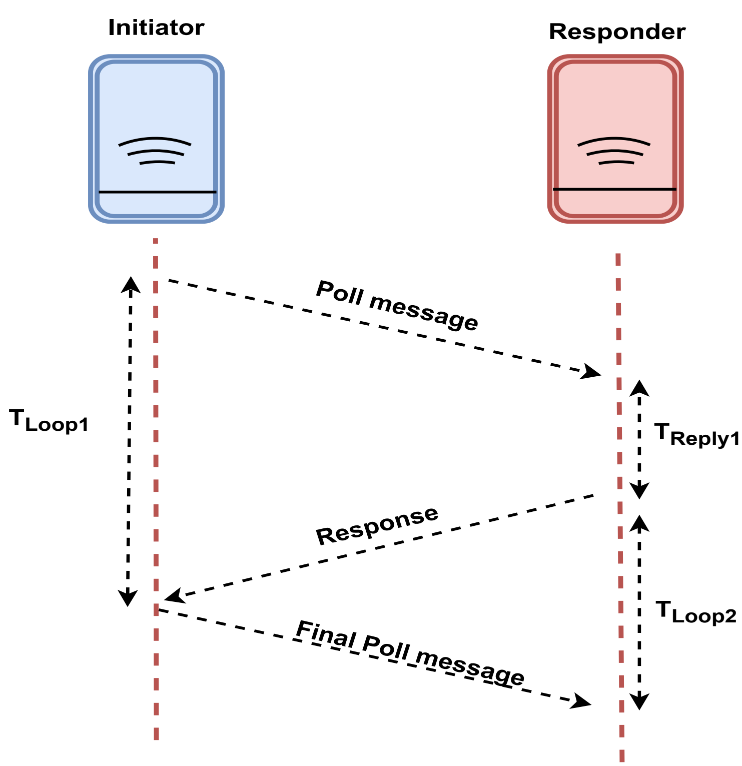


Figure 3.6: The principle of measuring distance through asymmetric double-sided two-way ranging.

As explained in Section 2.1.3 and depicted in Figure 2.3, it is clear that the BLE RSSI values vary over time. Consequently, using the raw RSSI measurements directly for constructing algorithms in wireless localization applications is not recommended. Hence, it is necessary to employ filtering techniques to preprocess the RSSI data, eliminate outliers, and enhance the system's performance. In this thesis, a Kalman filter is employed for preprocessing and smoothing the raw BLE RSSI data. In the case of UWB devices, which are capable of measuring precise range measurements in line-of-sight (LOS) conditions, filtering is not applied. However, in specific situations, filtering is employed to identify non-line-of-sight (NLOS) conditions using UWB measurements (Chapter 5).

3.5 Kalman Filter

The Kalman filter is an algorithm designed to estimate unknown variables based on a set of measurements in a noisy environment. It operates recursively by incorporating the previous measurement history. In this study, the Kalman filter is applied to smooth raw measurements, primarily focusing on RSSI data. The filter consists of two main steps: prediction and update. During the prediction step, the filter uses the previous state estimate (Eq. 3.25) and corresponding covariance matrix (Eq. 3.26) to generate a new estimate and corresponding covariance matrix for the current epoch. This step accounts for the evolution of the system over time. The update step involves computing the innovation sequence, which captures the difference between the new measurements and the currently estimated measurement. This difference, also known as the measurement residual, is then multiplied by the Kalman gain. The resulting value corrects the predicted state estimate for the current epoch. Simultaneously, the covariance state matrix error is updated using the observation model. The summary of Kalman filter algorithm is as follows:

Prediction Step:

$$\hat{x}_{k+1}^{-} = \phi_{k,k+1}\hat{x}_{k}^{+} \tag{3.25}$$

$$P_{k+1}^{-} = \phi_{k,k+1}P_{k}^{+}\phi_{k,k+1}^{T} + Q_k \tag{3.26}$$

Update Step:

$$\tilde{y}_k = (z_k - H_k\hat{x}_{k+1}^{-}) \tag{3.27}$$

$$K_k = P_{k+1}^{-} H_k^T (H_k P_{k+1}^{-} H_k^T + R_k)^{-1} \tag{3.28}$$

$$\hat{x}_{k+1}^{+} = \hat{x}_{k+1}^{-} + K_k \tilde{y}_k \tag{3.29}$$

$$P_{k+1}^{+} = (1 - K_k H_k) P_{k+1}^{-} \tag{3.30}$$

Where,

ϕ is the state transition matrix (3.25) define the time evolution of state estimates from epoch k to k+1

$\hat{x_k}$ is the state estimate

P_k is the state error covariance

Q_k is the system noise covariance

z_k is the new measurements

$\tilde{y}_k$ is the measurements residual

R_k is the measurement noise covariance

H_k is the measurement transformation matrix and K_k is the Kalman gain.

In the given equations, the $\hat{}$ symbol represents the estimated state variable. This means that $\hat{x}$ is an estimate of variable x. In addition, superscripts $-$ and $+$ indicate priori and posterior estimates in the predicted and updated stages, respectively. The Kalman gain is a very critical parameter that decides the amount of information from the innovation sequence to be considered for the final state update. From Equation 3.28, it is shown that the gain depends upon the uncertainty of the current state estimate and the measurement noise (herein variance of the RSSI signal). The innovation sequence or measurement residuals are zero mean, white, and Gaussian, which are used to detect outliers and blunders. Also, the covariance matrix of the innovation sequence can be derived from Equation 3.28 and is shown in Equation 3.31.

$$C_{vk} = (H_k P_{k+1}^{-} H_k^T + R_k) \tag{3.31}$$

Where $C_v k$ is the innovation sequence covariance matrix at the kth epoch and its uncertainty is the summation of both measurement noise (R_k) and propagation of predicted states uncertainty. It is to be noted that the updated error covariance is lower than the predicted error covariance (Equation

3.26 and 3.30), which shows the filter is more confident of the state estimate now after incorporating measurement in the updated step. The two steps process of the Kalman filter algorithm is illustrated in Figure 3.7.

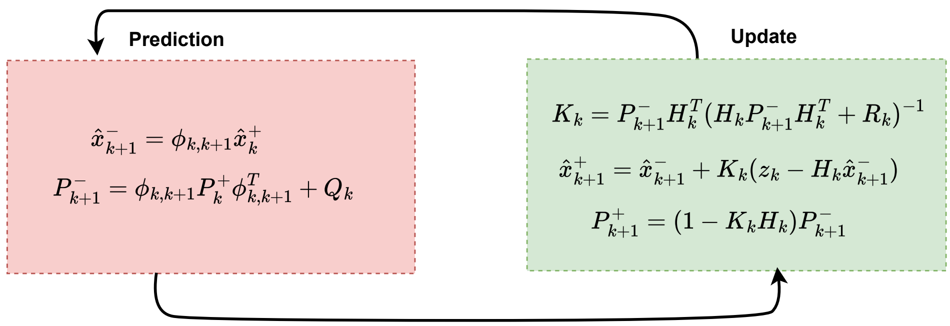


Figure 3.7: Schematic of the Kalman filter with prediction and update steps.

In this work, the 1-dimensional Kalman filter is used to remove blunders and smooth out the RSSI measurements. An RSSI measurement value is provided as input to the filter, and it estimates the next state of $\hat{x}$ (herein output RSSI) by following the prediction and updates step for every new RSSI sample. Since in this application, both transmitter and receiver positions are static mostly; therefore, the state transition matrix and measurement transformation matrix becomes an identity matrix. Moreover the values for process noise Q and measurement noise R are selected from the literature. A few research (Zhao et al., 2018; Kaduskar et al., 2020) investigated similar conditions, uses a small value for Q as most of the system noise is assumed to be added only during the measurement, and R is set to the variance of RSSI values. Hence, in this project, Q and R values are chosen as 0.008 and 3, respectively. As an example, a series of raw measurements of BLE RSSI collected from a fixed transmitter at the fixed receiver location is shown in Figure 3.8. The RSSI values presented in the figure reflect the combined samples of all primary channel measurements, which show more variation and noise.

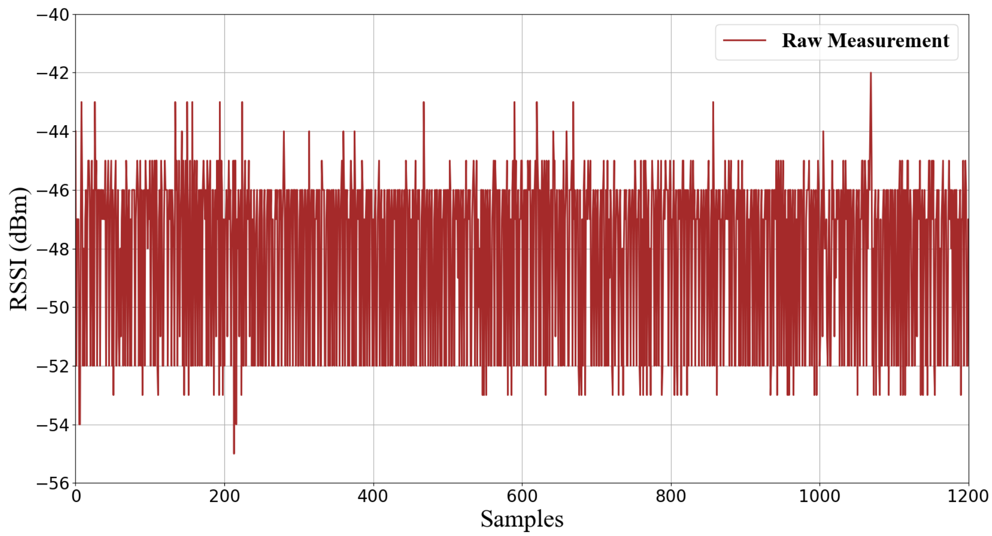


Figure 3.8: Characteristics of line-of-sight BLE RSSI measurements from a fixed single transmitter at a fixed position.

As described in section 2.1.3, the aggregated RSSI has a high standard deviation which makes it less stable as compared to separate channels RSSI. Furthermore, using aggregated RSSI in range estimation eventually leads to large position errors. Hence, RSSI samples are first separated out corresponding to each primary channel (Figure 3.9). The variation of RSSI values in each of the primary channels is reduced but still present. It can also be noticed that each primary channel produces unique RSSI samples for similar reasons explained before (Section 2.1.3). In addition, RSSI values in each primary channel preserve the fine-grained information, which is diminished in the aggregated RSSI. Therefore, Kalman filter is applied on each of the three separate channels to remove outliers and smooth the RSSI data further, as shown in Figure 3.10.

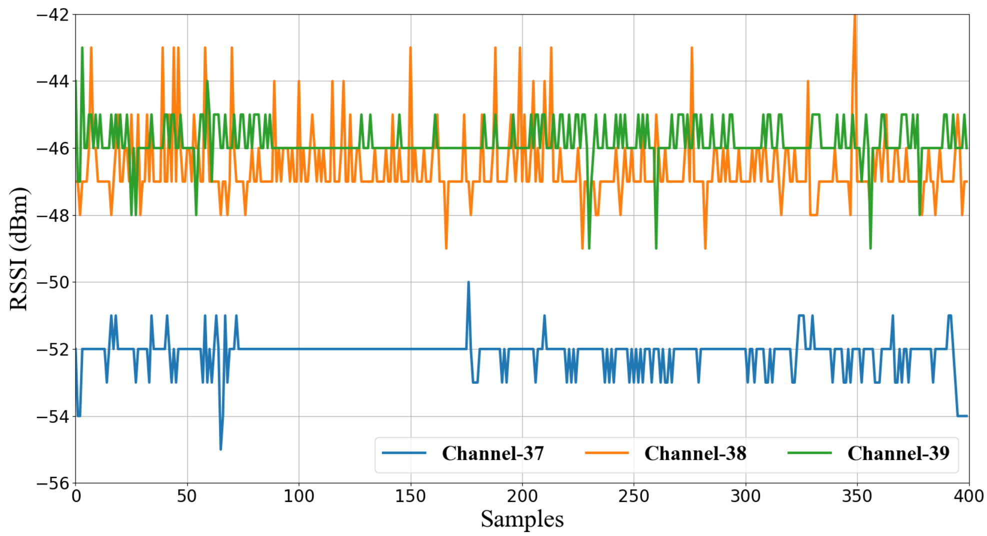


Figure 3.9: Advertising channel RSSI values from the aggregated raw measurements.

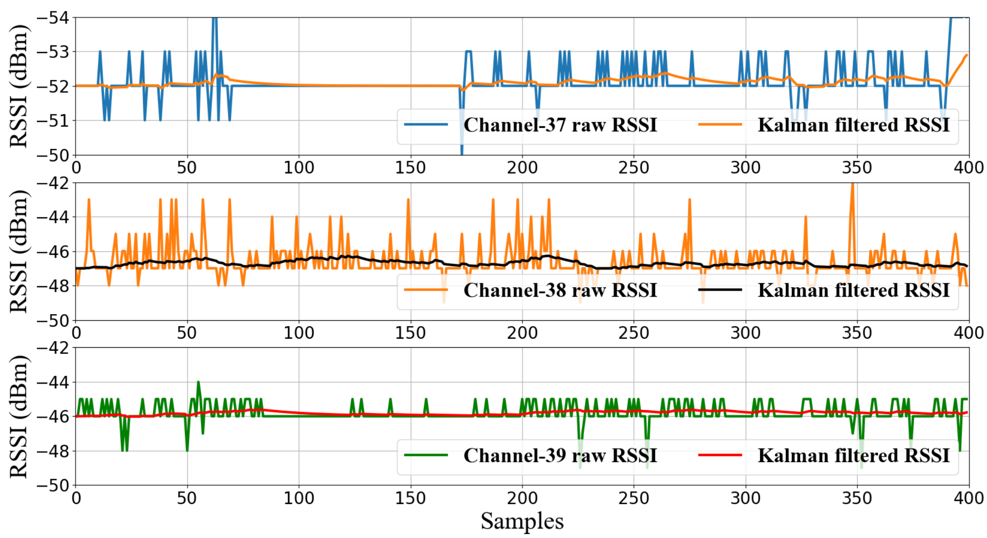


Figure 3.10: Applying Kalman filtering to separate channel raw RSSI measurements provides an illustrative example of reduced variation and improved stability.

3.6 Introduction of Hardware Components

This section provides an overview of the data collection system for capturing BLE and Ultra-wideband measurements. The consumer market has a wide range of commercially available development kits for Bluetooth Low Energy (BLE) and Ultra-wideband (UWB) technologies. Ubisense and beSpoon have emerged as pioneers in real-time locating system (RTLS) using UWB radios. The Ubisense sensor system employs Time-Difference-of-Arrival (TDoA) and Angle-of-Arrival (AoA) techniques, contributing to its superior performance, accuracy, and reliability. However, it's important to note that the Ubisense system requires cable connections between sensors to determine TDoA information. This aspect introduces bulkiness and adds complexity during installation and configuration. On the other hand, BeSpoon has made significant strides in UWB technology. They have successfully demonstrated the integration of UWB into smartphones, although they are still in the early stages of research and evaluation. No concrete hardware is available from BeSpoon, except for a general-purpose modular chip (UM100), which allows users to design their own solutions.

In addition to Ubisense and beSpoon, DecaWave is another well-known commercial manufacturer of real-time location systems based on Ultra-Wideband (UWB) technology. DecaWave has gained significant attention in the industry due to its introduction of a fully integrated UWB chip. The DecaWave DWM1001 module, based on the DW1000 Ultra-Wideband transceiver IC, provides asymmetric double-sided two-way ranging (TWR) time-of-flight measurements with an exceptional accuracy of 10 cm (DecaWave, 2017).

To evaluate the performance of UWB systems in indoor positioning under NLOS industrial environments, an extensive comparison was conducted between Ubisense, beSpoon, and DecaWave UWB devices. The analysis demonstrated that DecaWave's UWB system outperformed both Ubisense and beSpoon in terms of accuracy, establishing it as a superior choice in such settings (Jiménez Ruiz and Seco Granja, 2017). Due to their portable nature, ease of configuration, affordability, long battery life, and simplified installation process, Decawave UWB nodes are an ideal choice for the research work conducted in this project. Consequently, Decawave's Developmental Kit was chosen as the primary research tool. While there are various options available for the BLE data collection system, including developmental kits and Graphical User Interface (GUI) from Argenox, NXP

Semiconductor, Infineon Technology, and others, this project specifically uses the developmental kit from Nordic Semiconductor to collect the RSSI from three primary BLE channels.

3.6.1 DWM1001 UWB Ranging Module

Decawave development board DWM1001-DEV, manufactured by Decawave Ltd in Dublin, Ireland, is suitable for UWB ranging and positioning. The MDEK1001 development kit comprises twelve fully functional DWM1001-DEV boards or nodes. Each development board houses a DWM1001 module, which incorporates the DW1000 Ultra-Wideband (UWB) transceiver chip from Decawave Ltd, an nRF-52832 BLE radio from Nordic Semiconductor in Trondheim, Norway, and an STM-LIS2DH12 accelerometer from STMicroelectronics NV in Amsterdam, Netherlands. In other words, the DWM1001 module incorporates BLE radio, and the UWB transceiver and transmits both BLE advertising data and UWB-ranging information from a single module. The DW1000 chip, which transmits signals with picosecond accuracy (15 ps), employs the ADS-TWR technique, enabling precise range calculation between two UWB transceivers (Sidorenko et al., 2019). These features are crucial for fulfilling the hardware requirements of this project, particularly for data collection purposes.

The selection of DWM1001-DEV boards as transmitters in this project are motivated by their affordability, low power consumption, compact size, and portability. They can be conveniently positioned on tripods or mounted on walls. Specifically, in this project, the DWM1001-DEV transmitters are configured to broadcast BLE advertising data at a high rate, with a 20 ms interval between transmissions.

3.6.2 nRF-52840 BLE Hardware Board

The BLE receiver in this project uses an nRF-52840 development kit from Nordic Semiconductor, based in Norway. This hardware incorporates the nRF-52840 BLE radio, four buttons, 4 LEDs, the PCA10056 chip, and a Near Field Communication (NFC) antenna. It is worth noting that Decawave's BLE transmitter radio, the nRF-52832, is different from Nordic Semiconductor's BLE receiver radio, the nRF-52840. In other words, Decawave employs Nordic Semiconductor's BLE radio for transmitting BLE signals. The Department of Geomatics at the University of Calgary developed a software program called BLEAPPRSCSSW.ZIP using the Arm Mbed C++ development environment to collect RSSI measurements in the primary channels of BLE 4.0. This program is designed to work with the nRF-52840 hardware and serves as a reference for collecting BLE RSSI data in this thesis. The nRF-52840 receiver is configured to collect RSSI values from all advertising channels at a slower rate, with a scanning interval of 50 ms. The higher scanning interval ensures that the receiver captures all the available advertising packets from the BLE sources without any data loss.

Figure 3.11 (a) depicts the DecaWave DWM1001-DEV development board, featuring a DecaWave DWM1001 module highlighted in the orange color, while (b) showcases the DWM1001-DEV unit with a reflective target symbol. Figure 3.12 illustrates the Nordic nRF52840 development board, highlighting the PCA10056 chip and the nRF-52840 BLE receiver chip.

(a) **Decawave DWM1001 module and Microcontroller board.**

(b) **Decawave DWM1001-DEV board with a Reflective target symbol.**

Figure 3.11: Decawave DWM1001 module embedded with BLE transmitter and UWB transceiver chip (Decawave, 2017).

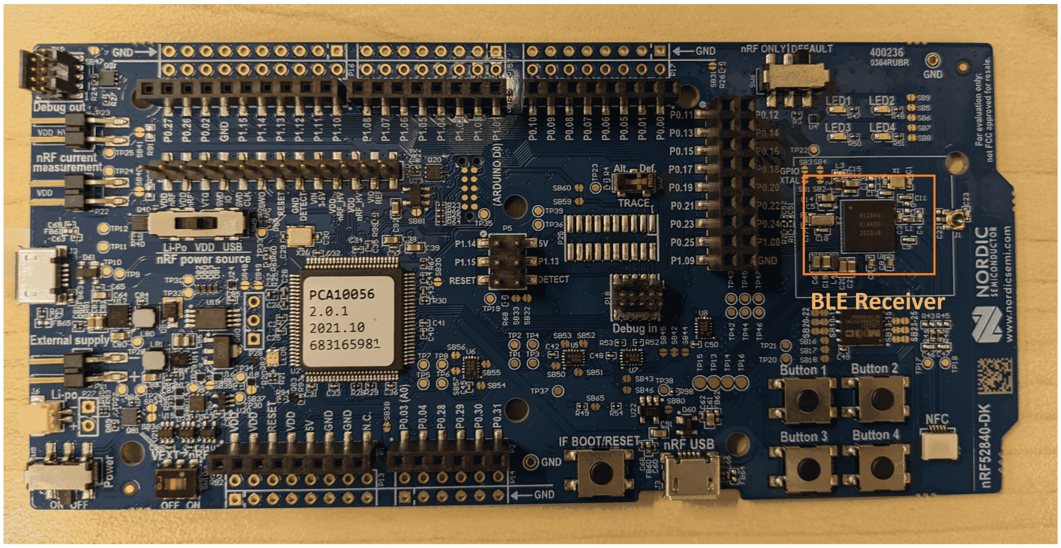


Figure 3.12: nRF-52840 BLE receiver sensor.

3.7 Assessing BLE RSSI and UWB Range in a Test Scenario

This section presents a test scenario to investigate the reliability of RSSI values in determining the distance between two connected nodes. The data collection system consists of two key hardware components, as discussed in Section 3.6. The first element is the Decawave DWM1001-DEV developmental kit, which includes an nRF52832 BLE radio and a DW1000 UWB chip. Consequently, it can function as a BLE source and a UWB transceiver. The second element is a separate nRF52840 developmental kit manufactured by Nordic Semiconductors, which is placed close to the second DWM1001-DEV unit is used as a BLE receiver. The nRF52832 BLE receiver records RSSI values from three primary channels at a rate of 50 ms. At the same time, the two DWM1001-DEV units engage in asymmetric double-sided two-way ranging, performing measurements at a rate of 100 ms.

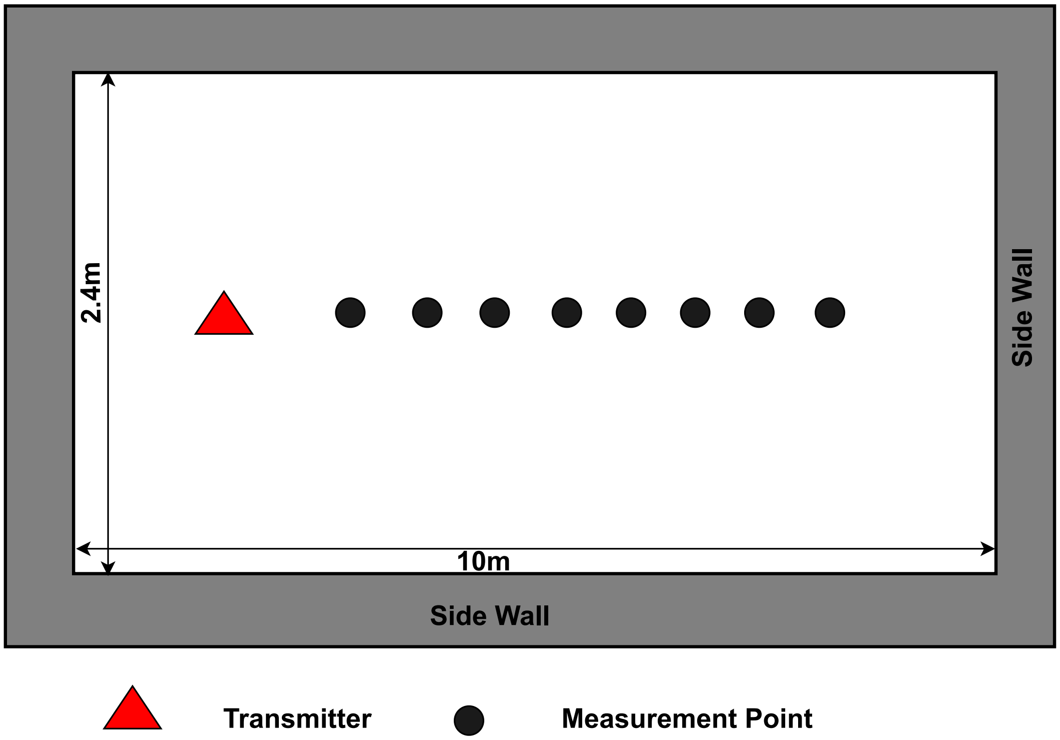


Figure 3.13: A test environment receiving BLE RSSI and UWB range data from a fixed transmitter at line-of-sight.

The system is tested in a closed environment of 2.4 m wide and 10 m in length located on the third floor of the CCIT building of the University of Calgary. The layout is illustrated in Figure 3.13. Measurements are collected in static mode in each measurement point ranging from 1 to 8 m. The distance between the BLE source and each successive measurement point is 1 m. It should be noted that RSSI values are integers and cannot represent decimal or fractional values, limiting their ability to provide high resolution for distinguishing small incremental changes in distance. RSSI is effective at discerning more considerable differences in signal power caused by greater distances (Dong and Dargie, 2012). Hence, avoiding testing RSSI values for small distance increments is advisable. Therefore, in this experiment, RSSI values are tested at 1-metre intervals up to a maximum distance of 8 metres. It should also be noted that the raw BLE RSSI measurements exhibit considerable fluctuations even in static mode. In contrast, peer-to-peer UWB range measurements remain stable and provide accurate measurements within the LOS range.

The observations indicate that RSSI values demonstrate constructive interference and a shift in path-loss exponent at certain distances. Furthermore, as the distance increases, the variability of RSSI values becomes more pronounced. During the data collection, BLE RSSI and UWB range measurements were logged simultaneously for over 15 minutes at each location. Only the samples with matching time instances for RSSI and range were considered from the collected data. Next, a 1-D Kalman filter was applied to the RSSI values of each primary channel to remove outliers and obtain more stable and smoothed RSSI data. The test results, including the average RSSI, average UWB distance, and actual distance for each measurement point, are summarized in Table 3.1.

A simple line-fitting model was employed to analyze the relationship between RSSI values (Y-axis) and the corresponding distances measured from a laser range finder and UWB (X-axis). This model is a reference curve illustrating the one-to-one relationship between RSSI and relative distance, and it is an alternative to the path loss model (Eq. 3.15). For simplicity, only one primary channel (channel 38) was used to establish the mapping between RSSI and distance. The mapping is illustrated in Figure 3.14.

Table 3.1: Channel-38 filtered clustered RSSI and measured peer-to-peer UWB range at LOS.

Laser distance (m)	Average RSSI (dBm)	Average UWB distance (m)
1	-46.82	0.99
2	-49.38	1.99
3	-53.38	2.99
4	-54.02	4.03
5	-56.02	5.04
6	-62.94	6.05
7	-61.02	7.04
8	-72.02	7.97

Finally, the line fit model is assessed to calculate the residual range error for each RSSI measurement. The time series residual error plot, which incorporates both the actual and measured UWB range, is depicted in Figure 3.15

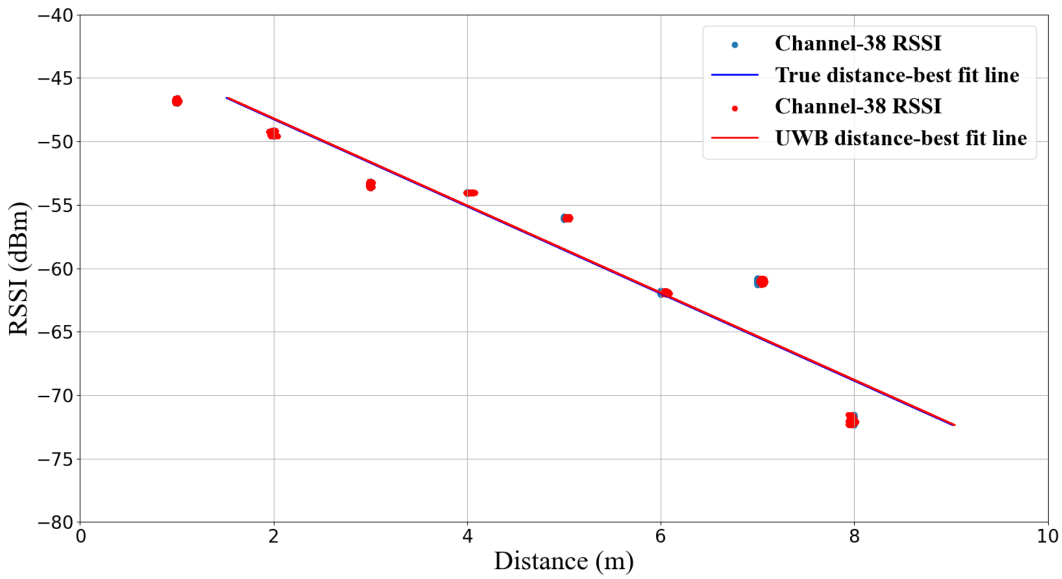


Figure 3.14: Line fitting of channel-38 RSSI measurements illustrating one-to-one relationship between BLE RSSI and reference UWB range.

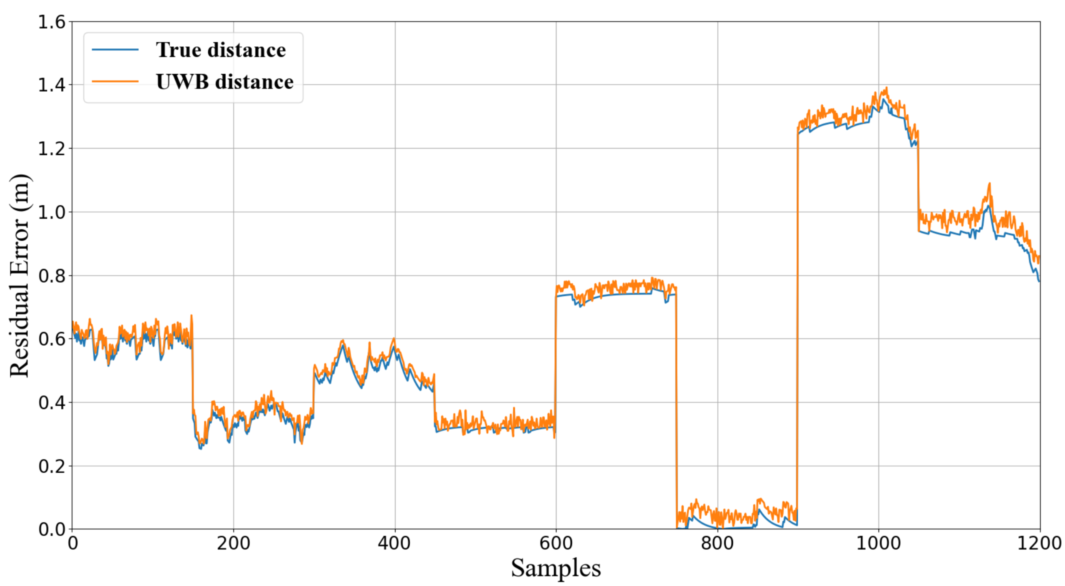


Figure 3.15: Residual error using line fitting model for ground truth and measured UWB range at line-of-sight from a single transmitter.

Several observations can be drawn from Figures 3.14, 3.15 and Table 3.1:

a) BLE RSSI values strongly correlate to the distance between transmitter and receiver nodes. The RSSI values tend to decrease proportionally to every 1 m increase of separation between the two nodes. However, this pattern changes in the presence of a strong multipath signal, leading to constructive interference that affects the received signal, as evidenced by the data presented in Table 3.1.
b) The mean value of the measured UWB ranges closely aligns with laser distance measurements at each location. A minimal deviation of approximately 1 cm is observed for shorter distances up to 3 m. In comparison, a slightly higher variation of up to 5 cm is observed at longer distances, specifically when 6 m away from the transmitter.

c) The line-fit model, which includes both the true (laser range finder) and UWB distance, adequately fits the RSSI data. However, focusing on shorter distances can improve the model's performance. In this example, the linear model provides an alternative approach to estimating RSSI path-loss exponents, serving as an alternative to the standard path-loss propagation model.

d) The linear model partially fits the data, increasing residual range errors in the metre range as the peer-to-peer distance increases. This discrepancy is from variations in estimated path loss exponents, compromising distance estimation accuracy. Therefore, the line fit model cannot accurately map RSSI to distance.

e) The line fit model shows similar magnitudes of residual range errors compared to actual distance and peer-to-peer UWB range, as depicted in Figure 3.15. This alignment is expected since the UWB range provides accurate measurements comparable to the true distance. These results suggest that using the UWB range could be more suitable for training BLE RSSI data in range estimation.

The test scenario highlights that the UWB range outperforms the standard path-loss model and the simple line fit model in determining the BLE RSSI path-loss exponent along the same measurement path. Moreover, both models are inadequate for accurate range estimation. The standard path-loss model is highly dependent on the propagation environment, while the line fit model fails to fit the data effectively. Recognizing these limitations and leveraging the benefits of UWB, this research introduces a novel algorithm. This algorithm uses UWB range measurements as the reference ground truth and employs artificial neural network (ANN) techniques to train BLE in converting RSSI into corresponding ranges. The following section provides an overview of the system model development, emphasizing the integration of the UWB range and BLE RSSI signals within the ANN algorithm.

3.8 Overview of Proximity Detection model

The general architecture of the proposed proximity detection algorithm is depicted in Figure 3.16 (a). The model comprises two hardware components: a stationary transmitter and a tag (or mobile receiver). The co-located receivers capture both UWB range and separate channel RSSI measurements, which are transmitted simultaneously to a computer through serial ports. Before further analysis, the data undergoes preprocessing using a 1-D Kalman filter to eliminate outliers and noise, explicitly focusing on enhancing the quality of RSSI data. This preprocessing step generates clean inputs that neural network models can effectively process. The ANN model, which incorporates a convolutional neural network and a random forest, is trained using supervised learning on a labeled dataset. The labeled data represents the UWB range, while the input consists of RSSI values from three advertising channels. After the training, the model predicts range outputs for the three input channels when presented with new test data. It is important to note that the proposed model assumes the involvement of expert users within a group who can supply UWB range data as the reference truth for training BLE in converting RSSI to the range.

The trained neural network model is designed for proximity detection between two BLE-enabled devices, such as smartphones, which only have access to each other's RSSI values. Figure 3.16 (b) illustrates a simplified schematic of the ANN model trained by expert users and how a general mobile user can use it to perform proximity detection with other users in a small, crowded area. As an example, the user collects sufficient RSSI values from two nearby BLE-enabled users and employs the trained model to predict the distance between them. Subsequently, the user can establish a safe proximity threshold for these users. If a user approaches this threshold, the model can detect the situation and issue a warning notification to that user.

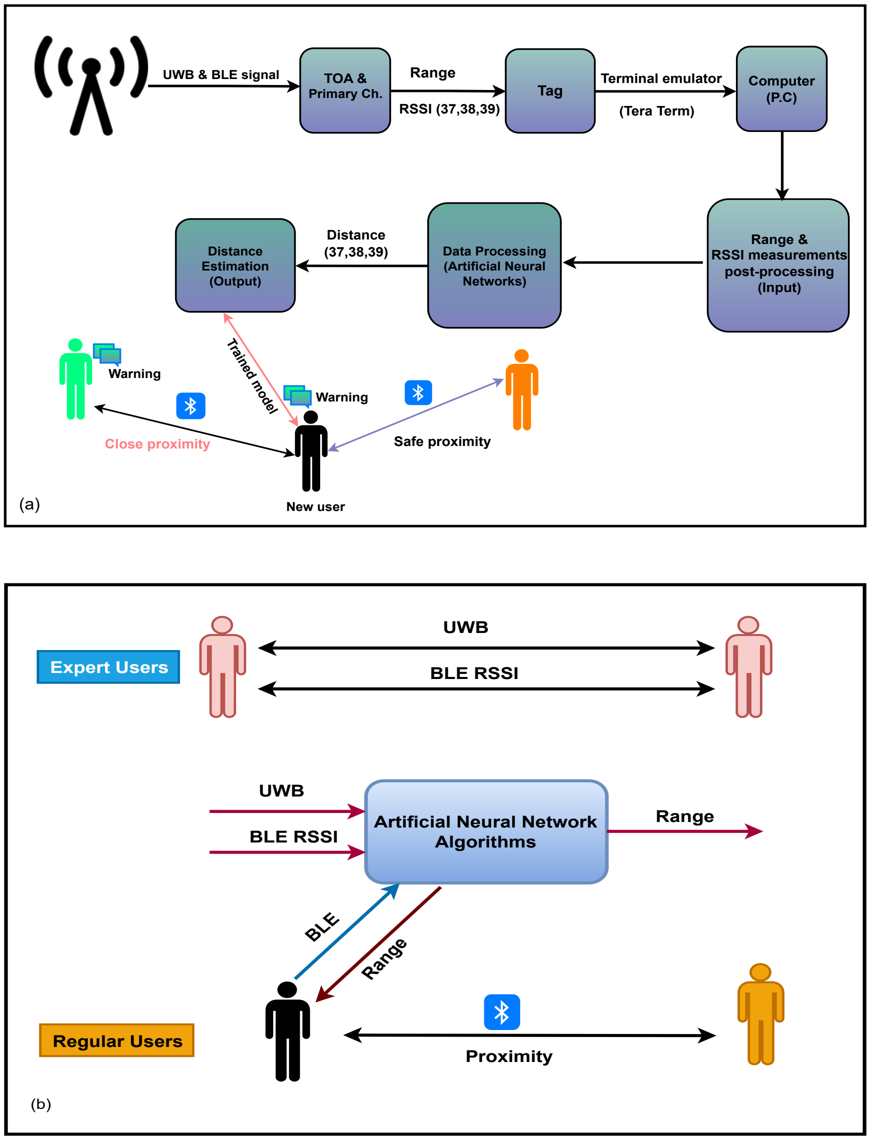


Figure 3.16: Architecture of the Proximity Detection Model: (a) Training of BLE RSSI using UWB range with ANN techniques, (b) Simplified schematic of ANN model used by BLE users for proximity application.

3.9 Summary

This chapter introduces the ranging techniques using BLE RSSI and UWB timing signal characteristics. It emphasizes the standard path-loss model in computing range estimation from RSSI values in a dynamic propagation environment. Furthermore, a concise overview of the Kalman filtering technique for processing and filtering time-series data is presented. Through a real-world example mapping RSSI to distance, the limitations of path-loss exponent estimation and the advantages of using UWB range as the reference ground truth is highlighted. The chapter proposes a general architecture for the proximity detection model, catering to BLE users in the presence of expert UWB users, using the ANN techniques. In the subsequent chapter, the software design and training of two neural network models are elaborated. It describes the data collection, training, and verification steps for the ANN algorithms. Additionally, it briefly covers the test environments and conditions in which the training dataset is prepared for training and testing of ANN models.

Chapter 4

Software Realization of System Model

4.1 Introduction

Chapter 3 highlighted the drawbacks of BLE RSSI in position estimation using the existing standard radio propagation model. It then derived relationships between RSSI and distance through a curve-fitting model. It demonstrated that a user position could not be reliably determined using received RSSI values from a fixed transmitter. On the contrary, the UWB range offered an added advantage in providing training data for the RSSI measurements and improving the feasibility of accurate distance estimation.

Following this discussion, this chapter focuses specifically on the research at hand and introduces two neural network models designed to estimate proximity between two BLE-enabled users. The primary objectives of this chapter are as follows:

- Training an artificial neural network model using BLE RSSI and UWB range through a supervised learning mechanism.
- Validate and assess the effectiveness of the proposed algorithm by evaluating the training dataset in the training environment.

The chapter begins by describing the test environment, experimental setup, and data collection procedures followed during the tests in Section 4.2. Subsequently, Section 4.3 provides a detailed explanation of the step-by-step procedures for training the ANN model. Furthermore, a brief analysis of some critical tuning parameters for each algorithm is presented while training the RSSI-range model and used throughout the tests.

4.2 Experimental Setup

The training data for the neural network model is collected in an empty office lab environment. More details of the training environment is provided in Section 4.2.1. Following the training phase, the model will be assessed in different environments to evaluate its performance (Section 5.2 and 5.3 of Chapter 5).

4.2.1 Empty Room Environment

The medium-sized empty office room is located on the 3rd floor of the G-block Engineering building at the University of Calgary (UofC). Although the room is unoccupied with ample space in the middle, there are furniture items such as tables, chairs, and cupboards positioned along the side walls. The shape of the room is the rectangular type with dimensions (length x breadth) approximately 6 m by 11 m, as shown in Figure 4.1. To establish a reference point, the top left of the room is considered the origin of the local coordinate system as illustrated in the 2D plan view of the room (Figure 4.2). In order to facilitate measurements and evaluations, the floor is marked with four columns of reference points (RPs), comprising a total of 34 RPs spaced 1 m apart. At one of these reference points, a Decawave DWM1001-DEV module is mounted on a tripod at a height of approximately 1.2 m. This module functions as a transmitter, capable of simultaneously transmitting UWB signals and BLE RSSI values across three advertising channels.

Figure 4.3 provides a visual representation of the user's measurement path, indicated as the second column from the right, which follows the North-South direction of the room. In order to capture both types of measurements, the test user is equipped with two co-located receivers: the nRF52840 DK module and the second DWM1001-DEV module. The nRF52840 unit which serves as a BLE receiver sensor, receives RSSI values across three advertising channels. On the other hand, the second Decawave module is responsible for collecting UWB range data. To ensure seamless data acquisition, these receivers are securely mounted on a metallic stand side by side. Additionally, the metallic body is conveniently placed on top of a movable cart, allowing for easy maneuverability and stable position in front of the user body throughout the measurement process.

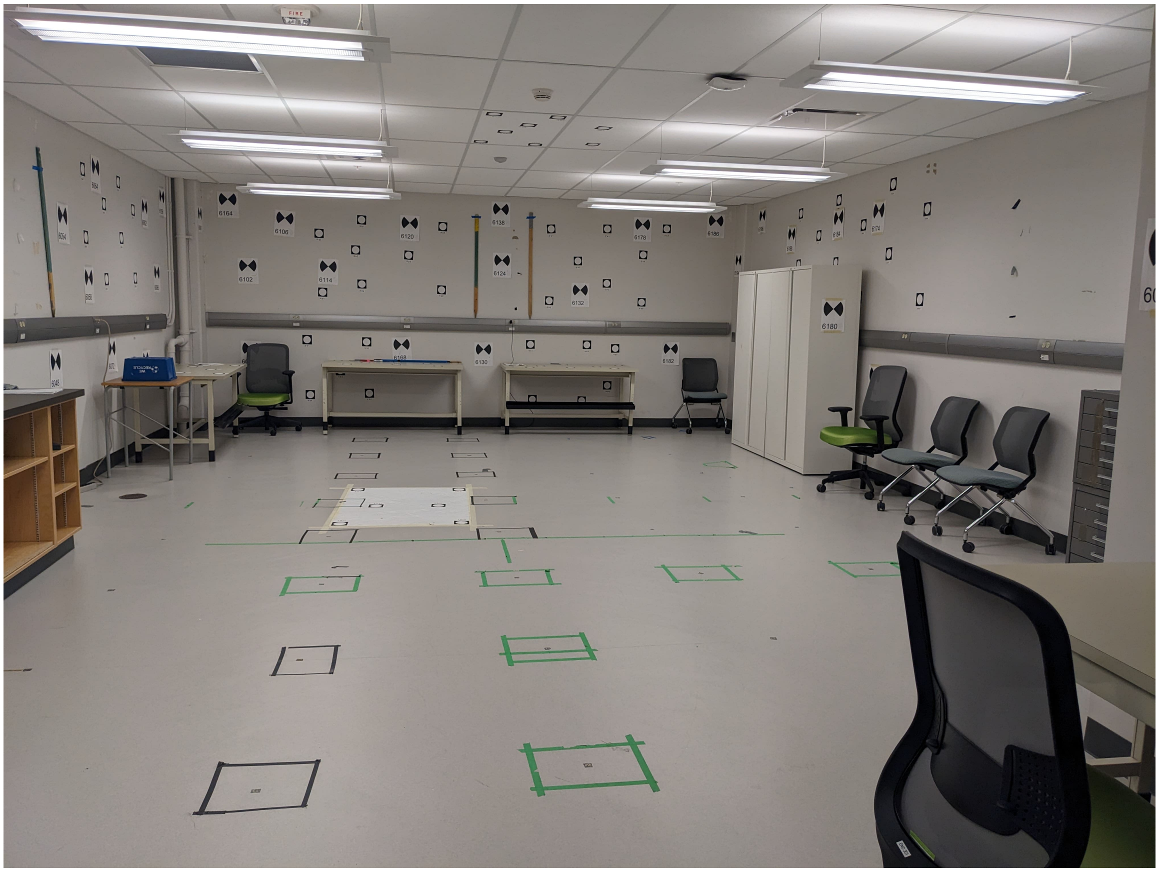


Figure 4.1: Empty room inside view.

The combined height of the movable cart and the two receivers on the metallic stand was approximately 1.15 m. This specific height was chosen to ensure that both the transmitter and the receiver are aligned at the same elevation. Moreover, this height reflects the average position at which users typically hold smartphones while operating them.

The data collection process involved capturing measurements in a LOS configuration between the transmitter and the test subject. As the test subject gradually moved away from the transmitter within the range of 1 m to 8 m, continuous RSSI and range values were collected. The RSSI measurements were logged at a rate of 50 ms, while the UWB range values were logged at a rate of 100 ms. To facilitate data logging, a laptop computer equipped with two serial ports was used. This allowed for simultaneous recording of the RSSI and range data and synchronized measurements using PC-based time stamp throughout the experiment.

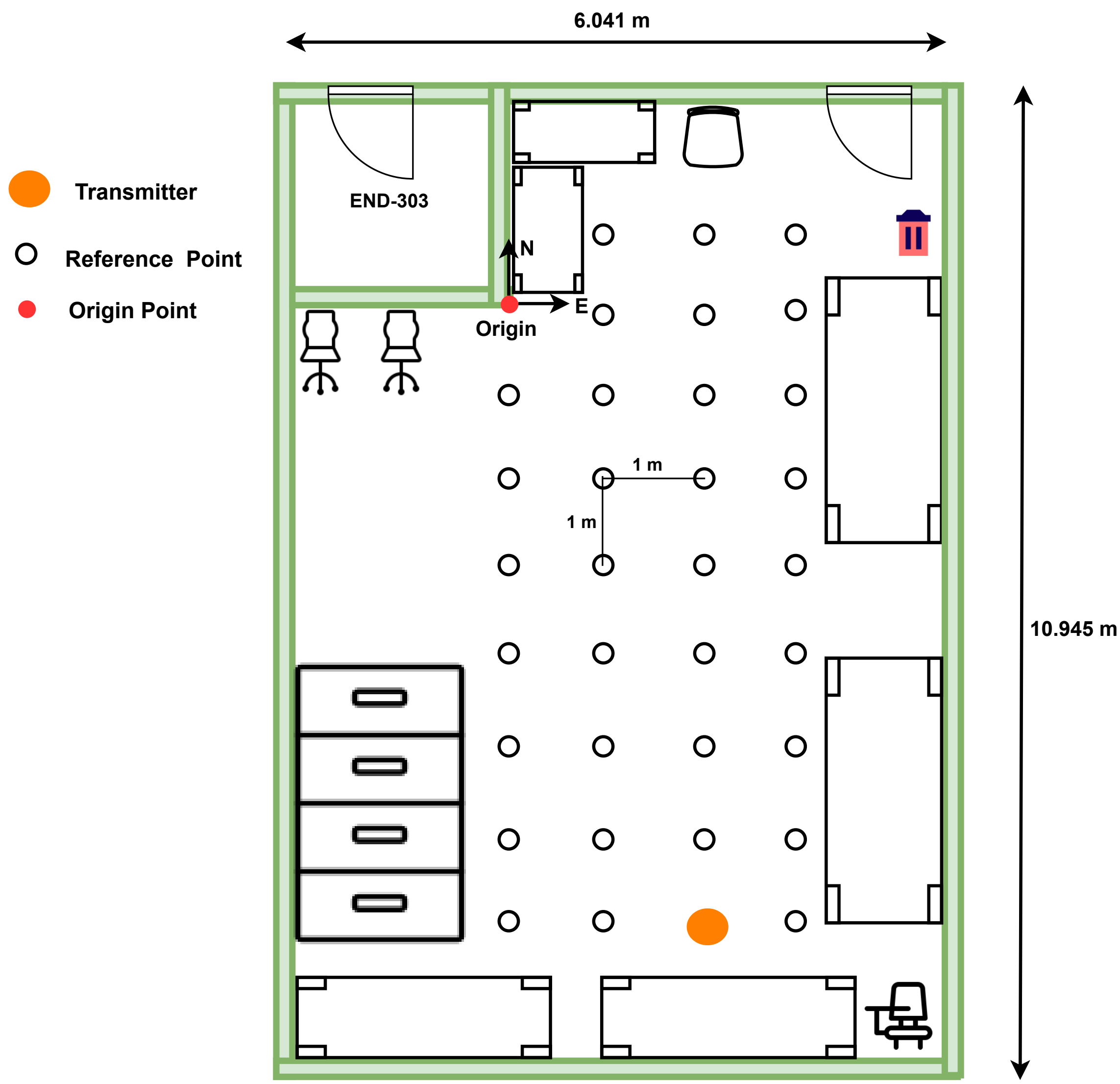


Figure 4.2: The 2D plan view of the training room.

Post-processing is specifically applied to the BLE RSSI values, as they are prone to exhibiting random variations due to their inherent characteristics. Additionally, as the user moves throughout the measurement path, the RSSI values experience fluctuations and outliers. A Kalman filter is employed to mitigate these effects and ensure clean data as requisite for neural network training. During the data collection, 2130 RSSI observations were recorded for each primary channel and their corresponding UWB range values. While neural networks generally benefit from larger datasets for training and producing reliable outcomes, this dataset size is considered as moderate and sufficient for training the ANN models.

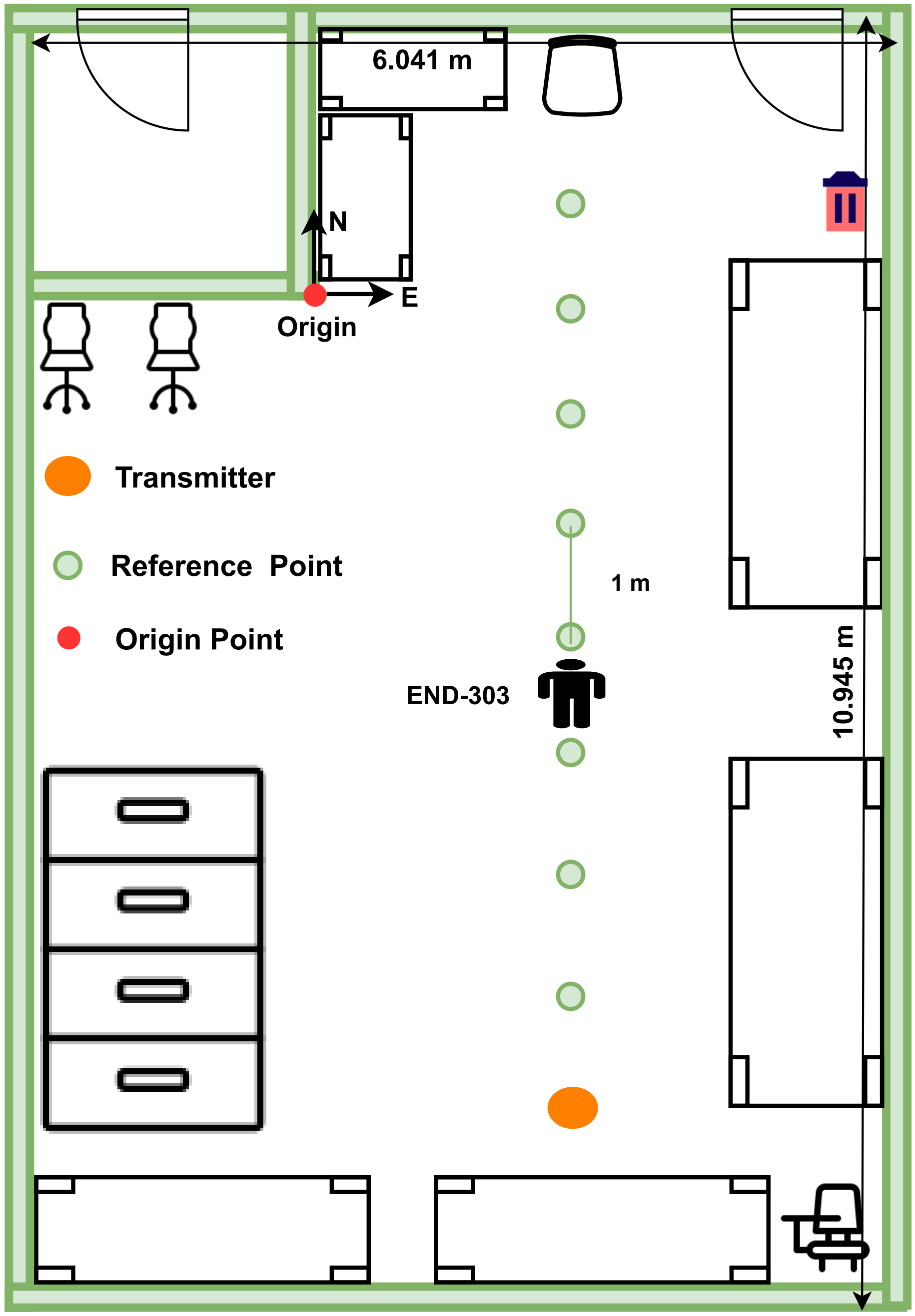


Figure 4.3: Test user training data (BLE RSSI & UWB range) collection measurement path at LOS.

4.3 Artificial-based Proximity Model

This section presents two artificial-based approaches, i.e., random forest and convolutional neural network model, to train a BLE RSSI to range model using UWB range as training data. From Section

2.3, a neural network consists of numerous neurons and layers capable of learning and extracting specific hidden features from the data. This makes ANN algorithms well-suited for handling complex and noisy input measurements. Additionally, with sufficient training data, ANN can effectively capture patterns in the data that may not be discernible to conventional decision-making techniques. This is because artificial-based methods employ powerful statistical algorithms to analyze the data, allowing them to learn the relationship between the input and output data.

The ANN model in this work is realized in Python using TensorFlow library. The step-by-step procedures for training an ANN model for predicting range output from RSSI measurements are mentioned below:

- Collect RSSI (herein filtered) of each BLE advertising channel and corresponding UWB range measurements from the fixed transmitter.
- Randomly shuffle both the input (RSSI) and output (range) data using the `train_test_split` method from the sklearn library.
- Use the significant percentage of the observations (for example, 90%) from the previous step for training the ANN model, using them as input and output combinations.
- Evaluate and test the performance of the model with the remaining 10 percent (based on percentage of data used in previous step) of measurements.
- Fine-tune the hyper-parameters of the model to improve its performance and generalization.

It is important to note that both the ANN models (RF and CNN) use the same training dataset but a different sample size for training and evaluating their respective algorithms. The training and test procedures of the ANN technique are explained more comprehensively as illustrated in Figure 4.4. Once the training phase is complete, the ANN models can provide three range outputs corresponding to the three advertising channel RSSI inputs. These range outputs are predicted distances between the transmitter and the receiver based on the received RSSI values. Alternatively, the model can be retrained to output a single distance estimate using three separate channel RSSI values (refer to Appendix A: Neural Network Re-Training for details).

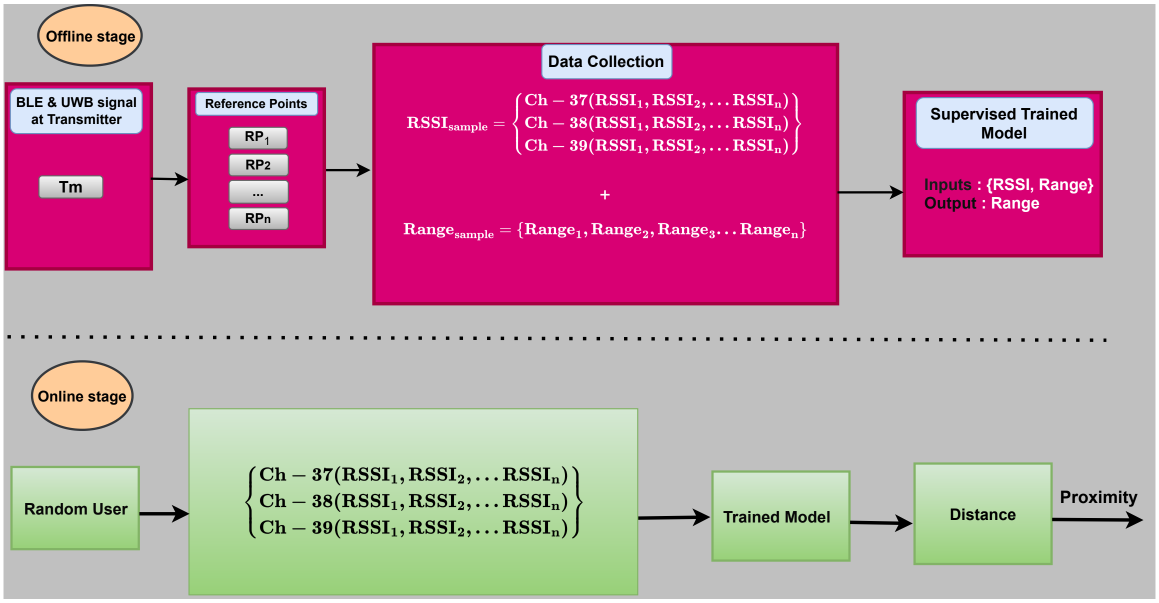


Figure 4.4: Artificial-based proximity model schematic using BLE RSSI and UWB range measurements.

As shown in the block diagram, the ANN technique involves two stages (offline and online) of model training and testing. In the offline stage, the RSSI values of three advertising channels and range measurements are used to train the model using the supervised learning techniques. The objective is to guide the model to learn the relationship between the RSSI inputs and the reference range outputs. As a result, the trained model generates three range values as outputs, each corresponding to the RSSI inputs from advertising channels. These range values represent distance estimates derived from the RSSI values, without using a regular log-distance model. In the online stage, the trained model is provided with only RSSI values (in all three channels) to predict distance as output (one for each RSSI channel). This enables the ANN technique to function as an RSSI-distance model or proximity model, capable of estimating distances using RSSI values alone.

4.4 Neural Network Training

Back-propagation is the most fundamental training process for any neural network algorithm. Using this algorithm, input data are first forward propagated through hidden layers with constant weight and biases and the output is produced at the outer layer neurons. The neural network then actually start learning from back propagating the error signal and determine the changes in weight and biases to minimize the error cost function and eventually produce accurate results. Error cost function is a function defined as the error between obtained values and expected values to measure the performance of the model for given data. The TensorFlow library automatically applies backpropgation during training of model. The mean squared error function (MSE) can be expressed as below:

$$J = \frac{1}{m}\sum_{i=1}^{m}(d_i - y_i)^2 = \frac{1}{m}\sum_{i=1}^{m}e_i^2 \tag{4.1}$$

where, d_i and y_i represent the expected and actual outputs for i^{th} trained sample, respectively, while e_i and m denote the sample error and the number of training samples, respectively.

4.4.1 Random Forest Model Training

The general training procedure is same as described in preceding section. Moreover, from section 2.3.5.1 it is described that random forest is an ensamble learning method where multiple decision trees are trained using different subsets of the training data. The final prediction of the model is obtained from aggregating the predictions of all the tress. In this case, the random forest model is trained using 100 decision trees, which is the default number of trees in the implemented Keras model. The training dataset consists of total 2130 samples for each advertising channel. In the proposed model, hundred decision tress are used to train 90% of the data i.e 1917 samples and each one of them is trained independently. Figure 4.5 shows the relationship between distance and signal power of the training samples. It can be observed that the BLE RSSI signal levels exhibit time-varying characteristics in the indoor environment. For peer to peer distance of 8 metre, the signal power changes up to -75 dBm. The remaining 10% of the training samples (213 samples) are reserved for testing the trained model. Figure 4.6 depicts the trained model output on test data and shows the same behaviour as training data samples. In order to verify the prediction of the model, the residual error is calculated using model output and the true label output (reference UWB range). The residual error obtained is illustrated in Figure 4.7. The error is concentrated around the mean value of zero with standard deviation of 0.08 m. This indicates the learning capability of RF algorithm by observing data and making quality prediction for time series regression problem as in this case.

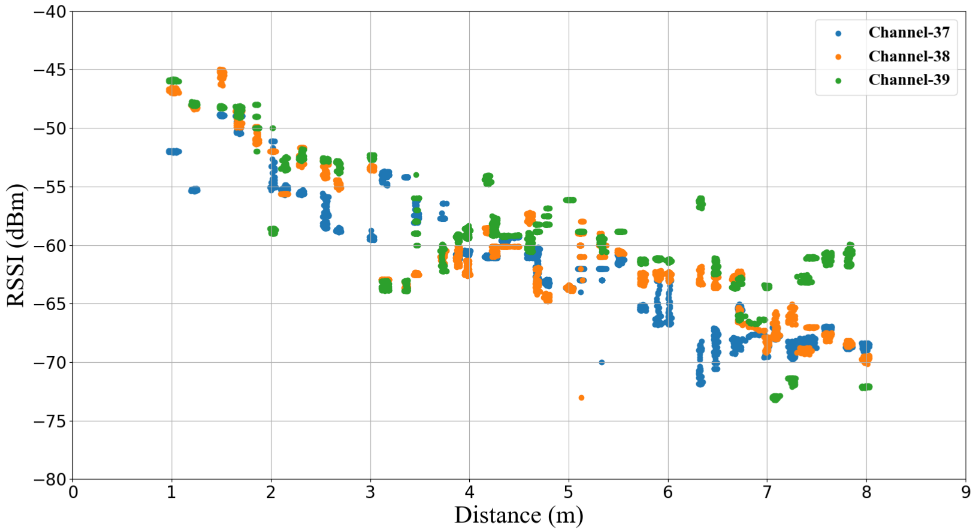


Figure 4.5: BLE RSSI and UWB range dataset for training random forest algorithm at the empty lab environment.

Figure 4.8 shows the cumulative distribution function (CDF) of distance estimation errors. The 90^{th} percentile of range error for each RSSI channel are : 0.038 m , 0.037 m and 0.035 m respectively.

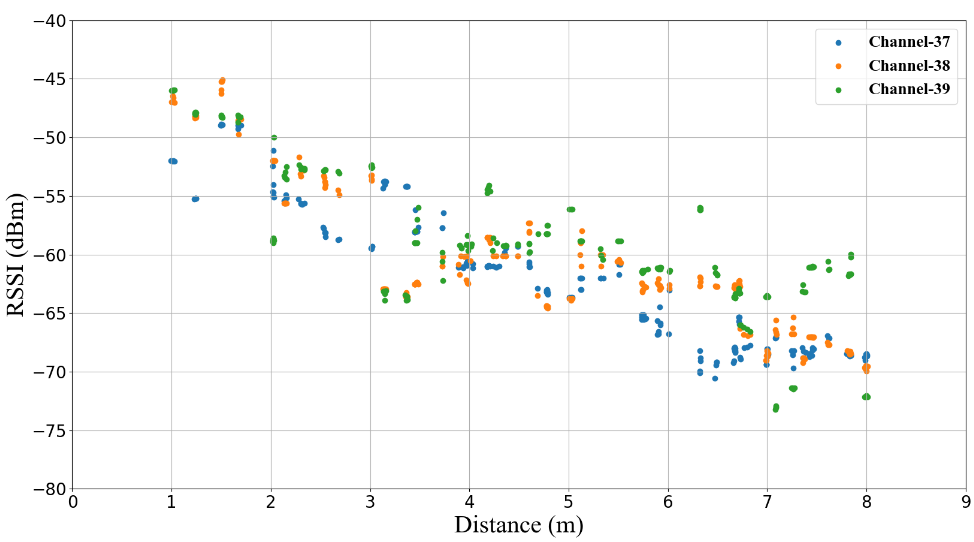


Figure 4.6: Predicted distance output of RF algorithm on test samples.

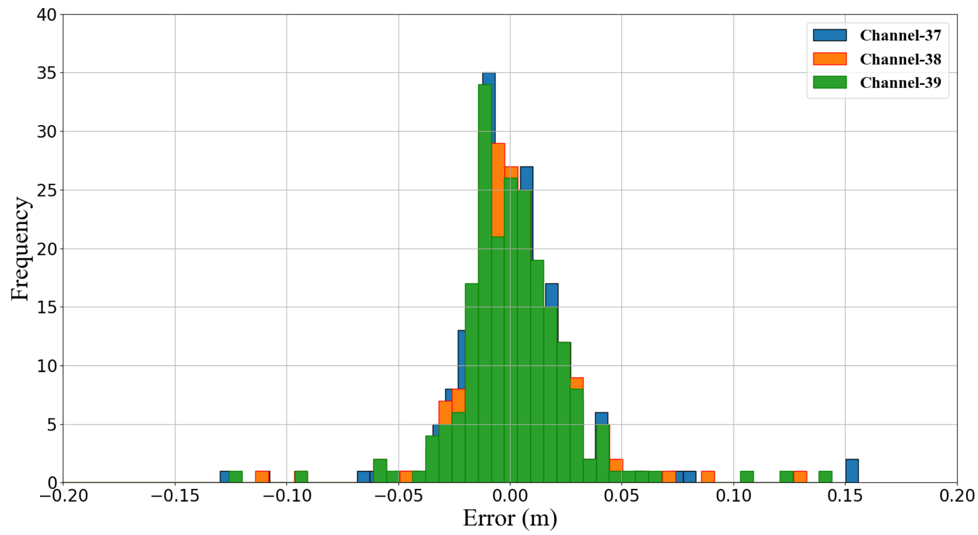


Figure 4.7: Residual error between predicted and reference distance.

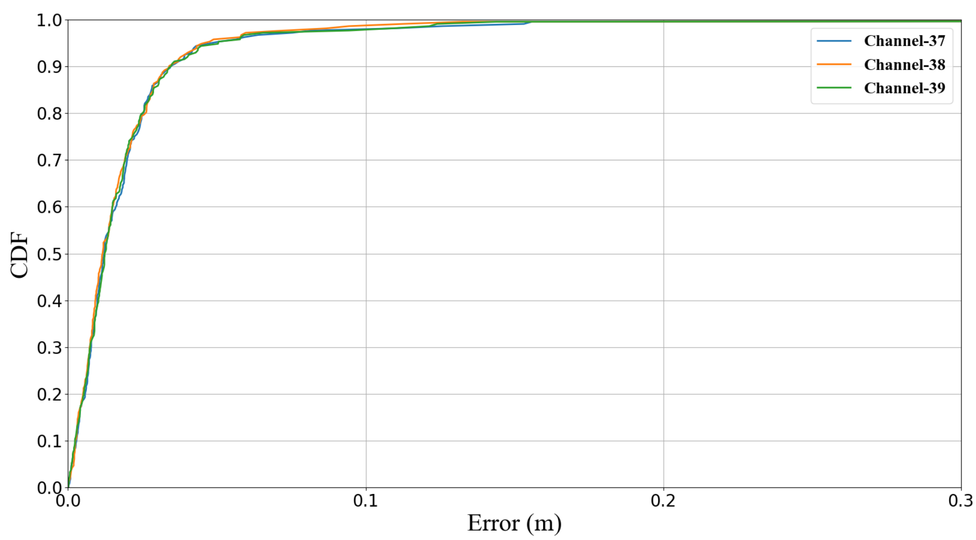


Figure 4.8: Cumulative distance error of each primary channel.

When using the random forest ensemble, it is not meaningful to plot the training accuracy or loss of each individual tree. Instead, the performance of the ensemble is evaluated using metrics such as accuracy, precision, F1-score, or R-squared, depending on the type of problem, be it classification or regression. Since the problem in this thesis falls into the regression category, the R-squared or R2 score and mean square error (MSE) are calculated for the test dataset. The R2 score is a metric that typically ranges from 0 to 1, where a value close to 1 indicates a perfect fit of the data, while a value of 0 suggests a poor fit. Similarly, the MSE represents the averaged squared difference between the predicted values and the actual values. The MSE measure is 0.01, which signifies a good performance. The performance of RF model using 100 decision trees on the test dataset are summarized in Table 4.1.

Table 4.1: RF model training parameter and performances.

No. of Trees	R2 score	MSE	Mean residual error (m)	CDF 90% (m)
100	1.00	0.01	0.01	0.037

4.4.2 Convolutional Neural Network Model Training

Training a convolutional neural network follows the same initial procedure described in random forest training. Like the RF training, the training dataset for the CNN model consists of the same dataset. The training dataset includes 2130 RSSI samples and corresponding UWB reference ranges. However, in this case, 80% of this data (1704 samples) is used to train the model while the remaining 20% of data (426 samples) is used to validate the model. This is to evaluate the model with the more test dataset. The CNN model is trained and implemented using Keras built-in sequential model. The training procedures are described as follows: the model has one convolutional layer as the first layer with 256 neurons (or filters of size 2), followed by a dropout layer of 0.1% to prevent overfitting, a flattened layer to convert the obtained values into a 1D array. This is followed by another dense layer of 128 neurons with ReLU activation function and a final dense layer with three outputs for each channel. The model is compiled using the mean square error loss function and Adam optimizer with a 0.001 learning rate. Lastly, an early stopping callback feature is employed to stop the training process if the validation loss does not improve for 10 epochs. The summary of the CNN training parameters is mentioned in Table 4.2

Figure 4.9 illustrates the training loss function plot of the CNN model. Here, the loss function is the MSE and a function of number of training epochs. The graph shows smooth behaviour of both training and validation loss curve after just a few epochs which indicates the network has come to its best learning phase with minimum error after a certain number of epochs.

Table 4.2: CNN model training parameters.

Parameters	Neurons	Dropout	Activation function
Conv1D	256	-	ReLU
Dropout	-	0.1	-
Flatten	-	-	-
Dense	128	-	ReLU
Dense	3	-	-

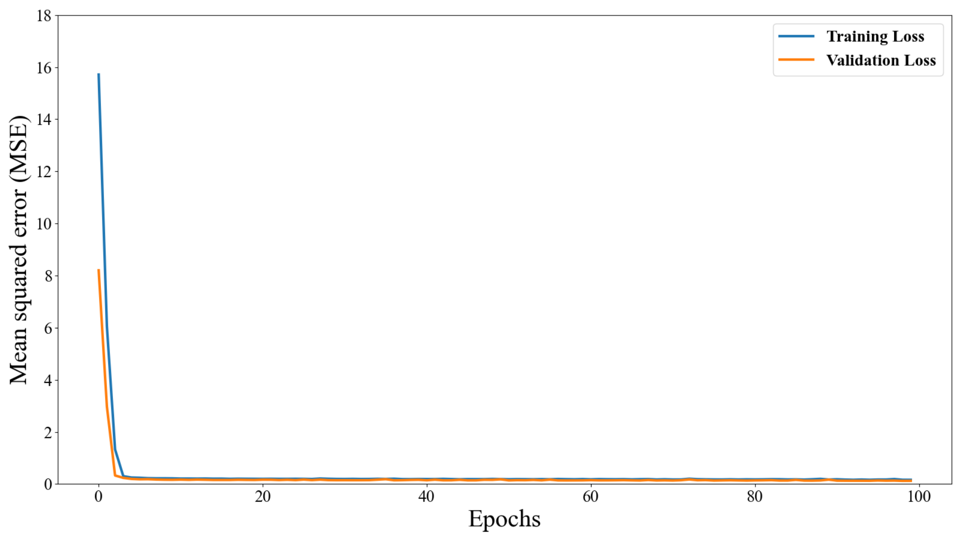


Figure 4.9: Training and validation loss of CNN using parameters in Table 4.2. The best validation performance is 0.134 at epoch number 8.

After training the model, the system performance is evaluated on test dataset which comprises of 426 samples. Figure 4.10 shows the model predicted results of three range outputs from input RSSI values. Comparing with the Figure 4.6, the CNN output results also follows the same behaviour. However, the residual error obtained using CNN network is larger than the RF algorithm. The CNN residual error is shown in Figure 4.11 with error ranging from -0.75 - 0.75 m. The mean residual error of all three channels are -0.13 m with standard deviation of 0.33 m. The cumulative range error is depicted in Figure 4.12 with 90% of error in all three channels is better than 0.60 m approximately.

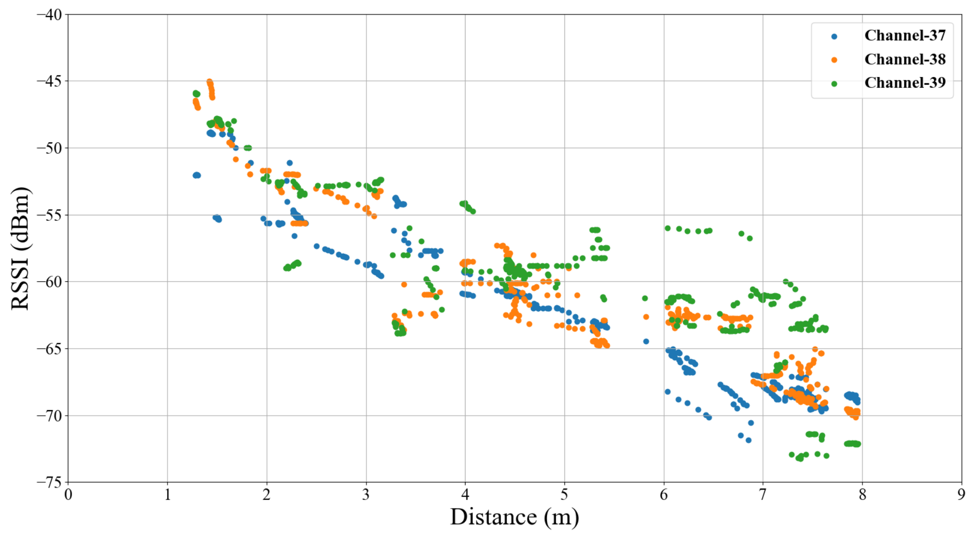


Figure 4.10: Predicted distance output of CNN model on randomly selected subset of the same test dataset at the empty lab environment.

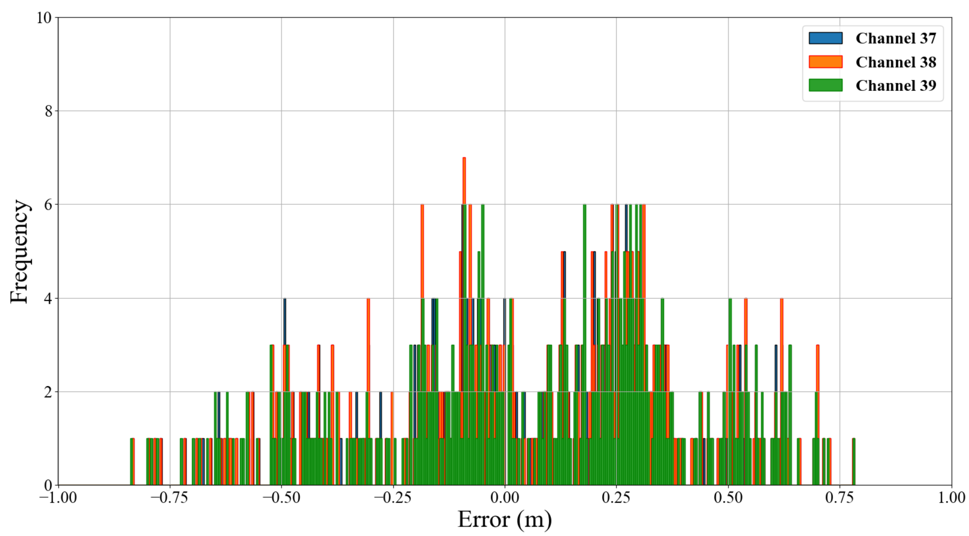


Figure 4.11: Residual error of CNN model between predicted and reference range at the empty lab environment.

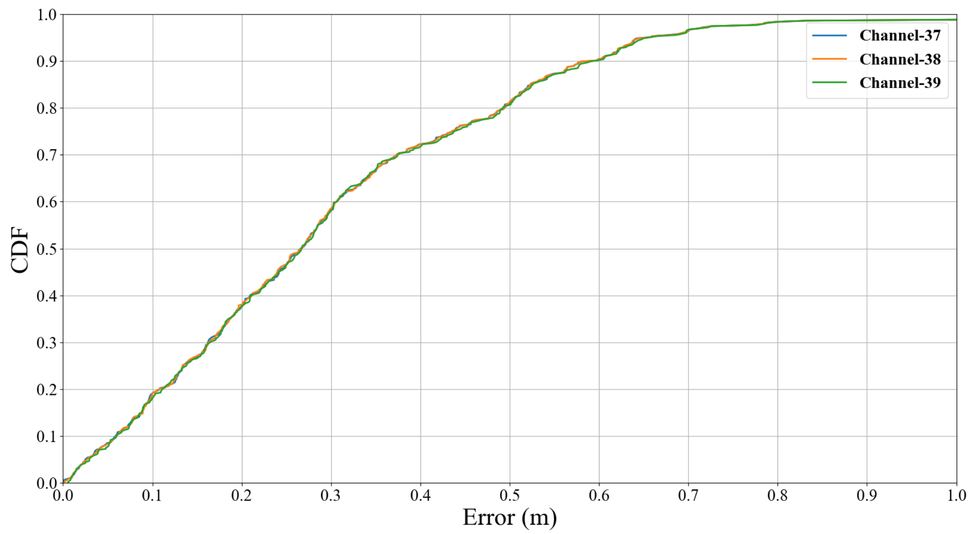


Figure 4.12: Cumulative distance error of each primary channel.

4.4.3 Random Forest versus Convolutional Neural Network

Random forest and Convolutional neural networks are both powerful ANN algorithms and differ in architecture and learning process. Random forest is a decision tree-based model which learns by splitting the data into smaller groups and combining each tree's predictions to make the final conclusion. In contrast, CNN is considered a deep learning model that uses convolutional layers to extract features from the input data and then uses fully connected layers to make a final prediction. The advantage of random forest is that it is generally faster to train and requires less computational resources than CNN. Moreover, random forests can handle large datasets with more features using parallelization. In contrast, CNN consumes more training time and computational resources due to its deep network architecture and the need for feature extraction. It can be noticed from Section 4.4.1 and 4.4.2 that the random forest test result predictions are comparatively better than the CNN results. In this case, despite the relatively smaller training dataset (2130 training samples), the random forest model performed reasonably well than CNN, which is primarily suitable for larger datasets with a spatial structure.

It is essential to mention that both RF and CNN architectures require hyperparameter tuning to improve their results. This may involve adjusting parameters such as the number of trees in the forest or the minimum number of samples needed to split a node for the random forest or adjusting the number of convolutional and dropout layers for feature extraction and regularization to overcome overfitting in CNN. However, the advanced hyperparameters tuning is not carried out in this work since the objective was to verify the benefits of using neural networks over classical methods to address the problem mentioned in section 1.4 of Chapter 1.

4.5 Summary

This chapter provides detailed explanations of two ANN algorithms, namely random forest and convolutional neural network, covering aspects such as their training environment, data collection, and the specific procedures employed for training. A comparison and analysis of the residual errors were conducted using test samples. Furthermore, the chapter discusses the advantages and conditions required for optimal application results. The chapter concludes by recommending hyper-parameter tuning as a practical approach to achieving an optimal model configuration.

Chapter 5

Experimental Results and Analysis

5.1 Introduction

The two artificial neural network techniques described in Chapter 4 were first tested on an empty open lab environment. After having a successful demonstration of range estimation from RSSI values on the test dataset, the ANN approach was applied to different field data.

This chapter focuses on evaluating the effectiveness of proposed algorithms, namely Random Forest and Convolutional Neural Network, in a typical office environment. The chapter begins by providing a detailed description of this new environment, including the test setup used for data collection and the specific test scenario under which the data were gathered. Both algorithms share similarities as they use RSSI values from the three BLE advertising channels to estimate range and proximity.

The chapter presents the estimated range and proximity results obtained from real field data for both systems. Furthermore, a comprehensive analysis is conducted on the range residual error, using the algorithm's output and UWB range values. This analysis aims to evaluate the performance of the algorithms when applied to field data. Throughout the chapter, the evaluation and analysis of the proposed ANN techniques in the specific office environment sheds light on their performance and suitability for real-world applications.

To assess the performance of the proposed algorithms in a larger and more demanding environment, experiments are conducted in two distinct office areas located on the third floor of a multi-story university building. These two office environments are:

END-309 : Rectangular size room with compact space (Section 5.2).

END-313 : Medium size square room, which is adjacent to END-309 (Section 5.3).

5.2 END-309 Office Environment

Figure 5.2 illustrates the inside view of END-309 office space. Unlike the empty training lab area (discussed in Section 4.2.1), this office room is equipped with chairs, tables, shelves, cubicles, metallic stands, and also users moving around, making it a more dynamic environment to conduct experiments.

Figure 5.1: The inside view of END-309 office area.

The dimension of this room (length x breadth) is approximately 9.6 m x 7.90 m. The 2D plan view of the office area, with 27 RPs (Reference Points) of known locations identified, is illustrated in Figure 5.2. The DWM1001-DEV module in this setup serves as a BLE source and a UWB transceiver. The two co-located receivers include another DWM1001-DEV module and one Nordic nRF52840 DK BLE module. The advertising and scan intervals of BLE and the UWB data rate remained unchanged from those used in the empty environment case.

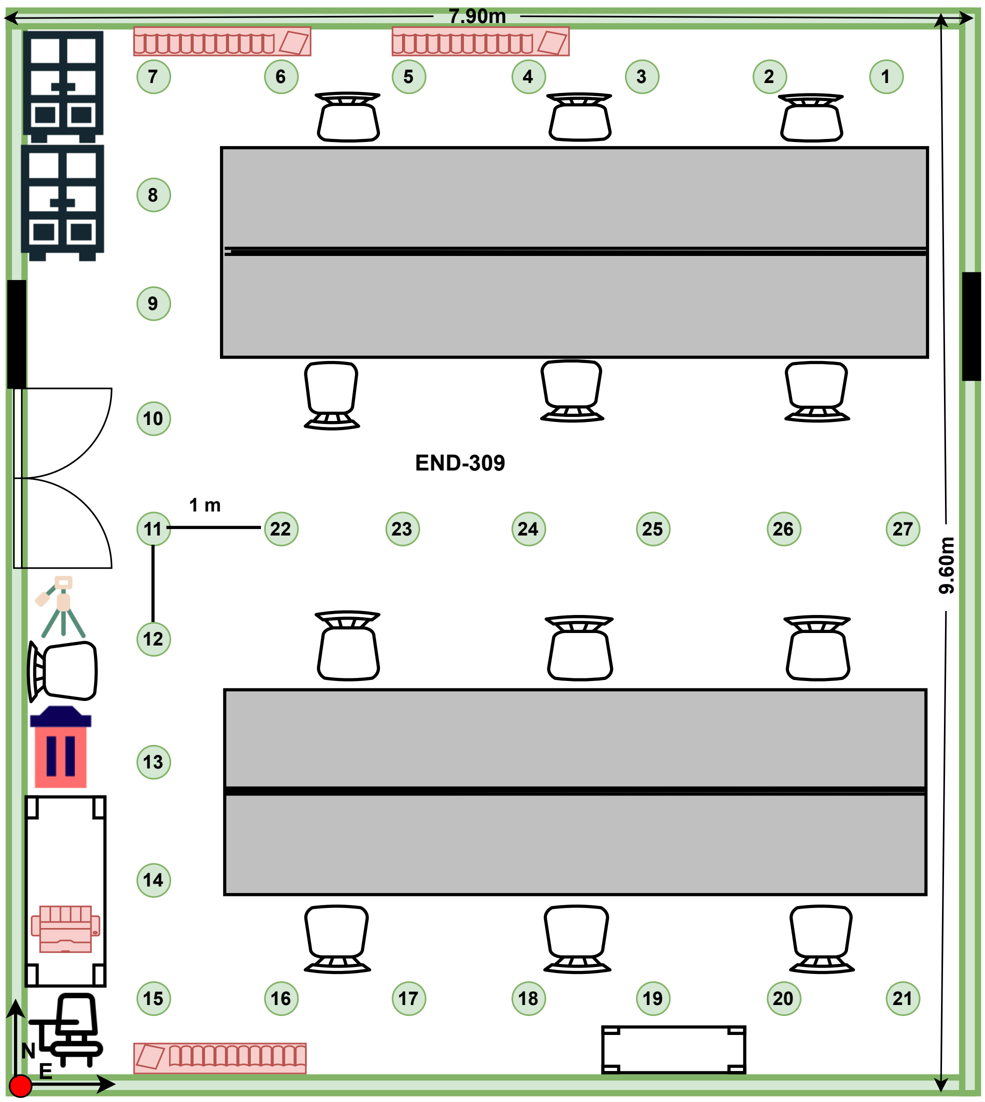


Figure 5.2: The blueprint of the END-309 office evaluation area.

To assess the performance of the trained ANN models without further training, evaluations are conducted in five distinct scenarios within the first office room. The five scenarios are:

Static case: Stationary conditions, where users remained in fixed positions.

Dynamic case: A source user who moves towards other static users at LOS.

Blind points: Areas within the office room where the BLE signal reception is obstructed.

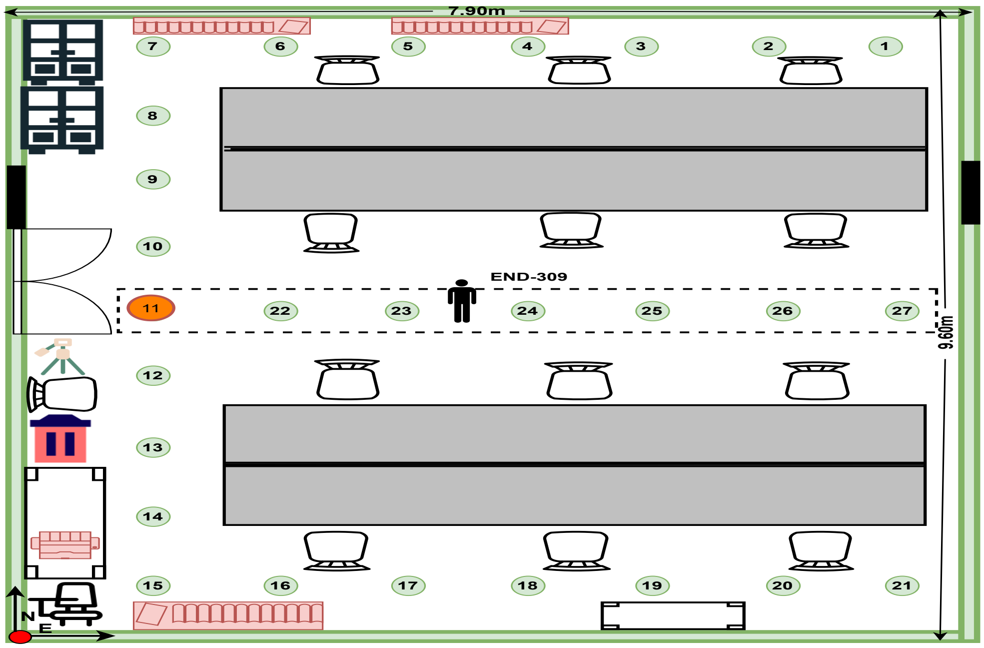
A new user: Detecting presence of a new BLE user within the room and estimating range.

Proximity detection: Classifying proximity of a user based on a threshold distance.

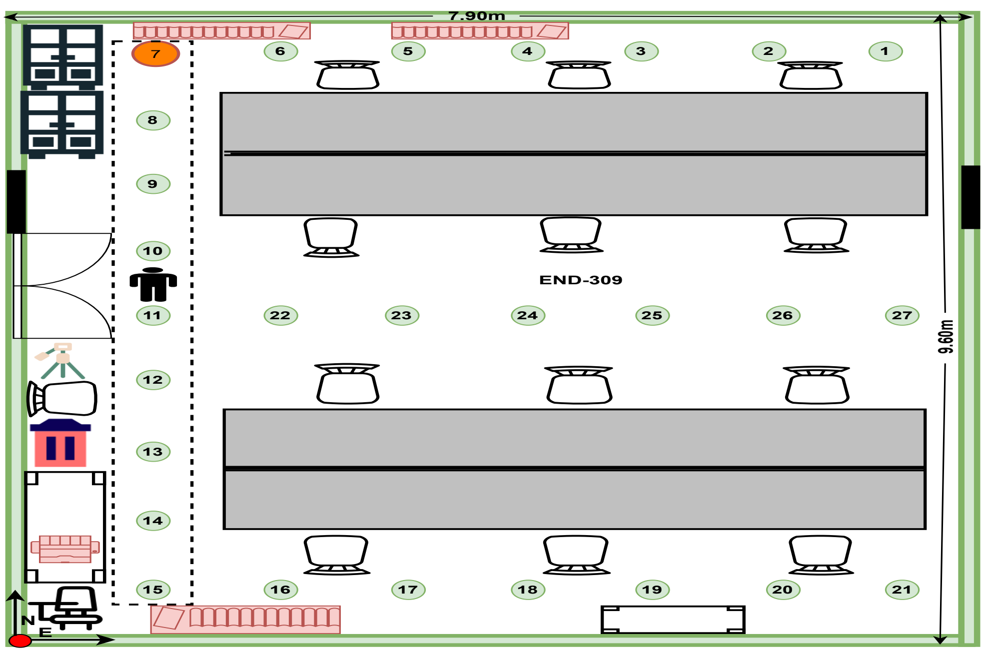
5.2.1 Range Estimation in Static Case

In the static scenario, the receiver remains stationary at each reference point (RP) while collecting measurements from the fixed transmitter. Initially, the transmitter is located at RP 11 and the receiver at RP 22, one metre from each other. The receiver switches between six successive RPs (from RP 22 to RP 27) in the East-West direction and collects RSSI and UWB range samples line-of-sight at each RP. This data collection process is repeated over 200 test epochs, allowing for data filtering or averaging before being processed by ANN. Due to the dynamic nature of the office environment and BLE signal characteristics, the RSSI measurements are subjected to fluctuations and contain outliers or errors. To mitigate this, the RSSI values undergo post-processing using a 1D-Kalman filter, effectively removing noise and outliers. The filtered RSSI values are then input into the trained ANN models for predicting distance. It is important to note that the ANN model relies on clean input data to ensure optimal functionality and produce accurate predictions.

Similarly, the test user measures RSSI and UWB range data at eight additional reference points in the North-South direction, with the transmitter positioned at RP 7. However, in this case, the measurement path is narrower, and there is a higher likelihood of signal obstruction from nearby objects. It is worth mentioning that, the ANN models which are previously trained only take three advertising channel RSSI values for distance estimation, while reference distance serves as ground truth to compute the residual error. The two static data collection scenarios and the test user measurement paths are illustrated in the 2D plan view of Figure 5.3. The histogram plot in Figure 5.4 shows the distance estimation errors of the RF and CNN models, evaluated in the East-West direction at six reference points for three BLE primary channels.



(a) **Test user in scenario #1**



(b) **Test user in scenario #2**

Figure 5.3: Blueprint of the measurement paths in Scenario #1 (East-West) and #2 (North-South) in the static case. The orange reference point represents the position of the transmitter in each scenario.

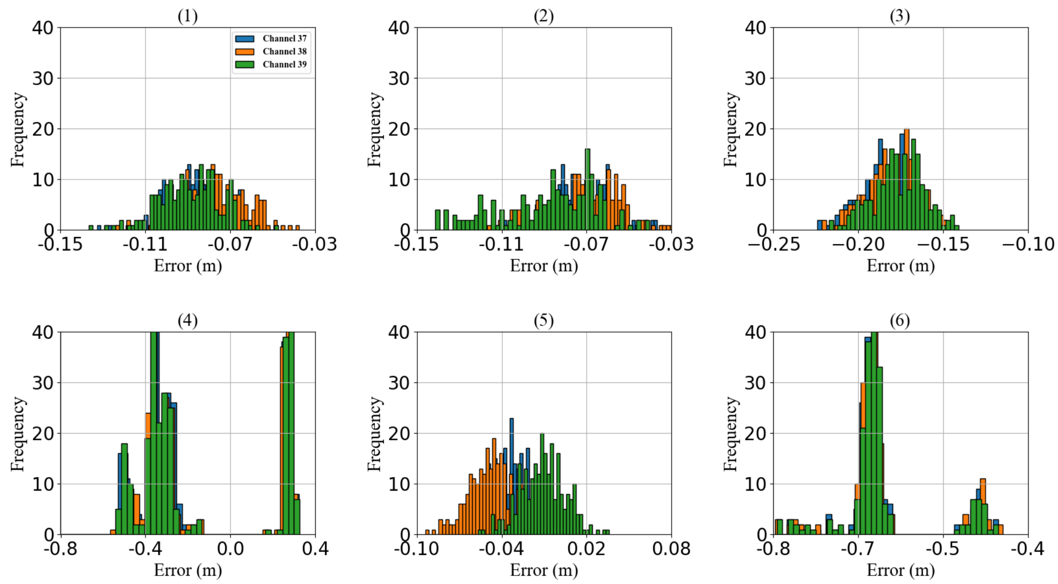


Figure 5.4: Histogram of distance estimation errors using RF model in the East-West direction at LOS in six reference points. Each subplot corresponds to residual error measured at (1) 1 m, (2) 2 m, (3) 3 m, (4) 4 m, (5) 5 m, and (6) 6 m respectively in sequence for each BLE advertising channel.

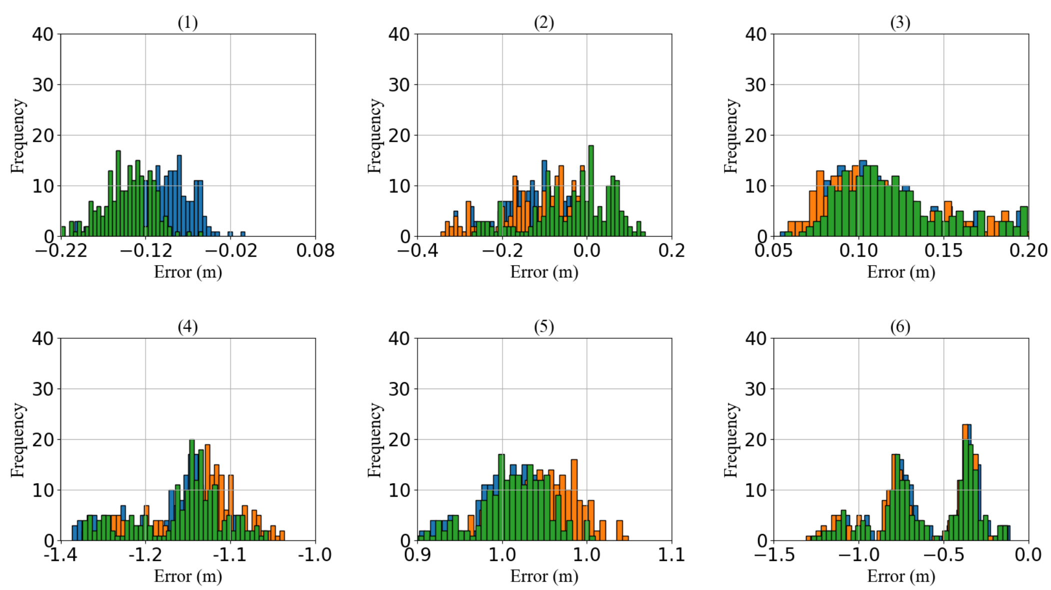


Figure 5.5: Histogram of distance estimation errors using CNN model in the East-West direction at LOS in six reference points. Each subplot corresponds to residual error measured at (1) 1 m, (2) 2 m, (3) 3 m, (4) 4 m, (5) 5 m, and (6) 6 m respectively in sequence for each BLE advertising channel.

The histogram results illustrate a close agreement between the RF and CNN models, particularly at short ranges i.e., 1 m, 2 m, and 3 m. Both models exhibit a range error of less than 25 cm within this range, with the estimated range often longer than the reference known distance.

The CNN method exhibits a range error between 1-1.5 m for distances of 4 m (case d) and 6 m (case f). It is important to consider that due to the random nature of RSSI values, especially between distances of 3 m and 6 m, the model's range estimation may deviate from the expected value. The dynamic environment presents a greater challenge for the model to determine accurate range estimates at these distances. Interestingly, in cases (d) and (e), the RF method outperforms the CNN method, indicating the performance of the regular ML algorithm over the deep neural network in this particular scenario. Table 5.1 summarizes the mean range estimation error and the error percent for three advertising channels for the RF and CNN models at six reference points.

Table 5.1: Mean estimated distance and error percent of the trained ANN models at LOS peer-to-peer distances in the East-West direction.

Distance (m)	RF		CNN	
	Estimated mean distance (m)	Error percent (%)	Estimated mean distance (m)	Error percent (%)
1	1.08	8	1.13	13
2	2.07	3.50	2.09	4.50
3	3.18	6	2.88	4
4	4.15	3.75	5.17	29.25
5	5.03	0.60	4.01	19.80
6	6.62	10.33	6.60	10

Similarly, Figure 5.6 and Figure 5.7 shows the histogram of distance estimation errors in BLE primary channels in the North-South direction using RF and CNN models for distances up to 8 metres. The RF method demonstrates a range error concentrated within 0.10 m, while the CNN approach shows an error of less than 0.15 m for peer-to-peer distances up to 2 m. Even at a user distance of 3 metres, the error remains below 0.4 m. However, as the user moves farther from the source, the range error increases to metres. As observed by both models, the range error exceeds one metre at a user distance of 4 metres. For user distances between 5-7 metres, the range error is still less than 1 metre using the RF method and less than 1.2 metres using the CNN approach. At a highest user distance of 8 m, the RF model estimates distances between 5-8 metres, while the CNN model estimates between 6-8 metres.

The RSSI values fluctuate more randomly as the test user moves farther from the transmitter, resulting in relatively larger predicted distances due to increased attenuation. This randomness is also observed here between 3-6 metres. Table 5.2 presents the estimated mean distance and the error percent for the RF and CNN models. It is noteworthy to compare the results to those reviewed in the literature. Gadhgadhi et al. (2019) showed distance estimation from RSSI values using neural network techniques (Gadhgadhi et al., 2019). Their method achieved an error of less than 1 m using a small training dataset of 17 mean RSSI values up to 10 metres. These results are comparable to the findings presented in this study with the RF ML algorithm. Moreover, the short distance proximity estimation is better compared to the results mentioned in sections 3.3.1 and 3.3.3. The RF-based approach showed better performance (less than 1 m mean error) than the neural network model (CNN) up to a distance of 6 m using a separate test dataset. The method shows an error spread of 1-3 m at user distance of 8 m. In contrast, the CNN method shows error spread of 2 m for distances up to 8 m for the same dataset.

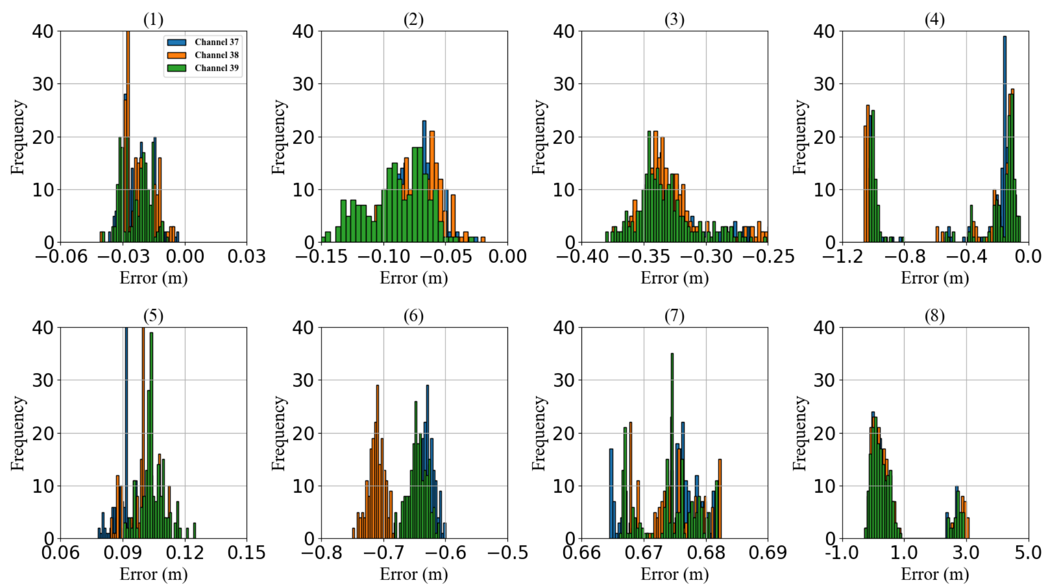


Figure 5.6: Histogram of distance estimation errors using RF model in the North-South direction at LOS in 8 reference points. Each subplot corresponds to residual error measured at (1) 1 m, (2) 2 m, (3) 3 m, (4) 4 m, (5) 5 m, (6) 6 m, (7) 7 m, and (8) 8 m respectively in sequence for each BLE advertising channel.

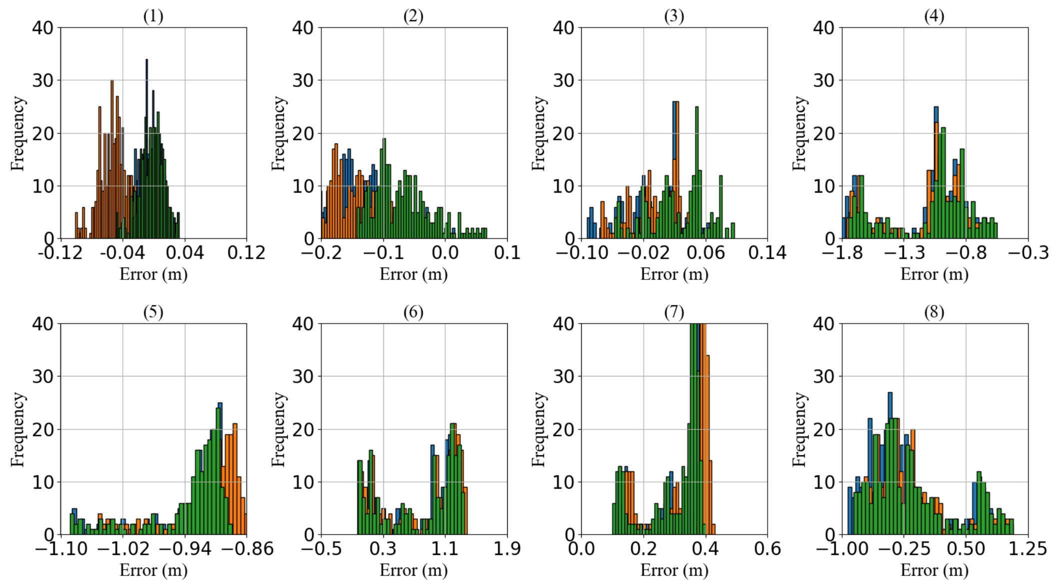


Figure 5.7: Histogram of distance estimation errors using CNN model in the North-South direction at LOS in 8 reference points. Each subplot corresponds to residual error measured at (1) 1 m, (2) 2 m, (3) 3 m, (4) 4 m, (5) 5 m, (6) 6 m, (7) 7 m, and (8) 8 m respectively in sequence for each BLE advertising channel.

Table 5.2: Mean estimated distance and error percent of trained ANN models at LOS peer-to-peer distances in the North-South direction.

Distance (m)	RF		CNN	
	Estimated mean distance (m)	Error percent (%)	Estimated mean distance (m)	Error percent (%)
1	1.02	2	1.02	2
2	2.08	4	2.11	5.50
3	3.33	11	2.99	0.33
4	4.46	11.50	5.11	27.75
5	4.90	2	5.92	18.40
6	6.66	11	5.21	13.16
7	6.32	9.71	6.68	4.57
8	7.40	7.50	8.15	1.86

5.2.2 Range Estimation in Dynamic Case

In the dynamic case, scenario #1 is considered again. The test user gradually moves at an approximate rate of 1 metre per minute from reference point 27 towards the transmitter, eventually reaching RP 23. This movement covers a distance of 4 metres, with the user stopping 2 metres away from the source. During this movement, 1200 samples for both RSSI and UWB measurements are collected. After comparing the measurements based on PC logged time, 800 samples are approximately selected that exhibit same time stamps. The dynamic RSSI values are input to the ANN models for predicting the user trajectory path. These predicted paths are compared with the ground truth obtained from UWB measurements. Figures 5.8 and 5.9 depict the dynamic user trajectory paths predicted by the ANN models using the measured RSSI values. As the user approaches the source, the predicted distances show a decreasing trend, reflecting the improvement in signal strength and resulting in shorter distance estimations by the models.

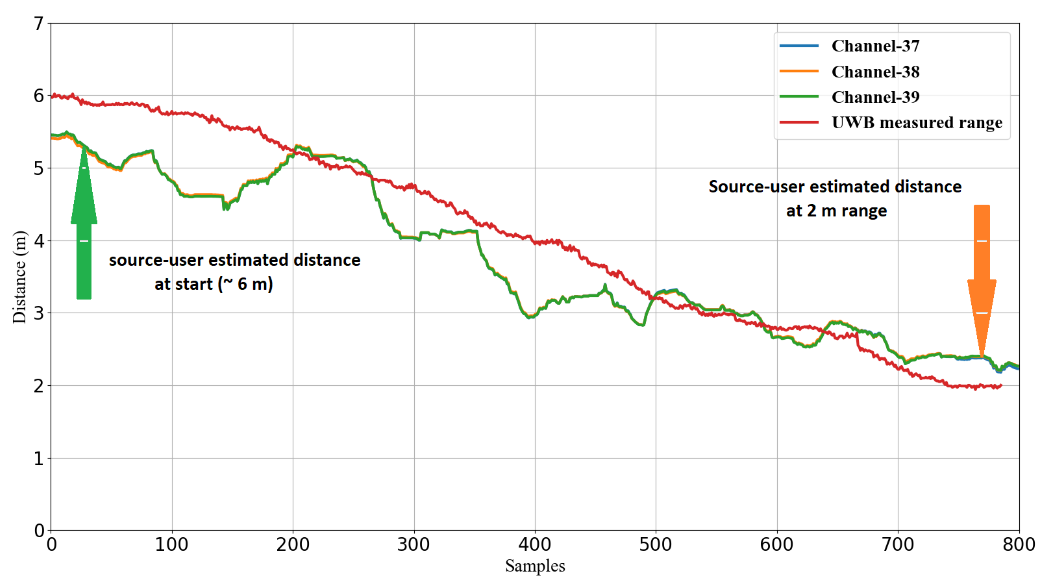


Figure 5.8: Comparison of estimated dynamic user trajectories in the East-West direction using RF model with UWB ground truth for BLE advertising channels.

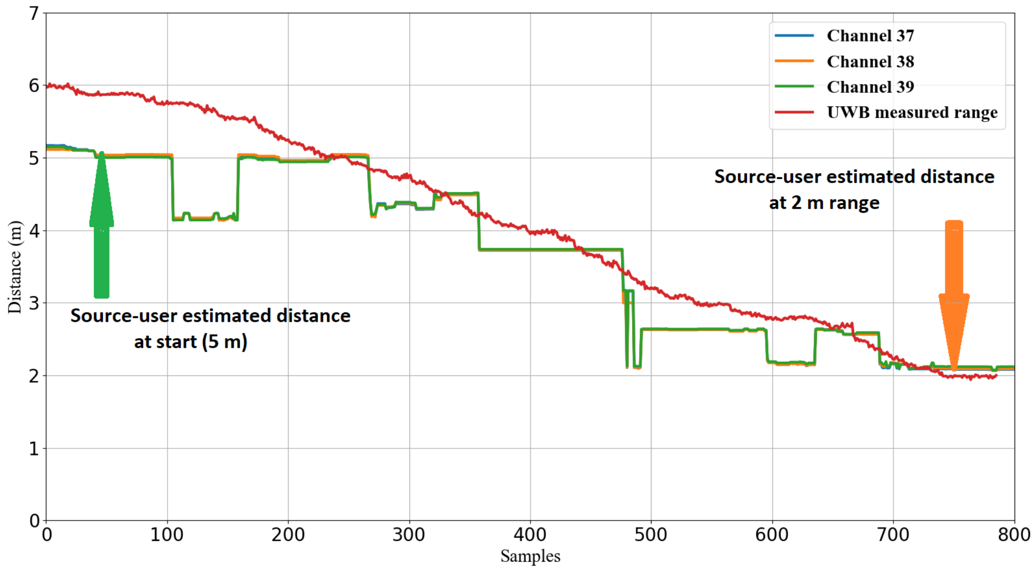


Figure 5.9: Comparison of estimated dynamic user trajectories in the East-West direction using CNN model with UWB ground truth for BLE advertising channels.

The RF approach shows smooth trajectory of the user movement, closely following the reference curve. On the other hand, although the CNN method may not demonstrate a smooth trajectory, it displays step transitions when the user is between 4-2 metres away from the source. However, when the user reaches a distance of 2 metres from the transmitter, both the ANN models accurately estimate a proximity that is close to 2 metres, as shown in the graphs. This test demonstrate that the ability of the AI techniques to project the trail of a slow moving user path based on a series of RSSI values.

5.2.3 Range Estimation in Blind Locations

In this scenario, the trained model is evaluated in highly constrained areas within the same office environment, specifically designated as blind test points. These blind test points are randomly selected cubicles separated by metallic and wooden sheets. The test assume that two users work in separate cubicles in an office environment, collecting and logging each other's RSSI values. For this specific test, the test user location is fixed at Desk number L, while measurements are collected from five different blind points (Desk No. A, C, I, G, K). Figure 5.10 provides a visual representation of the blind points and the test user location. It is important to note that in this case, five different DWM1001 DEV modules are placed at the center of each identified blind cubicle simultaneously. This placement signifies the presence of different users within the office area and their active communication with the test user. The test user collects RSSI and UWB measurements for approximately 5 minutes to gather more than 200 samples for each measurement type. The collected RSSI values are then fed to the ANN models to predict the range between the test user and other users in each desk.

The users being in highly blocked areas within the cubicle walls, direct ground truth distance between the transmitter and receiver is calculated using tap measure at right angles to a line on the floor. To assess the accuracy of the UWB measured range and the output from the ANN RSSI-to-range model, they are compared against the reference known distance. Figure 5.11 illustrates the measured UWB range and predicted estimated distances by two ANN models at three blind locations. The severe NLOS conditions affect the quality of the measurements resulting in a longer travel time for the UWB signal and an increased computed distance, as shown in the "UWB" column. From the figure, several outliers can be inferred. Hence a 1D Kalman filter is used to post-process the UWB data and detect NLOS scenarios. This imposes restrictions on the smart user on the reliability of the UWB data to be used for training in a new environment without filtering.

The mean filtered UWB range values at Desk A, C, and I are 8.97 m, 9.43 m, and 6.99 m, respectively. Due to attenuation, the ANN models predict longer distances for the RSSI values. In the first case, the RF method estimated a range between 6.8-7.5 m, while the CNN method showed a range between 7.4-8 m. Similarly, for the other two cases, the estimated distances by the ANN models closely align with the measured UWB range, with an accuracy better than 2 metres. However, in all three cases, the UWB measurements deviated from the true distances by approximately 3 metres, as these locations were highly constrained areas. Consequently, the maximum mean error between the CNN estimated distance output and the reference known distance varied within 2.5-3 m. On the other hand, the RF method achieved a mean error better than 2 metres (Table 5.3). It is important to note that the comparison presented here focuses on the UWB and ANN estimated ranges w.r.t ground distance, omitting the residual range error for simplicity.

The test results highlight the intelligence of the trained AI model in effectively predicting reliable distances from the input RSSI data even in highly-constrained complex environments.

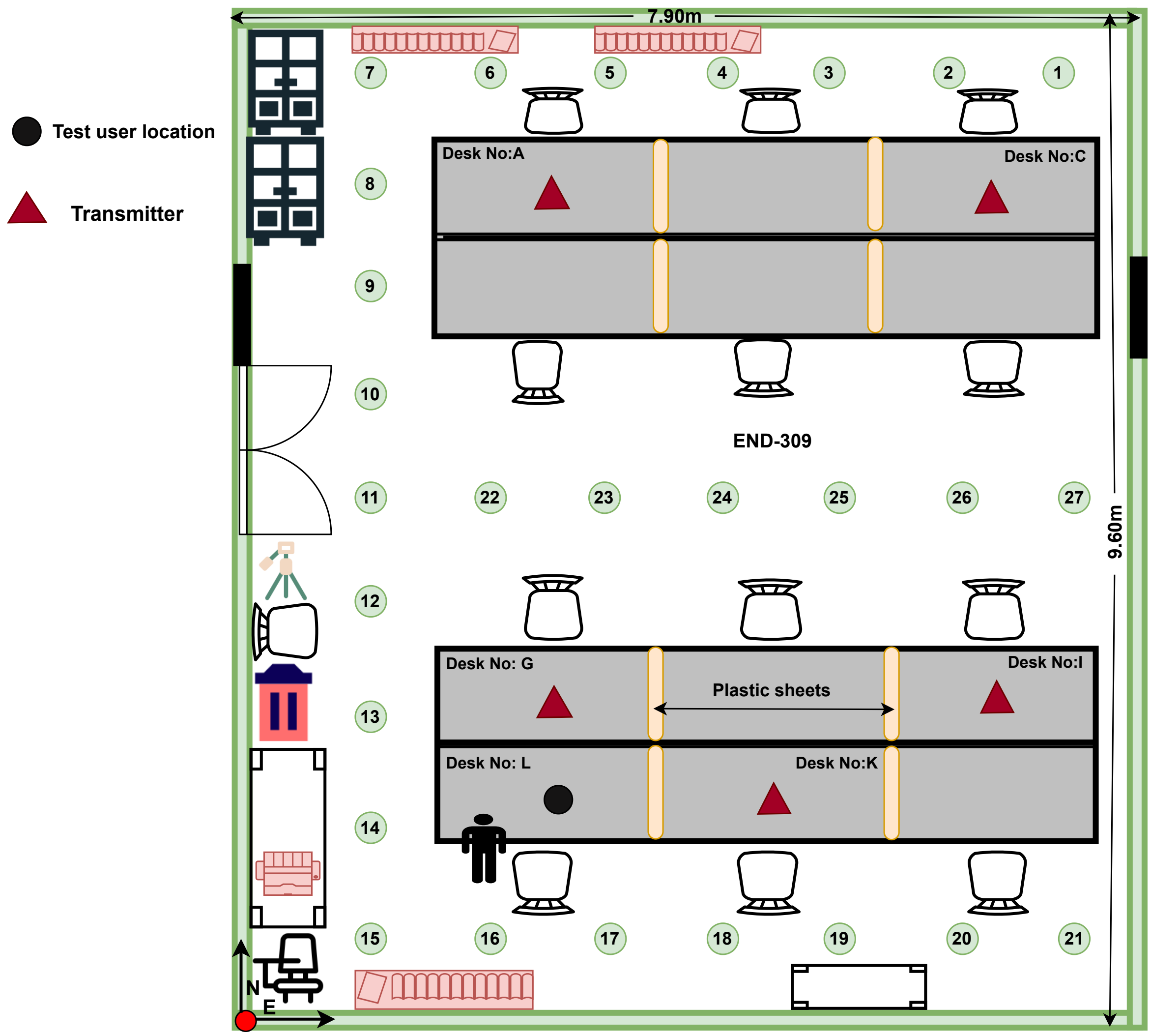


Figure 5.10: Schematic of five blind test point locations represented by Desks having Decawave transmitters in the END-309 office area. The test receiver is located at Desk L.

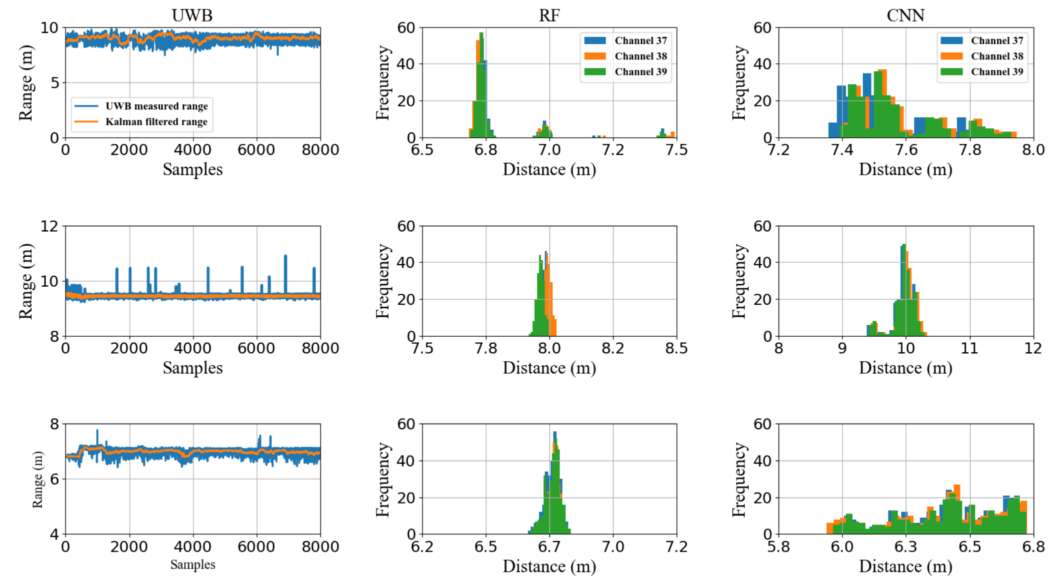


Figure 5.11: Comparison of UWB measured range and ANN-estimated distance at three blind locations (a) Desk A (b) Desk C and (c) Desk I (each location presented in one row).

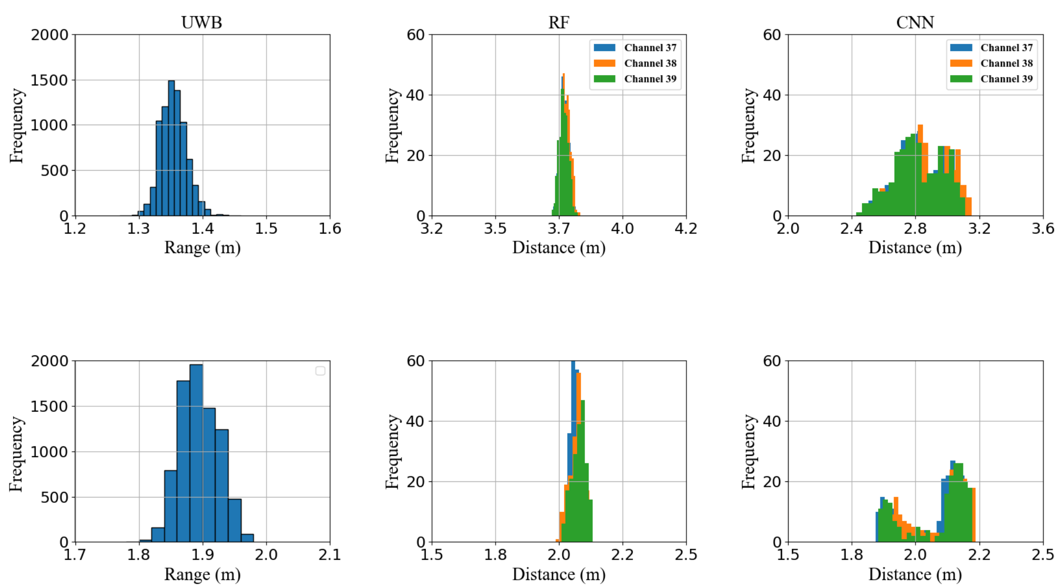


Figure 5.12: Comparison of UWB measured range and ANN-estimated distance at two nearby blind locations with different dividers. (a) Metallic body with a wooden wall divider, true reference range: 1m. (b) Plastic sheet divider between user and transmitter, true reference range: 2m (each location presented in one row).

Figure 5.12 depicts the case when two different materials obstruct the transmitter and receiver. In case (a), desks G and L are separated by a wooden wall with a metallic strip running within them. The actual distance between the two devices was approximately 1 m. Studies have shown that metallic objects can interfere with UWB signals, leading to significant range errors (Wang et al., 2018); the UWB range deviated approximately 30-40 cm from the true distance. In contrast, the RSSI to distance output showed a longer range between 2.4-3.8 metres,resulting in an overall difference of 2.5 metres compared to the UWB range.

In case (b), a thin plastic sheet separates desks G and K, and the distance between the two devices was approximately 2 m. The histogram of the UWB range is concentrated around the mean value of 1.92 m, and both the ANN models estimate a range of more than 2 metres. This behavior can be attributed to the ability of both UWB and BLE signals to pass through the thin plastic sheet, acting as a transparent medium at short distances (Flammini et al., 2009), in line with the principles of radio wave propagation and the properties of plastic materials. Table 5.3 summarizes the distance estimation results obtained from the ANN models and the UWB range measurements compared with the known distance.

Table 5.3: Comparison of UWB mean range and mean estimated distance output from trained RF and CNN models with reference distance at five blind locations within an office environment.

Reference	UWB		RF		CNN	
Distance (m)	KF mean range (m)	Std	Estimated mean distance (m)	Std	Estimated mean distance (m)	Std
5.39	8.97	0.20	6.78	0.15	7.56	0.14
7.03	9.43	0.01	7.98	0.01	9.96	0.18
4.01	6.99	0.07	6.71	0.03	6.46	0.21
1	1.35	0.02	3.73	0.03	2.83	0.15
2	1.92	0.03	2.07	0.06	2.09	0.11

(Row 4) Metallic and wooden divider between Tx and Rx.
(Row 5) Plain plastic sheet between Tx and Rx.

5.2.4 A Random New BLE User Sitting in a Location

In this case, a person with a DWM1001-DEV device as a BLE user enters the END-309 office and sits in a chair at RP 27 for some time. The test receiver at Desk No: G detects it and receives RSSI in three advertising channels along with UWB measurements. This case aims to demonstrate a typical scenario where an AI model detects a person's presence inside the room and estimates the distance between the person and the test subject. Figure 5.13 shows the sitting position of the user. The measured UWB range and the ANN estimated distance results for all three BLE channels are depicted in Figure 5.14. The UWB plot infers the presence of a few instances of multipath, as both the transmitter and receiver were in NLOS. Subsequently, a Kalman filter is applied to smooth the signal. The mean value of the filtered range is 5.87 m. Similarly, the RSSI values also exhibit deviations in the computed distance, as indicated by the ANN model. However, despite the blockage in the signal path, the UWB range and the RF estimated output show comparable results. The RF model estimates a distance between 5-6 metres, with the output around 5.8 m. On the other hand, the CNN method estimates a slightly shorter separation, ranging between 4.2-5.2 m. Table 5.4 provides a comparison between the distance estimated by the two ANN models and the UWB range with the known reference distance of the BLE user.

Figure 5.13: Location of random BLE user sitting in a place exchanging RSSI with other users.

Table 5.4: Comparison of UWB mean range and mean estimated distance of ANN models with reference distance of a random BLE user in an office environment.

Reference	UWB		RF		CNN	
Distance (m)	KF mean range (m)	Std	Estimated mean distance (m)	Std	Estimated mean distance (m)	Std
4.80	5.87	0.02	5.69	0.21	4.74	0.18

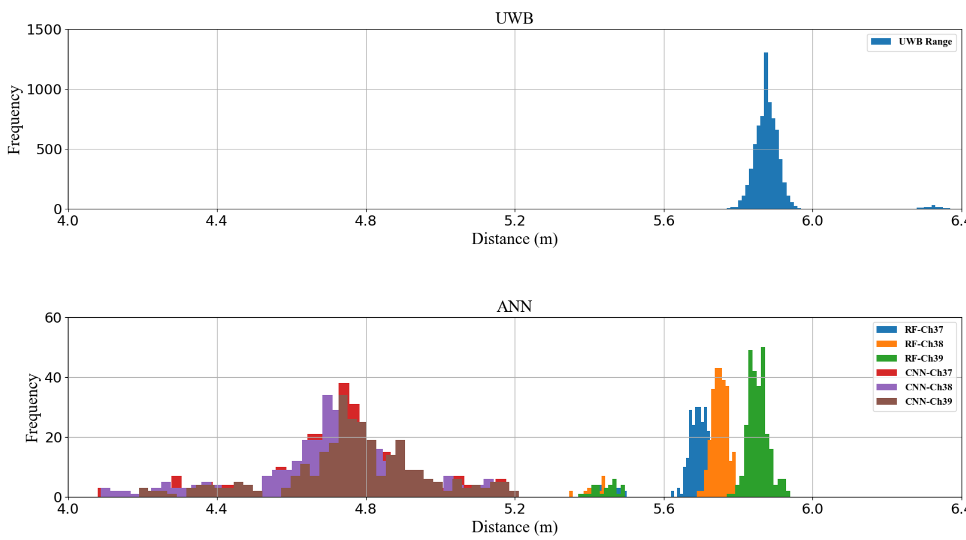


Figure 5.14: Comparison of measured UWB range and distance estimated by the ANN models for a random user sitting at a specific location relative to the test user.

In this particular case, the maximum range error between the ANN model output and the known distance is less than 1 metre.

5.2.5 Proximity Classification at Shorter Distance

In the current environment, the final category of tests focuses on proximity classification. This test involves evaluating the trained model's ability to classify instances when a mobile user is detected in close proximity to the test user based on a threshold value set by the model. The receiver remains fixed at one position while the source moves from distances greater than 2.5 metres to as close as 1 metre from the receiver. This combination is repeated in four different locations within the room under the same LOS condition. This specific test is similar to the dynamic case, but the distance between the source and receiver is limited to less than 3 metres. The measured RSSI and UWB measurements are collected over a relatively short period of 1-2 minutes. The shorter duration is because the user moves, and the distance between the source and receiver is limited. The RSSI samples are filtered, and corresponding UWB-matched instances are retained. The RSSI values as input to the ANN model give distance output and are compared with UWB ranges to give final classification results. It is to be noted that, here, only the RF model results are shown for the classification, and the other ANN model (CNN) results are not discussed.

Table 5.5 displays the proximity classification results of the trained RF model for four cases, using the estimated mean output of separate BLE channels and UWB range at a 1.5 m threshold. For classification, all instances of the ANN model continuous distance output values less than 1.5 m are considered 1.5 m (Label 1), indicating close proximity. On the other hand, distance values between 1.5-3 m are considered 2 m (Label 2), depicting relatively higher proximity between the users. The diagonal cell shows the results of the location of the mobile device predicted by the classification model correctly. The inaccurate prediction results are represented by the non-diagonal cells.

Table 5.5: Proximity classification results of RF method at LOS for four locations with a 1.5-metre threshold.

(1)

		Predicted: Yes	Predicted: No
Actual	Yes	82	43
	No	33	71

(2)

		Predicted: Yes	Predicted: No
Actual	Yes	85	38
	No	43	97

(3)

		Predicted: Yes	Predicted: No
Actual	Yes	61	37
	No	24	147

(4)

		Predicted: Yes	Predicted: No
Actual	Yes	86	22
	No	20	113

The performance of the ANN model is evaluated based on the number of missed detection and false alarm. The dataset comprises 1002 observations from four cases, with 434 samples less than 1.5 m and 568 samples more than 1.5 m. Table 5.6 provides a breakdown of the true and false classification samples processed by the ANN model.

Table 5.6: Summary of RF-based classification results for missed detection and false alarms at a 1.5 m threshold across four locations.

Status	**Samples**	**Percent (%)**
Correct identification	742	74.05
Missed detection (<1.5 m)	120	11.97
False alarm (>1.5 m)	140	13.97

The RF-based classification model achieves a 74% accuracy rate in correctly categorizing user epochs, regardless of whether they are below or above the 1.5-metre threshold. Specifically, 12% of the total observations, equal to 120 instances, were identified as missed detections (i.e., distances less than 1.5 metres), while 14% (140 samples) were falsely categorized as greater than 1.5 metres (high proximity). Although the latter case is more critical, comprising 13.97% of the total observations, it still represents a relatively small percentage of the overall dataset.

5.3 END-313 Office Environment

The END-313 office room serves as a second test environment for assessing the robustness and reliability of the ANN models in determining the range between users. Despite its smaller size compared to END-309, with approximate dimensions (length x breadth) of 9.6 m x 6.3 m and a square shape, it presents challenges due to various items such as books, boxes, chairs, and tables, restricting the available open space. The proximity of END-313 to END-309 enables the evaluation of scenarios involving two BLE users (herein students) situated in two different environments and exchanging RSSI values. Moreover, the wide dry wall separating the two rooms allows for analyzing the variability and strengths of UWB and BLE RSSI signals across rooms. The drywall is assumed to be composed of a combination of materials, including metal, wood, and concrete. Figure 5.15 illustrates a 2D blueprint of the END-309 and END-313 office environments together. In the adjacent room, eight reference points and several desks are designated as blind test points for evaluation.

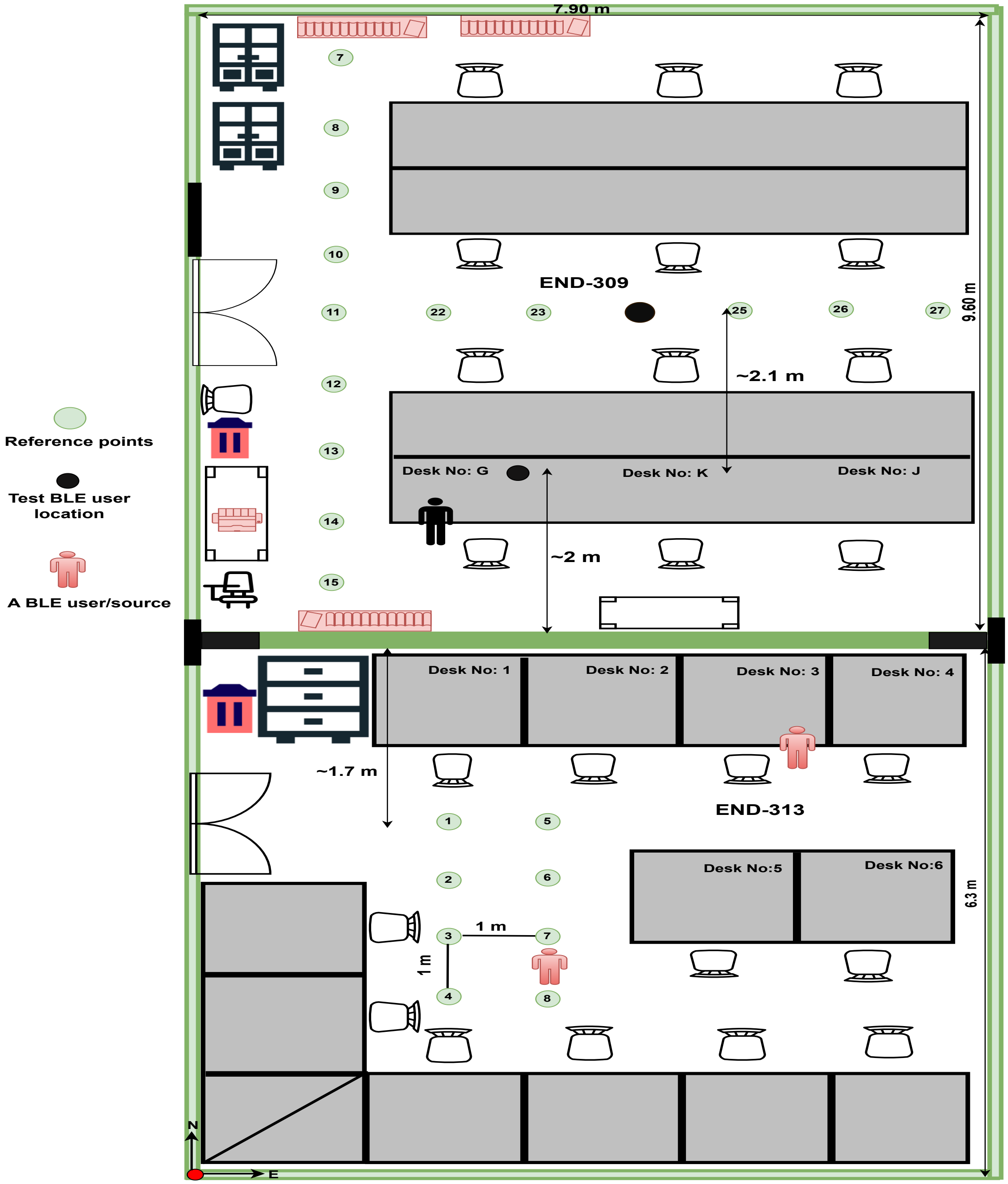


Figure 5.15: BLE proximity estimation scenario in two adjacent office environments (END-309 and END-313) with BLE users exchanging each other RSSI values.

The two specific environments described in Section 5.2 are again presented section wise.

(i) **Static case**

(ii) **Blind points**

5.3.1 Range Estimation in Static Case

The test case assume that both the test subject and the source at their respective positions remain stationary for certain duration. In this case, the test user (receiver) location is fixed at Desk number G in END-309, while the BLE source (only one DWM1001-DEV module), located in END-313, is positioned at different reference points (RP1 to RP4) at various time intervals. These RPs are 1 metre apart from each other. It is important to note that the exact thickness of the wall separating them is unknown (assumed 0.20-0.30 m wide based on visual inspection), and the presence of nearby wooden desks and PC monitors in END-313 causes severe blockage, resulting in signal degradation. Furthermore, the test user is approximately 2 m away from the thick wall, while the first reference point (RP1) is approximately 1.7 m away from the same wall. The test user collects more than 200 epochs of data containing RSSI and UWB measurements. For similar reasons described in Sections (2.1.3, 3.7 and 5.2.1) the RSSI values are post-processed using 1D Kalman filter.

The filtered RSSI values are used as input to the ANN methods to estimate ranges. Figure 5.16 shows the comparison of UWB measured range with the distance estimated from RSSI values using two ANN methods (RF and CNN) at four reference points. The mean value of the UWB range at the first reference point is 4.22 m. It can be observed that the UWB measured range closely aligns with the manually calculated distance of approximately 4 m, taking into account the relatively substantial large dry wall. The minor disparity between the measured and computed range at centimetre-level is attributed to the attenuation caused by the wall. These results align with previous findings that indicate UWB signals experience attenuation and distortion when propagating through walls due to their dispersive properties (Muqaibel et al., 2005). Moreover, as the source moves to subsequent reference points, each 1 m apart until reaching RP 4, resulting in approximately a total length of 7 m from the test user, the measured range increases by roughly 1 metre while maintaining a consistent level of deviation.

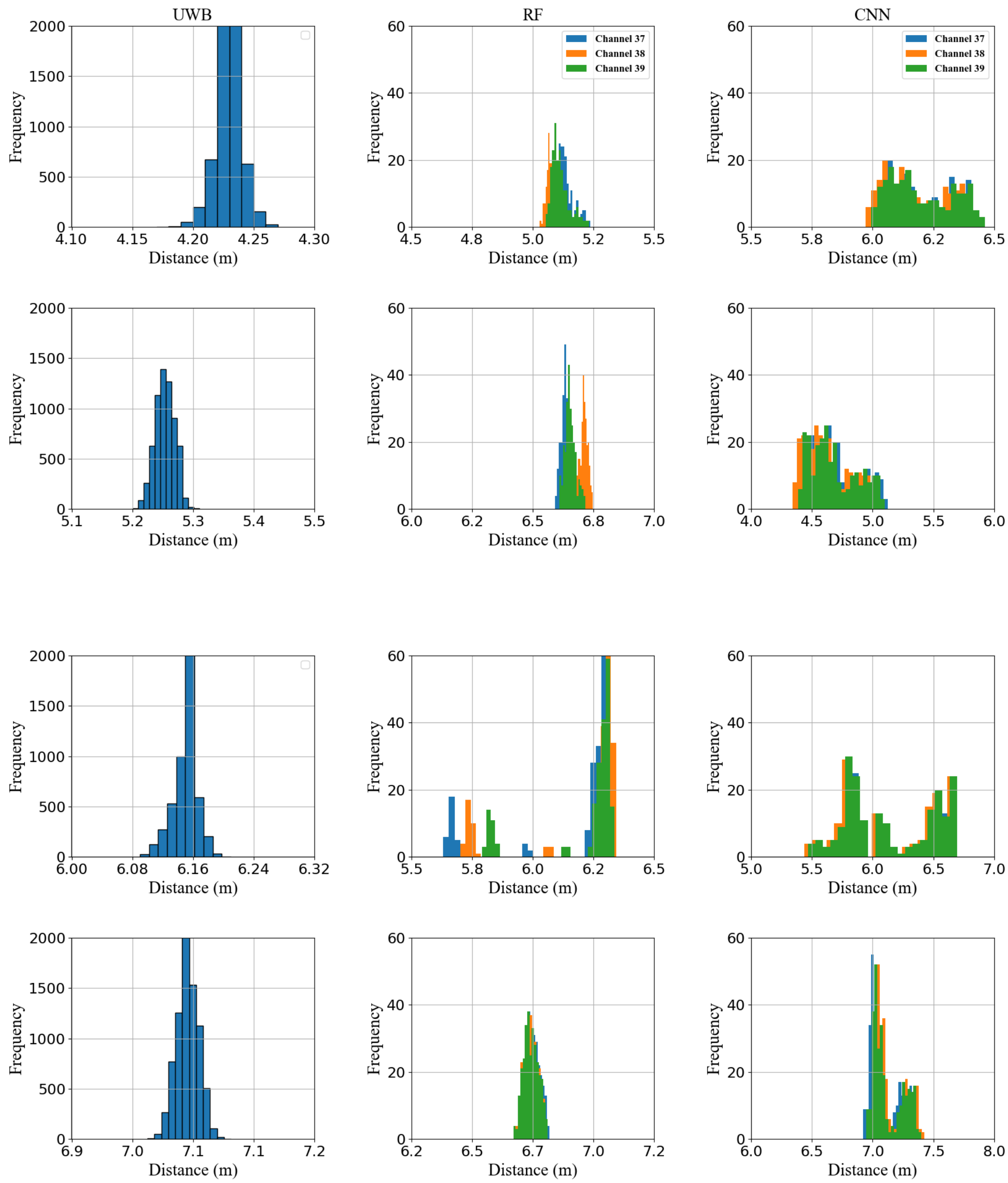


Figure 5.16: Comparison of measured UWB range with the trained ANN estimated distances (RF and CNN) in two adjacent office environments.

On the contrary, the ANN-based RSSI-to-distance model predicted longer distances in three advertising channels at the four reference points. When using the random forest method, the estimated distance falls within the range of the UWB measurements most of the time, with the highest range error difference of 1.5 metres occurring at reference point 2.

Similarly, with the CNN method, the range estimation is almost similar to that of the RF method, with minor differences in the distribution. The highest range error obtained from the method is less than 2 metres, obtained at RP 1. The mean and standard deviation of distances computed by UWB and the two ANN models, relative to the known reference distance, are summarized in Table 5.7. The table provides insights into the distance spread of the RF and CNN methods. It is worth noting that despite the significant obstruction caused by the wall, the RSSI values do not exhibit higher attenuation as expected with increasing distance between the transmitter and receiver. The RSSI values vary between (-60 to -70 dBm) even when the distance exceeds 7 metres. Consequently, the ANN-predicted results often fall between 5 to 6 metres, as observed in the plots above.

In summary, the error between the UWB measured range and the ANN-based RSSI-to-distance model is less than 2 metres, holding true even when considering the ground truth and the ANN model output. This indicates the robustness of the AI model in reliably predicting the proximity of BLE users when they are located in two adjacent office spaces. Moreover, this finding suggests that the proposed algorithm is more suitable than existing distance models in terms of its adaptability and predictive capabilities based on the available data. The AI model can be easily integrated into low-cost BLE smartphones, enabling proximity detection with other Bluetooth users.

Table 5.7: Comparison of UWB mean range and mean estimated distance of trained ANN models with reference distance of BLE user in two adjacent office environments.

Reference	**UWB**		**RF**		**CNN**	
Distance (m)	**KF mean range (m)**	**Std**	**Estimated mean distance (m)**	**Std**	**Estimated mean distance (m)**	**Std**
4	4.22	0.01	5.11	0.03	6.19	0.13
5	5.25	0.02	6.67	0.02	4.67	0.19
6	6.15	0.02	6.20	0.20	6.13	0.37
7	7.05	0.02	6.69	0.03	7.12	0.13

5.3.2 Range Estimation in Blind Locations

This test case further evaluates the ANN model's reliability in accurately determining the proximity between BLE users. It is a more rigorous evaluation compared to the one presented in Section 5.2.3, as it introduces an additional obstacle due to the wall between rooms, on top of the blind locations surrounded by PC monitors and cubicle walls. In this scenario, the test user remained at desk G, while the source was successively positioned at Desk numbers 3 and 4, and measurements were collected for each location. Furthermore, the test receiver was moved to RP 24 (shown in black at centre of END-309) for two instances, with the source positioned at desk number 5 in one case and RP 7 in the other. The last two instances represent scenarios where the users are located in the middle of their respective rooms. The histogram results, depicting the distances measured by UWB and the distances estimated by the ANN models, can be observed in Figure 5.17.

When the BLE source was positioned at Desk 3, the UWB measured range closely matched the ANN estimated range, estimating a separation of approximately 5.3-5.5 m between the users. However, when the source was moved to Desk 4, the ANN estimated distance was longer than the UWB range by approximately 1.5-2 m. This difference can be attributed to the cumulative effect of obstructions in the RSSI signals caused by a bench in END-309, walls, desks, and PC monitors in sequence. These obstructions increased signal attenuation, leading to the observed discrepancy between the two distance estimates.

Next, the test user was positioned in the centre of END-309, while the source was placed at Desk 5. The UWB range measurements ranged from 8.6-8.7 m, with a mean value of 8.64 m. It is important to note that the test user is now receiving the delayed UWB signal resulting in longer distance measurements. The RF approach showed distances between 6.8-7 m, while the CNN method estimated distances ranging from 6.8-8 m. These results are significant as the ANN estimated distances are comparable to the UWB range, with a difference of less than 2 m, taking into account NLOS conditions and the complex dynamic environment.

Furthermore, when the source was positioned at RP 7, the UWB range measurements distributed from 8.8-9.4 m, with a mean value of 8.83 m. The RF method estimated distance had a mean value of 7.97 m with a small spread. On the other hand, the CNN estimated range exceeded 9.5 m, with a spread of 1 m and a mean of 9.75 m. In all four cases the reference ground distance between Tx and Rx are measured using surveying of floors and a laser distance finder.

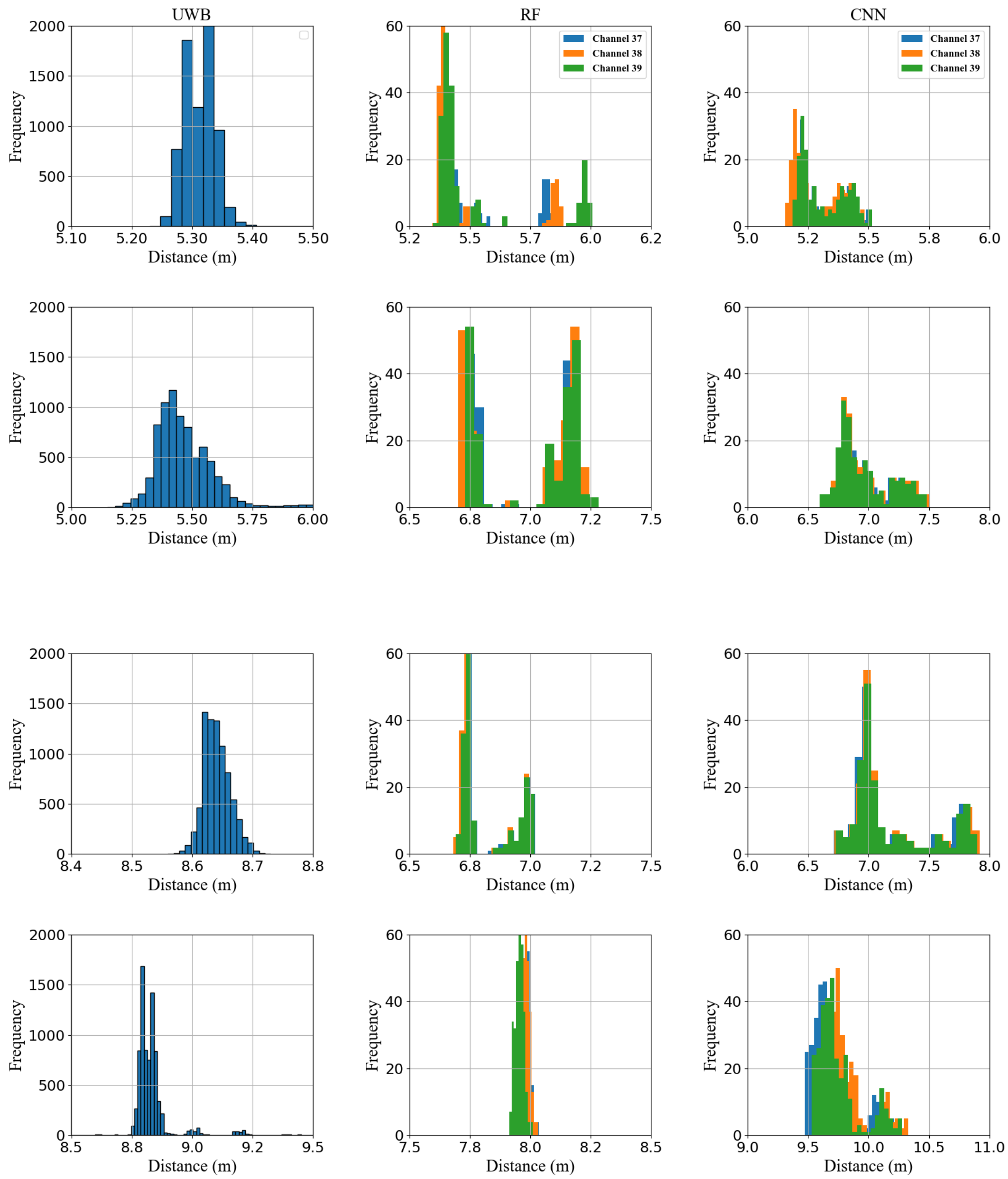


Figure 5.17: Comparison of measured UWB range with the trained ANN estimated distances (RF and CNN) at blind locations in two adjacent office environments.

The mean and standard deviation of distances computed by UWB and the two ANN models, relative to the approx reference distance, are summarized in Table 5.8. Again, it is essential to emphasize that despite signal blockage by the wall and increased distance between the transmitter and receiver, the BLE RSSI values did not experience significant signal attenuation. The lowest RSSI values were within the -70 to -75 dBm range, indicating a relatively strong signal. This can be attributed to constructive interference from the surrounding objects or the design characteristics of the BLE receiver sensor, which effectively maintains the signal strength. As a result, the ANN model's distance output lies up to the maximum allowed range of RSSI values. All three range estimates (UWB, RF, and CNN) for the last case exceeded the expected distance of 8 m, further demonstrating the strength of the ANN model's quality predictions of proximity between two BLE entities with some reasonable accuracy.

Table 5.8: Comparison of UWB mean range and mean estimated distance of trained ANN models with the reference distance at blind locations in two adjacent office environments.

Reference	UWB		RF		CNN	
Distance (m)	KF mean range (m)	Std	Estimated mean distance (m)	Std	Estimated mean distance (m)	Std
4.36	5.31	0.03	5.45	0.18	5.30	0.09
5.28	5.48	0.15	7.00	0.20	6.97	0.22
7.10	8.64	0.02	6.82	0.11	7.18	0.34
8.34	8.83	0.10	7.97	0.02	9.75	0.18

5.4 Summary

The chapter discusses the results of trained ANN models in two separate rooms without further training under real environmental conditions. In each user scenario, the models were provided with BLE advertising channel RSSI values as input and estimated the range as output for each input RSSI value. The results demonstrate a significant improvement in range estimation using the ANN model compared to the path-loss propagation model reported in the literature, particularly in static LOS conditions and blind locations. The results showed that the machine learning algorithm (RF) and the deep neural network model (CNN) achieved similar proximity results. Lastly, it was observed that there was no significant attenuation in the RSSI values beyond a certain distance between the transmitter and receiver using the current BLE receiver sensor.

Chapter 6

Conclusion and Recommendations

6.1 Introduction

The research work is divided into three main components. In the first section, an analysis was conducted on the characteristics of BLE advertising channel signals and UWB range measurements. These measurements were obtained from a fixed transmitter at a specific range. A peer-to-peer test was conducted to analyze the relationship between BLE RSSI and reference known distances up to 8 metres. A first-order line-fitting model was applied to establish a one-to-one mapping between RSSI and distance. The residual error was computed and analyzed. The influence of UWB range measurements as a source of training data for estimating BLE RSSI path-loss exponents for the same measurement path was analyzed by forming a relationship between them. The residual error (difference between the predicted range and the actual range) from the line-fitting model computed for ground truth and UWB range were analyzed and compared.

The second part of the research focused on training two artificial neural network algorithms (RF and CNN) using RSSI and UWB range data in an open environment. A similar peer-to-peer test was once again conducted to collect RSSI and UWB observations. The algorithm was verified and tested to estimate three distance outputs for each input advertising channel RSSI value. The residual error was computed on the test dataset, and the hyper-parameters of the model were adjusted.

The third and final part of the research focuses on evaluating the performance of the trained model in estimating range solely from RSSI without additional training across a range of diverse environmental scenarios. Using the ANN model, it is then possible to estimate proximity between BLE-enabled devices. A comprehensive series of tests were conducted in two office spaces located on the third floor of the CCIT building. The scenarios in the first office include static (LOS) and dynamic user situations (LOS), blind locations (NLOS), detection of new user presence (NLOS), and proximity classification at some threshold level (LOS). The performance of the model is furthermore scrutinized in the adjacent office rooms. Lastly, the effectiveness of the ANN algorithms are assessed for each scenario by computing the residual error using the model output and the reference truth.

6.2 Conclusions

The following conclusions can be drawn from this research work:

Using three separate BLE advertising channels offers improved performance in the range and positioning domain, as demonstrated in the work by Naghdi et al. (Naghdi and O'Keefe, 2019). Individually, the RSSI values from each advertising channel exhibit greater accuracy and stability than the aggregated channel. The results depicted in Figure 3.8 and Figure 3.9 highlight the superior accuracy and reduced fluctuation of RSSI values in the three advertising channels compared to the raw measurements in the aggregated channel mode. Furthermore, Figure 3.10 demonstrates that the variation in each channel can be further mitigated by implementing a Kalman filter.

Extensive analysis was conducted on the characteristics of UWB signals and ranging methods in conjunction with BLE RSSI. It is observed that that the UWB range, computed based on time of flight, offers higher accuracy than distance derived from RSSI using a radio propagation model. This suggests the potential use of UWB as a reliable source of ground truth and training data for converting BLE RSSI to range estimation. The experimental results depicted in Figure 3.14 and Figure 3.15 demonstrate the effectiveness of using UWB as a training source for BLE RSSI in range estimation. This approach eliminates the need for a standard radio model that is environment-specific and has limited accuracy of few metres. The residual error obtained from the line fit model was identical when compared to the reference ground truth and UWB range. Notably, the maximum error observed was approximately 1.4 metres up to 8 metre peer-to-peer distance.

Artificial neural network (ANN) techniques are recognized for their ability to process complex, non-linear, and dynamic data while learning patterns from extracted features. Once sufficiently trained, an ANN model can be applied to new environments, enabling accurate predictions. Therefore, two ANN models (RF, CNN) were trained with advertising channel RSSI and corresponding UWB range observations as input-output combinations. The training was conducted in an open lab environment with line-of-sight conditions, and the dataset consisted of a total of 2130 observations. The RF model was trained on 90% of the data and tested on the remaining 10%, while the CNN model was trained on 80% of the data and tested on the remaining 20% of samples. The residual error obtained from the RF approach was mainly centred around zero, with a standard deviation of 0.08 m. On the other hand, the CNN method had a mean value of -0.13 m and a standard deviation of 0.33 m.

The trained ANN models were applied to two representative office environments, considering various scenarios. These scenarios included static conditions, dynamic movement, blind locations, random user locations, and short-distance proximity classifications. These models take three advertising channel RSSI values as inputs and estimate range as output for each input channel. The residual error in static case is computed by comparing the predicted distance output from the ANN with the known distance. The test results obtained in each user scenario in each office environment can be summarized as follows:

END-309 Office:

1. Static case: The RF method demonstrated superior performance compared to the CNN method in estimating range from a static user, with a mean error of less than 1 m up to 6 m. In contrast, the CNN method exhibited mean error of 1.2 m for the same user distance (Table 5.1 and Table 5.2).

2. Dynamic case: The ANN techniques effectively traced the trajectory of a slow-moving user based on the received RSSI values. Both models exhibited the capability to detect a 2 m proximity threshold, with the RF-based method estimating a smoother trajectory compared to the CNN-based method.

3. Blind locations: The performance of the ANN models was evaluated in five highly NLOS constrained areas characterized by heavy blockage from cubicles and metallic shelves. The

methods proved effective in accurately estimating the range between two users in diagonal cubicles separated by a distance of 7 m. The maximum mean error using the RF approach in the blind locations was less than 2.7 m. The CNN method achieved a maximum mean error of 2.9 m. The test results (Table 5.3) provide further evidence that a thin plastic sheet acts as a transparent medium for RSSI signals at short distances.

4. Random user proximity: The test conducted to identify the presence of a BLE user in the room demonstrated the reliability of the ANN model in detecting and estimating distance. The RF method achieved a mean error of 0.90 m in estimating the user distance, while the CNN method had an error of 0.10 m in NLOS conditions (Table 5.4).

5. Proximity classification: The performance of the ANN model in classifying whether a slow moving user is in close or high proximity based on a 1.5 m threshold yielded overall good results. The RF classification model achieved a correct classification rate of 74.05%, while there were misclassifications in 26% of cases (Table 5.6).

END-313 Office:

1. Static case: The use of the ANN model to estimate the range or proximity between two static BLE users in adjacent office environments demonstrated moderate performance. Despite obstructions caused by dry walls and other objects, the RSSI-to-distance model achieved an accuracy of better than 2.5 metres. The RF method showed better performance than the CNN method in estimating proximity as illustrated in Table 5.7. Furthermore, it is observed that the accuracy of UWB range is limited by the interference of the dry wall.

2. Blind Locations: The evaluation of the ANN models in blind locations within the second office environment yielded reliable results with significantly smaller errors than the blind test case discussed in Section 5.2.3. This can be attributed to the variations in multipath and environmental conditions during different runs. Additionally, the BLE RSSI values remained within a specific lower limit even with strong NLOS conditions and blocked signal paths. As a result, the ANN models could consistently estimate distances within a specific range.

6.3 Application and Recommendation

The research concluded that by using a precise data source such as UWB, it is possible to train ANNs to estimate ranges in multiple BLE channels and identify instances when two BLE users are in proximity. One potential application of this research is in contact tracing, as necessitated during the global COVID-19 pandemic, where BLE RSSI can serve as an observable for proximity detection. While ANN models have shown to be a suitable alternative for converting BLE RSSI to range compared to standard radio propagation models in many cases, they still present challenges for real-world applications. The following are some limitations and suggestions for future research:

1. In the training of ANN models, additional constraints can be incorporated. Currently, the models are trained only using LOS observations. Therefore, better prediction estimates can be achieved in blind or constrained areas by including NLOS observations in the training process. The training dataset size can also be maximized as ANN requires massive data for effective results.

2. The proposed ANN system has primarily been tested in static and very slow-moving user scenarios. However, in real-world conditions, users often move at a relatively faster pace. Therefore, further research can be conducted to analyze the performance of the proximity model in different user movement scenarios.

3. One significant limitation of this work is that the model was trained and tested using a specific type of hardware (Decawave BLE chipset and antenna). However, it is important to recognize that different users may possess different smartphones with varying RSSI levels. As a result, future research could involve evaluating the ANN model using RSSI values from other smartphone models.

4. Although the ANN methods have been tested in NLOS and highly blocked areas, one crucial aspect that is not considered is the human body's impact as an obstacle in the signal path. In future work, it would be highly significant to experiment with scenarios where one or more human obstacles are present, considering real-world situations. Furthermore, further research can be conducted to specifically identify and process NLOS signals within the model.

5. Finally, to have a more performance evaluation of the proposed system, future work can involve testing the ANN models using a minimal amount of training data across diverse environments. This approach will offer valuable insights into the system's robustness and effectiveness in estimating reliable proximity across different locations.

References

Abbas, Z. and Yoon, W. (2015). A survey on energy conserving mechanisms for the internet of things: Wireless networking aspects. *Sensors*, 15(10):24818–24847.

Abdelrahman, R. B. M., Mustafa, A. B. A., and Osman, A. A. (2015). A comparison between ieee 802.11 a, b, g, n and ac standards. *IOSR Journal of Computer Engineering (IOSR-JEC)*, 17(5):26–29.

Aggarwal, P., Syed, Z., Niu, X., and El-Sheimy, N. (2008). A standard testing and calibration procedure for low cost mems inertial sensors and units. *The Journal of Navigation*, 61(2):323–336.

Aitenbichler, E. and Muhlhauser, M. (2003). An ir local positioning system for smart items and devices. In *23rd International Conference on Distributed Computing Systems Workshops, 2003. Proceedings.*, pages 334–339.

Amer, B. and Noureldin, A. (2016). Rss-based indoor positioning utilizing firefly algorithm in wireless sensor networks. In *2016 11th International Conference on Computer Engineering Systems (ICCES)*, pages 329–333.

Arbula, D. and Ljubic, S. (2020). Indoor localization based on infrared angle of arrival sensor network. *Sensors*, 20:1–32.

Bai, L., Ciravegna, F., Bond, R., and Mulvenna, M. (2020). A low cost indoor positioning system using bluetooth low energy. *IEEE Access*, 8:136858–136871.

Balaji, S. (2020). Binary image classifier cnn using tensorflow. *Techiepedia*. `https://medium.com/techiepedia/binary-image-classifier-cnn-using-tensorflow-a3f5d6746697`.

Banerjee, S. (2012). *Improving Accuracy in Ultra-Wideband Indoor Position Tracking through Noise Modeling and Augmentation*. PhD thesis, Clemson University.

Barshan, B. and Durrant-Whyte, H. (1995). Inertial navigation systems for mobile robots. *IEEE Transactions on Robotics and Automation*, 11(3):328–342.

Bhoyar, R., Ghonge, M., and Gupta, S. (2013). Comparative study on ieee standard of wireless lan/wi-fi 802.11 a/b/g/n. *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)*, 2(7):687–691.

Bohn, D. A. (1988). Environmental effects on the speed of sound. *Journal of the Audio Engineering Society*, 36(4):223–231.

Brital, A. (2021). Random forest algorithm explained. *Techiepedia*. `https://anasbrital98.github.io/blog/2021/Random-Forest/`.

Chai, S., An, R., and Du, Z. (2016/04). An indoor positioning algorithm using bluetooth low energy rssi. In *Proceedings of the 2016 International Conference on Advanced Materials Science and Environmental Engineering*, pages 274–276. Atlantis Press.

Chavez-Santiago, R. and Balasingham, I. (2014). Ultrawideband signals in medicine [life sciences]. *IEEE Signal Processing Magazine*, 31(6):130–136.

Chen, J., Ou, G., Peng, A., Zheng, L., and Shi, J. (2018). An ins/wifi indoor localization system based on the weighted least squares. *Sensors*, 18(5):1458.

Chen, Z., Zou, H., Yang, J., Jiang, H., and Xie, L. (2020). Wifi fingerprinting indoor localization using local feature-based deep lstm. *IEEE Systems Journal*, 14(2):3001–3010.

Cheng, L., Chang, H., Wang, K., and Wu, Z. (2020). Real time indoor positioning system for smart grid based on uwb and artificial intelligence techniques. pages 1–7.

Chiu, D. and O'Keefe, K. (2008). Seamless outdoor-to-indoor pedestrian navigation using gps and uwb. volume 1.

Chong, C.-C., Watanabe, F., and Win, M. Z. (2007). Effect of bandwidth on uwb ranging error. In *2007 IEEE Wireless Communications and Networking Conference*, pages 1559–1564.

Chu, C.-H., Wang, C.-H., Liang, C.-K., Ouyang, W., Cai, J.-H., and Chen, Y.-H. (2011). High-accuracy indoor personnel tracking system with a zigbee wireless sensor network. In *2011 Seventh International Conference on Mobile Ad-hoc and Sensor Networks*, pages 398–402.

Coppens, D., Shahid, A., Lemey, S., Van Herbruggen, B., Marshall, C., and De Poorter, E. (2022). An overview of uwb standards and organizations (ieee 802.15.4, fira, apple): Interoperability aspects and future research directions. *IEEE Access*, 10:70219–70241.

Coronel, P., Furrer, S., Schott, W., and Weiss, B. (2008). Indoor location tracking using inertial navigation sensors and radio beacons. In *The Internet of Things*, pages 325–340. Springer.

Dabove, P., Di Pietra, V., Piras, M., Jabbar, A. A., and Kazim, S. A. (2018). Indoor positioning using ultra-wide band (uwb) technologies: Positioning accuracies and sensors' performances. In *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pages 175–184.

DecaWave (2017). 3.5 - 6.5 ghz ultra-wideband (uwb) transceiver ic with 1 antenna port. *Decawave Limited*. `https://www.qorvo.com/products/p/DW1000#documents`.

Decawave (2017). Dwm1001-dev product datasheet. *Decawave products*. `https://www.qorvo.com/products/p/DWM1001-DEV#documents`.

DecaWave, A. (2015). The implementation of two-way ranging with the dw1000. *Decawave Limited*.

Dong, Q. and Dargie, W. (2012). Evaluation of the reliability of rssi for indoor localization. In *2012 International Conference on Wireless Communications in Underground and Confined Areas*, pages 1–6. IEEE.

Dotlic, I., Connell, A., and McLaughlin, M. (2018). Ranging methods utilizing carrier frequency offset estimation. In *2018 15th Workshop on Positioning, Navigation and Communications (WPNC)*, pages 1–6.

Ergen, S. C. (2004). Zigbee/ieee 802.15. 4 summary. *UC Berkeley, September*, 10(17):11.

Fang, S.-H., Wang, C.-H., Huang, T.-Y., Yang, C.-H., and Chen, Y.-S. (2012). An enhanced zigbee indoor positioning system with an ensemble approach. *IEEE Communications Letters*, 16(4):564–567.

Faragher, R. and Harle, R. (2014). An analysis of the accuracy of bluetooth low energy for indoor positioning applications. In *Proceedings of the 27th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2014)*, pages 201–210.

Faragher, R. and Harle, R. (2015). Location fingerprinting with bluetooth low energy beacons. *IEEE Journal on Selected Areas in Communications*, 33(11):2418–2428.

Flammini, A., Marioli, D., Sisinni, E., and Taroni, A. (2009). Design and implementation of a wireless fieldbus for plastic machineries. *IEEE Transactions on Industrial Electronics*, 56(3):747–755.

Fontana, R. (2004a). Recent system applications of short-pulse ultra-wideband (uwb) technology. *IEEE Transactions on Microwave Theory and Techniques*, 52(9):2087–2104.

Fontana, R. J. (2004b). Recent system applications of short-pulse ultra-wideband (uwb) technology. *IEEE Transactions on microwave theory and techniques*, 52(9):2087–2104.

Gadhgadhi, A., Hachaichi, Y., and Zairi, H. (2019). Distance estimation using polynomial approximation and neural network based on rssi technique. In *2019 19th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, pages 24–29. IEEE.

García, E., Poudereux, P., Hernández, , Ureña, J., and Gualda, D. (2015). A robust uwb indoor positioning system for highly complex environments. In *2015 IEEE International Conference on Industrial Technology (ICIT)*, pages 3386–3391.

Gezici, S. and Poor, H. V. (2009). Position estimation via ultra-wide-band signals. *Proceedings of the IEEE*, 97(2):386–403.

Giuliano, R., Cardarilli, G. C., Cesarini, C., Di Nunzio, L., Fallucchi, F., Fazzolari, R., Mazzenga, F., Re, M., and Vizzarri, A. (2020). Indoor localization system based on bluetooth low energy for museum applications. *Electronics*, 9(6).

Hoang, M. T., Yuen, B., Dong, X., Lu, T., Westendorp, R., and Reddy, K. (2019). Recurrent neural networks for accurate rssi indoor localization. *IEEE Internet of Things Journal*, 6(6):10639–10651.

Hsieh, C.-H., Chen, J.-Y., and Nien, B.-H. (2019). Deep learning-based indoor localization using received signal strength and channel state information. *IEEE access*, 7:33256–33267.

Ibrahim, M., Torki, M., and ElNainay, M. (2018a). Cnn based indoor localization using rss time-series. In *2018 IEEE Symposium on Computers and Communications (ISCC)*, pages 01044–01049.

Ibrahim, M., Torki, M., and ElNainay, M. (2018b). Cnn based indoor localization using rss time-series. In *2018 IEEE symposium on computers and communications (ISCC)*, pages 01044–01049. IEEE.

Ijaz, F., Yang, H. K., Ahmad, A. W., and Lee, C. (2013). Indoor positioning: A review of indoor ultrasonic positioning systems. In *2013 15th International Conference on Advanced Communications Technology (ICACT)*, pages 1146–1150.

Jedari, E., Wu, Z., Rashidzadeh, R., and Saif, M. (2015). Wi-fi based indoor location positioning employing random forest classifier. In *2015 international conference on indoor positioning and indoor navigation (IPIN)*, pages 1–5. IEEE.

Jiang, C., Shen, J., Chen, S., Chen, Y., Liu, D., and Bo, Y. (2020). Uwb nlos/los classification using deep learning method. *IEEE Communications Letters*, 24(10):2226–2230.

Jiang, Y. and Leung, V. C. (2007). An asymmetric double sided two-way ranging for crystal offset. In *2007 International Symposium on Signals, Systems and Electronics*, pages 525–528.

Jianyong, Z., Haiyong, L., Zili, C., and Zhaohui, L. (2014). Rssi based bluetooth low energy indoor positioning. In *2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 526–533.

Jiménez Ruiz, A. R. and Seco Granja, F. (2017). Comparing ubisense, bespoon, and decawave uwb location systems: Indoor performance analysis. *IEEE Transactions on Instrumentation and Measurement*, 66(8):2106–2117.

Kaduskar, A., Vengurlekar, O., and Shinde, V. (2020). Rssi filtering methods applied to localization using bluetooth low energy. *International Journal of Recent Technology and Engineering (IJRTE)*, 9(3):2277–3878.

Kim, H.-S., Kim, D.-R., Yang, S.-H., Son, Y.-H., and Han, S.-K. (2013). An indoor visible light communication positioning system using a rf carrier allocation technique. *Journal of Lightwave Technology*, 31(1):134–144.

Kim Geok, T., Zar Aung, K., Sandar Aung, M., Thu Soe, M., Abdaziz, A., Pao Liew, C., Hossain, F., Tso, C. P., and Yong, W. H. (2021a). Review of indoor positioning: Radio wave technology. *Applied Sciences*, 11(1).

Kim Geok, T., Zar Aung, K., Sandar Aung, M., Thu Soe, M., Abdaziz, A., Pao Liew, C., Hossain, F., Tso, C. P., and Yong, W. H. (2021b). Review of indoor positioning: Radio wave technology. *Applied Sciences*, 11(1).

Kok, M., Hol, J. D., and Schön, T. B. (2015). Indoor positioning using ultrawideband and inertial measurements. *IEEE Transactions on Vehicular Technology*, 64(4):1293–1303.

Lakkundi, V. (2006). Ultra wideband communications: History, evolution and emergence. *Acta Polytechnica*, 46(4).

Lee, C., Nam, Y.-S., Kim, J.-Y., and Cho, Y.-Z. (2009). One-way ranging method using reference-based broadcasting messages for wireless sensor networks. *IEICE transactions on communications*, 92(6):2271–2274.

Lee, H.-j. and Lee, M. C. (2006). Localization of mobile robot based on radio frequency identification devices. In *2006 SICE-ICASE International Joint Conference*, pages 5934–5939.

Li, G., Geng, E., Ye, Z., Xu, Y., Lin, J., and Pang, Y. (2018). Indoor positioning algorithm based on the improved rssi distance model. *Sensors*, 18(9):2820.

Li, S., Deng, Y.-Q., Zhu, Z.-L., Hua, H.-L., and Tao, Z.-Z. (2021). A comprehensive review on radiomics and deep learning for nasopharyngeal carcinoma imaging. *Diagnostics*, 11(9):1523.

Li, Z., Wang, S.-H., Fan, R.-R., Cao, G., Zhang, Y.-D., and Guo, T. (2019). Teeth category classification via seven-layer deep convolutional neural network with max pooling and global average pooling. *International Journal of Imaging Systems and Technology*, 29(4):577–583.

Liu, H., Darabi, H., Banerjee, P., and Liu, J. (2007). Survey of wireless indoor positioning techniques and systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 37(6):1067–1080.

Lu, Y.-M., Sheu, J.-P., and Kuo, Y.-C. (2021). Deep learning for ultra-wideband indoor positioning. In *2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 1260–1266.

Mackey, A., Spachos, P., Song, L., and Plataniotis, K. N. (2020). Improving ble beacon proximity estimation accuracy through bayesian filtering. *IEEE Internet of Things Journal*, 7(4):3160–3169.

Maglogiannis, I. and Hadjiefthymiades, S. (2007). Emerloc: Location-based services for emergency medical incidents. *International journal of medical informatics*, 76(10):747–759.

Mahfouz, M. R., Kuhn, M. J., Wang, Y., Turnmire, J., and Fathy, A. E. (2011). Towards sub-millimeter accuracy in uwb positioning for indoor medical environments. In *2011 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems*, pages 83–86.

Mahfouz, M. R., Zhang, C., Merkl, B. C., Kuhn, M. J., and Fathy, A. E. (2008). Investigation of high-accuracy indoor 3-d positioning using uwb technology. *IEEE Transactions on Microwave Theory and Techniques*, 56(6):1316–1330.

Mainetti, L., Patrono, L., and Sergi, I. (2014). A survey on indoor positioning systems. In *2014 22nd International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pages 111–120.

Majeed, K. and Hranilovic, S. (2021). Passive indoor visible light positioning system using deep learning. *IEEE Internet of Things Journal*, 8(19):14810–14821.

Makki, A., Siddig, A., Saad, M., Cavallaro, J. R., and Bleakley, C. J. (2016). Indoor localization using 802.11 time differences of arrival. *IEEE Transactions on Instrumentation and Measurement*, 65(3):614–623.

Mautz, R. (2012). *Application for Venia Legendi in Positioning and Engineering Geodesy*. ETH Zurich, Department of Civil, Environmental and Geomatic Engineering.

McCarthy, J. et al. (2007). *What is artificial intelligence.* Stanford University.

Minh Dang, L., Min, K., Wang, H., Jalil Piran, M., Hee Lee, C., and Moon, H. (2020). Sensor-based and vision-based human activity recognition: A comprehensive survey. *Pattern Recognition*, 108:107561.

Muqaibel, A., Safaai-Jazi, A., Bayram, A., Attiya, A., and Riad, S. (2005). Ultrawideband through-the-wall propagation. *IEE Proceedings-Microwaves, Antennas and Propagation*, 152(6):581–588.

Naghdi, S. and O'Keefe, K. (2019). Trilateration with ble rssi accounting for pathloss due to human obstacles. In *2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–8. IEEE.

Neirynck, D., Luk, E., and McLaughlin, M. (2016). An alternative double-sided two-way ranging method. In *2016 13th Workshop on Positioning, Navigation and Communications (WPNC)*, pages 1–4.

Nikoukar, A., Abboud, M., Samadi, B., Güneş, M., and Dezfouli, B. (2018). Empirical analysis and modeling of bluetooth low-energy (ble) advertisement channels. In *2018 17th Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net)*, pages 1–6.

Oh, J. H., Kim, D., and Lee, B. H. (2014). An indoor localization system for mobile robots using an active infrared positioning sensor. *Journal of Industrial and Intelligent Information*, 2(1).

Ostlin, E., Zepernick, H.-J., and Suzuki, H. (2010). Macrocell path-loss prediction using artificial neural networks. *IEEE Transactions on Vehicular Technology*, 59(6):2735–2747.

Parker, M. (2017). Chapter 12 - error-correction coding. In Parker, M., editor, *Digital Signal Processing 101 (Second Edition)*, pages 129–147. Newnes, second edition edition.

Parmar, R. (2018). Training deep neural networks. *Techiepedia*. `https://towardsdatascience.com/training-deep-neural-networks-9fdb1964b964`.

Pei, L., Liu, J., Chen, Y., Chen, R., and Chen, L. (2017). Evaluation of fingerprinting-based wifi indoor localization coexisted with bluetooth. *The Journal of Global Positioning Systems*, 15(1):1–12.

Qi, J. and Liu, G.-P. (2017). A robust high-accuracy ultrasound indoor positioning system based on a wireless sensor network. *Sensors*, 17(11).

Rahayu, Y., Abd Rahman, T., Ngah, R., and Hall, P. (2008a). Ultra wideband technology and its applications. pages 1 – 5.

Rahayu, Y., Rahman, T. A., Ngah, R., and Hall, P. (2008b). Ultra wideband technology and its applications. In *2008 5th IFIP International Conference on Wireless and Optical Communications Networks (WOCN '08)*, pages 1–5.

Rani, K. U. (2011). Analysis of heart diseases dataset using neural network approach. *arXiv preprint arXiv:1110.2626*.

Rovnakova, J., Svecova, M., Kocur, D., Nguyen, T. T., and Sachs, J. (2008). Signal processing for through wall moving target tracking by m-sequence uwb radar. In *2008 18th International Conference Radioelektronika*, pages 1–4.

Sadowski, S. and Spachos, P. (2018). Rssi-based indoor localization with the internet of things. *IEEE Access*, 6:30149–30161.

Sanpechuda, T. and Kovavisaruch, L. (2008). A review of rfid localization: Applications and techniques. In *2008 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, volume 2, pages 769–772.

Sarkar, T., Ji, Z., Kim, K., Medouri, A., and Salazar-Palma, M. (2003). A survey of various propagation models for mobile communication. *IEEE Antennas and Propagation Magazine*, 45(3):51–82.

Shan, G. and Roh, B.-H. (2018). Advertisement interval to minimize discovery time of whole ble advertisers. *IEEE Access*, 6:17817–17825.

Sidorenko, J., Schatz, V., Scherer-Negenborn, N., Arens, M., and Hugentobler, U. (2019). Decawave uwb clock drift correction and power self-calibration. *Sensors*, 19(13):2942.

Silva, B., Pang, Z., Åkerberg, J., Neander, J., and Hancke, G. (2014). Experimental study of uwb-based high precision localization for industrial applications. In *2014 IEEE International Conference on Ultra-WideBand (ICUWB)*, pages 280–285.

Sinha, R. S. and Hwang, S.-H. (2019). Comparison of cnn applications for rssi-based fingerprint indoor localization. *Electronics*, 8(9):989.

Siva, J., Yang, J., and Poellabauer, C. (2019). Connection-less ble performance evaluation on smartphones. *Procedia Computer Science*, 155:51–58.

Song, X., Fan, X., Xiang, C., Ye, Q., Liu, L., Wang, Z., He, X., Yang, N., and Fang, G. (2019). A novel convolutional neural network based indoor localization framework with wifi fingerprinting. *IEEE Access*, 7:110698–110709.

Subhan, F., Hasbullah, H., Rozyyev, A., and Bakhsh, S. T. (2011). Indoor positioning in bluetooth networks using fingerprinting and lateration approach. In *2011 International Conference on Information Science and Applications*, pages 1–9. IEEE.

Tabassam, A. A., Heiss, S., and Hoing, M. (2007). Bluetooth device discovery and hop synchronization by the eavesdropper. In *2007 International Conference on Emerging Technologies*, pages 1–5.

Tadakamadla, S. and Oelmann, B. (2006). Indoor local positioning system for zigbee, based on rssi.

Varshney, V., Goel, R. K., and Qadeer, M. A. (2016). Indoor positioning system using wi-fi bluetooth low energy technology. In *2016 Thirteenth International Conference on Wireless and Optical Communications Networks (WOCN)*, pages 1–6.

Wang, T. Q., Sekercioglu, Y. A., Neild, A., and Armstrong, J. (2013). Position accuracy of time-of-arrival based ranging using visible light with application in indoor localization systems. *Journal of Lightwave Technology*, 31(20):3302–3308.

Wang, X., Gao, L., Mao, S., and Pandey, S. (2017). Csi-based fingerprinting for indoor localization: A deep learning approach. *IEEE Transactions on Vehicular Technology*, 66(1):763–776.

Wang, Z., Li, S., Zhang, Z., Lv, F., and Hou, Y. (2018). Research on uwb positioning accuracy in warehouse environment. *Procedia Computer Science*, 131:946–951. Recent Advancement in Information and Communication Technology:.

Want, R., Hopper, A., Falcão, V., and Gibbons, J. (1992). The active badge location system. *ACM Trans. Inf. Syst.*, 10(1):91–102.

Wheeler, A. (2007). Commercial applications of wireless sensor networks using zigbee. *IEEE Communications Magazine*, 45(4):70–77.

Win, M. and Scholtz, R. (2000). Ultra-wide bandwidth time-hopping spread-spectrum impulse radio for wireless multiple-access communications. *IEEE Transactions on Communications*, 48(4):679–689.

Win, M. Z., Dardari, D., Molisch, A. F., Wiesbeck, W., and Jinyun Zhang, W. (2009). History and applications of uwb. Institute of Electrical and Electronics Engineers.

Wu, S., Sakamoto, T., Oishi, K., Sato, T., Inoue, K., Fukuda, T., Mizutani, K., and Sakai, H. (2019). Person-specific heart rate estimation with ultra-wideband radar using convolutional neural networks. *IEEE Access*, 7:168484–168494.

Yadav, R. and Malviya, L. (2020). Uwb antenna and mimo antennas with bandwidth, band-notched, and isolation properties for high-speed data rate wireless communication: A review. *International Journal of RF and Microwave Computer-Aided Engineering*, 30(2):e22033.

Yamano, K., Tanaka, K., Hirayama, M., Kondo, E., Kimuro, Y., and Matsumoto, M. (2004). Self-localization of mobile robots with rfid system by using support vector machine. volume 4, pages 3756 – 3761 vol.4.

Yang, O. (2007). Introduction to ultra wideband (uwb). *Techiepedia*. `https://hajim.rochester.edu/ece/sites/wcng//meetings/YangSp07.pdf`.

Yu, L., Laaraiedh, M., Avrillon, S., and Uguen, B. (2011). Fingerprinting localization based on neural networks and ultra-wideband signals. In *2011 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pages 184–189.

Zafari, F., Gkelias, A., and Leung, K. K. (2019). A survey of indoor localization systems and technologies. *IEEE Communications Surveys Tutorials*, 21(3):2568–2599.

Zhang, R., Xia, W., Jia, Z., and Shen, L. (2014). The indoor localization method based on the integration of rssi and inertial sensor. In *2014 IEEE 3rd Global Conference on Consumer Electronics (GCCE)*, pages 332–336.

Zhao, Z., Zhang, M., Yang, C., Fang, J., and Huang, G. Q. (2018). Distributed and collaborative proactive tandem location tracking of vehicle products for warehouse operations. *Computers & Industrial Engineering*, 125:637–648.

Zhou, Y., Law, C. L., Guan, Y. L., and Chin, F. (2011). Indoor elliptical localization based on asynchronous uwb range measurement. *IEEE Transactions on Instrumentation and Measurement*, 60(1):248–257.

Appendix A: Neural Network Re-Training

The two ANN models described in sections 4.4.1 and 4.4.2 were originally trained to produce three range outputs, each corresponding to one of the three BLE RSSI inputs. However, it is possible to retrain these models to produce a single range output based on all three RSSI input values. In sections A.1 and A.2, both models are retrained to generate a single range estimate using the three separate channel RSSI values. The retrained models are evaluated to estimate single distance output for the static test case in Appendix B.1.

A.1 Random Forest for Multi-channel BLE RSSI to Single Range Estimation.

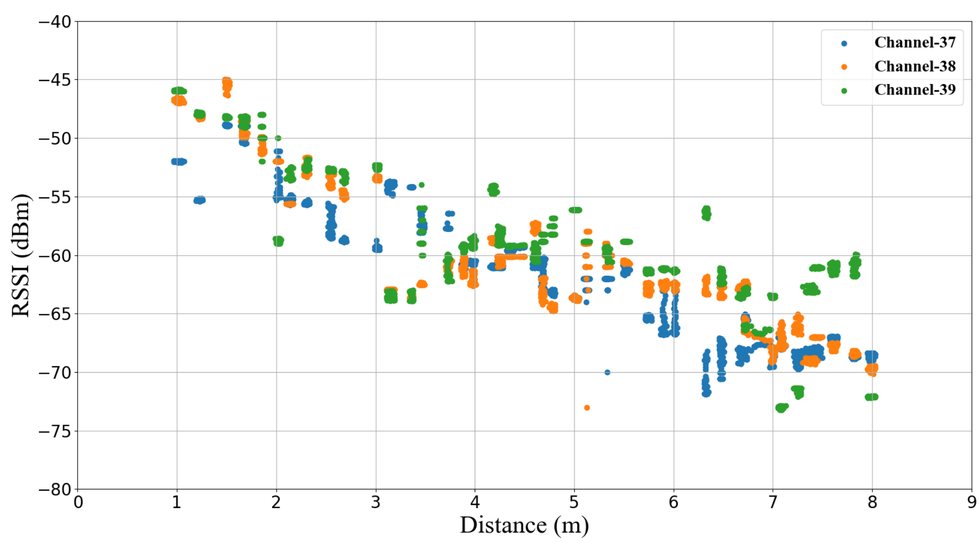


Figure A1.1: BLE RSSI and UWB range dataset for training random forest algorithm.

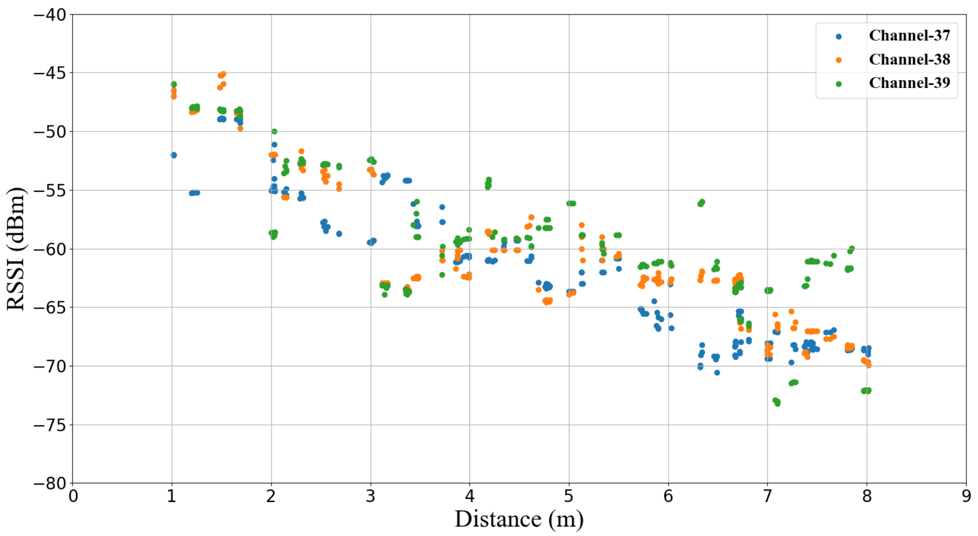


Figure A1.2: Predicted distance output of RF algorithm on test samples.

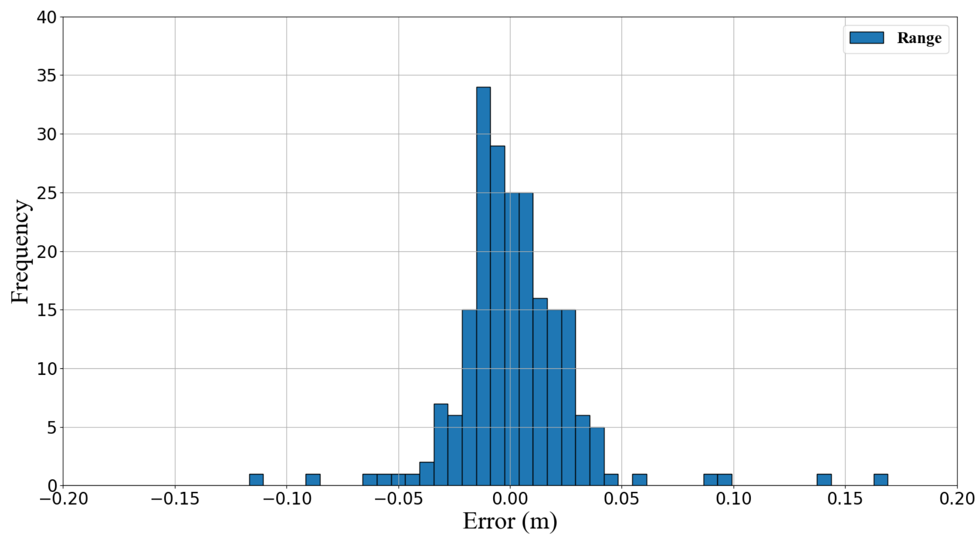


Figure A1.3: Residual error between predicted and reference distance.

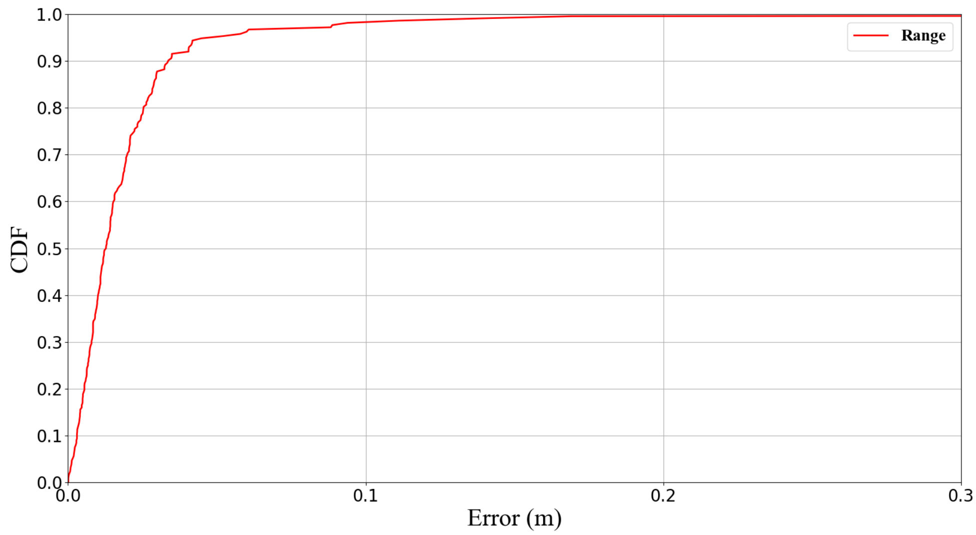


Figure A1.4: Cumulative distance error of single distance estimate.

A.2 Convolution Neural Network for Multi-channel BLE RSSI to Single Range Estimation.

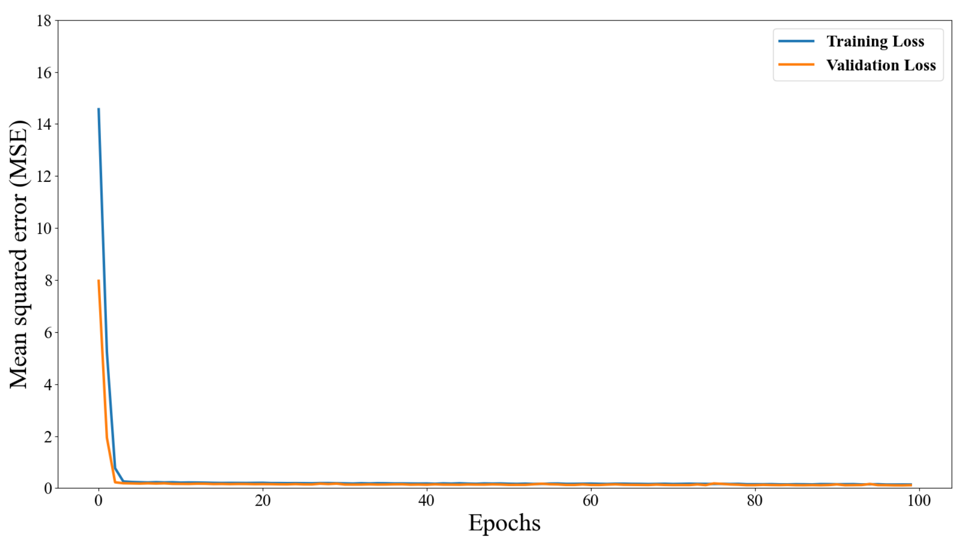


Figure A2.1: Training and validation loss of CNN using parameters in Table 4.2 (except only one neuron at the last dense layer). The best validation performance is 0.122 at epoch number 8.

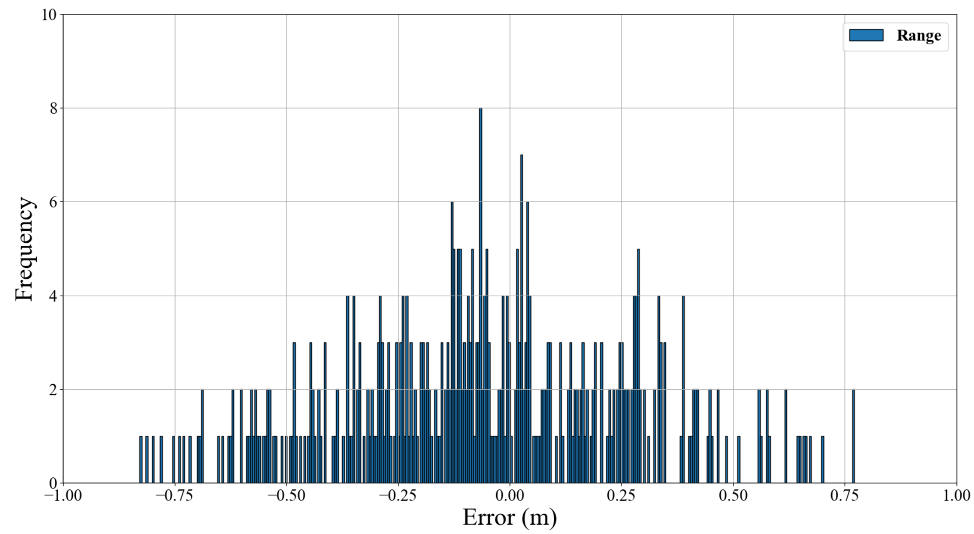


Figure A2.2: Residual error of CNN model between predicted and reference range.

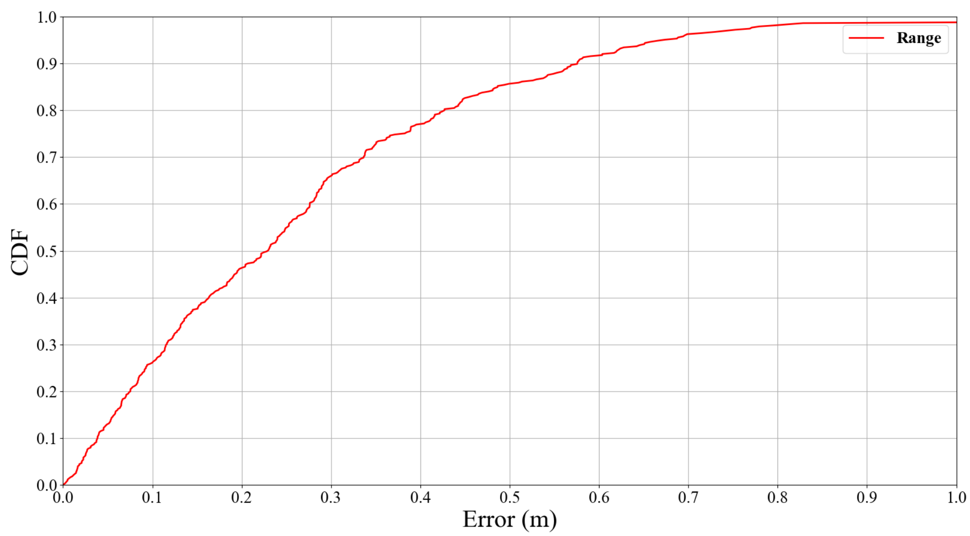


Figure A2.3: Cumulative distance error of single distance estimate.

Appendix B: Re-Trained Neural Network Results

B.1 Range Estimation in Static Case

This section shows the results of section 5.2.1 test using the two re-trained models to output single range estimate. The residual error is calculated using model predicted output and the known truth.

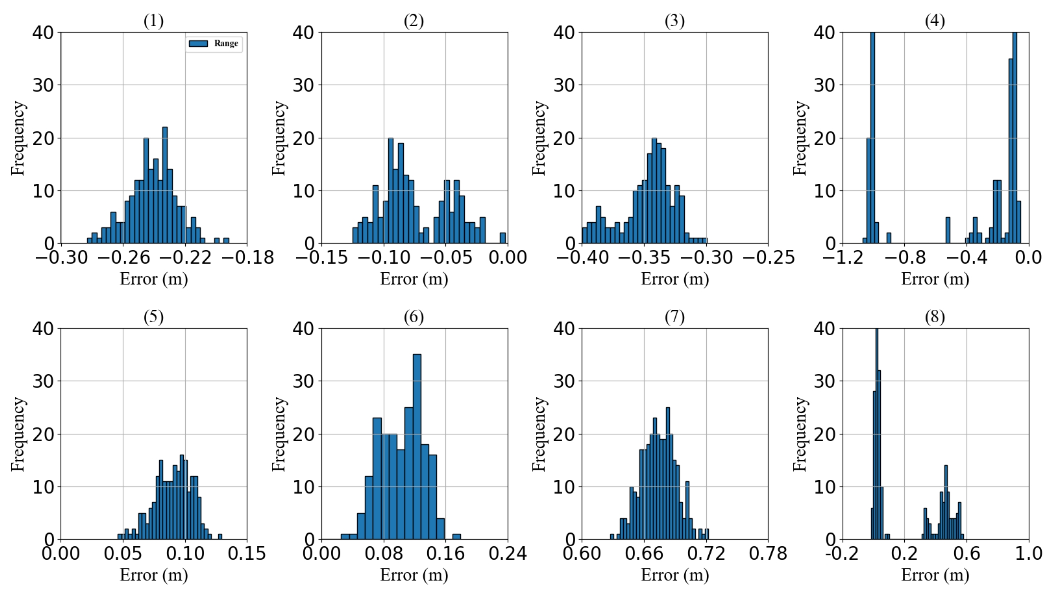


Figure B.1: Histogram of distance estimation errors using retrained RF model in the North-South direction at LOS in 8 reference points. Each subplot corresponds to residual error measured at (1) 1 m, (2) 2 m, (3) 3 m, (4) 4 m, (5) 5 m, (6) 6 m, (7) 7 m, and (8) 8 m respectively in sequence for each BLE advertising channel.

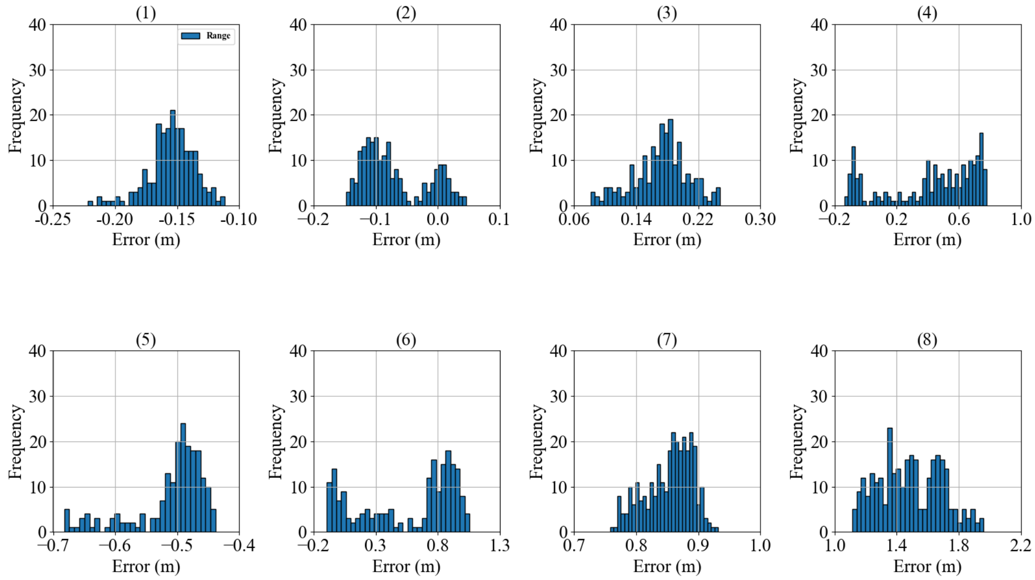


Figure B.2: Histogram of distance estimation errors using retrained CNN model in the North-South direction at LOS in 8 reference points. Each subplot corresponds to residual error measured at (1) 1 m, (2) 2 m, (3) 3 m, (4) 4 m, (5) 5 m, (6) 6 m, (7) 7 m, and (8) 8 m respectively in sequence for each BLE advertising channel.

The estimated single-distance outputs obtained through retraining the model closely resemble the results (Table 5.2) acquired using the originally trained model, with only a slight difference of a few centimetres. A summary of the results achieved through the retrained Random Forest and Convolutional Neural Network models is presented in Table B.1. Additionally, the retrained model can be applied to other test case scenarios for estimating single-distance outputs.

Table B.1: Estimated distance and error percent of re-trained ANN models at LOS peer-to-peer distances in the North-South direction.

Distance (m)	RF		CNN	
	Estimated mean distance (m)	Error percent (%)	Estimated mean distance (m)	Error percent (%)
1	1.22	22	1.15	15
2	2.07	3.5	2.06	3
3	3.35	11.6	2.83	27.66
4	4.45	11.25	3.58	10.5
5	4.90	2	5.51	9.8
6	5.90	1.5	5.44	9.33
7	6.32	9.57	6.14	12.28
8	7.81	8.8	6.15	23.12